

Odprti podatki

načrt za vzpostavitev
sistema odprtega dostopa
do raziskovalnih podatkov
v Sloveniji

Janez Štebe, Sonja Bezjak, Sanja Lužar

Janez Štebe, Sonja Bezjak, Sanja Lužar

Odperti podatki: načrt za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v Sloveniji

Izdajatelj in založnik: Fakulteta za družbene vede, Založba FDV

Ljubljana, 2013

Copyright ©



[To delo](#) avtorjev Janez Štebe, Sonja Bezjak, Sanja Lužar je ponujeno pod [Creative Commons Attribution 4.0 International](#)

Recenzenta: dr. Mojca Kotar in dr. Stojan Sorčan

Jezikovni pregled: dr. Tina Lengar Verovnik

Oblikovanje: Maja Ojsteršek

Način dostopa (URL):

http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odprti_podatki_zakljucno_porocilo/

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

001.891:02:004.7(497.4)(0.034.2)

ŠTEBE, Janez

Odperti podatki [Elektronski vir] : načrt za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v Sloveniji / Janez Štebe, Sonja Bezjak, Sanja Lužar. - El. knjiga. - Ljubljana : Fakulteta za družbene vede, 2013

ISBN 978-961-235-650-7 (pdf)

1. Bezjak, Sonja 2. Lužar, Sanja
269233664

Projekt	ODPRTI PODATKI – Priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji Raziskovalni projekt št. V5-1018 v okviru Ciljnega raziskovalnega programa »KONKURENČNOST SLOVENIJE 2006–2013«
Vodja projekta	Janez Štebe, predstojnik ADP
Vsebinski spremljevalki	Petra Tramte, Direktorat za znanost, MIZŠ Liljana Lučič, ARRS
Financer	MIZŠ, ARRS

OPEN DATA – Action Plan for the Establishment of a System of Open Access to Publicly Funded Research Data in Slovenia

SUMMARY OF A REPORT

With the purpose of clarifying how research data are treated currently in the country and to propose suggestions for future plans, considering the need for respecting current conditions, MIZS (Ministry of Education, Science and Sports) and ARRS (Slovenian Research Agency) issued a call for the grant on the “Target research programme” with the heading OPEN DATA – Action Plan for the Establishment of a System of Open Access to Publicly Funded Research Data in Slovenia. The project proposal of the Social Science Data Archives (Arhiv družboslovnih podatkov – ADP) of the University of Ljubljana was accepted for the grant.

An agreement about the importance of building open access to research data for development of science was reached in the international environment. OECD Principles and Guidelines for Access to Research Data from Public Funding, European Commission Policy Initiatives and other scientific stakeholders’ recommendations (LERU, Science Europe working group on Research Data, LIBER and ALLEA) are influential. The Republic of Slovenia has committed, with the recently acquired membership in the OECD international organisation, to follow its Guidelines regarding open access to research data from public funding. It has taken over certain responsibilities to follow the EU Initiatives within fixed timeframes and with dedicated financial support both regarding access to research data and in other areas of open access to scientific information. The Ministry of Education, Science and Sports (MIZS) supports the European collaborative development of scientific infrastructure according to the agreed ESFRI Roadmap which, irrespective of covering diverse scientific disciplines, has improved access to research data in common. European Structural Funds could be well used to incentivise the upholding of disciplinary research data centre infrastructures.

The research data life cycle encompasses a wide area of dealing with and managing data, starting with planning and creating data, editing, transforming, evaluating and selecting, primary analysis, registering and preserving data, and finally discovering it for access and secondary use. Different stages are performed by different actors with each of them having their own goals, needs, and habits. The actors come from various scientific fields. Researchers may not see beyond the limits of their own research area expertise to follow a wider state of discussion regarding open access to research data. Despite this, dedicated expert discussions have already been present for some time, and they offer some strong conclusions to lean on. They address problems about how to secure long-term digital preservation effectively and according to the highest quality standards, and problems about effective secondary use of research data. Questions, dilemmas and creative solutions offered are already mature in some environments

with developed research infrastructure to support the creation of quality research data and effective use of its potential for the research community. Institutional frameworks are created based on vivid and fruitful discussions about experiences gained, with the purpose of widening good practices into environments that still struggle with how to address the problems and find solutions. The creation of research data, its digital curation, and access to these data need carefully designed dedicated support service infrastructure to fully exploit the potential and thus justify the investments.

The report brings together partial results of the whole three-year project, jointly in one final publication. It consists of a shortened presentation of main findings of the first part of the research project that deals with the analysis of the current state and conditions in the country. This was as a starting point for further activities that followed during the length of the OPEN DATA project. A series of meetings and workshops was organised with the aim to search for an appropriate approach in introducing proposed solutions to the problems that were identified before. The proposal of certain solutions to the problems identified was already influenced by the insights gained on the situation in the international scene.

Part two of the current report then brings a wider overview of activities in the area of data service and data infrastructure development in the international environment, which was carried out as a parallel activity during the second part of the project to become informed about the current technological and organisational approaches being used globally. We analysed in detail the international initiatives which dealt with research data, emphasised good practice and presented experiences in other countries and in respective institutions.

The report finally presents the main results of the OPEN DATA project: A proposal for the action plan for the establishment of an open access system for research data from public funding. In the appendix, an open research data policy draft for Slovenia, consistent with the proposed action plan, is presented. After analysing the situation in Slovenia and comparatively presenting the international dynamic, the final aim of the project was the identification of emerging points for the future, a more elaborated system of access to research data. Propositions about such a system, which would be adapted to the institutional and human resources background of Slovenian scientific environments, were presented as a synthesis in the final action plan proposed. Thus, this is based on a realistic assessment of current conditions and potential. At the same time, it is designed as an incentive to open up opportunities for developing activities everywhere where there is willingness to take over responsibilities for tasks, which would introduce changes in the most creative, effective, and reasonable way.

The action plan foresees the need for an intermediate period for some disciplinary areas to start establishing research data management support services. Active promotion of ideas and clearly expressed willingness of those most accountable for development in science are needed to support their realisation. Only in such a case will the internationally prevailing agreement on the importance of open access to research data be able to spread into different disciplinary and institutional environments. The determination to introduce changes, accompanied by visible support in necessary activities, is a main responsibility of funding bodies and scientific policy makers. But at the same time, the support needs to be gained from the representatives of the scientific community, such as professional associations, councils, advisory bodies, the academy of science, heads of institutes, academic publishers, and other reputed individuals. The changes introduced need to be financially effective, fair, and orientated towards a visible improvement of

the quality of scientific research. Changes also need to be introduced to generate trust in scientific results from the general public's perspective, by providing greater transparency of scientific conclusions, which could be checked using publicly available data.

Computer technologies help to support new possibilities of generating, linking, sharing, and further use of data and other scientific artefacts in the new 'open science' paradigm. But technological changes alone are not sufficient to bring changes. Firstly we need to achieve expert, political, and societal consensus on what is worth preserving and sharing with others, and how to make this happen with a guarantee that those involved would be appropriately recognized, awarded, motivated, and protected.

Chapters' overview

The starting chapters present the goals and the purpose of the OPEN DATA project, and elaborate the reasons behind the project. Then we begin with the introduction to the analysis of the current situation in daily practice in the data life cycle in different disciplinary settings nationally. Some other studies, which analysed the needs and habits of researchers related to data provision infrastructure services, are presented as examples for the reference to the national study. Results are presented with a perspective of highlighting problems and simultaneously suggesting viable solutions from the perspective of different roles and responsibilities in data management in open access scenarios.

The presentation of the national situation is based on the interviews conducted with Slovenian researchers and other stakeholders with a role in research data provision. In order to gain deep understanding of the situation, we sought for immediate answers provided directly from those actors involved in data management, working on routine everyday research. We distinguished roles played and problems faced within data creation, data access and use, and those in providing wider support services. An inductive method of collecting data was selected as the most appropriate for the purpose of detecting problems in a complex environment. Thus, 22 lengthy semi-structured interviews were carried out. The discussions with interviewees from different disciplinary and work settings (research institutes, libraries, universities, etc.) followed the research data lifecycle scheme.

Three chapters present the following broader phases of the data life cycle: data creation, adding value and digital preservation, and access. In the introduction to the interviews we asked for clarification of the research data management planning and how actual data generation is processed. We explored the data types created, what measures were in place to achieve data quality, and the procedures in place and the criteria used when performing quality control. The next question we asked was about how data cleaning had been carried out, what other data transformations had been performed, and what documentation and metadata had been produced during the process. We asked about the importance of long-term digital preservation and tested the awareness about the best practices and technical solutions in the area of digital curation. We paid special attention to the possibilities of access and secondary use that were provided through support and collaboration of data creators and service providers in different phases of the data life cycle, where specialist service support, such as preservation and access, was organised for a wider range of users. We asked about barriers and motivations to secure open access to research data, both from the perspective of data creation and from the perspective of negotiating access to data, owned by others, when needed. We explored both the practices in the country and also experiences of those who used international data collections.

The first part of the report concludes with listing questions about accountability in tackling open data problems. During the conversations we assessed the awareness about problems discussed, we explored the dominant data sharing culture and habits, the level of cooperation in international settings, and readiness to accept changes. We encouraged the interviewees to offer solutions to problems as they see them, to propose the names of other stakeholders whose acts could help bring solutions, and asked about their image of an ideal open access system that would be both sustainable and of high quality. Which measures (among financial incentives, formal policies, expert solutions, and extended support service) they felt as most encouraging in opening research data?

We have also identified who already had taken some specific data support service role, and also explored the potential for taking more responsibilities in future scenarios. We gained understanding of the situation that Slovenian researchers and others involved were faced with in the scientific environment at that time, structural conditions that determined their work, and also explored how informed they were about various national and international initiatives and guidelines. We obtained information on what was their level of awareness about problems and solutions in open access to research data, and assessed their readiness to accept different solutions that might have been considered. The description of the conditions in different settings in the country as such formed a background for conceiving a future system of providing access to research data from public funding.

We tried to retain immediate experiences by keeping direct quotes in the end of each chapter of the previously published project working report. We summarized their statements and included them in the text where appropriate. The main groups of problems identified were again summarised in boxes in the end of each chapter to increase the visibility, together with the tentative solutions to the problems proposed. The selected problems and solutions suggested were further discussed throughout the length of the project. Among the key problems identified were the following:

- Research data have many different notions. The research data, which are typically created for the purpose of fulfilling the aims of a particular research project, are more neglected compared to reference collections. Research disciplines have large differences in approaches in research data management. Formalised protocols are not always consistently followed even in situations where researchers are aware of them.
- Research data custody and documentation appropriate to enable further use is mainly absent, and data are rarely handed over to dedicated research data centres. Big contrasts are presented in respecting metadata standard conventions.
- Rules and duties regarding research data management are mainly informal, and particular aspects of planning, creation, reformatting and publishing, and secondary use are usually not covered.
- There is a low level of research infrastructure services development to support data management.
- Data sharing culture is largely absent; there are different levels of awareness about the importance of data sharing among disciplines. The reward is not sufficient and other barriers for data sharing are quoted as an excuse for not participating. The funding agencies should set rules and policies, which would be clear and equal for all. They were explicitly requested in order to improve the situation, as well as to provide necessary conditions that would enable the compliance.

- The scientific evaluation is not established for quality research data as normal research project output for further use. The criteria and procedures to assure the evaluation and selection regarding digital preservation regimes and access arrangements are not elaborated. Additional work in research data processing for the purposes of open access needs to be additionally awarded.
- The problems are present in the lack of wider adoption of journal policies requiring research data deposit related to paper publication, and in strict research data citation format in publications.
- The copyright dilemmas, licencing, privacy protection, and other obstacles to free access to data which are related to ethical and juridical considerations, were often mentioned. Those dilemmas raise demand for professional advisory services.

As already explained, the report here only delivers a shortened presentation of the main findings of the first phase of the project. That phase was in its integrity presented in a separate project report titled “Current state in the area of research data management in Slovenia.”

What followed then in a report was an overview of many of the activities performed with the purpose to check and evaluate our preliminary findings in the context of the Slovenian research environment. Among other things, three workshops were dedicated to:

- The problems and solutions in the field of research data support services in Slovenia,
- Policies demanding publicly financed research data management planning,
- The prospects of advanced technologies supporting research data infrastructure building.

We tested the readiness of potential research data infrastructure and service developers to gain more recognition. The workshops confirmed that there is growing awareness in the research and scientific information services community about need to regulate research data management and access. Whereas the number of questions and challenges is growing, there is a lack of understanding about particular approaches to tackle the problems. But the existing islands of knowledge and experiences still could provide a good starting point and show potential for further development. A declaration of intent was already signed by the participants of the first workshop: [CREATING THE NATIONAL ENVIRONMENT OF OPEN ACCESS TO PUBLICLY FUNDED RESEARCH DATA IN SLOVENIA](#).¹ The first part of the report thus concludes with a list of groups and institutions identified as potential future service providers in the research data infrastructure in the country.

The second part of the monograph introduces many international initiatives and related published documents. We presented the active role of the EU Commission, which influences EU member states to get involved in common and coordinated activities in the research data management field. The documents and the initiatives are important for countries like Slovenia, mostly to encourage actions coordinated among countries which can lead to narrowing the gap in development in comparison to the most developed countries. Other internationally important initiatives and activities, going on in frameworks of various working groups and academic and professional associations, or expert connections such as documents of OECD, RCUK Common Principles, LERU Roadmap, LIBER, Science Europe, ALLEA etc., are presented in detail.

¹ See: <http://www.adp.fdv.uni-lj.si/odpp10D1/images/Recommendations.pdf>, 12. 2. 2014.

We presented a series of projects aimed at the improvement of research data infrastructure. Among those that are also oriented towards global integration of different research disciplines, we emphasised the GRDI 2020 project. We also mentioned a series of ongoing projects that are important for research data infrastructure upgrade at the European level from the ESFRI list, with emphasis on those that have representatives in the national environment as well.

In the next part of the report we gave an overview of currently active research data centres, also mentioning the criteria for evaluation of infrastructure service functionality. This includes evaluation of the suitability of a data centre facility for accepting certain data types, evaluation of expert support and advice in data management, etc.

It is then followed by the chapter in which activities in more advanced countries were presented. We started the overview with Great Britain as the country with the most valuable experience and tradition in tackling research data and, in many aspects, the leading example for other countries. We mentioned Germany and Nordic countries (e.g. the NordForsk group that deals with coordinated Nordic research policy development). Those are the countries that coordinate common activities in the Knowledge Exchange association in cooperation with Great Britain. The chapter ends with an overview of the situation in other countries around the world, with a special emphasis on the USA and Australia.

The international part of the report ends with presenting changes in scientific publishing related to research data access. We illustrate the situation by showing an overview of publishers and journals with already existing explicit requirements for open access to research data, which journal scientific publications are based on. The data citation environment is presented. The new type of so-called data journals was mentioned, which came into existence recently to promote further use of research data.

The last part of the publication presented the Action plan, which is formulated for the establishment of an open access system for research data from public funding as a final OPEN DATA project proposal to the Ministry of Education, Science and Sports. It also contained a proposal for the open research data policy for Slovenia as an appendix to the Action plan. The Action plan bullet points follow the main actors' roles, responsible for implementation of each of the target group aims:

- policy makers and funders, learned societies and other leading representatives of the scientific community, responsible for common policy and financial sustainability;
- research institutions, administration, actual researchers and information services working there as data creators, intermediate curators and data users;
- service providers of digital data curation, evaluation and access to data;
- scientific publishers.

We tried to optimise the establishment process of internationally integrated national open research data access systems by following the conclusion from current assessments of national conditions while taking into account numerous initiatives, conclusions and approaches around the globe.

The main problem that needs action among the leading scientific policy makers is to recognise research data as a public good with research potential beyond limits of particular research projects, which can lead to improved scientific development. The direction of activities planned

is to include open access to research data among the development priorities formulated in national strategic documents. The activity includes an approval of overarching national policy documents, stating responsibilities for research data access.

Tackling the problem of underdeveloped research data support infrastructure, the next goal set is to achieve more equal distribution of support infrastructure among scientific disciplines and to enable contacts and exchange with the international data centres. Common scientific information infrastructure should be extended to gain a comprehensive overview of the research data produced and on the access to them. Scientific funders need to support the sustainable operation of disciplinary data centres and secure coordination among other data service support providers to optimise end user service delivery. Training and support activities for researchers and support staff need to be provided.

The next field we need to concentrate on is to motivate researchers and research institutions to fulfil their part of the task of preparing research data for open access. Additional financial expenses during the project work need to be covered. Motivation could be reinforced also by recognising scientific achievements by including them in their personal bibliographies in the form of data published by a dedicated data centre, and by offering a comprehensive support infrastructure which reduces uncertainties and amount of work needed if carefully planned. The duty to include a data management plan in the project proposal is foreseen. The emphasis is also on making institutions accountable by their internal policy formulation and on building intermediate institutional support services.

The backbone of the future system of open access as imagined in the action plan is an establishment of data service research infrastructure accountable for delivering sustainable data support services. They should work on tasks that range from support in preparing DMP and in their fulfilment, in support for the registering, submission and acceptance of research data, further processing, evaluation and control of data and metadata, further supplemented by fulfilment of digital curation standards, and in organizing and promoting open access for data users. Disciplinary fields' autonomy and readiness to develop need to be respected in allowing further specialisation. Some disciplinary fields would need an intermediate period of time before the full functionality of a disciplinary research data centre is achieved, depending on the present state and capacity. Disciplinary data centres are built around specialists' knowledge about data types and related metadata standards and discipline-specific users' services. In order to develop efficient research data infrastructure, a layer of common central and integrating service needs to be built. The definition of that is a task of the coordination of disciplinary data centres. Optimal efficiency needs to be conceived in maximising the outcome for the scientific community. Sustainable services need to be directed into reducing additional burden on data creators.

To round up the whole system, clear and decisive requirements need to be imposed by a funder, following the principle of research data as a public good with few exceptions. The requirements, which impose a research data management plan in all publicly financed research projects, would fulfil that purpose. The research data as an output of a research project enters in the national registries of scientific production and counts in bibliographies that are integrated into the national scientific evaluation system (SICRIS). Open research data are the common obligation of all actors in a science, and cannot be the responsibility of only one of the segments, e.g., ethical principles alone. Instead it requires clearly elaborated requirements of all: individual researchers, their institutions and heads of the institutions, professional associations and

learned societies, other scientific community representatives, research data service providers, publishers, etc. National scientific policy and strategy need to provide conditions which are necessary to allow individuals and groups to participate more in a system of open research data access.

POVZETEK

Da bi preverili, kakšno je v Sloveniji stanje na področju raziskovalnih podatkov in oblikovali predloge za prihodnje načrte, ki bodo utemeljeni na realnih možnostih, je Ministrstvo za izobraževanje, znanost in šport v sodelovanju z ARRS na razpisu za predlagani Ciljni raziskovalni projekt: Odprti podatki – priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji izbralo Arhiv družboslovnih podatkov pri Univerzi v Ljubljani.

V mednarodnem okolju obstaja soglasje o pomenu vzpostavitve odprtega dostopa do raziskovalnih podatkov za razvoj znanosti. Zelo vplivna so Načela in usmeritve OECD, Priporočila Komisije EU in ostalih predstavnikov deležnikov v znanosti (LERU, Science Europe in njena delovna skupina Research Data, LIBER, ALLEA). S članstvom v mednarodni organizaciji OECD se je Slovenija zavezala, da bo sledila njenim smernicam na področju dostopa do raziskovalnih podatkov, financiranih iz javnih sredstev. Sprejela je konkretne obveznosti, vezane na roke in finančno ovrednotene, pri uresničevanju Priporočil Komisije EU in podpora pri njenih ostalih aktivnostih na področju odprtega dostopa do znanstvenih informacij. Resorno ministrstvo podpira z Evropo povezan razvoj raziskovalne infrastrukture po mednarodno usklajenem programu ESFRI, ki mu je ne glede na različna področja znanosti skupen izboljššan dostop do raziskovalnih podatkov. Sredstva iz strukturnih skladov bodo dobro naložena, če bodo namenjena za spodbujanje vzpostavljanja infrastrukture področnih podatkovnih središč.

Življenjski krog podatkov obsega široko področje ravnanja s podatki, od načrtovanja in ustvarjanja podatkov, pregledovanja, predelovanja, izbora, prve analize, evidentiranja, hrambe do dostopa in druge uporabe. Različne faze opravljajo akterji, ki imajo vsak svoje cilje, potrebe, navade, ki si med sabo lahko tudi nasprotujejo. Hkrati pa prihajajo z raznolikih znanstvenih področij. Raziskovalci, ki so največkrat omejeni na izkušnje lastnega raziskovalnega področja, pogosto ne vedo, kako širše misliti problematiko. In vendar so razprave o kakovostni hrambi in učinkoviti rabi podatkov v strokovni javnosti že dovolj dolgo in dovolj močno prisotne, da so vprašanja, dileme in odgovori nanje dokaj jasni. V okoljih z razvito raziskovalno infrastrukturo se pojavljajo domišljene rešitve, ki prispevajo k ustvarjanju kakovostnih podatkov in k učinkoviti izrabi potenciala, ki ga ti imajo za znanstveno skupnost. Na osnovi živahnih in plodnih razprav takšne izkušnje že dobivajo institucionalni okvir, katerega namen je, da se dobre prakse širijo tudi v okolja, v katerih so problemi manj opredeljeni in dobre prakse precej tuje. In sicer s ciljem, da bo za ustvarjanje podatkov, za njihovo hrambo in dostop poskrbljeno tako, da bo njihov potencial najbolje izkoriščen, s tem pa bodo tudi upravičena vložena sredstva.

Pričujoče poročilo v obliki monografije združuje rezultate celotnega projekta. Vsebuje povzetek prvega dela raziskovalnega projekta, se pravi analizo stanja in možnosti v državi, ki je bil kot izhodišče za nadaljnje aktivnosti že podrobno predstavljen kot delovno poročilo (Štebe in dr. 2012). Poročilo vključuje tudi pregled nadaljevalnih aktivnosti, ki so na naših tleh tekle v okviru projekta Odprti podatki, kot so posvetovanja in delavnice, namenjene iskanju in izpostavljanju primernih pristopov pri reševanju problemov, ki se pojavljajo ob vzpostavljanju okolja odprtega dostopa. V samostojnem delu poročila je nato prikazan obsežnejši pregled aktivnosti na področju vzpostavljanja podatkovnih storitev in podatkovnih infrastruktur v mednarodnem okolju, ki smo ga opravili v drugem delu projekta. Podrobneje smo preučili mednarodne pobude,

ki se ukvarjajo z raziskovalnimi podatki ter dobre prakse in izkušnje drugih držav oziroma posameznih ustanov. Določena izhodišča, ki se opirajo na mednarodno sceno, so za primerjavo kot možne rešitve nakazana že tudi v prvem delu, usmerjenem na zaznavo problemov doma.

Poročilo se konča s predstavitvijo predloga določil politik ravnanja z raziskovalnimi podatki in predloga akcijskega načrta za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov, financiranih z javnimi sredstvi, ki sta ključna rezultata projekta Odprti podatki. Poleg tega, da smo ugotavljali, kakšno je stanje v Sloveniji in primerjalno predstavili dogajanje v mednarodnem okolju, je bil eden od temeljnih ciljev projekta Odprti podatki poskus identifikacije nastavkov prihodnjega, bolj izpopolnjenega sistema dostopa do raziskovalnih podatkov. Ugotovitve in predloge takšnega sistema, prilagojenega na institucionalno in kadrovsko zaledje slovenskega okolja znanosti, smo kot sintezo predstavili v predlogu akcijskega načrta, ki tako temelji na realistični oceni možnosti in potencialov. Hkrati pa je zasnovan ambiciozno kot spodbuda in odpiranje možnosti za razvoj dejavnosti povsod tam, kjer obstaja pripravljenost, da prevzamejo naloge pri uvajanju sprememb na najbolj ustvarjalen, učinkovit in smiseln način v okviru svojih vlog in z njimi povezanih pristojnosti.

V akcijskem načrtu smo predvideli potrebno vmesno obdobje, da se na nekaterih področjih šele začne vzpostavljati dejavnost storitev podpore pri ravnanju z raziskovalnimi podatki. Potrebna je aktivna promocija idej in jasno izražena volja najodgovornejših za razvoj znanosti za njihovo uresničitev, da se soglasje o pomenu odprtega dostopa do raziskovalnih podatkov razširi v različna disciplinarna in institucionalna okolja. Spremembe je potrebno uvajati z odločnostjo in z izrecno podporo potrebnim aktivnostim, pri tem je odgovornost pri financerjih in oblikovalcih politik, hkrati pa je potrebno pridobiti podporo predstavnikov znanstvene skupnosti, kot so društva, sveti, akademija znanosti, vodstva inštitutov in drugi ugledni posamezniki. Spremembe morajo biti finančno učinkovite, pravične in usmerjene v očitno izboljšanje kakovosti raziskovanja. Spremembe pa so potrebne tudi za zagotavljanje zaupanja splošne javnosti v rezultate znanosti, ki se krepí z večjo preglednostjo ugotovitev, ki jih je mogoče preveriti na dostopnih podatkih.

Razvoj računalniških tehnologij že danes prinaša nove možnosti, ki kličejo k povezovanju, k deljenju in drugi rabi podatkov in drugih znanstvenih artefaktov znotraj pristopa »odprte znanosti«. Vendar pa same tehnološke rešitve še niso dovolj, saj moramo najprej doseči strokovni, politični in tudi družbeni dogovor o tem, kaj je vredno hraniti in deliti z drugimi ter kako to urediti, da bodo akterji ustrezno motivirani, nagrajani, zaščiteni.

Pregled poglavij

Izhodiščna poglavja osvetljujejo namen in cilje projekta Odprti podatki ter podlage, iz katerih ti cilji izhajajo. V prvem delu so kot uvod v pregled problematike raziskovalnih podatkov v Sloveniji danes na kratko predstavljeni primeri referenčnih raziskav, ki odpirajo vpogled v študije potreb, navad in praks raziskovalcev. Temu sledi niz poglavij, namenjenih izpostavljanju problemov in iskanju možnih rešitev na področju ravnanja z raziskovalnimi podatki v scenariju odprtega dostopa.

Prikaz stanja je zasnovan na intervjujih s slovenskimi raziskovalci in drugimi, ki imajo tako ali drugače opraviti z raziskovalnimi podatki. Da bi spoznali in preučili slovenske razmere, smo iskali odgovore iz prve roke, torej pri tistih akterjih, ki se s podatki ukvarjajo v svojem vsakodnevnem raziskovalnem procesu, bodisi kot ustvarjalci, uporabniki ali skrbniki. Zanimalo

nas je, s kakšnimi problemi se soočajo ustvarjalci podatkov in tisti, ki skrbijo za podatke, da bi ti bili širše dostopni. Ker gre za detekcijo problemov v kompleksnem okolju, je najustreznejša metoda zbiranja podatkov »od spodaj navzgor«. Tako smo opravili 22 polstrukturiranih intervjujev. Pogovore z intervjuvanci z različnih raziskovalnih področij in iz različnih delovnih okolij (raziskovalni inštituti, knjižnice, fakultete, državni uradi in drugo) smo vodili ob shemi, imenovani življenjski krog podatkov.

V treh poglavjih smo predstavili naslednje združene faze življenjskega kroga podatkov: ustvarjanje podatkov, dodana vrednost skupaj z vrednotenjem in prevzemom v digitalno skrbništvo ter iskanje, dostop in druga raba podatkov. V intervjujih smo ugotavljali, kako raziskovalci načrtujejo ravnanje s podatki in kako poteka samo ustvarjanje podatkov. Pri tem nas je zanimalo, kakšni tipi podatkov nastajajo, kako je poskrbljeno za kakovost podatkov, ali so oblikovani postopki in merila za doseganje kakovosti. Kako poskrbijo za čiščenje in preoblikovanje podatkov ter kako za ustvarjanje spremne dokumentacije in metapodatkov? Skušali smo izvedeti, kako pomembna je faza dolgotrajne digitalne hrambe in skrbništva, kakšne prakse in tehnične rešitve poznajo intervjuvanci. Posebno pozornost smo namenili zagotavljanju možnosti za drugo rabo, pri čemer sodelujejo tako ustvarjalci podatkov kot tudi ponudniki storitev, povezanih z življenjskim krogom podatkov, na primer storitev hrambe in organiziranja dostopa do podatkov za širši krog uporabnikov. Zanimalo nas je, kako je poskrbljeno za dostop do podatkov, kakšne ovire in zadržki stojijo na poti, ko raziskovalci skušajo priti do podatkov drugih; kakšne prakse obstajajo v Sloveniji in kakšne so izkušnje tistih, ki do podatkov dostopajo v mednarodnih zbirkah.

V zaključnem poglavju prvega dela poročila smo se posvetili odgovornosti v zvezi z reševanjem problemov na področju odprtih podatkov. Skozi pogovore smo identificirali, koliko so raziskovalci sploh ozaveščeni glede obravnavane problematike, zanimala nas je prevladujoča kultura, prakse, sodelovanje s tujino in pripravljenost na spremembe. Od intervjuvancev smo skušali izvedeti, kje vidijo možnosti za reševanje problemov, kdo bi lahko prispeval k rešitvam, kako bi moral biti urejen sistem, da bi lahko delovali čim bolj kakovostno in vzdržno; kateri ukrepi (finančni, v obliki politik, strokovne rešitve in podpora) bi spodbudili odpiranje raziskovalnih podatkov.

Zanimalo nas je, kdo so že sedaj nosilci posameznih nalog in kje so potenciali za v prihodnje. Poročilo je v tem delu poskus razumeti situacijo, v kateri se nahajajo slovenski raziskovalci in drugi vpleteni, v kakšnih strukturnih pogojih delujejo, koliko so seznanjeni s smernicami doma in v svetu, kako široko ali ozko je njihovo poznavanje problematike in kakšna je pripravljenost na sprejemanje rešitev. Opis dejanskega stanja, v danem času in prostoru, je bil zamišljen kot podlaga za oblikovanje prihodnje ureditve na področju ravnanja z raziskovalnimi podatki, financiranimi iz javnih sredstev.

Da bi ostali čim bolj v stiku s konkretnimi izkušnjami, smo dobesedne izjave intervjuvancev povzeli in jih smiselno vključili v besedilo, v dobesednem navedku pa jih navajamo na koncu poglavij že objavljenega delovnega poročila (Štebe in dr. 2012). Za boljšo preglednost smo zaznane probleme predstavili v okvirjih ob koncu poglavij in na istem mestu nakazali smer reševanja, izbrani problemi in rešitve pa so bili tudi predmet nadaljnjih aktivnosti. Med ključnimi problemi so se pokazali naslednji:

- Raziskovalni podatki so pojmovani zelo različno. Bolj se bo potrebno posvetiti raziskovalnim podatkom, ki nastajajo znotraj raziskovalnih projektov, za razliko od referenčnih zbirk. Raziskovalna področja se razlikujejo v pristopih pri ravnanju z raziskovalnimi podatki. Formalizirani protokoli se ne uporabljajo vedno povsem dosledno, četudi ponekod obstajajo in so raziskovalci z njimi seznanjeni. Pomanjkljiva je skrb za podatke in dokumentacijo za drugo rabo, redko je izročanje podatkov v pristojno podatkovno središče. Razlike so pri upoštevanju metapodatkovnih standardov.
- Odsotnost pravil glede ravnanja s podatki, od načrtovanja, ustvarjanja, preoblikovanja do objave in druge rabe.
- Nizka stopnja razvitosti raziskovalne infrastrukture storitev za podporo delu s podatki.
- Odsotnost kulture deljenja podatkov, različna stopnja zavedanja pomena po disciplinah, nezadostno nagrajevanje, druge ovire. Akterji pozivajo, naj financer za izboljšanje stanja določi jasna in za vse zavezujoča pravila ter omogoči njihovo izpolnjevanje.
- Znanstveno vrednotenje kakovostnih raziskovalnih podatkov za drugo rabo ni vzpostavljeno niti ni izoblikovanih meril in postopkov za zagotavljanje izbora, digitalne hrambe ter dostopa do raziskovalnih podatkov. Potrebno je nagrajevati dodatno delo pri pripravi podatkov za odprti dostop.
- Problemi so pri uveljavljanju obveznosti dostopa do raziskovalnih podatkov, povezanih z objavo, in pri navajanju raziskovalnih podatkov v publikacijah.
- Med ovirami za prosti dostop se običajno omenja vprašanja avtorstva, licenc in priznavanja pravic ter druge omejitve, ki izhajajo iz etičnih ali pravnih obvez v povezavi s podatki. Na tem področju je veliko nejasnosti in s tem povezanega povpraševanja po svetovalnih storitvah.

V pričujoči publikaciji, kot rečeno, podajamo skrajšan prikaz ugotovitev prve faze projekta, ki je bila v integralni obliki že predstavljena v ločenem projektne poročilu Opis stanja na področju raziskovalnih podatkov v Sloveniji.

Temu je sledilo več aktivnosti, s katerimi smo v slovenskem raziskovalnem okolju skušali izpostaviti aktualne teme in preveriti naše ugotovitve. Med drugim smo izvedli tri delavnice:

- Problemi in rešitve na področju podatkovnih storitev v Sloveniji,
- Politike ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi,
- Napredne tehnologije za vzpostavitev podatkovne infrastrukture v Sloveniji.

V okviru delavnic smo preizkušali tudi pripravljenost potencialnih razvijalcev prihodnjih podatkovnih storitev in infrastrukture, da se bolj uveljavijo. Delavnice so pokazale, da v raziskovalni skupnosti obstaja zavedanje o nujnosti urejanja področja raziskovalnih podatkov, narašča tudi število vprašanj in problemov s tega področja, medtem ko je poznavanje rešitev redkeje prisotno. Obstoječi otoki znanja in izkušenj pa vendarle ponujajo potencial za naprej. Prvi del poročila zaključujemo s pregledom skupin oziroma ustanov, ki smo jih zaznali kot potencialne nosilce prihodnjih podatkovnih storitev in infrastruktur.

Drugi del poročila začnemo s pregledom številnih mednarodnih iniciativ in objavljenih dokumentov. Predstavimo aktivno vlogo Evropske komisije, ki države članice neposredno spodbuja k skupnim in usklajenim aktivnostim na področju ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi. Dokumenti in pobude so pomembni za Slovenijo, ker spodbujajo koordinirano delovanje z drugimi državami in lahko prispevajo k zmanjšanju

zaostanka za najrazvitejšimi. V poročilu so izčrpno predstavljene tudi druge mednarodno pomembne pobude in aktivnosti, ki se odvijajo znotraj različnih akademskih združenj ali strokovnih povezav, kot so dokumenti OECD, Skupna izjava RCUK, deklaracije LERU, LIBER, Science Europe, ALLEA itd.

Predstavljamo tudi vrsto projektov s področja razvoja podatkovne infrastrukture. Med tistimi, ki si prizadevajo za povezovanje na globalni ravni na različnih znanstvenih področjih, smo izpostavili projekt GRDI 2020. Navedli smo več aktualnih projektov s seznama ESFRI, pomembnih tudi za izgradnjo podatkovne infrastrukture na evropski ravni, ki so s svojimi predstavniki zastopani v slovenskem prostoru.

V nadaljevanju poročila predstavljamo okolje delujočih podatkovnih središč skupaj s kriteriji, ki jih je potrebno upoštevati pri vzpostavljanju storitev in infrastrukture, npr. presoja primernosti podatkovnega središča za prevzem določene vrste raziskovalnih podatkov, strokovna podpora in svetovanje pri delu s podatki. Sledi poglavje, v katerem predstavljamo aktivnosti v posameznih državah, ki sodijo med naprednejše. Pregled začnemo z Veliko Britanijo, ki ima najbogatejše izkušnje in tradicijo ravnanja z raziskovalnimi podatki in je v veliko primerih zgled drugim državam. Omenimo tudi Nemčijo in nordijske države (npr. skupino NordForsk, ki pripravlja skupne nordijske raziskovalne politike), ki so skupaj z Veliko Britanijo del koordinacije Knowledge Exchange. Poglavje zaključimo s prekomorskimi državami, med katerimi smo največ pozornosti namenili ZDA in Avstraliji.

Mednarodni del poročila zaključimo s predstavitvijo zahtev po spremembah na področju znanstvenih objav. V ilustracijo ponudimo pregled vrste založb in revij, ki so s politikami že opredelile obveznost odprtega dostopa do podatkov, na katerih temeljijo objave v njihovih publikacijah. Nekatere založbe pa so vzpostavile svojevrstne publikacije, t. i. podatkovne revije, katerih namen je objava opisa podatkovnega vira z namenom spodbuditi možnost druge rabe raziskovalnih podatkov.

Publikacijo zaključujemo s predstavitvijo Akcijskega načrta in priloženega osnutka politik. Akcijski načrt je po točkah sestavljen tako, da so v njem opredeljeni glavni nosilci odgovornosti za posamezni sklop ciljev: to so oblikovalci politik in financerji, nato raziskovalne ustanove in na njih delujoči raziskovalci kot ustvarjalci, vmesni skrbniki ter uporabniki podatkov, izvajalci storitev prevzema digitalnega skrbništva za podatke, njihovega vrednotenja in dostopa ter na koncu tudi založniki in revije. Sledeč izsledkom analiz stanja pri nas in ob upoštevanju številnih pobud, ugotovitev in pristopov po svetu smo prišli do zaključkov, kako se čim bolj optimalno lotiti izgradnje mednarodno integriranega nacionalnega sistema odprtega dostopa do podatkov.

Pod točko odgovornosti MIZŠ in najvišjih predstavnikov znanstvene skupnosti in profesionalnih združenj smo obravnavali splošni problem premalo izkoriščenega potenciala odprtega dostopa do raziskovalnih podatkov kot javnega dobrega za napredovanje znanosti. Predlagane aktivnosti na tem področju gredo v smer vključitve odprtega dostopa kot razvojne prioritete v državne strateške dokumente. Aktivnosti vključujejo tudi sprejem krovne nacionalne politike na tem področju.

Izhajajoč iz razmeroma nizke stopnje razvitosti podatkovne infrastrukture, je naslednji cilj zagotoviti enakomernejšo zastopanost različnih področij znanosti pri podporni dejavnosti in vzpostavitvi povezave z mednarodnimi podatkovnimi središči. Z nacionalnega vidika je pomembno pridobiti celovit pregled nad ustvarjenimi raziskovalnimi podatki in njihovim

dostopom. Z vidika financerja je pomembna skrb za vzdržno delovanje področnih podatkovnih središč in koordinacija delovanja z ostalimi ponudniki s podatki povezanih storitev ter usposabljanje na različnih področjih dela s podatki.

Naslednje področje je motiviranje raziskovalcev in raziskovalnih ustanov za izpolnjevanje nalog odprtega dostopa do podatkov s financiranjem dodatnega dela, s priznavanjem znanstvenega prispevka v obliki objavljenih raziskovalnih podatkov v osebnih bibliografijah ter izgradnja celovite podporne storitvene dejavnosti. Predvidena je obveznost priprave Načrtov ravnanja z raziskovalnimi podatki. Poudarek je tudi na odgovornosti ustanov, da vzpostavijo svoje notranje politike in podporne storitve.

Hrbtenica prihodnjega sistema odprtega dostopa v akcijskem načrtu je vzpostavitev odgovornosti izvajalcev podatkovnih storitev v obliki raziskovalne infrastrukture za izvajanje nalog za podporo pri pripravi Načrta ravnanja z raziskovalnimi podatki (v nadaljevanju NRRP) in njihovem izvajanju, evidentiranju in podpori pri prevzemu raziskovalnih podatkov od ustvarjalcev, nadaljnji obravnavi podatkov, pregledu in dopolnitvi metapodatkov skladno s standardi digitalnega skrbništva ter zagotavljanju odprtega dostopa za uporabnike. Potrebno je upoštevati avtonomijo in pripravljenost za delovanje znotraj področij ved. Glede na različne potrebe in trenutni položaj glede delovanja podpornih dejavnosti pri ravnanju s podatki imajo področja na voljo prehodno obdobje, v katerem lahko izoblikujejo dejavnost področnih podatkovnih središč. Ta se vzpostavljajo ob poznavanju posebnosti glede tipov podatkov ter njim prilagojenih metapodatkov in področnih storitev. Razvoj podatkovne infrastrukture pa za učinkovito delovanje potrebuje tudi določene osrednje skupne in povezovalne storitve, ki se razvijajo v koordinaciji področnih podatkovnih središč. V vsaki fazi je potrebno upoštevati optimalno učinkovitost delovanja, ki se presoja z vidika koristi za znanstveno skupnost. Storitve je potrebno oblikovati tako, da prevzamejo bremena z ramen ustvarjalcev podatkov.

Potrebno pa je postaviti tudi zelo odločne zahteve s strani financerja v obliki politike obveznosti, da so, razen izjemoma, raziskovalni podatki javno dobro, zaradi česar je potrebno skrbno ravnanje v vseh fazah življenjskega kroga podatkov. Temu je namenjena obveznost priprave načrtov ravnanja s podatki, ki se evidentirajo in ki jih upoštevajo v povezavi z sistemom znanstvenega vrednotenja rezultatov projekta. Odprti raziskovalni podatki so skupna odgovornost vseh akterjev v znanosti, ki ne more biti prepuščena samo enem segmentu, npr. etičnim načelom, pač pa zahteva jasno opredeljene obveznosti tako posameznih raziskovalcev, njihovih ustanov in vodstev, strokovnih in znanstvenih društev ter drugih predstavnikov znanstvene skupnosti, izvajalcev s podatki povezanih storitev, založnikov. Naloga znanstvene politike je zagotavljanje potrebnih pogojev, da se posamezniki in skupine intenzivneje vključijo v sistem odprtega dostopa do podatkov.

KAZALO

OPEN DATA – ACTION PLAN FOR THE ESTABLISHMENT OF A SYSTEM OF OPEN ACCESS TO PUBLICLY FUNDED RESEARCH DATA IN SLOVENIA	III
SUMMARY OF A REPORT	III
POVZETEK.....	XI
KAZALO	XVII
ŠIRŠI NAMEN IN CILJI PROJEKTA ODPRTI PODATKI.....	1
<i>VSEBINSKI CILJI PROJEKTA</i>	<i>2</i>
<i>POMEN DEKLARACIJE IN NAČEL OECD.....</i>	<i>3</i>
PRIKAZ STANJA V SLOVENIJI	7
IZHODIŠČA ZA PRIKAZ STANJA NA PODROČJU RAZISKOVALNIH PODATKOV V SLOVENIJI.....	7
<i>ŠTUDIJE POTREB RAZISKOVALCEV</i>	<i>8</i>
<i>TIPIČNE OVIRE NA POTI DO ODPRTEGA DOSTOPA.....</i>	<i>10</i>
<i>POMEN IZOBRAŽEVANJA IN USPOSABLJANJA ZA RAVNANJE S PODATKI.....</i>	<i>12</i>
RAZISKOVALNA METODA	15
<i>OSNOVNI PODATKI O VZORCU.....</i>	<i>17</i>
<i>VPRAŠALNIK.....</i>	<i>19</i>
<i>ŽIVLJENJSKI KROG PODATKOV.....</i>	<i>20</i>
<i>FAZE ŽIVLJENJSKEGA KROGA PODATKOV</i>	<i>23</i>
PRIKAZ REZULTATOV V FAZI USTVARJANJA PODATKOV	25
<i>TIPI PODATKOV.....</i>	<i>25</i>
<i>VPRAŠANJA KAKOVOSTI PODATKOV V POGLEDIH INTERVJUANCEV</i>	<i>29</i>
<i>POMEN KAKOVOSTNIH PODATKOV IN METAPODATKOV</i>	<i>30</i>
<i>POSTOPKI IN MERILA ZA ZAGOTAVLJANJE KAKOVOSTI V FAZI USTVARJANJA</i>	<i>31</i>
<i>IZBOR PODATKOV NA PODLAGI KAKOVOSTI.....</i>	<i>34</i>
<i>IZBOR PODATKOV V RAZLIČNE REŽIME DOSTOPA</i>	<i>35</i>
PRIKAZ REZULTATOV V FAZI DODANA VREDNOST	37
<i>ČIŠČENJE, TRANSFORMACIJA PODATKOV IN METAPODATKI</i>	<i>39</i>
<i>SPRAVILO.....</i>	<i>42</i>
PRIKAZ REZULTATOV V FAZI ISKANJE, DOSTOP (OBJAVA), DRUGA RABA.....	47
<i>DOSTOP DO PODATKOV</i>	<i>47</i>
<i>ZADRŽKI IN OVIRE POVEZANE Z DRUGIMI FAZAMI ŽIVLJENJSKEGA KROGA PODATKOV</i>	<i>49</i>
<i>VREDNOTENJE PODATKOV.....</i>	<i>55</i>
ODGOVORNOST V ZVEZI Z REŠEVANJEM PROBLEMOV	60
<i>POTREBA PO KROVNI POLITIKI.....</i>	<i>60</i>
<i>USTREZNO VREDNOTENJE PODATKOVNIH VIROV</i>	<i>61</i>
<i>FINANČNA SREDSTVA.....</i>	<i>62</i>
<i>DOBER SISTEM IZVEDBE HRAMBE IN DOSTOPA DO PODATKOV</i>	<i>63</i>
ZAKLJUČKI ANALIZE POLOŽAJA ODPRTIH RAZISKOVALNIH PODATKOV V SLOVENIJI	66
POSVETOVALNE DEJAVNOSTI PROJEKTA ODPRTI PODATKI V SLOVENIJI	70
<i>DELAVNICE PROJEKTA ODPRTI PODATKI</i>	<i>70</i>
<i>POSVETI S STROKOVNJAKI V OKVIRU PROJEKTA ODPRTI PODATKI</i>	<i>78</i>
<i>NASTAVKI ZA RAZVOJ PODATKOVNIH STORITEV IN INFRASTRUKTURE V SLOVENIJI</i>	<i>79</i>

MEDNARODNI PREGLED DOGAJANJA NA PODROČJU DOSTOPA DO RAZISKOVALNIH PODATKOV ...	86
IZHODIŠČA.....	86
POBUDA EVROPSKE KOMISIJE.....	89
POBUDE MEDNARODNIH STROKOVNIH ZDRUŽENJ IN AKADEMSKIH POVEZAV	92
STROKOVNA ZDRUŽENJA IN PROJEKTI S PODROČJA RAZVOJA PODATKOVNE INFRASTRUKTURE	97
<i>SPLOŠEN PREGLED SKUPIN IN PROJEKTOV</i>	97
<i>PREDLOG EVROPSKE STRATEGIJE NA PODROČJU PODATKOVNIH VIROV</i>	101
<i>POVEZOVALNE IN GLOBALNE PODATKOVNE INFRASTRUKTURE</i>	103
<i>PROJEKTI GRADNJE IN IZPOLNJEVANJA PODROČNE RAZISKOVALNE PODATKOVNE</i>	
<i>INFRASTRUKTURE</i>	104
<i>PRESOJA PRIMERNIH PODATKOVNIH SREDIŠČ</i>	109
SREDIŠČA ZA STROKOVNO PODPORO IN SVETOVANJE PRI DELU Z RAZISKOVALNIMI PODATKI	112
PREGLED PO DRŽAVAH	116
RAZISKOVALNI PODATKI IN SPREMLJAJOČA INFRASTRUKTURA ZNANSTVENEGA PUBLICIRANJA....	129
<i>VREDNOTENJE RAZISKOVALNIH PODATKOV</i>	129
<i>POLITIKE ZNANSTVENIH REVIJ IN UREDNIŠTEV</i>	133
ZAKLJUČKI MEDNARODNEGA PREGLEDA.....	137
AKCIJSKI NAČRT	142
<i>ODGOVORNOST MIZŠ IN NAJVIŠJIH PREDSTAVNIKOV ZNANSTVENE SKUPNOSTI IN PROFESIONALNIH</i>	
<i>ZDRUŽENJ</i>	142
<i>RAZVOJ PODATKOVNE INFRASTRUKTURE</i>	148
<i>MOTIVIRANJE RAZISKOVALCEV IN RAZISKOVALNIH USTANOV ZA IZPOLNJEVANJE NALOG ODPRTEGA</i>	
<i>DOSTOPA DO PODATKOV</i>	152
<i>VZDRŽEVANJE STORITEV ZA PODPORO PRI DELU Z RAZISKOVALNIMI PODATKI, VREDNOTENJE, IZBOR</i>	
<i>IN PREVZEM, DIGITALNO SKRBNIŠTVO IN ZAGOTAVLJANJE DOSTOPA</i>	158
<i>PRISPEVEK IZDAJATELJEV ZNANSTVENIH PUBLIKACIJ ZA ODGOVORNO RAVNANJE Z RAZISKOVALNIMI</i>	
<i>PODATKI</i>	163
LITERATURA.....	165
PRILOGE	177
<i>PRILOGA 1: Prošnja za sodelovanje v projektu Odprti podatki</i>	177
<i>PRILOGA 2: Priporočilno pismo za sodelovanje v projektu Odprti podatki</i>	179
<i>PRILOGA 3: Izjavo o strinjanju s pogoji sodelovanja</i>	180
<i>PRILOGA 4: Vprašalnik</i>	181
<i>PRILOGA 5: Priporočila udeležencev delavnice Problemi in rešitve na področju podatkovnih storitev v</i>	
<i>Sloveniji</i>	196
<i>PRILOGA 6: Določila politike odprtega dostopa do raziskovalnih podatkov v Sloveniji. Priloga k</i>	
<i>akcijskemu načrtu</i>	200

ŠIRŠI NAMEN IN CILJI PROJEKTA ODPRTI PODATKI

Hitro naraščanje obsega raziskovalnih podatkov in razvoj tehnologij, ki omogočajo lahek dostop preko interneta, zahtevajo spremembo krovnih politik glede pravic in obveznosti na področju podatkovne izmenjave v znanstveni skupnosti. Ena od »gurujk« odprte znanosti Christine Borgman (2010) ugotavlja, da je prav v zadnjem času zanimanje za problematiko podatkov izjemno naraslo. Imamo celo vrsto aktualnih poljudnih člankov, ki predstavljajo izzive in možnosti podatkovne izmenjave. V reviji *Nature*² so leta 2008 posebno številko namenili problematiki »velikih podatkov«. Zanimalo jih je, kaj za sodobno znanost pomenijo velike in kompleksne podatkovne strukture ter kako se bodo z novimi izzivi spopadle inštitucije in obstoječa kultura na področju znanosti, saj so še zmeraj ukoreninjene v predelektronski dobi. Leto kasneje so v članku *Data's shameful neglect*³ obravnavali pomen hrambe in dostopa do podatkov, ki sta predpogoj napredka v znanosti, skrb za to pa naložili financerjem s pripravo strategij in podporo politik obveznosti, pa tudi raziskovalcem, ki naj skupaj poskrbijo za strojno in programsko opremo ter za ustanove, ki bodo zagotavljale infrastrukturne storitve za hrambo in dostop do podatkov na način podatkovnih knjižnic in podatkovnih arhivov. Zelo jasno se zavzemajo za to, da mora postati načrtovano ravnanje s podatki kot osnovni temelj sestavni del vseh korakov v znanosti. O podatkovni problematiki govorijo tudi bolj splošno usmerjeni časopisi in revije. Npr. ameriški *WIRED*⁴ preizkuša zamisli o spremembi paradigme pri pridobivanju znanja iz velike količine podatkov. *Economist*⁵ iz leta 2010 govori o vse večjem naraščanju količine podatkov, pa naj gre za podatke v velikih trgovskih verigah, knjižnicah ali na družabnih omrežjih, pri tem pa izpostavlja pomen poklicev prihodnosti, ki se bodo ukvarjali s podatki, kot so podatkovni znanstvenik, podatkovni znanstveni knjižničar ipd. Potrebe po podatkovnih veščinah in znanjih niso več stvar prihodnosti in samo specialistov, pač pa so že del vsakdana znanstvenikov, čeprav bo ponekod sprememba v spremljajoči miselnosti vidna šele z menjavo generacij. Naraščanje podatkov vodi v t. i. »četrto paradigmo« podatkovno usmerjenih znanosti (Hey in dr. 2009) in botruje nastanku novih področij, kot so »računalniška biologija« ali pa »digitalna humanistika« (Borgman 2010).

Cilj projekta Odprti podatki je, izhajajoč iz predstavljenega, jasen: pomagati pri oblikovanju rešitev za čim boljši izkoristek potenciala, ki ga nosijo raziskovalni podatki, nastali z javnimi sredstvi. V bolj odprtem in čim širšem dostopu do podatkov so posredne koristi tako za tiste, ki sodelujejo v znanstvenih projektih, ki ustvarjajo ali uporabljajo podatke, kot tudi za širšo javnost. Veliko se govori o potencialu odprte znanosti, tudi ljudske oz. državljanske znanosti, kjer lahko vsakdo črpa iz zakladnice znanja in prispeva vanjo. Prav nekatere inovacije na področju interneta, npr. sodelovanje skupnosti, so prinesle nove obete za večjo vključenost vseh zainteresiranih pri znanstvenih projektih. Odprta znanost, tudi v povezavi z načeli OECD,⁶ je

² Glej: <http://www.nature.com/nature/journal/v455/n7209/full/455001a.html>, 7. 8. 2012.

³ Glej: <http://www.nature.com/nature/journal/v461/n7261/pdf/461145a.pdf>, 7. 8. 2012.

⁴ Glej: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory, 7. 8. 2012.

⁵ Glej: <http://www.economist.com/node/15557443/print>, 7. 8. 2012.

⁶ *Principles and Guidelines for Access to Research Data from Public Funding* (2007). Dokument bomo v poročilu na kratko imenovali Načela OECD ali samo Načela.

blagovna znamka za različne skupinske pobude in gibanja, ki nasprotujejo vsakovrstnim »nepravičnim« omejitvam dostopa do znanstvenih objav, podatkov in orodij.

V okviru projekta smo želeli ugotoviti, kaj imamo v zamekih v Sloveniji na področju dostopa do raziskovalnih podatkov za namen nadaljnje znanstvene obravnave, kakšne so sedanje možnosti in potencial za prihodnje. In sicer je bil cilj tudi, da začnemo vzpostavljati interesne povezave in da pobudam pomagamo odpirati prostor za uveljavitev. Prepričanje o koristih odprtega dostopa se namreč vse bolj uveljavlja, tako da je pravi čas, da se projektom, povezanim z njimi, omogoči finančne, kadrovske in organizacijske pogoje za resno delovanje.

Osnovna načela so bolj ko ne preprosta. Vsakdo se bo strinjal, da je odprt dostop do podatkov zagotovilo bolj tehtnih zaključkov. Preverjamo lahko ugotovitve drug drugega, z večanjem dostopa se poveča množina odkritij, kadar zaradi kompleksnosti podatkov in njihovega združevanja spoznavni potencial ni izčrpan z ločenimi analizami, zlasti pa je kolektivno delo na podatkih odlična spodbuda za dialog, medsebojno kritiko in dopolnjevanje analiz sorodnih problemov. To pa je – popperjansko vzeto – zagotovilo napredovanja spoznanj. Na splošno lahko rečemo, da se poveča tako kakovost kot količina spoznanj za isto ceno (BRTF-SDPA 2010). Prednost odprtega dostopa do podatkov je tudi v tem, da se omogočijo povezave sicer ločenih podatkov ter povezave podatkov in znanstvenih objav v obliki besedil, tovrstna dodana vrednost pa zopet bogati njihov analitični potencial. Pogoj je, da so podatki za nadaljnjo uporabo ustrezno pripravljeni in primerno opisani. Ustrezati morajo principom objavljanja podatkov, oblikovanim po analogiji z objavami člankov. Objava podatkov (v pooblaščenem podatkovnem arhivu ali v podatkovni reviji) namreč zahteva ustrezno dokazilo o kakovosti ter dokumentacijo o nastanku in vsebini, ki je potrebna za njihovo smiselno drugo uporabo.

Glede na to samoumevnost o koristih je presenetljivo, zakaj podatki niso javni, še posebej če so nastali s podporo javnih sredstev. Velikokrat predstavlja oviro odsotnost kulture izmenjave podatkov, ki jo opažamo kljub deklariranim etičnim načelom. Ena pomembnih ovir so vsekakor dodatni napor in čas ter znanje, ki je potrebno za pripravo podatkov in njihovo objavo. Zato je potrebna podporna infrastruktura, ki predstavlja središče za posredovanje potrebnega znanja, omogoča generiranje spodbud s pripravo podatkov ter na koncu zagotavlja ustrezno mesto, kamor se podatki lahko izročijo za dolgoročno dostopnost z zagotovitvijo storitev trajnega skrbništva. To so lahko podatkovna središča in arhivi, specializirani po posameznih disciplinah, nacionalni in mednarodno povezani, za določeno vrsto podatkov in v posameznih fazah ravnanja s podatki so to lahko tudi splošni repozitoriji posameznih ustanov ali začasni podatkovni arhivi posameznih projektov ali posameznikov. Potrebne so tudi od zgoraj generirane spodbude, tako finančne kot moralne, ter v politikah zapisane obveznosti za raziskovalce, da bodo pripravljeni izročati podatke.

Če se zdaj vrnemo k izhodiščnemu vprašanju: kje v Sloveniji imamo vsaj v zamekih dostop do izvornih podatkov, ki bi se jih dalo uporabljati za nadaljnje analize? Medicina, naravoslovje, biologija, okolje, družboslovje in humanistika? Če da, kako odprt je dostop? Če ne, zakaj ne? In kaj lahko naredimo? V pričujočem projektu CRP oblikovanja akcijskega načrta dostopa do raziskovalnih podatkov iz raziskav, ki se financirajo iz javnih virov, smo razčiščevali prav ta vprašanja.

VSEBINSKI CILJI PROJEKTA

Če povzamemo, so bili vsebinski cilji projekta Odprti podatki sledeči. Najprej, ugotoviti, kakšno je trenutno stanje na področju ravnanja z raziskovalnimi podatki v Sloveniji. Preveriti, kdo so nosilci dejavnosti in posameznih nalog, ter identificirati nastavke prihodnjega, bolj izpopolnjenega sistema dostopa do raziskovalnih podatkov, skladnega z mednarodnim okoljem in sprejetimi obveznostmi države. Ugotavljali smo, kateri so glavni problemi, s katerimi se soočajo ustvarjalci podatkov in tisti, ki dolgoročno skrbijo za podatke, poskušali identificirati vrste podatkov in kaj se s podatki dogaja od nastanka pa do morebitne druge rabe,⁷ kakšne so povezave ter medsebojna izmenjava podatkov in storitev doma in v tujini. Zlasti nas je zanimalo, kakšna so medsebojna pričakovanja, kako posamezni nosilci vlog razumejo svoje obveznosti do financerjev in medsebojne odvisnosti, kadar nastopajo, včasih tudi v isti osebi, kot ustvarjalci, izvajalci storitev, povezanih s podatki, ter njihovi uporabniki. Ta spoznanja smo koristno uporabili pri snovanju predloga politik in akcijskega načrta, saj je pomembno, da ta temelji na konkretnih zgledih iz našega okolja in da so predlagane rešitve take, da obetajo široko podporo pri izvedbi.

Drugi, vzporedni cilj projekta je bil mobilizacija vpletene in zainteresirane javnosti. Odprti dostop je področje, ki predstavlja inovacijo. Da bi sprejeli novost, bo potrebno spremeniti način razmišljanja in delovanja. V okviru projekta smo si prizadevali za vključevanje tematik, ki zadevajo potencial prihodnje ureditve: to je bolj transparenten in široko zajemajoč sistem dostopa do raziskovalnih podatkov, z namenom promocije načel odprtega dostopa. Že danes lahko identificiramo nekatere šampione, ki vlečejo s svojimi zgledi. Te posameznike in skupine smo še posebej izpostavili ter izkoristili njihovo pripravljenost in angažiranost za uveljavljanje novosti na področju organiziranja podatkov in njihovega bolj učinkovitega izkoriščanja. S tem smo želeli vzbuditi ambicije, pripravljenost za delo in sodelovanje tudi pri tistih, pretežno ustvarjalcih podatkov, ki se do sedaj še niso odločili vstopiti na področje opravljanja storitev in svetovanja glede ravnanja s podatki, za kar je potrebno dodatno usposabljanje. Glede na trenutni ugled poklica »podatkovnega znanstvenika«, »podatkovnega arhivista« ali »podatkovnega knjižničarja« ter težavne karijerne odločitve (Swan in Brown 2008b) bodo tudi tukaj potrebne spodbude.

Tretji cilj projekta, načrtovanje strategije in politik, smo v zadnji fazi projekta zastavili tako, da smo v čim večji meri upoštevali in uporabili obstoječe nastavke dejavnosti ter povezave in zglede iz tujine. Zadnja faza projekta je obsegala poskus doseganja soglasja med nosilci vlog in včasih tudi nasprotujočih si interesov glede strateških ciljev in njihove implementacije, kar smo opravili skozi posvete in delavnice v okviru projekta. Ob pregledu aktivnosti v mednarodnem okolju so podana strokovna izhodišča, ki se jih da izkoristiti za konkretno ocenjevanje stroškov, kadrovskih potreb in dodatnega dela pri vzpostavljanju in nadgradnji robustne podatkovne infrastrukture, ko se bo tovrstna področna infrastruktura načrtovala in postavljala. Podani so zgledi, kako se v sodobnem okolju znanstvenega vrednotenja lahko zagotavlja nagrajevanje in motivacija za prispevek podatkov v odprti dostop, iz tega so izpeljani predlogi za oblikovanje politik in upoštevanje strokovnih standardov. Vse to bo prispevalo k zagotavljanju realističnega načrtovanja ter optimalnega izkoriščanja kadrovskih, finančnih in podatkovnih virov.

⁷ V tem poročilu uporabljamo pojem "druga raba", ki je različica termina "sekundarna" ali "ponovna" uporaba, analiza podatkov.

Potek projekta je sledil zastavljenim ciljem, pri čemer smo spodbujali tudi uporabo družabnih omrežij kot novega načina komuniciranja tudi na področju znanosti, ki ga tudi sami uporabljamo. Že na začetku smo ustvarili [blog](#)⁸ o projektu, ki je podlaga za medsebojno informiranje in morebitno sodelovanje v razpravah, z ambicijo, da začnemo graditi skupnost in povezovati zainteresirano javnost in angažirane posameznike. Blog je bil odprt za posamezne gostujoče prispevke, sicer pa je vseboval komentarje aktualnih dogodkov. Pričujoči projekt ima bistveno vlogo pri seznanjanju s koristmi in deljenju pozitivnih izkušenj odprtega dostopa do podatkov. Skupaj s spletno stranjo *Openaccess.Si*, v okviru katere v poglavju Odprti podatki s prispevki sodelujemo izvajalci projekta, je sam projekt predstavljal forum, namenjen promociji načel in podpori pri njihovem uresničevanju. Tudi v nadaljevanju, po formalnem zaključku projekta pričakujemo odzive in nove prispevke, saj je vzpostavitev tega sistema skupna naloga in odgovornost.

Posamezne faze izvedbe projekta, kot so predstavljene v pričujočem poročilu, sledijo zgoraj zastavljenim ciljem. Začeli bomo z ugotavljanjem stanja: kdo so nosilci dejavnosti oziroma njenih zametkov, kakšne so vloge, medsebojna pričakovanja in kako so opredeljene obveznosti. Gre za relacije med ustvarjalci podatkov in ustvarjalci raziskovalnih politik, financerji ter tistimi, ki že opravljajo ali bi lahko opravljali storitve prevzema, dolgotrajnega shranjevanja in dostopa do podatkov, seveda upošteva potrebe, probleme in interese sedanjih ali prihodnjih uporabnikov podatkov in z njimi povezanih storitev.

V drugi fazi projekta, katere povzetek sledi v nadaljevanju, smo podali pregled dobrih praks glede dostopa do raziskovalnih podatkov ter številnih pobud in načrtov za reševanje problemov v mednarodnem okolju.

Ugotovitve glede stanja doma smo presojali z vidika dobrih praks in modelov v tujini ter iz tega izpeljali izhodiščne predloge za reševanje zaznanih problemov. Temu je sledila organizirana razprava o podanih predlogih kot preizkus sprejemljivosti za različne javnosti in nosilce vlog, tako na različnih raziskovalnih področjih kot tudi v različnih fazah življenjskega kroga podatkov.⁹ Cilj je bil doseganje strokovnega in političnega soglasja glede predlaganih rešitev, čemur sledi končno poročilo o izsledkih projekta in predlog za akcijski načrt.

Pri uresničevanju zastavljenega smo uporabljali različne metode za vključevanje ciljnih skupin in za zbiranje informacij, kot so intervjuji in posebej pripravljena poročila o delu posameznih skupin in projektov, kombinirali smo tudi prispevke in komentarje na družabnih omrežjih in organizirali posamezna posvetovanja na način delovnih srečanj po izbranih tematikah. Ob tem smo si prizadevali za intenzivnejše vključevanje v mednarodna omrežja, situacijo in načrte v mednarodnem okolju pa smo spremljali tudi s študijem literature, zlasti poročil o tekočih projektih in predstavitev z aktualnih konferenc.

POMEN DEKLARACIJE IN NAČEL OECD

Pri vlaganju v razvoj sistema odprtega dostopa do podatkov je potrebno pazljivo načrtovanje na ravni strateških odločitev nacionalne porabe sredstev za dodatno delo, ki je potrebno. Načela OECD, ki so podlaga pričujočega projekta, pokrivajo vse najpomembnejše vidike načrtovanja: tehnološke, institucionalne, upravljavske, finančne, proračunske, pravne, vidike nacionalnih

⁸ Glej: <http://www.adp.fdy.uni-lj.si/blog/2010/novo/projekt-odprti-podatki/#axzz2t1EnCDIn>, 11.2.2014

⁹ Več o tem glej poglavje Življenjski krog podatkov.

politik ter vidik kulture, navad in praks, in kot taka predstavljajo odlično izhodišče za razvoj. Sama Načela niso tako podrobno razdelana, da bi nudila konkretne aktivnosti za uresničitev zastavljenih ciljev. Poudarjajo namreč celostno zasnovo in obravnavo posameznih vidikov ob nujnosti prilagajanja nacionalnim razmeram. Slednje je naloga pričujočega projekta. Zagotavljanje uspeha akcijskega načrta je dolgoročno povezano s tem, kako izhodišča razumejo in sprejemajo različni nosilci vlog v sistemu dostopa do podatkov. Zato smo si že v okviru izvedbe projekta prizadevali pridobiti podporo tako na strani nosilcev moči, ki lahko vplivajo na spremembe, kot na strani nosilcev storitev in njihovih uporabnikov, da spremembe posvojijo. Izhodišče in temeljni dokument za pripravo pregleda stanja v Sloveniji na področju dostopa do odprtih podatkov predstavljajo Načela OECD, na katera se opirajo tudi kasnejša priporočila Evropske komisije državam članicam. Dokument izvira iz Deklaracije,¹⁰ v kateri so leta 2004 članice OECD pripoznale in potrdile potrebo po dostopu do raziskovalnih podatkov. Temeljni cilj je izboljšati učinkovitost in izkoristek nacionalnih in globalnega znanstvenega sistema. Boljši dostop do podatkov pomaga maksimizirati raziskovalni potencial novih digitalnih tehnologij in omrežij ter zagotavlja večji izkoristek javnih sredstev, vloženih v raziskovanje. Pomeni pa tudi potencialni vir znanja, ki je potrebno pri soočanju z izzivi, ki pretijo človeštvu, so zapisali v Načela. Dokument je zasnovan tako, da v temelju določa izhodišča in pogoje, s katerimi bi bilo mogoče zastavljene cilje uresničiti na čim bolj ekonomičen način in ne da bi ovirali razvoj s težavnimi obveznostmi in pravili ali nalagali nove stroške (OECD 2007, 13).

Načela so pomembna kot dokument, ki naj države članice OECD in druge vodi pri oblikovanju strategij, politik in načrtov odprtega dostopa do podatkov. V preambuli dokumenta je jasno opredeljeno, v čem so smisel, pomen in prednosti odprtega dostopa. Dejansko težo tega dokumenta predstavlja soglasje o pomenu cilja najširšega deljenja raziskovalnih podatkov, ki v obliki sprejetih načel države zunaj in znotraj OECD zavezuje, da po svojih zmožnostih težijo k njegovi uresničitvi. Pomen tega dokumenta se tako kaže v novo zasnovanih državnih ali disciplinarnih strategijah in politikah, ki se nanj sklicujejo. Primer implementacije načel OECD, ob eksplicitnem sklicevanju nanj, je denimo dokument Economic & Social Research Council (ESRC) v Veliki Britaniji. Slednji v politikah dostopa do podatkov zahteva, da je v raziskavah, financiranih z javnimi sredstvi, ravnanje s podatki načrtovano že v prijavi projekta, zahtevajo pa tudi dejansko ravnanje s podatki v času izvedbe raziskovalnih projektov s ciljem, da bo ob koncu projekta zagotovljen dostop do podatkov (*ESRC Research Data Policy* 2010).

Načela OECD predstavljajo pomemben temeljni dokument, ki pomaga državam z različno stopnjo razvitosti, da se začno zavedati problemov in iskati rešitve. Namenjena so financerjem in politikom, raziskovalcem, raziskovalnim ustanovam in državnim raziskovalnim agencijam, da bi razumeli pomen odprtega dostopa do raziskovalnih podatkov. Dokument ne ponuja receptov in univerzalnih rešitev, prinaša pa celostni pristop k reševanju »problema« odprtih podatkov. zgoraj omenjena Christine Borgman (2011, 2) opozarja na kompleksnost vprašanj »podatkov, raziskovalnih praks, inovacij, spodbud, ekonomike, intelektualne lastnine in javnih politik, ki so povezane s problemom («*conundrum*«, op. a.). deljenja podatkov – z zahtevnim problemom, ki postavlja izzive«.

S tem želimo reči, da se zavedamo težavnosti tega problema, da želimo pokazati na njegove razsežnosti in kompleksnost, preiskati nekatere ponujene rešitve kot zgled, da pa je na koncu

¹⁰ Declaration on Access to Research Data from Public Funding (OECD, 2004).

iskanje in doseganje rešitev stvar akterjev, ki so vpleteni v problem, njihovega znanja in motiviranosti ter pogojev, ki omogočajo ravnanje, skladno z omenjenimi načeli.

Projekt Odprti podatki – Priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji smo zasnovali na predpostavki, da je odprti dostop do raziskovalnih podatkov nekaj dobrega. Že v izhodiščih smo se zavedali dejstva, da je vprašanje za veliko vpletenih nekaj novega in da bo treba poiskati področja, storitve in zglede, ki bodo tudi tiste z manj izkušnjami in znanjem spodbudili k uresničevanju odprtega dostopa.

Želeli smo zasnovati sistem, ki bo pravičen, učinkovit, ki ne bo povečeval bremen ustvarjalcev in bo izpostavil možnosti integracije storitev, pravil, tehnologij, orodij, hkrati pa ohranil posebnosti pristopov in narave podatkov ter upošteval potrebe znanstvene skupnosti. Kajti kot pravi večkrat citirana avtorica Margaret Henty (v Feijen 2011): »S strani raziskovalcev se največkrat pojavi zahteva po storitvah, ki so na voljo, ko jih potrebujejo, a so oblikovane tako, da ne posegajo v ustvarjalno delo, ki je bistvo raziskovalnega procesa.« Storitve podatkovne infrastrukture naj bi bile po možnosti zasnovane tako, da jih opaziš samo tedaj, kadar npr. zaradi tehničnih težav pride do njihovega izpada delovanja, se pravi, nemoteče vtkane v rutino raziskovalnega vsakdana.

Pomembno izhodišče za doseganje projektnih ciljev je podpora Ministrstva za izobraževanje, znanost in šport, ki razume in podpira načela OECD. Predstavniki Ministrstva so sodelovali pri različnih aktivnostih projekta Odprti podatki, zlasti pri promociji in pri oblikovanju ustrezne politike zahtev glede načrtov ravnanja s podatki, nagrajevanja dodatnih naporov ustvarjalcev podatkov ter oblikovanja okvirov za vzpostavitev storitvenih infrastrukturnih dejavnosti.

Ministrstvo je z razpisom projekta presodilo, da za pripravo strategije na tem področju potrebuje strokovna izhodišča. Za pripravo strokovnih izhodišč in osnutka akcijskega načrta so potrebne konkretne izkušnje, ki jih ima izvajalec projekta z izvajanjem storitev dostopa do raziskovalnih podatkov, potrebno je poglobljeno poznavanje situacije v drugih državah in vzpostavljen stik z delujočimi podatkovnimi infrastrukturami na določenem znanstvenem področju. Zlasti pa je potreben akcijsko usmerjen pregled stanja in možnosti, potencialov, ki temelji na vključevanju vseh relevantnih akterjev že v fazi priprave prihodnje strategije. Ključni akterji so, prvič, pripadniki znanstvene skupnosti, ki jih je potrebno seznaniti z možnostmi odprtega dostopa do znanstvenih podatkov in prednostmi tega za izboljšanje znanstvenih odkritij, po drugi strani pa jih je potrebno seznaniti z nalogami, ki jih je potrebno opraviti, da bodo prednosti prišle do izraza. Pripadniki znanstvene skupnosti so po tem, ko se seznanijo z načeli odprtega dostopa, v večini primerov pripravljene ta načela podpreti in se dejavno vključiti v njihovo izpeljavo. Drugič, tu so akterji in organizacije, ki opravljajo storitve infrastrukturnih podatkovnih središč – organiziranih disciplinarno ali v okviru ustanov. Ti zagotavljajo ustrezno podporo ustvarjalcem podatkov, bodisi s svetovanjem in usmerjanjem ter usposabljanjem za ravnanje s podatki v fazi ustvarjanja in priprave ali pa z opravljanjem storitev po prevzemu podatkov v svoje skrbništvo. Tretji akter pa so izvajalci in oblikovalci znanstvenih politik, ki morajo zagotoviti ustrezen okvir predpisov in motivacije za izpeljavo nalog zagotavljanja odprtega dostopa, zlasti pa podati ustrezne finančne in organizacijskega podlage za delo. Naloga pričujočega projekta je bila tako pripraviti predlog, ki bo vzdržal s strokovnega vidika, vključeval pa bo tudi zaveze in konsenz glede rešitev s strani omenjenih akterjev.

V okviru pričujočega projekta smo se dotaknili vseh vidikov iz Načel OECD, ki izhajajo iz ključnega pojmovanja odprtosti dostopa do podatkov kot javne dobrine (Kaul, Grunberg in Stern 1999), ter izpostavili pregled dobrih praks, izkušenj in zgledov. S tem smo preizkusili pripravljenost našega sistema za izpolnjevanje teh načel, in sicer z namenom, da se prilagodijo realnosti trenutnega stanja glede finančnih, organizacijskih, strokovnih (znanja) in tehnoloških pogojev. Kot je že zahteva besedila razpisa, je bila naloga projekta »pregled situacije v Sloveniji in strokovna uskladitev med relevantnimi akterji v Sloveniji, vključno z nosilci raziskovalne politike, glede ključnih korakov v nadaljnjem razvoju (pregled nad zametki storitev odprtega dostopa do podatkov v Sloveniji, organiziranje razprave in povezava med strokovnjaki iz ključnih raziskovalnih okolij zaradi izmenjave informacij, znanj in tehnoloških rešitev za povezovanje storitev na nacionalnem nivoju in mednarodno)«.

Iz Načel OECD posebej povzemamo eno od ugotovitev:

Ključna ovira pri doseganju maksimalne vrednosti na podlagi vlaganj v raziskovalne podatke je v posameznih vejah znanosti odsotnost načrtovanja in izvajanja ustreznih nalog, namenjenih dokumentaciji in arhiviranju podatkovnih datotek. Pri načrtovanju projektov in programov je potrebno na vseh ravneh neposredno izpostaviti problematiko podatkov v najzgodnejši fazi, zato da je mogoče že pri pripravi projekta predvideti finančne in tehnične zahteve izhodiščne priprave in nadaljnega izvajanja skrbništva nad podatki. Poskrbeti je treba za podporo pobudam in za podporo strokovnemu usposabljanju na vseh področjih ravnanja s podatki. (OECD 2007, 19)

Ob tem, ko je izpostavljena tudi potreba po usposabljanju za delo s podatki,¹¹ je poudarek na načrtovanju in zagotavljanju pogojev za izvajanje sprotnega skrbništva nad podatki v okviru tekočih raziskovalnih projektov, kar v večini primerov predstavlja minimalen dodatni napor, ki prinaša koristi že sproti pri notranji organizaciji podatkov v okviru projekta. Če se to odloži na čas po koncu projekta, pa je lahko že zamujeno, saj včasih stvari za nazaj ni mogoče spreminjati ali reproducirati. To je smer, v katero gredo politike glede zahtev dostopa do raziskovalnih podatkov po svetu, očitno v soglasju z Načeli OECD. Aktualna je pobuda financerjev znanosti z naslovom *Dissemination and Sharing of Research Results*¹² v ZDA in Veliki Britaniji, da v prijavih na raziskovalne projekte zahtevajo načrt ravnanja s podatki in dostop do podatkov (Jones 2009). Načrt je kot del prijave podvržen recenzentskemu postopku. Britanski Digital Curation Centre (med drugimi podobnimi) pripravlja pripomoček *DMP Online*¹³ in navodila za pripravo načrta *Management Plans*.¹⁴ Britanski arhiv družboslovnih podatkov (UK Data Archive) ponuja gradiva in svetovanje že pri pripravah projektov in v zgodnji fazi njihovega izvajanja.¹⁵ Značilnost tega modela je, da financer, obenem ko zahteva načrt ravnanja s podatki, v projektu zagotovi tudi sredstva za to delo in predvidi sankcije v primeru neizpolnjevanja. Očitno obstaja vse več soglasja tako na ravni raziskovalnih politik med državami in med različnimi financerji znotraj držav, da tovrstne rešitve predstavljajo optimalno porazdelitev odgovornosti in zahtev med različnimi vlogami v krogu ravnanja s podatki.

¹¹ Problematika je lepo povzeta v enem od poročil *Digital Curation Center* (DCC) (Swan in Brown 2008).

¹² Glej: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>, 7. 8. 2012.

¹³ Glej: <http://dmponline.hatii.arts.gla.ac.uk/>, 7. 8. 2012.

¹⁴ Glej: <http://www.dcc.ac.uk/resources/data-management-plans>, 7. 8. 2012.

¹⁵ *CREATE & MANAGE DATA*, <http://www.data-archive.ac.uk/create-manage/>, 7. 8. 2012.

PRIKAZ STANJA V SLOVENIJI

IZHODIŠČA ZA PRIKAZ STANJA NA PODROČJU RAZISKOVALNIH PODATKOV V SLOVENIJI

Zagotavljanje storitev odprtega dostopa do raziskovalnih podatkov zahteva veliko sodelovanja. Obstaja vrsta modelov, po katerih je mogoče razvijati odprti dostop, ob tem pa je pomembno vprašanje, kako so opredeljene strategije in politike države na področju znanosti, kakšne so strategije raziskovalnih ustanov in ponudnikov podatkovnih storitev. Tudi različna disciplinarna področja imajo različne pristope in zahteve. En sam recept verjetno ne ustreza vsem. Poleg ohranjanja posebnosti raziskovalnih področij pa se je na ravni oblikovanja nacionalnih in mednarodnih strategij potrebno izogniti podvajanju infrastrukturnih podatkovnih storitev in omogočiti njihovo združljivost, npr. pri agregiranju za namene iskanja in dostopa do gradiv s strani uporabnikov. Analiza trenutnih razmer na področju ustvarjanja in zagotavljanja dostopa do raziskovalnih podatkov je ponudila uvide glede prednostnih nalog in njihovih potencialnih izvajalcev, upošteva dosežane aktivnosti in kapacitete.

V analizi in prikazu stanja smo se oprli na model vlog, pravic, odgovornosti in povezav posameznih nosilcev v fazah življenjskega cikla podatkov (primerjaj Green in Gutman 2006; Lyon 2007; Ruussalepp 2008; Jones 2009). Namen tovrstnega prikaza je poudariti medsebojne odvisnosti, pri katerih izpolnjevanje nalog enega pogojuje možnosti delovanja drugega, izpostaviti dobre rešitve, ki bi se lahko z enega prenesle na druga področja, opozoriti na pomanjkljivosti in nakazati smernice za njihovo odpravo. Podporo takšnemu pristopu k presoji delovanja in načrtovanju storitev podatkovnega dostopa najdemo tudi v Načelih OECD in dokumentih britanskega Research Councils (RCUK).¹⁶ Tudi zadnja poročila iz ZDA (Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age 2009; Interagency Working Group on Digital Data 2009) poudarjajo celovito skrb za kakovostno znanstveno produkcijo podatkov ter skrb za shranjevanje in uporabo, s katero so opredeljene vse vloge, tako za strokovnjake s področja ravnanja s podatki kot za raziskovalce in financirje raziskav.

Pregled je narejen s predpostavko o potrebi po močnih področnih podatkovnih središčih tam, kjer obstajajo možnosti za njihovo oblikovanje in delovanje. Za enkrat razen Arhiva družboslovnih podatkov takega osrednjega podatkovnega središča v Sloveniji na drugih področjih nimamo. Sicer so mogoče tudi druge rešitve, kot je model samoarhiviranja ali pa model repozitorijev digitalnih gradiv ustanov. Naša teza gre v smer, da je obstoj področnih podatkovnih središč hrbtenica celovitega sistema zagotavljanja dostopa do raziskovalnih podatkov, tako da bomo v prikazu stanja proučevali možnosti in nastavke za oblikovanje tovrstnih osrednjih področnih podatkovnih storitev v posameznih okoljih. Ostale storitve, kot so

¹⁶ Prvi princip skupnega dokumenta Znanstvenih svetov v Veliki Britaniji (RCUK) se glasi: "Vloga in odgovornost raziskovalcev, raziskovalnih ustanov in financirjev mora biti opredeljena čim bolj jasno /.../ za zagotovitev zavedanja in izpolnjevanja odgovornosti ustvarjalcev in uporabnikov raziskovalnih podatkov" (RIN 2008, 8). Podobno predvidevajo že omenjena Načela OECD o glede preglednosti, formalne odgovornosti in profesionalizma (2007).

institucionalni arhivi in mednarodne podatkovne infrastrukture, so lahko dopolnili ali vmesne rešitve, medtem ko je robusten in vzdržen sistem sodelovanja raziskovalcev in nacionalno dejavnih področnih podatkovnih središč edini zmožen učinkovito zagotoviti trajno hrambo, dostop in izkoriščanje najbolj kakovostnih podatkovnih virov za znanost.

Green in Gutmann (2006) na podlagi situacije v ZDA, kjer je sicer zelo razvita mreža digitalnih repozitorijev po ustanovah, izpostavljata prednosti področnih podatkovnih središč, kot je Interuniversity Consortium for Political and Social Research (ICPSR), ki delujejo že dolga desetletja ter so prepoznavni in uveljavljeni znotraj znanstvene skupnosti. Namreč, repozitoriji ustanov nastopijo svojo vlogo pri koncu raziskovalnega cikla: pretežno prevzemajo že v končni obliki pripravljene publikacije in druga gradiva, nudijo malo ali nič podpore glede priprave podatkov, metapodatkov in njihove izrabe, prevzem je pretežno avtomatiziran in gradiva spremljajo metapodatki brez nujnih podrobnosti za razumevanje konteksta nastanka in vsebinskih lastnosti gradiv, z minimalnim preverjanjem na ravni skladnosti formatov. V Sloveniji sicer močno zaostajamo tudi na področju repozitorijev ustanov. Za razliko od teh pa so področna podatkovna središča osredotočena na izpolnjevanje potreb posameznih disciplinarnih področij tako glede vsebin kot oblik in načina dostopa do podatkovnih gradiv, s svojimi storitvami pa se raztezajo v vse faze življenjskega kroga podatkov, od nastanka, hrambe, objavljanja do druge rabe. Celota teh storitev pa je poudarjeno tematika pričujočega pregleda.

Obstoj močnega področnega podatkovnega središča, kot je ADP, zagotavlja kritično maso ohranjanja in razvijanja strokovnega znanja za ravnanje s podatki. To je lahko dobra podlaga za koordiniran nastanek novih področnih podatkovnih središč za področja, na katerih tega še ni, saj je prilagojen slovenskim razmeram v znanosti. Pomembno je, da na področju dostopa do raziskovalnih podatkov v Sloveniji ne razvijamo nečesa povsem novega, pač pa upoštevamo obstoječe izkušnje nekaterih že delujočih infrastrukturnih projektov. Treba je ohraniti sedanje stanje in ga nadgraditi v smeri kakovosti delovanja v okviru celovitega sistema medsebojno prepletenih vlog in odgovornosti. V okviru izvajanja pričujočega projekta je bila naša naloga prenesti izkušnje na druga področja, na katerih so podatkovne storitve še v zametkih, ter zagotoviti prilagajanje potrebam uporabnikov in posebnostim posameznih disciplin, poskrbeti pa tudi za integracijo strokovnih pristopov, orodij in storitev tam, kjer je to bolj optimalno z vidika porabe sredstev, človeških naporov in ravni zagotavljanja kakovosti storitev (Štebe 2010).

ŠTUDIJE POTREB RAZISKOVALCEV

Preden se lotimo podrobnega prikaza rezultatov naše lastne študije, naj opozorimo, da obstaja vrsta podobnih študij potreb in možnosti za oblikovanje storitev odprtih podatkov, ki smo jih upoštevali pri zasnovi in interpretaciji rezultatov. Martin Feijen (2011) je v poročilu z naslovom *What researchers want* pregledno povzel rezultate različnih raziskav po svetu. Zanimale so ga potrebe raziskovalcev na področju ravnanja s podatki in možnosti dostopa do njih. Posebej izpostavlja študije, nastale v okviru projektov *Parse* (2009), *e-IRG* (2009), in druge, ki izpostavljajo zgoraj omenjeno ugotovitev o nujnosti ravnanja in priprave podatkov že v teku trajanja projektov (Feijen 2011, 4). Pravne težave in morebitna zloraba ali napačna raba podatkov so med najpogosteje omenjenimi ovirami (Feijen 2011, 11). Zlasti pravne dileme so strokovno področje splošne narave, ki je tipično izven obsega znanj, ki jih posedujejo raziskovalci specialisti, zato bi bilo potrebno posebno svetovanje in poenotena navodila s strani

ustanov, ki bi pomagale premostiti te ovire. Glede druge, sicer tudi v ostalih študijah večkrat omenjene ovire, to je skrbi ustvarjalcev zaradi možnosti napačne rabe podatkov, če ti pridejo v roke nezadostno seznanjenim uporabnikom, se nakazuje potreba po večji vključenosti raziskovalcev ustvarjalcev podatkov v usposabljanje in svetovanje drugim uporabnikom pri uporabi njihovih podatkov. To so že naloge podatkovnih storitev, ki bi jih glede na svoje kompetence najbolje lahko prevzeli ustvarjalci podatkov v sodelovanju s podatkovnimi središči. Model za izmenjavo znanj pa bi lahko bil oblikovanje (npr. tudi virtualne) skupnosti uporabnikov določene vrste podatkov, ki jih v nekaterih državah gojijo na način srečevanj uporabnikov in ustvarjalcev določenih podatkov.¹⁷

Pri pripravi nacionalnih strategij in akcijskega načrta so koristne tudi ugotovitve, povzete iz poročila *UK Research Data Service* (RIN/BL 2010), ki prav tako sloni na raziskavi o različnih akterjih v življenjskem krogu podatkov. Kar potrebujejo raziskovalci, je proaktivno vključevanje storitev knjižnic in raziskovalne administracije. Temeljna potreba je čim večja prilagoditev politik, zahtev in storitev obstoječim, pogosto neformalnim omrežjem in potekom raziskovalnih projektov, ki lahko vključujejo orodja in znanja, vezana na discipline. Se pravi, da je politika vključevanja in podpore načeloma najbolje sprejeta, če je vezana na uspešna nastajajoča orodja in prakse znotraj raziskovalnih skupnosti samih. Izhajajoč iz teh ugotovitev, smo iskali in potrjevali obstoj podobnih iniciativ tudi v okviru naših znanstvenih okolij, s tem da zagovarjamo tezo, da bi bilo potrebno del politik na tem področju usmeriti ravno v spodbujanje inovativnih pristopov podpore podatkovnim storitvam in skupnostnim orodij (kot so orodja *Taverna workflow*),¹⁸ ki nastajajo pri sodelovanju znotraj discipline ali pa v okviru interdisciplinarnega sodelovanja znanstvenikov.¹⁹ Na splošno je v večini študij izražena velika potreba po dodatnem usposabljanju delujočih znanstvenikov v veščinah ravnanja s podatki in digitalne hrambe. Na tem mestu izpostavljamo točke, ki so bile vodilo našega pristopa: potreba po prednostnem vzpostavljanju podatkovnih storitev glede na povpraševanje dajalcev podatkov, potreba po postopnem uvajanju novosti za discipline in področja, kjer je očitna takojšnja korist (npr. sodelovanje in modeli podatkovne izmenjave iz družboslovja v druge discipline, kot je medicina), pragmatične in postopne širitve storitev ter previdnost pri postavljanju ciljev z vidika izvedljivosti. Pomembno je namreč dobro predstaviti ideje ter koristi in zagotoviti potrebno podporo in sodelovanje, zlasti na institucionalni ravni (Feijen 2011, 23–25, glej tudi Williford in Henry 2012, pogl. 2.4.2).

V zgoraj omenjenem projektu *UK Research Data Service* so si zadali, da bodo skušali oceniti izvedljivost in stroške razvoja ter vzdrževanja podatkovnih storitev za univerze v Veliki Britaniji, raziskava pa je zanimiva tudi zaradi kombiniranja uporabe več metod zbiranja podatkov. Anketnemu vprašalniku so na vsaki od izbranih univerz sledile fokusne skupine. Namen skupinskih pogovorov je bil ugotoviti, kakšne zahteve imajo univerzitetni raziskovalci glede hrambe podatkov – zato so bolj poglobljeno analizirali odgovore, pridobljene s pomočjo spletnega vprašalnika. Poleg ankete in fokusnih skupin je raziskava obsegala še 37 individualnih intervjujev z raziskovalci z Univerze v Oxfordu in delavnico, na kateri so razpravljali o ugotovitvah, pridobljenih z intervjuji.

¹⁷ *The Labour Force Survey: past, present and future*: <http://ukdataservice.ac.uk/news-and-events/eventsitem/?id=3576>, 7. 8. 2012.

¹⁸ Glej: <http://www.taverna.org.uk/>, 7. 8. 2012.

¹⁹ Primerjaj projekte digitalne humanistike: <http://www.diggingintodata.org/>, 7. 8. 2012.

TIPIČNE OVIRE NA POTI DO ODPRTEGA DOSTOPA

Obstaja vrsta študij, ki situacije v državi predstavljajo na način anketne poizvedbe o problemih in podpori odprtemu dostopu. Tak primer prestavlja priložnostna anketna raziskava z naslovom *Open Access to and Reuse of Research Data – The State of the Art in Finland* (Kuula in Borg 2008), ki se prav tako kot naš projekt sklicuje na podmene Načel OECD. Ugotavljata, da je druga raba podatkov v dokumentih OECD privzeta kot nekaj samo po sebi umevnega. Realno stanje glede podatkovne izmenjave in podporne raziskovalne infrastrukture v manj razvitih okoljih je seveda ravno nasprotno. V tej študiji so ugotovili, da je osnovni problem, ki ga je bilo mogoče pričakovati tudi v okviru naše raziskave, majhna ozaveščenost in seznanjenost z možnostmi, ki jih ponuja odprti dostop do raziskovalnih podatkov. Raziskava je pokazala na tipične ovire in pomisleke. Raziskovalce skrbi nenamerna zloraba podatkov in posledične napake. Pravica glede intelektualnega lastništva podatkov ni jasno opredeljena, ovira je tekmovalnost za akademske položaje in financiranje. Ugotovili so tudi odsotnost politik, navodil in naporov na ravni ustanov, kar raziskovalce ustvarjalce podatkov in uporabnike pušča v negotovosti. To kaže na širši problem, da so pri vzpostavitvi in delovanju sistema odprtega dostopa do podatkov pomembne vmesne podporne strukture. To je obvezna tema za področje implementacije prihodnjih nacionalnih politik, ker je očitno, da bi se reševanja problemov pri delu s podatki lahko lotili že na ravni raziskovalnih ustanov, kjer podatki nastajajo. Tam bi se opravila tudi selekcija pomembnih gradiv, oblikovale bi se spodbude in ozko specializirani svetovalni centri ter nekatere vmesne storitve v skrbi za podatke, preden se jih preda v pooblaščen področna podatkovna središča.

Kuula in Borg (2008, 27) posebej izpostavljata vrsto koristi, ki sledijo implementaciji odprtega dostopa in so jih anketiranci prepoznali kot take:

- finančno: bolj učinkovita izraba raziskovalnih podatkov vodi v maksimiranje finančnega izkoristka glede na velike vložke v ustvarjanje podatkov;
- znanstveno: odpiranje znanosti, kontinuirano ustvarjanje podatkov za zagotovitev njihove ažurnosti, ponovljivost in kontrola raziskav, večje sodelovanje v raziskovalni skupnosti;
- raziskovalnoetično: npr. upoštevanje pravic sodelujočih posameznikov med ustvarjanjem in rabo podatkov, kar je zlasti izpostavljeno pri medicinskih in družboslovnih raziskavah;
- pravno: upoštevanje prava med ustvarjanjem, rabo in hrambo podatkov;
- demokratično: natančna opredelitev položaja in odgovornosti raziskovalnih subjektov, raziskovalcev, raziskovalnih organizacij in financerjev;
- enakost: promocija enakega dostopa do raziskovalnih podatkov;
- družbeno: izboljšanje kakovosti in stroškovne učinkovitosti informacijskih virov;
- politično: OECD obvezuje države članice.

Med predlogi za implementacijo odprtega dostopa je tudi ustrezno vrednotenje raziskovalnih podatkov, se pravi, da bi raziskovalni podatki dosegli podoben status, kot ga imajo raziskovalne objave. Raziskovalci so večinoma naklonjeni temu, da se podatke obravnava kot samostojni izdelek, zanje je to dodatna motivacija za zgledno pripravo podatkov in nagrada za vloženo znanje, ki ga sami po sebi vsebujejo raziskovalni podatki, neodvisno od publikacije ali v povezavi z njo, saj predstavljajo dodatni vir informacij, ki jih publikacija ne vsebuje. Raziskovalna publikacija je samorazlagalna, medtem ko raziskovalni podatki zahtevajo dodatno razlago in

dokumentacijo, zlasti če niso ustrezno arhivirani. Za razliko od objav lahko podatki vključujejo občutljive in zaupne informacije. Raba podatkov običajno zahteva dovoljenje, publikacija ne. Pri publikacijah so lastništvo in avtorske pravice običajno razjasnjene, pri podatkih so pogosto nejasne. Objave so v znanstveni skupnosti dostopne prosto ali proti plačilu. Pri podatkih gre za več stopenj odprtosti (od popolne odprtosti do tega, da so podatki zaprti zaradi posebnosti, povezanih s specifičnim podatkovnim virom). Publikacija je že pripravljena za drugo rabo, podatki terjajo dodatne postopke (Kuula in Borg 2008, 28). Tema je zelo aktualna in zasluži podrobnejšo obravnavo, saj je povezana s specifičnimi vprašanji izbora in znanstvene kakovosti podatkov, ustrezne infrastrukture za registracijo identifikatorjev podatkov in podporo citiranju podatkov. Prav zato predstavlja enega od izzivov za reševanje tudi v okviru preiskovanja pogojev in možnosti v Sloveniji.²⁰

Izhajajoč iz Načel OECD so se vprašanja odprtih podatkov lotili tudi v Čilu. Njihov osnovni cilj je bil urediti, okrepiti in zagotoviti odprti dostop do podatkov in nasploh do raziskovalnih informacij, zbranih in ustvarjenih v Čilu, financiranih z javnimi sredstvi, z namenom, da spodbudijo izmenjavo podatkov, ustvarjenih v državi, in da ohranijo raziskovalno podatkovno zapuščino. Pod drobnogled so vzeli institucionalni okvir, človeški kapital, zapuščino in politike.²¹ Poročilo je zanimivo, ker kaže tipične ugotovitve za okolja, ki nimajo razvite raziskovalne podatkovne infrastrukture, kot je po vseh predvidevanjih tudi naše. Sami smo pri pregledu situacije izhajali iz tovrstnih pričakovanj; nekatera so tudi eksplicitno potrjena.

Britanski RIN, Joint Information Systems Committee (JISC) in Natural Environment Research Council (NERC) so leta 2007 naročili raziskavo, ki se je osredotočila na navade in prakse raziskovalcev v odnosu do raziskovalnih podatkov. Posebej jih je zanimalo, kako je razširjena kultura deljenja podatkov, kakšne so infrastrukturne omejitve pri objavah podatkov, kakšen je učinek pobud, katere politike spodbujajo objave podatkov in kakšna je nasploh pripravljenost za objavljanje podatkovnih zbirk. Ugotovili so, da se področja med seboj precej razlikujejo in da je glavni problem pomanjkanje usklajenosti med disciplinami, v smislu ravnanja, politik in potreb. Swan in Brown (2007, 13) sta zapisala: »Kakršno koli rešitev problema bomo opredelili, bo zahteve in prakse potrebno prilagoditi vsaki posamezni raziskovalni disciplini.«

Posebno pozornost so med drugim posvetili tudi vprašanju, zakaj nekateri raziskovalci objavljajo svoje podatke, drugi pa ne. Raziskovalci, ki objavljajo podatke, so kot spodbude poleg ostalih dejavnikov navedli upoštevanje in pričakovanje s strani financerja (Swan in Brown 2007, 25–26). Med tipičnimi ovirami, zaradi katerih ne objavljajo podatkov, so navedli pomanjkanje izkušenj, časa in sredstev za delo s podatki, pravne ali etične ovire, nepoznavanje možnosti arhiviranja, tekmovalnost, strah pred zlorabo ali napačno rabo, prepričanje, da nihče ne potrebuje njihovih podatkov, omejene nagrade ali pa nagrad sploh ni (Swan in Brown 2007, 26–28).

Prisotnost podobnih razlogov, ki ovirajo deljenje podatkov, bomo po pričakovanjih odkrili tudi med slovenskimi raziskovalci. Verjetno pa bomo ugotovili tudi, da v naših razmerah raziskovalci ne navajajo veliko razlogov, ki bi bili v prid deljenju podatkov, kar je vsekakor posledica odsotnosti delujočih primerov in pozitivnih zgledov, ki bi bili širše uveljavljeni.

²⁰ Primerjaj opis ciljev projekta *Open AIRE Plus* v mednarodnem pregledu in druge iniciative v zvezi z infrastrukturo za citiranje podatkov.

²¹ Izvlečke te raziskave je Patricia Muñoz Palma predstavila na *Board on International Scientific Organizations* aprila 2011 v Washingtonu, http://sites.nationalacademies.org/PGA/biso/PGA_061353, 7. 8. 2012.

POMEN IZOBRAŽEVANJA IN USPOSABLJANJA ZA RAVNANJE S PODATKI

Za premagovanje ovir na poti k odprtim podatkom je še kako pomembno poznavanje navad, praks in kulture raziskovalcev.²² Poleg kakovostnih podatkov in kakovostnih postopkov za ravnanje s podatki je pomemben del pozornosti treba usmeriti tudi na usposabljanje strokovnjakov, ki ravnajo z njimi. Posamezne ustanove že govorijo o rešitvah. Na Univerzi v Oxfordu so v svojih [politikah](#)²³ zapisali, da sta ustrezna odgovornost in ozaveščenost zaposlenih pomembna cilja za doseganje kakovostnih podatkov. V središče so postavili idejo, da morajo vsi zaposleni prepoznati potrebo po skrbi za kakovost podatkov in razumeti, da lahko k temu sami prispevajo. Da se morajo zavedati individualne odgovornosti, ki mora biti vključena v ustvarjanje, hrambo, analizo in poročanje o podatkih. Prav vsak zaposleni se mora zavedati posledic slabe kakovosti podatkov, ki vplivajo ne le na njegovo ozko področje, pač pa tudi na delo drugih oddelkov in ustanov, ki so z njim povezane. Vsak je svojemu vodji dolžan poročati o težavah v zvezi s kakovostjo podatkov in vodja je dolžan ukrepati. Vsak zaposleni mora poznati politike, povezane s kakovostjo in zaščito podatkov. Z institucionalizacijo praks postanejo zahteve in odgovornosti transparentne in njihovo uresničevanje del vsakodnevne prakse raziskovalnega procesa.

Čeprav se morajo vsi zavedati pomena in odgovornosti za doseganje kakovostnih podatkov, gre posebna vloga tistim, ki se poleg znanstvenega ustvarjanja ukvarjajo tudi s storitvami, povezanimi s podatki. V obstoječih pogojih je zato smotrno poiskati centre in nosilce znanja, ki bi lahko bili ustrezno izhodišče za oblikovanje in sistematizacijo novega poklica, ter tovrstna znanja vključiti v univerzitetno izobraževanje znanstvenikov, prilagojeno njihovi disciplini. Eno od področij, o katerem se veliko razpravlja v zvezi s spremenjenimi poklicnimi usmeritvami, je tudi vloga knjižničarjev v novem okolju politik in zahtev odprtega dostopa do podatkov. V Ameriškem združenju knjižničarjev (Association of Research Libraries, v nadaljevanju ARL) se zavzemajo, da bi knjižnice že danes lahko služile kot infrastrukturna podpora. Nekaj njihovih argumentov:

- knjižnice posedujejo strokovno razumevanje politik in načel, povezanih z odprto izmenjavo znanstvenih in izobraževalnih informacij, ter poznajo vloge, ki bi jih lahko opravljali institucionalni repozitoriji za zagotavljanje izmenjave;
- imajo izkušnje z razvojem in podporo orodij za distribucijo in iskanje informacij, tako znotraj discipline kot tudi širše;
- imajo izkušnje z razvojem in podporo poslovnih in tehničnih strategij za dolgoročno arhiviranje;
- razumejo tiste vidike znanstvene informacije, ki zadevajo arhiviranje in življenjski krog podatkov ter pomen dolgoročnega zagotavljanja dostopa in uporabnosti (Agenda 2007, 13).

Čeprav se knjižničarji zavedajo svojih omejitev, se zavzemajo za to, da bi v prihodnosti razvili orodja in veščine, specifične za področje ustvarjanja podatkov, dostopa in druge rabe. Zavzemajo se, da bi bili vse bolj udeleženi v postopkih nastajanja podatkov, saj bi tako lažje in bolje prispevali k izvajanju programa storitev infrastrukturne podatkovne podpore. Konkretni cilj je razviti takšna znanja, sposobnosti in veščine, da bodo knjižničarji mogli prispevati k e-znanosti

²² Za več informacij je zanimiva študija uporabnikov, ki so jo na vzorcu 1329 znanstvenikov pripravili Tenopir in dr. (2011).

²³ *Data Quality and Data Quality Assurance Policy*, http://www.admin.ox.ac.uk/pras/aboutus/data_quality/, 7. 8. 2012.

in k oblikovanju novih vlog in modelov storitev ter postali aktivni udeleženci v procesu zasnove in razvoja raziskovalne infrastrukture, vključujoč sistem in storitve, ki bodo podpirale celotni življenjski krog raziskovanja (Agenda 2007, 16–17; glej tudi Lossau 2012). Ta dokument je zanimiv tudi zato, ker obsega strukturo ciljev, strategij in akcij o tem, kaj konkretno bi bilo potrebno storiti – npr. ustanoviti delovno skupino, organizirati sestanke, pripraviti razpise za pilotne projekte, oblikovati učne načrte in usposabljanja.

Izoblikovali so delovno skupino in leta 2009 izvedli raziskavo med članicami ARL. Namen je bil »spoznati, kako lahko knjižnice prispevajo k e-znanstvenim dejavnostim«, ter »opredeliti organizacije in ustanove, ki imajo podoben interes« (Soehner in dr. 2010, 7). Rezultati raziskave so pokazali, da bo potrebno namesto enega samega sistema storitev razviti koordiniran in sintetiziran sklop, ki bo šele tvoril podporni sistem e-znanosti (Soehner in dr. 2010, 7–8). Med najbolj perečimi problemi v tej raziskavi izpostavljajo: pomanjkanje sredstev, ustrezno usposobljenost kadra in strokovnjakov s področja ravnanja s podatki ali hrambe in pomanjkanje poenotenih smernic. Čeprav so izrazili zaskrbljenost zaradi težav in omejitev, pa je med ameriški in kanadskimi knjižničarji mogoče zaznati navdušenje nad novimi vlogami v akademskem raziskovalnem procesu (Soehner in dr. 2010, 8). V primerjavi z ugotovitvami te študije je v izhodišču naše študije vtis, da gre pri nas za izrazitejšo pomanjkanje institucionalnih politik na področju vprašanj e-znanosti in za marginalno vlogo knjižnic pri podpori projektov e-znanosti. Knjižnice ostajajo omejene na tradicionalni servis dostopa do kupljenih revij in drugih tiskanih in e-publikacij ter vodenja raziskovalnih bibliografij.

Po drugi strani se tudi pri nas soočamo s podobnimi problemi, kot so ti, ki so jih prepoznali med ameriški in kanadskimi knjižničarji: neurejeno financiranje razvoja in projektov na teh področjih, nezadostno strokovno znanje in usposabljanje kadrov pri ravnanju in skrbi za podatke. Veliko se pričakuje od rezultatov projektov, kot je *DataNet*,²⁴ se pravi razvojnih projektov in projektov, usmerjenih v oblikovanje strategij, politik in rešitev dobrih praks. Razvijajo se tudi programi in študijske smeri na fakultetah.

S tem problematičnim področjem, kjer gre za iskanje obstoječih znanj splošnih informacijskih strokovnjakov, kot so knjižničarji, in njihovo transformacijo za podporo podatkovnim storitvam, se bomo v poročilu še ukvarjali, saj se kaže kot premalo izkoriščen potencial, ki obenem predstavlja novost in kreativnost pri iskanju in določanju vlog pri raziskovalni podpori: raziskovalci in ustvarjalci prevzemajo del vlog podatkovnih storitev, knjižničarji se vključujejo v aktivnosti raziskovalnih projektov, postajajo del raziskovalnega tima. Dolgoročno to predstavlja tudi izziv za oblikovanje izobraževalnih programov in študijskih smeri na fakultetah. Vzorčni primer predstavljajo učni moduli [RDM Train](#),²⁵ financirani s strani britanske organizacije JISC.²⁶ Gre za pilotne primere tečajev in učnih gradiv za različna področja znanstvenega raziskovanja. Namenjeni so podiplomskim študentom, ki v okviru tečajev pridobijo osnovna znanja s področja raziskovalnih podatkov, največji del pozornosti je usmerjen na ravnanje s podatki. Za enkrat so na voljo moduli za področja zdravstvenih ved/medicine²⁷, umetnosti²⁸, arheologije²⁹,

²⁴ Sustainable Digital Data Preservation and Access Network Partner, NSF:
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503141, 11.2.2014.

²⁵ Glej: <http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmtrain.aspx>, 11.2.2014.

²⁶ <http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmtrain.aspx>, 7. 8. 2012.

²⁷ DATUM for Health,
<http://www.northumbria.ac.uk/sd/academic/ceis/re/isrc/themes/rmarea/datum/health/materials/?view=Standard>, 7. 8. 2012.

psihologije³⁰, geoznanosti, družbenih in političnih ved³¹. Vzorčni primeri predstavljajo zgled, kako je mogoče problematiko vključiti v redno izobraževanje novih raziskovalnih kadrov. Za splošno področje informacijskih ved podoben program razvijajo v okviru [projekta RDMRose](#)³², programi usposabljanja pa se pojavljajo tudi med knjižničarji v ZDA³³.

²⁸ *Cairo* je podiplomski učni modul, namenjen ustvarjalnim umetniškim disciplinam, <http://www.projectcairo.org/node/9>, 7. 8. 2012.

²⁹ Projekt *DataTrain* za arheologijo, <http://archaeologydataservice.ac.uk/learning/DataTrain>, 7. 8. 2012.

³⁰ *Data Management Training* za psihologe, <http://www.dmtpsych.york.ac.uk/>, 7. 8. 2012.

³¹ *MANTRA, Research Data Management Training* je tečaj, namenjen podiplomskim študentom s področja geoznanosti, družbenih in političnih ved ter klinične psihologije. Več o tem na: <http://datalib.edina.ac.uk/mantra/introduction.html>, 7. 8. 2012.

³² Glej: <http://www.sheffield.ac.uk/is/research/projects>, 7. 8. 2012.

³³ Glej: <http://americanlibrariesmagazine.org/features/08212012/what-s-new-lis-schools>, 7. 8. 2012.

RAZISKOVALNA METODA

Mark Parsons (2010) poudarja, da ni vprašanje, ali odpreti dostop do podatkov ali ne, pač pa da je potrebno bolj podrobno razdelati, kdaj, v katerih primerih in kako odpreti podatke, se pravi pripraviti politike, ki upoštevajo stvarno situacijo na posameznih področjih. Kakor je izraženo v Načelih OECD glede fleksibilnosti, je pri pripravi potrebno upoštevati raznolikost različnih ravni in področij. V tem duhu Parsons izpostavlja potrebo po pristopu »od spodaj navzgor«, po politikah, ki bodo razumele stanje, saj bodo izoblikovane skozi razumevanje problematike, kot jo vidijo raziskovalci in drugi, ki se ukvarjajo s podatki.

Tako bo mogoče spoznati raznovrstnost ravnanja z znanstvenimi podatki, ki je pogojeno in omejeno s posameznimi področji. Učinkovit akcijski načrt je mogoče pripraviti le tako, da poznamo razmišljanje in delovanje akterjev na različnih raziskovalnih področjih ter v različnih fazah življenjskega kroga podatkov. Zato je za oblikovanje politik in strategij na tem področju potrebno koordinirano razpoznavanje problemov in njihovo reševanje z različnih perspektiv. V enem od zaključkov poročila *Shared Responsibilities in Sharing Research Data* iz Padove (ESF 2008) je eksplicitno poudarjeno, da morajo politike izmenjave raziskovalnih podatkov »ostati blizu znanosti in znanstvenikom«, sicer se lahko primeri, da bodo podatkovne storitve same sebi namen.

Da bi se temu izognili, si je treba prizadevati za širšo vključitev raziskovalcev v razvoj politik in zmogljivosti na področju dela s podatki. Odgovoriti si je potrebno na pomembna vprašanja. Ali bodo raziskovalci izročili svoje podatke, če ni na voljo ustrezne infrastrukture, kjer bi znali ravnati z njimi? In obratno, ali je smiselno veliko in trajno vlagati v raziskovalno infrastrukturo, če repozitoriji ostajajo prazni zaradi pomanjkanja pripravljenosti, da bi raziskovalci deponirali svoje podatke? Takšna vprašanja nakazujejo, da morajo politike in financiranje hoditi z roko v roki z raziskovalci, ustvarjalci in uporabniki podatkov (ESF 2008, 24–25). Ne glede na raznolikost politik in pristopov pa Parsons (2010) poudarja, da morajo vsi nujno vključevati odprti dostop in dolgotrajno digitalno hrambo.

Da bi dosegli odprto in etično ter zakonsko podprto podatkovno ureditev, bo treba uporabiti oba pristopa: »od zgoraj navzdol« in »od spodaj navzgor«. »Od spodaj navzgor« bodo raziskovalne skupnosti morale opredeliti nove etične zahteve čim bolj odprtega deljenja podatkov, ki bo temeljilo na sodelovanju in pravičnosti glede na prispevke. »Od zgoraj navzdol« bo politike treba uskladiti z nacionalnimi zakonskimi okviri, izhajajoč iz primata načel odprtosti in etične rabe, ki v znanosti temelji predvsem na zaupanju, se pravi, da bo potrebno oblikovati ustrezno interpretacijo in prilagoditev zakonodaje (Parsons 2010, 4).

Skladno z logiko pristopa »od spodaj navzgor« smo si za temeljno metodo ugotavljanja izhodiščnega stanja na področju ustvarjanja, vrednotenja, shranjevanja, objavljanja in uporabe raziskovalnih podatkov izbrali metodo intervjuja s strokovnjaki z različnih znanstvenih področij. Da bi čim bolj spoznali razmere, pogoje in okoliščine, v katerih nastajajo podatki, se hranijo in ponovno uporabljajo, smo izbrali metodo polstrukturiranega intervjuja. Prednost izbrane metode je možnost prilagajanja vprašanj in medsebojnega pojasnjevanja za boljše razumevanje. Z odpiranjem vprašanj »od spodaj navzgor« so imeli intervjuvanci večjo možnost izraziti pomisleke in svoje predloge glede obravnavane problematike.

Okvirne tematike, ki smo jih pokrili z intervjuji, smo črpali iz obstoječih zgledov in jih prilagodili posameznim situacijam. Američanka Francine Berman (2010) npr. predlaga, kako bi bilo mogoče podatke popisati na način popisa prebivalstva. Tako zbrani oziroma popisani podatki bi po njenem mnenju omogočili učinkovitejše načrtovanje javne infrastrukture za hranjenje digitalnih podatkov. Pri popisu podatkov iz raziskav, financiranih iz javnih sredstev, naj bi si pomagali z naslednjimi vprašanji:

- Koliko je posameznih podatkovnih datotek in kakšna je njihova velikost (v bitih)?
- Kakšne vrste podatkovnih datotek (besedilo/video/avdio ...) se pojavljajo na posameznih (znanstvenih) področjih?
- Kakšne so potrebe po hrambi (potrebe uporabnikov/obveznosti)?
- Kakšne storitve in orodja se uporabljajo v povezavi s temi podatki (npr. sekvenčna analiza genov, vizualizacija podatkov in drugo)?

Zastavljena vprašanja predstavljajo izhodišče naše raziskave, ob strani smo pustili le prvo vprašanje, saj smo se z izborom metode in eksplorativnim pristopom bolj osredotočili na primere kot pa na kvantitativni popis stanja.

Tudi raziskava, ki so jo Harley in dr. (2010) izvedli na Univerzi v Kaliforniji, nam je bila za zgled kot pristop k izvajanju intervjujev in fokusnih skupin. Podobno kot mi so se odločili za priložnostno vzorčenje. Udeležencem so pred izvedbo intervjuja poslali kratek povzetek vprašanj, na katera so kasneje odgovarjali; med intervjujem so dovolili, da so intervjuvanci govorili tudi o drugih temah, če so s tem dodatno osvetlili obravnavano tematiko. Z obsežno raziskavo so pokrili različna raziskovalna področja in jih predstavili kot študije primera.

Osnovna zahteva tudi pri nas je bila pokriti več različnih znanstvenih disciplin in spoznati različne faze življenjskega kroga podatkov. Zato je bilo treba ustvariti nabor sodelujočih, ki prihajajo iz različnih raziskovalnih okolij in imajo s podatki različne izkušnje, kot ustvarjalci, skrbniki ali uporabniki. Prvi korak pri izboru intervjuvancev je bil sestanek s predstavniki raziskovalnih inštitutov, ki ga je 19. januarja 2011 organiziralo Ministrstvo za visoko šolstvo, znanost in tehnologijo. Sestanka se je udeležilo okoli 30 udeležencev, ki so na seznam vpisali svoje kontaktne podatke. Ti so bili en vir za pripravo seznama potencialnih intervjuvancev, druge smo poiskali z namiznim raziskovanjem, tretje preko osebnih poznanstev in s priporočili drugih. Končno smo izoblikovali seznam 85 oseb, ki so bile potencialno zanimive za sodelovanje.

Med temi smo izbrali 47 kandidatov, in sicer tako, da smo pokrili čim bolj različna raziskovalna področja. Kandidate za intervju smo pisno obvestili o vsebini projekta Odprti podatki, o predvidenem poteku raziskave ter o ciljih. Zaposlili smo jih, da po svojih močeh čim dejavneje sodelujejo in tako pomagajo pri odstranjevanju ovir na poti k prostemu dostopu do podatkov in k spodbujanju čim širše rabe podatkov.³⁴ Pisno prošnjo je spremljalo priporočilo Ministrstva za visoko šolstvo, znanost in tehnologijo z dne 17. 12. 2010.³⁵ Po pisnem obvestilu smo stopili v stik s kandidati še osebno in se z njimi skušali dogovoriti za intervju.

Nekateri so sodelovanje zavrnil, druge smo izpustili, ko se je izpolnila kvota predstavnikov z njihovega področja. Spet tretje smo dodatno vključili po priporočilih intervjuvancev, če so jih ti

³⁴ Priloga 1 – Prošnja za sodelovanje v projektu Odprti podatki.

³⁵ Priloga 2 – Priporočilno pismo za sodelovanje v projektu Odprti podatki.

označili za kompetentne za določeno področje, ki ga do tedaj še nismo pokrili. Intervjuje smo izvedli z 22 kandidati.

Protokol intervjuja

Intervjuji so potekali v prisotnosti dveh predstavnikov ADP, eden je vodil pogovor, drugi po potrebi kaj dodatno povprašal, beležil, skrbel za delovanje snemalne naprave in drugo. Večina intervjujev je potekala individualno in brez motenj. Le v enem primeru je bil del intervjuja opravljen v prisotnosti nadrejene osebe, v nekaj primerih je prišlo do manjših nenačrtovanih motenj (obiski), ki so za krajši čas prekinili potek pogovora, a ga niso zmotili. V začetku intervjuja smo intervjuvancem dali v podpis *Izjavo o strinjanju s pogoji sodelovanja*³⁶, s katero so potrdili, da soglašajo s sodelovanjem v projektu.

Sledilo je izpolnjevanje osnovnih podatkov o intervjuvancu. Vpisali smo ime in priimek, organizacijo, v kateri intervjuvanec deluje, njegovo znanstveno področje, položaj, izobrazbo, leta dela in letnico rojstva. Potem smo z nekaj uvodnimi vprašanji intervjuvanca vpeljali v temo pogovora. V začetnem delu intervjuja smo ga oziroma jo prosili, naj na razpredelnici označi, katere vloge opravlja na področju podatkov in v katerih fazah življenjskega kroga podatkov sodeluje.³⁷

Uvodna vprašanja so bila iztočnica za nadaljevanje pogovora o fazah življenjskega kroga podatkov. Intervju je v povprečju trajal eno uro. Intervjuvanci so se pozitivno odzvali na pogovor in večinoma izkazali interes za morebitno nadaljnje sodelovanje.

Vse intervjuje smo zvočno posneli, vse posnetke prepisali, očistili in opremili z internetnimi povezavami na obravnavane projekte, formate, primere, zglede ... 11 intervjuvancev se je odločilo, da želijo ohraniti anonimnost,³⁸ trije so ostali neopredeljeni. Intervjuji se pripravljajo za spravilo kot raziskovalni podatki projekta v podatkovnih arhivih ADP, kjer bodo v ustrezno anonimizirani obliki dostopni uporabnikom za druge raziskovalne ali izobraževalne namene.

OSNOVNI PODATKI O VZORCU

Od 31. 3. 2011 do 10. 6. 2011 smo opravili 22 intervjujev. Intervjuvanci prihajajo iz 17 različnih disciplin:

- biokemija in molekularna biologija,
- biologija,
- fizika,
- medicinske veze in statistika,
- geodezija,
- gradbeništvo,
- ekonomija,
- jezikovne tehnologije,
- slovenistična literarna veda,
- glasbena akustika,

³⁶ Priloga 3 – Izjava o strinjanju s pogoji sodelovanja.

³⁷ Priloga 4 – Vprašalnik.

³⁸ "Potrdite, ali izrecno želite, da je vaše ime navedeno v poročilih in drugih gradivih projekta, ki se navezujejo na izrečeno v okviru intervjuja."

- zgodovina,
- arheologija,
- geografija, kartografija,
- urbanizem (specializirana knjižnica),
- socialna antropologija,
- socialno delo,
- statistika.

Večina intervjuvancev je zaposlenih na Univerzi v Ljubljani in na posameznih specializiranih raziskovalnih inštitutih. Vsi intervjuji so bili opravljeni v Ljubljani.

Univerza v Ljubljani:

- Filozofska fakulteta,
- Biotehniška fakulteta,
- Fakulteta za socialno delo,
- Ekonomska fakulteta,
- Fakulteta za matematiko in fiziko ter
- Fakulteta za gradbeništvo in geodezijo.

Inštituti:

- Inštitut za novejšo zgodovino,
- Urbanistični inštitut RS,
- ZRC SAZU, Inštitut za slovenski jezik,
- ZRC SAZU, Inštitut za arheologijo,
- ZRC SAZU, Glasbenonarodopisni inštitut,
- Inštitut za biostatistiko in medicinsko informatiko, Medicinska fakulteta,
- Institut Jožef Stefan,
- Geodetski inštitut Slovenije in
- Inštitut za ekonomska raziskovanja.

Glede na spol so bili intervjuvanci približno enako razdeljeni. Več kot polovica vprašanih je pri izobrazbi dosegla stopnjo doktorata, ostali imajo končano najmanj univerzitetno izobrazbo. Najmlajši intervjuvanec je leta 2011 štel 35 let, najstarejši pa 75. V povprečju so bili intervjuvanci stari nekaj manj kot 47 let.

Povprašali smo jih tudi po položaju, ki ga zasedajo v organizaciji, kjer so zaposleni. Med intervjuvanci je sledeča struktura:

- 2 asistenta,
- 1 izredni profesor,
- 2 redna profesorja,
- 2 prodekana,
- 1 samostojni strokovni sodelavec,
- 1 raziskovalec,
- 2 znanstvena sodelavca,
- 1 višji znanstveni sodelavec,
- 1 znanstveni svetnik,
- 1 direktor inštituta,

- 1 pomočnik direktorja inštituta,
- 1 svetovalec v direktorjevi pisarni,
- 2 vodji skupine,
- 1 vodja znanstvenoinformacijskega centra,
- 1 vodja projektov in
- 2 vodji knjižnice.

Izkazalo se je, da so intervjuvanci pogosto hkrati nosilci različnih vlog. Med 18, ki so odgovorili na vprašanje, katere vloge opravljajo na področju podatkov, se jih je 12 opredelilo za ustvarjalce, 14 jih je reklo, da opravljajo storitve s področja podatkov, 11 se jih je opredelilo za uporabnike podatkov in 9 jih dela na področju vodstva in oblikovanja politik.

Med 18, ki so opredelili svoje vloge, je 5 takih, ki so nosilci izključno ene same vloge – in sicer opravljajo storitve na področju podatkov. 3 so odgovorili, da opravljajo dve od naštetih vlog, 5 jih je povedalo, da opravljajo tri vloge, in 5, da so nosilci vseh naštetih vlog.

VPRAŠALNIK

V zasnovi vprašalnika smo izhajali iz naslednjih vprašanj. Kje so glavne pomanjkljivosti ustvarjanja, dostopa in hrambe raziskovalnih podatkov, ki bi jih bilo potrebno izboljšati? Ali gre za pomanjkanje znanja, vodil, pripravljenosti, odgovornosti? Kako akterji vidijo drug drugega: identificiranje medsebojnih odvisnosti in vlog, odgovornosti. Zanimalo nas je poznavanje in uporaba dobrih praks in standardov, ki bi lahko služili za določanje pomanjkljivosti pri ustvarjanju podatkov in medsebojno izmenjavo izkušenj za izboljšave.

Pri intervjujih smo izhajali iz pričakovanj, da so na nekaterih področjih odgovori na vprašanja že zelo jasno izoblikovani, medtem ko se bo drugje pokazala zelo nizka stopnja ozaveščenosti in razgledanosti. Na videz preprosto vprašanje, kje bi lahko objavili svoje podatke, ima lahko jasen odgovor – institucionalni repozitorij, domače ali svetovno področno podatkovno središče, podatkovni repozitorij revije, na domači strani raziskovalca – lahko pa ugotovimo, da intervjuvanci o tem ničesar ne vedo. Denimo eno od področij, ki je v razvitem svetu vse bolj izpostavljeno, tj. zahteva, da objavo v reviji spremljajo tudi podatki oziroma da mora biti omogočen dostop do podatkov, je poznana le redkim ustvarjalcem podatkov, in sicer tistim, ki so ob objavi ali branju publikacije naleteli na tovrstna spremljajoča gradiva. Tudi če za določeno področje obstaja možnost spravila in storitev dostopa do podatkov, sta zahtevnost priprave podatkov, pripravljenih po določenih standardih, in pomanjkanje časa oviri za pripravljenost izkoristiti to možnost. Pogosto je pripravljenost za omogočanje dostopa do podatkov povezana tudi s transparentnostjo vseh postopkov in analiz. Kadar so problematični statistični postopki in drugi vidiki priprave ali dokumentacije podatkov, je nizka tudi pripravljenost izpostaviti in narediti vidno celoto vsega, kar je v ozadju ugotovitev v člankih.³⁹

Zaradi tega smo v intervjuje vključevali tudi hipotetična vprašanja, ali bi bili intervjuvanci sposobni in ali bi jim bilo v interesu prevzeti vlogo opravljanja storitev drugih področij, bodisi v smislu centrov za svetovanje, npr. pri pripravi DMP⁴⁰, ali izmenjave dobrih praks, ali pa tudi kot storitve hrambe in dostopa do podatkov. Kdo bi se lahko s tem ukvarjal? Knjižnica, raziskovalci,

³⁹ Za več o tem, ko so statistične slabosti razlog za neobjavo podatkov, glej Heather Piwowar (2011) in Wicherts in dr. (2011).

⁴⁰ *Data Management Planning* oziroma Načrti ravnanja s podatki.

specializirani centri, splošni repozitorij za podatke? Kakšni so nastavki te dejavnosti: kdo so favoriti, kje so že vzpostavljene povezave doma in v tujini, uporaba dobrih praks in standardov.

Pri oblikovanju strukture intervjuja smo, kot rečeno, izhajali iz podobnih študij drugje, tako da smo med intervjujem imeli pred seboj seznam vprašanj in problemov, ki smo se jih želeli dotakniti. Nismo vnaprej pričakovali, da bomo dobili zelo drugačne odgovore kot drugod. Bolj je bilo pomembno, da iz prve roke dobimo predstavo o stanju in možnih zadržkih ter ovirah pri vzpostavitvi prihodnjega sistema in politik odprtih podatkov. S tem smo sledili tudi Parsonsovemu (2010, 4) stališču, da je ena od najpomembnejših kratkoročnih strategij to, da morajo podatkovne politike vključevati jasno opredelitev vlog, odgovornosti in virov. Zato je potrebno opredeliti mehanizme za deponiranje podatkov v arhive ter identificirati strokovnjake za hrambo in izmenjavo podatkov. Prvi korak do oblikovanja kakovostnih politik je temeljito poznavanje dejanskega stanja na področjih potreb, ovir, zadržkov, predstav, pričakovanj, izkušenj, razmerij med akterji ...

ŽIVLJENJSKI KROG PODATKOV

Opredelitve pojmov

Preden preidemo k predstavitvi vprašalnika, moramo opredeliti pojme, s katerimi bomo poskušali pojasniti stanje na področju raziskovalnih podatkov, in izpostaviti kategorije, za katere po zgledih tujih študij sklepamo, da pomagajo soustvarjati življenjski potek raziskovalnih podatkov: kaj se dogaja s podatki od načrtovanja in konceptualizacije naprej. Za čim bolj jasno predstavo o življenjski poti podatkov bomo opredelili vloge in njihove nosilce, objekte oziroma predmete in pojme, ki se nanašajo na podatke in življenjski krog podatkov.

Začeli bomo kar z opredelitvijo tega, kaj so raziskovalni podatki. Sledeč Načelom in usmeritvam OECD, so **raziskovalni podatki** na splošno digitalno berljivi zapisi o dejstvih, ki predstavljajo osnovno podlago za znanstveno raziskovanje in ki v okviru znanstvene skupnosti veljajo kot ustrezno sredstvo za preverjanje veljavnosti raziskovalnih spoznanj. Poudarek je na dostopu do čim bolj osnovnih podatkov, pripravljenih in dostopnih v takšni obliki, da jih je mogoče uporabiti za različne druge namene. Za boljšo predstavitev problematike opredeljevanja raznolikih vrst podatkov v nadaljevanju povzemamo nekaj podobnih splošnih definicij podatkov, ki sta jih v *IASSIST Quarterly* povzela Yoon in Tibbo (2011):

V The National Research Council (1999, 15) podatke opredeljujejo kot »dejstva, številke, črke in simbole, ki opisujejo predmete, zamisli, pogoje, situacije ali druge dejavnike«.

V The National Science Board (2005, 13) pojem »podatki« uporabljajo za »kakršno koli informacijo, vključujoč besedilo, številke, podobe, video ali filme, avdio, programsko opremo, algoritme, enačbe, animacije, modele, simulacije in drugo«.

V The National Science Foundation razlikujejo štiri tipe podatkov:

- podatki opazovanj (npr. vreme, človeško ravnanje),
- računalniško ustvarjeni podatki (rezultati računalniških modelov in simulacij),
- eksperimentalni podatki (iz laboratorijskih študij),
- zapisi (s strani vlad, podjetij, javnega in zasebnega življenja) (povzeto po Borgman 2010, 19).

Tu je še nekaj različnih definicij raziskovalnih podatkov, s katerimi si avstralske univerze pomagajo pri ravnanju z raziskovalnimi podatki (ANDS 2011):

»Pojem raziskovalni podatki obsega dejstva, opažanja, podobe, rezultate računalniških programov, zapise, meritve ali izkustva, na katerih temelji argument, teorija, preizkus, hipoteza ali kak drug raziskovalni izsledek,« so zapisali v svojih politikah o ravnanju s podatki na The Queensland University of Technology.

Na The University of Melbourne so raziskovalne podatke opredelili kot »dejstva, opažanja ali izkustva, na katerih temelji argument, teorija ali preizkus. Podatki so lahko numerični, opisni ali vizualni. Podatki so lahko surovi ali analizirani, eksperimentalni ali podatki opazovanj. Podatki vključujejo: laboratorijske in terenske zapiske, primarne raziskovalne podatke (vključujoč raziskovalne podatke v tiskani ali računalniški obliki), vprašalnike, avdiokasete, videokasete, modele, fotografije, filme, testne odgovore. Raziskovalne zbirke lahko vključujejo: diapozitive, artefakte, primerke in vzorce. Izvirne informacije o podatkih lahko vključujejo: kako, kdaj in kje so bili zbrani podatki ter s čim (npr. naprava). Obsega lahko tudi programsko kodo, s katero so ustvarili, komentirali ali analizirali podatke«.

Na The Monash University raziskovalne podatke razumejo kot »podatke, posnetke, dokumente ali druge dokaze, ne glede na vsebino ali obliko (to je tiskana, digitalna, fizična ali druga oblika), ki vključuje raziskovalna opažanja, ugotovitve ali izsledke, vključujoč primarno gradivo in analizirane podatke«.

Na Griffith University razlikujejo med raziskovalnimi podatki in primarnim gradivom. Raziskovalni podatki so: zapiski v obliki števil, simbolov, besedila, podob ali zvokov, uporabljeni kot primarni vir za raziskavo in v raziskovalni skupnosti sprejeti kot nujni za vrednotenje raziskovalnih ugotovitev.

Pri oblikovanju definicije, kaj so raziskovalni podatki, si je mogoče pomagati tudi z opredelitvijo, kateri podatki vanje ne sodijo:

- administrativni podatki: četudi se nekateri nanašajo na raziskovalne projekte in jih je mogoče videti kot raziskovalne podatke, zanje večinoma poskrbijo v okviru upravnih postopkov znotraj inštitucije (menedžment, strategije);
- učna gradiva: so del učnega procesa na univerzi, običajno zanje skrbijo neodvisno;
- raziskovalne objave: mogoče jih je videti kot podatke, vendar je zanje večinoma dobro poskrbljeno s strani založnikov (ANDS 2011).

Odprti dostop do raziskovalnih podatkov, podprtih z javnimi sredstvi, pomeni, skladno z [mednarodnimi priporočili OECD](#)⁴¹ in [Evropske komisije](#)⁴², po možnosti brezplačen dostop preko interneta za katerega koli uporabnika, razen v posebnih primerih potreb po varovanju.

⁴¹ *OECD Principles and Guidelines for Access to Research Data from Public Funding*, <http://www.oecd.org/sti/scienceandtechnologypolicy/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm>, 7. 8. 2012.

Pomemben pojem, s katerim operiramo v nadaljevanju, je tudi vloga in z njo povezana odgovornost. **Vloga** se uresničuje skozi navade in običaje, omejujejo jo pravila, norme, standardi, protokoli, licence, ki določajo in urejajo raziskovalno dejavnost na posameznem področju. Od tega, kako so opredeljene in strukturirane vloge, je odvisna raven kulture in organiziranosti deljenja podatkov.

Zgled je lahko stališče Research Information Network (RIN), predstavljeno v *Stewardship of digital research data – principles and guidelines* (2008), da morajo biti vloge in odgovornosti raziskovalcev, raziskovalnih ustanov in financerjev kolikor je mogoče podrobno in natančno opredeljene. Akterji morajo skupaj oblikovati kodeks delovanja, ki bo zagotavljal, da bodo ustvarjalci in uporabniki podatkov izpolnjevali svoje odgovornosti skladno z opredeljenimi načeli.

Swan in Brown (2008, 2) ugotavljata, da se je večina znanstvenikov priučila vloge, vezane na podatke, šele ob delu. Pridobiti so si morali veliko različnih veščin. Ob področni strokovni usposobljenosti in računalniškem znanju, ki sta za znanstvenika bolj ali manj samoumevna, za najpomembnejšo štejejo sposobnost »prevajanja« oziroma prenosa potreb in praks znanstvenikov v jezik tehnoloških orodij in potem spet nazaj v prakso.

Vloge se izražajo in uresničujejo z akterji, ki delujejo na področju raziskovalnih podatkov. Nosilci vlog uresničujejo pravice, dolžnosti in obveznosti ali pa tudi ne. Iz položaja nosilca vloge v znanstvenoraziskovalni skupnosti izvira stopnja moči ter njegove prednosti in slabosti pri izvajanju nalog. Glavni nosilci vlog so, po zgledu Liz Lyon (2007) in Interagency Working Group on Digital Data (2009):

- ustvarjalec podatkov: podatkovni znanstvenik, raziskovalna enota, inštitut;
- izvajalec storitev: skrb za podatke, katalogiziranje, opremljanje z metapodatki za uporabo, spravilo, povezovanje, ki jih prevzemajo specializirani področni podatkovni arhivi, mednarodni arhivi, digitalni institucionalni ali nacionalni repozitoriji, knjižnice;
- uporabnik (drugi raziskovalci, študentje, širša javnost);
- vodstvo ustanov in oblikovalci politik, financerji.

Eden od projektnih ciljev je ugotoviti dobre plati obstoječih nastavkov storitev, jih podpreti in okrepiti ter v njihovi smeri začrtati oblikovanje prihodnjih storitev. Pri pregledu stanja smo pozornost namenili razmerju moč/nemoč v povezavi s pristojnostmi in medsebojnimi pričakovanji o tem, kdo lahko izpolni določene naloge. Zgledi dobrih praks lahko služijo v procesu opolnomočenja manj razvitih področij.

Področje medsebojne izmenjave oziroma deljenja podatkov predpostavlja tudi **odgovornost** različnih nosilcev vlog. Ne zadostuje želja razvijati in spodbujati dejavnosti storitev prevzema, skrbi za podatke in dostopa do njih, saj lahko tovrstna infrastruktura ostane neizkoriščena, če ni spodbud za raziskovalce ustvarjalce podatkov. Zato je pomembno, da oblikovanje politik in strategij na tem področju zajema koordinirano razpoznavanje problemov in njihovo reševanje z različnih perspektiv. Poleg tega mora vključevati vse akterje na tem področju, tudi financerje.

Idealnotipsko bomo na področju sistema dostopa do podatkov ločili naslednje institucionalne nosilce aktivnosti: na eni strani so to nosilci in izvajalci raziskovalnih politik, kot sta na

⁴² Policy Initiatives, <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301>, 7. 8. 2012.

nacionalni ravni Ministrstvo za izobraževanje, znanost in šport (MIZŠ) in Javna agencija za raziskovalno dejavnost RS (ARRS) ter vodstva inštitucij in drugi organi odločanja, ter na drugi strani člani znanstvene skupnosti, ki nastopajo v različnih vlogah, kot uporabniki ali kot ustvarjalci in dajalci podatkov, kot povezovalni člen pa nastopajo nosilci podatkovnih infrastrukturnih storitev.

Objekti. Za razumevanje in razlago procesov, ki spremljajo življenjski krog podatkov, je pomembnih več objektov, pojmov, vsebin, gradiv:

Najprej seveda *podatki*, ki jih delimo na: surove, predelane in izvedene. Podatke morajo spremljati *metapodatki*, ki so sistematično predstavljena in strojno berljiva dokumentacija o podatkih in so ključnega pomena za razumevanje podatkov in za drugo rabo (Vipavc Brvar in Klep, 2004). *Publikacije* so gradiva, kjer so običajno predstavljene ugotovitve, temelječe na podatkih. Vse bolj postajajo pomembne tudi kot mesto za predstavitev podatkov kot takih, npr. kot podatkovna objava⁴³. *Orodja* za ustvarjanje, prikaz in analizo podatkov zadevajo tehnični vidik ustvarjanja podatkov in ravnanja z njimi. *Standardi, dobre prakse, navodila* o kakovosti, metapodatkih in hrambi obsegajo normativni vidik, napisana in nenapisana pravila znanstvene skupnosti, znanstvenih skupin ali posameznih znanstvenikov; neposredno so pomembni za ustvarjanje in drugo rabo podatkov. Nad vsem tem pa bdijo *politike, strategije, načrti, projekti*, ki kratkoročno in/ali dolgoročno vplivajo na celo področje skrbnega ravnanja s podatki. Iz njih izvirajo *pravice, dolžnosti in odgovornosti* akterjev. Iz njih pa izhajajo tudi *nagrade, spodbude in sankcije*, ki lahko okrepijo določene aktivnosti in zmanjšajo druge. Na kratko povedano, tehnični vidik in kulturo ustvarjanja in rabe podatkov, ki se zrcalita v življenjskem krogu podatkov, je mogoče razložiti le s poznavanjem vsebin in razmerij med nosilci vlog in objekti.

FAZE ŽIVLJENJSKEGA KROGA PODATKOV

Življenjska doba podatkov je običajno daljša, kot je trajanje projekta, v okviru katerega so nastali, saj lahko raziskovalci podatke uporabijo ali nadgradijo v sorodnih drugih projektih, lahko jih pa uporabijo tudi povsem drugi raziskovalci, so zapisali v UK Data Archive. Življenjsko pot podatkov vidijo kot preprost krog, v katerem si sledijo naslednje faze: ustvarjanje podatkov, obravnava podatkov, analiza, hramba, zagotavljanje dostopa do podatkov in druga raba podatkov.⁴⁴

V Digital Curation Centre (DCC) so oblikovali bolj podroben življenjski krog podatkov, ki vključuje več stopenj življenjske poti podatkov in upošteva različen značaj aktivnosti. V ospredje so postavili stopnje, ki zagotavljajo uspešno hrambo in zaščito podatkov, od zasnove do postopkov hrambe. Njihov model je namenjen načrtovanju aktivnosti v organizacijah.

Opredelili so aktivnosti, ki so posebej pomembne, zato ker potekajo ves čas življenjskega kroga podatkov: zagotavljanje metapodatkov, načrtovanje digitalne hrambe, spremljanje skupnosti in udeležba pri razvoju standardov, orodij ter shranjevanje in digitalno skrbništvo v teku celotnega življenjskega kroga podatkov.

⁴³ Primer podatkovne publikacije na način opisa podatkovnega vira glej Štebe (2013).

⁴⁴ Za shematski prikaz glej: *Research Data Lifecycle*, <http://www.data-archive.ac.uk/create-manage/life-cycle>, 7. 8. 2012.

Poleg dejavnosti, ki potekajo ves čas, so pomembne tudi tiste, ki potekajo v določenem zaporedju. Prva med njimi je konceptualizacija (načrtovanje ustvarjanja podatkov, vključno s hrambo). Sledijo ji: ustvarjanje podatkov ali prejem podatkov od drugih; ovrednotenje in izbor; sprejem podatkov v hrambo; aktivnosti hrambe, spravilo; dostop, raba in druga raba; predelava podatkov.

Spregledati pa ne smemo niti občasnih aktivnosti, ki so prav tako pomembne za ustrezno hrambo in vzdrževanje podatkov: prenos podatkov, ki niso bili izbrani za hrambo; ponovna ocena vrednosti podatkov in dodatne aktivnosti za izboljšanje kakovosti; pretvorba podatkov v druge formate ipd.⁴⁵

Na osnovi tovrstnih modelov smo pripravili poenostavljen slikovni prikaz, ki nam je služil kot opora za oblikovanje vprašalnika in izvedbo polstrukturiranega intervjuja.

Življenjski krog podatkov, kot smo ga obravnavali v intervjujih, je strukturiran po fazah, kot sledi: načrtovanje, financiranje, ustvarjanje, dodana vrednost, izbor/vrednotenje kakovosti in uporabnosti, spravilo, analiza/objava rezultatov v članku, dostop/iskanje/druga raba.



Slika1: Življenjski krog podatkov

Posamezne faze so služile kot osnova za oblikovanje področnih sklopov vprašanj, ki jih bomo predstavili v nadaljevanju. Za vsak sklop bomo najprej predstavili konceptualno izhodišče, nato pa s pomočjo povzemanja in navajanja iz intervjujev oblikovali zaključke. Sklope zaključkov bomo predstavili tudi ločeno v okvirčkih na način problemov, ki smo jih identificirali, ter možnih ukrepov za njihovo rešitev, ki jih bomo kasneje implementirali tudi v akcijski načrt.

Med oblikovanjem vprašalnika se je izkazalo, da je določene sklope smiselno združiti, med pogovori pa se je tudi dogajalo, da intervjuvanci večinoma govorijo o več fazah hkrati, zato se je nemalokrat struktura pogovora sprostil in je sledila temam, ki so bile intervjuvancu bližje ali jih je zmožgel bolje opredeliti, pojasniti ali preprosto izraziti svoje stališče.

⁴⁵ Za shematski prikaz glej: *The DCC Curation Lifecycle Model*, <http://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>, 7. 8. 2012.

PRIKAZ REZULTATOV V FAZI USTVARJANJA PODATKOV

Prva faza življenjskega kroga podatkov obsega *opis nastajanja podatkov*, z njimi povezana orodja, postopke in navodila. Dvanajst od osemnajstih intervjuvancev je odgovorilo, da je ustvarjanje podatkov ena od njihovih vlog. Izkazalo se je, da se intervjuvanci med seboj zelo razlikujejo po tem, kakšne vrste in tipe podatkov ustvarjajo in kakšne metode pri tem uporabljajo ter s kakšno infrastrukturo razpolagajo.

Zanimalo nas je, kako nastajajo podatki in metapodatki: za kakšne vrste podatkov gre in kaj se z njimi dogaja? Kako jih ustvarjajo in/ali kako z njimi ravnaajo? Od intervjuvancev smo skušali izvedeti, *kako vidijo odgovornost* (lastno in drugih) v zvezi z reševanjem problemov. Čemu so zavezani s pravili: zahtevam projekta glede kakovosti, eksplicitnim merila kakovosti, ki jih upoštevajo pri ustvarjanju podatkov; beleženju karakteristik glede kakovosti pri nastajanju podatkov? Zanimalo nas je, kakšna je odgovornost do naročnika, financerja, vodstva, uporabnika: ali že v načrtu opredelijo, kako in kakšne vrste podatkov bodo ustvarili? Za koga ustvarjajo podatke: zase, za ožji krog, za druge? Ali odgovornost do drugih uporabnikov razumejo v smislu skrbi za kakovost? Intervjuvance smo prosili, naj navedejo primere problemov: zanimalo nas je, ali identificirajo probleme, poznajo rešitve, vidijo vlogo standardov, ukrepov različnih akterjev. Na tej osnovi smo izluščili nosilce problemov in nosilce zglede reševanja, ki bodo v nadaljevanju osnova za preučevanje možnosti in pogojev za uveljavljanje dobrih praks.

Zanimalo nas je, kateri *problemi* iz faze ustvarjanja podatkov so povezani z drugimi fazami življenjskega kroga, ali jih akterji upoštevajo že v tej fazi ali ne, ali načrtujejo, kaj se bo zbiralo in za kakšen namen, ali upoštevajo vrednotenje in izbor, da bi z načrtovanjem dosegli kakovostne podatke. Ali upoštevajo zahteve hrambe in ali razrešujejo probleme na poti do zagotavljanja odprtega dostopa?

TIPI PODATKOV

Že na uvodnem srečanju predstavnikov različnih slovenskih inštitutov, ki ga je januarja 2011 v sodelovanju z Ministrstvom za visoko šolstvo, znanost in tehnologijo organiziral izvajalec projekta, in tudi kasneje, v začetnem delu intervjujev, smo sogovornikom vsakič znova zastavili osnovno vprašanje, kaj so raziskovalni podatki za različne akterje in na različnih znanstvenih področjih. Ali denimo numerične simulacije predstavljajo podatke? Kako je s podatki v humanistiki, kjer običajno uporabljajo drugačno terminologijo, zgodovinarji denimo govorijo o virih? Po začetni negotovosti smo se običajno hitro sporazumeli glede namenov projekta, tako na omenjenem sestanku kot pri intervjujih so izražali naklonjenost in podporo ciljem. Udeleženci prvega sestanka so na nas naslovili razmeroma veliko vprašanj, pripomb in predlogov, ki smo jih v nadaljevanju obdelali tudi v okviru intervjujev – npr. določitev meril kakovosti podatkov, vprašanja omejevanja dostopa do individualnih podatkov, dostopnost naravoslovnih podatkov v primerjavi z družboslovnimi, izraženi so bili pomisleki glede avtorskih pravic.

Nekaj osnovnih dilem se je nanašalo, kot omenjeno, na samo izhodiščno definicijo »raziskovalnih podatkov«, ki bi intervjuvancem pomagala zamejiti osrednjo temo pogovorov. Z naše strani je bil podan predlog, naj raziskovalci sami opredelijo, katere podatke na svojih področjih štejejo kot pomembne in pod kakšnimi pogoji so jih pripravljene deliti. Prav tako pa smo nosilce aktivnosti na začetku intervjujev zaprosili, naj sami podrobneje opišejo, s kakšnimi podatki se ukvarjajo pri svoji dejavnosti.

Na kratko bomo strnili, da se definicije nanašajo na obliko podatkov in na nosilce podatkov. V nekaterih primerih opisujejo tudi, na kakšen način so podatki vključeni v raziskovalni proces. V nadaljevanju pogledimo, kako raziskovalne podatke vidijo udeleženci naše raziskave.

Jaz si ne predstavljam, kaj je za vas podatek. Običajno si pod tem predstavljam neko surovo številko ali pa neke vrednosti, ker pač s tem največ delam. (Znanstvenik s področja geografije, N08)

Mi nimamo podatkov, mi se ukvarjamo z viri. (Znanstvenik s področja zgodovine, N010)

Podatek ni samo grob podatek. Lahko je knjiga, ki jo je Charles Dickens napisal, lahko je študija, ki jo je Charles Dickens napisal, lahko so njegove lingvistične analize ... Nekaj, kar jaz uporabim kot podatek in naredim analizo, bo nekdo drug mojo analizo uporabil kot podatek. (Nina Vodopivec, raziskovalka s področja socialne antropologije)

Z intervjuji smo uspešno pokrili več različnih področij tako v družboslovju in humanistiki kot v naravoslovju, kar se odraža v širokem naboru zelo raznolikih tipov podatkov, ki jih zbirajo v posameznih inštitucijah.⁴⁶

Izkazalo se je, da že sama opredelitev, kaj je podatek, v različnih znanstvenih disciplinah ni povsem samoumevna. Nabor zabeleženih tipov podatkov je temu primerno obsežen. Intervjuvanci naše raziskave so navedli različne tipe podatkov, s katerimi se srečujejo pri raziskovalnem delu:

- numerični podatki, meritve,
- rezultati numeričnih modelov,
- ekonomski modeli,
- anketni podatki,
- intervjuji,
- dnevniški zapisi,
- delovna poročila,
- slike,
- grafi,
- risbe,
- besede (besedilni korpus),
- pisna gradiva,
- statistični letopisi,
- popisi prebivalstva,
- bibliografske podatkovne zbirke,
- zvočni zapisi,

⁴⁶ Iz tega sledi sklep, da bo bodoče "izvajanje storitev skrbništva nad podatki potrebno prilagoditi močno raznolikim pod-disciplinarnim karakteristikam podatkov in navadam deljenja podatkov." (Cragin in dr. 2010)

- video,
- programska oprema,
- strojna oprema,
- aplikacije,
- spektri,
- senzogrami,
- lidarski podatki,⁴⁷
- herbariji,
- zbirke živali,
- vzorci tkiv.

Na osnovi odgovorov intervjuvancev na vprašanje, s kakšnimi podatki se ukvarjajo, smo ugotovili, da podatke pogosto predstavljajo običajni numerični podatki v obliki tabel, podatkovnih baz ipd., veliko pa je tudi pisnega, slikovnega, avdio in video gradiva. Ugotovili smo, da obstaja tudi kar nekaj bolj posebnih zbirk podatkov, kot so npr. herbariji, zbirke živali, vzorci tkiv, lidarski podatki, spektri, numerični modeli ipd. Kar nekaj intervjuvancev je v povezavi z opredelitvijo podatkov navedlo tudi svoje produkte v obliki programske ali strojne opreme in aplikacije, s pomočjo katerih se podatki generirajo. V našem naboru najdemo predstavnike podatkov za vsako od kategorij iz zgoraj omenjene tipologije *National Science Board report*: opazovanja, simulacije, eksperimenti in zapisi.

Pogosto se podatke deli tudi na t. i. referenčne zbirke in ostale podatke (*Long-Lived Digital Data Collections* 2005). Nekateri naši intervjuvanci so spočetka svoje razumevanje pojma podatki omejevali na referenčne zbirke,⁴⁸ ki so na določenih področjih najbolj opazne. Velikokrat se je šele med intervjujem izčistilo razumevanje, da so lahko za druge raziskovalce pomembni tudi vsi ostali podatki, ki se ustvarjajo znotraj raziskovalnih projektov. Med pogovorom so se ostrile razlike med tem, kaj je podatek, kaj podatkovna zbirka (vodja znanstvenoinformacijskega centra 1, 2, 3/ 5.7).⁴⁹ Na različnih področjih so se pokazale razlike v pojmovanju primarnega gradiva, razlike med surovimi podatki in predelanimi podatki, dokumentacijo, zapiski... Na primer razlikovanje med arheološkim najdiščem in podatki, povezanimi z lokacijo, najdbo, izkopavanjem in drugim (raziskovalec s področja arheologije, 4/5.7). Včasih je bilo potrebno zastaviti dodatno vprašanje, da smo v pogovoru prišli do osnovne ravni raziskovanja, do primarnih podatkov (raziskovalec s področja arheologije, 5/5.7). Tovrstne razlike v pojmovanju smo upoštevali pri pripravi akcijskega načrta, ki mora biti dovolj fleksibilen, da se v njem lahko prepoznajo raznolika področja.

Glede na metodo zbiranja je tipe podatkov mogoče razdeliti v naslednje skupine:

- podatki, ki jih zberejo raziskovalci sami,
- podatki, ustvarjeni z napravami, in
- podatki, katerih osnova za generiranje so drugi viri, npr. vzorci, kartoteke (znanstvenik s področja medicinskih vez in statistike, 6/5.7).

⁴⁷ Podatki, pridobljeni z metodo zračnega laserskega skeniranja.

⁴⁸ Npr. SARNAS (Seznam Arheoloških najdišč Slovenije) in druge zbirke, dostopne preko spletnega modela ARZENAL, več o tem na: <http://www.arzenal.si/sobe/zbirke>, 7. 8. 2012.

⁴⁹ Citati iz intervjujev so objavljeni v Opis stanja na področju raziskovalnih podatkov v Sloveniji, Poročilo projekta Odprti podatki – priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji (Štebe in dr. 2012), dostopno preko: http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odpp10_opis_stanja/, 17. 1. 2014.

V nekaterih inštitucijah je poleg ustvarjanja lastnih podatkov pogosta praksa tudi druga raba podatkov, zlasti je razširjena raba podatkov, ki so jih ustvarile za to uradno zadolžene javne ustanove (SURS, ARSO, APJES, Geodetska uprava RS, različni javni arhivi), ter nadaljnje delo z njimi, tudi preurejanje in združevanje. Intervjuvanci so omenili več problemov, povezanih s tem področjem. Denimo, kakšna je dostopnost do podatkov po tem, ko jih je raziskovalec dopolnil in preuredil za svoje namene. V tem pogledu je pogosto sklicevanje na stroge pogoje, pod katerimi so raziskovalci prišli do podatkov. Te pogoje je mogoče razumeti tudi kot izgovor, da podatkov v principu ni mogoče posredovati naprej. Npr. na področju kartografije in geografije so se zavezali, da bodo podatke uporabljali samo za lastne potrebe. Kar lahko delijo z drugimi, so novonastali produkti (npr. karte), ne smejo pa deliti oziroma posredovati surovih podatkov (znanstvenik s področja geografije, 7/5.7). Podobno velja tudi na področju ekonomije, kjer imajo dostop do določenih podatkov, vendar jih ne smejo »prenašati naokoli«, kot se je izrazil intervjuvanec (Boris Majcen, znanstvenik s področja ekonomije, 8/5.7). Poleg tovrstnih omejitev pa raziskovalce omejujejo tudi drugi pogoji. Na področju geodezije raziskovalci npr. lahko dobijo koordinate. Če bi želeli podatke, ki segajo na področje varovanja osebnih podatkov, se pogoji zaostrijo (znanstvenik s področja geodezije, 9/5.7). Podobno velja tudi na področju ekonomije, kjer se raziskovalci ukvarjajo s preučevanjem podatkov o posameznikih ali podjetjih, za kar obstajajo posebej določeni pogoji (Boris Majcen, znanstvenik s področja ekonomije, 10/5.7).

Rešitve problemov oziroma vprašanj, kot npr. pogoji dostopa do podatkov v javnih in državnih ustanovah ter do registrov obstoječih podatkov, po katerih je mogoče iskati, obstajajo. Za to je denimo mogoče pooblastiti področne podatkovne arhive, ki bodo posredovali pri dostopu do podatkov. Primer takega sodelovanja že poteka med ADP in SURS.⁵⁰

Intervjuvanci deloma poznajo tudi rešitve, ki so uveljavljene v tujini. Npr. intervjuvanec s področja ekonomije pravi, da imajo v ZDA raziskovalci takojšen dostop do statističnih podatkov, na katere morajo pri nas čakati tudi do pol leta. Želi si, da bi tudi v Sloveniji bile mogoče takšne rešitve (Polona Domadenik, znanstvenica s področja ekonomije, 11/5.7).

Poseben sklop predstavljajo podatki različnih (mednarodnih) centrov za hrambo in distribucijo specifičnih vrst podatkov, kot so GenBank, UniProt, Protein Data Bank, Amadeus ipd. Nekateri podatki iz takšnih baz so prosto dostopni, do drugih lahko raziskovalci dostopajo le ob plačilu.

V Sloveniji je pomembno upoštevati tudi razlike med področji oziroma delitev na »velike« (»*big science*«) in »male« znanosti (»*small science*«).⁵¹ Veliko vrst podatkov izhaja iz domen »malih« znanosti in so zaradi tega v okviru specializiranih podatkovnih arhivov pogosto zapostavljeni. Raziskovalec, ki ima izkušnjo z dvema različnima raziskovalnima področjema (jezikoslovje in fizika), odgovarja, da je na področju fizike delcev zavest glede hrambe in ravnanja s podatki visoka in problemi drugačni kot na področju družboslovja (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 12/5.7). Za »velike« znanosti je po navadi dobro poskrbljeno, uveljavljeni so standardizirani zapisi in skrb za podatke je dobro

⁵⁰ Glej podrobnosti na: <http://www.adp.fdv.uni-lj.si/projekti/uradne-statistike/>, 7. 8. 2012.

⁵¹ "Trenutno obstoječa znanja s področij ravnanja s podatki in skrbništva v akademskem okolju zadevata dva problema. Prvič, stalna raba fizike delcev visokih energij in/ali astronomije kot referenčne in primerjalne točke prikrije kompleksnost ustvarjanja in komuniciranja, ki je značilno za male znanosti in poddiscipline. (npr. Chompalov in dr. 2002). Drugič, osredotočanje na univerzalne rešitve glede nalog skrbništva nad podatki običajno prinaša zgolj modele skupnih storitev in modele na najbolj splošni ravni" (Cragin in dr. 2010, 4025). Podoben je tudi zaključek drugih avtorjev: "Veliki preostanek (*the long tail*, op. a.) malih podatkovnih datotek predstavlja poseben izziv – to, kako heterogene podatkovne datoteke priversti skupaj" (Theesen in Patterson 2011, 43).

vzpostavljena, podatki so zmeraj na voljo v elektronski obliki. Področje je zelo napredno, hkrati pa zelo ozko opredeljeno, samo za določen niz podatkov.

Okvir 1: P-R1

PROBLEM:

Podatki so zelo raznoliki, raziskovalci različno dojemajo, kaj so podatki in kateri so pomembni v njihovem raziskovalnem postopku. Opaziti je odsotnost jasne definicije, kaj podatki sploh so.

Področje »velikih« znanosti je precej urejeno in ne terja iskanja novih rešitev, še zlasti ker so uveljavljene in sprejete mednarodne rešitve. Opazno manj je urejeno področje »malih« znanosti.

REŠITEV:

Pri oblikovanju prihodnjih načrtov bo potrebno upoštevati kompleksnost in različne zahteve različnih področij.

Pri iskanju rešitev se bo potrebno posebej osredotočiti na »male« znanosti, saj v nacionalnem okolju, ki ga pokrivamo, prevladujejo prav te.

Pri iskanju rešitev so smiselne razprave med raziskovalci in nosilci storitev z različnih področij. Področno povezovanje je treba dopolniti z naprednimi zgledi in izkušnjami »velikih« znanosti, pri katerih so definicije, postopki in rešitve že preverjene in uveljavljene, kolikor so prenosljivi na druga področja.

Odkvisno od tega, ali storitve ravnanja in dostopa do podatkov že nastajajo ali ne, je smiselno oblikovanje razvojnih strategij različnih hitrosti. Na področjih, na katerih so akterji izrazili interes, da so pripravljene prevzeti naloge, povezane s podatkovnimi storitvami, je potrebno omogočiti razvoj in ne čakati na področja, ki zaostajajo.

VPRAŠANJA KAKOVOSTI PODATKOV V POGLEDIH INTERVJUANCEV

Vprašanje kakovosti je še eden od poudarkov, ki smo si jih izbrali za pregled problemov odprtega dostopa do raziskovalnih podatkov. Kakovost se nanaša tako na elementarno kakovost samih podatkovnih gradiv kot tudi na kakovost metapodatkov in storitev, povezanih s pripravo podatkov (npr. časovni rok za dostop, stopnja čiščenja in dodane vrednosti podatkov).

Ocena kakovosti podatkov je lahko odvisna tudi od tega, ali so podatki ustvarjeni za ozek namen, npr. za izdelavo raziskovalnega poročila v okviru lastnega projekta, ali pa so namenjeni za širšo uporabo, npr. za odprti dostop. Kakovost podatkov kot taka obsega več vidikov. Pomeni lahko temeljno zmanjševanje napake s ciljem, da bi se z meritvami približali pravi vrednosti. Zajema pa tudi širše organizacijske vidike, ki vključujejo poznavanje pravil in postopkov, načrtovanje in usposabljanje za kakovostne storitve v okolju ravnanja s podatki: od nastanka preko dodane vrednosti do zagotavljanja dostopa in rabe pri generiranju raziskovalnih poročil (primerjaj Biemer in Lyberg 2003).

POMEN KAKOVOSTNIH PODATKOV IN METAPODATKOV

Tukaj navedena izhodišča za ocenjevanje kakovosti podatkov in metapodatkov bodo služila za primerjavo z zaključki, ki smo jih dobili na podlagi intervjujev, kjer je bil vidik kakovosti v ospredju vseh faz izdelave in skrbi za podatke.

Najširša opredelitev kakovosti podatkov v okviru sistema odprtega dostopa je po kriteriju uporabnosti podatkov za določen namen. Neuporabnost je lahko posledica napak pri ustvarjanju podatkov, neustrezne dokumentacije ali neprečiščenih podatkov. Kakovostna storitev zagotavljanja dostopa pomeni, da je uporabnost gradiv maksimizirana, pri tem pa so upoštevani različni pogoji: kulturni, pravni in tehnični (van der Graaf in Waaijers 2011).

Ravnanje s podatki mora biti načrtovano, vodeno, o njem teče komunikacija in je tudi ovrednoteno. Prav tako je pomembno vprašanje ocenjevanja vrednosti podatkov za nadaljnjo izrabo in izbor, ki upošteva povezljivost različnih tipov podatkov, potencial, ki ga podatki nosijo s seboj. Nekateri izčrpajo svojo vrednost že v prvi objavi in nato služijo bolj kot spremljajoča objava za zagotovilo preverljivosti.

Kakovostni podatki in kakovostne podatkovne storitve so v ospredju različnih mednarodnih institucionalnih pobud. European Strategy Forum on Research Infrastructure (ESFRI) v *Poročilu o ravnanju s podatki* iz leta 2009 kot enega pomembnejših ciljev na poti do e-znanosti izpostavlja kakovostne podatke in tudi kakovostne metapodatke. Poudarjajo, da mora biti zapisovanje metapodatkov dovolj prilagodljivo, da je mogoče dodajati nove elemente, različne tipe, ki bodo omogočali raznoliko rabo. Poročilo se podrobno posveča temam, kot so: raba, obseg, izvor, trajanje, agregiranje, standardizacija, povezljivost, kakovost, čas in razpoložljivost, s ciljem, da bi dosegli boljši izkoristek podatkovnega potenciala kot končnega merila kakovosti.

Podobne zaključke glede vprašanj kakovosti povzema tudi ekspertiza Kevina Ashleyja, direktorja Digital Curation Center, Velika Britanija. Med drugim izpostavlja pogosto neusklajenost ali parcialnost meril kakovosti, uporabljenih na različnih vrstah podatkov. »Vsi akterji v vseh fazah življenjskega kroga podatkov skrbijo za kakovost podatkov, študije pa kažejo, da imajo le redko enaka merila kakovosti: jaz želim obsežne podatke, ti jih želiš pravočasno, nekdo tretji želi točne podatke in četrti hoče prosti dostop.« Velikokrat je sicer te zahteve mogoče uskladiti, kdaj pa so si tudi v nasprotju, saj na primer obsežnih podatkov ni mogoče ustvariti hitro in poceni (Ashley 2011).

Pomemben vidik širše razumljene kakovosti storitev odprtega dostopa predstavlja upoštevanje in vključevanje več področij življenjskega kroga podatkov: določanje vsebine zbirk podatkov (načrtovanje obsega, opredelitev kriterijev za ocenjevanje kakovosti v smislu uporabnosti za drugo rabo in izbor), ustvarjanje podatkov (kako bodo podatki generirani) in dodana vrednost (čiščenje, dokumentacija, priprava na shranjevanje).

Merila strokovnosti je potrebno upoštevati tudi pri določanju pogojev dostopa. V določenem segmentu storitev je potrebno vsebinsko, ožje poznavanje načel stroke, ki ga je mogoče pridobiti s strokovnim usposabljanjem in v sodelovanju z izkušenimi raziskovalci. V tem se kaže velik pomen ožjih strokovnih področij pri opredeljevanju in izvajanju ukrepov za doseganje kakovosti.

POSTOPKI IN MERILA ZA ZAGOTAVLJANJE KAKOVOSTI V FAZI USTVARJANJA

Na splošno lahko rečemo, da so udeleženci naše raziskave izrazili veliko skrb za kakovost v fazi ustvarjanja podatkov, tako glede zavedanja pomena kakovosti kot glede poznavanja postopkov in meril. Na nekaterih področjih ustvarjalci natančno poznajo standardizirane postopke in se pri ustvarjanju držijo predpisanih protokolov (znanstvenik s področja biologije, 13, 14/5.7). Konsenz, ki so ga sprejeli v stroki, ni le ustno pravilo ali običaj, pač pa je opredeljen v pisni obliki kot neke vrste pravilnik ali protokol, ki ga je strokovnjak dolžan izpolniti. To pomeni, da pravila niso odvisna zgolj od osebnih navad, prepričanj, izkušenj posameznih nosilcev vlog, pač pa so formalizirana in institucionalizirana. V takšnih primerih imajo ustvarjalci preverjeno podlago, iz katere črpajo. Imajo odgovore na vprašanja: kaj je kakovosten postopek (npr. eksperiment), kako ga kakovostno izvesti ter kaj so kakovostni podatki in kako jih ustvariti. Zanimiv je primer s področja biologije, kjer je skrb za kakovost izjemno dodelana, do ravni certificiranja laboratorijev za določene meritve. V takem primeru je sistemsko poskrbljeno, da podatki dosežejo zahtevano kakovostno raven (znanstvenik s področja biologije, 15/5.7).

Posamezniki, skupine, laboratoriji uporabljajo tudi znanja in postopke, ki niso vselej formalizirani, pač pa se jih mora raziskovalec priučiti ali jih pridobiti z lastnimi izkušnjami. Nenapisana pravila so skrb vsakokratnega ustvarjalca in/ali njegove skupine (Gregor Anderluh, znanstvenik s področja biologije, 16, 17/5.7). Primer s področja ekonomije kaže, da so ponekod vzpostavljeni načini ocenjevanja kakovosti podatkov v smislu napake. Razumevanje je vzpostavljeno na ravni bližnjega poznavanja ustvarjalca podatkov, temelji pa na merilih stroke ter v tem okviru opredeljenih in uveljavljenih načinov preverjanja in zagotavljanja kakovosti (Polona Domadenik, znanstvenica s področja ekonomije, 18/5.7).

Med intervjuvanci so tudi takšni, ki imajo izkušnje z mednarodnimi zbirkami, npr. zbirko genskih podatkov. Na področju, s katerega prihaja intervjuvanec, so razvili dodatne aplikacije za zagotavljanje kakovosti podatkov, namenjenih hrambi. Pred oddajo podatkov je raziskovalec primoran iti skozi postopek, v katerem podatke opremi z ustreznimi metapodatki. Ta postopek je v pomoč pri preverjanju kakovosti razpoložljivih podatkov tako ustvarjalcem kot tudi uporabnikom (Gregor Anderluh, znanstvenik s področja biologije 19, 20/5.7).

Okvir 2.: P-R 2

PROBLEM:

Opazen je velik razkorak med ustvarjalci, ki delajo v okolju formaliziranih, področno in mednarodno standardiziranih pogojev, in drugimi, ki so pri delu prepuščeni lastni iznajdljivosti in izkušnjam.

Po drugi strani obstaja relativno visoka ozaveščenost o pomenu kakovosti na vseh ravneh ustvarjanja podatkov.

REŠITEV:

Neenake stopnje standardiziranosti so priložnost za deljenje izkušenj, dobrih praks in zgledov, ki lahko ustvarjalce z manj izdelanimi postopki motivirajo k večji standardizaciji. Izmenjava izkušenj bo spodbudila proces povečanja kakovosti, ki je eden od pogojev odprtega dostopa, in povečala možnosti bolj širokega izkoriščanja ustvarjenih podatkov. Ena od možnosti za

zbliževanje in izmenjavo dobrih praks je oblikovanje področnih mrež okoli problematike »odprtih podatkov«, ki naj presegajo posamezne laboratorije, oddelke, katedre idr. (primer spletišče SIDIH).

Področna srečanja lahko spodbudijo izmenjavo izkušenj in povečajo ozaveščenost o pomenu hrambe in dostopa do podatkov, ki presegata interes in potrebe vsakega posameznega raziskovalca in skupine.

Sicer pa je odnos do kakovosti pogosto izrazito usmerjen v ciljne produkte, denimo na področju jezikoslovja na slovarske produkte. Gradivo izbirajo premišljeno in ga sistematično čistijo (znanstvenik s področja jezikoslovja, 21, 22/5.7). Na področju hrambe glasbenih posnetkov ustvarjajo metapodatke vzporedno z nastajanjem in urejanjem podatkov in upoštevajo priporočila, sprejeta na mednarodni ravni. Poleg pomena kakovosti se zavedajo tudi dejstva, da je ta zmeraj kompromis med tehničnimi in vsebinskimi zmožnostmi (znanstvenik s področja zvočnega arhiviranja, 23, 24/5.7). Kakovost podatkov je po navadi ocenjena realistično, je nek »kompromis« v okviru možnega (Gregor Anderluh, znanstvenik s področja biologije, 25/5.7).

Na področju socialnega dela so podatki praviloma namenjeni konkretnemu projektu, zato v fazi ustvarjanja praviloma ne razmišljajo o drugi rabi niti nimajo razvitih orodij, ki bi jim bila pri tem v pomoč. Metapodatke in drugo dokumentacijo razumejo le tisti, ki so neposredno vključeni v postopek ustvarjanja podatkov. V takih primerih je kakovost podatkov, kar zadeva drugo rabo, nizka oziroma omejena. Za kontrolo poskrbijo tako, da zbrane in urejene podatke pred rabo dajo v preverbo intervjuvancem in tako odpravijo morebitne napake, ki bi bile posledica neustreznega razumevanja med spraševanim in spraševalcem (znanstvenik s področja socialnega dela, 26/5.7).

Merila, kaj so dobri podatki, temeljijo na določeni samoumevnosti, prepričanju o dobrih praksah, ki so vzpostavljene v okviru znanstvene skupnosti. Nekdo se priuči ravnanj, ki zagotavljajo dobre podatke, z izkušnjami in z usposabljanjem za člana znanstvene skupnosti (raziskovalec s področja arheologije, 27/5.7). Zagotovilo kakovosti podatkov je sama kakovost raziskovalca – njegova usposobljenost in znanje sta osnova za zaupanje, da dela kakovostno (vodja znanstvenoinformacijskega centra, 28/5.7).

Okvir 3: P-R3

PROBLEM:

Medtem ko so na določenih področjih raziskovalci oskrbljeni s standardiziranimi orodji in načini, so na drugih področjih prepuščeni lastni iznajdljivosti. Na ravni poznavanja postopkov in meril za zagotavljanje kakovosti obstajajo velike razlike. Ovire v dostopu, napake, neustrezna ali pomanjkljiva dokumentacija, neprečiščeni podatki idr. zmanjšujejo kakovost podatkov in otežujejo rabo ustvarjenih podatkov.

Na nizko kakovost vplivajo tudi tehnični, kulturni in pravni pogoji, ki so nemalokrat prepuščeni v presojo posamezni raziskovalni skupnosti ali posameznemu raziskovalcu.

REŠITEV:

Tako kot se morajo ustvarjalci podatkov spoznati z vsebino preučevanega področja, da bi lahko ustvarjali podatke, bo potrebno usvajati tudi znanja in veščine, ki zadevajo samo ravnanje s

podatki. Čeprav so med področji, kot so denimo družboslovje, humanistika, naravoslovje, medicina, velike in na videz nepremostljive razlike, kar zadeva vsebine in metode preučevanja, jih na področju ustvarjanja in ravnanja s podatki zadevajo vsaj deloma podobni problemi (kakovost, selekcija, hramba, dostop ...).

Rešitev problemov je treba iskati v smeri strokovnega usposabljanja, v obliki delavnic, tečajev, poletnih šol idr.

Dolgoročno je treba razmišljati o tem, da naj ravnanje s podatki postane obvezen del izobraževanja ustvarjalcev in uporabnikov podatkov ter strokovnega usposabljanja.

Zanimiv je tudi pristop, ki ga uveljavlja pobuda za pridobivanje potrdila o ustreznosti podatkov za dolgotrajno uporabo –*Data Seal of Approval* (DSA) (tabela 1). Za razliko od nekaterih drugih tovrstnih postopkov preverjanja, ki so usmerjeni na procese digitalne hrambe v ožjem smislu (glej pregled v Štebe in Vipavc Brvar 2011), ta pristop izraziteje poudarja vlogo dajalca pri ustrezni pripravi podatkov, tako da ti izpolnjujejo pogoje za vstop v sistem dolgotrajne hrambe.

Tabela 1: Prva tri vodila DSA

1. Ustvarjalec podatkov v podatkovni repozitorij izroči raziskovalne podatke, opremljene s spremljajočimi informacijami, na osnovi katerih je mogoče ovrednotiti znanstveno kakovost oddanega gradiva, skladno z merili posameznih disciplin in etičnimi zahtevami.
2. Ustvarjalec podatkov poskrbi, da je gradivo oddano v obliki, ki jo priporoča podatkovni repozitorij.
3. Ustvarjalec podatkov poskrbi, da so raziskovalni podatki opremljeni z metapodatki in da izpolnjujejo zahteve podatkovnega repozitorija.

Vir: [Data Seal of Approval](#)⁵²

V Sloveniji predstavlja primer podpore pri izročanju podatkov delovanje ADP. ADP je za dajalce podatkov pripravil podrobna navodila in obrazce za pripravo in oddajanje podatkov v arhiv. Usposabljanje za pripravo podatkov je tematika občasnih seminarjev, ki so namenjeni dajalcem (Štebe in Vipavc Brvar 2008). V sodelovanju s partnerskimi organizacijami ADP pomaga pri pripravi vodičev za dajalce (primerjaj podobne vodiče na [CESSDA](#)⁵³, [ICPSR](#)⁵⁴, [DATA PASS](#)⁵⁵, [UK DA](#)⁵⁶). S pomočjo tovrstnih navodil lahko vnaprej preprečimo nastanek težav pri izročanju podatkov v podatkovni center, ki bi se lahko pojavile ob nejasnih določilih obsega prostovoljnega pristanka in zaupnosti podatkov v okviru kvalitativnih družboslovnih raziskav (primerjaj Štebe in dr. 2011). Vprašanje zaupnosti je v širšem smislu povezano tudi z dostopnostjo podatkov statističnih uradov. V zvezi s tem je naloga dajalca, da (skupaj s podatkovnim središčem) opredeli postopke za anonimizacijo in kriterije glede zahtev zaupnosti. Kadar dajalec podatkov določi omejitve glede namena uporabe ali kategorije uporabnika, je naloga podatkovnega središča, da omogoči ustrezno registracijo uporabnikov in nadzira dostop do gradiv, upoštevajoč opredeljene omejitve gradiv glede dostopa in kako se te skladajo s pravicami različnih kategorij uporabnikov. Tovrstni sistem zahteva nadaljnjo uskladitev in

⁵² Glej: <http://www.datasealofapproval.org/>, 7. 8. 2012.

⁵³ Glej: <http://www.cessda.org/sharing/>, 7. 8. 2012.

⁵⁴ Glej: <http://www.icpsr.umich.edu/icpsrweb/ICPSR/access/deposit/>, 7. 8. 2012.

⁵⁵ Glej: <http://www.icpsr.umich.edu/icpsrweb/content/DATAPASS/citations.html>, 7. 8. 2012.

⁵⁶ Glej: <http://www.data-archive.ac.uk/deposit>, 7. 8. 2012.

sodelovanje na ravni mednarodne izmenjave in dostopa do podatkov, kar je ena od nalog projekta *Data Without Boundaries* (DwB), pri katerem sodelujejo partnerji CESSDA in Statističnih uradov v Evropi. Možna rešitev pri integraciji sistemov dostopnosti in različnih režimov dostopa do podatkov je, da se poenoti in med sabo poveže samo metapodatkovne opise, namenjene iskanju in presoji uporabnosti podatkovnih gradiv za sekundarno analizo s strani uporabnika (glej LeClere 2006), kar se nanaša na nekatere podatke statističnega urada. Pri tem pa fizično mesto dostopa še zmeraj ostane pri dajalcu.

IZBOR PODATKOV NA PODLAGI KAKOVOSTI

Vrednotenje in izbor podatkov na podlagi kakovosti nista samoumevna na način, kot je uveljavljen pri publikacijah, kjer obstajajo ustaljene uredniške politike za revije ali monografije in kjer recenzenti s svojimi ocenami zagotovijo kakovosten izbor. Problema smo se dotaknili tudi v intervjujih, in sicer smo spraševali, ali je mogoče izbor na podlagi kakovosti podatkov formalizirati po analogiji z recenzentskim postopkom, ki velja za znanstvene članke. Vtis je, da je o tem, kaj so dobri oziroma uporabni podatki, v okviru znanstvene skupnosti doseženo neko splošno soglasje.

Intervjuvanci pa so izpostavili tudi pogled, da so lahko zanimivi tudi slabi podatki, saj povedo več, kot če podatka ni in je potrebno znova investirati čas in energijo v iskanje, odkrivanje, ustvarjanje ... (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 29/5.7) Zato odgovor, da je upravičeno hraniti zgolj kakovostne podatke, ni sprejemljiv za vse. Izraženo je bilo mnenje, da je tudi npr. neformalen nadzor nad prispevki na način delovanja Wikipedije lahko zagotovilo kakovosti. V takem primeru je kakovost stvar soglasja, ki ga je mogoče doseči v diskusiji med zainteresiranimi in z uporabo moči strokovnih argumentov; to pomeni, da zanjo ni osnova stroga uredniška politika, recenziranje ipd., kot je uveljavljeno na področju znanstvenega publiciranja (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 30, 31/5.7).

Med intervjuvanci je prisotna skepsa do zunanjega ocenjevanja in vrednotenja tega, kaj so dobri podatki. Nekateri so zadržani do tega, da bi podatke na način urednikovanja publikacij obravnavala znanstvena knjižnica na ustanovi. Menijo, da mora biti to prepuščeno samemu raziskovalcu, razen v primerih, ko sprejem članka v objavo avtomatsko pomeni tudi potrditev ustrezne vrednosti podatkov. Tako kot v fazi ustvarjanja podatkov so merila za ocenjevanje kakovosti podatkov za izbor stvar ožjih strokovnih kompetenc področnih znanstvenikov. Knjižničarji sami menijo, da niso usposobljeni za vrednotenje, ki zadeva vsebinsko kakovost podatkov. Po drugi strani pa vidijo realno možnost v tem, da oblikujejo in vzdržujejo nadzor nad formalnimi kriteriji in omejitvami (vodja znanstvenoinformacijskega centra, 32, 33/5.7).

Izkazalo se je, da bi knjižničarji kvečjemu lahko bili podporniki sistema, sami pa tega projekta ne morejo začeti, saj se ne čutijo kompetentne niti ne poznajo raznolikosti področij in se ne morejo odločati o vsebini in kakovosti (vodja znanstvenoinformacijskega centra, 34, 35, 37/5.7, vodja knjižnice, 36/5.7). Prispevek knjižnice je lahko v tem, da so strokovnjaki za splošne metapodatke. Denimo tako, da zahtevajo minimalen, še zadosten formalni opis podatkov, ki vključuje zgodovino nastanka in življenja podatkov v okviru življenjskega kroga podatkov, se pravi dokumentacijo, po kateri je podatkovno gradivo določeno po svoji enkratnosti, ne presojajo pa drugih vidikov kakovosti.

Najpomembnejši zaključek glede kakovosti podatkov v fazi ustvarjanja je, da so podatki lahko tudi slabši, toda še vedno koristni. Podatki so lahko samo toliko dobri, kot je to mogoče doseči ob omejitvah glede usposobljenosti raziskovalca in pomočnikov, aparature in finančnih možnosti. Poleg tega je pomembno zagotoviti razumevanje dosežene ravni kakovosti podatkov kot konteksta, ki pogojuje njegovo uporabo. Prav neprečiščeni in pomanjkljivo dokumentirani podatki pogosto predstavljajo omejitev za dostop, saj druga raba zahteva še bolj eksplicitno predstavitev različnih vidikov, potrebnih za poglobljeno razumevanje nastanka in specifične narave podatkov.

Okvir 4: P-R 4

PROBLEM:

Vrednotenje in izbor podatkov med intervjuvanci nista nekaj samoumevnega. Nekateri zagovarjajo možnost, da se uredniško politiko nadomesti z izmenjavo argumentov znotraj zainteresirane javnosti in doseženim soglasjem.

Intervjuvanci niso enotni glede tega, ali bi lahko nekdo zunanji ocenjeval in vrednotil kakovost podatkov. Knjižničarji, ki bi lahko bili potencialni nosilci te vloge, se za enkrat čutijo nemočne in nepripravljene za vrednotenje podatkov.

REŠITEV:

Rešitve je mogoče iskati v zgledih iz tujine, saj je razprava o recenziranju in selekcioniranju podatkov trenutno zelo aktualna. Narašča število podatkovnih revij, v katerih sta izbor in sprejem objav omejena na formalne zahteve, ki jim morajo podatki zadostiti. To pomeni, da obstaja konsenz glede zahtev, ki morajo biti izpolnjene, da so podatki ustrezni za objavo, s tem pa tudi za drugo rabo.

Ena od možnih rešitev je razmislek v smeri prilagajanja uredniških politik, kot jih poznajo tradicionalne revije. Novi modeli recenziranja, ki jih preizkušajo v nekaterih revijah z odprtim dostopom in v podatkovnih revijah, npr. PLOS, kažejo na potrebe po splošnem spreminjanju tradicionalnih uredniških politik in recenzentskih postopkov, ne samo na področju vrednotenja podatkov.

V tem trenutku je mogoče problem zgolj določiti, možne rešitve pa so trenutno še v odprti debati. To vprašanje je povezano s širšim vprašanjem vrednotenja v znanosti, ki se prevečkrat omejuje na golo preštevanje objav, citatov idr. Potrebno bo spodbuditi tovrstne razprave, se vanje vključevati in uporabiti modele, ko bodo ti vzpostavljeni.

IZBOR PODATKOV V RAZLIČNE REŽIME DOSTOPA

Znotraj zbirk se podatki pogosto razlikujejo glede na intenzivnost obravnave, ki so je bili deležni, ki je običajno povezana tudi z merili kakovosti. Vseh podatkov ni mogoče pripraviti do ravni, da so prečiščeni in razumljivi brez dodatne razlage (npr. muzikološka zbirka). Zato se teži k reprezentativnemu izboru podatkov in k temu, da se izbrano posreduje v prečiščeni obliki, z dodano vrednostjo, se pravi z interpretacijami glede okoliščin in dodatnimi razlagami, ki gradivo postavijo v širši kontekst, potreben za razumevanje. Ta oblika dostopa do gradiv je že monografska publikacija, ki jo spremlja izbrano in prečiščeno podatkovno gradivo. To pomeni,

da poleg zbiranja podatkov ustvarjalci pokrijejo tudi druge faze, vse do integrirane objave besedila in spremljajočih podatkov. Ne gre samo za ustvarjanje podatkov, pač pa tudi za izdelavo končnih proizvodov (znanstvenik s področja geografije, 38/5.7).

Kadar podatke spremljajo pomanjkljivi metapodatki, ki so nastali sproti ob raziskovalnem delu, naj bi bil dostop omejen. Ker gre za gradiva, ki so nastala v preteklosti, ne da bi med raziskovanjem razmišljali, da jih bodo nekoč dali v rabo širši javnosti, so glede tega razumljivi pomisleki (znanstvenik s področja zvočnega arhiviranja, 39, 40/5.7). Ena od možnosti je, da se selekcionira, katera gradiva ostanejo za interno rabo in katera bodo dostopna javnosti. Intervjuvanec s področja hrambe glasbenih gradiv je izpostavil pomen sodelovanja z uporabniki. Gradivo, ki ni ali ne more biti ustrezno opremljeno z metapodatki, da bi omogočalo samostojno drugo rabo, terja, da se ustvarjalci podatkov in zbirk vključijo tudi v naslednje faze življenjskega kroga podatkov (znanstvenik s področja zvočnega arhiviranja, 41/5.7). Gre za podoben problem, kot smo ga omenili v uvodu pri razpravi o pomislekih, ovirah zaradi strahu pred zlorabami, napačnimi interpretacijami in so ga zaznali tudi v drugih sorodnih študijah.⁵⁷

Na podlagi pogovorov z intervjuvanci ugotavljamo, da se ustvarjalci podatkov in zbirk zavedajo odgovornosti v zvezi s kakovostjo podatkov, kakor se zavedajo tudi tega, da je zaradi odsotnosti standardnih postopkov druga raba precej omejena.

Okvir 5: P-R 5

PROBLEM:

Ustvarjalci podatkov in zbirk opravljajo svoje vloge skladno z razpoložljivimi možnostmi in se zavedajo, da so rezultati kompromis.

V obstoječih okoliščinah so zahteve hrambe v veliki meri prepuščene trenutni presoji posameznikov in skupin, razen v primerih, kadar svoje podatke oddajo v obstoječe zbirke, ki imajo svoje standarde in protokole.

REŠITEV:

Rešitev je oblikovanje kriterijev za vrednotenje pomena podatkov, kadar so podatki deležni posebne obravnave, npr. da se uvrstijo v zbirko podatkovnega centra. V okviru podpore načrtu ravnanja z raziskovalnimi podatki je potrebno stremeti k temu, da se zagotovi sredstva in pogoje za delovanje področnih podatkovnih centrov, ki bodo skrbeli za koordinirano pripravo »referenčnih« podatkov in opravljali storitev njihovega skrbništva ter posredovanja zainteresiranim drugim uporabnikom.

Zagotoviti je treba tudi rešitve in modele za podatke, ki so manjšega pomena, kar zadeva večkratno rabo. Npr. da se za revije oblikuje skupen repozitorij za podatke povezane z objavami, kot je na primer [DRYAD](http://datadryad.org/).⁵⁸

⁵⁷ »V razpravi o deljenju podatkov predstavljajo tveganja za zlorabo podatkov (in druge ovire) pomembno temo« (Cragin in dr. 2010).

⁵⁸ Glej: <http://datadryad.org/>, 7. 8. 2012.

PRIKAZ REZULTATOV V FAZI DODANA VREDNOST

Fazi zbiranja podatkov običajno sledijo določeni postopki, s katerimi zbrane surove podatke ustrezno pripravimo za nadaljnjo uporabo. Gre za predelovanje podatkov in za čiščenje oziroma odstranjevanje nepotrebne, odvečne, motečega, vse to pa po možnosti spremlja dokumentacija vsega, kar se s podatki dogaja, v metapodatkovnih zapisih. To fazo imenujemo dodana vrednost, kjer torej en del predstavljata predvsem čiščenje in transformacija podatkov, za kar raziskovalci poskrbijo že zaradi lastnih potreb, da podatke lahko analizirajo, drugi del pa je opremljanje podatkov z metapodatki, v čemer večina raziskovalcev ne vidi neposredne koristi. To se kaže v slabši skupni kakovosti podatkov z vidika možnosti druge rabe. Zato je še posebej pomembno, da se ustvarjalci podatkov zavedajo, katere spremljajoče informacije so potrebne za drugo rabo.

Intervjuvance smo prosili, da nam opišejo *stanje* v tej fazi življenjskega kroga podatkov ter katera orodja, postopke in navodila uporabljajo v tej fazi. Bolj natančno smo skušali izvedeti, kako poteka proces spreminjanja podatkov in dodajanja metapodatkov ter kakšne standarde pri tem upoštevajo, če sploh katere. Ali uporabljajo surove podatke ali jih pred analizo obogatijo? Kaj se dogaja s podatki od nastanka do točke spravila z vidika dodane vrednosti?

Zanimalo nas je, kakšni problemi pestijo fazo dodane vrednosti in ali jih intervjuvanci upoštevajo in rešujejo ali ne? Ali upoštevajo načrtovanje kakovosti, zahteve dolgotrajne digitalne hrambe, ali je zagotovljena sledljivost postopkov in transformacij? Ali razrešujejo probleme za zagotavljanje dostopa: lažja dostopnost, odstranjene ovire zaupnosti in podobno.

Da bi opredelili *odgovornosti* v zvezi z reševanjem problemov, smo jih prosili, da opišejo svoje dolžnosti, pravice, medsebojna pričakovanja in odvisnost nosilcev vlog. Kdo jih opravlja? Sami ali pooblaščen podatkovno središče, repozitorij ali specialist v okviru organizacije, centralni nacionalni digitalni arhiv, mednarodni centri ...?

Zanimalo nas je, kakšen je pomen politik, vodil, standardov pri opredeljevanju odgovornosti ter kakšne so ovire, sankcije, spodbude in nagrade. Ali dodana vrednost prinese nagrado? Ali dobi izvajalec povrnjene stroške? Ali so opredeljene obveznosti glede dodane vrednosti, metapodatkov (na ravni organizacije, discipline, nacionalno)? Kdo bi moral opredeliti obveznosti? Ali je za doseganje ciljev potrebno opredeliti obveznosti – enotna raziskovalna politika? Skušali smo identificirati *nosilce dobrih praks*, zato smo jih vprašali, kje dobijo navodila oziroma napotke za usposabljanje za delo na tem področju. So to morda področna središča za strokovno svetovanje, kdo je vodilni na tem področju, kje so nastavki dejavnosti?

Pokazalo se je, da na nekaterih področjih obstajajo že uveljavljeni standardi, na drugih morda bolj neformalni dogovori, na tretjih so stvari popolnoma prepuščene raziskovalcu. Intervjuvanec s področja zgodovine pravi, da so pogovori o metapodatkih na tem področju precej novi, o tem se govori šele kratek čas (znanstvenik s področja zgodovine, 1/6.3). Raziskovalec s področja naravoslovnih ved pravi, da v ustanovi nimajo enotnega predpisa o dokumentiranju, da se raziskovalec tega znanja priuči od mentorja med uvajanjem v raziskovalno delo (raziskovalec s področja naravoslovnih ved, 2/6.3), in dodaja, da so na tem področju podatki brez ustrezne

dokumentacije za drugega neuporabni (raziskovalec s področja naravoslovnih ved, 3/6.3). Podobno ugotavlja tudi znanstvenik s področja ekonomije (Polona Domadenik, znanstvenica s področja ekonomije, 4/6.3). Na področju zvočnega arhiviranja se zavedajo prednosti standardizacije, vendar še niso našli prave rešitve (znanstvenik s področja zvočnega arhiviranja, 5/6.3). Znanstvenik s področja gradbene informatike pravi, da je že zapis v XML jeziku korak v pravo smer k večji uporabnosti, hkrati pa bi se mu zdelo koristno, da bi država na področju predstavitve gradbenih objektov uporabljala standarde (znanstvenik s področja gradbene informatike, 6, 7/6.3). Na področju biologije so v splošnem že sprejeta pravila in standardi glede tega, kaj so kakovostni podatki (Gregor Anderluh, znanstvenik s področja biologije, 8/6.3), k njihovem upoštevanju pa dodatno prispevajo akreditirani laboratoriji, ki od ustvarjalcev podatkov zahtevajo določeno kakovostno raven (znanstvenik s področja biologije, 9/6.3). Tudi na področju meteorologije je mogoče govoriti o ustrezni standardizaciji podatkov (Nedjeljka Žagar, znanstvenica s področja meteorologije, 10/6.3), na področju fizike delcev pa poteka vzpostavitev infrastrukture, ki bo omogočala učinkovito hrambo in izmenjavo podatkov velikega obsega (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 11/6.3).

Znanstvenik s področja zvočnega arhiviranja ugotavlja, da raziskovalci deloma že sami po svoji vesti in skladno z utečenimi navadami pripravljajo spremno dokumentacijo, vendar pa jim za kaj bolj doslednega zmanjkuje časa (znanstvenik s področja zvočnega arhiviranja, 12/6.3). Pomanjkanje časa, pa tudi ustreznih znanj je ena od ključnih ovir, ki znižuje kakovost dokumentacije v fazi ustvarjanja. Za kakovostno in kontrolirano pripravo podatkov bi potrebovali veliko ur in ustrezno tehnologijo, česar v obstoječem sistemu preprosto nimajo, dodaja znanstvenik s področja zvočnega arhiviranja (znanstvenik s področja zvočnega arhiviranja, 13/6.3).

Intervjuvanci so izrazili potrebo po strokovni pomoči oziroma po kadrih, ki bi jih razbremenili potrebnega in nujnega dela. Znanstvenik s področja biologije ugotavlja, da bi za to potrebovali dodatno službo (znanstvenik s področja biologije, 14/6.3), podobno zaključuje tudi intervjuvanec s področja ekonomije (Polona Domadenik, znanstvenica s področja ekonomije, 15/6.3). Več o tem bomo spregovorili nekoliko kasneje.

Splošno prevladujoče stanje je, da znotraj inštitucije običajno ni službe, ki bi skrbela za arhiviranje, niti ni skupne politike glede postopkov dolgotrajne digitalne hrambe. Skrb za hrambo podatkov in ostalih gradiv je tako večinoma stvar metodologije posameznih raziskovalcev ali raziskovalne skupine, ki med delom običajno sami oblikujejo svoja neformalna pravila. Je pa izražen interes, da se podatki hranijo čim bližje nastanku in da raziskovalec ohrani nadzor nad tem, kaj se s podatki dogaja (znanstvenik s področja geografije, 16/6.3). Takšen model skrbništva nad podatki s strani specialistov se v določenem pogledu uveljavlja tudi v okviru disciplinarnih arhivov, tak primer je mreža podatkovnih središč⁵⁹ pod okriljem britanskega Natural Environment Research Council ([NERC](http://www.nerc.ac.uk))⁶⁰, ki ima skupen katalog raziskovalnih podatkov⁶¹.

⁵⁹ Glej: <http://www.nerc.ac.uk/research/sites/data/>, 7. 8. 2012.

⁶⁰ V predstavitvi na svoji spletni strani so zapisali, da so njihovo največje premoženje podatki, ki jih sami ustvarijo. Podatkovna središča zagotavljajo podporo in usmerjanje pri ravnanju s podatki, odgovorni so za dolgoročno skrb in zagotavljajo dostop do podatkov. Krovna politika določa, da so vsi podatki, ki so financirani s strani NERC, dostopni javno in brez omejitev. Od raziskovalcev zahtevajo pripravo načrtov ravnanja s podatki. Oblikovali so seznam

Tudi iz različnih študij, narejenih drugod, je mogoče ugotoviti, da je samo manjši del podatkov deležen obravnave s strani podatkovnega središča, prav tako pa tudi, da je uporabnost podatkov pogosto omejena na ožji krog uporabnikov ali pa je predvidena dokaj redka prihodnja druga uporaba (Feijen 2011, 19–20). Te ugotovitve so pomembne pri vprašanju vrednotenja in ocenjevanja potenciala polne uporabe raziskovalnih podatkov. Vsekakor so tu problemi in koristi nekoliko drugačni kot pri klasičnih raziskovalnih poročilih v obliki publikacij. Pogosto so raziskovalnemu projektu zunanji podatki samo za dodatno evidenco ali pa primerjavo in umerjanje inštrumentov, vendar tudi kot taki nepogrešljiv del v mozaiku sestavljanja novih kakovostnih spoznanj.⁶²

ČIŠČENJE, TRANSFORMACIJA PODATKOV IN METAPODATKI

Z intervjuvanci smo se pogovarjali o tem, ali morajo surove podatke, preden jih lahko uporabijo, kakor koli obdelati in s kakšnimi postopki to storijo. Iz odgovorov sledi, da je večino podatkov potrebno vsaj minimalno obdelati, predno so uporabni za analizo, na nekaterih področjih pa so brez dodatne obdelave celo povsem neuporabni.

Surovi podatki so osnova, v nadaljnjih raziskovalnih postopkih, tj. z rabo različnih metod in tehnologij, pa se spreminjajo, preoblikujejo, čistijo. Končni proizvod so lahko podatki v različnih verzijah, ki so uporabni za različne raziskovalce. Za hrambo podatkov, torej za arhiviranje, niso pomembni le surovi podatki in transformirani podatki, pač pa tudi zapisi o transformaciji, programska oprema in druge spremljajoče informacije, ki zagotavljajo sledenje spremembam. Problem je tudi uporaba večinoma licenčnih programov, saj bi se transparentnost v smislu dolgotrajne hrambe povečala z rabo odprtokodnih programov, pri katerih lahko arhiv lažje preveri in ohrani izvorno kodo. Sledenje smernicam ISO OAIS standarda in na njih temelječih dobrih praksah pri oblikovanju in vzdrževanju informacijskih paketov prav tako velja kot garancija za kakovostno hrambo (Griffin in dr. 2012).

Kakšni so ti postopki, je zelo odvisno od tipa podatkov, so pa večinoma pravila za čiščenje in transformacijo podatkov vsaj neformalno opredeljena kot uveljavljene prakse na posameznih inštitutih. Podatke, ki jih generira neka naprava, ki mora že sama po sebi zadostiti določenim standardom, je običajno potrebno le vizualno prečistiti, pri čemer se izločijo artefakti, mora pa biti oseba, ki to počne, zadostno usposobljena. Na področjih, kjer pogosto kombinirajo več različnih podatkovnih baz, se pred analizo redno preverja konsistentnost podatkov in v primeru nenavadnih odstopanj se za konkretne primere poskuša ugotoviti razlog odstopanja (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 17/6.3). Pri zgodovinarjih je zelo uveljavljena praksa preverjanja pravilnosti posameznega podatka v različnih zgodovinskih virih. Raziskovalci, ki izvajajo poglobljene kvalitativne intervjuje, potem ko naredijo prepis intervjuja, intervjuvance pogosto prosijo za avtorizacijo besedila (medtem ko protokola za natančnost pri transformaciji zvočnega zapisa v besedilo nimajo) in s tem zagotovijo kar najmanj izgub informacij, ki nastanejo v procesu transformacije (znanstvenik s področja socialnega dela, 18/6.3). Postopek čiščenja je tako posebej zahteven pri avdio-vizualnih gradivih in besedilih (besedilni korpusi).

kriterijev, po katerem je mogoče določiti, kateri podatki so smiselni za dolgoročno hrambo. Vir:

<http://www.nerc.ac.uk/>, 7. 8. 2012.

⁶¹ Glej: <http://data-search.nerc.ac.uk/>, 7. 8. 2012.

⁶² Glede distinkcije med podatki v ozadju in v ospredju, ko gre za različne uporabe (kar je podatek v ozadju za nekoga, je za drugega primarna podlaga analiz in objave), glej npr. Wallis in dr. (2012).

Na nekaterih področjih pa so standardi za kakovost podatkov še posebno ustaljeni in jasni, npr. v naravoslovju obstajajo politike obravnavanja podatkov in je kakovost podatkov mogoče preverjati tudi za nazaj, ko imamo na podlagi analize podatkov že rezultat (Gregor Anderluh, znanstvenik s področja biologije, 19/6.3). Za zbiranje prostorskih podatkov obstaja ISO standard za kakovost (raziskovalec s področja geodezije, 20/6.3), za zbiranje zvočnih in video gradiv obstajajo priporočila za delo na terenu, ISO, AES, IASA (znanstvenik s področja zvočnega arhiviranja, 21/6.3), ponekod pa standarde za kakovost določi kar skupina strokovnjakov v metodologiji, ki jo pripravijo na začetku projekta (znanstvenik s področja zgodovine, 22/6.3). Vsi intervjuvanci dobro razumejo, da je za zagotavljanje uporabnosti podatkov v dostopu potrebna dodatna dokumentacija.

Za določen segment podatkov, tistih, ki so namenjeni za mednarodne zbirke (npr. GenBank, UniProt, Protein Data Bank, Amadeus), je zagotovljena popolna dokumentacija nastajanja podatkov. Postopki dodane vrednosti (normalizacije, čiščenja) podatkov so dokumentirani; oddajo se tako surovi podatki kot normalizirani podatki. Spremljajoča kakovostna, izčrpna in standardizirana informacija o poteku eksperimentov je v tem segmentu del vzpostavljene prakse. Tovrstne mednarodne povezave so pomembne tudi zaradi izmenjav znanja pri ustvarjanju in dokumentaciji ter vrednotenju raziskovalnih podatkov.

V določenih krogih, v katerih je kultura izmenjave podatkov že dalj časa prisotna, npr. genomske raziskave, je ravno potreba po izmenjavi podatkov spodbudila vzpostavitev in uporabo metapodatkovnih standardov.⁶³ Raziskovalci takšnih znanosti se zato zelo dobro zavedajo problemov protokolov in točnega zapisa, poleg tega pa so tudi zavezani k njihovi uporabi pri objavi podatkov v mednarodnih zbirkah. Takšne podatkovne zbirke imajo običajno za vnos podatkov že narejene aplikacije, ki zahtevajo določene metapodatke (Gregor Anderluh, znanstvenik s področja biologije, 23/6.3).

Podoben primer dobre prakse je tudi področje meteorologije, kjer so protokoli za zbiranje podatkov in standardi za metapodatke že zelo dolgo splošno sprejeti in uporabljeni (Nedjeljka Žagar, znanstvenica s področja meteorologije, 24/6.3). Njihovo uveljavitev in vsesplošno uporabo je prav tako spodbudila potreba po medsebojni izmenjavi tovrstnih podatkov, ki pa se kaže tudi na drugih področjih, npr. v gradbeništvu, kjer informacijsko bolj razvite države zahtevajo dokumentacijo v standardnih oblikah (ISO-Step, IFC) (znanstvenik s področja gradbene informatike, 25/6.3).

Pomembnosti metapodatkov se dobro zavedajo tudi intervjuvanci, ki se ukvarjajo z avdio in video gradivom. Tovrstne podatke običajno spremljajo vsebinsko zelo bogati metapodatki v obliki zapisnikov, ki vsebujejo bolj ali manj strukturirana polja za tehnične in vsebinske informacije. Brez tega bi bila uporabnost gradiva znatno manjša (znanstvenik s področja zvočnega arhiviranja, 26/6.3).

Na nekaterih področjih, na katerih še ni uveljavljenih standardov raziskovanja in dokumentacije podatkov, so se posamezne inštitucije znašle na različne načine: za zapis metapodatkov uporabljajo kar Dublin Core, ki ga za potrebe svoje zbirke podatkov dopolnijo z dodatnimi polji, metapodatke zapisujejo v nestandarden XML zapis, ki ga je razmeroma enostavno brati, v eni od inštitucij pa so se lotili tudi priprave internega pravilnika o organizaciji raziskovalnega dela, ki izhaja iz načel projektnega menedžmenta (znanstvenik s področja socialnega dela, 27/6.3).

⁶³ Standard MIAME, <http://www.mged.org/Workgroups/MIAME/miame.html>, 7. 8. 2012.

V povezavi z zagotavljanjem metapodatkov smo naleteli tudi na več zanimivih manjših projektov, kot sta npr. projekt Fototeka, ki je vključeval digitalizacijo lastnega arhiva slik in opremljanje z metapodatki (vodja znanstvenoinformacijskega centra, 28/6.3), ter jezikovni korpus, ki uporablja metapodatkovni standard TEI (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 29/6.3).

Kljub temu, da že obstaja kar nekaj dobrih praks za specifične tipe podatkov, pa je dokumentacija podatkov v fazi nastajanja najpogosteje še vedno prepuščena navadam in potrebam samega raziskovalca. Znanje in strogost glede priprave in sprotne natančne dokumentacije podatkov je priučeno znanje, stvar notranje kulture raziskovalne skupine ali celo pazljivosti, discipline posameznega raziskovalca; pogosto metapodatke predstavlja raziskovalčev prosti opis eksperimenta, ki podatke spremlja v obliki delovnega poročila ali programske dokumentacije in nima vnaprej definiranih polj, po katerih bi podatke lahko klasificirali in iskali v katalogu.

Podatki, ki jih generirajo naprave, so že ob nastanku opremljeni z osnovnimi metapodatki, pri novejših aparaturah je mogoče vpisati tudi veliko dodatnih informacij, kar pa se raziskovalci le redko uporabljajo.

Eden od razlogov, da večina organizacij nima predpisanih standardov za metapodatke, je, da so področja pogosto zelo kompleksna in si odgovorni ne predstavljajo, kako bi zagotovili poenotene predpise za različne discipline. Za mnenje smo povprašali tudi knjižničarje, kot strokovnjake na področju bibliografije, ki pa rešitve pri opisovanju podatkov prepuščajo raziskovalcem, sami se namreč ne vidijo v kakšni bolj aktivni vlogi, ne pri opisovanju in pripravi podatkov ne pri njihovem vrednotenju (vodja znanstvenoinformacijskega centra, 30/6.3). Prispevek knjižničarjev bi lahko bil, da se usposobijo za sprejemanje različnih vrst gradiv, vključno z raziskovalnimi podatki, v zajem nacionalne bibliografije. Uveljavljeni bibliografski standard COMARC je po mnenju nekaterih dovolj širok, da bi lahko pokrival tudi raziskovalne podatke.

Očitno je, da obstajajo v znanstveni skupnosti tudi znanja in izkušnje s pripravo podatkov za podatkovni arhiv, saj smo naleteli na kar nekaj dobrih praks z različnih področij. Področja, kjer so te prakse vzpostavljene, lahko služijo kot nastavek za reševanje problemov na drugih področjih. Te dobre prakse bi bilo potrebno izpostaviti in na njih graditi, vključno z usposabljanjem za uporabo standardov pri opisovanju, kot je MIAME. Za določene, t. i. bolj »tradicionalne« eksperimente, pa je pristop bolj »akademski« v tem smislu, da postopki niso tako strogi, marsikaj se v hitenju k analizi zapiše bolj na hitro, se opusti.

Okvir 6: P-R 6

PROBLEM:

Inštitucije večinoma (še) niso oblikovale in sprejele pravil za uporabo metapodatkovnih standardov, tako je dokumentacija podatkov največkrat prepuščena samim raziskovalcem.

Večina raziskovalcev dokumentiranje podatkov še vedno vidi le kot dodatno delo, za katerega nimajo ne časa ne sredstev. Nekateri ne vidijo prednosti zaradi uporabe standarda oziroma menijo, da zanje niso dovolj velike, da bi v to vlagali svoj čas in znanje.

Ugotovili smo tudi, da znanje sicer obstaja, vendar ni spodbude s strani institucij in financerja, da bi se dejansko uporabilo.

REŠITEV:

Rešitev je treba iskati v smeri povezovanja skupin ekspertov na disciplinarnih področjih, ki bodo svetovali (dobri zgledi in prakse) pri pripravi metapodatkov že v fazi ustvarjanja podatkov. Hkrati pa bodo svetovali tudi pri iskanju specializiranih storitev hrambe in dostopa do podatkov. Potrebno bo spodbuditi področja, na katerih obstajajo znanja in razvojne dejavnosti, ter vzpostaviti evidenco ekspertnih skupin in njihovih kompetenc in poskrbeti za institucionalizacijo zagotavljanja njihovih svetovalnih storitev.

S tem se odpira potreba po oblikovanju novega poklica in novega kadra. Hkrati bo treba preučiti, kakšen del nove storitve bodo lahko prevzeli knjižničarji. Knjižničarje, kot prejemnike gradiva v repozitorije, bi bilo potrebno dodatno usposobiti glede klasifikacij, tipov podatkovnih gradiv in za podporo pri pripravi podatkov za objavo.

SPRAVILO

Produkti ustvarjanje in transformacija podatkov skupaj z njihovo zgodovino so predmet spravila podatkov, skrbi za njihovo dolgotrajno razpoložljivost tako za ustvarjalce kot tudi za drugo rabo. Intervjuvance smo prosili, da nam opišejo korake v fazi spravila. Zanimalo nas je, kaj se dogaja s podatki, kakšne vrste in formati podatkov nastopajo; kje podatke hranijo; kakšna znanja obstajajo na področju hrambe; katera orodja, metapodatke, standarde in postopke uporabljajo. Ali imajo načrte za hrambo? Ali poznajo *Open Archival Information System* (OAIS)?

Nadalje nas je zanimalo, kako opredeljujejo probleme, povezane z drugimi fazami življenjskega kroga podatkov. Ali je shranjevanje namenjeno bolj za lastne potrebe ali pa upoštevajo možnost, da ob koncu projekta podatki služijo tudi drugim? Kakšne so prednosti hrambe na obstoječi lokaciji ali centralizirano po disciplinah? Katere ciljne vsebine (*intellectual content*) so kot bistvene namenjene hrambi? Ali je pomen podatkov strokovno ovrednoten, v smislu dolgotrajne vrednosti za projekt in za druge? Kaj se zgodi s podatki majhne vrednosti? Ali gre za »enkratne« podatke, ki jih je mogoče z lahkoto reproducirati? Ali je zagotovljeno ohranjanje vrednosti podatkov? Ali je razpoložljivost podatkov ogrožena zaradi odsotnosti hrambe? Ali je v okviru sistema hrambe zagotovljena sledljivost izvora podatkov in postopkov v predhodnih fazah?

Intervjuvance smo prosili, naj pojasnijo, kako vidijo odgovornost v zvezi z reševanjem problemov v fazi spravila podatkov. Kje vidijo primerno mesto za odlaganje podatkov? Kakšen je pomen politik, vodil in standardov pri opredeljevanju odgovornosti? Ali imajo merila in vrednotenje kakovostne hrambe? Kaj so ovire, sankcije, spodbude in nagrade? Kdo plača hrambo, kdo skrbi za usposabljanje glede hrambe? Kakšni so pravni problemi hrambe: prenos pravic, zaščita podatkov? Kateri so primeri dobrih praks, zgledov, ki bi lahko bili izhodišča za razvoj te faze podatkov?

V ustanovah, iz katerih prihajajo naši intervjuvanci, ne samo da ustvarijo ogromno raznovrstnih podatkov, pač pa obstaja tudi množstvo različnih sistemov hrambe tako z vidika formatov kot same organizacije spravila gradiva in skrbi za dolgotrajno hrambo. Če govorimo o elektronskih gradivih, se ta največkrat hranijo na osebnih računalnikih, včasih pa se dodatno prenesejo še na kakšen drug medij ali na strežnik, za katerega se periodično dela varnostna kopija (Polona Domadenik, znanstvenica s področja ekonomije, 31/6.3). Ponekod imajo tudi vzpostavljen sistem sistematičnega shranjevanja gradiv na skupni strežnik, do katerega lahko dostopajo vsi

zaposleni, običajno uporabljajo sistem kronološkega shranjevanja (znanstvenik s področja biologije, 32/6.3, raziskovalec s področja arheologije, 33/6.3). Pri gradivih v fizični obliki, teh je bilo manj, pa prevladuje praksa, da se podatki hranijo v lastnih prostorih ali pa v za to namenjenih ustanovah (npr. muzejske zbirke, vzorci tkiv, zbirka jamskih živali ipd.) (znanstvenik s področja favnistike jamskih nevretenčarjev, 34/6.3, znanstvenik s področja medicinskih vez in statistike, 35/6.3).

Če se vrnemo nazaj k podatkom v elektronski obliki, se v povezavi z različnimi mediji, na katerih te podatke hranijo, pojavlja velik problem zastarelosti medijev. Pri starejšem gradivu, ki obstaja ali v fizični obliki ali pa na zastarelem mediju, so se marsikje lotili tudi digitalizacije, ki pa je zelo dolgotrajna in zahtevna (znanstvenik s področja zvočnega arhiviranja, 36/6.3, znanstvenik s področja zgodovine, 37/6.3, raziskovalec s področja arheologije, 38/6.3).

Tudi na področju formatov največji problem predstavlja zastarelost, zato marsikje poskušajo uporabljati čim bolj standardne formate (ASCII) ter programe in orodja, ki so uporabni na dolgi rok (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 39/6.3). Pogosto se podatki hranijo v formatu, ki je vezan na program, v katerem podatke analizirajo. Podatkovne baze ponekod hranijo v tekstovnem formatu ali Excelu, zraven pa hranijo tudi sintakso in merski instrument (znanstvenik s področja socialnega dela, 40/6.3). Včasih so formati specifični za projekt ali pa so lastni posamezni napravi za generiranje podatkov. Ko se z aparaturo zamenja tudi format, običajno proizvajalci poskrbijo za možnost izvoza v bolj univerzalen format, npr. ASCII (Gregor Anderluh, znanstvenik s področja biologije, 41/6.3).

Na nekaterih področjih so za potrebe lastnega shranjevanja sami razvili (odprtokodne) aplikacije, na primer program Neva (aplikacija za shranjevanje, iskanje in prikazovanje gesel), ki deluje vzporedno s spletno aplikacijo Eva, ki podatke v obliki HTML-ja prikazuje na spletu (znanstvenik s področja jezikoslovja, 42/6.3).

V redkih primerih, kakršno je področje meteorologije, za hrambo podatkov poskrbijo operativne nacionalne službe (ARSO), shranjevanje podatkov pa je urejeno na svetovni ravni (WMO) (Nedjeljka Žagar, znanstvenica s področja meteorologije, 43/6.3).

Pri spravilu gradiva težavo pogosto predstavljajo izredno velike datoteke, ki zavzamejo ogromno prostora, npr. podatki v obliki slik v visoki ločljivosti, digitalizirano besedilo v formatu TIF, avdio datoteke ipd. Zaradi prostorskih omejitev podatke tudi brišejo (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 44/6.3).

Največji problem, ki so ga intervjuvanci izpostavljali, je ta, da v okviru projekta trenutno ni mogoče poskrbeti za dolgoročno hrambo, saj se z zaključkom projekta zaključijo tudi financiranje (znanstvenik s področja geografije, 45, 46/6.3). Več intervjuvancev je izrazilo željo, da bi bila na višji ravni (nacionalni, mednarodni) vzpostavljena storitev dolgotrajne hrambe, kot nek centraliziran sistem, sami namreč nimajo ustreznega znanja za to. V povezavi s tem se bojijo prevelikih stroškov. Nekateri izražajo pričakovanje, da bi lahko razmere glede hrambe digitalnega gradiva, tudi kar zadeva raziskovalne podatke, urejal obstoječi krovni Zakon o varstvu dokumentarnega in arhivskega gradiva ter arhivih (ZVDAGA) (raziskovalec s področja arheologije, 47/6.3).

Veliko strokovnega znanja glede ravnanja s podatki in na njih zgrajenih storitvah sicer obstaja, a to ostaja razpršeno. Ohranjanje in večja raba tega vira intelektualnega kapitala bi lahko bila

predmet oblikovanja nacionalnih usmeritev glede razvoja. Prav človeški kapital je največja vrednost, ki se pogosto zaradi neustrezne organiziranosti izgublja, poleg tega da odsotnost usposobljenih skrbnikov predstavlja tudi ogrožanje dragocenih podatkovnih virov (primerjaj Ashley 2011). Ta tema, se pravi samoocena strokovne usposobljenosti za različne faze zagotavljanja storitev pri ravnanju s podatki, in izražena potreba po tovrstni podpori je ena od izpostavljenih tem pri določanju potencialov tudi za stanje v naši državi. V smeri načrtovanja prihodnje ureditve gre s tem v zvezi tudi za vprašanje, kako zagotoviti prenos te lokalne strokovnosti na širši krog uporabnikov, npr. na način svetovanja ali pa na osnovi modela skupnosti, ki jo gradijo uporabniki določenih podatkov. V naši raziskavi je načelno naklonjenost in pripravljenost za prevzem svetovalne vloge pokazal denimo intervjuvanec s področja medicinskih ved in statistike.

Intervjuvanci se zavedajo, da je vprašanje digitalne hrambe in skrbništva, če je zastavljeno na trdnih strokovnih temeljih, precej kompleksno področje, ki presega kompetence delujočih znanstvenikov ter tudi časovne in finančne zmožnosti tekočih projektov (znanstvenik s področja zgodovine, 48/6.3). Iz tega izhaja potreba po centraliziranih in strokovno vodenih storitvah za zagotavljanje ustreznih pristopov ter uporabe orodij in medijev. Tovrstno vlogo bi lahko zagotavljali strokovnjaki, ki so sodelovali na [Slovenski konferenci o trajnem ohranjanju digitalnih virov](#),⁶⁴ tehnološke rešitve in strojna oprema pa so vsakič stvar optimalnih izbir danega trenutka. Tudi v primerjavi z drugimi raziskavami, ki smo jih omenili v uvodu (npr. Parse.Insight), je splošni vtis, da je vprašanje digitalnega skrbništva in digitalne hrambe šele začelo postajati ena od pomembnih sestavin ravnanja s podatki znotraj življenjskega kroga podatkov. Pomembne so pobude za strokovno usklajevanje nacionalnih pristopov.⁶⁵ Razvoj gre v smeri, da bodo raziskovalci potrebovali pomoč tistih strokovnjakov, ki imajo posebna znanja in izkušnje s področja ravnanja s podatki in ki razumejo obseg in vsebine življenjskega kroga podatkov. Se pravi, da je potrebno razlikovanje med strokovnimi kompetencami ustvarjalca podatkov in strokovnim znanjem za opravljanje storitev skrbništva in ravnanja s podatki. Slednja so posebna znanja, čeprav pogosto v tesni povezavi z ožjimi strokovnimi področji posameznih znanstvenih disciplin – kako dokumentirati, urejati, hraniti, da bodo podatki dostopni in uporabni tudi za druge, ne le za ustvarjalce. Pri tem se pojavlja potreba po dveh vlogah: prva se nanaša na samo izvajanje storitve, druga pa kot pomoč v smislu svetovanja o dobrih praksah in zgledih.

Sicer pa smo na posameznih področjih identificirali kar nekaj zbirk, ki so nastale v okviru različnih projektov, primarno z namenom, da lahko do njih dostopajo različni uporabniki, in jih je mogoče razumeti kot zametke dobrih praks:

- v [Centru za kartografijo favne in flore](#)⁶⁶ generirajo biološke podatke (razširjenost živalskih in rastlinskih vrst), lahko bi jih vzeli kot referenco za to, kako poskrbeti za posredovanje podatkov drugim;
- [Eva in Neva](#)⁶⁷ – jezikoslovci so razvili lastno aplikacijo za shranjevanje, iskanje in prikazovanje gesel, sestavljeno iz dveh delov: Neva teče na strežniku, drugi del, Eva, pa skrbi za prikazovanje v obliki HTML-ja preko spletnih brskalnikov (koda je odprta, lahko

⁶⁴ Glej: <http://www.nuk.uni-lj.si/todv/konferenca/>, 7. 8. 2012.

⁶⁵ Glej: *Aligning National Approaches to Digital Preservation: An Action Assembly* Biblioteca de Catalunya (National Library of Catalonia), 18.–20. november 2013, Barcelona, Španija, <http://www.educopia.org/events/ANADPII>, 3. 1. 2014.

⁶⁶ Glej: <http://www.ckff.si/>, 7. 8. 2012.

⁶⁷ Glej: <http://www.laze.org/>, 7. 8. 2012.

se dodajajo nova polja, če je potrebno). Prikaz za splošno rabo (on-line dostop) je omejen, za lokalno rabo pa lahko raziskovalci dobijo tudi daljša besedila. Program Eva uporabljajo tudi biologi, arheologi na [Inštitutu za raziskovanje Krasa](#),⁶⁸

- jezikoslovci uporabljajo tudi Wikiverzo (kot vadnico, ni dostopna javnosti), [Wikivir](#)⁶⁹ (za tekste) in [Wikipedijo](#)⁷⁰ (za besedilno zbirko);
- [slovarske in besedilne zbirke](#), objavljene na spletu;⁷¹
- [Gigafida](#) – korpus slovenskega jezika;⁷²
- slovensko leposlovje na spletu;⁷³
- slovenski zgodovinski roman;⁷⁴
- slovenska kmečka povest;⁷⁵
- Zgodovina Slovenije – [Sistory](#)⁷⁶ je portal, na katerem so poleg digitalnih objavljene tudi digitalizirane zbirke podatkov (slike, tiskano gradivo, ustno) s področja zgodovine. Pri objavljanju določenih gradiv je pomembno vprašanje avtorskih pravic;
- [ARKAS](#)⁷⁷ je osnovna baza arheoloških najdišč Slovenije, v njej so združene podatkovne, tekstovne baze in kartografske podlage. Enota je arheološko najdišče, ki ga opišejo na različne načine, da ga lahko povežejo z drugimi bazami, ki niso njihove. Ponuja možnosti iskanja. So eni prvih v Sloveniji, ki so se lotili zametka informacijskega sistema v digitalni obliki, ki naj bi predstavljal eno infrastrukturo za celotno arheološko dejavnost;
- na področju humanistike in umetnosti je povezovanje digitaliziranih podatkov v precejšnjem zagonu. Mednarodna pobuda [DARIAH](#)⁷⁸ je projekt, s katerim skušajo podpreti in okrepiti digitalno raziskovanje;
- v povezavi z DARIAH se razvija portal slovenske znanstvene digitalne humanistike, imenovan [SIDIH](#).⁷⁹ In sicer s ciljem, povezati različne baze podatkov ter poenotiti metapodatke za različna gradiva;
- primer specializiranega spletišča s področja humanistike je [Etnoinfolab](#),⁸⁰ ki predstavlja slikovno, zvočno in filmsko gradivo s področja etnologije in kulturne antropologije;
- [DEDI](#)⁸¹ – Enciklopedija naravne in kulturne dediščine na Slovenskem;
- [Arzenal](#)⁸² je portal, na katerem inštituti ZRC SAZU objavljajo vsebine, ki jih želijo dati v uporabo javnosti; metapodatki za gradiva so v besedilni obliki.

Takšne zbirke običajno obsegajo tudi različne vidike podatkovnih storitev (npr. prevzemajo gradiva več avtorjev, omogočajo dostop itd.). Sredstva za financiranje vzdrževanja in nadgrajevanja teh zbirk so v glavnem pridobljena s projekti, zaradi česar ni dolgoročno zagotovljenega financiranja in je ogrožena kontinuiteta, kar skrbniki izpostavljajo kot hud problem.

⁶⁸ Glej: <http://izrkc.zrc-sazu.si/#v>, 7. 8. 2012.

⁶⁹ Glej: http://sl.wikisource.org/wiki/Glavna_stran, 7. 8. 2012.

⁷⁰ Glej: http://sl.wikipedia.org/wiki/Glavna_stran, 7. 8. 2012.

⁷¹ Glej: <http://bos.zrc-sazu.si/>, 7. 8. 2012.

⁷² Glej: <http://www.fidaplus.net/>, 7. 8. 2012.

⁷³ Glej: <http://slovenskaliteratura.ff.uni-lj.si/sl.html>, 7. 8. 2012.

⁷⁴ Glej: http://lit.ijs.si/zgr_dons.html, 7. 8. 2012.

⁷⁵ Glej: <http://lit.ijs.si/skp1.html>, 7. 8. 2012.

⁷⁶ Glej: <http://www.sistory.si/>, 7. 8. 2012.

⁷⁷ Glej: <http://arkas.zrc-sazu.si/>, 7. 8. 2012.

⁷⁸ Digitalna infrastruktura za humanistiko in umetnost: <http://www.dariah.eu/>, 7. 8. 2012.

⁷⁹ Glej: http://www.sidih.si/kdo_smo, 7. 8. 2012.

⁸⁰ Glej: <http://www.etnoinfolab.org/>, 7. 8. 2012.

⁸¹ Glej: <http://www.dedi.si/>, 7. 8. 2012.

⁸² Glej: <http://www.arzenal.si/>, 7. 8. 2012.

Okvir 7: P-R 7

PROBLEM:

Največji problem na področju hrambe je, da stvari niso urejene na ravni institucionalnih politik in je skrb za hrambo v največji meri stvar metodologije vsakega posameznika.

Večina raziskovalcev je povedala, da v okviru projekta nimajo zagotovljenih finančnih sredstev, s katerimi bi lahko zagotovili dolgoročno hrambo.

Poleg tega se srečujejo še z drugimi ovirami, kot so hitra zastarelost medijev in formatov, pomanjkanje prostora ipd.

REŠITEV:

Potrebno bo oblikovati politike, ki bodo upoštevale načrtovanje ravnanja z raziskovalnimi podatki. Tako bo tudi financiranje dela s podatki že vnaprej načrtovano in kot tako tudi zagotovljeno.

Pri premagovanju tehnoloških rešitev je smiselna mednarodna integracija: katalogi, orodja itd. Na to temo bi lahko vodili kontinuirano evidentiranje orodij in tehnoloških pristopov ter uvajanje novosti na obeh področjih: podpora za pripravo in ravnanje s podatkih v posameznih fazah življenjskega kroga podatkov in za razširjanje podatkov⁸³.

⁸³ Primerjaj delavnico v okviru projekta SERSCIDA v delovnem sklopu WP4: *Building capacities for establishment of social science digital data archives*: http://www.adp.fdv.uni-lj.si/serscida_wp4_lj2013_workshop/, ki smo jo organizirali v Ljubljani.

PRIKAZ REZULTATOV V FAZI ISKANJE, DOSTOP (OBJAVA), DRUGA RABA

DOSTOP DO PODATKOV

Intervjuvance smo prosili, da nam opišejo, kateri podatki so dostopni. Zanimalo nas je, komu, pod kakšnimi pogoji so dostopni ter kakšnih oblik, formatov in v kakšnih verzijah. Razumljivo je, da odvisno od tega, kako več je uporabnik, lahko prilagodimo ponudbo podatkovnih gradiv, pripravimo orodja in vmesnike, ki olajšajo iskanje in presojo relevantnosti podatkov ter njihovo uporabo. Katera orodja, metapodatki, standardi, postopki so uporabljeni za iskanje in dostop? Ali so podatki evidentirani in kako podrobno so opisani v katalogu? Ter po drugi strani, kakšna je praksa glede omejevanja dostopa?

Zanimalo nas je, ali so v predhodnih fazah priprave podatkov izpolnjeni pogoji za zagotavljanje dostopa do podatkov. Ali imajo podatki vrednost za druge uporabnike: preverjanje objav, raziskovalni potencial za nova odkritja in objave, kombiniranje več podatkovnih virov med sabo? Ali se spremlja in beleži druga raba? Kakšen je postopek in kakšna so merila? Ali ima uporabnik možnost oceniti vrednost in uporabnost podatkov? Ali obstaja interes za podatke drugih? Če da, za kakšen namen? Ali je zagotovljena sledljivost izvora podatkov in postopkov v predhodnih fazah, s čimer bi bila zagotovljena uporabnost podatkov za druge, dostop do metapodatkov, dokumentacije?

Vprašali smo jih tudi, kako vidijo odgovornosti v zvezi z reševanjem težav na tem področju. Kje bi bilo primerno mesto za dostop do podatkov? Kdo bi lahko opravljal storitve zagotavljanja dostopa (knjižnice)? Katere funkcije bi moral izpolnjevati? Zanimalo nas je, kakšen je pomen politik, vodil, standardov pri opredeljevanju odgovornosti. Ali imamo merila kakovostne podatkovne objave? Ali poznajo in podpirajo strategije in politike za zagotavljanje dostopa?

Kaj so ovire in sankcije za premagovanje: nezmožnost lociranja podatkov s strani uporabnikov, zaustavljeno financiranje projekta s strani naročnika? In kaj so spodbude in nagrade: kaj intervjuvanec pričakuje kot avtor? Citiranje, zabeleženo objavo? Kaj so pravni problemi dostopa: avtorske pravice, etična občutljivost podatkov? Kakšen je pogled na obveznost dostopa do podatkov s strani naročnika: javno financirani projekti, zahteve EU, revije ipd. Posebej nas je zanimala kultura izmenjave ter pogled glede omejevanja in pogojev dostopa. Ali ima interesent za podatke opredeljene pravice glede dostopa? Kakšen je odnos do načel OECD: kdo in kako zagotavlja izpolnjevanje na nacionalni ravni? Ali vidi v tej vlogi sebe in druge? Je to nekaj, kar velja za samoumevno ali neizvršljivo?

Dostop do podatkov je na različnih področjih raziskovanja neenako razvit, kar se kaže v različni ozaveščenosti glede pomena te faze življenjskega kroga, pa tudi v različnih rešitvah. Tako imamo na eni strani visoko profesionalizirane in standardizirane rešitve, povezane z mednarodnim sistemom. Na drugi strani so iskanje, dostop in druga raba podatkov prepuščeni iznajdljivosti posameznega raziskovalca ali skupine, rešitve pa so pogojene tudi z močjo in prednostmi, ki izvirajo iz njegovega položaja. Opaziti pa je mogoče tudi zametke dostopa, ki so jih razvili znotraj posamičnih ustanov in temeljijo na delu in znanju posameznikov, ki so se zgledovali po tujini ali iznašli svoje rešitve. Na primer v nadaljevanju govorimo o izkušnji Glasbenonarodopisnega

inštituta ZRC SAZU, ki znanje črpa iz mednarodnih izkušenj, vendar je pri svojem delu samostojen ter inventiven pri razvoju postopkov in storitev.

Na področju naravoslovja, kjer je mednarodno sodelovanje pogosta praksa, so uveljavljena rešitev osrednji mednarodni podatkovni arhivi, specializirani za določeno področje, ali mednarodne podatkovne banke, kot je npr. Worldwide Protein Databank (Gregor Anderluh, znanstvenik s področja biologije, 1/7.5).⁸⁴ Gre za repozitorije s podrobno določenimi protokoli za delo in rabo, iz tega pa izhajajo utečeni postopki in uveljavljeni standardi hrambe in dostopa (Gregor Anderluh, znanstvenik s področja biologije, 2, 3/7.5, znanstvenik s področja biologije, 4/7.5). Mednarodne banke so najboljša rešitev za določene vrste podatkov, vendar je taka rešitev znana le ozkemu krogu poznavalcev oziroma specialistov, tako kadar gre za ustvarjanje in s tem povezano možnost deponiranja podatkov kot kadar gre za možnost pridobivanja in njihovo ponovno uporabo.⁸⁵ Tudi na področju naravoslovja so znane in bolj pogoste prakse posameznikov in skupin, da s podatki, ki so jih sami ustvarili, ravna po lastni presoji in se vedejo kot neke vrste lastniki podatkov. Torej se sami odločajo, s kom in kdaj jih bodo delili ter v kakšni obliki in kolikšnem obsegu.

Na področju slovenistike že obstajajo prakse odprtega dostopa. Nastale so na pobudo posameznih ustvarjalcev zbirk, ki so v delo vložili velik osebni angažma (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 5/7.5, Miran Hladnik, znanstvenik s področja slovenistične literarne vede, intervju NO1, citat št. 6). Gre za zanesenjake, ki delajo tudi brez finančne podpore. Iščejo lastne rešitve, tudi po sistemu t. i. robinhudovstva, kadar ni jasne politike ali obstoječih rešitev, se odločajo v smeri večjega dostopa (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 7/7.5). Težijo k večji promociji svojih podatkov, da ti niso rezervirani za ekspertno skupino, ampak so v širši rabi, tako med študenti kot tudi v laični javnosti (znanstvenik s področja jezikoslovja, 8/7.5). Določena besedila, gre večinoma za žanrska literarna besedila starejšega datuma, ki se množično ustvarjajo v digitalni obliki (skeniranje, prepisovanje), so namenjena tudi širšemu javnemu dostopu – pri tem pa je kompromis med količino podatkov in kakovostjo (odsotnost napak) bolj na strani prvega, lahko so tudi slabše narejena. Druga besedila so namenjena pretežno za znanstveno rabo in predstavljajo reprezentativne »kritične« izdaje, so opremljena in pregledana po t. i. visokih znanstvenih merilih s težnjo po kar se da doslednem odpravljanju napak, tam je stopnja natančnosti neprimerno večja, a je veliko manj besedil (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 9, 10/7.5).

Na področju humanistike je mogoče opaziti zametke posamičnih podatkovnih središč, na primer: Arkas – arheologija, Zgodovina Slovenije – Sistory – zgodovina, DEDI – naravna in kulturna dediščina, Arzenal – virtualna zakladnica nacionalne dediščine ... Ta področja se soočajo s težavami vzpostavljanja ali vzdrževanja sistema storitev generiranja in izkoriščanja gradiv v digitalni obliki. Identificirali smo velik interes tako vodstev kot posameznih središč, ki

⁸⁴ Hierarhija naraščanja vrednosti in stalnosti podatkovnih arhivov je videti tako, da so čisto na vrhu glavni mednarodni viri (npr. World Wide Protein Bank), drugo raven, ki pomeni tudi nižjo stopnjo, predstavljajo nacionalna podatkovna središča (npr. UK Data Archive), še nižje, na tretji ravni se nahajajo inštitutski repozitoriji (npr. ePrints Soton), najnižje pa so umeščene individualne zbirke. Vrednost zbirk narašča od dna proti vrhu piramide, od individualnega pomena k skupnostnemu. Od spodaj navzgor naraščajo tudi odgovornost in zahteve, povezane z dostopom do podatkov. V isti smeri narašča tudi obseg infrastrukture, s tem pa tudi pomen standardov, stalnosti in vzdržljivosti sistema. (The Royal Society Science Policy Centre, 2012)

⁸⁵ Na tem področju sta zgled dobre prakse spletni strani Databib, <http://databib.org/>, in Re3data.org, <http://www.re3data.org/>, ki sta namenjeni iskanju podatkovnih središč. Gre za storitev identificiranja in lociranja repozitorijev raziskovalnih podatkov.

opravljajo storitve v zametkih, da se delo nadaljuje in da se nadgradi v smeri nacionalnih in mednarodnih povezav v okviru projektov DARIAH, SI-DIH ipd., z ambicijami intenzivneje vstopiti tudi na področje nove paradigme »digitalne humanistike«. ⁸⁶ Da gre za živahno in ambiciozno zastavljene načrte in ideje, kaže tudi predlagani projekt za Center odličnosti – Humanistično središče za novo družbo iz leta 2009, ki je bil interdisciplinarno zasnovan, toda žal ni bil odobren. Na področju socialnega dela je faza iskanja, dostopa in ponovne uporabe podatkov manj razvita (znanstvenik s področja socialnega dela, 11/7.5). Omejena je s praktičnimi razsežnostmi kvalitativnih raziskav in podatkov, ki jih je težko standardizirati. Poleg tega veliko metapodatkov sploh ne beležijo, so izgubljeni ali založeni po predalih. Omejujoče je tudi prepričanje raziskovalcev, da podatkov ni mogoče uporabljati, če nisi sodeloval pri ustvarjanju (znanstvenik s področja socialnega dela, 12/7.5).

Ker ni enotnega sistema, po katerem bi bili določeni postopki, pravice in odgovornosti v fazi iskanja, dostopa in druge rabe, so na določenih področjih podatki slabše dostopni in ostajajo neizkoriščen potencial, ki je prepuščen znanju, zmogljivostim in sposobnostim posameznih ustvarjalcev podatkov. Ker dosedanji sistem podatkovne menjave temelji na medsebojnem zaupanju, poznanstvih in osebni presoji raziskovalcev, so pogoji menjave stvar individualnih dogovorov, posledično pa so podatki pridržani za ožji segment uporabnikov. To je značilna ugotovitev, ki se pojavlja tudi v drugih podobnih študijah: ⁸⁷ odsotnost institucionalizacije storitev izmenjave podatkov z opredeljenimi protokoli, politikami, opredeljenimi pravicami in dolžnostmi pusti prostor za podatkovne izmenjave na ozki interesni podlagi neposrednih koristi vpletenih posameznikov.

ZADRŽKI IN OVIRE POVEZANE Z DRUGIMI FAZAMI ŽIVLJENJSKEGA KROGA PODATKOV

Ustvarjalci in dajalci podatkov se v fazi življenjskega kroga, ki obsega iskanje, dostop in ponovno uporabo raziskovalnih podatkov, srečujejo z različnimi zadržki in ovirami, ki jim preprečujejo, da bi podatke dali v ponovno rabo, ali jim onemogočajo, da bi prišli do podatkov, ki so jih ustvarili drugi. V tem podpoglavju smo podrobneje opredelili zaznane težave in predstavili, kako jih vidijo uporabniki in ustvarjalci. Ustvarjalci in uporabniki z različnih področij so kot zadržke in ovire navajali naslednje: konkurenčnost, tekmovalnost, slab zgled državnih ustanov, pomen velikega osebnega angažmaja ustvarjalcev, vprašanje avtorstva, vprašanje kakovosti.

Tekmovalnost in konkurenčnost

Poleg tehničnih, finančnih, pravnih in strokovnih zadržkov prosto rabo podatkov oziroma odprti dostop omejuje tudi kompetitivna kultura, ki je posledica poudarjeno točkovnega sistema pridobivanja sredstev za raziskovanje. Številni ustvarjalci in uporabniki druge vidijo kot tekmece, zato sodelovanje med njimi ni običajna praksa. Med razlogi, zakaj podatkov, ki so jih sami ustvarili, ne želijo dati, izpostavljajo vrednost podatkov oziroma neizkoriščen potencial. Za nekatere niti denarna nagrada ne bi bila dovolj, da bi podatke delili, saj menijo, da so pomembnejše objave, ki jih lahko ustvarjalec generira na podlagi monopola nad podatki

⁸⁶ Glej: http://www.sidih.si/digitalna_humanistika, 7. 8. 2012.

⁸⁷ Glej: *Patterns of information use and exchange: case studies of researchers in the life sciences*, <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/patterns-information-use-and-exchange-case-study>, 3. 1. 2014.

(znanstvenik s področja biologije, 13/7.5). Do deljenja podatkov so raziskovalci veliko bolj odprti v tisti fazi življenjskega kroga, ko so prepričani, da so obdelali in objavili že vse, kar so imeli v načrtu. Iz tega sledi zahteva po reformi obstoječih sistemov točkovanj, ovrednotiti bo potrebno tudi objave v netradicionalnih oblikah, zlasti še odprto dostopne objave in raziskovalne podatke, ki naj bodo obravnavani enakovredno kot ostale znanstvene objave.

Podatki so osnova za objavo in kot taki nosijo potencial za prihodnje delo (nova sredstva, objave), so konkurenčna prednost. V tem smislu imajo za ustvarjalce pomembno vrednost (Polona Domadenik, znanstvenica s področja ekonomije, 14/7.5). Za opisovanje zadržkov ustvarjalci pogosto uporabljajo pojme, kot so trg, konkurenca, tekmeč, tekma, kar je posledica prenosa tržnih modelov delovanja v znanstveno in strokovno sfero (Boris Majcen, znanstvenik s področja ekonomije, 15/7.5, raziskovalec s področja naravoslovnih ved, 16/7.5, vodja knjižnice, 17/7.5).

Tekmovalnost se kaže v izraziti individualizaciji pri prijavi in izvajanju projektov. Intervjuvanci se zavedajo, da tržni modeli (pritisk denarja in časa) negativno vplivajo na znanstveno dejavnost, in pričakujejo, da bodo prihodnje politike prinesle rešitve, ki bodo izpostavile vrednote, lastne znanosti, in motivirale raziskovalce k večjemu sodelovanju. Podatki so eden od elementov, ki vabijo k medpodročnemu izkoriščanju, pri tem pa je pomembno pravično ovrednotenje prispevkov in naporov pri ustvarjanju in pripravi podatkov.

Pomen velikega osebnega angažmaja raziskovalcev

Intervjuvanci so odnos do deljenja podatkov pogosto opisovali emotivno in s subjektivnega vidika. Pogosto so razlagali, da ob pozivu k deljenju podatkov občutijo ljubosumje (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 18/7.5). Ljubosumje so opisovali kot logičen odziv in posledico tega, da ustvarjanje podatkov ni ustrezno ovrednoteno. Ker so v proces vložili veliko osebnega truda, se ustvarjalci vedejo, kot bi bili lastniki podatkov. Če za podatke niso bili plačani, ne vidijo razloga za deljenje, saj so vanje vložili lastna sredstva (znanje, trud in čas). Kadar so vložili »gromozanski osebni angažma«, so prepričani, da jim tega nihče ne more poplačati (Polona Domadenik, znanstvenica s področja ekonomije, 19/7.5).

Intervjuvanec, ki je ustvarjalec in uporabnik podatkov, je opozoril na prisotnost občutka lastništva nad zbirkami, iz katerega izhaja ljubosumje in posledično omejevanje odprtega dostopa (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 20/7.5). Pomembno vlogo igra preračunljivost v smislu lastnega vložka in povračila zanj (Polona Domadenik, znanstvenica s področja ekonomije, 21/7.5). Če so za ustvarjanje podatkov prejeli plačilo, se bolj zavedajo odgovornosti in dolžnosti deljenja. Zavedajo se, da so podatki končni produkt dela, za katerega so bili plačani (znanstvenik s področja favnistike jamskih nevretenčarjev, 22/7.5).

Okvir 8: P-R 8

PROBLEM:

Med intervjuvanci prevladuje mnenje, da so podatki sredstvo za cilj, to je morebitna objava. Podatki so dragoceno gradivo, ker nosijo potencial za publikacijo. So tudi valuta za menjavo in sodelovanje pri pisanju člankov, monografij.

Zaradi takega pojmovanja podatkov je razširjena težnja po njihovi monopolizaciji. Prevladuje ljubosumje, občutek, da je podatek lastnina in da ima raziskovalec, ki je zbiral podatke, monopol nad možnimi odkritji, ki bi sledili iz podatkov.

Kultura odprte izmenjave podatkov še ni vzpostavljena. Ni sistema, ki bi natančno določal vloge, odnose, odgovornosti, dolžnosti in pravice ustvarjalcev in uporabnikov. Zato je akterjem prepuščeno, da se v raziskovalnem postopku ravna po lastnih občutenjih in osebnih presojah.

REŠITEV:

Potrebno je poskrbeti za bolj jasno določene vloge in odgovornosti nosilcev. Za večjo pretočnost podatkov in večje medsebojno sodelovanje akterjev v življenjskem krogu podatkov bo potrebno upoštevati problem lastništva podatkov. Prepričanja, da so ustvarjalci in skrbniki podatkov zaradi vloženega dela edini, ki morejo odločati o ravnanju z njimi, bo mogoče premostiti z ustreznim vrednotenjem podatkov.

Najbolj očitni ukrep, ki se ponuja v tem segmentu, je reforma sistema vrednotenja znanstvenih prispevkov in financiranja projektov, ki bo izpostavila zahtevo po dostopu do raziskovalnih podatkov, bodisi povezanih s publikacijo ali pa kot samostojnega produkta, ter s tem ustrezne zahteve po navajanju podatkovnih virov v primerih druge rabe. To naj gre v smeri, da se prizna avtorske pravice in se uredi ekvivalent nagrad za objavo podatkov po načelih, ki veljajo za objavo člankov.

Državne ustanove so slab zgled

Več nezadovoljstva v zvezi s fazo iskanja, dostopa in uporabe podatkov so izkazali intervjuvanci s področja ekonomskih ved. Opozorili so na dve slabosti, ki otežujeta delovni proces. Ena je komercializacija podatkov v državnih ustanovah, druga so dolgotrajne procedure, ki ovirajo dostop. Tak sistem so navajali kot negativen zgled in kot argument za restriktivno vedenje, kadar so sami postavljeni v vlogo dajalca (Polona Domadenik, znanstvenica s področja ekonomije, 23/7.5, Boris Majcen, znanstvenik s področja ekonomije, 24/7.5).

K dodatnemu premisleku glede prihodnjih politik vodijo tudi opažanja o dosedanjih rešitvah s področja hrambe in dostopa do podatkov. Ustvarjalci, skrbniki in uporabniki so pri svojem delu pogosto omejeni z obstoječimi modeli v širšem okolju, na katere kot uporabniki nimajo neposrednega vpliva. Intervjuvanci so kot primer slabe prakse izpostavili tako državne (npr. SURS) kot tudi mednarodne ustanove (npr. OECD, Svetovna banka), ki zaračunavajo dostop do podatkov, čeprav so bili ti podatki že plačani iz javnih sredstev (Polona Domadenik, znanstvenica s področja ekonomije, 25, 26, 29/7.5, Boris Majcen, znanstvenik s področja ekonomije, 27, 30/7.5, znanstvenik s področja geografije, 28/7.5).

Takšne rešitve so slab zgled in krepijo zadržke raziskovalcev, kadar se sami odločajo o prostem dostopu do lastnih gradiv. Tisti, ki morajo plačati dostop do podatkov oziroma morajo podatke kupiti, so se negativno odzvali na pobudo, da bi z drugimi delili svoje podatke, in so v vlogi dajalca podatkov zavzeli restriktivno stališče. To še ne pomeni, da podatkov ne bodo dali, vendar imajo negativno stališče do krovne ureditve, ki bi odprtost terjala le od raziskovalcev. Izpostavljen je bil očitek, da so državne ustanove slab zgled in da s svojimi praksami ne delujejo motivacijsko za raziskovalce (Polona Domadenik, znanstvenica s področja ekonomije, 31/7.5).

Po drugi strani pa so navajali tudi posamične primere dobrih praks iz tujine ter razlike med ameriškim in evropskim sistemom (Polona Domadenik, znanstvenica s področja ekonomije, 32/7.5). V ZDA imajo davkoplačevalci prost dostop do podatkov, ki so bili plačani z javnimi sredstvi, saj že kot davkoplačevalci plačujejo infrastrukturno mrežo in hrambo. Intervjuvanci so opozorili tudi na »luknjo« v sistemu: kjer bi morali pri slovenskih inštitucijah za dostop do podatkov opraviti določene postopke, lahko preko ameriških strani do teh istih podatkov prosto dostopajo. V tem smislu je slovenski sistem veliko bolj restriktiven in za uporabnike omejujoč (Nedjeljka Žagar, znanstvenica s področja meteorologije, 33/7.5).

Okvir 9: P-R 9

PROBLEM:

Intervjuvanci so na podlagi slabih domačih praks in na podlagi dobrih primerov iz tujine izrazili eksplicitna pričakovanja, da se uredijo splošna načela glede dostopa do podatkov, ki jih producira javni sektor.

REŠITEV:

Na več mestih smo izpostavili, da bi lahko del storitev in nalog evidentiranja, iskanja, in dostopa do podatkov javnega sektorja prevzela področna akademska podatkovna središča, kot je na primer UK DA, ki vodi storitve »Varnega dostopa do zaupnih mikropodatkov«⁸⁸, ali pa »podatkovno pribežališče« v ZDA, ki ga vzdržuje NORC.⁸⁹

Na splošno lahko k dvigu pozitivne klime deljenja podatkov prispeva aktivnejše vključevanje države v iniciative »odprtih državnih/javnih podatkov« (primerjaj *Open government data*).

Vprašanje avtorstva

Precej zadržkov in različnih stališč se je pojavilo v zvezi z vprašanjem avtorstva. En vidik zadeva avtorstvo v smislu priznanja zasluge za ustvarjalni napor, drugi kot oviro za prosto deljenje in dostop do podatkov in del, kadar avtorske pravice to omejujejo.

Poleg dilem v zvezi z avtorskimi pravicami se pojavlja tudi vprašanje, kako sploh opredeliti avtorstvo, kajti iz tega izvirajo pravice in obveznosti v zvezi s hrambo, dostopom in drugo rabo. To vprašanje pride do izraza še zlasti pri uporabnikih, ki dodajo podatkom novo vrednost (npr. izdelava zemljevidov). Tak primer so tudi slovarji in druge zbirke. Kdo je po prečiščenju, transformaciji lastnik podatkov in kaj lahko da v prosto rabo (znanstvenik s področja geografije, 34/7.5)? Ker se to področje pogosto navaja kot vir negotovosti ali pa tudi kot izgovor za to, da se omejuje dostop do določenih podatkov, bi bilo potrebno ustvariti pregledna navodila glede problematike avtorskih pravic in modelov licenc. Obstaja kar nekaj mednarodnih svetovalnih strani in projektov, ki pripravljajo orodja in podporo za zagotavljanje ustreznih citatov in referenc na vire, skupaj z njihovim ohranjanjem avtentičnosti glede na izvor (stalni identifikatorji in modeli citiranja, ki jih vsebujejo), ki tudi omogočajo sledljivost. Za podatke je priporočen poseben seznam variant licenc OA, ki ne omejujejo njihove rabe bolj, kot je nujno.⁹⁰

⁸⁸ Glej: <http://securedata.data-archive.ac.uk/>, 7. 8. 2012.

⁸⁹ Glej: <http://www.dataenclave.org/index.php/home/set/visitor>, 7. 8. 2012.

⁹⁰ Avtorske pravice in licence: <http://www.openaccess.si/avtorske-pravice-in-licence/> ter Viri in orodja za delo s podatki: <http://www.openaccess.si/viri-in-orodja-za-delo-s-podatki/>, 7. 8. 2012.

Vprašanje ravnanja z vsebinami, ki imajo različno avtorsko zaščito (filmi, umetniška dela, zgodovinski viri, glasba ...), na posameznih področjih predstavlja resno oviro za strokovno delo, recimo na področju humanistike (Nina Vodopivec, raziskovalka s področja socialne antropologije, 35, 36/7.5). Podatke predstavljajo dokumenti ali dela, za katera ni jasno, ali še zapadejo pod omejitve avtorskih pravic, njihovi »skrbniki« pa ne vedo, kaj sploh lahko delijo z drugimi in pod kakšnimi pogoji. Obstaja močan interes in pripravljenost za reševanje problemov, s tem da je potrebno preseči ozko razumevanje avtorstva, ki znanost zapira. Avtorske pravice so za to, da omogočajo javni dostop do »kulturne dediščine« in ne da ga omejujejo. Intervjuvanec opozarja na prisotnost t. i. avtorskega napuha, ki ga vidi kot oviro na poti k odprtim podatkom. Povezuje jo z logiko ekonomskega profita, za katero meni, da se bo morala spremeniti, saj ne sodi v znanstveno etiko (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 37, 38/7.5).

Vprašanje avtorstva je problematično tudi pri ustvarjanju mednarodnih zbirk, saj so različne države to vprašanje različno rešile. Eden od sogovornikov predlaga, da bi na mednarodni ravni ponudili zgolj metapodatke, dostop do podatkov pa bi vsaka država uredila skladno z lastnimi politikami (znanstvenik s področja zvočnega arhiviranja, 39/7.5).

Okvir 10: P-R 10

PROBLEM:

V zvezi z vprašanjem avtorstva je jasno izraženo eksplicitno povpraševanje po znanju, rešitvah, nasvetih, protokolih in dobrih praksah. Problem je najbolj jasno artikuliran na področjih, kjer se pojavljajo zametki storitev hrambe in izmenjav gradiv.

REŠITEV:

Eden od nujnih in logičnih korakov je koordinirano reševanje problema oziroma celega področja, ki obsega širše vidike avtorskih pravic v digitalnem okolju. Vprašanje avtorstva kot tako izkazuje resno potrebo po politikah, s katerimi bodo določena pravila in obveznosti ustvarjalcev podatkov, posledično pa bodo razrešene tudi dileme skrbnikov zbirk in zadržki v zvezi z drugo rabo.

Ena od teženj je v vse širšem pojavljanju t. i. podatkovnih revij, ki v sodelovanju z uveljavljenimi repozitoriji ali podatkovnimi arhivi, ki podatkom dodeljujejo stalne identifikatorje in zagotavljajo dolgoročno skrbništvo, pripravljajo recenzentsko pregledane članke o ustvarjenih podatkih.

Strah pred izgubo nadzora

Več intervjuvancev je izpostavilo strah, da bodo kot dajalci podatkov izgubili nadzor nad gradivom. Prepričani so tudi, da bo širši dostop povečal tveganje za zlorabe. Na določenih področjih je previdnost posledica strokovne podlage (npr. arheologija), na drugih je posledica odsotnosti praks in nepoznavanja preverjenih sistemov (znanstvenik s področja geografije, 40, 41/7.5). Tukaj bo potrebno seznanjanje s pozitivnimi izkušnjami. Na primer, da se široki javnosti ponudijo prečiščeni podatki, selekcionirani kot »smetana« podatkovne produkcije, za bolj zahtevne uporabnike z ustreznimi pravicami pa se zagotovi nadzorovan dostop.

Posamezna področja so regulirana tako, da so podatki dostopni le registriranim raziskovalcem, ki izrazijo svoj namen. Ena od možnosti je vloga oziroma prošnja, ki jo preveri in oceni za to

določena oseba ali komisija (raziskovalec s področja arheologije, 42/7.5). Na občutljivih področjih, kot je denimo medicina, imajo etično komisijo, ki odloči, ali bodo vlagatelju odobrili dostop in pod kakšnimi pogoji (znanstvenik s področja medicinskih vez in statistike, 43, 44/7.5). Dodatna rešitev je lahko tudi izjava oziroma pogodba, ki vlagatelja obveže, da z gradivom ravna pod natančno določenimi pogoji in je za kršitve lahko tudi kazensko odgovoren.

Okvir 11: P-R 11

PROBLEM:

Ustvarjalci in uporabniki so menjavam bolj naklonjeni in imajo manj zadržkov in strahov na tistih področjih, na katerih je druga raba podatkov ustaljena praksa. Drugod so skeptični in pri ravnanju zadržani oziroma odprtemu dostopu niso naklonjeni.

REŠITEV:

Pozitivne izkušnje lahko služijo kot primer dobre prakse na tistih področjih, na katerih imajo akterji manj izkušenj z drugo rabo podatkov, in lahko pomagajo tako na tehnološki ravni kot tudi na ravni ureditve razmerij med akterji (ustvarjalci, skrbniki, uporabniki).

Pri ozaveščanju naj večjo vlogo odigrajo osrednja nacionalna podatkovna središča. Vloga osrednjih nacionalnih središč za ustvarjanje ali za dostop do določene vrste podatkov je lahko tudi svetovalna, lahko posredujejo in organizirajo delavnice za veščine uporabe določenih vrst podatkov, delijo dobre primere analiz in posredujejo orodja⁹¹ za analizo. Predvsem pa imajo urejene protokole in preverjene obrazce za zagotavljanje nadzora nad izročanjem in dostopom do podatkov.

Okvir 12: P-R 12

PROBLEM:

Med splošnimi zaključki lahko izpostavimo presenetljivo podobnost zaznave problemov in odgovornosti za kakovost in zagotavljanje dostopnosti v različnih strokah. Morda se nakazuje, da je v humanistiki ter bioloških vedah in medicini (*life sciences*) manj kompetitivnosti kot v naravoslovju in družboslovju ter več poudarka na vrednosti priprave »referenčnih« zbirke podatkov, ki koristijo celotni skupnosti. Pri prvih so problemi specifični, izpostavljamo odsotnost razlikovanja med referenčnimi zbirkami in popisi gradiv ter surovimi podatki v digitalni obliki posameznih najdišč, vrst, eksperimentov ipd.

REŠITEV:

Več pozornosti bi bilo treba usmeriti k nalogam, s katerimi bi zagotovili dostop do surovih gradiv, ki so lahko podlaga za nadaljnje raziskave, oblikovati bi bilo treba ustrezne politike zajema gradiv ter regulacije pravic dostopa do gradiv in storitev v okviru dostopa.

Potrebno bo nadaljnje delo pri identificiranju potencialov in nastavkov storitvenih dejavnosti na teh področjih, npr. na področju naravoslovja bi lahko del nalog prevzele znanstvene knjižnice ali podobni skupni servisi velikih inštitutov in univerz, npr. z oblikovanjem splošnega odprto dostopnega repozitorija objav in podatkov; pri oblikovanju zahtev in politik naj imajo

⁹¹ Glej Gesis, Leibniz Institute for Social Sciences: <http://www.gesis.org/en/services/data-analysis/official-microdata/microcensus/microdata-tools/>, 7. 8. 2012.

večjo vlogo področne revije.

V družboslovju bi bilo potrebno preveriti sprejemljivost sodelovanja z že vzpostavljenimi storitvami Arhiva družboslovnih podatkov na sedaj manj zastopanih področjih, kot so ekonomija, psihologija ipd.

VREDNOTENJE PODATKOV

Zadnji sklop vprašanj je namenjen vrednotenju podatkov. Ker smo vprašanje o vrednotenju vključili že v vseh predhodnih fazah, smo ga ob koncu zastavili le še po potrebi. Zanimalo nas je, ali se upošteva vrednotenje podatkov in ali je zagotovljeno v predhodnih fazah priprave podatkov, ali se pri vrednotenju upošteva pomen za druge.

Na področju razumevanja odgovornosti (pravic in dolžnosti) nas je zanimalo, kje bi bilo primerno mesto za vrednotenje podatkov: ob nastanku, shranjevanju, dostopu, objavi? Kakšen je pomen politik, vodil, standardov: ali obstajajo merila vrednotenja, kakšne so ovire, sankcije, finančne zahteve; in po drugi strani spodbude, nagrade? Ali se kakovost podatkov spleča, ali je nagrajena?

Odsotnost skupnih politik in standardov iskanja, dostopa in druge rabe je vidna tudi na ravni povezovanja. Pomanjkljivi ali neobstoječi metapodatki namreč raziskovalce odvrtaajo od tega, da bi izkoristili potencial obstoječega gradiva. V nekontroliranih ali vsaj neenotnih pogojih, kakor so značilni za številna področja ustvarjanja podatkov, potencialni uporabniki menijo, da je druga raba praktično nemogoča (znanstvenik s področja biologije, 45, 46/7.5). Vendar so zadržki ali skepsa glede možnosti uporabe pogosto pomanjkljivo utemeljeni in ne opravijo preizkusa ovrednotenja v okviru potrjenih načrtov ravnanja z raziskovalnimi podatki. Priročnik o zadržkih v zvezi z deljenjem podatkov, ki so pogosti med ustvarjalci podatkov, in odgovorih nanje, so pripravili v UK Data Archive (glej tabela 2).

Med spodbudami za dostop do podatkov je tudi možnost izrabe v smislu povezovanja lastnih podatkov s podatki od drugod. Kakovost priprave podatkov se v tem pogledu vrednoti skozi uporabo ustreznih oznak vsebine in oblike podatkov, ki jo je zlasti v mednarodnem okolju včasih težko poenotiti (raziskovalec s področja arheologije, 47/7.5). Uporaba standardov in priporočil kljub vsemu lahko pomaga pri povezovanju (znanstvenik s področja biologije, 48/7.5).

Povezovanje podatkov, primerjave in raba skupnih orodij (naprav, modelov ...) prispevajo k izkoriščanju potenciala, ki ga nosijo podatki. V več primerih se je to povezovanje izkazalo koristno ne le med tistimi, ki uporabljajo uveljavljene mednarodne zbirke (znanstvenik s področja biologije, 49/7.5), pač pa tudi med domačimi raziskovalci. Tak primer so denimo humanistične vede, kjer se med seboj povezujejo arheologi, zgodovinarji, geografi. Izmenjava podatkov znižuje stroške, kar je še posebej pomembno, kadar so v igri drage metodologije (npr. fotografiranje iz zraka) (raziskovalec s področja geodezije, 50/7.5).

K večji povezljivosti podatkov lahko prispevajo specializirana podatkovna središča, ki imajo razvit sistem hrambe in izkušnje z drugo rabo. Lahko so tudi pomoč v smislu, da izdelajo analize ali poročila.

Tabela 2: Zakaj ne bi delili podatkov (zadržki in odgovori nanje)

	Zakaj ne bi delili podatkov?	Zakaj je vredno deliti podatke?
1.	Moji podatki nikogar ne zanimajo.	Raziskovalci želijo imeti dostop do podatkov iz različnih študij, metodologij in področij. Zapiski nekega vrtnarja lahko postanejo pomemben vir za raziskovanje meteoroloških sprememb.
2.	Želim objaviti, preden podatke pokažem drugim.	Dostop do podatkov je mogoče urediti tako, da ima ustvarjalec podatkov prednost (npr. časovni embargo).
3.	Nimam časa ali denarja, da bi pripravil podatke za deljenje.	Ravnanje s podatki mora postati sestavni del raziskovalnega postopka. Načrtovati je potrebno pred začetkom raziskovanja, s čimer se prihrani čas in denar.
4.	Če bom respondente vprašal za dovoljenje za deljenje podatkov, ne bodo želeli sodelovati v raziskavi.	Respondentom je potrebno razložiti, kaj pomeni arhiviranje in deljenje podatkov. Običajno jih to ne skrbi tako, kot si mislijo raziskovalci. Če dovoljenja še nimate, ga je mogoče pridobiti tudi naknadno.
5.	Ukvarjam se z zelo občutljivo raziskavo. Podatki preprosto ne morejo biti dostopni drugim.	Najprej je treba vprašati respondente, ali dovolijo deljenje. Druga pot je anonimizacija podatkov. Tretji način je embargo za določeno obdobje. Poleg tega je dostop do podatkov običajno omejen, npr. samo na raziskovalce.
6.	Imam kvantitativne podatke, povezovanje spremenljivk bo razkrilo identiteto udeležencev.	Kvantitativne podatke je mogoče anonimizirati s postopki anonimizacije, npr. zaokrožanje zgornjih kategorij distribucije, odstranitev spremenljivk, nadzorovan dostop do določenih spremenljivk (kot so poštna številke ali imena naselij).
7.	Imam avdio-vizualne podatke in jih ne morem anonimizirati, zato teh podatkov ne morem deliti.	Vizualne podatke je mogoče anonimizirati z zameglitvijo obrazov in s popačenjem glasu. To je lahko zamudno in drago, zato je smiselno prej preveriti, ali kljub vsemu dovolijo dostop, morda pod omejenimi pogoji.
8.	Obljubil sem, da bom po koncu raziskave podatke uničil.	Zakaj ste kaj takega obljubili? Vsekakor se je v prihodnje temu bolje izogniti. Kljub vsemu premislite, ali se obljuba oziroma zaveza nanaša na vse podatke.
9.	Podatke sem zbral z obljubo strogo zaupno.	V prihodnje se takim nepotrebnim obljubam izognite. Z anonimizacijo je mogoče zaščititi istovetnost, popolne zaupnosti nikoli ne morete zagotoviti.
10.	Moja podatkovna zbirka in nastali prepis (transkript) sta v tujem jeziku.	To ne bi smel biti problem. Podatkovni arhiv sprejme tudi transkripte v tujem jeziku, čeprav so prevodi bolj zaželeni.
11.	Nemogoče je anonimizirati prepise, saj se bo izgubilo preveč pomembnih informacij.	Stopite v stik s podatkovnim arhivom in skušali vam bodo pomagati. Boljša rešitev je lahko omejen dostop.

12.	Moja podatkovna zbirka vključuje podatke, ki sem jih kupil, zato ne morejo biti javni.	Potrebno je ugotoviti, kdo je lastnik avtorskih pravic in pridobiti ustrezna dovoljenja. Raziskovalec mora biti seznanjen z licencami in pogoji rabe podatkov.
13.	Drugi raziskovalci ne bodo razumeli mojih podatkov – ali jih bodo uporabili za napačne namene.	Za razumevanje podatkov je treba ustvariti dobro dokumentacijo in zagotoviti informacije o raziskovalnem kontekstu.
14.	Podatki so zaščiteni kot intelektualna lastnina.	Potrebno je pridobiti dovoljenje lastnika pravic. Najbolje v začetku raziskovalnega projekta, mogoče tudi za nazaj.

Vir: (Corti in dr. 2011, 11–12)

Katere podatke shranjevati

Vprašanje prostega dostopa do podatkov odpira vrsto dodatnih vprašanj, ki segajo tudi na druga področja življenjskega kroga podatkov. Eno takih je, katere podatke ohraniti in katere zavreči. Ker politike hrambe še niso razvite, tovrstni premisleki zadevajo predvsem ustvarjalce podatkov in zbirk, posledično pa vplivajo tudi na možnost dostopa in druge rabe.

Večinoma se ustvarjalci podatkov in zbirk strinjajo, da ni smiselno hraniti vsega, in se sprašujejo o tem, kje postaviti mejo. Ponekod velja prepričanje, da je pomembno ohraniti vse. Na posameznih področjih so raziskovalci prepričani, da je vsak podatek vreden hrambe, saj je bil vanj vložen nek trud (znanstvenik s področja socialnega dela, 51, 52/7.5, Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 53/7.5). V družboslovju in humanistiki, kjer se pojavi nenehno spreminjajo, vlada prepričanje, da je dolgoročna hramba toliko bolj pomembna. A je problem, kako podatke ustrezno opremiti, pripraviti za drugo rabo in zagotoviti dostop. Gre tudi za problem t. i. »*big data*«, s katerim se soočamo v zadnjem obdobju vsesplošno razširjenega generiranja digitalnih vsebin, kjer je težko postaviti mejo, kaj še hraniti.⁹²

Intervjuvanci so večkrat izpostavili predlog, da bi dostop do gradiv uredili glede na to, kdo je financer. Strinjajo se, da o dostopu do gradiv, ki so financirani z zasebnimi sredstvi, odloča plačnik, ki je tudi lastnik teh podatkov (Gregor Anderluh, znanstvenik s področja biologije, 54/7.5). Seveda je tudi na teh področjih ob ustrezni ureditvi mogoče doseči dogovor o možnostih dostopa za raziskovalne namene, ki ga kot uravnotežen pristop omenjajo tudi Načela OECD.⁹³ Veliko več pomislov so imeli pri tem, kako opredeliti in določiti mejo dostopa do podatkov, plačanih iz javnih sredstev (znanstvenik s področja geografije, 55/7.5, Nedjeljka Žagar, znanstvenica s področja meteorologije, 56/7.5, znanstvenik s področja biologije, 57/7.5).

⁹² Glej npr. JCDL 2012: *Panel 1: Big Data Is Already Here, and It's Not Always What We Think*: <http://jcdl2012.info/panels>, 7. 8. 2012.

⁹³ Glej Načelo E, *OECD Principles and Guidelines for Access to Research Data from Public Funding* (2009). Načelo E, *Protection of intellectual property*, pravi: Zaradi naraščanja javno-zasebnega partnerstva pri financiranju raziskovalnih in drugih dejavnosti, povezanih s podatki, obstaja potreba po uravnoteženih javno-zasebnih dogovorih, ki bodo omogočili širok dostop do raziskovalnih podatkov, kjer je to mogoče. Samo dejstvo, da je v podatkovne zbirke vključen tudi zasebni sektor, ne bi smelo biti razlog za omejevanje dostopa do podatkov. Premisliti bi bilo potrebno ukrepe, ki spodbujajo nekomercialen dostop in rabo ter hkrati ščitijo komercialne interese, npr. zakasneli ali delni dostop do podatkov. Glej tudi: Knowledge Exchange: <http://www.knowledge-exchange.info/Default.aspx?ID=461>, 7. 8. 2012. Glej: *Improving knowledge transfer between research institutions and industry across Europe: embracing open innovation. Implementing the Lisbon agenda* (2007): http://ec.europa.eu/invest-in-research/pdf/download_en/knowledge_transfe_07.pdf, 7. 8. 2012.

Večinoma se intervjuvanci torej strinjajo, da ne morejo biti dostopni vsi podatki. Nekateri menijo, da potem ne bi bilo več pregleda, kaj je zanesljiv podatek in kaj ne. Iz tega stališča izvira jasna potreba po storitvah, ki bi pomagale pri presoji, kaj je vredno hraniti in česa ne. Vprašanje je, ali je mogoče zagotoviti enoten, skupen servis, saj so področja raziskovanja in ustvarjanja podatkov raznolika, s tem pa tudi ovire in dileme za hrambo.

Za določena gradiva so zadostna rešitev splošni repozitoriji (npr. Dryad), za druga se kaže potreba po razvoju specializiranih, področnih podatkovnih središč. Zlasti na področjih, kjer se že kažejo zametki hrambe in izmenjav gradiv, je smotrno povezovanje in vzpostavljanje rešitev, ki bodo omogočile integracijo (npr. poenotenje registracije in pravic dostopa za uporabnike), kar bo dodatno spodbudilo izmenjavo in povečalo možnosti izkoristka obstoječega potenciala.

Strinjajo se glede tega, da je smotrno hraniti kakovostne podatke (znanstvenik s področja geografije, 58/7.5). Kakovost pa ni edino merilo, po katerem bi raziskovalci presojali o hrambi. Nekateri podatki so za stroko vse preveč dragoceni in jih ne želijo deliti s širšo javnostjo, saj se, kot je bilo že večkrat omenjeno, bojijo zlorab (znanstvenik s področja geografije, 59/7.5). Zavzemajo se, da bi bilo treba določiti nacionalne prioritete, za katere bi bila zagotovljena sredstva za vzdrževanje, hrambo in dostop. Kar se drugih gradiv tiče, so nekateri mnenja, naj ostane v presoji posameznikov, kako bodo ravnali z njimi (znanstvenik s področja geografije, 60/7.5).

Eden od predlogov je bil, naj se hranijo tisti podatki, po katerih je veliko povpraševanje. Pojavilo se je tudi mnenje, ki takšnemu razvrščanju podatkov ostro oporeka, z utemeljitvijo, da privilegiranje gradiv vzpostavlja hierarhijo moči (Nina Vodopivec, raziskovalka s področja socialne antropologije, 61, 62, 63/7.5). Zato je pri postavljanju prioritet bolj kot trenutno povpraševanje treba upoštevati dolgoročne strategije, nacionalne posebnosti in podobno. V prihodnjih politikah dostopa do podatkov mora biti ovrednoten njihov pomen, v konsenzu z ustvarjalci z različnih področij pa je treba določiti merila pomembnosti na dolgi rok.

Prihodnje politike morajo vključevati smernice in rešitve glede vzpostavljanja splošnih in področnih podatkovnih storitev, ki sledijo sprejetim nacionalnim prioritetam, kakor so opredeljene v dokumentih razvoja znanosti.⁹⁴ Usposobiti bo potrebno kader, ki bo po načelih knjižničarjev zagotavljal hrambo, vzdrževanje in dostop (znanstvenik s področja biologije, 64/7.5).

Okvir 13: P-R 13

PROBLEM:

Raziskovalci, ki tako ali drugače rokujejo s podatki, nimajo enotnega mnenja, kaj bi bilo smiselno hraniti in kaj ne. Nimajo prave predstave o tem, kakšne naj bodo pravice in odgovornosti ustvarjalcev in uporabnikov v fazi iskanja, dostopa in hrambe raziskovalnih podatkov, plačanih iz javnih sredstev.

REŠITEV:

S politikami bo potrebno določiti tudi obveznosti in pravice ustvarjalcev podatkov in zbirk. Eden od predlogov je, da bi prihodnje politike vključevale možnost embarga za krajše časovno

⁹⁴ Glej: Načrt razvoja raziskovalnih infrastruktur 2011-2020 (NRRI) in Nacionalna raziskovalna in inovacijska strategija Slovenije 2011 - 2020 (RISS).

obdobje, znotraj katerega bi imel ustvarjalec pravico, da podatke izkorišča za potrebe svojega dela, po preteku tega obdobja pa jih mora dati na voljo javnosti. S tem bi ustvarjalcu zagotovili določeno obdobje, ko bi lahko objavljali in izkoriščali potencial zbranega gradiva.

Problem rabe računalniških tehnologij

Področja, ki niso a priori vezana na digitalne podatke, se soočajo z dodatnimi, in sicer tehničnimi oziroma tehnološkimi ovirami. Na primer na področju slovenistike se ustvarjalci podatkov soočajo s predsodki, ki zadevajo digitalizacijo besedil (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 65/7.5). Navajajo nezadostno infrastrukturno podporo pri rabi računalniške tehnologije (npr. dodaten prostor na strežnikih). Zadržki so povezani z različnim pogledom znotraj vede, ki raziskovalce deli na tradicionaliste, ki gradijo na natančnem branju, in na druge, ki literarno znanost gradijo na podatkovnih zbirkah. Po ocenah intervjuvanca slednji nimajo veliko podpore.

Čeprav tega primera ne gre posploševati, pa vendarle kaže na to, da je potrebno dodatno pozornost namenjati vedam, ki so v tehničnem smislu v »zaostanku« in jih je zato težko primerjati z drugimi disciplinami, pri katerih je raba računalniških tehnologij in digitalnih podatkov že desetletja stalnica. Eden izmed intervjuvancev je izpostavil primer bioinformatike, kjer je prisotnost oziroma udeležba informatikov na tem področju dvignila ozaveščenost in povečala možnosti avtomatizacije (znanstvenik s področja biologije, 66/7.5). Ena od sogovornic je izpostavila pomen prisotnosti mlajših sodelavcev, ki se šele uvajajo v delo in so pripravljeni na novosti, prav nasprotno starejši zaposleni velikokrat nasprotujejo novostim (znanstvenik s področja socialnega dela, 67/7.5).

Okvir 14: P-R 14

PROBLEM:

Neenaka stopnja razvitosti infrastrukture, rabe računalniških tehnologij in sodobnih pristopov pomeni, da se raziskovalci soočajo z različnimi konflikti in težavami nerazumevanja.

REŠITEV:

Načelo pri tem naj bi bilo, da se nove rešitve in pristopi uvajajo postopno in da inovativni pristopi dobijo priložnost na način spodbujanja in financiranja pilotnih projektov, iz katerih zrastejo prihodnje storitvene dejavnosti dolgoročnejšega značaja.

ODGOVORNOST V ZVEZI Z REŠEVANJEM PROBLEMOM

Ustvarjalci in uporabniki podatkov se zavedajo, da so del širše strokovne in znanstvene javnosti, znotraj katere delujejo. Hkrati pa se zavedajo tudi obveznosti in omejitev, ki jih zastavlja financer (znanstvenik s področja geografije, 1/8.5).

POTREBA PO KROVNI POLITIKI

Intervjuvanci so izpostavili potrebo po skupni politiki, s katero bi fazo iskanja, dostopa in ponovne rabe podatkov uredili od zgoraj. Zavedajo se, da bodo od spodaj navzgor sami težko dosegli spremembe, zato je zaželena krovna rešitev, ki bi bila dosežena v sporazumu s predstavniki različnih področij (Boris Majcen, znanstvenik s področja ekonomije, 2/8.5, znanstvenik s področja socialnega dela, 3/8.5, znanstvenik s področja favnistike jamskih nevretenčarjev, 4/8.5).

Ena od sogovornic, ki je bila na vodstvenem položaju, pravi, da je skušala vplivati na zaposlene in vpeljati nove prakse, vendar sama v kolektivu težko kaj spremeni. Izrazila je potrebo po zunanji podpori, ne le infrastrukturni, ampak tudi pri ozaveščanju. Kot je že bilo omenjeno na drugih mestih, so nekatera področja neozaveščena v smislu digitalizacije, pa tudi deljenja in ponovne rabe; še zlasti tista, na katerih ne rokujejo s standardiziranimi podatki.

Pri politikah je mogoče govoriti o spodbujevalnih ali kaznovalnih (Andrej Blejec, znanstvenik s področja biologije, 5/8.5). Intervjuvanci vidijo spodbude predvsem v smislu, da bi morale biti dodatno delo tudi dodatno ovrednoteno (znanstvenik s področja socialnega dela, 6/8.5). Kjer je potrebno, bi bilo nujno zagotoviti finančna sredstva, ki bi delo sploh omogočila. Poskrbeti pa bi bilo treba tudi za dosledno navajanje podatkov; ena od možnih rešitev je po sistemu citiranja, kot je uveljavljeno za (znanstvene) publikacije (znanstvenik s področja medicinskih vez in statistike, 7/8.5). Kaznovalne pa so politike v smislu, da je s pravili določeno, da mora ustvarjalec poskrbeti tudi za dostopnost podatkov, na primer da je zahteva vključena že v projektno dokumentacijo in da ga k temu zavezuje pogodba (znanstvenik s področja gradbene informatike, 8/8.5).

Že omenjeni Feijen (2011, 8) glede koristi in razlogov za shranjevanje in dostopa do podatkov ugotavlja, da za raziskovalce po navadi ni dovolj zgolj postaviti obvezo, pač pa se pogosto raje podredijo zahtevam, če jih razumejo in če so izraženi utemeljeni razlogi. Ti naj bodo vidni skozi koristi in prednosti bolj odprtega dostopa do podatkov, npr. vrednost podatkov zaradi težavnega zbiranja in s tem povezani stroški, njihova edinstvenost in možnost preverjanja objavljenih ugotovitev, zaradi česar bi bilo škoda podatke zavreči zgolj po enkratni uporabi. Raziskovalci v večini niso za administrativno določeno obveznost, pač pa podpirajo postavljanje načelnih spodbud. Tudi naša izkušnja z intervjujev je bila, da raziskovalci ustvarjalci podatkov večinoma pozitivno gledajo na možnost dostopa do podatkov, če ne z drugih vidikov, pa zaradi možnosti preverjanja raziskovalne integritete pri objavljenih rezultatih. V nasprotju s tu navedenimi ugotovitvami pa so bili naši intervjuvanci večinoma naklonjeni tudi jasnim obveznostim za

izročanje podatkov in dostopu pod pogojem, da so obveznosti enake za vse in da se zagotovi ustrezna finančna in moralna podpora, tj. priznanje avtorstva.

USTREZNO VREDNOTENJE PODATKOVNIH VIROV

Ustrezno vrednotenje raziskovalnega dela predstavlja pomemben del raziskovalnega postopka, katerega del so tudi raziskovalni podatki. Raziskovalci se bolj ali manj strinjajo, da denarne nagrade niso ključna motivacija za delo, in razmišljajo o sistemu, ki bi spodbujal kakovostno delo in ga tudi nagrajeval (Polona Domadenik, znanstvenica s področja ekonomije, 9/8.5). Realen ukrep, ki je v tujini že uveljavljen, je obvezno citiranje podatkovnih virov. Obvezno citiranje podatkovnih virov ustvarjalce spodbuja, da svoje gradivo delijo z drugimi po načelu, da bo ponovna raba koristila tudi njim (Jan Jona Javoršek, raziskovalec s področja eksperimentalne fizike osnovnih delcev, 10/8.5, Nedjeljka Žagar, znanstvenica s področja meteorologije, 11/8.5).

Na področjih, na katerih raziskovalcem primanjkuje objav, so do dajanja podatkov še bolj restriktivni. Jasno je bilo izraženo pričakovanje, da bodo podatke oddajali, če bodo ti ustrezno ovrednoteni s točkami za znanstveno uspešnost, kar je v njihovem interesu (znanstvenik s področja biologije, 12/8.5, Polona Domadenik, znanstvenica s področja ekonomije, 13/8.5). Primer dobre prakse s tega področja so revije, ki zahtevajo, da ustvarjalci oddajo tudi podatke, ti pa so ovrednoteni tako, da se štejejo kot avtorstvo, iz česar izhajajo točke, ki so kazalec znanstvene uspešnosti raziskovalca.

Z uvedbo obveznega citiranja podatkov oziroma podatkovnih virov bo vrednost priznana tudi celemu področju ustvarjanja podatkov in zbirk (Andrej Blejec, znanstvenik s področja biologije, 14/8.5). Predstavlja pozitivno spodbudo za delo, ki med raziskovalci velja za manjvredno in težaško in ki se ga številni sploh ne lotijo, če to ni neizbežno (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 15/8.5).

Intervjuvanec je slikovito popisal problem, ko podatkovne zbirke nimajo statusa ekvivalenta znanstvene dejavnosti:

Trenutno je stanje teh zbirk podobno statusu strokovnih publikacij, slovarju in kar je podobnega, tisto se pa v naših krogih relativno nizko vrednoti, ker preprosto nima statusa znanstvene dejavnosti, ampak strokovne dejavnosti in je takoj polovičko manj vredno. To seveda ni spodbudno. (Miran Hladnik, znanstvenik s področja slovenistične literarne vede)

Problem je v neozaveščenosti znanstvene skupnosti, katere predstavniki določajo merila za vrednotenje, in pri knjižničarjih, ki še premalo poznajo posebnosti kategorije »znanstvenih podatkov«, kot je opredeljena v merilih. Medtem ko bo v znanstveni skupnosti potrebno sprožiti razprave, ki bodo dvignile ozaveščenost glede problema in rešitev, pa bo pri knjižničarjih potrebno vpeljati ustrezno usposabljanje.

Vrednotenje podatkovnih virov po načelu publiciranja člankov ustvarjalce spodbudi, da že v fazi ustvarjanja upoštevajo možnosti ponovne rabe, s tem pa povečajo potencial gradiva in sami stimulirajo drugo rabo, kar dodatno poveča plačilo za vloženo delo. Na kratko, obveznost citiranja lahko prispeva k dvigu kakovosti gradiv (znanstvenik s področja socialnega dela, 16/8.5).

Zgled dobre prakse na tem področju je mednarodna organizacija [DATACITE](#),⁹⁵ ki obvezno citiranje spodbuja z naslednjimi utemeljitvami:

- obvezno citiranje poenostavlja pregled nad ponovno rabo in preverjanjem istovetnosti podatkov,
- mogoče je slediti vplivu podatkov,
- ustvari mehanizem za prepoznavanje in vrednotenje prispevka ustvarjalcev podatkov.

Večje priznavanje vrednosti podatkov v okviru vrednotenja znanstvenega ustvarjanja je eden od najpomembnejših ukrepov in kot tak mora biti vključen v prihodnje politike. Te morajo obsegati ustrezno strokovno vrednotenje vloženega dela in nastalega gradiva.

Okvir 15: P-R 15

PROBLEM:

V slovenskem raziskovalnem okolju je trenutno mogoče govoriti o delovanju neformalnega sistema, po katerem konkurenčne prednosti izvirajo iz monopoliziranja in restriktivnega ravnanja s podatkovnimi viri, iz česar sledi, da podatkovni viri ostajajo neizkoriščen potencial.

REŠITEV:

Potrebno bo vzpostaviti ustrezno vrednotenje, ki bo na sistemski ravni spodbudilo sodelovanje in nadomestilo obstoječi neformalni sistem.

FINANČNA SREDSTVA

Intervjuvanci se strinjajo, da bo odprti dostop do podatkov mogoče uresničiti samo, če bo to eksplicitna zahteva financerja (raziskovalec s področja naravoslovnih ved, 17/8.5). Zaželena je sistemska rešitev na ravni objavljenih politik, ki se potem izvajajo v obliki pogodbenih obveznosti, sprejetih pri prijavi v obliki načrtov ravnanja s podatki in pri njihovem izvajanju v teku projekta in ob zaključku. Intervjuvanci so eksplicitno izrazili, da je zahteva realna le, če bo plačnik zagotovil finančna sredstva (raziskovalec s področja naravoslovnih ved, 18/8.5). Potreba po dodatnih finančnih sredstvih, ki bi omogočila zbiranje, ureditev, hrambo, vzdrževanje in dostop, je ena od pomembnejših zahtev za uresničenje akcijskega načrta projekta odprtih podatkov. Namreč, raziskovalci se večinoma soočajo z razmerami, ko nimajo sredstev za hrambo, kaj šele, da bi lahko poskrbeli za dostop do teh podatkov ali zbirk. Zato prevladuje stališče, da bi sredstva za ureditev in hrambo podatkov morala biti zagotovljena že v projektu. Zahtevo utemeljujejo s tem, da gre za dodatno strokovno znanje, ki ga je treba usvojiti in ga prenesti v prakso, za vse akterje pa pomeni tudi dodatno delo, ki ga z obstoječimi kadri in kapacitetami ne zmorejo opraviti (Polona Domadenik, znanstvenica s področja ekonomije, 19/8.5, Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 20/8.5).

Financiranje je tudi nasploh problem. Če je raziskovalno delo odvisno od tega, ali bodo dobili projektna sredstva ali ne, ne morejo kontinuirano zbirati določenih podatkov. Če ni projekta, ni denarja in se zbiranje ustavi. Zlasti je to problem pri podatkih, katerih kakovost je pogojena s posodabljanjem (znanstvenik s področja geografije, 21, 22/8.5). Pri oblikovanju finančne sheme projektov je potrebno upoštevati tudi posebnosti posameznih področij, npr. nekateri potrebujejo

⁹⁵ DataCite, <http://www.datacite.org/>, 7. 8. 2012.

več tehnične opreme, drugi več kadrov (Boris Majcen, znanstvenik s področja ekonomije, 23, 24/8.5, vodja znanstvenoinformacijskega centra, 25/8.5).

Iz tega sledi potreba, da morajo prihodnje politike vključevati jasno finančno shemo, ki bo že pri prijavi projekta predvidela dodatna sredstva. Tako bodo prijavitelji že pri pripravi na raziskavo presojali vrednost podatkov za ponovno rabo in ovrednotili stroške priprave za hrambo in dostop, s čimer se bo povečal izkoristek potenciala, ki ga nosijo gradiva. Vključevanje dodatnih sredstev je mogoče skozi model obveznih načrtov ravnanja z raziskovalnimi podatki, ki ga spodbujajo sodobne raziskovalne politike. Neustrezno ureditev pri načrtovanju sredstev in poročanju o porabi, ki administrativno obremenjuje projekte danes, bo potrebno reformirati v smeri načrtovanja in porabe za projektne določene naloge, katerih sestavni del so tudi naloge skrbi za podatke (znanstvenik s področja geografije, 26/8.5).

Poleg prihodnjih projektov je potrebno upoštevati tudi stanje obstoječih gradiv. Sredstva bo potrebno zagotoviti tudi za že zbrana gradiva, ki jih je potrebno zavarovati pred zastarevanjem, uničenjem ... To pa je pogosto povezano z uvajanjem postopkov digitalne hrambe, ki zahtevajo denarna sredstva, čas in strokovno znanje.

DOBER SISTEM IZVEDBE HRAMBE IN DOSTOPA DO PODATKOV

Prihodnje politike morajo zagotavljati učinkovitejše sisteme, ki bodo omogočali hrambo in dostop do podatkov. Če je sistem dober, deluje spodbujevalno, saj olajša delo. V nasprotnem primeru pa kaznovalno, saj pomeni dodatno obremenitev za raziskovalce.

Pričakovanje novih rešitev je bilo še posebej močno izraženo na tistih raziskovalnih področjih, na katerih vključevanje v mednarodne repozitorije ni uveljavljena praksa. Tak primer so podatkovni viri, ki imajo nacionalno vrednost, zaradi nacionalnih posebnosti pa je vključevanje v mednarodne zbirke omejeno in oteženo (Miran Hladnik, znanstvenik s področja slovenistične literarne vede, 27, 28, 30/8.5, znanstvenik s področja socialnega dela, 29/8.5). Intervjuvanci so izrazili potrebo in pričakovanje po dodatnih službah oziroma storitvah, ki bodo imele pregled nad podatkovnimi viri in tudi rešitve za posebnosti posameznih področij (etika, anonimnost, pravice, zaščita).

Okvir 16: P-R 16

PROBLEM:

Številna področja se soočajo s težavami, ki izvirajo iz vzpostavljanja ali nadgradnje sistema podatkovnih storitev (družboslovje, humanistika in posamezna področja naravoslovja). Ustvarjalci in uporabniki se ukvarjajo s tehničnimi, pravnimi, administrativnimi in finančnimi težavami. Določena vprašanja pa so povsem vsebinske narave in so kot taka vezana na konkretno stroko oziroma področje.

REŠITEV:

Ker gre za isto fazo razvoja in podobne težave, je smotrno razmišljati o povezanem in koordiniranem iskanju rešitev. Pogled od spodaj navzgor kaže, da so nekatere rešitve lahko skupne za vse akterje (npr. administrativne, pravne, finančne), druge bodo morale biti bolj ali manj diferencirane, usklajene in prilagojene področnim problematikam. Pomembno je, da so

rešitve usklajene s hkratnim potekom spremenjenih norm in običajev znotraj posameznih disciplin in da upoštevajo posebne zahteve in veljavne pristope v disciplinah, tako da se zagotovi široka podpora njihovem uresničevanju. Zato je pomembno spremljati in posodabljati uveljavljene standarde na posameznih področjih in uvajati spremembe skladno z razvojem strok.

Na področjih, kjer ne poznajo sistemskih rešitev, vlada dodatno nezaupanje glede kakovosti podatkov (raziskovalec s področja naravoslovnih ved, 31/8.5, vodja znanstvenoinformacijskega centra, 32/8.5, znanstvenik s področja favnistike jamskih nevretenčarjev, 33/8.5). Iz tega izhaja potreba po vzpostavitvi sistema, ki bo presegal subjektivne ocene in merila kakovosti ustvarjenega ter vpeljal standardizirane rešitve, v katerih kakovost podatkov na pregleden način jamčijo licence ali drugi standardi kakovosti.⁹⁶ Razviti bi bilo treba ekspertna znanja in kader, ki presoja o kakovosti (Polona Domadenik, znanstvenica s področja ekonomije, 34, 36/8.5, Nina Vodopivec, raziskovalka s področja socialne antropologije, 35/8.5, raziskovalec s področja naravoslovnih ved, 37/8.5). Zgled so lahko mednarodne zbirke,⁹⁷ ki jih uporabljajo naravoslovci.⁹⁸

Ena od možnosti je skupni sistem, ki bi pokrival vse univerze in inštitute v državi ter skrbel za organizacijsko in vsebinsko raven (vodja knjižnice, 38, 39/8.5), tega pa bi ustrezno dopolnjevala področna podatkovna središča. Oblikovanje središč zahteva usposabljanje kadra, specializiranega za hrambo in dostop.

Raziskovalci so izrazili potrebo po razumni rešitvi, ki bo omogočala enostavno hrambo, spodbujala sodelovanje ter ustvarjalce servisirala v primerih dilem, vezanih na vprašanja hrambe in dostopa do podatkovnih virov.

Okvir 17: P-R 17

PROBLEM:

Že danes obstaja vrsta ovir za bolj kakovostno delo, zato so se raziskovalci v pogovoru največkrat strinjali, da od financerja pričakujejo razumno rešitev, ki ne bo povečala obsega administrativnih in drugih birokratskih del ter z dodatnimi obremenitvami zmanjševala kakovosti dela. Izražena je bila tudi potreba po krovni politiki, ustreznem vrednotenju dodatnega dela in učinkovitem sistemu hrambe, ki bi zagotavljal dostopnost podatkov. Raziskovalci imajo kvečjemu parcialen pregled nad problematiko, ki je podkrepjen z lastnimi izkušnjami. Velikokrat ne poznajo uveljavljenih in utečenih rešitev, ki jih poznajo akterji na drugih področjih ali v tujini.

REŠITEV:

Prihodnje politike bodo morale poskrbeti za sistematizacijo načrtnega ravnanja z raziskovalnimi podatki, to pomeni, da bo že v fazi priprave projekta potrebno vključiti in upoštevati pomen hrambe in dostopa do podatkov. Zato je smiselno, da zahteva po spremembi

⁹⁶ Primerjaj pravkar izdano publikacijo *Allining national approach to digital preservation* (ur. McGovern 2012).

⁹⁷ Npr. že prej omenjene zbirke: GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>, 12. 2. 2014), UniProt (<http://www.uniprot.org/>, 12. 2. 2014), Protein Data Bank (<http://www.pdb.org/>, 12. 2. 2014), Amadeus (<http://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/Amadeus.aspx>, 12. 2. 2014).

⁹⁸ Zanimiv prispevek na to temo sta na konferenci IASIST predstavila Downs in Chen (2012).

in predlogi pridejo od zgoraj navzdol. Izraženo je bilo pričakovanje, da zahteve postavi financer, vendar v dialogu z raziskovalci.

ZAKLJUČKI ANALIZE POLOŽAJA ODPRTIH RAZISKOVALNIH PODATKOV V SLOVENIJI

Predstavitev analize pogojev in možnosti za vzpostavitev nacionalnega sistema zagotavljanja odprtega dostopa do raziskovalnih podatkov smo začeli s prikazom aktualnosti te tematike v strokovni in poljudni literaturi. O aktualnosti in novosti, inovativnem potencialu možnosti odprtega izkoriščanja raziskovalnih podatkov govorijo tudi že omenjeni izrazi, kot so »četrt paradigma« (ur. Hey in dr. 2009), »digitalna humanistika«, »veliki podatki« ipd. Vendarle gre bolj za spremenjene predstave o nas samih in načinu, kako dojemamo vlogo podatkov v prostem dostopu v dobi interneta, kot pa za povsem drugačen način ustvarjanja znanstvenih odkritij. Se pravi, da ne začenjamo iz nič, ampak kot smo izpostavili v našem pristopu, iščemo in se opiramo na že obstoječe nastavke in dejavnosti, ki so v večini primerov tesno vpete v okolje ustvarjanja podatkov. Gre torej za to, da izboljšave in novosti v to okolje vtremo čim bolj nemoteče in na način, da bodo imeli vsi vpleteni od tega dejansko nekaj pozitivnega. Izboljšave, četudi niso revolucionarne, pa tako kot vsaka sprememba potrebujejo obdobje prilagajanja in usposabljanja, preden bodo sprejete in bodo zaživele.

Že v pristopu k projektu smo omenili, da naš namen ni zbrati seznam ukrepov, ki naj bi jih Ministrstvo uvedlo od zgoraj. Izhodišče, da že v prvi fazi projekta, katere rezultate tu predstavljamo, namenimo posebno pozornost ugotavljanju, kje in kako različni nosilci vlog v znanosti vidijo probleme in možne rešitve, izhaja iz strateške usmeritve, da je potrebno predvsem ustvariti možnosti in okoliščine, v katerih se bodo vsi vpleteni na svojih področjih, ki jih tudi najbolje poznajo, lotili ukvarjanja s problemi odprtega dostopa do podatkov. In sicer po korakih: od tega, da se problemov sploh začnejo zavedati, da najdejo primerne kapacitete in organizacijske možnosti, da se z njimi ukvarjajo, do tega, da prispevajo svoje pobude tudi za izgradnjo institucionalnega okolja z ustreznimi politikami in načeli, ki spodbujajo in tudi zavezujejo k ravnanju v prid odprtega dostopa. Tako so tudi rešitve, ki smo jih izluščili iz prve faze projekta, kljub temu da so oprte na študij literature ter na obsežne in poglobljene intervjuje, samo predlogi, dani v premislek in nadaljnjo dopolnitev vsem, ki jih zadevajo. Na nosilcih znotraj posameznih področij je, da se prepoznajo v okrepljenih vlogah v dejavnostih, povezanih s podatki, da se usposobijo in se s svojo dejavnostjo vzpostavijo kot pomemben člen znotraj posameznih disciplinarnih področij, katerih potrebe in možnosti so sposobni uvideti zaradi vsebinske bližine. Ministrstvo in drugi oblikovalci politik lahko dajejo svoje pobude, obenem pa je njihova vloga, da se odzivajo na izražene pobude iz znanstvenih skupnosti ter omogočijo organizacijske in institucionalne okvire za njihovo uresničevanje. Pri tem učenju seveda zelo pomagajo zgledi iz bolj razvitih okolij in meddisciplinarna zgledovanja, kadar gre za sorodne probleme, kar pa zopet terja prilagajanje in aktivno prevzemanje, se pravi nek dodaten angažma. Bistveno pri tem je, da se vzpostavi okolje, v katerem bodo prevladale koristi za znanstveno skupnost v smislu napredovanja spoznanj, kot tudi v smislu nagrad za vložene dodatne napore (ustrezne priprave podatkov) in nadomestil za žrtve (npr. izgubljanje monopola nad ustvarjenimi podatki) izboljšane dostopa do raziskovalnih podatkov.

Glavna ovira, ki smo jo identificirali že iz pregleda sorodnih študij drugje, je lahko kapaciteta in dobro delovanje usposobljenih podpornih dejavnosti storitev, povezanih z organizacijo dostopa do podatkov. Pri nas je še v zametkih, v razvitih okoljih pa je že dobro vzpostavljen sistem institucionalnih repozitorijev objav v obliki znanstvenih besedil, ki se vse bolj usmerjajo tudi na usposabljanje za podporo bolj kompleksnim oblikam znanstvenih digitalnih objektov, med katere sodijo tudi raziskovalni podatki. Glede na to, da segment institucionalnih digitalnih repozitorijev pri nas ni tako razvit, bo potrebno institucionalno podporno dejavnost za skrb za podatke razvijati in vzpostavljati drugje. Predvsem to lahko pomeni precejšnje spremembe v nalogah knjižničarjev in njihovo večjo vključenost v vse faze ustvarjanja podatkov ter poteka raziskovalnih projektov na inštitutih in univerzah. Pomembna je lahko tudi izgradnja centraliziranih storitev nad ravno raziskovalnih ustanov (npr. na področju usposobljenosti in orodij za podporo digitalne hrambe). Ključno pa je vzpostavljanje in podporno delovanje specializiranih področnih storitev nacionalnih podatkovnih središč ali arhivov (v povezavi s sorodnimi evropskimi in svetovnimi). Zgled za slednje so nekatere že delujoče povezave na različnih področjih, od družboslovja in humanistike do naravoslovja in biologije. Obenem je pomembno, da se delujoči znanstveniki usposobijo v vlogah podatkovnih znanstvenikov, za kar je potrebna dodatna spodbuda za zagotavljanje enakopravnosti pri napredovanju in ugledu.

Ustvarjanje podatkov

Spoznanja, ki smo jih oblikovali s pregledom intervjujev, se začnejo z ugotovitvami, da so podatki lahko pojmovani zelo različno glede na kontekst uporabe, da jih lahko klasificiramo na različne načine glede na tipe ter da se tudi v fazi ustvarjanja generirajo zelo raznoliki formati in zapisi podatkov. Posebej izpostavljamo pomembnost razlikovanja med referenčnimi zbirkami, ki običajno nastajajo s kolektivnimi prispevki in predstavljajo celovit pregled na določenem področju, ter raznolikimi raziskovalnimi podatki, ki nastajajo znotraj raziskovalnih projektov v teku raziskovanja. Prav na slednje se pogosto pozablja in je zanje slabo poskrbljeno, zato jim je potrebno posvetiti enako mero pozornosti in jih ovrednotiti z vidika nadaljnje uporabnosti.

Pomembne so ugotovitve, da so za znanost zanimivi tudi podatki, ki nastajajo drugod; zlasti tisti, ki nastajajo v javnih ustanovah, bi lahko s svojo večjo dostopnostjo in možnostmi za dodelave in nadaljnje razširjanje v obogatenih oblikah predstavljali zgled tudi za ožje znanstveno ustvarjanje. To sicer ni predmet tega projekta, je pa področje, ki je vsebinsko tesno povezano z njegovimi cilji, namreč vprašanje »odprtih javnih podatkov« (*Open government data*).

Enako je pomembno, da se pozornost posveti ne samo »big data«, ampak tudi prevladujočemu ustvarjanju malih, po obliki in vsebini lahko zelo kompleksnih vrst podatkov. Poskrbeti je potrebno za možnosti izmenjave dobrih praks med področji, obenem pa je jasno, da zelo uniformirane rešitve in robustni pristopi s področja velikih podatkov ne bodo vselej uporabni za kompleksno problematiko malih podatkov. (Okvir 1: P-R1)

V vseh fazah ravnanja s podatki je vprašanje skrbi za kakovost eno ključnih. Na tem področju je mogoče nadzor vzpostaviti s formaliziranimi protokoli, opredeljenimi v pisni obliki, ki pa jih, po izkušnjah intervjuvancev, v vsakdanjih raziskovalnih postopkih ne uporabljajo vedno povsem dosledno, četudi ponekod obstajajo in so raziskovalci z njimi seznanjeni. Tako da se to vprašanje navezuje tudi na širšo problematiko vrednotenja podatkov in načrtovanja, kakšne standarde kakovosti bo nekdo uporabil, tudi glede na to, ali bodo podatki kasneje na voljo v sistemu odprtega dostopa. Zagotovila glede kakovosti morajo biti v tem primeru še bolj pregledna, kot je običajna praksa. (Okvir 2: P-R 2)

Pomembno je tako tudi naslednje spoznanje (Okvir 3: P-R3). Četudi je v vzpostavljenih znanstvenih okoljih skrb za kakovost podatkov na visoki ravni, kar je povezano s ciljem zagotoviti kakovostne zaključke v analizi in objavi, je skrb za podatke in njihovo dokumentacijo (načrtovanje, varovanje, zagotavljanje ustreznega mesta hrambe in dostopa) področje, na katerem se bodo morali raziskovalci dodatno usposobiti; šele tako bo zagotovljena ustrezna kakovostna raven priprave podatkov za uporabo kot celote.

Vrednotenje in izbor podatkov je sam po sebi posebna naloga, ki ima po eni strani za podatke enak pomen kot vrednotenje publikacij skozi postopek recenziranja, se pravi, gre za ustrezno priznavanje znanstvenega prispevka, po drugi strani pa terja drugačne in prilagojene kriterije in postopke. Tudi tu je pomembno tako usposobiti nosilce vlog, kot so založniki, knjižničarji in drugi predstavniki znanstvene skupnosti, kot prilagoditi politike financerjev znanosti. Med drugim je pomembno, da recenzenti prijav na raziskovalne projekte poznajo načela priprave načrtov ravnanja s podatki (NRRP) in da to znajo ustrezno ovrednotiti. Dodatno delo in predvideni naporji pa morajo biti tudi ustrezno finančno in kadrovsko podprti s strani financerja (Okvir 4: P-R 4, Okvir 5: P-R 5).

Dodana vrednost podatkov

Zlasti problem spodbud in usposobljenosti za pripravo ustreznih metapodatkov, ki predstavljajo zgodovino nastanka in dokumentacijo značilnosti samih raziskovalnih podatkov, se je pokazal kot izrazito odvisen tudi od strokovne usposobljenosti in podpore ustreznih podatkovnih arhivov na posameznih področjih. Če ni motivacije za izročanje podatkov v podatkovno središče, potem ni niti znanja niti pripravljenosti, da se nekdo temu področju dodatno posveti (Okvir 6: P-R 6).

Področje digitalne hrambe je v svetu v zadnjem obdobju zelo napredovalo, tako v tehnološkem pogledu kot v pogledu temu prilagojenih postopkov, preverjenih s pomočjo standardov in pod nadzorom politik. Temu ustrezno se opazi zaostajanje pretežne večine »nespecialistov« za to področje, kar delujoči znanstveniki običajno so. Izraženo je zavedanje, da je ta problem prisoten, hkrati pa so prevladujoče »ad hoc« rešitve in nemoč, da se ta problem ustrezno uredi na način, da se zmanjša nevarnost izgubljanja in uničevanja podatkov. Tudi tu je vloga specialistov in centralizirane podpore nepogrešljiva, hkrati pa je potreben dodaten napor vseh ostalih (Okvir 7: P-R 7).

Odprti dostop do podatkov

Zadnja faza v življenjskem krogu podatkov je možnost nadaljnje analize dostopnih podatkov s strani drugih raziskovalcev, ki je podprta z ustreznimi storitvami iskanja, registracije in dovoljenjem dostopa. Da so podatki dostopni drugim, pa jih mora biti ustvarjalec pripravljen dati na razpolago na ustrezno mesto, ki ureja omenjene storitve pri dostopu. Kljub nekaterim zametkom te dejavnosti je to področje kritično podhranjeno zaradi prevladujoče kulture zapiranja in monopoliziranja podatkov. Seveda problem ni specifičen le za naše okolje, tako da so tudi rešitve, o katerih smo razpravljali z našimi intervjuvanci, že ponujene in se v drugih okoljih postopoma uveljavljajo. Zlasti gre za vse bolj uveljavljen ukrep, da se odprto dostopni podatki vrednotijo v okviru bibliografij, da se obravnavajo enako kot druge znanstvene objave ter da se zagotovi tudi obveznost navajanja med viri v znanstvenih člankih, v katerih so bili podatki uporabljeni. Obenem se vzpostavlja tehnološko in s pomočjo trajnih identifikatorjev podprta infrastrukturna storitev zagotavljanja poenotenega naslavljanja podatkovnih virov,

podobno kot z znanstvenimi publikacijami, bodisi z njimi povezano, ali v obliki podatkovnih revij, ali pa skozi robustno zasnovane področne podatkovne arhive.

Ob tem, ko se uvaja ustrezna tehnološka infrastruktura, pa je pomembno začeti s spremembami pri razumevanju vloge in odgovornosti vseh vpletenih, od ustvarjalcev, dejavnosti storitev do uporabnikov in financerjev. Naši intervjuvanci pozivajo financerja, naj določi jasna in za vse zavezujoča pravila ter omogoči njihovo izpolnjevanje. Na ta način zagotavljajo, da bo problem nizke kulture deljenja podatkov hitro rešen (Okvir 8: P-R 8).

Enako je tudi na že omenjenem področju dostopa do javnih podatkov potrebno zagotoviti večje vključevanje raziskovalnih potreb, skupaj tudi s potrebno posredovalno vlogo področnih podatkovnih arhivov, katerega zgled je lahko sporazum med ADP in SURS ter skupno delo pri pripravi podatkov za dostop preko »varne sobe« in »oddaljenega varnega dostopa« (Okvir 9: P-R 9).

Vprašanja avtorstva, licenc in priznavanja pravic ter druge omejitve, ki izhajajo iz etičnih ali pravnih obvez v povezavi s podatki, so področje, ki se običajno omenja med ovirami za odprti dostop. Tudi to je področje, ki je ozko strokovno, na katerem pa lahko rešitve, ko se enkrat oblikujejo in opredelijo, pomagajo praktikom, ki ustvarjajo ali posredujejo podatke drugih. Tako je povezava in forum za izmenjavo dobrih praks med ponudniki storitev dostopa do raziskovalnih podatkov logična rešitev, prav tako tudi v povezavi z vprašanji zaupnosti podatkov in drugimi pravnimi problemi (Okvir 10: P-R 10).

Svetovanje in usposabljanje tako ustvarjalcev kot uporabnikov sta pglavitni vlogi obstoječih, že delujočih ponudnikov storitev dostopa do raziskovalnih podatkov ali tistih, ki se šele vzpostavljajo v tej dejavnosti. Zelo pogosto so ustvarjalci podatkov izpostavili problem in skrb, da bodo podatki uporabljeni na neustrezen način, uporabniki pa, da je podatke težko razumeti in da so orodja za uporabo preveč zapletena. To je mogoče rešiti tudi s pomočjo večje izmenjave znanj in izkušenj v skupnosti, npr. na način srečevanja ustvarjalcev in uporabnikov določenih podatkov. Tudi ostali pomisleki in ovire, ki jih navajajo posamezni intervjuvanci in se pojavljajo v javnosti, se lahko rešijo skozi uveljavljen, delujoč sistem zagotavljanja izročanja podatkov in dostopa preko za to pooblaščenih arhivov podatkov, z uporabo vnaprej določenih postopkov in protokolov, npr. pri pridobivanju izrecnega pristanka za sodelovanje pri raziskavi, skupaj s postopki anonimizacije podatkov, kadar gre za podatke o ljudeh. Na drugih področjih so v ospredju bolj vprašanja licenc, pri katerih je prav tako mogoče predvideti modele, ki zagotavljajo morebitno izključno komercialno izkoriščanje, hkrati pa ne omejujejo znanstvene rabe. Poznavanje in promocija tovrstnih pristopov in modelov lahko pomembno pripomore tudi k širšemu sprejemanju idej odprtega dostopa do podatkov in njegovega uveljavljanja.

Ob koncu polstrukturiranega intervjuja smo intervjuvance povprašali, ali bi bili pripravljene sodelovati v delavnicah in na katerih področjih iz Načel OECD bi bili pripravljene sodelovati v razpravah (tehnologija, standardi; digitalna hramba; organizacija, odgovornost, politike; avtorske in druge pravne omejitve, zasebnost, patenti; kultura izmenjave, etika; interoperabilnost, učinkovitost; kakovost podatkov; ocenjevanje storitev in nadzor nad izvajanjem; obveznosti načrtov ravnanja s podatki in njihov nadzor ter vrednotenje izpolnjevanja v okviru projektov). Prikaz izvedenih delavnic je vključen v nadaljevanju poročila.

POSVETOVALNE DEJAVNOSTI PROJEKTA ODPRTI PODATKI V SLOVENIJI

Z delavnicami in posvetovanji smo v okviru projekta Odprti podatki skušali ugotoviti, kako načrtovati prihodnje aktivnosti na področju vzpostavljanja sistema odprtega dostopa do raziskovalnih podatkov, da bodo te uresničljive in učinkovite. Takšno obliko dela smo izbrali s ciljem, da bi akterje čim bolj aktivneje vključili v postopek priprave načrtov za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v njihovih okoljih. Interaktiven pristop je zagotovil možnost vplivanja na predlagane rešitve in usklajevanje načrtov z realnimi zmožnostmi. V okviru delavnic smo obravnavali različna področja življenjskega kroga podatkov in različne faze ravnanja z raziskovalnimi podatki. Izvedli pa smo tudi več posamičnih posvetov s strokovnjaki, katerih znanje in izkušnje segajo na različna področja ravnanja s podatki.

DELAVNICE PROJEKTA ODPRTI PODATKI

Od decembra 2012 do maja 2013 smo izvedli tri delavnice. Delavnice so bile načeloma odprte za zainteresirano javnost, vendar smo predhodno opravili preliminarni izbor in po e-pošti povabili akterje, za katere smo domnevali, da bi jih tema in področje morala zanimati zaradi njihovega področja dela. Na drugo delavnico smo povabili predstavnike ustvarjalcev podatkov in predstavnike s položajem moči odločanja, ki bodo igrali pomembno vlogo pri sprejemanju politik in akcijskega načrta za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov, financiranih z javnimi sredstvi. Na tretjo delavnico smo povabili razvijalce podatkovne infrastrukture, s ciljem, da že danes predstavijo obstoječa orodja in znanja ter se prijavijo za prevzemanje prihodnjih vlog pri vzpostavljanju nacionalnega sistema podatkovne infrastrukture.

Pri pripravi na delavnice smo izhajali iz poročila Opis stanja na področju raziskovalnih podatkov v Sloveniji ter pri tem sledili problemom in rešitvam, ki smo jih kot pomembne izluščili na podlagi opravljenih intervjujev.

Delavnica 1: Problemi in rešitve na področju podatkovnih storitev v Sloveniji, 15. december 2012

Prva delavnica je potekala kot nadaljevanje raziskave, ki smo jo predstavili v poročilu Opis stanja na področju raziskovalnih podatkov v Sloveniji. V poročilu ugotavljamo, da na področju ravnanja z raziskovalnimi podatki v Sloveniji obstaja vrsta ovir, ki otežujejo postopek vzpostavljanja sistema odprtega dostopa do podatkov, vzporedno pa nastajajo, se razvijajo in razširjajo pobude in prakse, ki predstavljajo rešitve na tem področju. Prvo delavnico smo namenili vprašanju podatkovnih storitev, ki predstavljajo pomemben del vseh faz življenjskega kroga podatkov, torej od načrtovanja raziskave pa do končne oddaje raziskovalnih rezultatov za namen druge rabe. Podatkovne storitve najprej obsegajo delo z ustvarjalci podatkov in poleg oddaje oziroma prevzema raziskovalnih rezultatov obsegajo tudi svetovanje glede priprave raziskovalnih podatkov na način, ki omogoča kakovostno hrambo in deljenje podatkov za namen druge rabe.

Pomemben del podatkovnih storitev predstavlja hramba podatkov, storitev, ki jo je mogoče primerjati s tradicionalnimi arhivskimi storitvami – skrb za ustreznost formatov, metapodatkov, izbor, vrednotenje ... In končno, delo z uporabniki raziskovalnih podatkov, ki obsega zagotavljanje dostopa in usposabljanje uporabnikov za ustrezno drugo rabo podatkov.

Shema 1: Vloge deležnikov na področju izvajanja podatkovnih storitev



Vir: Poročilo z delavnice Problemi in rešitve na področju raziskovalnih podatkov v Sloveniji

Delavnico Problemi in rešitve na področju podatkovnih storitev v Sloveniji smo zasnovali kot poskus evidentiranja obstoječih nosilcev podatkovnih storitev, ki lahko že danes predstavljajo zglede in dobre prakse na področju ravnanja z raziskovalnimi podatki. Vendar pa tudi kot poskus iskanja potencialnih prihodnjih nosilcev podatkovnih storitev, ki bi v oblikovanju novih storitev lahko prepoznali svoj prihodnji interes za poklicno sodelovanje.

Na delavnico smo povabili predstavnike ustvarjalcev državnih statistik, kot so SURS, IVZ in Arhiv RS, ki imajo veliko izkušenj s protokoli, standardi, metapodatki, notranjimi politikami idr. Povabili smo tudi akterje, okoli katerih že nastajajo nastavki podatkovnih storitev. To pomeni, da že imajo izkušnje z ravnanjem s podatki (tako z dajalci kot tudi z uporabniki podatkov) in da bi se iz njih v prihodnosti lahko razvili področni centri, specializirani za podatkovne storitve (npr. Glasbenonarodopisni inštitut ZRC SAZU). Na povabilo so se odzvali tudi predstavniki NUK-a, Arhiva RS in Univerze v Ljubljani; gre za večje ustanove, ki bi v prihodnosti lahko igrale pomembnejšo vlogo pri centralizirano koordiniranem izvajanju trajne hrambe in svetovanja na tem področju. Delavnice se je udeležilo tudi precejšnje število knjižničarjev iz različnih področnih znanstvenoraziskovalnih knjižnic.⁹⁹

V uvodnem delu delavnice smo k besedi povabili tri govorce, ki so na kratko predstavili svoje videnje stanja na področju podatkovnih storitev v Sloveniji. Matjaž Kragelj, NUK, je predstavil [Primer dobre prakse na področju strategij digitalne hrambe](#).¹⁰⁰ Mojca Kotar, Univerza v Ljubljani, Univerzitetna služba za knjižnično dejavnost, je predstavila [Odprti dostop do rezultatov javno financiranih raziskav v Evropi: kaj pa Slovenija?](#).¹⁰¹ Drago Kunej, Glasbenonarodopisni inštitut ZRC SAZU, je predstavil [Avtorske in sorodne pravice na primeru avdiovizualnega gradiva GNI ZRC SAZU](#).¹⁰²

⁹⁹ NIB in BF, Oddelek za biologijo, Filozofska fakulteta, Oddelek za geografijo, Filozofska fakulteta, Oddelek za etnologijo in kulturno antropologijo, Centralna ekonomska knjižnica, Fakulteta za družbene vede, Osrednja družboslovna knjižnica Jožeta Goričarja, Pedagoška fakulteta, Urbanistični inštitut Republike Slovenije.

¹⁰⁰ Glej: <http://www.adp.fdv.uni-lj.si/odpp10D1/presentations/kragelj.pptx>, 11. 2. 2014.

¹⁰¹ Glej: <http://www.adp.fdv.uni-lj.si/odpp10D1/presentations/kotar.pptx>, 11. 2. 2014.

¹⁰² Glej: <http://www.adp.fdv.uni-lj.si/odpp10D1/presentations/kunejPPT.pdf>, 11. 2. 2014.

Po uvodnih treh predavanjih smo 31 udeležencev in udeleženk razdelili v tri delovne skupine. V prvi skupini smo se posvetili vprašanju, ki zadevajo delo z ustvarjalci podatkov. Izhajali smo iz dela poročila Opis stanja na področju raziskovalnih podatkov, v katerem ugotavljamo, da je pomemben del odgovornosti glede ravnanja z raziskovalnimi podatki na ustvarjalcih podatkov, ki pravijo, da nimajo ne znanja, časa, kapacitet in ne sredstev, da bi razvili strategije, merila, pravila ..., potrebna za ravnanje z raziskovalnimi podatki. Zastavlja se vprašanje, kdo bi lahko prevzel vzpostavitev in izvajanje podatkovnih storitev ter kako bi ta postopek potekal. V prvi sekciji smo skušali odgovoriti na vprašanje, kaj lahko ponudimo ustvarjalcem podatkov. V razpravi je sodelovalo enajst udeležencev z različnih raziskovalnih ustanov: Glasbenonarodopisnega inštituta ZRC SAZU, Filozofske fakultete – Oddelek za etnologijo in kulturno antropologijo, Pedagoške fakultete, Inštituta za varovanje zdravja (IVZ), Statističnega urada RS (SURS), Narodne in univerzitetne knjižnice (NUK), Fakultete za družbene vede in ADP.

Udeleženci so ugotovili, da imajo ustanove, specializirane za ravnanje s podatki (SURS, IVZ), urejene notranje politike ravnanja s podatki, poznajo protokole in standarde. Poznajo tudi rešitve za ravnanje z občutljivimi podatki (npr. varna soba, registracija itd.). Na drugih ustanovah eksplicitnih pravil ni, vendar so udeleženci navajali, da ravnajo skladno z internimi dogovori in z drugimi pravili, ki se deloma nanašajo tudi na področje ravnanja s podatki. Udeleženci smo se strinjali, da bo v prihodnje potrebno spodbuditi oblikovanje politik ravnanja s podatki na ravni ustanov. Pri določanju pravil ravnanja s podatki in oblikovanju področnih podatkovnih centrov, ki bi prevzeli vlogo nosilcev posameznih storitev, naj ključno vlogo odigra država, saj sedanji zametki nosilcev storitev z obstoječimi kadri in kapacitetami niso sposobni širiti svojega delovanja. Strinjali smo se tudi, da je za hitrejše sporazumevanje, iskanje rešitev in izmenjavo izkušenj potrebno stremeti k oblikovanju informacijske točke, ki bi delovala posredniško med akterji na področju ravnanja z raziskovalnimi podatki.

Druga sekcija se je nanašala na politike hrambe oziroma pripravo gradiv za hrambo in dostop. V prvem delu poročila smo ugotovili, da ima digitalno skrbništvo tri splošne sestavine, ki jih je potrebno razviti, da bomo lahko govorili o zaupanju vrednem sistemu. To so priprava strategij, načrtov, organiziranje in izvajanje postopkov v institucionalnem okolju ter robustne tehnološke podlage. V drugi sekciji smo prišli do naslednjih zaključkov: da je potrebno oblikovati in institucionalizirati center za koordinirano izvajanje storitve trajne hrambe in svetovanja, kar bi vsaj deloma lahko prevzel NUK. Na inštitucijah, na katerih poteka ustvarjanje podatkov, bodo pomembno vlogo glede priprave podatkov za oddajo v področni podatkovni center lahko odigrali knjižničarji, raziskovalci in tehnično osebje. Udeleženci so med drugim sklenili tudi, da sta pomembni tako centralizacija storitev kot tudi oblikovanje področnih storitev pooblaščenih področnih podatkovnih centrov.

V drugi sekciji so sodelovali udeleženci z naslednjih ustanov: Statističnega urada RS, Moderne galerije, Filozofske fakultete UL – Oddelek za geografijo, Osrednje humanistične knjižnice, Urbanističnega inštituta, NUK, Centralne knjižnice Ekonomske fakultete UL, Knjižnice kmetijskega inštituta, Biološke knjižnice – Oddelka za biologijo (Biotehniška fakulteta) in iz ADP.

V tretji sekciji smo udeleženci obravnavali vprašanje politike dostopa oziroma smo skušali odgovoriti na vprašanje, kaj lahko v okviru podatkovnih storitev ponudimo uporabnikom. V poročilu ugotavljamo, da stanje na področju dostopa do podatkov v Sloveniji še ni sistemsko urejeno in da dostop največkrat poteka preko individualnih dogovorov med raziskovalci, iz česar sledi, da podatki pogosto ostanejo zaprti v ozke raziskovalne skupine in s tem tudi slabo

izkoriščeni. Zato je ena od pomembnih nalog prihodnjih politik, da se uredi dostop do podatkov. Pri teh pa je potrebno upoštevati utemeljene razloge za omejitev dostopa, hkrati pa si prizadevati za čim bolj odprt dostop do podatkov.

V diskusiji smo na kratko obdelali vse teme, ki so bile predstavljene v predstavitvi tretje sekcije (dostop; licence, citiranje, vrednotenje; metapodatki, agregacija, skupni katalogi). Vprašanja, za katera so udeleženci izrazili večje zanimanje oziroma so se nanašala na njihovo bolj specifično področje dela, smo nekoliko bolj razdelali in jih podrobneje obdelali.

Glavne dileme in ugotovitve teh bolj poglobljenih diskusij lahko povzamemo v štirih točkah:

1. Obstajajo centralne storitve za avtentikacijo (preverjanje pristnosti) in avtorizacijo uporabnikov (AAI), ki delujejo po principu, da se uporabniki pri njih registrirajo ter tako storitvi posredujejo informacije o sebi in opredelijo svoj status. Ponudniki storitev, kot so področna podatkovna središča, tem servisom zaupajo in nato na podlagi uporabnikovega statusa, registriranega pri takšnem centralnem servisu za avtentikacijo, uporabniku omogočijo dostop do podatkov. Pojavilo se je vprašanje, ali bi bil takšen sistem ustrezen za vse oblike podatkov ter ali bi se podatkovni centri lahko priključili takšnim storitvam.

Izpostavljeno je bilo tudi vprašanje, ali je dostop do podatkov, za katerega je potrebna avtentikacija, sploh še odprti dostop v pravem pomenu. Razmišljanja so šla v smeri, da takšen sistem dostopa upošteva načela odprtega dostopa ter da je za določene vrste podatkov potrebno zagotoviti določene omejitve dostopa. Bi pa bilo morda smiselno, kadar je to mogoče, pripraviti nekoliko prilagojeno verzijo gradiva, ki ne bi potrebovala tako strogih omejitev in bi pokrila potrebe manj zahtevnih uporabnikov. Ugotovili smo, da glede omejitev dostopa pri podatkih stvari ne moremo enostavno prenesti s področja zagotavljanja odprtega dostopa za publikacije, saj so bile slednje že v fazi nastanka namenjene objavi, medtem ko za podatke v veliki večini to ne velja.

2. Slovenija na področju vzpostavitve sistema odprtega dostopa nekoliko zaostaja za večino evropskih držav, vendar smo se sodelujoči strinjali, da lahko na to v trenutni situaciji gledamo kot na prednost: priporočila za vzpostavitev sistema dostopa (dobre prakse) iz tujine že obstajajo in so preizkušena, zato jih moramo le prevzeti in jih prilagoditi našim nacionalnim potrebam (npr. priporočila [EUDAT](http://www.eudat.eu/)¹⁰³).
3. Ugotovili smo, da je stanje glede dokumentacije metapodatkov večinoma še neurejeno, vendarle pa ponekod obstajajo interne tehnične dokumentacije za pripravo metapodatkov, ki so bile razvite za lastne potrebe. Takšne prakse bi lahko razširili tudi na področja, kjer to še ni uveljavljeno.
4. Nujno je, da so metapodatki interoperabilni, pri čemer je potrebno zagotoviti minimalne standarde, ki bodo to zagotavljali. Določitev teh minimalnih standardov je lahko naloga knjižničarjev in ne bi smela predstavljati večjih težav, medtem ko bi moralo biti vse ostalo, kar se tiče vsebine metapodatkov, v domeni ekspertov na posameznih področjih.

V tretji sekciji so sodelovali udeleženci z naslednjih ustanov: Univerze v Ljubljani – Univerzitetna služba za knjižnično dejavnost in Univerzitetna služba za informatiko, Ekonomske fakultete – Centralna ekonomska knjižnica, Instituta Jožef Stefan, –Laboratorij za umetno inteligenco,

¹⁰³ Glej: <http://www.eudat.eu/>, 11. 2. 2014.

Narodne in univerzitetne knjižnice, Univerze v Mariboru – Fakulteta za elektrotehniko, računalništvo in informatiko in Inštitut za računalništvo, ZRC SAZU – Inštitut za slovensko narodopisje, Pedagoške fakultete – knjižnica, Statističnega urada RS, IZUM-a, IVZ-ja, Zdravstveno podatkovnega centra.

Udeleženci so pokazali veliko mero pripravljenosti na sodelovanje pri izgradnji sistema odprtega dostopa do raziskovalnih podatkov. Medtem ko so se obstoječi nosilci posamičnih podatkovnih storitev zavedali posamičnih problemov in omejitev ter poznali rešitve, so knjižničarji premišljevali, na kakšne načine bi lahko pristopili in prevzeli posamične vloge in odgovornosti. Ob koncu delavnice smo udeleženci sprejeli skupno [izjavo s priporočili](#),¹⁰⁴ ki kličejo k vzpostavljanju nacionalnega okolja odprtega dostopa do podatkov iz javno financiranih raziskav.

Delavnica 2: Politike ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi, 15. marec 2013

Druga delavnica je v središče razprave postavila vprašanje krovnih politik ravnanja z raziskovalnimi podatki, model podatkovnega središča po načelu krovnih politik ter vprašanje raziskovalnih politik in politik znanstvenih revij. Na dogodek smo povabili predstavnike ustvarjalcev podatkov, predstavnike znanstvenih revij in predstavnike financerjev, torej akterje, ki lahko vplivajo na sprejem politik na področju ravnanja z raziskovalnimi podatki. Na povabilo se je odzvalo deset udeležencev z naslednjih ustanov: ARRS, Fakultete za socialno delo, Filozofske fakultete – Oddelek za bibliotekarstvo, Gozdarskega inštituta RS, Kmetijskega inštituta Slovenije, Oddelka za sadjarstvo in vinogradništvo, ZRC SAZU, NUK, Centra za razvoj knjižnic, revije Tekstilec (Naravoslovnotehniška fakulteta), revije Teorija in praksa (FDV) ter predstavniki ADP, ki smo vodili delavnico. Povabili smo tudi več predstavnikov različnih znanstvenih svetov ved, vendar žal ni bilo odziva.

V prvem delu delavnice smo udeležencem predstavili osnutek predloga politik ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi. Posebej smo izpostavili naslednje tri točke:

1. Obveznosti ustvarjalcev podatkov, ki so dolžni zagotoviti pravočasen, maksimalno odprt dostop do podatkov.
2. Infrastrukturne podporne storitve in storitve digitalne hrambe, ki so naloga podatkovnih centrov.
3. Odločevalce na vseh ravneh odločanja, katerih odgovornost je oblikovanje politik in nadzor ter zagotovilo pogojev za njihovo izvajanje.

Osnutek politik temelji na modelu, ki se imenuje načrt ravnanja z raziskovalnimi podatki oz. v izvirniku *Data Management Planning* in predvideva načrtno ravnanje s podatki od samega začetka, torej od načrtovanja raziskave, skozi cel postopek ustvarjanja podatkov do hrambe in končno druge rabe raziskovalnih podatkov. Načrt ravnanja s podatki pa je mogoče uresničiti le preko natančno določenih pravic in odgovornosti deležnikov. Zato krovne ali nacionalne politike že v osnutku vsebujejo:

1. Dolžnosti raziskovalcev
2. Odgovornost ustanov in opredeljevanje lastnih notranjih politik
3. Določila glede delovanja področnih podatkovnih centrov

¹⁰⁴ Glej: <http://www.adp.fdv.uni-lj.si/odpp10D1/images/Priporocila%20z%20delavnice.pdf>, 11. 2. 2014.

4. Obveznosti in pravice javnih financerjev
5. Notranje politike domačih znanstvenih periodičnih publikacij

Politike, kakor so zasnovane, predvidevajo vzpostavitev področnih podatkovnih centrov, ki bodo prevzeli naslednje naloge: prevzem, obravnavo in pripravo podatkov za dolgotrajno hrambo, svetovanje in usposabljanje pri pripravi načrtov ravnanja z raziskovalnimi podatki. Na podlagi pripravljenih in registriranih načrtov ravnanja z raziskovalnimi podatki bodo spremljali in potrjevali načrtovane aktivnosti. Njihove naloge so tudi vzpostavljanje infrastrukture za registracijo uporabnikov, zagotavljanje enoličnih identifikatorjev in ustrezno citiranje podatkov, skrb za promocijo, svetovanje in usposabljanje, organiziranje storitev vrednotenja podatkov po vnaprej objavljenih kriterijih in postopkih ter specifične storitve, odvisne od posameznega področja in zahtev uporabnikov. Kot primer edinega področnega podatkovnega centra v Sloveniji smo predstavili ADP, ki bi lahko bil za zgled pri ustanavljanju prihodnjih tovrstnih ustanov v Sloveniji.

Del politik se nanaša na obveznosti znanstvenih revij, financiranih z javnimi sredstvi. Te namreč preko svojih notranjih, uredniških politik določajo pravila citiranja, kadar gre za drugo rabo raziskovalnih podatkov. Da pri slovenskih znanstvenih revijah na tem področju še niso sprejeli pravil citiranja podatkov, gre vsaj deloma pripisati tudi dejstvu, da financer tega ne zahteva niti ne spremlja izvajanja. Citiranje podatkov pomeni vsaj delno povrnitev truda vloženega v ustvarjanje podatkov, prinese pa tudi možnost preverjanja podatkov. Z udeleženci druge delavnice smo govorili o obeh možnostih, da bi pobuda za citiranje prišla od zgoraj navzdol, torej s strani financerja, ali pa od spodaj navzgor, s samih revij oziroma njihovih uredništev.

Tako kot že prej v raziskavi, ki smo jo predstavili v Opisu stanja na področju raziskovalnih podatkov v Sloveniji, smo tudi v okviru druge delavnice ugotovili, da je močno prisotno nepoznavanje stanja na področju ravnanja z raziskovalnimi podatki, iz česar izhajajo tudi zadržki do odprtega dostopa. Med glavnimi zadržki je slaba razvitost storitev, ki bi bile v pomoč ustvarjalcem podatkov pri pripravi metapodatkov. Raziskovalci se zavedajo, da bi potrebovali nove storitve, nove kadre in dodatna sredstva, če bi želeli izvajati načrt ravnanja z raziskovalnimi podatki in tako prispevati k sistemu odprtega dostopa do raziskovalnih podatkov. V razpravi so večkrat izpostavili strah, da bodo dodatne zadolžitve pomenile manj časa in sredstev za raziskovanje ter več za dodatno administriranje. Med zadržki so izpostavili tudi strah pred napačno rabo podatkov, pred zlorabami in neupravičeno rabo – zlasti so izpostavili strateški in nacionalni pomen posameznih podatkov, v smislu da bodo pri odprtem dostopu manjše raziskovalne skupine, ki prevladujejo v Sloveniji, na slabšem v primerjavi z večjimi. Več časa so namenili tudi vprašanju lastništva podatkov in časa, v katerem bi bili po politikah dolžni podatke predati v hrambo in drugo rabo. Udeleženci, ustvarjalci podatkov, so podatke videli kot njihovo osebno last, saj so v ustvarjanje vložili osebno znanje in še zlasti ker so podatki neizkoriščena surovina, ki jim bo koristila na poti izgradnje osebne kariere znanstvenika. V razpravi se je za pokazalo, da je še kako pomembno raziskovalcem pojasniti prednosti odprtega dostopa.

Delavnica 3: Napredne tehnologije za vzpostavitev raziskovalne podatkovne infrastrukture v Sloveniji, 22. maj 2013

Na zadnji od treh delavnic smo predstavili stanje na področju raziskovalne podatkovne infrastrukture v Sloveniji. Prvi dve delavnici sta bili v večji meri namenjeni razpravi med udeleženci, tretja je bila zasnovana kot nabor predstavitev, ki so jih izvajali vabljeni govorniki. Naš

cilj je bil povabiti strokovnjake, ki se že danes ukvarjajo z razvojem podatkovne infrastrukture in lahko drugim ponudijo rešitve, služijo kot zgled.

Na povabilo se je odzvalo šest govorcev, po dva s fakultet za računalništvo in informatiko, po dva z Instituta Jožef Stefan in po en z ARNES-a in Morske biološke postaje Piran. Sledi seznam predstavitev po vrstnem redu, kot so potekale na delavnici:¹⁰⁵

1. *Predstavitev možnosti vključitve raziskovalnih podatkov v nacionalno infrastrukturo odprtega dostopa*, Milan Ojsteršek, Laboratorij za heterogene računalniške sisteme, Fakulteta za elektrotehniko, računalništvo in informatiko, UM
V okviru projekta ODUN poteka izgradnja repozitorijev univerz v Ljubljani, Mariboru, Novi Gorici in na Primorskem. Namen projekta je omogočiti odprt dostop do intelektualne produkcije slovenskih univerz. Univerzitetni repozitoriji so v prvi vrsti namenjeni zagotavljanju odprtega dostopa do objav, vanje pa bo mogoče vključiti tudi raziskovalne podatke. Projekt se konča 30. avgusta 2013. (Iz predstavitve M. Ojsterška, 22. 5. 2013)
2. *SeaDataNet: Ravnanje z oceanografskimi in morskimi podatki v okviru panevropske infrastrukture*, Branko Čermelj, Nacionalni inštitut za biologijo, Morska biološka postaja Piran
Morska biološka postaja Piran je vključena v SeaDataNet –panevropsko infrastrukturo, namenjeno ravnanju morskih in oceanografskih podatkov. Gre za vzpostavitev evropskih zbirk oziroma katalogov morskih projektov, podatkovnih zbirk, organizacij, opazovalnih sistemov v morskem okolju ter križarjenj. V infrastrukturo je trenutno povezanih 83 centrov, ki razvijajo skupne standarde za »morske« podatke, imajo skupno strategijo, licenco in orodja, ki se med seboj »razumejo«. Centri imajo enoten spletni portal. Eden izmed glavnih ciljev je zagotavljanje celovitega pregleda in dostop do podatkov, ki jih hranijo podatkovni centri. (Iz predstavitve B. Čermelja, 22. 5. 2013)
3. *Enotna prijava in napredni avtentikacijski sistemi*, Rok Papež, ARNES
Dostop do podatkov velja za eno pomembnejših storitev na področju vzpostavljanja sistema odprtega dostopa do raziskovalnih podatkov. V predstavitvi je bil govor o različnih oblikah dostopa in tveganjih, ki izvirajo iz njih. Predstavljeni so bili problemi množice uporabniških imen in gesel ter napredni mehanizmi avtentikacije, kot sta dvojna prijava in uporaba certifikatov za prijavo. Kot rešitev je bila predstavljena ločitev avtentikacije in storitve.
4. *Uporaba odprtokodnih rešitev pri zagotavljanju podatkovne infrastrukture*, Tomaž Erjavec, Odsek za tehnologije znanja, Institut Jožef Stefan
Govorec je predstavil »infrastrukturo« za področje (slovenskih) jezikovnih virov. Na portalu ijs.nl, ki obstaja že od leta 1995, teče več projektov. Med njimi: projekt JOS, ki je namenjen spodbujanju razvoja tehnologij za slovenski jezik, korpusi, slovarji, leksikoni, registri, digitalne knjižnice ... Pri uresničevanju projektov uporabljajo odprtokodne rešitve, ki prinašajo več prednosti:
 - odprta koda: brezplačna, možnost popraviljanja napak in dograjevanja;

¹⁰⁵ Predstavitve so dostopne na: <http://www.adp.fdv.uni-lj.si/odpp10D3/presentations.html>, 11. 2. 2014.

- odprti standardi: jasen, dobro dokumentiran, od implementacije neodvisen in trajen način zapisa besedil;
- odprte licence: enostavna in predvsem čim širša nadaljnja uporaba razvitih (jezikovnih) virov, zgrajenih z državnim financiranjem. (Iz predstavitve T. Erjavca, 22. 5. 2013).

5. *Iskanje po multimedijskih podatkih: primer na arhivu EtnoMuza*, Matija Marolt, Laboratorij za računalniško grafiko in multimedije, Fakulteta za računalništvo in informatiko, UL

Velika količina multimedijskih vsebin z majhnim deležem metapodatkov predstavlja poseben izziv. Na primeru arhiva EtnoMuza –digitalnega arhiva slovenske ljudske glasbe in plesa GNI ZRC SAZU –je bilo predstavljeno orodje za iskanje neposredno po gradivih. Med prvimi koraki je bila vzpostavitev arhiva, ki obsega orodje za vnos metapodatkov ter spletni vmesnik za iskanje in brskanje po metapodatkih. Naslednji korak je prinesel razvoj naprednih tehnologij, s katerimi je mogoče iskati po notografijah in zvočnih posnetkih, o katerih ni metapodatkov ali jih je zelo malo.

Predstavljeno je bilo orodje za avtomatsko prepoznavanje vsebine. Členitev gradiva na: govor, solo petje, večglasje, instrumental, pritrkavanja. Možnost členitve vsebine na:

a) petje in inštrumentali:

- iskanje po melodiji, parametrih melodije,
- sestav, število glasov, spol,

b) govor:

- število govorcev, spol,

c) pritrkavanja:

- iskanje po vzorcih,
- število zvonov. (Iz predstavitve M. Marolta, 22. 5. 2013)

6. *Slovenska infrastruktura in omrežja evropskega raziskovalnega prostora*, Jan Jona Javoršek, Institut Jožef Stefan in SLING

V zadnji predstavitvi je bilo predstavljeno mrežno računalništvo, predvsem pa potencial grida oz. kdaj, kako in zakaj ga uporabiti. Grid je vmesna programska oprema, ki jo je smiselno uporabiti, ko imamo opravka z veliko količino podatkov, ko izvrševanje nalog zahteva veliko procesorsko moč ali ko gre za veliko število nalog. Grid omogoča lažjo uporabo mrežnih računalniških kapacitet (enotni, standardizirani vmesniki, samodejna izbira, gruče, upravljanje z nalogami in podatki), uporabo novih tehnologij (*high throughput computing*, nameščanje programske opreme, virtualizacija na različnih stopnjah ipd.) ter lažje znanstveno-raziskovalno sodelovanje (*virtual research communities*).

Predstavljena je bila tudi Slovenska iniciativa za nacionalni grid (SLING). Ključna razloga zanjo sta vzpostavitev stalne nacionalne infrastrukture in vključitev v mednarodno infrastrukturo EGI. (Iz predstavitve J. J. Javorška, 22.5.2013)

Delavnica je imela 34 udeležencev, ki so prišli z različnih ustanov, med njimi so bili štirje predstavniki treh različnih ministrstev, trije z dveh fakultet za računalništvo in informatiko, sedem predstavnikov iz nacionalne in različnih področnih knjižnic, po nekaj predstavnikov

zasebnih in javnih podjetij ter raziskovalnih inštitutov in fakultet ter po en predstavnik SURS in ZRC SAZU. Ker so potrebe po naprednih podatkovnih tehnologijah velike, je bil velik odziv udeležencev na predstavljene rešitve pričakovan. Udeleženci so z vprašanji aktivno sodelovali in ustvarjali tok razprave, ki je potekal od prvih seznanitev s posameznimi rešitvami do izmenjave dobrih izkušenj in navezovanja prihodnjih stikov.

Zaključiti je mogoče, da v Sloveniji že imamo strokovnjake, ki se na svojih področjih poglobljeno ukvarjajo z razvojem podatkovnih tehnologij in lahko predstavljajo pomembna središča pri vzpostavljanju sistema odprtega dostopa do raziskovalnih podatkov. Ker prihajajo z različnih raziskovalnih področij (naravoslovje, humanistika, družboslovje), poznajo problematiko različnih tipologij podatkov in tehnoloških ovir ter prednosti tehnologije, s katero delajo.

Udeleženci so pozdravili srečanje in poudarili potrebo po nadaljevanju tovrstnih aktivnosti, s katerimi področne ali parcialne rešitve vstopijo v skupno raziskovalno okolje. Omeniti še velja, da smo delavnico organizirali skupaj z Geodetskim inštitutom, in sicer s skupnim naslovom *Odpiranje raziskovalnih in prostorskih podatkov v Sloveniji*. Delavnici Napredne tehnologije za vzpostavitev raziskovalne podatkovne infrastrukture v Sloveniji je sledila delavnica Geodetskega inštituta z naslovom Razvoj geostoritev z odprtimi prostorskimi podatki.

POSVETI S STROKOVNJAKI V OKVIRU PROJEKTA ODPRTI PODATKI

Pri iskanju problemov in rešitev na poti k vzpostavljanju sistema odprtega dostopa do raziskovalnih podatkov smo opravili tudi več posvetov s strokovnjaki, katerih znanja in izkušnje segajo na področje raziskovalnih podatkov.

Urad Informacijskega pooblaščenca, 15. januar 2013

V pogovoru s slovenskimi raziskovalci smo prišli do ugotovitve, da jih je pogosto strah, da bi z deljenjem podatkov (odprti podatki) kršili pravico do zasebnosti oz. varovanja osebnih podatkov. Zato so raje restriktivni in je to eden od argumentov, zakaj podatkov, financiranih z javnimi sredstvi, ne delijo z drugimi. Na delavnici Problemi in rešitve na področju podatkovnih storitev v Sloveniji (5. december 2012) smo v sekciji Delo z ustvarjalci podatkov prišli do sklepa, da je problemov, ki zadevajo varovanje osebnih podatkov, še več – že znani primer besedilnega korpusa; primer NUK, ki arhivira splet (po analogiji, kot arhivira časopise). Pojavljajo se posamezniki, ki iz različnih vzrokov zahtevajo odstranitev svojih imen in priimkov. Za raziskovalne namene so splošne rešitve, kot npr. izbris/zakrivanje vseh osebnih podatkov, lahko škodljive, saj vrednost podatkovne zbirke močno okrnijo ali pa sploh onemogočijo širšo znanstveno rabo gradiva, ki je bilo ustvarjeno z javnimi sredstvi.

Tudi udeleženci delavnice so opozorili na razkorak med zahtevami financerja raziskovalnih projektov, ki npr. zahteva odprti dostop do podatkov, in zakonodajno prakso urada Informacijskega pooblaščenca, ki raziskovalne podatke obravnava enako kot podatke drugih javnih ustanov in lahko deljenje podatkov tudi sankcionira.

Da bi premostili strahove, da bi druge in sebe zavarovali pred morebitnimi kršitvami ter da bi predstavili posebne interese pri ravnanju s podatki znanstvene skupnosti, smo se dogovorili, da se posvetujemo z uradom Informacijskega pooblaščenca (IP). Na sestanku smo se srečali 15. januarja 2013, in sicer z namenom, da bi oblikovali delovno skupino, ki bi se v prihodnje ukvarjala s tem, kako zadostiti dvema interesoma: varovanju osebnih podatkov in maksimalni

izrabi podatkovnih gradiv, ki so nastala z javnimi sredstvi. V IP so na voljo za morebitna odprta vprašanja, ki bi jih obravnavala tovrstna delovna skupina.

Urad za intelektualno lastnino, 17. januar 2013

Raziskava med ustvarjalci podatkovnih zbirk in izvajalci podatkovnih storitev je pokazala, da je ena od ovir pred deljenjem podatkov strah, da bi kršili avtorske pravice, ki v digitalnem okolju ostajajo odprto vprašanje brez dokončnih odgovorov. Strah pred kršitvami so potrdili tudi udeleženci naših delavnic. Zato je vprašanje avtorskih pravic v digitalnem okolju zagotovo eno od obveznih vprašanj na poti vzpostavljanja sistema odprtega dostopa do raziskovalnih vprašanj.

Pri iskanju rešitev na tem področju smo se obrnili tudi na Urad za intelektualno lastnino. Sestanek je bil seznanitvene narave in poskus iskanja možnosti oblikovanja svetovalne storitve na področju avtorskih pravic. Ena od možnosti bi bila, da velike raziskovalne ustanove skupaj zaposlijo pravnik, ki skrbi za usposabljanje raziskovalcev in razreševanje dilem na področju intelektualne lastnine. Obstajajo pa tudi možnosti sodelovanja pri pilotnih projektih, ki so namenjeni iskanju rešitev na tem področju.

Javna agencija za raziskovalno dejavnost Republike Slovenije, 10. julij 2013

Pomemben del vzpostavljanja kulture deljenja podatkov predstavlja ustrezno vrednotenje druge rabe raziskovalnih podatkov. Znanstveno ovrednotenje namreč pomeni priznanje za raziskovalčev napor, vložen v ustvarjanje raziskovalnih podatkov, in če je ustrezno opredeljeno, ugodno vpliva na potek njegove znanstvene kariere. Sestanek pri direktorju ARRS Franciju Demšarju je bil seznanitvene narave, z njim smo predstavnik financerja slovenske raziskovalne dejavnosti skušali seznaniti s problemi slovenske znanstvene skupnosti in predstavili predlagane smeri iskanja rešitev, ki se uveljavljajo tudi v mednarodnem, evropskem raziskovalnem okolju. ARRS pobude mednarodnih krogov spremlja in se zaveda, da se bo tudi v slovenskem okolju potrebno zavzeti za spremembo sistema v smeri, ki bo pripoznala raziskovalne podatke kot samostojne produkte v raziskovalnem okolju.

NASTAVKI ZA RAZVOJ PODATKOVNIH STORITEV IN INFRASTRUKTURE V SLOVENIJI

Na podlagi izvedenih aktivnosti je mogoče ugotoviti, da imamo na eni strani široko raziskovalno skupnost, ki je slabše seznanjena z vprašanjem odprtega dostopa do raziskovalnih podatkov. Posledično smo v teku naših aktivnosti pogosto odgovarjali na osnovna vprašanja, kot kaj so prednosti odprtega dostopa, kako je vprašanje urejeno v drugih državah, zakaj je deljenje podatkov v interesu znanosti in drugo. Po drugi strani pa ugotavljamo, da tudi v slovenski raziskovalni skupnosti že obstajajo otoki znanja in izkušenj na področju podatkovne infrastrukture in s podatki povezanih storitev.

Za prihodnje načrtovanje tukaj navajamo seznam ustanov, ki so bile v teku aktivnosti prepoznane kot potencialni prihodnji nosilci splošnih podatkovnih storitev oziroma infrastruktur:

- svetovanje na področju dela z občutljivimi podatki (varovanje osebnih podatkov, anonimizacija, omejen dostop): Statistični urad Republike Slovenije;

- storitev razvoja in dodeljevanja trajnih identifikatorjev: Nacionalna univerzitetna knjižnica (NUK);
- svetovanje ne področju vrednotenja in izbora podatkovnih datotek ter vpisa v sistem SICRIS in COBISS: Arhiv družboslovnih podatkov v sodelovanju z Osrednjo družboslovno knjižnico Jožeta Goričarja in IZUM;
- centralizirano koordinirano izvajanje trajne hrambe in svetovanja s tega področja: NUK in Laboratorij za heterogene računalniške sisteme, Fakulteta za elektrotehniko, računalništvo in informatiko, Univerza v Mariboru;
- storitev registriranega dostopa (prijava in avtentikacijski sistemi) AAI: ARNES;
- velika količina podatkov, malo metapodatkov: Slovenska iniciativa za nacionalni grid (SLING), Laboratorij za računalniško grafiko in multimedije, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani;
- uporaba odprtokodnih rešitev pri zagotavljanju podatkovne infrastrukture: Odsek za tehnologije znanja, Institut Jožef Stefan.

Kot že večkrat povedano, pa bo potrebno spodbujati tudi razvoj področnih podatkovnih središč. Kot primer že vzpostavljenega in vrsto let delujočega podatkovnega središča velja izpostaviti Arhiv družboslovnih podatkov, v nadaljevanju pa še nabor drugih jeder, prepoznanih kot možnih zasnov prihodnjih področnih podatkovnih središč.

Družboslovje: Arhiv družboslovnih podatkov (ADP)

ADP¹⁰⁶ je področno podatkovno središče za družboslovje, ki hrani zbirko podatkov, zanimivih za družboslovne analize, s poudarkom na problemih, povezanih s slovensko družbo. Prednost imajo vsebinsko pomembnejše in metodološko dobro izdelane raziskave, ob tem še posebej podatki, ki jih je mogoče primerjati v različnih časovnih obdobjih, in mednarodni primerjalni podatki, pri katerih je upoštevana Slovenija. Ciljni uporabniki so domači in tuji raziskovalci, učitelji in študentje, ki so podatkovno in statistično pismeni¹⁰⁷ za samostojno ravnanje z gradivi. Za dostop do gradiv, ki jih hrani Arhiv družboslovnih podatkov, se je potrebno registrirati.¹⁰⁸ Arhiv od dajalcev sprejema podatke, ki dosegajo osnovna merila kakovosti in vsebinske zanimivosti za nadaljnje analize. Izročeni podatki štejejo kot znanstvena objava po merilih ARRS.

ADP hrani približno 600 raziskav z različnih družboslovnih področij. V glavnem gre za kvantitativne, nekaj pa je tudi kvalitativnih raziskav. ADP poleg podatkovne datoteke za posamezno raziskavo hrani tudi druga spremljajoča gradiva, potrebna za razumevanje podatkov. Običajno so to kopije originalnih vprašalnikov, lahko pa tudi šifranti (npr. standardna klasifikacija poklicev), obvestilna pisma vzorčnim enotam, navodila anketarjem itd.

Od leta 1998 je ADP član evropskega združenja družboslovnih arhivov – [CESSDA](#).¹⁰⁹ Novoustanovljena evropska infrastrukturna enota iz pobude ESFRI, to je CESSDA AS, katere ustanovni član je tudi Republika Slovenija, predstavlja organizacijsko in strokovno nadgradnjo podatkovne infrastrukture na področju družboslovja, s ciljem zagotoviti pogoje za vzpostavitev dobro opredeljene, celovite in med seboj povezane podatkovne infrastrukture za družboslovje tako na nacionalni kot na mednarodni ravni. ADP je s strani RS v [Načrtu razvoja raziskovalnih](#)

¹⁰⁶ *About the Social Science Data Archive in Slovenia*, <http://www.openaire.eu/en/component/content/article/76/402>, 15. 1. 2014.

¹⁰⁷ Glej: http://sl.wikipedia.org/wiki/Podatkovna_pismenost, 7. 8. 2012.

¹⁰⁸ Glej: <http://adp.fdv.uni-lj.si/registracija/>, 7. 8. 2012.

¹⁰⁹ Glej: <http://www.cessda.net>, 7. 8. 2012.

[infrastruktur 2011–2020](#)¹¹⁰ pooblaščen za ustrezno raven izvajanja nacionalnih storitev CESSDA AS.¹¹¹

Opaziti pa je mogoče tudi jedra, iz katerih bi se v prihodnosti lahko razvila področna podatkovna središča, saj imajo ustanove že specifična znanja, vezana na področno metodologijo, izdelana merila kakovosti, poznajo vrste podatkov. V nadaljevanju predstavljamo nekaj takih jeder, ki imajo potencial, da bi lahko prevzeli večje obveznosti na specifičnih področjih storitev.

Biologija: Morska biološka postaja Piran (MBP)

Morska biološka postaja deluje od leta 1969, ko so jo ustanovili kot samostojno enoto Inštituta za biologijo Univerze iz Ljubljane. Znotraj MBP deluje tudi instrumentalni center Morska biološka postaja Piran, katerega osnovni namen sta pretok znanja in nudenje velike raziskovalne opreme uporabnikom. Z vidika ravnanja z raziskovalnimi podatki je posebej zanimivo sodelovanje v projektu [SeaDataNet](#).¹¹² Gre za vseevropski projekt razvoja in zagotavljanja infrastrukture, namenjene označevanju, zagotavljanju in ravnanju z oceanografskimi podatki in podatki o morskem okolju. Med storitvami, ki jih zagotavljajo partnerji projekta, je tudi dostop do podatkov in najrazličnejših produktov (fizikalne, kemične, geološke, batimetrične in biološke lastnosti) ter zagotavljanje varnosti, dolgoročno shranjevanje in skrbništvo nad podatki. V infrastrukturi SeaDataNet so povezani podatkovni centri, ki zagotavljajo (ustvarjajo) podatke, skrbijo za njihovo kakovost in medsebojno povezljivost ter skrbijo, da so podatki na voljo vsem uporabnikom preko osrednjega portala.

Njihove izkušnje s področja iskanja podatkov, varovanja podatkov, prenosa podatkov, ogleda podatkov, nadzora sistema in vzdrževanja sistema, kakor tudi znanja s področja povezljivosti podatkov so lahko v pomoč pri vzpostavljanju področnih podatkovnih središč.¹¹³

Zgodovina: spletna stran Zgodovina Slovenije – Sistory

Zgodovina Slovenije – Sistory je spletna stran Inštituta za novejšo zgodovino (INZ) v Ljubljani. Izobraževalni in raziskovalni portal slovenskega zgodovinopisja je javno dostopen od konca leta 2008. V štirih vsebinskih/področnih sklopih je uporabnikom na voljo obsežen nabor zgodovinskih in zgodovinopisnih vsebin z adekvatnimi analitičnimi sistemi in ustreznim znanstvenim aparatom.

Gradnja portala še poteka in vključuje prenos gradiva v novo medijsko in tehnološko obliko oziroma digitalizacijo in predstavitev gradiva na spletni strani, s tem pa tudi skrbi za podaljšanje njegove relevance. V letu 2011 so v INZ skenirali 13.000 strani, v fazi digitalizacije pa je 43.000 strani arhivskega gradiva.

Leta 2011 je INZ portal nadgradil v spletni sistem, ki je primerljiv z evropskimi in svetovnimi trendi in sistemi, ki omogočajo ustvarjanje raziskovalnih podatkov in njihovo uporabo v raziskovalnem okolju. Spletna stran je kompatibilna z ostalimi sorodnimi evropskimi spletnimi portali oziroma z arhitekturo evropske mreže za humanistiko (DARIAH), v prvi fazi pa predstavlja predvsem izhodišče sodelovanja z ZRC SAZU pri vzpostavitvi nacionalne digitalne

¹¹⁰ Glej: <http://www.arhiv.mvzt.gov.si/fileadmin/mvzt.gov.si/pageuploads/pdf/znanost/nacrt-RI.pdf>, 7. 8. 2012.

¹¹¹ Več na: <http://www.adp.fdv.uni-lj.si/blog/2014/blog/council-of-european-social-science-data-archives-arhiv-druzboslovnih-podatkov/>, 15. 1. 2014.

¹¹² Glej: <http://www.seadatanet.org/>, 15. 1. 2014.

¹¹³ Glej: <http://www.mbss.org/portal/index.php>, 15. 1. 2014.

infrastrukture za humanistiko in umetnost z akronimom SI-DIH, ki je opredeljen z osnutkom Načrta razvoja RI 2011–2020 Ministrstva za visoko šolstvo, znanost in tehnologijo.

Sami pravijo, da je spletni portal Zgodovina Slovenije – SIstory po nadgradnji dokončno presešel vlogo repozitorija, knjižnice ali arhiva gradiv v klasičnem pomenu besede in je povsem primerljiv z evropskimi in svetovnimi trendi in sistemi, ki omogočajo ustvarjanje raziskovalnih podatkov in njihovo uporabo v raziskovalnem okolju. Verificirane vsebine, objavljene na spletnem portalu Zgodovina Slovenije – SIstory, so pritegnile raziskovalne in izobraževalne kadre različnih disciplin doma in v tujini, hkrati pa so vzbudile zanimanje tudi med osnovnošolci in dijaki ter ne nazadnje med laično publiko.¹¹⁴

Umetnost: arhiv in zbirke Glasbenonarodopisnega inštituta ZRC SAZU

Glasbenonarodopisni inštitut je bil ustanovljen leta 1934 in je kot tak najstarejši inštitut ZRC SAZU. Ena od temeljnih nalog tega inštituta je »sestavljanje čim popolnejše zbirke slovenske glasbene folklore«. Poleg slovenskih ljudskih pesmi v arhivu hranijo raznovrstne zapise ljudske glasbe drugih narodov in razne dokumentarne posnetke. V njihovih zbirkah je mogoče najti: pisni arhiv (rokopisi, pesmarice, zapuščine, transkripcije zvočnih posnetkov, zapisniki, dokumentacija delovanja inštituta in starejših akcij zbiranja ljudskih pesmi), zvočni arhiv (posnetki na voščenih valjih, na gramofonskih ploščah, na magnetofonskih trakovih, na kasetah, na digitalnih zvočnih kasetah (DAT) in posnetki na zgoščenkah), vizualno dokumentacijo (filmi s folklorno tematiko, videokasete različnih formatov, fototeka in diateka) ter zbirko ljudskih inštrumentov.

Bogatenje zbirke je bilo in je še zmeraj močno povezano z razvojem snemalnih naprav in možnosti, ki jih je imel na voljo inštitut, hkrati pa je rezultat povezovanja inštituta navzven.

Danes je temeljna dejavnost Glasbenonarodopisnega inštituta poleg rednega raziskovalnega dela in objavljanja raziskav predvsem izdajanje temeljnih in drugih zbirk ljudskih pesmi, sistematično izdajanje ljudskih plesov, izpopolnjevanje, hranjenje in objavljanje izjemnega arhiva na kasetah in zgoščenkah, spremljanje sodobne dejavnosti, povezane z glasbenim in plesnim izročilom, in strokovno posredovanje na tistih področjih, kjer se folklorizem oplaja ob pesemski, glasbeni in plesni dediščini.¹¹⁵

Humanistika: spletna stran SI-DIH

Na področju humanistike predstavlja pomembno novost portal SI-DIH, ki je skupni projekt Inštituta za novejšo zgodovino in Znanstvenoraziskovalnega centra Slovenske akademije znanosti in umetnosti. Spletna stran je še v izgradnji, osnovo pa v tej fazi predstavljajo vsebine portalov Zgodovina Slovenije – SIstory (INZ) in Arzenal (ZRC SAZU). Iskalnik poizveduje po različnih repozitorijih oziroma arhivih institucij ali društev s področja humanistike in umetnosti. Trenutno so v iskanje vključeni naslednji repozitoriji:

- [Arzenal](#),¹¹⁶ Virtualna zakladnica nacionalne dediščine,
- Etnoinfolab, Oddelek za etnologijo in socialno antropologijo, Filozofska fakulteta UL,
- [SIstory](#),¹¹⁷ Zgodovina Slovenije,

¹¹⁴ Glej: <http://www.sistory.si/>, 14. 1. 2014.

¹¹⁵ Glej: <http://gni.zrc-sazu.si/#v>, 14. 1. 2014.

¹¹⁶ Glej: <http://www.arzenal.si/>, 14. 1. 2014.

- [Inštitut za slovensko narodopisje](#),¹¹⁸ ZRC SAZU,
- monografije [Založbe ZRC](#).¹¹⁹

V prihodnje bodo mrežo partnerjev širili, saj so si zadali postati vozlišče slovenske znanstvene digitalne humanistike.

SI-DIH primarno vključuje digitalna gradiva, hkrati pa ponuja tudi gradiva starejšega nastanka, ki so bila digitalizirana z namenom, da bodo javno objavljena in tako dostopna širšemu krogu uporabnikov. Za enkrat je raziskovalcem na voljo spletni iskalnik po digitalnih vsebinah, v prihodnosti pa ga bodo skušali čim bolj prilagoditi potrebam humanistike. Poleg tega je njihov namen spodbuditi raziskovalce k uporabi novih tehnologij ter izkoriščanju možnosti, ki jih te ponujajo. In sicer s ciljem, da bo »mogoče zbrano znanje ponovno interpretirati ali inovativno pristopiti k novim raziskavam.« Kar zadeva tehnologijo, so v predstavitvi strani zapisali tako:

Iskalnik je zasnovan na principu trajne povezave med repozitoriji, ki posredujejo metapodatke, in agregatorjem (iskalnikom SI-DIH), ki te metapodatke hrani in redno posodablja v svoji računalniški bazi. Iskalnik SI-DIH omogoča hitro in učinkovito iskanje, prezentacija metapodatkov pa je omejena na osnovno metapodatkovno shemo (Dublin jedro) in povezavo na lokalni repozitorij, kjer ima uporabnik praviloma na voljo bogatejše metapodatkovne opise in napredne prezentacijske ali analitične tehnologije. Iskalnik ima predvsem posredniško vlogo, uporabniku pomaga najti podatek ter mu ponudi pot, kako do podatka priti.¹²⁰

SI-DIH že ponuja možnost hranjenja digitaliziranih zbirk za tiste organizacije ali skupine, ki nimajo pogojev za vzdrževanje strežniškega repozitorija, v katerem bi lahko shranili posamezne zbirke ali arhive, ki so nastali med njihovim delom; s tem SI-DIH prevzema tudi vlogo digitalnega arhiva.

Portal SI-DIH je partner nastajajoče evropske infrastrukturne mreže za digitalno humanistiko DARIAH in aktivno sodeluje pri njeni izgradnji.¹²¹

Medicina: Inštitut za biomedicinsko informatiko (IBMI)

IBMI je ustanovila Medicinska fakulteta¹²², in sicer leta 1973 s ciljem, da se bo ustanova posvečala informacijam, ki so pomembne za raziskovalno delo. Že od samih začetkov na inštitutu ločujejo strokovni in izobraževalni oziroma raziskovalni program. Jedro strokovnega programa tvori slovenska nacionalna bibliografska zbirka Biomedicina Slovenica, ki vsebuje podatke o slovenskih znanstvenih in strokovnih dokumentih s področja biomedicine.

V okviru inštituta deluje biostatistični center, ki vodi raziskovalni program in sodeluje pri raziskovalnih projektih s poudarjeno metodološko in statistično podporo pri delu na raziskovalnih podatkih s področja medicine, svetuje tako pri ustvarjanju kot pri analizi. Znotraj inštituta deluje tudi Infrastrukturni center za bibliometrijo, ki med storitvami nudi:

¹¹⁷ Glej: <http://www.sistory.si/?language=sl>, 14. 1. 2014.

¹¹⁸ Glej: <http://isn.zrc-sazu.si/index.php?q=sl>, 14. 1. 2014.

¹¹⁹ Glej: <http://zalozba.zrc-sazu.si/sl/predstavitev#v>, 14. 1. 2014.

¹²⁰ Glej: <http://www.sidih.si/?page=2>, 11. 2. 2014.

¹²¹ Glej: <http://www.sidih.si/>, 7. 8. 2012.

¹²² Glej: <http://www.mf.uni-lj.si/>, 7. 8. 2012.

- biostatistične analize v Biostatističnem centru,
- bibliometrične raziskave in analizo citiranosti,
- avtomatizacijo gradnje spletnih sistemov za ravnanje s podatki,
- odkrivanje znanja v podatkovnih zbirkah,
- vzdrževanje bibliografij slovenskih biomedicinskih raziskovalcev in drugo.

Pomemben del dejavnosti infrastrukturnega centra sodi na področje ravnanja s podatki. Razvijamo in nudimo pomoč pri uporabi sistema za podporo spletnemu zbiranju podatkov in ravnanju z njimi. Sistem lajša gradnjo relacijskih zbirk poljubno kompleksnih podatkovnih objektov in gradnjo orodij za njihovo spletno zbiranje.¹²³

Center za funkcijsko genomiko in biočipe (CFBGC)

CFBGC organizacijsko sodi pod okrilje Inštituta za biokemijo, med ustanovitelji pa so Slovenski konzorcij za biočipe, združenje slovenskih akademskih institucij, raziskovalnih inštitutov, kliničnih centrov in farmacevtske industrije. Deluje od leta 2005 in je del [Mreže raziskovalnih infrastrukturnih centrov Univerze v Ljubljani](#) (MRIC UL).¹²⁴ Med glavne aktivnosti centra sodijo:

- temeljne, aplikativne in klinične postgenomske raziskave,
- upravljanje in vzdrževanje obstoječe opreme,
- nabava nove opreme,
- usposabljanje študentov na vseh študijskih ravneh.

So del evropske infrastrukturne enote ELIXIR¹²⁵. Svoje storitve pa center ponuja članom slovenskega konzorcija za biočipe in tudi drugim. Nudijo več ravni storitev od popolne (zagotovitev osebja, materiala in opreme), usposabljanje (za delo na določeni napravi), omogočajo neodvisno delo (na napravi).¹²⁶

Odsek za tehnologije znanja

Odsek za tehnologije znanja, Institut Jožef Stefan, se ukvarja z jezikovnimi viri in tehnologijami, vezanimi na slovenski jezik in druge evropske jezike. Podatkovni viri in storitve, ki so pod licenco Creative Commons dostopni na [Strežniku za naravne jezike](#)¹²⁷/Natural language server, so namenjeni čim širši rabi v znanosti, predvsem družboslovju in humanistiki.

Strežnik nl.ijs.si deluje od leta 1995 in predstavlja infrastrukturo za področje jezikovnih virov. Na strežniku je mogoče dostopati do korpusov za empirično jezikoslovje, slovarjev, registrov in digitalne knjižnice. Osnovna programska oprema deluje na odprtokodnih rešitvah, vse je prosto dostopno za pregledovanje, večina pa je tudi prosto dostopna za prenos. V nadaljevanju navajamo nekaj orodij, ki so jih razvili na Institutu Jožef Stefan ali v projektih, v katerih je IJS sodeloval kot partner.

V okviru projekta [IOS jezikoslovno označevanje slovenskega jezika](#)¹²⁸ so razvili označene korpusne slovenskega jezika in pridružene vire, namenjene spodbujanju razvoja jezikovnih

¹²³ Glej: <http://ibmi.mf.uni-lj.si/ibmi/index.html>, 7. 8. 2012.

¹²⁴ Glej: http://www.uni-lj.si/raziskovalni_in_razvojni_projekti/mreza_raziskovalnih_infrastrukturnih_centrov.aspx, 14. 1. 2014.

¹²⁵ Glej: <http://www.elixir-europe.org/about/elixir-slovenia>, 14. 1. 2014.

¹²⁶ Glej: <http://cfgbc.mf.uni-lj.si/default.html> in <http://www.mf.uni-lj.si/cfgb>, 14. 1. 2014.

¹²⁷ Glej: <http://nl.ijs.si/e-zrc/>, 14. 1. 2014.

¹²⁸ Glej: <http://nl.ijs.si/ios/josMSD-sl.html>, 14. 1. 2014.

tehnologij za slovenski jezik. Preko vstopnih točk na spletni strani [NoSketch Engine](#)¹²⁹ je na strežniku dostopno [večje število korpusov](#),¹³⁰ od katerih je večina slovenskih, vendar so tudi v drugih jezikih, predvsem angleščini in japonščini, na voljo so tudi orodja za analizo korpusov. Na spletni strani z naslovom [jaSlo](#)¹³¹ so dostopni japonsko-slovenski viri za učenje japonščine, npr. slovarji, korpusi in literatura. [ezb](#)¹³² je repozitorij elektronskih znanstvenih besedil in vsebuje prosto dostopne sodobne elektronske izdaje različnih elektronskih znanstvenih besedilnih virov, med njimi elektronske znanstvene monografije, [ezMono](#),¹³³ Neznane rokopise slovenskega slovstva 17. in 18. stoletja ([NRSS](#)),¹³⁴ Slovenski biografski leksikon ([SBL](#))¹³⁵ in Elektronske znanstvenokritične izdaje slovenskega slovstva ([eziss](#)).¹³⁶

Odsek za tehnologije znanja sodeluje v CLARIN, evropski infrastrukturi za skupne jezikovne vire in tehnologije.¹³⁷

¹²⁹ Glej: <http://nlp.fi.muni.cz/trac/noske>, 14. 1. 2014.

¹³⁰ Glej: <http://nl.ijs.si/noske/index.html#table>, 14. 1. 2014.

¹³¹ Glej: <http://nl.ijs.si/jaslo/index-sl.html>, 14. 1. 2014.

¹³² Glej: <http://ezb.ijs.si/>, 14. 1. 2014.

¹³³ Glej: <http://ezb.ijs.si/fedora/get/ezmono:ezmono/VIEW/>, 14. 1. 2014.

¹³⁴ Glej: <http://ezb.ijs.si/fedora/get/nrss:nrss/VIEW/>, 14. 1. 2014.

¹³⁵ Glej: <http://ezb.ijs.si/fedora/get/sbl:sbl/VIEW/>, 14. 1. 2014.

¹³⁶ Glej: <http://nl.ijs.si/e-zrc/>, 14. 1. 2014.

¹³⁷ Glej: <http://nl.ijs.si/> in <http://www.adp.fdv.uni-lj.si/odpp10D3/presentations.html>, 14. 1. 2014.

MEDNARODNI PREGLED DOGAJANJA NA PODROČJU DOSTOPA DO RAZISKOVALNIH PODATKOV

IZHODIŠČA

Podatki so temeljna sestavina raziskovalnega procesa, njihova kakovost, dostopnost in preglednost pa pogojujejo kakovost raziskovalnih zaključkov v približevanju spoznanjem in v uporabnosti rezultatov. Povečujeta se informacijska pismenost populacije in dostop do svetovnega spleta, hkrati pa se v vseh znanostih množi število podatkov v digitalni obliki, npr. genska zaporedja v bioinformatiki, podatki iz astronomskih opazovanj, prostorski podatki itd. (Azberger in dr., 2004), s čimer podatki vse bolj pridobivajo pomen. Povečan pomen podatkov se kaže pri uredniških politikah pomembnejših znanstvenih revij z različnih področij, ki postavljajo zahteve, da se ob znanstvenem članku objavijo tudi podatki, ki so bili podlaga za objavljeno znanstveno delo. V mnogih najrazvitejših državah¹³⁸ pripravljajo in uvajajo sisteme odprtega dostopa do podatkov, kar gre pogosto v smeri usklajevanja različnih storitev posameznih disciplin in iskanja koristi iz združevanja storitev.

Večina razvitih držav se spopada s problemom, kako posamezne dobro delujoče storitve na določenih področjih, npr. v Veliki Britaniji na področju podatkov z Natural Environment Research Council (NERC) ali Economic and Social Research Council (ESRC), razširiti na druga področja in uskladiti raven storitev ter možnosti čezdisciplinarnih uporab podatkov¹³⁹. Medtem ko je v Sloveniji problem velika razpršenost in nepovezanost obstoječih podatkovnih storitev, kjer je večina podatkov še vedno shranjena v okviru projekta, kjer so nastali, ter je zelo malo znotrajdisciplinarnega povezovanja kaj šele meddisciplinarno, kadar gre za drugo rabo in dostop do podatkov. V splošnem to pomeni zelo nizko razvitost načina ravnanja s podatki in nizko razvitost podatkovnih storitev. Posledice takšnega stanja so uničevanje podatkov zaradi neustrezne hrambe in pomanjkljive dokumentacije, posledično je podatke težko najti in pridobiti za drugo rabo, kar vse znižuje njihovo izkoriščenost, hkrati pa zmanjšuje zagotovila glede njihove kakovosti že v izvoru. Glede na to, da so stroški za izdelavo podatkov po navadi zelo visoki, pomeni to veliko izgubo javnih sredstev, porabljenih za raziskovanje.

V večini razvitih držav, ki jih jemljemo za zgled, so formalne zaveze zelo jasne in tudi opredeljene v ustreznih dokumentih politik na ravni države in v okviru delovanja posameznih ustanov. V Sloveniji bo urejanje ena od prioritet kot zagotovilo institucionalnega zaledja dejavnosti zagotavljanja odprtega dostopa do podatkov. Naša teza je, da bo to v Sloveniji še dokaj nerazvito področje v prvih letih vzpostavljanja delovanja že nudilo široke demonstracijske učinke, ki bodo potrjevali koristi uvajanja odprtega dostopa.

¹³⁸ ZDA, Avstralija, Velika Britanija, Nizozemska, Nemčija ...

¹³⁹ Npr. uvajanje storitev podatkovnega arhiva v agrikulturni *Rural Economy of Land Use* (<http://www.relu.ac.uk/>, 21. 1. 2014) po zgledu UKDA.

Mark Parsons (2010, 2) ugotavlja, da so podatki zgodovinsko gledano veljali za zasebno last raziskovalca ali za nacionalno korist, ki mora biti zaščitena ali jo je mogoče izkoriščati kot proizvod. V ZDA in nekaterih drugih državah se tak pogled vse bolj spreminja. Tako družbi kot znanosti najbolj koristi, če so podatki skupno javno dobro. V znanstveni skupnosti je vse večji poudarek na odprtih podatkih. Na več ravneh in na različnih področjih je mogoče opaziti splošen trend k večji odprtosti podatkov in k manj omejitvam. Nekateri so prepričani, da je prosti dostop do podatkov temelj za udeležbo v družbi znanja in da bi odprti dostop moral biti privzeto pravilo in ne izjema (Parsons 2010, 1). Da ne gre le za prepričanja osamljenih posameznikov ali znanstvenih skupin, dokazujejo mednarodne iniciative in projekti.

Mednarodni pregled začinjamo s pobudami Evropske komisije (v nadaljevanju EK), ki je v zadnjih letih zelo aktivna, kar zadeva odprti dostop na področju raziskovanja, hkrati pa države članice neposredno spodbuja k skupnim in usklajenim aktivnostim. Očitno je, da zahteve EU izhajajo iz rekapitulacije strokovnih spoznanj in pristopov vodilnih držav na področju urejanja razmer pri dostopu do raziskovalnih podatkov, tako da je nadvse primerno te pobude čim resneje upoštevati. V pregled vključujemo različne načelne usmeritve raznih uglednih mednarodnih strokovnih združenj, tako da se potrjuje vtis, da so vprašanja dostopa do raziskovalnih podatkov dejansko v ospredju pozornosti tako združenj financerjev znanosti in oblikovalcev politik, predstavnikov znanstvene skupnosti kot tudi izvajalcev storitev, povezanih s podatki. Dokaze o razvejanosti in intenzivnosti dejavnosti v zadnjih letih bomo podkrepili s pregledom različnih strokovnih organizacij, projektov in infrastrukturnih omrežij, posvečenih delu s podatki, katerih rezultat je množica poročil, orodij in delujočih storitev – vseh s ciljem, da bi pregled lahko koristil pri vzpostavljanju in nadgradnji sistema odprtega dostopa v posamezni državi tako, da bo ta povezan z mednarodnimi okoljem in da bo zagotavljal vzdržnost. Pregled zaključujemo z izbranimi zgledi posameznih držav pri načrtovanju in izvajanju aktivnosti, iz katerih je razvidno, da je skrb za vzpostavitev in razvoj raziskovalne podatkovne infrastrukture deležna tako finančne podpore za vzdrževanje obstoječih in razvoj novih storitev kot podpore v obliki strategij in politik ter promocije odprtega dostopa.

Zahteve specifikacije projekta Odprti podatki iz razpisa se, kot omenjeno, sklicujejo na priporočilo Odbora OECD za znanstveno in tehnološko politiko iz leta 2006,¹⁴⁰ ki je namenjen nacionalnim vladam. Problemski okvir projekta torej predstavlja dejstvo, da se je Slovenija z vstopom v OECD zavezala k izpolnjevanju določil omenjenih Načel. Pobuda je izjemno pomembna prav zato, ker za razliko od znanih pobud širjenja odprtega dostopa do znanstvenih publikacij na splošno¹⁴¹ v ospredje postavlja prav specifično raziskovalne podatke za znanstveno rabo. Dogajanje na področju urejanja problematike dostopa do raziskovalnih podatkov je v zadnjem času zares intenzivno, saj poleg posameznih držav iniciativo prevzemajo tudi razni mednarodni forumi, vključno z EK,¹⁴² kar bomo podrobneje predstavili v nadaljevanju.

Kot že rečeno, so Načela OECD lahko odlična strateška podlaga za pripravo nacionalnih načrtov, vendar jih je pri tem potrebno izpeljati tako, da bo upoštevana specifična situacija v Sloveniji. Pomembna je tudi prilagoditev posameznim raziskovalnim področjem, o čemer izrecno govori načelo o fleksibilnosti:

¹⁴⁰ Recommendation of the Council concerning Access to Research Data from Public Funding, 14. december 2006, C (2006)184.

¹⁴¹ Npr. Berlinska pobuda, Budimpeštanska deklaracija ...

¹⁴² Glej: <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301> in http://cordis.europa.eu/fp7/ict/e-infrastructure/publications_en.html, 7. 8. 2012.

Fleksibilnost zahteva, da upoštevamo hitre in pogosto nepredvidljive spremembe v informacijski tehnologiji, značilnosti posameznih raziskovalnih področij in raznolikost raziskovalnih sistemov: pravnih sistemov in kulture držav članic. Pri oblikovanju dogovorov o dostopu do podatkov, pri razvoju politik za spodbujanje odprtega dostopa ter pri implementaciji teh načel in vodil je treba upoštevati pomen nacionalnih, družbenih, ekonomskih in institucionalnih posebnosti. (OECD 2007, 15).

Kakor ugotavljajo tudi avtorji Načel (Arzberger in dr. 2004), je eno nerazdelnih vprašanj, kako uvajati rešitve iz razvitih držav v manj razvitih državah s specifičnimi institucionalnimi okviri.

POBUDA EVROPSKE KOMISIJE

Evropska komisija je zelo dejavna na področju skrbi za raziskovalne podatke. Objavila je več dokumentov, v katerih so jasno izraženi pogledi na razvoj področja, izvaja pobude za usmerjanje aktivnosti držav članic in podpira razne projekte¹⁴³ za gradnjo medpodročno povezanih podatkovnih storitev. Prelomnico je prinesel dokument z naslovom [Communication on Scientific Information in the Digital Age: Access, Dissemination and Preservation](#) (COM 2007, 56),¹⁴⁴ na podlagi katerega je Svet EU izdal priporočila z naslovom [Council Conclusions on scientific information in the digital age: access, dissemination and preservation](#).¹⁴⁵ V omenjenih dokumentih so v sklopu znanstvenih informacij, za katere se predpostavlja odprti dostop, posebej izpostavljeni raziskovalni podatki. Časovno in s cilji so opredeljene priporočene aktivnosti tako na ravni EK kot držav članic. EK se med drugim zavezuje k nadaljnji podpori različnih projektov s področij izgradnje podatkovne infrastrukture in storitev digitalne hrambe¹⁴⁶, k organizaciji dialoga med vključenimi deležniki tako na nacionalni ravni kot na ravni EU¹⁴⁷ in da bo zahtevala odprti dostop do rezultatov raziskav, vključno z raziskovalnimi podatki, ustvarjenih v okviru projektov, podprtih s strani EU. Države članice so bile med drugim dolžne do leta 2008 pripraviti jasne politike glede dostopa do znanstvenih informacij, izhajajočih iz raziskav, financiranih z javnimi sredstvi, in opredeliti s tem povezane proračunske postavke. Vse s ciljem, da se uporabnikom zagotovi brezplačen dostop do znanstvenih informacij preko interneta, ter ob upoštevanju ekonomsko vzdržnih pristopov.

Sledilo je poročilo z naslovom [National open access and preservation policies in Europe](#),¹⁴⁸ ki povzema uresničevanje priporočil državam. Avtorji ugotavljajo, da je na obravnavanem področju prišlo do opaznega napredka. Posamezne države so pobude za izboljšanje na področju odprtega dostopa vključile v strategije razvoja znanosti. Med drugimi je omenjena Slovenija, ki je naročila pripravo pričujočega akcijskega načrta vzpostavitve sistema odprtega dostopa do raziskovalnih podatkov in ga vključila v strategijo Nacionalna raziskovalna in inovacijska strategija Slovenije 2011–2020 (RISS)¹⁴⁹. Vzpostavljeno je bilo pričakovanje, da se bodo na tej podlagi v naslednjih letih odvijali intenzivni premiki s konkretnimi rezultati, kar je bilo kot dober zgled predstavljeno v nadaljevanju pri pregledu nacionalnih situacij.

¹⁴³ *Navigation path: European Commission>Research & Innovation>Science in Society>Policy Initiatives>Open Access>Relevant Eu-Funded Projects.* <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1302&lang=1>, 11. 12. 2013.

¹⁴⁴ Glej: http://ec.europa.eu/research/science-society/document_library/pdf_06/communication-022007_en.pdf, 12. 2. 2014.

¹⁴⁵ Glej: <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301>, 14. 8. 2013.

¹⁴⁶ Primerjaj razpis iz zadnjega sklica:

http://www.mizs.gov.si/si/obzorje2020/tematski_sklopi_in_delovni_programi/european_and_research_infrastructures/, 11. 12. 2013.

¹⁴⁷ Eden od predlogov za krepitev nacionalnih strategij in struktur za dostop in diseminacijo znanstvenih informacij: "V razpravi o znanstvenih informacijah povezati glavne deležnike (znanstveniki, financerji, knjižnice, znanstveni založniki)" (dokument *Conclusions on scientific information in the digital age: access, dissemination and preservation* 2007, 5).

¹⁴⁸ Glej: http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-report-2011_en.pdf, 14. 8. 2013.

¹⁴⁹ Za več o tem glej poročilo *National open access and preservation policies in Europe* 2011, str. 17. Poročilo je zanimivo, ker predstavlja izhodiščno točko za oceno stanja in napredka države, kakršna je Slovenija, kjer je celovit razvoj podatkovnih storitev še v povojih. V njem so namreč nakazane poti, po katerih korakajo najnaprednejše države, ki so bile v takratnem obdobju večinoma v fazi dokončevanja strateških premislekov.

Serija dokumentov in pobud na ravni EU se zaključuje s [Priporočilom Komisije z dne 17. 7. 2012 o dostopu do znanstvenih informacij in njihovem arhiviranju](#).¹⁵⁰ Dokument se glede doseganja odprtega dostopa do raziskovalnih podatkov (in publikacij) in presoje trenutne situacije sklicuje na splošne usmeritve, zapisane v strategijah do leta 2020, to sta [Digital Agenda for Europe](#)¹⁵¹ in [Innovation Union](#).¹⁵² V spremnem dokumentu [Communication Towards better access to scientific information: Boosting the benefits of public investments in research](#)¹⁵³ so izoblikovana nadaljnja priporočila državam članicam za izvajanje dejavnosti, s katerimi se bo izboljšal dostop in hramba znanstvenih informacij.

Ti dokumenti so za Slovenijo pomembni zaradi poudarjene usmeritve, da se dejavnost na področju dostopnosti znanstvenih informacij krepi koordinirano z drugimi državami. Zaostajanje kake od držav bo postalo bolj vidno. Izpostavljajo naslednja načela, k doseganju katerih morajo biti usmerjene politike odprtega dostopa do raziskovalnih podatkov posameznih držav:

- čim zgodnejši in brezplačen dostop, ki omogoča drugo rabo;
- politike obsegajo vse javno financirane raziskave, kar podkrepijo z znanimi argumenti – očiten je prihranek zaradi nepodvajanja in lažjega dostopa, cilji pa so skladni tudi z zahtevami po večji dostopnosti znanstvenih informacij, kamor sodi tudi izboljšana kakovost in dostopnost raziskovalnih podatkov ter s tem večja preglednost znanstvenih ugotovitev;
- ureditev stanja na področju zaupanja vredne digitalne hrambe znanstvenih informacij in podatkov.

Z zgoraj naštetimi načeli so povezane določene obveznosti držav, izpolnjevanje katerih komisija tudi stalno preverja, zaradi tega morajo države jasno opredeliti strategije in politike na področju razširjanja podatkov in odprtega dostopa do raziskovalnih podatkov ter z njimi povezane postavke v proračunu. Med zahtevami so: ureditev dostopa do raziskovalnih podatkov na način, da jih je mogoče brez težav najti, da se nahajajo na stalnem mestu, tako da jih je mogoče povezovati z drugimi podatki in z besedilnimi objavami v revijah, ter da vsebujejo dodatne informacije (metapodatke), ki omogočajo ustrezno ovrednotenje in smiselno uporabo. Države naj predvidijo tudi pomoč in podporo ustvarjalcem podatkov pri uresničevanju zahtev politik, nagrade za vložen napor pri pripravi raziskovalnih podatkov za dostop ter spodbude za izobraževanje novega strokovnega poklica za delo s podatki (npr. podatkovni znanstvenik, podatkovni knjižničar).

Zgoraj predstavljene usmeritve EK so dobile podporo tudi februarja 2013 na zasedanju Evropskega sveta v Bruslju, ki je potekalo na temo [The Competitiveness \(Internal Market, Industry, Research and Space\)](#)¹⁵⁴ in na katerem so razpravljali o odprtem dostopu do znanstvenih informacij, ki izvirajo iz javno financiranih raziskovalnih projektov.¹⁵⁵ Spregledati ne gre niti

¹⁵⁰ V angleščini: *Recommendation on access to and preservation of scientific information*, 2012, http://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf, 14. 8. 2013.

¹⁵¹ Glej: <http://ec.europa.eu/digital-agenda/>, 26. 11. 2013.

¹⁵² Glej: http://ec.europa.eu/research/innovation-union/index_en.cfm, 26. 11. 2013.

¹⁵³ Glej: http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf, 14. 8. 2013.

¹⁵⁴ Glej: http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/intm/135548.pdf, 11. 2. 2014.

¹⁵⁵ Glej: http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/intm/135548.pdf, 22. 1. 2014.

ugodnega mnenja – »*favourable opinion*« iz januarja 2013, ki ga je podal Evropski ekonomsko-socialni odbor.¹⁵⁶

EK je svojo pobudo podkrepila z izidom pogosto citiranega poročila z naslovom *Riding the wave* (Evropska komisija 2010). Poročilo je pomembno, ker z avtoriteto skupine uglednih avtorjev na poljuden način predstavi že sedaj obstoječo stvarnost množine in kompleksnosti nastajajočih podatkov tako v znanosti kot izven nje ter njihov premalo izkoriščen potencial. O tem poročilu so tekle številne razprave na različnih posvetovalnih in odločevalskih forumih, s čimer je bila zagotovljena širša podpora za okrepljeno dejavnost na tem področju. Poročilo in njegova vizija nemotene dostopa do podatkov, ki so med sabo povezani in vključeni v celovito infrastrukturo na svetovni ravni, sta denimo tudi podlaga pred kratkim izdanega akcijskega načrta štirih držav, vključenih v skupino Knowledge exchange¹⁵⁷ z naslovom *A Surfboard for Riding the wave*, ki ga lahko beremo kot takojšen odziv na pobudo (Graaf 2011; več o slednjem v nadaljevanju). Pobude in zahteve EK so s tem pridobile visoko stopnjo legitimnosti, iz česar izhaja dodatna dolžnost držav članic, da sledijo skupaj zastavljenim načelom.

Med evropskimi pobudami je potrebno omeniti še idejo t. i. »*Fifth Freedom*« – *free movement of knowledge*, ki raziskovalne podatke vidi kot integralni del evropskega raziskovalnega okolja. Podatkovna infrastruktura predstavlja osrednjo stično točko mednarodne raziskovalne infrastrukture iz okvira European Strategy Forum on Research Infrastructures ali na kratko ESFRI¹⁵⁸ (primerjaj Graaf in Waaijers 2011, 8–9) Poudarek na izgradnji inovativne raziskovalne infrastrukture s podlago odprto dostopnih raziskovalnih podatkov je vključen v akcijski okvir tekočega sklica evropskih raziskovalnih projektov, predstavljen skozi stebre zaupanja in digitalne hrambe ter tokov ustvarjanja in uporabe podatkov ter z njimi povezanih storitev znotraj področnih skupnosti in širše ([A Framework for Action](#)).¹⁵⁹

¹⁵⁶ *Towards better access to scientific information: Boosting the benefits of public investments in research* - COM(2012) 401 final (Evropska komisija 2012).

¹⁵⁷ Pri tej koordinacijski pobudi sodelujejo Velika Britanija, Danska, Nizozemska in Nemčija, od leta 2013 tudi Finska.

¹⁵⁸ Do sedaj so izšli trije dokumenti na temo *Roadmap for Research Infrastructures*, in sicer leta 2006, 2008 in 2010, naslednji dokument predvidoma izide leta 2015. Vir:

http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri-roadmap, 7. 8. 2013.

¹⁵⁹ Glej: http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/framework-for-action-in-h2020_en.pdf, 7. 8. 2013.

POBUDE MEDNARODNIH STROKOVNIH ZDRUŽENJ IN AKADEMSKIH POVEZAV

Različna strokovna in interesna združenja, ki delujejo na akademskem področju, opozarjajo na probleme pri ravnanju z raziskovalnimi podatki in na zahteve odprtega dostopa, s čimer izražajo svoj interes in željo po sodelovanju. Ne le da so združenja objavila vrsto načelnih izjav in usmeritev, njihovi predstavniki tudi sodelujejo v različnih javnih razpravah in sodelujejo pri projektih, ki so namenjeni uresničevanju sprejetih zavez.

League of European Research Universities

Katrien Maes, predstavnica uglednega združenja League of European Research Universities (LERU)¹⁶⁰ je aprila 2013 na delavnici o problematiki bibliografskega vpliva raziskovalnih podatkov v okviru skupine KE¹⁶¹ izpostavila pomen in vlogo univerz pri vprašanju dostopa do raziskovalnih podatkov.¹⁶² Po njenem prepričanju morajo univerze razviti lastne strategije in akcijske načrte na področju ravnanja z raziskovalnimi podatki, spodbujati deljenje podatkov, zagotavljati pogoje za pripravo načrtov ravnanja z raziskovalnimi podatki, ki so obvezni v vse več državah, zagotavljati usposabljanje raziskovalcev in tistih, ki opravljajo storitve, povezane s podatki, ter delo s podatki upoštevati kot pomembno pri napredovanju. Razvijati bi morale tudi tehnične in vsebinske podporne službe, katerih namen bo delo na raziskovalnih podatkih. Če so podatki pod režimom omejenega dostopa, naj bodo na voljo vsaj metapodatki, v okviru katerih so opisani razlogi za omejitve in strožji pogoji dostopa. Te in druge usmeritve evropskih univerz so predstavniki združenja zajeli v dve strani dolgem dokumentu z naslovom [Open Research Data](#).¹⁶³

Pod okriljem LERU so nastale tudi smernice ([Roadmap](#))¹⁶⁴ za rektorje članic.¹⁶⁵ Z dokumentom so pokrili naslednja področja: obseg stroškov in kako se spopasti z njimi, infrastruktura in potencial za deljenje infrastrukture med evropskimi univerzami, človeški viri in spodbude za deljenje podatkov, ki lahko pozitivno vplivajo na raziskovalčevo napredovanje v karieri, kulturne spremembe in odpravljanje ovir, ki stojijo na poti odprtega dostopa. Pomembno je usklajeno delovanje na vseh področjih. Kljub oviram, ki jih bo potrebno premisliti in upoštevati, pa je uradno stališče LERU, da si je potrebno prizadevati za to, da so odprti podatki pravilo in ne več izjema, kakor je sedaj ([Ayris, 2013, LIBER and LERU participate in EC consultation on Open Research Data](#)).¹⁶⁶ Smernice kot najbolj primeren pristop vključujejo nalogo oblikovanja

¹⁶⁰ Ena od temeljnih nalog LERU od ustanovitve leta 2002 je promocija temeljnih raziskav evropskih univerz s ciljem, da lahko pomembno vplivajo na družbeni razvoj. Danes LERU, ki ga sestavlja 21 raziskovalno močnih univerz, igra pomembno vlogo pri oblikovanju raziskovalnih politik v EU. Glej: <http://www.leru.org/index.php/public/home/>, 7. 8. 2013.

¹⁶¹ Delavnica: *Making Data Count: Research data availability and research assessment*, 11. –12. april 2013, Berlin.

¹⁶² Povzeto po zapisu Alexa Balla, enega vodilnih strokovnjakov za vprašanja dostopa do raziskovalnih podatkov, <http://www.dcc.ac.uk/blog/making-data-count>, 7. 8. 2013.

¹⁶³ Glej: <http://www.leru.org/files/general/Open%20Access%20to%20Research%20Data-FINALdocx.pdf>, 7. 8. 2013.

¹⁶⁴ Glej: http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf, 11. 2. 2014.

¹⁶⁵ *LERU Roadmap for research data*,

http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf, 22. 1. 2014.

¹⁶⁶ Glej: <http://www.libereurope.eu/news/liber-and-leru-participate-in-ec-consultation-on-open-research-data>, 7. 8. 2013.

strategij, politik in strokovnih skupin za delovanje na posameznih univerzah, pri tem pa je izpostavljena ozaveščenost in podpora s strani vodstva. Dokument vsebuje tudi posamezne zglede za ilustracijo pristopov in kot celota povzema vsa pomembna aktualna področja skrbi za dostop do raziskovalnih podatkov, kot so vprašanja vrednotenja, stroškov vzpostavljene infrastrukture, izobraževanja, citiranja itd.

Opazen je premik ravni odgovornosti za uresničevanje zavez na ustanove. Raziskovalci niso prepuščeni sami sebi, pač pa je odgovornost za ustrezno lokalno podporo dodeljena ustanovam. Napredne univerze po Evropi že upoštevajo tovrstne usmeritve in kot eden od sestavnih delov z drugimi sodelujejo pri vzpostavljanju odprtega dostopa: s financerji na področju politik sodelujejo glede obveznosti in odgovornosti, s področnimi izvajalci podatkovnih storitev pa skušajo okrepiti izmenjave dobrih praks in sodelujejo pri ozaveščanju raziskovalne skupnosti o pomenu tega vidika raziskovanja, k čemur se bomo še vrnili ob pregledu dogajanj v posameznih državah. V LERU se poleg na že omenjene pobude EK sklicujejo tudi na priporočila združenja raziskovalnih knjižnic, ki jih prikazujemo v nadaljevanju.

Pobude združenj raziskovalnih knjižnic

Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries ([LIBER](#))¹⁶⁷ je še ena od uglednih strokovnih organizacij na ravni Evrope, ki je izdala svoja priporočila glede nalog pri organiziranju dela s podatki z naslovom [Ten recommendations for libraries to get started with research data management](#).¹⁶⁸

Organizacija zastopa interese evropskih raziskovalnih knjižnic, njihovih univerz in raziskovalcev. Na posvetu EK *Odpiranje raziskovalnih podatkov* (julij 2013) so predstavili poglede organizacije, da bodo knjižnice morale svojo vlogo prenoviti tako, da bo vanjo vključena tudi skrb za raziskovalne podatke (ustvarjanje, deljenje in druga raba). Novo vlogo knjižničarjev vidijo predvsem v podpori pri pripravi načrtov za ravnanje z raziskovalnimi podatki ter pri vzpostavljanju tezavrov in ontologij za opisovanje podatkov (Ayriz 2013).¹⁶⁹

Ameriško združenje knjižničarjev Association of research Libraries ([ARL](#))¹⁷⁰ se v poročilu [Agenda for Developing e-science in research libraries – ARL](#) (Agenda 2007)¹⁷¹ odločno zavzema za to, da bi knjižničarji postali pomemben člen oblikovanja strategij na področju e-znanosti in nasploh na področju dostopa do podatkov. Knjižnice bi s svojim znanjem in izkušnjami že danes lahko služile kot infrastrukturna podpora, pravijo pri ARL, vendar pa bi v prihodnosti morali razviti orodja in veščine, specifične za področje ustvarjanja podatkov, dostopa in druge rabe. Zavzemajo se, da bi bili vse bolj udeleženi v postopkih nastajanja podatkov, saj bi tako laže in bolje prispevali k izvajanju programa storitev podatkovne podpore. Konkretni cilj je razviti znanja s področja e-znanosti, oblikovati nov model vlog in storitev ter postati aktivni udeleženci v procesu zasnove in razvoja raziskovalne infrastrukture, ki bo podpirala celoten življenjski krog raziskovanja (Agenda 2007, 16–17; glej tudi Lossau 2012).

¹⁶⁷ LIBER vključuje več kot 400 nacionalnih, univerzitetnih in drugih knjižnic iz več kot 40 držav, na seznamu članic sta tudi Narodna in univerzitetna knjižnica (NUK) in Univerzitetna knjižnica Maribor (UKM), <http://www.libereurope.eu/>, 7. 8. 2013.

¹⁶⁸ Glej:

<http://www.libereurope.eu/sites/default/files/The%20research%20data%20group%202012%20v7%20final.pdf>, 22. 1. 2014.

¹⁶⁹ Glej: <http://www.libereurope.eu/news/liber-and-leru-participate-in-ec-consultation-on-open-research-data>, 7. 8. 2013.

¹⁷⁰ Glej: <http://www.arl.org/index.php>, 7. 8. 2013.

¹⁷¹ Glej: http://www.arl.org/bm~doc/arl_escience_final.pdf, 26. 11. 2013.

Velik interes raziskovalnih knjižnic za delo na področju raziskovalnih podatkov kaže tudi svetovno združenje Libraries in the digital age ([LIDA](#)).¹⁷² Na raznih svetovnih kongresih je ta tema posebej izpostavljena. Tako je bila denimo tej temi posvečena sekcija na zadnjem posvetu združenja LIDA in tudi vabljeno predavanje Liz Lyon z naslovom *Incremental Change or Revolution? Libraries and the Informatics Transform*, v katerem avtorica predlaga možne vloge in odgovornosti zaposlenih v raziskovalnih knjižnicah pri odprtem dostopu do raziskovalnih podatkov.¹⁷³ Z njimi programsko sodeluje posebna delovna skupina združenja [SPARC](#) za odprte podatke.¹⁷⁴

Javni financerji znanosti (raziskovalne agencije, fundacije, sveti)

Pri vzpostavljanju sistema odprtega dostopa bodo pomembno vlogo igrali financerji in s financiranjem povezana telesa. European Heads of Research Councils (EUROHORCs), neodvisno združenje javnih financerjev znanosti, je skupaj z European Science Foundation (ESF) predstavilo [načrte razvoja znanosti v Evropi](#),¹⁷⁵ ki kot pomembno sestavino vključujejo urejanje dostopa do raziskovalnih podatkov. Združenje je prenehalo delovati leta 2011, večina njihovih aktivnosti je prešla na Science Europe, kjer je pobudo prevzela posebna delovna skupina za raziskovalne podatke.¹⁷⁶ Rezultate njihovega dela še pričakujemo, bo pa to glede na sestavo prav gotovo zelo pomembna spodbuda in obveznost za oblikovanje nacionalnih in področnih načrtov delovanja odpiranja podatkov.

Omeniti velja tudi določilo o obveznem odprtem dostopu vseh objav in raziskovalnih podatkov iz sofinanciranih raziskav, ki ga je decembra 2007 izdal Evropski raziskovalni svet pod naslovom [ERC Scientific Council Guidelines for Open Access](#).¹⁷⁷ V dokumentu je kot posebna točka izpostavljen pomen dostopa do najbolj osnovnih raziskovalnih podatkov, ki omogoča neodvisno verifikacijo rezultatov in kar je še pomembneje, omogoča varno hrambo, nove analize in uporabo podatkov.

Na tem področju posebej odmeva odločitev [Ameriške agencije za financiranje znanosti](#) (NSF)¹⁷⁸, ki je v svojih politikah jasno opredelila odgovornosti raziskovalcev in raziskovalnih ustanov glede ravnanja z raziskovalnimi podatki. Od raziskovalcev pričakuje, da bodo primarne podatke, vzorce, zbirke in drugo podporno gradivo, ki je bilo ustvarjeno ali zbrano v okviru dela, financiranega s strani NSF, delili z drugimi v okviru razumnih stroškov in v razumnem času, kar opredeljujejo na naslednji način: »Vloge, oddane ali veljavne 18. januarja 2011 in po tem datumu, morajo vključevati dopolnilni dokument, ki ne obsega več kot dve strani in je označen z naslovom 'Načrt ravnanja s podatki'. Ta dopolnilni dokument naj opiše, kako bo vloga zadostila politikam NSF glede diseminacije in deljenja raziskovalnih podatkov.«¹⁷⁹ V nadaljevanju bomo pri pregledu dogajanj po drugih državah pokazali, da je tovrstna usmeritev vsesplošna.¹⁸⁰

¹⁷² Glej: http://ozk.unizd.hr/lida/2012/theme_two/, 26. 11. 2013.

¹⁷³ Glej: <http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/liz-lyon-lida-keynote-croatia.pdf>, 26. 11. 2013.

¹⁷⁴ Glej: <http://sparc.arl.org/issues/open-data>, 22. 1. 2014.

¹⁷⁵ Glej: <http://www.kowi.de/Portaldata/2/Resources/fp/EUROHORCs-ESF-RoadMap.pdf>, 11.2.2014.

¹⁷⁶ European Science Foundation in EUROHORC. 2009. *EUROHORCs and ESF Vision on a Globally Competitive ERA and their Road Map for actions*. Dostopno prek: <http://www.kowi.de/Portaldata/2/Resources/fp/EUROHORCs-ESF-RoadMap.pdf>, 22. 1. 2014. Inicijativo je prevzela delovna skupina za *Research Data* na Science Europe: <http://www.scienceeurope.org/policy/working-groups>, 22. 1. 2014.

¹⁷⁷ Glej: <http://erc.europa.eu/documents/erc-scientific-council-guidelines-open-access>, 10. 2. 2014.

¹⁷⁸ Glej: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>, 22. 1. 2014

¹⁷⁹ *Dissemination and Sharing of Research Results*, <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>, 26. 11. 2013.

¹⁸⁰ Primerjaj pravkar izdano študijo *European Landscape Study of Research Data Management*,

Pobude akademij znanosti

European Federation of Academies of Science and Humanities (ALLEA), katere članica je tudi Slovenska akademija znanosti in umetnosti (SAZU), je objavila deklaracijo [Open Science for the 21st Century](#),¹⁸¹ s katero svoje članice – posredno seveda tudi člane in članice akademij posameznih držav – zavezuje k podpori pri širjenju razprav in iskanju usklajenih rešitev, skladnih s pobudo EK na področju dostopa do raziskovalnih podatkov.

Poleg tega sta ALLEA in Europe Science Foundation marca 2011 skupaj objavili *The European Code of Conduct for Research Integrity*. Kodeks eksplicitno izpostavlja naslednja načela, h katerim zavezuje raziskovalce, javne in zasebne raziskovalne organizacije, univerze in financerje:

- poštenost pri komuniciranju,
- zanesljivost pri izvajanju raziskave,
- objektivnost,
- nepristranskost in neodvisnost,
- odprtost in dostopnost,
- dolžnost biti skrben,
- poštenost pri navajanju referenc in točkovanju ter
- odgovornost za prihodnje znanstvenike in raziskovalce.

Vrednote in etične norme znanstvenega dela so v sodobnem času v veliki meri podprte prav z zagotavljanjem odprtega dostopa do podatkov. Med dobrimi praksami dokument posebej opozarja na delo s podatki: »Vsi primarni in sekundarni podatki se morajo ohranjati v varni in dostopni obliki, dokumentirani in arhivirani za daljše obdobje. Na razpolago morajo biti kolegom.« (European Code of Conduct for Research Integrity 2011, 6)

Akademije igrajo pomembno vlogo tudi pri usklajevanju pristopov med znanstvenimi področji, pri dopolnjevanju etičnih kodeksov, pri prilagajanju politik revij, kjer so ugledni člani pogosto v vlogi urednikov ali izdajateljev, ter zlasti pri vzpostavljanju vzdržno delujočih področnih podatkovnih središč, postavljenih na strokovnih temeljih. Skupno sprejete usmeritve na poti k bolj odprti znanosti bodo gotovo tudi stvar razprave in dela Slovenske akademije znanosti in umetnosti.

Posebej zgleden primer angažiranja pri odprtih podatkih je ameriška akademija National Academy of Sciences (NAS), v okviru katere deluje poseben odbor Board on Research Data and Information (BRDI).¹⁸² Skupaj z delovno skupino CODATA ICSTI so izvedli številne projekte, eden takih je aktualni projekt Standardi in prakse citiranja podatkov (*Data Citation Standards and Practices*), organizirali so mednarodni simpozij in delavnico z naslovom [Developing Data Attribution and Citation Practices and Standards](#)¹⁸³ ter izdali več poročil.¹⁸⁴ Trenutno poteka tudi

<http://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM%20landscape%20report%20vs1%204-14.08.13.pdf>, 11. 12. 2013.

¹⁸¹ Glej: <http://www.allea.org/Pages/ALL/31/397.bGFuZz1FTkc.html>, 25. 11. 2013.

¹⁸² Glej: <http://sites.nationalacademies.org/PGA/brdi/index.htm>, 26. 11. 2013.

¹⁸³ Glej: http://sites.nationalacademies.org/PGA/brdi/PGA_064019, 26. 11. 2013.

¹⁸⁴ Npr.: "Developing Data Attribution and Citation Practices and Standards", dostopno prek:

http://www.nap.edu/catalog.php?record_id=13564, 28. 1. 2014, in "Out of Cite, Out of Mind: Citation of Data: The Current State of Practice, Policy, and Technology", dostopno prek:

https://www.istage.jst.go.jp/article/dsj/12/0/12_OSOM13-043/article, 28. 1. 2014.

izbor projektov na temo "[Raba podatkov za javno dobro](#)",¹⁸⁵ pri izboru ocenjujejo nove ideje, orodja, postopke, modele in rezultate.

Nasploh so najuglednejši predstavniki znanstvenih disciplin v različnih državah prvi pri pobudah za ureditev področja dostopa do podatkov, denimo v Veliki Britaniji, kjer je The Royal Society izdala zelo odmevno poročilo [The Science as an open enterprise](#).¹⁸⁶ Poročilo jasno izraža potrebo po izrabi velikih količin podatkov s pomočjo modernih tehnologij in na način, ki bo omogočal odprtost. Poročilo opredeljuje naloge, ki jih bo potrebno izvesti, da bi lahko izkoristili obstoječi potencial. Med temi posebej izpostavlja naslednjih šest področij:

- Znanstveniki morajo biti bolj odprti drug do drugega, pa tudi do javnosti in medijev.
- Bolje je treba ovrednotiti delo, povezano s podatki: zbiranje, analiziranje in komuniciranje.
- Skupni standardi za deljenje informacij morajo postati široko uporabni.
- Objava podatkov v formatu, ki omogoča drugo rabo in podpira ugotovitve, mora biti obvezna.
- Potrebno je zagotoviti več strokovnjakov za področje ravnanja in podpore pri rabi digitalnih podatkov.
- Potrebno je razviti nova orodja oz. programske opreme za analizo naraščajočega obsega podatkov.

Združenje založnikov

International Association of Scientific, Technical and Medical Publishers ([STM](#))¹⁸⁷ združuje založnike z vsega sveta in ima 120 članov iz 21 držav, med njimi so organizacije znanja, univerzitetne založbe, zasebne družbe ... V okviru združenja deluje tudi STM Research Group, ki ima dva ključna cilja: izmenjavo informacij o novih pobudah na področju integracije raziskovalnih publikacij in raziskovalnih podatkov ter razprave, ki vključujejo najboljše prakse na tem področju. Skupina se posveča vprašanjem povezovanja podatkov in publikacij, citiranju podatkov, trajnih identifikatorjev, recenziranju podatkov, hrambi podatkov in drugemu. V t. i. [Bruseljski deklaraciji](#)¹⁸⁸ je združenje med drugim zapisalo, da naj bodo raziskovalni podatki prosto dostopni vsem raziskovalcem, založniki naj spodbujajo objavo raziskovalnih podatkov in podatkovne baze, ki so del članka, naj bodo, kadar je to le mogoče, prosto dostopne drugim raziskovalcem.

¹⁸⁵ BRDI DATA AND INFORMATION CHALLENGE, *Using Data for the Public Good*, http://sites.nationalacademies.org/PGA/brdi/PGA_085255, 26. 11. 2013.

¹⁸⁶ *Science as an open enterprise, final report*, <http://royalsociety.org/policy/projects/science-public-enterprise/report/>, 26. 11. 2013.

¹⁸⁷ Glej: <http://www.stm-assoc.org/about-the-association/>, 20. 12. 2013.

¹⁸⁸ Glej: <http://www.stm-assoc.org/brussels-declaration/>, 20. 12. 2013.

STROKOVNA ZDRUŽENJA IN PROJEKTI S PODROČJA RAZVOJA PODATKOVNE INFRASTRUKTURE

SPLOŠEN PREGLED SKUPIN IN PROJEKTOV

Širšo sliko o intenzivni dejavnosti EU kaže ustanovitev [Oddelka za odprti dostop](#)¹⁸⁹ znotraj DG CONNECT, Direktorat C.3 za odličnost v znanosti, Enota za digitalno znanost v sodelovanju z [Oddelkom za podatkovno osrediščeno znanost in tehnologijo](#)¹⁹⁰ znotraj Direktorata C.1 za eInfrastrukturo. Napredno razmišljanja je razvidno iz sprejetja politik obveznosti za raziskovalne projekte v novem sklicu ter pregled razvojnih infrastrukturnih projektov, ki jih EU financira v okviru programa Obzorje 2020. Zanimivi so nekateri poudarki iz zadnje konzultacije v letu 2013. Denimo, pri izpostavljanju prioritet opozarjajo, da je t. i. malim podatkom potrebno posvetiti vsaj toliko ali celo več pozornosti kot t. i. velikim podatkom.¹⁹¹ Pri opredeljevanju, za katere tipe podatkov poskrbeti, so bila izražena mnenja, da naj se čim manj izključuje in omejuje, saj so lahko koristni tudi podatki iz eksperimentov, ki so vodili k negativnim zaključkom, in jih je kot take smiselno hraniti. Eden od komentatorjev opaža, da ni bilo čutiti posebnega pritiska v smeri podaljševanja obdobja embarga nad dostopom, izražena pa je bila skrb za ustrezno obravnavo občutljivih osebnih podatkov. V nasprotju z logiko omejevanja pod pretvezo skrbi za »varovanje« podatkov je anekdotična opazka enega od govorcev, da pacienti dajejo podatke za medicinske raziskave ne za nek ozek krog raziskovalcev, ampak za koristi celotne družbe (primerjaj Milia in dr. 2012).¹⁹² Med ključnimi ukrepi za krepitev kulture izmenjave so z vidika raziskovalca omenjali spodbude v obliki nagrajevanja in priznavanja, npr. z omogočanjem citiranja, prav tako pa z orodji in postopki, ki raziskovalcu olajšajo izročitev podatkov. Prisotni knjižničarji so potrdili svojo pripravljenost za nudenje podpore, infrastrukture in usposabljanja pri ravnanju s podatki, zlasti pri pisanju načrtov ravnanja s podatki.

Skupin in projektov, ki se ukvarjajo z izzivi raziskovalne podatkovne infrastrukture, tako tekočih kot zaključenih, je resnično veliko, zato tudi ni nenavadno, da iste skupine kot centri znanja nastopajo v množici projektov, hkrati pa je videti težnjo k boljšemu povezovanju različnih skupin in projektov, denimo medpodročno ali pa glede na funkcionalnost.

Alliance for Permanent Access (APARSEN)¹⁹³ deluje kot združenje različnih organizacij, katerega poslanstvo je razviti skupno vizijo in okvir za trajnostno organizacijsko infrastrukturo

¹⁸⁹ *Open Access to scientific information*, <http://ec.europa.eu/dgs/connect/en/content/open-access-scientific-information>, 23. 1. 2014.

¹⁹⁰ *eInfrastructures: data centric science and engineering*

<http://ec.europa.eu/dgs/connect/en/content/einfrastructures-data-centric-science-and-engineering>, 23. 1. 2014.

¹⁹¹ *European Commission Public Consultation on Open Research Data: some impressions from the day*, <http://www.libereurope.eu/blog/european-commission-public-consultation-on-open-research-data-some-impressions-from-the-day>, 14. 8. 2013.

¹⁹² *Mine, Yours, Ours? Sharing Data on Human Genetic Variation*, <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0037552>, 14. 8. 2013.

¹⁹³ Glej: <http://www.alliancepermanentaccess.org/>, 26. 11. 2013.

stalnega dostopa do znanstvenih informacij. Med njenimi cilji so: podpora razvoja vzdržne evropske infrastrukture za digitalne informacije; biti strateški partner EK in vlad držav članic, da bi tako okrepili evropske in nacionalne strategije in politike na področju digitalnih informacij v znanosti; biti platforma za deležnike iz sveta znanosti in znanstvenih informacij tako, da jim omogoča medsebojno sodelovanje in sodelovanje z drugimi organizacijami; krepiti vlogo evropskih deležnikov na globalni ravni s ciljem zagotoviti dolgoročno ohranjanje in dostop do znanstvenih rezultatov; postati globalni center odličnosti na področju digitalnega ohranjanja.¹⁹⁴

Pod vodstvom Davida Giaretta je združenje APARSEN zelo uspešno pri pridobivanju različnih projektov s strani EU. Objavili so vrsto poročil¹⁹⁵ s področij vprašanj kakovosti raziskovalnih podatkov ter povezovanja različnih storitev, kot so digitalni identifikatorji, sledljivost, shranjevanje in ocenjevanje stroškov, ki jih opravljajo podatkovna središča znotraj okvira modela [Open Archival Information System \(OAIS\)](#).¹⁹⁶

V koordinaciji APARSEN je bil izveden projekt [Permanent Access to the Records of Science in Europe](#)¹⁹⁷ oziroma na kratko *PARSE.Insight*. Cilj tega projekta je bil »osvetliti dolgoživost in ranljivost digitalnih raziskovalnih podatkov in se osredotočiti na tiste dele infrastrukture, ki so potrebni kot podpora za trajnost in razumljivost digitalnih objektov iz raziskav«. ¹⁹⁸ Projekt je potekal od 2008 do 2010, v tem času je nastalo več pomembnejših dokumentov, poročil in prikazov stanja ter načrtov za prihodnost, med temi npr.: *Science Data Infrastructure Roadmap* (junij 2010). Slednji je zanimiv zaradi poudarka na potrebnih organizacijskih, finančnih in tehnoloških ukrepih za vzdrževanje podatkovne infrastrukture za dolgotrajno hrambo po OAIS.

Opportunities for Data Exchange (ODE) je še eden od projektov, pri katerih sodeluje APARSEN skupaj z devetimi partnerji iz sedmih držav, ki pokrivajo različna področja ravnanja s podatki: ustvarjanje podatkov, raziskovalna oprema, založniki, knjižnice in podatkovna središča. Projekt ODE financira Evropska unija (PF7), njegov namen je zagotoviti sprejemanje premišljenih odločitev glede organiziranja sistemov druge rabe raziskovalnih podatkov. V okviru projekta so izvedli intervjuje z različnimi deležniki (posamezniki in skupinami) in na tej osnovi pripravili izvlečke za ustvarjalce podatkov, podatkovna središča, knjižničarje, financerje, založnike, poleg tega še več poročil, ki izpostavljajo spodbude in ovire na poti k deljenju podatkov.¹⁹⁹ Projekt je zanimiv zaradi pristopa, s katerim se na pregleden način primerja posamezne vloge in kako se te dopolnjujejo v procesu izboljševanja pogojev za deljenje podatkov: raziskovalci, podatkovna središča, izdajatelji revij, financerji znanosti in oblikovalci politik. V [zaključnem poročilu ODE](#) (26. 10. 2012)²⁰⁰ so zajeti napotki, ki jih morajo izpolnjevati posamezni nosilci vlog, če naj

¹⁹⁴ About APARSEN, <http://www.alliancepermanentaccess.org/index.php/about/>, 26. 11. 2013.

¹⁹⁵ Denimo poročilo Pampel in dr. (2012), *Report on Peer Review of Research Data in Scholarly Communication*. Dostopno prek: <http://epic.awi.de/30353/>, 23. 1. 2013; glej tudi http://www.alliancepermanentaccess.org/wp-content/uploads/2012/09/APARSEN_WS_OAT_Vienna2012_DataQuality_Pampel.pdf, 23. 1. 2013.

¹⁹⁶ *Reference Model for an Open Archival Information System*, <http://public.ccsds.org/publications/archive/650x0m2.pdf>, 23. 1. 2014. Dokument *Reference Model for an Open Archival Information System (OAIS)* je ključna opora za organiziranje sistemov dolgotrajne hrambe, in sicer na način jasne opredelitve informacijskih paketov in njihovega nadzora: potek sprejema (SIP), arhiviranja (AIP) in dostopa (DIP). Skupina za digitalno hrambo pri iniciativi za e-kulturo pripravlja slovenski prevod priročnika, kar kaže na njegov pomen.

¹⁹⁷ Glej: <http://www.parse-insight.eu/project.php>, 11. 2. 2014.

¹⁹⁸ *PARSE.Insight*, <http://www.parse-insight.eu/project.php>, 7. 8. 2013.

¹⁹⁹ Glej: <http://www.alliancepermanentaccess.org/index.php/community/current-projects/ode/outputs/>, 7. 8. 2013.

²⁰⁰ D6.1 *Summary of the Studies, Thematic Publications and Recommendations*, http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/11/ODE-WP6-DEL-0001-1_0.pdf, 23. 1. 2014.

opravljajo nalogo, ki jim je namenjena. Zelo zanimiva so tudi poročila, kjer so predstavljeni pogledi uspešnih posameznikov pri delu z raziskovalnimi podatki na različnih področjih in kjer se dotikajo tako ovir kot priložnosti za razvoj dejavnosti.

Združenju APARSEN je po strukturi podobna **Digital Preservation Coalition (DPC)**²⁰¹ — le da je bolj vezana na področje Velike Britanije. Obojim je skupno, da poudarjajo vprašanja dolgotrajnega digitalnega shranjevanja in z njim povezanega »zaupanja«, ki ga je mogoče testirati s pomočjo formalnih postopkov certificiranja procesov ter organizacijskih in upravljaljskih postopkov organizacij, ki delujejo na tem področju. Med aktualnimi dogodki DPC je bila delavnica z naslovom *Getting Started in Digital Preservation: Extra Stop in Dublin*,²⁰² ki se je odvila 1. 11. 2013. S sklopom predstavitev študij primerov in vaj so se udeleženci seznanili z rabo tehnik ocenjevanja, kriznim menedžmentom in načrtovanjem, ki pomagajo ohranjati digitalne zbirke. Obravnavali so vrsto težav, povezanih z digitalno hrambo, udeležencem so posredovali znanja za delo z digitalnimi zbirkami, predstavili so jim načrtovanje ohranjanja in pripravo osnovnega načrta ohranjanja ter kako sodelovati z drugimi na področju digitalne hrambe. Dogodek so priredili v sodelovanju z Digital Repository of Ireland, Health Sciences Library Group in Academic and Special Libraries Section of the Library Association of Ireland.

Poleg splošnih se razvija tudi cela vrsta manjših pobud, vezanih na posamezna disciplinarna področja. Primer je mednarodna pobuda za deljenje podatkov, imenovana **Opening Science to Society**,²⁰³ ki je v središče svojega poslanstva postavila širše vidike tega problema, zlasti pomen izobraževanja za prihodnost in družbene posledice (povezovanje znanosti in družbe). Locirana je v Italiji in se pretežno ukvarja z vprašanji dostopa do podatkov s področja biologije in genetike. Vzdržuje spletno mesto in organizira posamezne razprave. Septembra 2013 je potekala zanimiva delavnica z naslovom *Scientific data sharing*,²⁰⁴ na kateri so obravnavali vprašanja splošne narave, kot npr. politike in kulturo izmenjave podatkov, in področne primere dobrih praks.

Heritage portal²⁰⁵ je osrednja evropska pobuda za področje kulturne dediščine. Njen član je tudi slovenski Zavod za varstvo kulturne dediščine Slovenije. Portal je nastal leta 2011 s ciljem, da pomaga raziskovalcem z različnih področij priti do informacij s področja kulturne dediščine, da ti izmenjajo svoja spoznanja in se povežejo. Poleg področij, ki po tradiciji igrajo pomembno vlogo, kot npr. zgodovina umetnosti, arheologija in konzervatorstvo, je vse več raziskovalcev z drugih področij: informacijska tehnologija (IT), kemija, fizika in biologija ter inženirji.²⁰⁶

Pri odpiranju podatkov so vse pomembnejši tudi regijski projekti, eden takih je **HOMER**,²⁰⁷ ki teče na območju Sredozemlja. Glavni cilj projekta je prispevati k odpiranju potenciala informacij javnega sektorja, pri čemer imajo prednost naslednja področja: kmetijstvo, turizem, okolje, energija in kultura. Več aktivnosti tega projekta je potekalo tudi v Sloveniji, pod okriljem Geodetskega inštituta.

²⁰¹ Glej: <http://www.dpconline.org/>, 26. 11. 2013.

²⁰² Getting Started in Digital Preservation: Extra Stop in Dublin, <http://www.dpconline.org/events/details/66-getting-started-in-digital-preservation-extra-stop-in-dublin?xref=72>, 26. 11. 2013.

²⁰³ Glej: <https://sites.google.com/site/openingsciencetosociety/>, 11. 2. 2014.

²⁰⁴ Glej: <https://sites.google.com/site/openingsciencetosociety/news/anagni-programme>, 7. 8. 2013.

²⁰⁵ Glej: <http://www.heritageportal.eu/>, 11. 2. 2014.

²⁰⁶ A data infrastructure for digital cultural heritage: characteristics, requirements and priority services; <http://www.heritageportal.eu/About-Us/Who-we-are/#sthash.m00aptUG.dpuf>, 23. 12. 2013.

²⁰⁷ Glej: <http://homerproject.eu/>, 12. 12. 2013.

Obstaja še cela vrsta projektov ali skupin, ki so zanimive po svoji sestavi in dejavnosti in zaradi naprednih pristopov pri vprašanih ravnanja z raziskovalnimi podatki. Ena od precej aktivnih in povezovalnih je bila v preteklosti evropska pobuda **e-Infrastructure Reflection Group (e-IRG)**.²⁰⁸ e-IRG sestavljajo uradni predstavniki ministrstev za znanost iz držav EU, njeno temeljno poslanstvo je tlakovati pot skupni evropski e-infrastrukturi. Glavni cilj je podpreti oblikovanje političnega, tehnološkega in administrativnega okvira za preprosto in stroškovno učinkovito skupno rabo elektronskih virov v Evropi. V skupini ima svojega predstavnika tudi Slovenija.²⁰⁹ *e-IRG White Paper* iz leta 2013 priporoča, da pri uvajanju sprememb na področju e-infrastrukture tudi odločevalci prevzamejo svoje odgovornosti, tako na ravni vlad držav članic EU kot tudi na ravni EK (*e-IRG White Paper* 2013, 5).²¹⁰ Posebej aktiven je v tej skupini Peter Wittenburg z Max Planck Institute for Psycholinguistics, ki je tudi osrednja figura enote ESFRI za jezikovne vire **CLARIN**.²¹¹

Naslednika te iniciative sta **ICORDI**²¹² s poudarkom na čezatlantskem povezovanju in drugi trenutno še aktiven projekt *European Data Infrastructure* (EUDAT),²¹³ pri katerem pilotno delujejo predstavniki nekaterih mednarodnih podatkovnih središč z več področij znanosti²¹⁴. Posebnost pristopa te skupine je izpostavljeni tehnološki vidik vprašanj sodelovanja med različnimi raziskovalnimi skupnostmi in podatkovnimi centri, se pravi iskanje možnosti povezovanja in prehodov med disciplinarnimi otopki z namenom skupne graditve nekaterih storitev, kjer za to obstajajo možnosti. V zadnjem obdobju se je težišče sodelovanja pri vprašanih podatkovne infrastrukture preneslo pod okrilje na novo ustanovljene **Research Data Alliance**.²¹⁵ ki združuje tako predstavnike ustanov kot posameznike in financirje znanosti s celega sveta.

Omenimo še obstoj **Open Knowledge Foundation (OKF)**,²¹⁶ ki je neprofitna organizacija in od svoje ustanovitve v letu 2004 promovira odprte podatke in nasploh odprte vsebine. Prizadeva si za prost in odprt dostop do gradiv, možnost razširjanja in druge rabe gradiv. Iz njihove organizacije je izšlo več orodij, npr. CKAN – sistem za ravnanje s podatki, ki z zagotavljanjem orodij za objavo, deljenje, iskanje in uporabo podatkov omogoča dostop do podatkov.²¹⁷ OKF je skupaj z **European Journalism Centre**²¹⁸ izdal *Data Journalism Handbook* – obsežen priročnik za novinarje na temo uporabe podatkov.²¹⁹ Na njihovi spletni strani deluje tudi **Open Data Commons** – sklop orodij, namenjenih pomoči pri zagotavljanju in uporabi odprtih podatkov.²²⁰

²⁰⁸ Glej: <http://www.e-irg.eu/>, 11.2.2014.

²⁰⁹ Člani, <http://www.e-irg.eu/about-e-irg/members.html>, 26. 11. 2013.

²¹⁰ Glej: http://www.e-irg.eu/images/stories/dissemination/e-irg_white_paper_2013_short_version.pdf, 23. 1. 2014.

²¹¹ Glej: <http://www.clarin.eu/>, 23. 1. 2014.

²¹² Glej: <http://www.icordi.eu/>, 23. 1. 2014.

²¹³ Glej: <http://www.eudat.eu/>, 11.2.2014.

²¹⁴ Poleg zgoraj omenjenega še EPOS, ENES, LIFEWATCH, **VPH**, **INCF**.

²¹⁵ Glej: <https://rd-alliance.org/node>, 11.2.2014.

²¹⁶ Glej: <http://okfn.org/>, 11.2.2014.

²¹⁷ Glej: <http://ckan.org/#sthash.PuIvoF7K.dpuf>, 7. 8. 2013.

²¹⁸ Glej: <http://www.ejc.net/>, 12. 2. 2014.

²¹⁹ Glej: <http://datajournalismhandbook.org/>, 7. 8. 2013.

²²⁰ Glej: <http://opendatacommons.org/>, 7. 8. 2013.

PREDLOG EVROPSKE STRATEGIJE NA PODROČJU PODATKOVNIH VIROV

Ta primer izpostavljam zato, ker se tudi na nacionalni ravni pogovarjamo o tem, katere so lahko skupne plasti in kaj bo naloga področnih podatkovnih centrov, da se prepreči podvajanje storitev in da se oblikujejo tudi centri za transverzalne podatkovne storitve.

Konzorcij **Partnership for Accessing Data in Europe** (PARADE) si je prizadeval za izgradnjo storitev ravnanja s podatki različnih raziskovalnih skupnosti. Vanj so bile vključene številne skupnosti in tudi nacionalni partnerji s področja podatkovne infrastrukture. Njihov cilj je bil povezati različne evropske podatkovne pobude in tako vzpostaviti panevropsko sodelovanje. Konzorcij je pripravil [Belo knjigo o evropski strategiji na področju podatkovnih virov](#),²²¹ v kateri predlaga razvoj trajne, multidisciplinarne panevropske podatkovne infrastrukture, temelječe na potrebah uporabniških skupnosti. Ocenili so, da je za doseganje evropske podatkovne strategije najpomembnejša izgradnja medsebojnega zaupanja vseh akterjev. Za to so pomembni naslednji cilji:

- razvoj in izboljšanje skupnih podatkovnih storitev,
- oblikovanje in skrb za tesno sodelovanje med uporabniki podatkovne infrastrukture,
- oblikovanje skupne in trajne evropske podatkovne infrastrukture, ki bo upoštevala zahteve različnih uporabniških skupnosti (*Strategy for a European Data Infrastructure* 2009, 1).

Tudi iz tega dokumenta je razvidno, da se utrjuje prepričanje, da je potrebno povezano razvijati evropske podatkovne infrastrukture na različnih področjih in vzdrževati podporo za njihove nacionalne komponente. V dokumentu predstavljajo seznam potreb, ki jih zaznava devet različnih raziskovalnih skupnosti oz. pobud:²²² dolgoročno arhiviranje, iskanje, dostop, hramba, procesiranje podatkov in delovni poteki, podatkovna združenja, dostopnost in visoka stopnja zanesljivosti, avtentikacija, avtorizacija, trajni identifikatorji, metapodatki, spletne storitve, interoperabilnost in standardi, omrežja področnih vozlišč (Koski in dr. 2009, 10).²²³ V prihodnje bo ključnega pomena raznolikost storitev, povezanih s podatkovno infrastrukturo. Spodbuja se tudi nastanek globalne »sodelujoče raziskovalne infrastrukture«, ki vključuje tako generične kot disciplinarne vidike. T. i. osrednje aktivnosti so lahko plast, na katero se nato krožno nalagajo bolj posebne disciplinarne storitve (Slika 2).

²²¹ Strategy for a European Data Infrastructure, White Paper, http://www.grdi2020.eu/Pages/SelectedDocument.aspx?id_documento=a852120d-5f20-4fce-b758-0416c01fbb56, 23. 1. 2014.

²²² To so: CLARIN, LIFEWATCH, DIFRE, ELIXIR, INCF, Climate, Space, STFC, EDIFIS.

²²³ Glej: <http://www.csc.fi/english/pages/parade>, 7. 8. 2013.



Slika 2: Področne in skupne osrednje podatkovne storitve Vir: Koski in dr. 2009, 14

V poročilu *Strategy for a European Data Infrastructure* je izpostavljena ideja, kako povezovati različne skupnosti uporabnikov in infrastrukturne podatkovne storitve skozi različne plasti. Slika 1 prikazuje različne ravni in načine dopolnjevanja storitev, kjer so podlaga skupne generične storitve, potem pa tudi take, ki se razvijajo za različne discipline v pooblaščenih področnih središčih. Med generične storitve štejejo avtentikacije in avtorizacije uporabnikov po protokolu AAI, sistem stalnih identifikatorjev ipd.

Za ključno ugotovitev strategije, ki jo je pripravil konzorcij PARADE, lahko štejemo to, da so različni projekti (omenjajo *GÉANT Expert Group report*, *EGI2020*, *PRACE Roadmap*, *EUDAT Roadmap*, *Riding the Wave* ter *e-IRG Strategy report* in *e-IRG Roadmap*) pokazali, kakšne so potrebe, in dosegli soglasje med akterji glede potrebnih razvojnih korakov. Zdaj je čas za akcijo (*Strategy for a European Data Infrastructure* 2009, 6).

Skladno s tem predlagajo, da EU skupaj z nacionalnimi vladami okrepi uporabniške skupnosti ter da se te usposobijo za novo okolje mednarodnega dostopa do podatkov in po drugi strani promovirajo uporabo strukturnih sredstev (sic!) za razvoj e-infrastrukture na področjih s primanjkljajem (*Strategy for a European Data Infrastructure* 2009, 9–10). Ta ideja je zelo aktualna prav zaradi pogostih poudarkov glede neenakomernih razvojnih dosežkov na različnih disciplinarnih področjih in v različnih državah. Pristope k izkoriščanju strukturnih sredstev pri podpori širitve nacionalnih povezav v projekte ESFRI so predstavili na eni od delavnic projekta CoPoRI²²⁴.

Očitno je, da se lahko z usmerjenim vlaganjem v širitev in krepitev vključenosti večjega števila držav na podlagi dobrih praks in zgledov iz bolj razvitih držav v krajšem času doseže razvojne preskoke. Tak pristop je tudi ena od prednostnih nalog Council of European Social Science Data Archives (CESSDA AS) kot evropske infrastrukture, nastale na podlagi pobud ESFRI. Nekateri projekti s seznama ESFRI imajo izrecno potrebo po krepitvi povezav na podlagi podatkovnih storitev, kot so interoperabilnost, metapodatki, digitalna hramba, trajni identifikatorji ipd. (*Strategy for a European Data Infrastructure* 2009, 10). Na regionalni ravni ima podobno vlogo v nadaljevanju predstavljeni projekt SERSCIDA.

²²⁴ Workshop on Research Infrastructures and Structural Funds, <https://indico.desy.de/conferenceDisplay.py?confId=7693>, 7. 8. 2013.

POVEZOVALNE IN GLOBALNE PODATKOVNE INFRASTRUKTURE

Z vidika vzpostavljanja strokovnih podlag za odpiranje nekaterih ključnih problemov odprtega dostopa do raziskovalnih podatkov je zanimiv tudi projekt [Global Research Data Infrastructures](#) (GRDI 2020).²²⁵ Združuje in povzema nekatere osrednje osebnosti in izpostavlja primere za zgled prihodnjih usmeritev na področju dela z raziskovalnimi podatki, kot je denimo že omenjeno poročilo [Data Quality and Curation](#)²²⁶ Kevina Ashleyja o vprašanih ocenjevanja kakovosti raziskovalnih podatkov v povezavi z njihovo dostopnostjo.

Med poročili vsebuje izčrpen pregled iniciativ in projektov iz zadnjega obdobja, kot so pilotni razvojni projekti v sklopu programa *DataNet* iz ZDA, npr. *DataOne* in *teragrid*, ter evropske infrastrukturne projekte s poudarkom na skrbi za raziskovalne podatke. Poročilo GRDI 2020 z naslovom [Towards a 10 Year Vision for Global Research Data](#)²²⁷ izhaja iz prepričanja, da je med različnimi akterji pri zagotavljanju obstojne, učinkovite in vzdržne podatkovne infrastrukture potrebno vzpostaviti zaupanje, saj so izzivi, s katerimi se soočajo predstavniki različnih disciplin, tako tehnološki kot organizacijski. Naštevajo tudi več načinov, na katere se lahko podatkovna infrastruktura razvija, npr. na način virtualnih raziskovalnih soban (VRS) in orodij za dostop, vezanih na posamezna znanstvena področja.

V okviru projekta vizij je GRDI 2020 podpiral pobude, ki so se nanašale na podatke, pomembne za namene interdisciplinarne znanosti. Nastale so smernice za naslednjih deset let, ki predvidevajo podatkovno infrastrukturo glede na potrebe raziskovalne skupnosti. Tak je bil tudi poskus *World Data System* (WDS), ki ga je ustanovil International Council for Science (ICSU) z naslednjimi cilji: omogočati univerzalen in enak dostop do znanstvenih podatkov, podatkovnih storitev, proizvodov in informacij, katerih kakovost je zagotovljena, zagotoviti dolgoročno financiranje, spodbujati skladnost s sprejetimi standardi ter zagotoviti mehanizme za lažji in boljši dostop do podatkov in podatkovnih proizvodov (*WDS Goals*).²²⁸

V okviru delovanja GRDI 2020 velja izpostaviti tudi poročilo Marka Parsonsa, ki je po novem generalni sekretar svetovne RDA, z naslovom [Data Policy and Open Access](#):²²⁹ politike morajo biti zasnovane tako, da služijo ciljem organizacije, ki podatke ustvarja ali financira, hkrati pa morajo biti oblikovane tako, da služijo širšim interesom znanstvenega raziskovanja. V zaključku avtor priporoča, da naj se pri zasnovi nacionalnih podatkovnih politik upošteva:

- podatkovne politike morajo jasno določati vloge in odgovornosti deležnikov;
- načrti ravnanja s podatki morajo biti obvezni sestavni del vlog za projekte, to naj zahteva financer;
- za implementacijo politik mora financer zagotoviti, da se vključijo vsi deležniki;
- osnovno usposabljanje za ravnanje s podatki mora biti vključeno v izobraževalni proces znanstvenikov;

²²⁵ Glej: <http://www.grdi2020.eu/>, 12. 2. 2014.

²²⁶ Glej: http://www.grdi2020.eu/Pages/SelectedDocument.aspx?id_documento=201287b9-9f1c-4626-8b62-44cfe418707a, 12. 2. 2014.

²²⁷ Glej: <http://www.grdi2020.eu/Repository/FileScaricati/f43bfd7-3dac-4cf6-a506-03b07d01b19f.pdf>, 12. 2. 2014.

²²⁸ Glej: <http://www.icsu-wds.org/organization/intro-to-wds>, 7. 8. 2013.

²²⁹ Glej: <http://www.grdi2020.eu/Repository/FileScaricati/e31a1aab-b01e-4e7e-9b10-0fd93d4b710f.pdf>, 12. 2. 2014.

- financerji, revije, recenzenti in akademije morajo poskrbeti za ustrezno ovrednotenje dela ustvarjalcev podatkov;
- skrbniki podatkov morajo razviti tesno delovno sodelovanje z ustvarjalci podatkov (*Data Policy and Open Access*, 5–6).

V prispevku z naslovom [A conceptual framework for managing very diverse data for complex, interdisciplinary science](#)²³⁰ Parsons s soavtorji ugotavlja, da se pogosto bolj poudarja dostop kot hramba podatkov, kar je napačno, saj sta na ta način zanemarjeni dolgoročna razpoložljivost in uporabnost podatkov. Avtorji poudarjajo, da morajo politike obenem zahtevati ustanavljanje in vzdržnost podpornih ustanov, ki so usposobljene in primerne za opravljanje nalog podatkovnih storitev po modelu podatkovnih središč. Ugotavljajo namreč, da so mednarodna podatkovna središča (kot primer navajajo World Data Centre) kot mesto spravila v veliko primerih nezanesljiva in da je položaj največkrat tak, da za celo vrsto disciplin ni ne primernega mednarodnega ne nacionalnega podatkovnega središča. »Čeprav močno podpiramo mehke pobude, kot npr. citiranje podatkov, s katerim je povezano ovrednotenje prispevka ustvarjalcev podatkov, IPY ugotavlja, da je veliko močnejši ukrep to, da je oddaja podatkov v odprti arhiv obveza, naložena s strani financerja, in da je zahteva podprta s strani pooblaščenih arhivov. Nadalje, ta striktni ukrep je učinkovitejši, ko obstaja dober delovni odnos med arhivom in ustvarjalcem podatkov« (Parsons in dr. 2011, 6). Pomemben poudarek namenjajo standardnemu zapisu podatkov. Skupina, ki je sodelovala pri projektu *International Polar Year* (IPY), ugotavlja, da je razvoj standardov tehnična vaja, tej mora slediti vaja v socialnih veščinah – ki šele omogoči, da udeleženci usvojijo standarde in jih začnejo uporabljati (Parsons in dr. 2011, 6).

V združenem poročilu tako povzemajo ugotovitve, da raziskovalna podatkovna infrastruktura sama po sebi ne zadostuje, obsegati mora organizacijske, tehnične in socialne vidike, ki zagotavljajo povezano znanstveno delo. Vsi trije vidiki morajo biti predmet načrtovanja in izdelave, sicer podatkovna infrastruktura ne bo izpolnila svojih nalog ([Towards a 10 year vision for global data infrastructure](#) 2012, 13).²³¹

V viziji so zapisali, da mora globalna raziskovalna podatkovna infrastruktura za to, da ustvarja in vzdržuje zanesljivo, operativno, digitalno znanstveno okolje, podpirati ustvarjanje in ohranjanje storitvenih okolij, ki omogočajo učinkovitost storitev podatkovnega ekosistema: okolje za beleženje nastanka, okolje iskanja podatkov, okolje citiranja podatkov, okolje za izbor orodij, okolje iskanja podatkov, okolje integracije podatkov, okolje deljenja podatkov, okolje povezovanja podatkov, okolje ontologije/taksonomije ravnanja, okolje prenosljivih podatkov, okolje storitev/orodij prenosljivih podatkov, okolje podpore znanstvenega delovnega poteka, okolje politik ravnanja ([Towards a 10 year vision for global data infrastructure](#) 2012, 21–24).²³²

PROJEKTI GRADNJE IN IZPOPOLNJEVANJA PODROČNE RAZISKOVALNE PODATKOVNE INFRASTRUKTURE

²³⁰ Glej: <http://iis.sagepub.com/cgi/content/long/37/6/555http://iis.sagepub.com/content/37/6/555.full.pdf+html>, 12. 2. 2014.

²³¹ Glej: <http://www.grdi2020.eu/Repository/FileScaricati/e2b03611-e58f-4242-946a-5b21f17d2947.pdf>, 12. 2. 2014.

²³² Glej: <http://www.grdi2020.eu/Repository/FileScaricati/e2b03611-e58f-4242-946a-5b21f17d2947.pdf>, 12. 2. 2014.

*Support for Establishment of National/Regional Social Sciences Data Archives (SERSCIDA)*²³³ je eden od tekočih projektov, ki jih financira EK. Kot prvega ga izpostavljamo zato, ker kaže, da je mogoče z relativno majhnim projektom postaviti dobre temelje za vzpostavitev podatkovnih področnih središč. Naloga projekta je vzpostaviti pogoje za ustanovitev družboslovnih podatkovnih središč v regiji, ki vključuje Hrvaško, Bosno in Hercegovino ter Srbijo. Pripravljajo se povsem operativni načrti ustanovitve in delovanja enot. Pazljivo načrtovanje organizacijskih ukrepov in pridobivanje podpore s strani financerjev spremljajo pregledi situacij, obiski z namenom pridobivanja izkušenj, pilotno vzpostavljanje tehnične podlage. Pristop, ki ga ubirajo v projektu, je lahko uporaben za druge države in za druga področja kot model, ki ga je mogoče prilagoditi danim situacijam. Pri projektu kot partner sodeluje tudi slovenski ADP.²³⁴ Projekt SERSCIDA ima jasno navezavo na osrednjo koordinacijsko točko vzpostavljanja panevropske raziskovalne infrastrukture *European Strategy Forum on Research Infrastructures (ESFRI)*, saj predstavlja sodelovanje partnerjev iz razvitih okolij iz enote za dostop do raziskovalnih podatkov CESSDA s področja družboslovja in partnerjev iz bivše Jugoslavije.

ESFRI je, kot rečeno, osrednja iniciativa za evropsko povezano gradnjo večjih znanstvenih infrastrukturnih enot. Slovenija se je do seznama predlogov opredelila s pripravo [Načrta razvoja raziskovalnih infrastruktur](#).²³⁵ Med projekti, ki so vključeni v ta načrt, mnogi izrazito poudarjajo skrb za podatke. Ta naloga – skrb za podatke – kot stična točka predstavlja možnost za bolj koordinirano mednarodno delovanje infrastrukturnih enot, pa tudi da se ta poudarek na raziskovalnih podatkih prenese v nacionalna okolja, v katerih imajo infrastrukturne enote svoje predstavnike.

Vzorčen primer poskusa tovrstne koordinacije na ravni Evrope je projekt [Data Infrastructure for the Social Sciences and Humanities \(DASISH\)](#),²³⁶ ki povezuje pet raziskovalnih infrastruktur ESFRI s področja družboslovja in humanistike (CLARIN, DARIAH, CESSDA, ESS, SHARE). Cilj je, da se v okviru projekta opredelijo sinergije v razvoju infrastrukture in da se spodbudi skupne aktivnosti. Projekt [BioMedbridges](#)²³⁷ združuje raziskovalne infrastrukture ESFRI s področja biologije in medicine (ELIXIR, BBMRI, EATRIS, ECRIN, InfraFron.er in INSTRUCT). [Common Operations of Environmental Research Infrastructures \(ENVRI\)](#)²³⁸ združuje ESFRI raziskovalne infrastrukture s področja okoljskih znanosti (EISCAT_3D, EMSO, EPOS, Euro-Argo, ICOS, LifeWatch) – gradi skupne storitve in programsko opremo, pri tem se soočajo z vprašanji, kot so zajem podatkov, standardizacija metapodatkov, ravnanje z obsežnimi količinami podatkov, delovni proces in vizualizacija podatkov. [Cluster of Research Infrastructures for Synergies in Physics \(CRISP\)](#),²³⁹ v katerem so združene raziskovalne infrastrukture s področja fizike, astronomije in analitičnih zmožnosti (ESRF, GSI-FAIR, ILL, SLHC and GANIL-SPIRAL2). CRISP je lahko primer tehnološkega premika na področju fizike. Njegovo poslanstvo je spodbuditi in omogočiti partnerjem, da se povežejo pri izmenjavi znanja in se na učinkovit način dopolnjujejo. Pri tem sodelujejo strokovnjaki s področij pospeševalnikov, detektorjev, eksperimentov, informacijske tehnologije in ravnanja s podatki. Kot povezovalni projekt velja omeniti [D4Science-](#)

²³³ Glej: <http://www.serscida.eu/en/>, 23. 1. 2014.

²³⁴ Glej nedavne aktivnosti znotraj projekta SERSCIDA v organizaciji ADP: <http://www.adp.fdv.uni-lj.si/projekti/serscida/>, 23. 1. 2014.

²³⁵ Glej: <http://www.arhiv.mvzt.gov.si/fileadmin/mvzt.gov.si/pageuploads/pdf/znanost/nacrt-RI.pdf>, 23. 1. 2014.

²³⁶ Glej: http://dasish.eu/about_dasish/, 12. 2. 2014.

²³⁷ Glej: <http://www.biomedbridges.eu/>, 12. 2. 2014.

²³⁸ Glej: <http://envri.eu/home>, 12. 2. 2014.

²³⁹ Glej: <http://www.crisp-fp7.eu/about-crisp/>, 12. 2. 2014.

[II](#),²⁴⁰ katerega cilj je zagotoviti mehanizme za boljše sodelovanje različnih e-infrastruktur s podatkovnimi infrastrukturami, ki delujejo avtonomno in tako ustvarjajo jedro e-infrastrukturnega ekosistema, npr. s področja biodiverzitete, ribištva in fizike osnovnih delcev.

Obstajajo tudi različni projekti, usmerjeni v tehnološko podporo posameznim področjem znanosti. Isatools ([ISA](#))²⁴¹ – odprtokodna orodja za spremljanje metapodatkov, pomagajo pri vodenju zelo raznolikih eksperimentov na področju »znanosti o življenju«, okolja in biomedicine, ki terjajo eno ali kombinacijo več tehnologij. Orodja pomagajo zagotoviti bogate opise eksperimentalnih metapodatkov (kot npr. značilnosti vzorca, vrsta tehnologij in meritev), kar omogoča ponovno rabo podatkov in odkritij.

V Sloveniji predstavljajo potencial za razvoj podatkovne raziskovalne infrastrukture projekti s seznama ESFRI, v katerih aktivno sodelujejo slovenski raziskovalci. Ti so skladno z *Načrtom razvoja raziskovalnih infrastruktur*²⁴² zadolženi za primerno vzpostavitev nacionalne komponente:

[SHARE](#) (Raziskava zdravja, staranja in življenja upokojencev v Evropi (*Survey of Health, Ageing and Retirement in Europe*)): Inštitut za ekonomska raziskovanja ([IER](#))²⁴³

[ESS](#) (Evropska družboslovna raziskava (*European Social Survey*)): Center za raziskovanje javnega mnenja in množičnih komunikacij ([CJMMK](#), FDV)²⁴⁴

[CESSDA](#) (Svet evropskih arhivov družboslovnih podatkov (*Council of European Social Science Data Archives*)): Arhiv družboslovnih podatkov ([ADP](#), [UL](#))²⁴⁵

[DARIAH](#) (Digitalna raziskovalna infrastruktura za umetnost in humanistiko (*Digital Research Infrastructure for the Arts and Humanities*)): [Inštitut za novejšo zgodovino](#)²⁴⁶ in [ZRC SAZU](#)²⁴⁷

[CLARIN](#) (Infrastruktura za skupne jezikovne vire in tehnologije (*Common Language Resources and Technology Infrastructure*)): Institut Jožef Stefan ([IJS](#))²⁴⁸

[EATRIS](#) ((Evropska infrastruktura za translacijske raziskave v medicini (*European Advanced Translational Research Infrastructure in Medicine*)): Fakulteta za farmacijo ([FFA UL](#))²⁴⁹

[ELIXIR](#) (Evropska infrastruktura za vede o življenju in biološke informacije (*European Life sciences Infrastructure for Biological Information*)): Medicinska fakulteta ([ME, UL](#))²⁵⁰

²⁴⁰ Glej: <http://www.d4science.eu/>, 12. 2. 2014.

²⁴¹ Glej: <http://isa-tools.org/>, 12. 2. 2014.

²⁴² »Primeri raziskovalnih infrastruktur so: velike raziskovalne postavitve, zbirke, knjižnice, baze podatkov, biološki arhivi in zbirke, visokozmogljiva oz. širokopasovna komunikacijska omrežja, raziskovalna plovila, teleskopi, satelitske in letalske opazovalne zmogljivosti, mreže računskih zmogljivosti (HPC), čiste sobe, obalne opazovalne postaje, sinhrotroni in pospeševalniki ipd.« (*Načrt razvoja raziskovalnih infrastruktur*, str. 3).

²⁴³ Glej: <http://www.ier.si/>, 12. 2. 2014.

²⁴⁴ Glej: <http://www.cjm.si/>, 12. 2. 2014.

²⁴⁵ Glej: <http://www.adp.fdv.uni-lj.si/>, 12. 2. 2014.

²⁴⁶ Glej: <http://www.inz.si/index.php>, 12. 2. 2014.

²⁴⁷ Glej: <http://www.zrc-sazu.si/sl/node>, 12. 2. 2014.

²⁴⁸ Glej: <http://www.ijs.si/>, 12. 2. 2014.

²⁴⁹ Glej: <http://www.ffa.uni-lj.si/>, 12. 2. 2014.

²⁵⁰ Glej: <http://cfghbc.mf.uni-lj.si/default.html>, 12. 2. 2014.

LifeWatch (Znanstvena in tehnološka infrastruktura za podatke o biotski raznovrstnosti in opazovalnice (*Science and Technology Infrastructure for Biodiversity Data and Observatories*)): [ZRC SAZU](#)²⁵¹

EPOS (Evropski sistem opazovanja plošč (*European Plate Observing System*)): [ZRC SAZU](#)²⁵²

CERN (Evropska organizacija za jedrske raziskave (*European Organization for Nuclear Research*)): IJS/CERN

PRACE (Partnerstvo za napredno računanje v Evropi (*Partnership for Advanced Computing in Europe*)): [Fakulteta za strojništvo](#), UL²⁵³

ILL (Inštitut Laue Langevin (*Institute Laue Langevin*)): Kemijski inštitut ([KI](#))²⁵⁴

C-ERIC (Srednjeevropski konzorcij raziskovalnih infrastruktur (*Central-European Research Infrastructures Consortium*)): Kemijski inštitut ([NMR Center](#))²⁵⁵

FAIR²⁵⁶ (Center za raziskave z antiprotoni in ioni v Evropi (*Facility for Antiproton and Ion Research in Europe*)): [IJS](#)²⁵⁷

Izpostaviti velja napredni pristop enote ELIXIR, pri katerem gre za poskus pomagati znanstvenikom po Evropi pri deljenju in ohranjanju podatkov ter podporo pri obstoječih virih, kot so podatkovne baze in računalniška tehnologija. Raziskovalna področja oziroma raziskovalni projekti se med seboj razlikujejo, tako npr. v okviru projekta CERN podatki nastajajo na enem mestu, v primeru projekta ELIXIR pa je velika količina podatkov ustvarjena v številnih laboratorijih. Velika količina t. i. velikih podatkov nastaja v računalniških oblakih. Z uporabo te tehnologije raziskovalcem ni treba na novo kupovati svoje opreme in jo vzdrževati. Kadar jim njihova oprema ne zadošča, jo lahko dopolnijo s tisto v oblaku. Ustvarijo lahko virtualne sobane za podatke, programsko opremo in za rezultate, do katerih lahko vsi dostopajo, če želijo, pa lahko dostop tudi omejijo na določeno skupino uporabnikov. Ideja je, da se omogoči čim širši dostop do računalniških zmogljivosti, tako da lahko raziskovalno delo izvajajo tudi tisti, ki jim njihove računalniške kapacitete tega ne omogočajo (Marx 2013).²⁵⁸

Nekateri drugi projekti s seznama ESFRI imajo opredeljeno nacionalno kontaktno točko, tudi če država ni polnopravna članica. Tako je med sodelujočimi organizacijami infrastrukture *Biobanking and Biomolecular Resources Research Infrastructure* ([BBMRI](#))²⁵⁹ med drugim naveden tudi [Inštitut za anatomijo](#),²⁶⁰ Medicinska fakulteta UL.²⁶¹

²⁵¹ Glej: <http://izr.k.zrc-sazu.si/sl/programi-in-projekti/life-watch#v>, 12. 2. 2014.

²⁵² Glej: <http://izr.k.zrc-sazu.si/sl/programi-in-projekti/epos-evropski-sistem-opazovanja-plo%C5%A1%C4%8D-european-plate-observing-system#v>, 12. 2. 2014.

²⁵³ Glej: <http://www.fs.uni-lj.si/>, 12. 2. 2014.

²⁵⁴ Glej: <http://www.ki.si/>, 12. 2. 2014.

²⁵⁵ Glej: <http://www.nmr.ki.si/index.html>, 12. 2. 2014.

²⁵⁶ Glej: <http://www.fair-center.eu/partners/si-slovenia.html>, 12. 2. 2014.

²⁵⁷ Glej: <http://www.fair-center.eu/partners/si-slovenia.html>, 23. 1. 2014.

²⁵⁸ Marx (junij 2013), dostopno prek:

http://www.nature.com/nature/journal/v498/n7453/full/498255a.html?WT.ec_id=NATURE-20130613, 23. 1. 2014.

²⁵⁹ Glej: <http://bbmri.eu/> in <http://bbmri.eu/associated-organisations>, 12. 12. 2013.

²⁶⁰ Glej: <http://www.eurobiobank.org/en/partners/partners.htm#15>, 12. 2. 2014.

²⁶¹ Glej: <http://www.eurobiobank.org/en/partners/partners.htm#15>, 9. 1. 2014.

Poglejmo še nekaj pomembnejših projektov s tega področja. Evropski projekt [OpenAIRE](#)²⁶² si prizadeva za izgradnjo podporne infrastrukture za raziskovalce, kar zajema vzpostavitev in izvajanje e-infrastrukture za ravnanje z različnimi oblikami znanstvenih objav, hkrati pa proučuje problematiko njihovega povezovanja z raziskovalnimi podatki (prakse, pobude, modeli, tehnologije ...). OpenAIRE predstavlja tudi koordinacijsko točko za izpolnjevanje zahtev glede obvezne priprave načrtov ravnanja z raziskovalnimi podatki, ki jo napoveduje EK pri novem sklicu raziskovalnih projektov Obzorje 2020. Pri projektu je pomembna vloga Irine Kučma kot regionalne zastopnice in Mojce Kotar v vlogi slovenske predstavnice pri svetovanju in posredovanju relevantnih informacij. Poudarki so na vzpostavitvi storitev, ki bodo preko šifer raziskovalnih projektov podpora za povezovanje podatkov in publikacij kot znanstvenih rezultatov. Predpisali so določila glede standardnega zapisa Dublin Core in z njim povezanega protokola [OAI-PMH](#).²⁶³ Skratka, tudi pri tem projektu povezovanja odprto dostopnih repozitorijev je v ospredju zanimanje, kako v obstoječo infrastrukturo vključiti raziskovalne podatke in kako na tej podlagi razvijati nadaljnje storitve za uporabnike, za lažje iskanje, prehajanje in povezovanje različnih virov informacij.

Nekateri drugi projekti, financirani s strani EK, ki imajo med cilji vprašanje dostopa do raziskovalnih podatkov, so še:

- *Policy RECommendations for Open Access to Research Data in Europe* ([RECODE](#))²⁶⁴
- *Mediterranean Open Access Network* ([MEDOANET](#)),²⁶⁵ ki je povezan s SeaDataNet, pri katerem sodeluje [slovenska Morska biološka postaja Piran](#).²⁶⁶
- *Establishing an Efficient Network for Cancer Communication in Europe* ([EUROCANCERCOMS](#))²⁶⁷
- *Digital Repository Infrastructure Vision for European Research* ([DRIVER II](#))²⁶⁸
- *Euro-VO Astronomical Infrastructure for Data Access* ([EUROVO-AIDA](#))²⁶⁹
- *European organisation for international research information* ([EuroCRIS](#))²⁷⁰
- *Opportunities for Data Exchange* ([ODE](#))²⁷¹
- *ORCID and DATAcITE Interoperability Network* ([ODIN](#))²⁷²

Nemogoče je podrobno predstaviti vse, je pa pomembno, da se v prihodnje ohrani pregled in da se spremljajo aktivnosti, če se dotikajo posameznih področij delovanja slovenskih infrastrukturnih enot. To velja tudi za dva projekta, ki sta imela v začetku leta 2014 začetni srečanja:²⁷³ *Fostering Open Science Training for European Research* ([FOSTER](#))²⁷⁴ in *Open Access Policy Alignment Strategies for European Union Research* ([PASTEUR4OA](#)),²⁷⁵ ki se oba vsebinsko tesno navezujeta na aktivnosti, predvidene v tu predstavljenem akcijskem načrtu.

²⁶² Glej: <https://www.openaire.eu/en/about-openaire/openaire-project/objectives>, 12. 2. 2014.

²⁶³ Glej: <http://www.openarchives.org/pmh/>, 12. 2. 2014.

²⁶⁴ Glej: <http://recodeproject.eu/>, 12. 2. 2014.

²⁶⁵ Glej: <http://www.medoanet.eu/>, 12. 2. 2014.

²⁶⁶ Glej: <http://www.mbss.org/portal/index.php>, 12. 12. 2013.

²⁶⁷ Glej: <http://www.eurocancercoms.eu/>, 12. 2. 2014.

²⁶⁸ Glej: <http://www.driver-repository.eu/>, 12. 2. 2014.

²⁶⁹ Glej: <http://cds.u-strasbg.fr/twikiAIDA/bin/view/EuroVOAIDA>, 12. 2. 2014.

²⁷⁰ Glej: <http://www.eurocris.org/Index.php?page=homepage&t=1>, 12. 2. 2014.

²⁷¹ Glej: <http://www.alliancepermanentaccess.org/index.php/community/current-projects/ode/>, 12. 2. 2014.

²⁷² Glej: <http://odin-project.eu/>, 12. 2. 2014.

²⁷³ Glej: <http://openaccess.sdum.uminho.pt/?p=2844>, 23. 1. 2014.

²⁷⁴ Glej: <http://www.fosteropenscience.eu/>, 12. 2. 2014.

²⁷⁵ Glej: <http://www.pasteur4oa.eu/>, 12. 2. 2014.

PRESOJA PRIMERNIH PODATKOVNIH SREDIŠČ

Odperta znanost zahteva tudi trajno informacijsko infrastrukturo, ki bo znanstvenikom v podporo pri deljenju njihovih raziskovalnih podatkov ter bo zagotavljala dostop in možnost ponovne uporabe podatkov tudi za prihodnje generacije (Pampel in dr. 2013).²⁷⁶

Ugotovili smo že, da vedno več priznanih strokovnih revij ob znanstvenih objavah uveljavlja tudi zahtevo po odprtem dostopu do raziskovalnih podatkov, ki so služili kot podlaga za objavljen članek. Večina naprednih nacionalnih ali področnih politik prav tako določa priporočeno mesto izročanja najpomembnejših raziskovalni podatkov. V svojih priporočilih svetujejo, da avtorji uporabijo katero od uveljavljenih podatkovnih središč. Tako bodo podatki bolj vidni, preprosto dostopni preko trajnega identifikatorja mesta dostopa in zagotovljena bo dolgoročna hramba.

Podatkovna središča, ki sprejemajo specifične tipe podatkov, kot so na primer molekularna zaporedja (npr. [DDBJ](#),²⁷⁷ [GenBank](#),²⁷⁸ [MG-RAST](#)),²⁷⁹ imajo pogosto visoko standardizirana pravila glede tipov podatkov, formatov in kakovosti. Ustvarjalci podatkov se morajo s temi pravili seznaniti in zagotoviti njihovo izpolnjevanje, saj je prav ta formalni vidik pogosto v ospredju pri izboru raziskovalnih podatkov za sprejem v podatkovno središče.

Nekateri podatkovni repozitoriji so namenjeni predvsem prevzemu preprosto strukturiranih raziskovalnih podatkov, ki spremljajo objavo v znanstveni publikaciji. Ti običajno pokrivajo širši nabor različnih tipov podatkov, so manj standardizirani (npr. [Dryad](#),²⁸⁰ [KNB](#),²⁸¹ [PANGAEA](#)),²⁸² obstajajo pa tudi repozitoriji, ki sprejemajo kakršne koli vrste podatkov (npr. [Figshare](#)).²⁸³

Priporočila ustvarjalcem podatkov pri izbiri repozitorija običajno vključujejo nasvet, naj razmislijo, kje drugi raziskovalci iz iste znanstvene discipline predajajo svoje podatke, in se tako seznanijo s pristopom deljenja podatkov v področni skupnosti, s čimer povečajo možnost prepoznavnosti in uporabe svojih podatkov. Razmislijo naj o pogojih uporabe, pravicah, podeljevanju licenc in možnosti citiranja, pri čemer je za prepoznavnost raziskovalnih podatkov zelo pomembno, da repozitorij podatkom dodeljuje enolične identifikatorje.

Pri odločanju za ustrezno mesto predaje raziskovalnih podatkov, če to ni nacionalno ali mednarodno podatkovno središče, priporočeno s strani politik, lahko pomaga podrobnejši opis in pregled specialnih storitev, ki jih ponujajo obstoječi ponudniki podatkovne infrastrukture. Temu sta namenjena registra ponudnikov storitev v podporo dostopu do raziskovalnih podatkov [Re3data.org](#)²⁸⁴ in [Databib](#).²⁸⁵ Gre za orodji, s katerima lahko iščemo različne specializirane zbirke raziskovalnih podatkov.

²⁷⁶ *Making Research Data Repositories Visible: The re3data.org Registry*, <http://dx.doi.org/10.7287/peerj.preprints.21v1>, 23. 1. 2014.

²⁷⁷ Glej: <http://www.ddbj.nig.ac.jp/>, 12. 2. 2014.

²⁷⁸ Glej: <http://www.ncbi.nlm.nih.gov/genbank/>, 12. 2. 2014.

²⁷⁹ Glej: <http://metagenomics.anl.gov/>, 12. 2. 2014.

²⁸⁰ Glej: <http://datadryad.org/>, 12. 2. 2014.

²⁸¹ Glej: <http://knb.ecoinformatics.org/index.jsp>, 12. 2. 2014.

²⁸² Glej: <http://www.pangaea.de/>, 12. 2. 2014.

²⁸³ Glej: <http://figshare.com/>, 12. 2. 2014.

²⁸⁴ Glej: <http://service.re3data.org/search/>, 12. 2. 2014.

²⁸⁵ Glej: http://databib.org/index_subjects.php, 12. 2. 2014.

Opisi repozitorijev v registru Re3data.org²⁸⁶ vključujejo splošne informacije, pristojnosti, politike, pravne vidike, tehnične in metapodatkovne standarde, standarde kakovosti ter storitve.

Podoben pregled ponuja tudi Databib.²⁸⁷ Tudi tu gre za register podatkovnih repozitorijev, kjer so za vsak posamezen repozitorij na voljo naslednje informacije: naziv, URL, teme, ki jih podatki pokrivajo, opis, možnosti dostopa, datum začetka delovanja, lokacija, pravila, vezana na ponovno uporabo podatkov, pogoji za predajo podatkov in tip repozitorija. Po vseh teh parametrih je omogočeno tudi iskanje.

Glede na način organizacije delimo podatkovne repozitorije, kakor sledi:

Institucionalni

Institucionalne repozitorije vodijo raziskovalne institucije ali univerze. Pri slednjih gre za multidisciplinarne repozitorije, kakršen je na primer Edinburgh Datashare²⁸⁸ – spletni digitalni repozitorij za podatke raziskav z različnih področij, ki so nastale na Univerzi v Edinburgu.

Področni

Značilnost področnih repozitorijev je, da so namenjeni posebnim vrstam podatkov, ki nastanejo znotraj določene znanstvene discipline. Primera takšnega repozitorija sta npr. GenBank²⁸⁹ – javna podatkovna baza nukleotidnih zaporedij, in PANGEA²⁹⁰ – odprto dostopna podatkovna knjižnica, ki si prizadeva za hrambo, objavo in distribucijo geoloških podatkov.

Večdisciplinarni

Služijo za podatke z več področij. Primer takšnega repozitorija je Figshare,²⁹¹ ki raziskovalcem omogoča objavljanje različnih rezultatov raziskav, ki jih je mogoče na enostaven način citirati, deliti in poiskati, pri čemer uporablja licenco Creative Commons.²⁹² Eden bolj uveljavljenih je na primer tudi Dryad²⁹³ – mednarodni repozitorij za podatke, ki so podlaga recenzirani znanstveni in medicinski literaturi, namenjen je predvsem podatkom, za katere ni specializiranega podatkovnega središča.

Smiselna pa je tudi delitev na področne, nacionalne in mednarodne repozitorije. V večkrat omenjenem poročilu *Surfboard for Riding the Wave* so naredili primerjavo med štirimi državami, vključenimi v skupino KE, in ugotovili, da ima največ repozitorijev Velika Britanija, to je 17, v Nemčiji jih je devet, na Nizozemskem osem in na Danskem trije. Ob tem ugotavljajo, da večinoma ni zadovoljivega pregleda, kakšna podatkovna infrastruktura je na voljo po posameznih državah (Graaf 2011, 24).

Za vzpostavljanje pregleda in ažuriranje evidenc o obstoju podatkov, tudi če ti še niso dostopni, je izrednega pomena sistematizirano poročanje o »pojavi« novih podatkov. Potrebno je vzpostaviti dolžnost prijave in dolžnost prevzema metapodatkov s strani področnih podatkovnih središč, ki omogoči zagotavljanje registracije obstoječih podatkovnih »objav« (mednarodni repozitoriji, podatki ki spremljajo publikacije, itd.). Data Seal of Approval²⁹⁴ v svojih navodilih

²⁸⁶ Glej: <http://service.re3data.org/search/>, 12. 2. 2014.

²⁸⁷ Glej: <http://databib.org/>, 12. 2. 2014.

²⁸⁸ Glej: <http://datashare.is.ed.ac.uk/>, 12. 2. 2014.

²⁸⁹ Glej: <http://www.ncbi.nlm.nih.gov/genbank/>, 12. 2. 2014.

²⁹⁰ Glej: <http://www.pangaea.de/>, 12. 2. 2014.

²⁹¹ Glej: <http://figshare.com/>, 12. 2. 2014.

²⁹² Glej: <http://creativecommons.org/>, 12. 2. 2014.

²⁹³ Glej: <http://datadryad.org/>, 12. 2. 2014.

²⁹⁴ Glej: <http://datasealofapproval.org/en/>, 12. 2. 2014.

navaja dolžnosti za raziskovalce ustvarjalce podatkov, za institucije izvajalke storitev in za uporabnike podatkov. Institucijam, ki izvajajo podatkovne storitve nalaga naslednje dolžnosti:²⁹⁵

- Podatkovni repozitorij ima posebno poslanstvo na področju digitalnega arhiviranja in širjenja le-tega.
- Podatkovni repozitorij zagotavlja skladnost s pravnimi regulativami in pogodbami, ki vključujejo – kadar je to potrebno – zaščito ljudi.
- Podatkovni repozitorij zagotavlja dokumentacijo obravnav in postopkov za delovanje na področju ohranjanja podatkov.
- Podatkovni repozitorij ima načrt za dolgotrajno ohranjanje svojih digitalnih gradiv.
- Arhiviranje je del življenjskega kroga podatkov.
- Podatkovni repozitorij predpostavlja, da bodo ustvarjalci podatkov odgovorni, kadar gre za dostop in razpoložljivost digitalnih objektov.
- Podatkovni repozitorij omogoča uporabnikom, da najdejo in uporabijo podatke, ter jih ustrezno navajajo.
- Podatkovni repozitorij zagotavlja integriteto digitalnih objektov in metapodatkov.
- Podatkovni repozitorij zagotavlja avtentičnost digitalnih objektov in metapodatkov.
- Tehnična infrastruktura eksplicitno podpira naloge in vloge opisane v mednarodno sprejetih arhivarskih standardih, kot OAIS. (Navodila, povezana z repozitoriji, DSA)²⁹⁶

Nedavna pobuda predstavnikov področnih podatkovnih repozitorijev v ZDA, da se zagotovi vzdržnejše financiranje, izpostavlja osrednjo vlogo, ki jo imajo področna središča pri ekosistemu povezovanja in širjenja dobrih praks in pristopov, ki zagotavljajo znanstveno integriteto in kakovost ter dostopnost podatkov.²⁹⁷

²⁹⁵ Primerjaj Delavnico na konferenci IDCC 2013 na temo presoje podatkovnih repozitorijev »Data publishing, peer review and repository accreditation: everyone a winner?« na

<http://www.dcc.ac.uk/events/idcc13/workshops#sthash.zewgusyq.dpuf> ter poročilo projekta PREPARDE: http://proj.badc.rl.ac.uk/preparde/attachment/wiki/DeliverablesList/D5_1_Generic%20Repository%20Accreditation%20guidelines_v8.pdf, 23. 12. 2013.

²⁹⁶ Guidelines Related to Repositories, <http://data.sealofapproval.org/en/information/guidelines/>, 23. 12. 2013.

²⁹⁷ Sustaining Domain Repositories for Digital Data: A Call for Change from an Interdisciplinary Working Group of Domain Repositories, <http://www.icpsr.umich.edu/icpsrweb/ICPSR/support/announcements/2013/09/sustaining-domain-repositories-for>, 9. 1. 2014.

SREDIŠČA ZA STROKOVNO PODPORO IN SVETOVANJE PRI DELU Z RAZISKOVALNIMI PODATKI

Na področju odpiranja raziskovalnih podatkov pomembne poti utirajo tudi številne mednarodne pobude in projekti, ki niso neposredno povezani z Evropsko unijo. Med njimi tudi naslednji.

Committee on Data for Science and Technology

Committee on Data for Science and Technology ([CODATA](#))²⁹⁸ je ustanovil leta 1966 International Council for Science ([ICSU](#)),²⁹⁹ in sicer s ciljem, da bi skrbel za spodbujanje in promocijo, vrednotenje in diseminacijo zanesljivih raziskovalnih podatkov, pomembnih za znanost in tehnologijo v dobro družbe. Aktivnosti potekajo v okviru delovnih skupin, konferenc,³⁰⁰ delavnic, publikacij in sodelovanja z drugimi sorodnimi organizacijami – z njimi skušajo razvijati in širiti znanje o podatkih in o delu s podatki, kot so standardi, etična vprašanja, tehnološko sodelovanje. Izdajajo posebno revijo [Data Science Journal](#).³⁰¹ Ena od delovnih skupin pri CODATA je imela zelo opazno vlogo pri pripravi Deklaracije OECD. Pričakovati je, da se bodo aktivnosti tega foruma z nastopom novega izvršnega direktorja Simona Hodsona³⁰² in pa novim strateškim planom 2013–18 ([CODATA Strategic Plan 2013–2018](#))³⁰³ okrepile v nadaljevanju dolge tradicije spodbujanja kakovosti in dostopnosti raziskovalnih podatkov. Po napovedih se bo CODATA kot partner dejavno vključeval znotraj nove [Research Data Alliance](#).³⁰⁴

Knowledge Exchange

Mednarodna skupina Knowledge Exchange ([KE](#))³⁰⁵ je zanimiva zaradi svojega organizacijskega modela, saj združuje financerje znanosti, odgovorne za najvišjo raven politik, ter strokovne organizacije s ciljem, da se skupaj lotijo posameznih izzivov. Financirajo se iz članarine držav. Kot skupina imajo boljše možnosti za uspeh pri pridobivanju mednarodnih in meddržavnih projektov. Skupina sama iz lastnih sredstev financira vrsto projektov in z njimi povezanih poročil, ki obravnavajo aktualne probleme dostopa do podatkov v akademskem okolju. Njihov program dela bi bilo koristno preveriti ob pripravah tematik nacionalnih razvojnih projektov po posameznih področjih, saj ponujajo tematsko dobro domišljene poudarke (npr. storitve podpore načrtom ravnanja z raziskovalnimi podatki) in koristna gradiva, ki služijo kot podlaga za izboljšave na področju lokalnih podatkovnih storitev. V KE kot partnerji sodelujejo finski [CSC - IT Center for Science](#),³⁰⁶ danski [Denmark's Electronic Research Library](#),³⁰⁷ nemški [German Research Foundation](#),³⁰⁸ britanski [JISC](#)³⁰⁹ in nizozemski [SURF](#).³¹⁰ Partnerji sodelujejo na področju

²⁹⁸ Glej: <http://www.codata.org/>, 12. 2. 2014.

²⁹⁹ Glej: <http://www.icsu.org/>, 12. 2. 2014.

³⁰⁰ Zadnja je bila v Taipeiju leta 2012 <http://www.codata2012.com>, naslednja bo 2014 v Indiji.

³⁰¹ Glej: <http://www.codata.org/dsj/>, 12. 2. 2014.

³⁰² Twitter: @simonhodson99

³⁰³ Glej: http://www.codata.org/CODATA_Strategic_%20Plan_2013-2018.pdf, 12. 2. 2014.

³⁰⁴ Glej: <https://rd-alliance.org/node>, 12. 2. 2014.

³⁰⁵ Glej: <http://knowledge-exchange.info/>, 12. 2. 2014.

³⁰⁶ Glej: <http://www.csc.fi/english>, 12. 2. 2014.

³⁰⁷ Glej: <http://www.deff.dk/>, 12. 2. 2014.

³⁰⁸ Glej: <http://www.dfg.de/>, 12. 2. 2014.

odprtega dostopa, raziskovalnih podatkov, raziskovalnih orodij in tehnologij ter interoperabilnostnih standardov, in sicer s širšo vizijo, da bi izboljšali evropsko raziskovalno okolje. Poleg tega da skupina organizira strokovna posvetovanja na posamezne teme, povezane z vprašanji odprtega dostopa do raziskovalnih podatkov,³¹¹ delujejo tudi na področju usklajevanja strategij in politik. V tem okviru lahko ponovno izpostavimo že omenjeni akcijski načrt štirih držav iz leta 2011 (Graaf 2011). Na osnovi različnih poročil, ki analizirajo stanje, ugotavljajo, da praksa in kultura deljenja podatkov nista dovolj razširjeni zaradi različnih ovir, zlasti zaradi premajhne motiviranosti raziskovalcev in premajhne podpore, ki so je deležni (Graaf in Waaijers 2011, 8).

Pregled področij, ki jih pokrivajo podatkovna središča, je izmed štirih partnerskih držav skupine Knowledge Exchange najboljše v Veliki Britaniji. Skupna ugotovitev poročila je, da obstaja še vrsta področij, ki niso pokrita. Partnerji napovedujejo kot prioriteto ukvarjanje z vprašanji vrzeli v pokritosti ter z vprašanji medsebojnega povezovanja obstoječih podatkovnih storitev.

Digital Curation Centre

Veliko svetovno pomembnih centrov znanja, ki delajo na področjih, povezanih z dostopom do raziskovalnih podatkov, prihaja iz Velike Britanije. Med njimi je še ena zelo ugledna organizacija za strokovna vprašanja skrbnega ravnanja z raziskovalnimi podatki, to je Digital Curation Centre (DCC),³¹² ki so ga ustanovili leta 2004 na pobudo britanskega JISC-a, s ciljem postati nacionalni center za izzive na področju digitalne hrambe, ki presega zmožnosti posamezne ustanove. DCC, lociran na Škotskem, je edinstven s svojo produkcijo orodij in vodičev glede digitalne hrambe, posvečenih raziskovalnim podatkom. Vsaki dve leti organizirajo veliko konferenco *International Digital Curation Conference (IDCC)*,³¹³ v okviru katere so predstavljeni najnovejši pristopi in dosežki na področju. Predstavljajo zelo uporabne vodiče oz. priročnike.³¹⁴ V skladu z aktualnim dogajanjem so v zadnjem obdobju zelo aktivni pri podpori pisanju NRRP. Na tem področju so poleg priročnikov in vodil razvili *on-line* pripomoček za izdelavo NRRP, ki ga zelo široko uporabljajo tako v Veliki Britaniji kot širše in v različnih disciplinah.³¹⁵ Na spletni strani so dostopni in nastajajo v serijah vodiči in priročniki s področja digitalne hrambe na teme, kot so politike financiranj in ustanov, licenciranje podatkov, metapodatki ...

DCC je sodeloval pri številnih projektih, med njimi npr. pri britanskem projektu sodelovanja pri ustanavljanju mednarodnega repozitorija *Dryad*³¹⁶. Zelo zanimiv je tudi aktualni projekt *Digital Curator Vocational Education Europe*³¹⁷ – namenjen izobraževanju in usposabljanju za prihodnje poklice na področju ravnanja s podatki (v knjižnicah, muzejih, arhivih in drugih oddelkih za

³⁰⁹ Glej: <http://www.jisc.ac.uk/>, 12. 2. 2014.

³¹⁰ Glej: <http://www.surf.nl/en>, 12. 2. 2014.

³¹¹ Nekaj aktualnih tematik, ki so jih obravnavali, je vprašanje merjenja vpliva podatkov, modeli in pristopi k računanju stroškov podatkovne infrastrukture ter ločeno tudi načini financiranja, pa o načinih in zahtevah glede ravnanja s podatki ... (<http://www.knowledge-exchange.info/Default.aspx?ID=284>, 23. 1. 2014).

³¹² Glej: <http://www.dcc.ac.uk/>, 12. 2. 2014.

³¹³ Glej: <http://www.dcc.ac.uk/events/international-digital-curation-conference-idcc>, 12. 2. 2014.

³¹⁴ Izpostavljene teme: licenciranje, vrednotenje, politike, citiranje, načrt za ravnanje s podatki (<http://www.dcc.ac.uk/resources/how-guides>, 7. 8. 2013).

³¹⁵ Zelo zanimivo je orodje DMP Tool, ki deluje kot obrazec z usmeritvami, ki so prilagojene zahtevam področnega financiranja, <http://www.dcc.ac.uk/news/tools-data-management-planning>, 7. 8. 2013.

³¹⁶ Glej: <http://datadryad.org/>, 7. 8. 2013.

³¹⁷ Pri projektu poleg DCC sodeluje več centrov po Evropi, ki se ukvarjajo z vprašanji razvoja in širjenja znanja s področja digitalne hrambe na akademskem področju, kot so Digital Preservation Coalition (DPC), *Nestor* z nemškega govornega področja, italijanski Digital Renaissance Foundation – New Technologies for Cultural Heritage in drugi.

kulturno dediščino), saj je tovrstno usposabljanje med poudarki večine pobud za ureditev področja odprtega dostopa do podatkov. Njihovi projekti so še CARDIO – orodje za razvoj strategij ravnanja s podatki na ravni raziskovalnih oddelkov oziroma skupin,³¹⁸ PREPARDE – pozornost posveča recenziranim podatkovnim revijam s področja geoznanosti³¹⁹ in drugi.

Med združenji velja še enkrat omeniti tudi Digital Preservation Coalition (**DPC**).³²⁰ Člani tega so organizacije, ki vključujejo strokovnjake z različnih področij digitalne hrambe. Ti izdajajo vodiče pod naslovom *Technology Watch Reports*.³²¹ V letu 2013 so izšli vodiči na teme metapodatkov za hrambo, uporabe PDF formata za podatkovno hrambo ter shranjevanje CAD datotek (računalniško podprto načrtovanje). Slednjega je napisal Alex Ball, eden od vodilnih strokovnjakov na področju hrambe.³²²

DataCite

[DataCite](#)³²³ so leta 2009 ustanovili [British Library](#),³²⁴ danski [Technical Information Center](#),³²⁵ nizozemski [TU Delft Library](#),³²⁶ kanadski [Canada Institute for Scientific and Technical Information](#) (NRC-CISTI),³²⁷ [California Digital Library](#),³²⁸ [Purdue University](#)³²⁹ in nemška [National Library of Science and Technology](#).³³⁰ Osnovni cilji mednarodne organizacije so: vzpostaviti lažji dostop do raziskovalnih podatkov, spodbujati, da raziskovalni podatki postanejo ovrednoten del znanstvenih prispevkov, podpirati arhiviranje podatkov, ki obsega verifikacijo rezultatov in možnost druge rabe za prihodnje raziskave. V tehničnem smislu svojim uporabnikom ponujajo generiranje trajnih identifikatorjev, povezanih s podatki znotraj sistema DOI, s pomočjo katerega je mogoče slediti drugi rabi posamezne raziskave. Zagotavljajo registracijo metapodatkov, povezanih s posameznim DOI, in storitve na tej podlagi, kot so iskanje, agregiranje, oblikovanje izpisa podatkovnega citata ipd. Skrbijo pa tudi za promocijo same teme, s tem ko odgovarjajo na vprašanja, zakaj in kako citirati raziskovalne podatke.³³¹ Združenje običajno deluje tudi preko nacionalnih izpostav, ki skrbijo za registracijo oznak DOI za raziskovalne podatke v nacionalnih okvirih. Povezani so tudi z iniciativama [DataBib](#)³³² in [re3data.org](#).³³³

Research Data Alliance

Seveda pa ne moremo izpustiti delovanja Research Data Alliance (RDA), ki obeta, da postane forum za sodelovanja posameznikov iz različnih disciplinarnih okolij v svetovnem merilu. Organiziran je na način delovnih skupin, ki trenutno oblikujejo načrte dejavnosti za naslednje

³¹⁸ *Collaborative Assessment of Research Data Infrastructure and Objectives*, <http://www.dcc.ac.uk/projects/cardio>, 7. 8. 2013.

³¹⁹ *Peer REview for Publication & Accreditation of Research Data in the Earth Sciences*, <http://www.dcc.ac.uk/preparde/preparde#sthash.gLWCFI0t.dpuf>, 7. 8. 2013.

³²⁰ Glej: <http://www.dpconline.org/>, 12. 2. 2014.

³²¹ Glej: <http://www.dpconline.org/advice/technology-watch-reports>, 12. 2. 2014.

³²² Glej: <http://www.dpconline.org/newsroom/latest-news/1039-preserving-cad-new-tech-watch-report-enables-continuous-access-to-design?q=alex+ball>, 2. 12. 2013.

³²³ Glej: <http://www.datacite.org/>, 12. 2. 2014.

³²⁴ Glej: <http://www.datacite.org/britishlibrary>, 12. 2. 2014.

³²⁵ Glej: <http://www.dtic.dtu.dk/english>, 12. 2. 2014.

³²⁶ Glej: <http://www.datacite.org/tudelft>, 12. 2. 2014.

³²⁷ Glej: http://www.datacite.org/CISTI_eng, 12. 2. 2014.

³²⁸ Glej: <http://www.datacite.org/DataCiteUS>, 12. 2. 2014.

³²⁹ Glej: <http://www.datacite.org/purdue>, 12. 2. 2014.

³³⁰ Glej: <http://www.uni-hannover.de/en/universitaet/organisation/einrichtung/tibub/index.php>, 12. 2. 2014.

³³¹ Glej: <http://www.datacite.org/whatdowedo>, 7. 8. 2013.

³³² Katalog raziskovalnih podatkovnih baz.

³³³ Nemška pobuda s ciljem oblikovati globalni register raziskovalnih podatkovnih repozitorijev.

obdobje, ter se sofinancira s strani Evrope, ZDA in Avstralije. Sama organizacija izhaja iz več predhodnih projektov (npr. *The Data Access and Interoperability Task Force* – [DAITF](#))³³⁴ in predstavlja združevanje moči na globalni ravni.³³⁵

Serijske publikacije

Leta 2012 je pod okriljem [EDUCOPIA Institute](#)³³⁶ izšel zbornik z naslovom *Aligning National Approaches to Digital Preservation*,³³⁷ katerega namen je razširiti in poglobiti poglede, ki so si jih avtorji izmenjali na konferenci v Talinu leta 2011, in sicer s ciljem večjega povezovanja med pobudami na področju digitalnega ohranjanja, še posebej na področju ohranjanja kolektivnega kulturnega spomina. Leta 2013 so kot nadaljevanje začetega dela priredili [konferenco](#)³³⁸ v Barceloni z istim naslovom ter izpostavili pomen usklajevanja delovanja skupnosti, virov in usposobljenosti za digitalno ohranjanje. V tej pobudi je posebej izpostavljeno mednarodno sodelovanje in uskladitev delovanja različnih deležnikov, kot taka pa sodi med tiste, pri katerih bi bilo potrebno nadalje sodelovati ob izpopolnjevanju skrbi za digitalna gradiva nacionalnega pomena.

Obstaja vrsta revij, ki objavljajo specializirane objave na temo raziskovalnih podatkov. Poleg že omenjenih posameznih člankov v revijah *Nature* in *PLoS* je vredno posebej izpostaviti [International Journal of Digital Curation](#) (IJDC)³³⁹ in pa [Data Science Journal](#).³⁴⁰ Te moramo ločevati od podatkovnih revij, zvrsti, ki jo obravnavamo na drugem mestu, kot je npr. [Scientific Data](#)³⁴¹ ali pa [Earth System Science Data](#).³⁴² Na DCC imajo pregled strokovnih revij, ki vključujejo področje ravnanja z raziskovalnimi podatki.³⁴³

Revija *IJDC* je na primer eno celotno številko (let. 8, št. 2 (2013)) namenila vprašanjem načrtovanja ravnanja z raziskovalnimi podatki. Naj izpostavimo le nekaj obravnavanih tem, ki so splošnega pomena: ravnanje s podatki in načrtovanje ohranjanja za t. i. velike znanosti, ravnanje z občutljivimi podatki, zaupanje v digitalne repozitorije, standardizacija, povezovanje podatkov in objav, standardni identifikatorji in drugo.³⁴⁴

³³⁴ Glej: <http://www.daitf.org/>, 12. 2. 2014.

³³⁵ Glej: <http://www.rd-alliance.org/> za delovne skupine in dokumente, ki jih obravnavajo.

³³⁶ Glej: <http://educopia.org/>, 12. 2. 2014.

³³⁷ Glej: <http://educopia.org/publications/ANADP>, 12. 2. 2014.

³³⁸ Glej: <http://educopia.org/events/ANADPII>, 12. 2. 2014.

³³⁹ Glej: <http://www.ijdc.net/>, 12. 2. 2014.

³⁴⁰ Glej: <http://www.codata.org/dsj/>, 12. 2. 2014.

³⁴¹ Glej: <http://www.nature.com/scientificdata/>, 12. 2. 2014.

³⁴² Glej: <http://www.earth-system-science-data.net/>, 12. 2. 2014.

³⁴³ Glej: <http://www.dcc.ac.uk/resources/curation-journals>, 3. 12. 2013.

³⁴⁴ Glej: <http://www.ijdc.net/index.php/ijdc/issue/view/16>, 3. 12. 2013.

PREGLED PO DRŽAVAH

Omenili smo že integrativno vlogo pobud EK pri harmonizaciji pristopov glede politik in uresničevanja zavez držav članic. Med zgoraj naštetimi projekti in organizacijami jih je tudi mnogo takih, ki delujejo na podlagi partnerstva vodilnih strokovnih organizacij ali predstavnikov vlad z namenom približevanja in izmenjave dobrih praks, kot je skupina Knowledge Exchange. V nadaljevanju predstavljamo nekatere poudarke nacionalnih ali regionalnih pristopov pri odpiranju raziskovalnih podatkov.

Velika Britanija

Velika Britanija je v marsikaterem pogledu vodilna pri vzpostavljanju okolja odprtega dostopa do raziskovalnih podatkov. Po tradiciji imajo razvito infrastrukturno podporo in podporo ustanov. Financerji so v svojih politikah jasno določili zahteve glede raziskovalnih podatkov kot javnega dobrega, razširjeno je svetovanje in pomoč pri izpolnjevanju zahtev glede priprave načrtov ravnanja z raziskovalnimi podatki, poteka vrsta pilotnih projektov za inovativne pristope in orodja, sledijo tudi jasno zastavljenim strateškim usmeritvam.

Strategije razvoja pri različnih segmentih obravnave raziskovalnih podatkov so zajete v večjem številu študij, od katerih kot posebej odmevno omenjamo poročilo z naslovom *Stewardship of Digital Research Data: A Framework of Principles and Guidelines: Responsibilities of Research Institutions and Funders, Data Managers, Learned Societies and Publishers*,³⁴⁵ ki upošteva vidik ustvarjalcev podatkov, uporabnikov podatkov in financerjev raziskav.³⁴⁶ Research Information Network (RIN),³⁴⁷ ki je to poročilo izdal, izpostavlja pomanjkanje koordinacije pri ravnanju z raziskovalnimi podatki in na področju zagotavljanja dostopa. To, podobno kot poročilo National Science Board (NSB) iz ZDA z naslovom *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*,³⁴⁸ uvaja model vlog in odgovornosti različnih poklicev in med sabo povezanega okolja različnih ustanov podpore deljenju podatkov. Pogosto se navaja tudi poročilo *Roles and responsibilities, The Skills, Role and Career Structure of Data Scientists and Curators: An Assessment of Current Practice and Future Needs, Report to the JISC*,³⁴⁹ ki sta ga pripravila Alma Swan in Sheridan Brown. Slednje izpostavlja problem izobraževanja za podatkovne poklice in pa karierne možnosti.

Med britanskimi združenji in financerji, ki podpirajo projekte sodelovanja pri razvoju na področju podatkovne infrastrukture, omenjamo tudi Joint Information Systems Committee (JISC).³⁵⁰ V svoji strategiji 2007–2009 so zapisali, da je prvi cilj doseči izvirno in vzdržno informacijsko in komunikacijsko tehnologijo, infrastrukturo, storitve in prakse, ki podpirajo

³⁴⁵ Glej:

http://www.google.si/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CCwQFjAA&url=http%3A%2F%2Fwww.rin.ac.uk%2Fsystem%2Ffiles%2Fattachments%2FStewardship-data-guidelines.pdf&ei=WnrOUsvyNqWgyQPO7YHACQ&usg=AFQjCNF0TO3ZHQP0011CRIWJ4Z6-Ex1Cw&sig2=udc8_QLFhnE4JzSyjLcGbw&bvm=bv.59026428.d.bGQ, 12. 2. 2014.

³⁴⁶ Glej: <http://www.rin.ac.uk/system/files/attachments/Stewardship-data-guidelines.pdf>, 7. 8. 2013.

³⁴⁷ Glej: <http://www.rin.ac.uk/>, 12. 2. 2014.

³⁴⁸ Glej: <http://www.nsf.gov/pubs/2005/nsb0540/>, 7. 8. 2013.

³⁴⁹ Glej: <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf>, 12. 2. 2014.

³⁵⁰ Glej: <http://jisc.ac.uk/>, 12. 2. 2014.

ustanove pri uresničevanju njihovih poslanstev.³⁵¹ V ta okvir sodi serija projektov *Discovery*,³⁵² katere naloga je bila podpreti boljši dostop do pomembnejših podatkov v knjižnicah, muzejih in arhivih in zagotoviti nove storitve na področju raziskovanja in izobraževanja v Veliki Britaniji.³⁵³

Fundacija *JISC*³⁵⁴ znotraj širšega sklica *Digital infrastructure: Research management programme*³⁵⁵ že več let financira vrsto razvojnih projektov s področja ravnanja z raziskovalnimi podatki³⁵⁶. Cilj enega od aktualnih projektov je, da olajša iskanje in dostop do obstoječih zbirk raziskovalnih podatkov v britanskih raziskovalnih ustanovah in področnih podatkovnih središčih. Šestmesečni pilotni projekt partnerjev iz DCC in UK DA bo pokazal možnosti agregiranja metapodatkovnih opisov raziskav in dostopa z enega mesta.³⁵⁷

JISC je v svojih politikah zapisal, da je ena od prioritet promocija in podpora kakovostnemu ravnanju z raziskovalnimi podatki, kamor sodi tudi izobraževanje in usposabljanje za delo s podatki. Sredstva za to namenja npr. storitvam pod imenom *Jorum*.³⁵⁸ Britanski Jorum je namenjen zbiranju in deljenju odprtih izobraževalnih virov in omogoča njihovo drugo rabo. Na spletni strani so objavili gradivo za promocijo načrtovanja ravnanja z raziskovalnimi podatki.³⁵⁹ Splošno uvajanje dela s podatki v izobraževalne programe in standarde s področja informacijske pismenosti poudarjajo tudi v Researcher Development Framework (*RDF*),³⁶⁰ ki ga razvija Vitae. Research Information Network (*RIN*)³⁶¹ skupaj s podporo *JISC*³⁶² razvija priročnike in vodiče ter usposabljanja skozi Data Management Skills and Support Initiative (*DaMSSI*).³⁶³ Obstaja še vrsta drugih projektov, ki so namenjeni usposabljanju knjižničarjev in drugih podpornih storitev, npr. pri ARIADNE: *Upskilling Liaison Librarians for Research Data Management*,³⁶⁴ ali National Information Standards Organization (*NISO*):³⁶⁵ *Webinar: Research Data Curation, Part 2: Libraries and Big Data*,³⁶⁶ ali *MANTRA*:³⁶⁷ spletni tečaj za ravnanje z digitalnimi podatki v raziskovalnem procesu.

V Veliki Britaniji imajo tudi razvejano okolje področnih podatkovnih središč ter zahtev in politik področnih financierjev znanosti. V Research Council UK (*RCUK*)³⁶⁸ so določili skupna načela, ki določajo obveznosti raziskovalnih ustanov in raziskovalcev pri izročanju raziskovalnih podatkov. Ta so razvidna iz politik, ki se nanašajo na ravnanje s podatki, in predstavljajo širši, skupni okvir vsakega posameznega raziskovalnega sveta oz. njegove politike ravnanja s podatki.

³⁵¹ *Managing Research Data*, <http://www.jisc.ac.uk/whatwedo/programmes/mrd.aspx>, 12. 12. 2013.

³⁵² Glej: <http://discovery.ac.uk/>, 12. 2. 2014.

³⁵³ Glej: <http://discovery.ac.uk/>, 23. 1. 2014.

³⁵⁴ Glej: http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement/managingresearchdata.aspx, 12. 2. 2014.

³⁵⁵ Glej: http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement.aspx, 12. 2. 2014.

³⁵⁶ Glej: <http://www.jisc.ac.uk/whatwedo/programmes/umf.aspx>, 3.12.2013.

³⁵⁷ *Scoping approaches to a UK research data registry (DCC blog)*, <http://www.dcc.ac.uk/blog/scoping-approaches-uk-research-data-registry>, 23. 1. 2014.

³⁵⁸ Glej: <http://www.jorum.ac.uk/about-us/>, 12. 2. 2014.

³⁵⁹ Glej: <http://dspace.jorum.ac.uk/xmlui/handle/10949/18267/browse?type=dateissued>, 2. 12. 2013.

³⁶⁰ Glej: <http://www.vitae.ac.uk/researchers/428241/Researcher-Development-Framework.html>, 12. 2. 2014.

³⁶¹ Glej: <http://www.rin.ac.uk/>, 12. 2. 2014.

³⁶² Glej: <http://www.jisc.ac.uk/>, 12. 2. 2014.

³⁶³ Glej: <http://www.rin.ac.uk/our-work/researcher-development-and-skills/data-management-and-information-literacy>, 12. 2. 2014.

³⁶⁴ Glej: <http://www.ariadne.ac.uk/issue70/cox-et-al>, 12. 2. 2014.

³⁶⁵ Glej: <http://www.niso.org/about/>, 12. 2. 2014.

³⁶⁶ Glej: http://www.niso.org/news/events/2013/webinars/data_curation, 12. 2. 2014.

³⁶⁷ Glej: <http://datalib.edina.ac.uk/mantra/>, 12. 2. 2014.

³⁶⁸ Glej: <http://www.rcuk.ac.uk/pages/home.aspx>, 7. 8. 2013.

Načela imajo izjemen pomen pri uskladitvi politik ločenih področnih raziskovalnih svetov in jih kot njihovo skupno krovno usmeritev pogosto navajajo. Načela so zastavljena, kakor sledi:

- javno financirani raziskovalni podatki so javno dobro, ustvarjeni v javnem interesu in naj bodo kot taki odprto dostopni, s čim manj omejitvami;
- institucionalne politike ravnanja s podatki in tiste, vezane na posamezne projekte, naj bodo skladne z ustreznimi standardi in dobrimi praksami skupnosti, tako naj bodo podatki s prepoznano dolgoročno vrednostjo dostopni in uporabni tudi za nadaljnje raziskave;
- podatki naj bodo ustrezno opremljeni z metapodatki in odprto dostopni raziskovalcem, da jih bodo ti lahko poiskali in do njih dostopali za svoje raziskave;
- RCUK upošteva, da obstajajo pravne, etične in komercialne prepreke, ki jih je potrebno premisliti pred oddajo podatkov, in priporoča, da se vnaprej že pri načrtovanju ustvarjanja podatkov upoštevajo raziskovalne politike in prakse, ki rešujejo tovrstne probleme;
- RCUK ustvarjalcem podatkov zagotavlja pravico, da lahko za določeno obdobje uporabljajo podatke zase, preden gredo ti v odprti dostop – na ta način RCUK prizna napor, vložen v ustvarjanje podatkov; trajanje obdobja se razlikuje po raziskovalnih področjih, kjer je potrebno, se v politikah posameznih svetov dodatno določijo posebnosti;
- da bi dosegli priznanje intelektualnega napora raziskovalcev, ki ustvarjajo, hranijo in delijo raziskovalne podatke, morajo vsi raziskovalci uporabniki navajati vir podatkov in upoštevati določila in pogoje, pod katerimi so podatki dostopni;
- za izvajanje nalog pri ravnanju s podatki in pri deljenju podatkov, financiranih z javnimi sredstvi, je dovoljeno, da se uporabijo javna sredstva. Pri ravnanju s sredstvi, namenjenimi za podatkovne storitve, je potrebno zagotoviti vzdržnost, da se z omejenimi sredstvi doseže maksimalne koristi za raziskovanje.³⁶⁹

Poudarki so torej, da so podatki načeloma javno dobro, upoštevaje pravne in druge omejitve, vendar je potrebno po najboljših močeh zagotoviti odprti dostop ter najbolje že v fazi načrtovanja predvideti različne ovire in omejitve in pripraviti načine njihovega premagovanja. Poudarek je na pripravi načrtov ravnanja z raziskovalnimi podatki in na ustrezni motivaciji ustvarjalcev ter zagotavljanju vzdržne infrastrukture skrbnikov podatkov.

Večina javnih financerjev v Veliki Britaniji je ta določila že vključila v svoje področne politike (Jones 2009, 4).³⁷⁰ Med zgledi že dolgo uveljavljenih in doslednih so politike Economic and Social Research Council ([ESRC](#)),³⁷¹ Medical Research Council ([MRC](#))³⁷² in Science of the Environment ([NERC](#)).³⁷³ Na področju naravoslovja je usmeritev v doseganje ciljev po poteku prehodnega obdobja, in sicer so v vmesnem času raziskovalne ustanove zavezane, da uredijo svoje notranje politike in postopke tako, da bodo izpolnjene krovne zahteve.

³⁶⁹ Glej: <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>, 7. 8. 2013.

³⁷⁰ *A report on the range of policies required for and related to digital curation*, http://www.dcc.ac.uk/sites/default/files/documents/reports/DCC_Curation_Policies_Report.pdf, 12. 12. 2013.

³⁷¹ Glej: <http://www.esrc.ac.uk/>, 7. 8. 2013.

³⁷² Glej: <http://www.mrc.ac.uk/index.htm>, 12. 2. 2014.

³⁷³ Glej: <http://www.nerc.ac.uk/>, 12. 2. 2014.

V Royal Society's Science Policy Centre so se specializirali za svetovanje s področja politik, njihovi nasveti so namenjeni odločevalcem v Veliki Britaniji, Evropi in drugod. Glede na to, da se opirajo na bogate izkušnje in razvejano dejavnost v Veliki Britaniji, ni presenečenje, da tudi njihovo poročilo z naslovom *Science as an open enterprise*³⁷⁴ koncizno osvetljuje potrebo po ohranjanju načel odprtosti in uporabi podatkov na načine, ki lahko povzročijo revolucijo odprte znanosti. V poročilu uporabljajo pojem »razumljive odprtosti«: take, da vključuje iskanje in znano dostopnost, razumljivost za zunanjo presojo, podatki morajo biti tudi pregledni z vidika zanesljivosti, vključevati morajo zagotovila o usposobljenosti ustvarjalcev. Se pravi, da morajo biti podatki objavljeni s spremljajočimi izčrpnimi metapodatki.

Izpostavljajo tudi integriteto in poštenost raziskovanja, saj se na ta način lažje odkrije, kar je v znanosti nekakovostno. Predvidevajo, da se bodo cilji dosegli z uvajanjem naslednjih sprememb:

- proč od mišljenja, da so podatki zasebni kotiček,
- širitev kriterijev za ovrednotenje znanstvenega prispevka tako, da vključujejo uporabne odprto dostopne podatke,
- razvoj skupnih standardov,
- povezava razumljive odprtosti v povezavi s publikacijami,
- vzpostaviti novo generacijo podatkovnih znanstvenikov, sposobnih ravnanja in nudenja podpore pri uporabi podatkov,
- razvoj novih orodij za pripravo in izrabo podatkov.³⁷⁵

Poročilo vsebuje napredna priporočila skupine zelo uglednih predstavnikov ustanov in disciplin, ki so pazljivo formulirana okrog posameznih poudarkov. In sicer, da naj financerji znanosti poskrbijo za prenos raziskovalnih podatkov s projektov, ki jih podpirajo, s tem da prepoznajo prispevek tistih, ki na najboljši način skrbijo za uporabnost in prenos podatkov, da pokrivajo stroške priprave podatkov in metapodatkov v okviru projekta in da zagotavljajo sodelovanje ostalih pri trajnosti dostopa do datotek.

Vselej so dopuščene izjeme, npr. kadar gre za industrijsko izkoriščanje patentov. Po drugi strani je jasno, da je potrebno spodbujati makroekonomske učinke večje odprtosti in širšega izkoriščanja razpoložljivih informacij. Zato spodbujajo razmislek o večji medsebojni odprtosti in preglednosti podatkov tako javnega kot zasebnega sektorja, kar lahko privede do večjih javnih koristi in izboljša ekonomsko učinkovitost. Pomembno je upoštevati izjeme glede popolne odprtosti dostopa tudi z vidika varovanja zasebnosti, kjer pristopi anonimizacije ne rešijo vseh problemov, so pa pogoji in načini dostopa do podatkov lahko usklajeni z naravo podatkov in nameni uporabe, tako da se kljub potrebni regulaciji dostop do podatkov ne zapre povsem, kakor tudi v drugih primerih izjem zaradi varovanja koristi države ali skupin.

Med poudarki so tudi naslednji nasveti: da bi dosegli spremembe, potrebujemo primere, ki jasno kažejo na koristi, s tem se krepi podpora za širjenje kroga ljudi, ki se ukvarjajo s podatki, in za potrebna vlaganja v podporna orodja. Pristop tega poročila je zelo koristen tudi za države, ki šele vstopajo v pripravo strategij na področju odprtega dostopa. Za začetek je potrebno razširiti razumevanje o koristih in prepričanje glede potrebnih korakov, vključno z obveznostmi razumljivega objavljanja podatkov ob publikaciji itd. (*Science as an open enterprise* 2012).

³⁷⁴ Glej: <http://royalsociety.org/policy/projects/science-public-enterprise/report/>, 12. 2. 2014.

³⁷⁵ Glej: <http://royalsociety.org/policy/projects/science-public-enterprise/report/>, 7. 8. 2013.

Razna poročila navajajo vrsto primerov, kako se vlaganja v podatkovno infrastrukturo in njeno vzdrževanje večkratno povrnejo. V poročilu *Proposal and Business Plan for the Initial Pathfinder Development* (2010) je poudarjeno, da podatkovno infrastrukturo omogočimo tako, da spodbudimo raziskovalce, da naredijo svoje raziskovalne podatke dostopne in uporabne. Poskrbeti je potrebno za usposabljanje raziskovalcev, da bodo pridobili veščine ravnanja s podatki. Poskrbeti pa je potrebno tudi za financiranje in za infrastrukturo, kar omogoči skupne podatkovne in podporne storitve, ki predstavljajo hrbtenico podatkovne infrastrukture. Poročilo vsebuje še vrsto podobnih nasvetov, kako izgraditi učinkovito podatkovno infrastrukturo na področjih, ki so manj razvita.

V Veliki Britaniji, ki je zelo aktivna na področju promocije odprtih podatkov, se poleg splošnih pojavljajo že specifične zahteve, ki vključujejo individualne financerje, pa tudi spodbude posameznim ustanovam, naj uredijo svoje politike. Na precejšnjo aktivnost Velike Britanije na tem področju kaže pregled ustanov, ki so politike ravnanja s podatki uveljavile v zadnjih letih. Ustanove so politike sprejele med letoma 2009 in 2013, med seboj pa se razlikujejo po tem, kako poglobljene so te politike.³⁷⁶ Posebej aktualen je pristop, ki ga je ubral Engineering and Physical Sciences Research Council ([EPSRC](http://www.epsrc.ac.uk/)),³⁷⁷ ki je zadalžil institucije, da pripravijo svoje notranje politike in akcijske načrte za njihovo uveljavitev. Kot posledica tega poziva sedaj narašča število univerz v procesu preverjanja in dopolnjevanja prvih osnutkov politik in načrtov. Nekatere, npr. Univerza v Edinburgu, politike ravnanja s podatki predstavljajo v nekaj kratkih točkah. Te večinoma pojasnijo smisel in pomen politik, opredelitev raziskovalnih podatkov in zahteve v zvezi z ravnanjem s podatki. Svojo politiko in akcijski načrt utemeljujejo na nekaterih pilotnih projektih. Eden takih je *CERIF for Datasets* ([C4D](http://www.cerif.ac.uk/)),³⁷⁸ projekt, ki ga financira JISC in ki preizkuša razširitev in integracijo standarda Cerif na raziskovalne podatke. Uporabljajo nekatera druga prosto dostopna orodja, kot je [IRODS](http://www.irods.org/).³⁷⁹ Univerza v Glasgowu v zadnjih letih izvaja večje število projektov (usposabljanje, vodstvo in podpora), s ciljem razviti pogoje za izvajanje ravnanja s podatki in oblikovati politike ter tehnično infrastrukturo za izvajanje načrta ravnanja s podatki.³⁸⁰ Nekatere druge so natančneje izdelane in poleg splošnih definicij in usmeritev opredeljujejo še odgovornosti, lastništvo, hrambo in skrbništvo, odpis in uničenje, dostop ter drugo (npr. Univerza v Southamptonu). Seznam univerz iz Velike Britanije, ki so sprejele politike ravnanja z raziskovalnimi podatki, se začneja z Univerzo v Edinburgu, ki je bila med zgodnejšimi, z začetki že leta 2011.³⁸¹

Posebej je potrebno izpostaviti, da ima Velika Britanija razvito podporo v obliki storitev področnih podatkovnih središč. Večkrat smo že omenjali Economic and Social Data Service ([ESDS](http://www.esds.ac.uk/)),³⁸² [NERC](http://www.nerc.ac.uk/),³⁸³ British Atmospheric Data Centre ([BADC](http://www.badc.nerc.ac.uk/)),³⁸⁴ na področju ostalih znanosti pa

³⁷⁶ Politike posameznih raziskovalnih svetov so na voljo tukaj: <http://www.dcc.ac.uk/resources/policy-and-legal/funders-data-policies>, 7. 8. 2013.

³⁷⁷ Glej: <http://www.epsrc.ac.uk/Pages/default.aspx>, 12. 2. 2014.

³⁷⁸ Glej: http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement/managingresearchdata/infrastructure/c4d.aspx, 12. 2. 2014.

³⁷⁹ Glej: https://www.irods.org/index.php/IRODS:Data_Grids,_Digital_Libraries,_Persistent_Archives,_and_Real-time_Data_Systems, 12. 2. 2014.

³⁸⁰ Glej: <http://www.gla.ac.uk/services/datamanagement/rdm-at-gu/>, 7. 8. 2013.

³⁸¹ Glej: <http://www.ed.ac.uk/schools-departments/information-services/about/policies-and-regulations/research-data-policy>, 7. 8. 2013.

³⁸² Glej: <http://www.esds.ac.uk/>, 12. 2. 2014.

³⁸³ Glej: <http://www.nerc.ac.uk/>, 12. 2. 2014.

³⁸⁴ Glej: <http://badc.nerc.ac.uk/home/index.html>, 12. 2. 2014.

delujejo: Archaeology Data Service ([ADS](http://archaeologydataservice.ac.uk/)),³⁸⁵ Chemical Database Service ([CDS](http://cds.dl.ac.uk/)),³⁸⁶ British Oceanographic Data Centre ([BODC](http://www.bodc.ac.uk/)),³⁸⁷ Environmental Information Data Centre ([EIDC](http://www.ceh.ac.uk/sci_programmes/env_info.html)),³⁸⁸ European Bioinformatics Institute ([EBI](http://www.ebi.ac.uk/)),³⁸⁹ National Cancer Data Repository ([NCDR](http://www.ncin.org.uk/collecting_and_using_data/national_cancer_data_repository/)),³⁹⁰ National Geoscience Data Centre ([NGDC](http://www.bgs.ac.uk/services/ngdc/home.html)),³⁹¹ UK Solar System Data Centre ([UKSSDC](http://www.ukssdc.ac.uk/))³⁹² in nekateri drugi.

Napredna vloga Velike Britanije pri širjenju dobrih praks pri odprtem dostopu do raziskovalnih podatkov je vidna tudi z delovanjem v okviru skupine partnerskih držav, vključenih v že omenjeno povezavo Knowledge Exchange.

Nemčija

V Nemčiji je več raziskovalnih ustanov, združenih v [Alliance of German Science Organisations](https://www.allianzinitiative.de/en/core_activities/research_data/principles/),³⁹³ objavilo skupna načela, podobno kot jih RCUK promovira v britanski raziskovalni skupnosti. Načela nosijo naslov [Principles for the Handling of Research Data](http://www.allianzinitiative.de/en/core_activities/research_data/principles/)³⁹⁴ in posebej poudarjajo, da je potrebno upoštevati razlike med znanstvenimi disciplinami, opredeliti in priznati vrednost prispevanja raziskovalnih podatkov v odprti dostop ter spodbuditi razvoj podporne infrastrukture. Med ustanovami podpisnicami načel so Deutsche Forschungsgemeinschaft (DFG), partner KE, Alexander von Humboldt Foundation, German Academy of Sciences Leopoldina, German Academic Exchange Service (DAAD), Fraunhofer-Gesellschaft, Helmholtz Association, German Rectors' Conference (Hochschulrektorenkonferenz – HRK), Leibniz Association, Max Planck Society, Wissenschaftsrat (German Council of Science and Humanities).

DFG je leta 2009 objavil priporočila z naslovom [Recommendations for Secure Storage and Availability of Digital Primary Research Data](http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf).³⁹⁵ Dokument priporoča upoštevanje tehničnih in organizacijskih pogojev, ki so pomembni za ohranjanje primarnih podatkov in jih morajo področja individualno razviti. Predlaga oblikovanje delovnih skupin po področjih, ki bodo pripravile predloge, kdo in kako bi ohranjal podatkovno gradivo. Izpostavlja tudi pomen citiranja podatkov in samoiniciativnega deljenja podatkov s strani raziskovalcev ter pomen kontrole kakovosti podatkov. Oblikovali so sedem priporočil, kakor sledi:

1. Primarni raziskovalni podatki so rezultati, ki nastanejo v znanstveni raziskavi z eksperimenti, meritvami, raziskavo ali anketo in so osnova za znanstveno objavo.
2. Opredeliti je potrebno področno specifičen organizacijski koncept, ki uravnava učinkovito in vzdržno ohranjanje podatkov.
3. Raziskovalni podatki se ohranjajo skladno z opredeljenimi standardi.
4. Podatki so označeni z osebnim imenom in shranjeni z imenom raziskovalca.
5. Če je le mogoče, vsak znanstvenik poskrbi, da so njegovi raziskovalni podatki prosto na voljo na medpodročni ravni.
6. Podatki so opisani z metapodatki.

³⁸⁵ Glej: <http://archaeologydataservice.ac.uk/>, 12. 2. 2014.

³⁸⁶ Glej: <http://cds.dl.ac.uk/>, 12. 2. 2014.

³⁸⁷ Glej: <http://www.bodc.ac.uk/>, 12. 2. 2014.

³⁸⁸ Glej: http://www.ceh.ac.uk/sci_programmes/env_info.html, 12. 2. 2014.

³⁸⁹ Glej: <http://www.ebi.ac.uk/>, 12. 2. 2014.

³⁹⁰ Glej: http://www.ncin.org.uk/collecting_and_using_data/national_cancer_data_repository/, 12. 2. 2014.

³⁹¹ Glej: <http://www.bgs.ac.uk/services/ngdc/home.html>, 12. 2. 2014.

³⁹² Glej: <http://www.ukssdc.ac.uk/>, 12. 2. 2014.

³⁹³ Glej: <https://www.humboldt-foundation.de/web/press-release-2013-03.html>, 12. 2. 2014.

³⁹⁴ Glej: http://www.allianzinitiative.de/en/core_activities/research_data/principles/, 12. 2. 2014.

³⁹⁵ Glej: http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf, 12. 2. 2014.

7. Kakovost podatkov je preverjena. ([Recommendations for Secure Storage and Availability of Digital Primary Research Data](#) 2009, 2–3)³⁹⁶

Z vidika DFG je za uspeh koordiniranega sistema informacijske infrastrukture zlasti pomembno tesno sodelovanje deležnikov, da se zagotovi izpolnitev interesov raziskovalcev, potrebno je tudi razmišljati o interoperabilnosti in integraciji z mednarodnimi strukturami, da se presežejo regionalne in državne meje. Tovrstna načela so predstavili v strategiji [Taking Digital Transformation to the Next Level, The Contribution of the DFG to an Innovative Information Infrastructure for Research](#) (2012, 2–3).³⁹⁷ Omeniti velja tudi zaključno poročilo projekta RADIESCHEN (RADISH) z naslovom [Final report of the project and roadmap for the development of a research data infrastructure in Germany](#),³⁹⁸ ki govori o pogojih za vzpostavitev interdisciplinarne raziskovalne podatkovne infrastrukture.

Kot primer pilotnih projektov na področju informacijske infrastrukture izpostavljamo tudi novi nemški program, financiran s strani DFG, imenovan [Information Infrastructures for Research Data](#).³⁹⁹ Program bo pomagal identificirati in vpeljati posebne zahteve za prihodnje strukture, ki bodo pomagale pri ravnanju z raziskovalnimi podatki in pri delu raziskovalnih podatkovnih repozitorijev. Pri financiranju bodo imeli prednost predvsem projekti, ki razvijajo koncepte in rešitve za čezregijske in vzdržne celovite informacijske infrastrukture na enem področju ali medpodročno.

Nemški podatkovni forum Rat für Sozial und Wirtschaftsdaten ([RatSWD](#))⁴⁰⁰ je leta 2004 ustanovilo ministrstvo za izobraževanje in znanost. Deluje kot neodvisno telo, ki ga tvorijo raziskovalci z univerz, kolidžev in drugih neodvisnih raziskovalnih ustanov ter predstavniki pomembnejših ustvarjalcev podatkov. Cilj RatSWD je izboljšati raziskovalno infrastrukturo in tako prispevati h konkurenčnosti na mednarodni ravni. Do sedaj jim že uspelo odpreti in izboljšati dostop do obstoječih podatkov in izboljšati sinergije med znanostjo in ustvarjalci podatkov.⁴⁰¹

Na zrelost razumevanja ravnanja s podatki kaže nedavno izdani zbornik z naslovom [Handbuch Forschungsdatenmanagement](#),⁴⁰² ki obsega šestnajst prispevkov. V njem avtorji razpravljajo o življenjskem krogu podatkov, politikah, vezanih na podatke, načrtih ravnanja z raziskovalnimi podatki, metapodatkih in standardih, podatkovnih analizah, o vizualizaciji, arhiviranju raziskovalnih podatkov in drugem.

³⁹⁶ Glej: http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf, 12. 2. 2014.

³⁹⁷ Glej: http://www.dfg.de/download/pdf/foerderung/programme/lis/strategy_paper_digital_transformation.pdf, 12. 2. 2014.

³⁹⁸ Glej: http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:124265:3/component/escidoc:124264/ProjectRadieschen_Synthesis_EN.pdf?mode=download, 12. 2. 2014.

³⁹⁹ Glej: http://www.dfg.de/en/research_funding/announcements_proposals/archive/2013/info_wissenschaft_13_19/index.html, 12. 2. 2014.

⁴⁰⁰ Glej: <http://ratswd.de/en/ratswd/members>, 12. 2. 2014.

⁴⁰¹ *Tasks and Structure*, <http://ratswd.de/en/ratswd/tasks-structures>, 3. 12. 2013.

⁴⁰² Glej: http://www.forschungsdatenmanagement.de/?page_id=2, 11. 2. 2014

Nordijske države

Ambiciozno je zastavljeno tudi poročilo in načrt prihodnjega bolj intenzivnega sodelovanja [Reinforced Nordic collaboration on data resources](#).⁴⁰³ Poročilo omenja vrsto iniciativ na Finskem, Danskem in na Švedskem, ki se ukvarjajo z vprašanji raziskovalnih podatkov in obravnavajo potrebne spremembe zakonodaje s področja varovanja podatkov zaradi direktive EU, dostop do registrskih podatkov, dolgotrajno hrambo in širitev obstoječega pokritja disciplinarnih podatkovnih središč. Skupina [NordForsk](#)⁴⁰⁴ si je zadala pripravo skupne nordijske raziskovalne politike in sodelovanja ter obenem koordinacijo nacionalnih strategij in politik. (Sandberg 2012, 23).

V spremembah zakonodaje s področja varovanja podatkov kot posledice uvajanja direktive EU vidijo priložnost, da se odpre pooblaščen dostop do podatkov za raziskovalce preko državnih meja. V nordijskih državah je že tradicionalno pomembna tudi usmerjevalna vloga etičnih komisij, ki zagotavljajo, da je ustvarjanje in dostop do podatkov skladno z zakonodajo.

Tudi na **Danskem** potekajo aktivnosti in posvetovanja na področju dostopa do raziskovalnih podatkov. Open Access Committee priporoča, da se ustanovi medpodročna organizacija, ki bo povezala različne akterje in jih spodbujala k sodelovanju na področju primarnih raziskovalnih podatkov. Predvsem pa mora odgovoriti na vprašanje, kako arhivirati danske raziskovalne podatke, da bo poskrbljeno za njihov sedanji in prihodnji obstoj.⁴⁰⁵ Leta 2011 so pripravili [Danish Research Infrastructure Roadmap](#).⁴⁰⁶ V njem govorijo o načinu financiranja, ki upošteva posebnosti podatkovnih infrastruktur, kot npr. večji konstantni stroški vzdrževanja delovanja na dolgi rok v primerjavi s stroški vzpostavitve. Poseben poudarek dajejo vprašanjem dolgotrajnega digitalnega shranjevanja. Ustanovili pa so tudi skupino za interdisciplinarno povezovanje, v kateri sodelujejo raziskovalci, knjižnice, podatkovni arhivi, univerze (Sandberg 2012, 36).

Finsko ministrstvo za izobraževanje in kulturo je leta 2009 oblikovalo državno medresorsko delovno skupino z namenom, da bo pripravila smernice za razvoj dostopa in hrambe podatkovnih virov za raziskovalne namene. Vključili so e-infrastrukturne rešitve, predloge, ki vključujejo prevzemanje odgovornosti, sodelovanja in koordinacije med različnimi akterji. Osnovni namen projekta je bil izoblikovati splošno razumevanje sedanje situacije na Finskem ter jo primerjati z mednarodnimi priporočili in ureditvami drugih držav. Delovna skupina je leta 2011 pripravila smernice za uporabo elektronskih virov pri raziskavah. Na Finskem so na nacionalni ravni dali poudarek izdelavi strategij in politik na ravni ustanov. Pri tem poudarjajo, da je potrebno upoštevati in nadgraditi obstoječe razpršene dejavnosti, ne pa jih razvijati na novo. Vsi vpleteni, od ministrstva do raziskovalnih organizacij in izvajalcev storitev, morajo pripraviti dolgoročne načrte delovanja in politike.

Finnish Information Centre for Register Research ([ReTki](#))⁴⁰⁷ skuša pomagati pri težavah, s katerimi se srečujejo raziskovalci, ko skušajo dostopati do podatkov, pomagajo s promocijo rabe

⁴⁰³ *Nordic co-operation* vključuje Dansko, Finsko, Islandijo, Norveško, Švedsko, Ferske otoke, Grenlandijo in islandsko otočje.

⁴⁰⁴ Glej: <http://www.nordforsk.org/en>, 12. 2. 2014.

⁴⁰⁵ *Recommendations for Implementation of Open Access in Denmark*, http://www.kulturstyrelsen.dk/fileadmin/publikationer/publikationer_engelske/open_access_2010/html/chapter03.htm, 13. 12. 2013.

⁴⁰⁶ Glej: <http://fivu.dk/en/publications/2011/danish-roadmap-for-research-infrastructure-2011/>, 12. 2. 2014.

⁴⁰⁷ Glej: <http://rekisteritutkimusen.wordpress.com/>, 12. 2. 2014.

javnih registrov za raziskovalne potrebe, zlasti na področju družboslovja in medicine, ponujajo informacije o teh registrih in usposabljanja za rabo.

Na **Švedskem** so oblikovali komisijo Database Infrastructure Committee (DISC), katere naloga je usklajevanje obstoječih in novih kakovostnih raziskovalnih podatkovnih baz. DISC je pooblastil Swedish National Data Service (SND), ki ima svoje korenine v že dlje časa delujočem družboslovnem podatkovnem arhivu, za koordinacijo področij družboslovja, humanistike in zdravja. SND si prizadeva vzpostaviti dober pregled nad podatkovnimi bazami na teh področjih in združevati nekatere storitve. Spodbuja tudi delo na vsebinskih projektih sodelovanja sorodnih področij, kjer se obetajo največje koristi: medicina, družboslovje in humanistika, vključno z vprašanji dostopa do registrskih podatkov. Kot primer tovrstnih projektov so načrti rabe registrov in sistema osebnih identifikacijskih števil, ki vzpostavljajo pogoje za raziskovanje pomembnih interdisciplinarnih tem na področju znanosti o življenju in drugih področjih, npr. povezovanje podatkov med družbenimi okoliščinami, stanovanjem in zdravjem.⁴⁰⁸

Nizozemska je, kot že povedano, ena od članic Knowledge Exchange. Aktivnosti na področju odpiranja raziskovalnih podatkov te države so razvidne na več načinov. Npr. Royal Netherlands Academy of Arts and Sciences ([KNAW](http://www.knaw.nl))⁴⁰⁹ kot financer zahteva, da mora vsak predlog za raziskovalni projekt vsebovati tudi načrt ravnanja s podatki. V navodilih raziskovalcem nalagajo, da morajo opisati, kdaj in kam bodo podatke trajno shranili, kdaj bodo podatki na voljo v odprtem dostopu, in če ne bodo dostopni, za to zahtevajo utemeljitev. Zagotoviti pa morajo tudi metapodatke. Nizozemska organizacija za znanstveno raziskovanje Netherlands Organisation for Scientific Research ([NWO](http://www.nwo.nl))⁴¹⁰ med pogoji za subvencioniranje zahteva tudi podpis pogodbe, ki vključuje dogovor o oddaji podatkov. Raziskovalcem je na voljo priročnik z navodili za pripravo projekta [Preparing Data for Sharing; Guide to Social Science Data Archiving](#).⁴¹¹

Med aktivnimi na tem področju je tudi že prej omenjena fundacija SURF, ki promovira uporabo trajnih identifikatorjev URN-URI, kar je tudi tema skupine DataCite pri poenotenju storitev v podporo lociranju raziskovalnih podatkov ob uporabi različnih sistemov trajnih identifikatorjev. V bibliografskem okolju je najbolj razširjen sistem DOI, vendar so za raziskovalne podatke z vidika vodenja različic in granularnosti lahko bolj primerni drugi standardi. Sam SURF ima ambiciozen program inovacijskih razvojnih projektov storitev na raziskovalnih podatkih. Med podatkovnimi središči sta posebej izpostavljeni [3TU.Datacentrum](http://www.datacentrum.3tu.nl)⁴¹² na področju dela z raziskovalnimi podatki s tehničnih univerz in [DANS](http://www.dans.knaw.nl)⁴¹³ s področja družboslovja in humanistike. Slednji je koordiniral tudi pripravo *Data Seal of Approval (DSA)*,⁴¹⁴ enega od pristopov potrjevanja skladnosti z merili dobro delujočih storitev s področja raziskovalnih podatkov. Tovrstnih standardov in orodij je sicer še nekaj.⁴¹⁵ DSA velja za manj zahtevno za izpolnjevanje in se uveljavlja kot začetna metoda za preverjanje delovanja zaupanja vrednih arhivov.

⁴⁰⁸ Life Sciences, <http://www.government.se/sb/d/16288/a/202590>, 3. 12. 2013.

⁴⁰⁹ Glej: <https://www.knaw.nl/en>, 12. 2. 2014.

⁴¹⁰ Glej: <http://www.nwo.nl/en/research-and-results/programmes/incentive+fund+open+access>, 12. 2. 2014.

⁴¹¹ Glej: <http://www.dans.knaw.nl/content/categorieen/publicaties/dans-data-guide-8>, 12. 2. 2014.

⁴¹² Glej: <http://datacentrum.3tu.nl/en/home/>, 12. 2. 2014.

⁴¹³ Glej: <http://www.dans.knaw.nl/en>, 12. 2. 2014.

⁴¹⁴ Glej: <http://datasealofapproval.org/en/>, 12. 2. 2014.

⁴¹⁵ V okviru projektov Digital Curation Centre (DCC) in Digital Preservation Europe (DPE) so razvili metodo ocenjevanja digitalnih repozitorijev na osnovi analize tveganja – orodje se imenuje [DRAMBORA](#). Najbolj dodelan, pa

Prekomorske države

V ZDA je več skupin in projektov, ki so namenjeni izboljšavam in načrtovanju podatkovne infrastrukture. Tu lahko omenimo odmevno študijo delovne skupine [Blue Ribbon Task Force](#)⁴¹⁶ z naslovom [Sustainable Economics for a Digital Planet: Assuring Long-Term Access to Digital Information](#),⁴¹⁷ v kateri podajajo splošna načela in dejavnosti za dolgoročno ekonomsko vzdržnost vzpostavljanja trajnega dostopa do digitalnih informacij. Program [DataNet](#)⁴¹⁸ je ena največjih investicij v zadnjem času in je pobuda, s katero naj bi integrirali knjižnice in znanje arhivov, kiberinfrastrukturo, računalniške in informacijske znanosti ter področne strokovnosti.

Pomembna je tudi usmerjevalna vloga National Digital Information Infrastructure and Preservation Program ([NDIIPP](#)).⁴¹⁹ Poslanstvo NDIIP je razviti nacionalne strategije zbiranja, arhiviranja in ohranjanja digitalnih vsebin. Nudijo bogastvo informacij za ozaveščanje o pomenu dolgotrajne hrambe na različnih področjih, od vsakdanjega življenja do kulturnih in znanstvenih ustanov.

Pri National Science Foundation ([NSF](#))⁴²⁰ so leta 2005 ustanovili Office for Cyberinfrastructure (OCI), skozenj pa spodbudili postopek strateškega načrtovanja za agencijske naložbe v kiberinfrastrukturo. Vizijo razvoja so opredelili v dokumentu [Cyberinfrastructure Vision for 21st Century Discovery](#)⁴²¹ (2007). Tudi poročilo delovne skupine Interagency Working Group on Digital Data z naslovom [Harnessing the Power of Digital Data for Science and Society](#) (2009)⁴²² predstavlja jasno zastavljeno vizijo potrebnih korakov za večje izkoriščanje potenciala, ki ga predstavljajo raziskovalni podatki v odprtem dostopu.

National Science Board ([NSB](#)),⁴²³ ki predstavlja skupino uglednih predstavnikov znanstvenih področij, je objavil vplivno poročilo z naslovom [Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century](#)⁴²⁴. V okviru National Science Board deluje tudi skupina za znanstvene podatke in uveljavlja načela odprtega dostopa do podatkov. *NSB Committee on Strategy and Budget Task Force on Data Policies* v stališču [Statement of Principles](#)⁴²⁵ poudarja, da morajo biti v politikah jasno opredeljene vloge, odgovornosti in sredstva, to označuje kot 3 R (*Roles, Responsibilities, Resources*). Ugotavljajo, da medtem ko se v politikah hitro razvijajo vidiki, ki zadevajo tehnični del, se pogosto zanemarja socio-ekonomski vidik, in poudarjajo, da brez slednjega ni mogoče pričakovati uspeha na tem področju. Prav tako poudarjajo pomen obvezne priprave načrta ravnanja z raziskovalnimi podatki.

Poteka tudi vrsta razvojno naravnanih projektov in aktivnosti, ki jih organizirajo združenja in društva. Eno takih je bila denimo delavnica z naslovom [Hot Topics in Big Data: What You Need to](#)

tudi najzahtevnejši pa je ISO 16363/TDR oz. njegova klasična različica [Trusted Digital Repositories and Audit Checklist \(TRAC\) standard](#).

⁴¹⁶ Glej: <http://brtf.sdsc.edu/>, 12. 2. 2014.

⁴¹⁷ Glej: <http://brtf.sdsc.edu/about.html>, 12. 2. 2014.

⁴¹⁸ Glej: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503141, 12. 2. 2014.

⁴¹⁹ Glej: <http://www.digitalpreservation.gov/>, 12. 2. 2014.

⁴²⁰ Glej: <http://www.nsf.gov/>, 12. 2014.

⁴²¹ Glej: <http://www.nsf.gov/pubs/2007/nsf0728/>, 12. 2. 2014.

⁴²² Glej: http://www.nitrd.gov/About/Harnessing_Power_Web.pdf, 13. 12. 2013.

⁴²³ Glej: <http://www.nsf.gov/nsb/>, 12. 2. 2014.

⁴²⁴ Glej: <http://www.nsf.gov/pubs/2005/nsb0540/>, 12. 2. 2014.

⁴²⁵ Glej: <http://www.nsf.gov/nsb/committees/dp/principles.pdf>, 12. 2. 2014.

[Know Now](#)⁴²⁶ v soorganizaciji [CENDI](#),⁴²⁷ [NFAIS](#)⁴²⁸ in [FEDLINK](#),⁴²⁹ ki je potekala 11. decembra 2011.

Temu sledi tudi operativna raven, kakor kaže nedavno izdani [Memorandum for the Heads of Executive Departments and Agencies](#) (26. junij 2013)⁴³⁰ glede dostopa do raziskovalnih podatkov kot javne dobrine. V strategiji izpostavijo, da je potrebno podpreti obstoječe arhive in spodbuditi zasebno-javna partnerstva. Podpreti je treba pristop, ki optimizira iskanje, spodbuja inovativnost in interoperabilnost, obenem pa zagotavlja dolgotrajno skrbništvo raziskav, financiranih s strani države. Izpostavljajo motiviranje raziskovalcev, da izpolnijo svoje obveze, ter zagotovitev sredstev za izvajanje zastavljenih načrtov. Vsaka raziskovalna agencija mora šest mesecev po izdaji tega memoranduma oddati načrt za izvajanje uradu Office of Science and Technology Policy ([OSTP](#)).⁴³¹ Agencije morajo pri načrtovanju upoštevati potrebe in posebnosti svojih področij, in sicer tako, da v pripravo načrta vključijo raziskovalce, univerze, knjižnice, založnike, uporabnike in civilno družbo. V načrtu morajo kot mesta predaje podatkov promovirati javne akademske podatkovne arhive, spodbujati sodelovanje, razvijati pristope za dostop do podatkov, podpirati usposabljanja in izobraževanja glede načrtovanja ravnanja s podatki in drugo. Širšo sliko o dokaj naprednem trenutnem položaju politik financierjev v ZDA daje prikaz [De-Mystifying the Data Management Requirements of Research Funders](#).⁴³² V njem Dianne Dietrich, Trisha Adamus, Alison Miner in Gail Steinhart ugotavljajo, da bo vse več financierjev zahtevalo od raziskovalcev, da pripravijo načrte ravnanja z raziskovalnimi podatki, vključno s tem, da bodo vedeli, kje bodo hranili podatke, in da bo poskrbljeno za dostop do kakovostnih in dobro dokumentiranih digitalnih podatkov. NSF kot financier je navodila za prijavitelje projektov objavil v rubriki [Dissemination and Sharing of Research Results](#).⁴³³

Že prej omenjeni [BRDI](#)⁴³⁴ je program Ameriške akademije, ki deluje povezovalno v mednarodnem in nacionalnem okolju ter organizira posvete o različnih aktualnih vprašanih ravnanja za raziskovalnimi podatki. Med njihovimi zadnjimi aktivnostmi je bil v februarju 2013 simpozij z naslovom [Finding the Needle in the Haystack](#)⁴³⁵ o strategijah odkrivanja raziskovalnih podatkov na spletu,⁴³⁶ septembra 2013 pa simpozij z naslovom [Privacy in a Big Data World](#),⁴³⁷ na katerem so govorili o nacionalnih in mednarodnih pristopih na temo zasebnost in t. i. veliki podatki ter o spremembah v razumevanju zasebnosti.^{438, 439} Med zadnjimi publikacijami velja omeniti poročilo z delavnice z naslovom [The Future of Scientific Knowledge Discovery in Open Networked Environments](#) (2012).⁴⁴⁰ Na delavnici so se posvetili vprašanju raziskovanja s pomočjo računalniških tehnologij, kot so npr. podatkovno rudarjenje, umetna inteligenca, grid ipd. Omeniti velja še povzetek prispevkov s simpozija [The Case for International Sharing of](#)

⁴²⁶ Glej: http://www.cendi.gov/activities/12_11_2012_CENDI_NFAIS_FEDLINK.html, 12. 2. 2014.

⁴²⁷ Glej: <http://www.cendi.gov/>, 12. 2. 2014.

⁴²⁸ Glej: <http://nfais.org/>, 12. 2. 2014.

⁴²⁹ Glej: <http://www.loc.gov/flicc/>, 12. 2. 2014.

⁴³⁰ Glej: <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-16.pdf>, 12. 2. 2014.

⁴³¹ Glej: <http://www.whitehouse.gov/administration/eop/ostp/about>, 12. 2. 2014.

⁴³² Glej: <http://www.isrl.org/12-summer/refereed1.html>, 12. 2. 2014.

⁴³³ Glej: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>, 28. 1. 2014

⁴³⁴ Glej: <http://sites.nationalacademies.org/PGA/brdi/index.htm>, 12. 2. 2014.

⁴³⁵ Glej: http://sites.nationalacademies.org/PGA/brdi/PGA_081268, 12. 2. 2014.

⁴³⁶ Glej: http://sites.nationalacademies.org/PGA/brdi/PGA_081268, 18. 12. 2013.

⁴³⁷ Glej: http://sites.nationalacademies.org/PGA/brdi/PGA_084312, 12. 2. 2014.

⁴³⁸ Glej: http://sites.nationalacademies.org/PGA/brdi/PGA_084312, 18. 12. 2013.

⁴³⁹ Več o aktivnostih na: *Chronological list of past events*, http://sites.nationalacademies.org/PGA/brdi/PGA_046395, 4. 12. 2013.

⁴⁴⁰ Glej: http://sites.nationalacademies.org/PGA/brdi/PGA_053371, 12. 2. 2014.

[Scientific Data: A Focus on Developing Countries: Proceedings of a Symposium](#) (2012)⁴⁴¹ ter povzetek delavnice z naslovom [For Attribution – Developing Data Attribution and Citation Practices and Standards](#) (2012),⁴⁴² ki govori o potrebi po razvoju tehnologij in storitev, ki bodo omogočile citiranje podatkov.

Med aktualnimi razvojno usmerjenimi projekti podatkovne raziskovalne infrastrukture iz ZDA je tudi projekt [DataOne](#),⁴⁴³ ki ga financira U.S. National Science Foundation. Cilj projekta je ustvariti inovativno okoljsko znanost s pomočjo tehnologij, ki omogočajo odprt, trajen, stabilen dostop do podatkov opazovanja Zemlje. Organizacija svojim uporabnikom zagotavlja varno infrastrukturo in orodja za iskanje, ravnanje, hrambo, deljenje in analiziranje znanstvenih podatkov. Ponuja tudi usposabljanja in izobraževanja ter skupnostno podporo.⁴⁴⁴ Posebnost pristopa *DataOne* je, da prepozna pomen ravnanja s podatki v okoljih storitev, namenjenih posameznim skupnostim praks. Te integrira skozi repliciranje in izkoriščanje obstoječih zmogljivosti med posameznimi repozitoriji za hrambo. Cilj je povečanje zmogljivosti, zasnovano odprto v prihodnost, hkrati pa ima koristno izrabo za raziskovalce že sedaj. Temelji na izrabi in prilagajanju obstoječih programskih rešitev in strojne infrastrukture, ki izhaja iz skupnostnih pristopov, s poudarkom na odprtokodnem programu (Treloar in dr. 2012, 198) Bill Michener, predstavnik *DataOne*,⁴⁴⁵ je v treh točkah predstavil akcijski načrt, po katerem bi dosegli večje upoštevanje prispevka raziskovalcev, ki sodelujejo pri deljenju raziskovalnih podatkov:

1. raziskovalcem zagotoviti nagrado za deljenje podatkov, ki je vezana na identifikacijske sheme, kot npr. DOI⁴⁴⁶ in ORCID⁴⁴⁷, ter storitve, kot npr. ImpactStory⁴⁴⁸;
2. raziskovalcem zagotoviti orodja, ki bodo poenostavila deljenje podatkov;
3. z izobraževanjem, usposabljanjem in politikami financiranja zagotoviti učinkovito spremembo kulture (Ball 2013).⁴⁴⁹

Avstralska vlada je leta 2004 objavila nacionalno strategijo [Backing Australia's Ability – Building our Future through Science and Innovation](#)⁴⁵⁰, kasneje še dve različici *Strategic Roadmap for Australian Research Infrastructure* (2006 in 2008).

Prav znotraj razvoja raziskovalne infrastrukture je nastala podenota za raziskovalne podatke Australian National Data Service ([ANDS](#)).⁴⁵¹ Podlaga za delovanje te ustanove je izčrpna študija [Towards the Australian Data Commons](#). Enota ima razvejan program dejavnosti, podprt z izdatnimi vlaganji. Glavna značilnost je »kombinacija nacionalnih storitev in koherentne raziskovalne infrastrukture na ravni ustanov, združena s sposobnostjo izkoriščanja infrastrukture z orodji na podlagi uveljavljenih politik in usposobljenosti« (Treloar in dr. 2012, 181).⁴⁵² Storitve, ki jih zagotavljajo, dejansko predstavljajo integracijo informacijskih storitev in

⁴⁴¹ Glej: http://www.nap.edu/catalog.php?record_id=17019, 12. 2. 2014.

⁴⁴² Glej: http://www.nap.edu/openbook.php?record_id=13564, 12. 2. 2014.

⁴⁴³ Glej: <http://www.dataone.org/data-management-planning>, 12. 2. 2014.

⁴⁴⁴ Glej: <http://www.dataone.org/>, 7. 8. 2013.

⁴⁴⁵ Glej: <http://www.dataone.org/>, 7. 8. 2013.

⁴⁴⁶ Glej: <http://www.doi.org/>, 7. 8. 2013.

⁴⁴⁷ Glej: <http://orcid.org/>, 7. 8. 2013.

⁴⁴⁸ Glej: <http://impactstory.org/>, 7. 8. 2013.

⁴⁴⁹ Glej: <http://www.dcc.ac.uk/blog/making-data-count#sthash.1Fgyb0HB.dpuf>, 7. 8. 2013.

⁴⁵⁰ Glej: <http://pandora.nla.gov.au/pan/32884/20080307-0000/backingaus.innovation.gov.au/index.html>, 19. 12. 2013.

⁴⁵¹ Glej: <http://www.ands.org.au/>, 12. 2. 2014.

⁴⁵² Treloar in dr. (2012), <http://www.dlib.org/dlib/september12/mavernik/09mavernik.print.html>, 23. 1. 2014.

objektov, kot so administrativne informacije o raziskovalnih projektih in z njimi povezanih entitetah, agregiranje metapodatkov, s tem da obstajajo tudi disciplinarno specifični metapodatki in orodja, podatki se lahko fizično nahajajo v nacionalnem ali mednarodnem podatkovnem arhivu. Vzpostavili so tudi poseben program, ki skrbi za urejanje in izboljšave pri dostopnosti podatkov javnega sektorja, kot je Biro za meteorologijo, Biro za statistiko Avstralije itd. (Treolar in dr. 2012, 189)

V strateških smernicah za razvoj raziskovalne infrastrukture v Avstraliji z naslovom *Strategic Roadmap for Australian Research Infrastructure* (2011)⁴⁵³ ugotavljajo, da so v zadnjih letih napredovali, vendar bodo morali okrepiti podporo na ravni načrtovanja ravnanja s podatki ter poskrbeti za integracijo z mednarodnimi pristopi ter s širšim okoljem znanstvene informacijske infrastrukture. Predlagajo cilj, da bi bili do leta 2020 zajeti vsi ustvarjeni podatki in da bi jih bilo mogoče odkriti že v fazi ustvarjanja.

Zelo ambiciozen je tudi avstralski program *Research Data Commons*, ki vzpostavlja tehnično in organizacijsko integracijo prostorsko razpršenih podatkovnih središč in njihovih storitev. Pomembno vlogo igrajo tudi univerze, ki izdajajo svoje vodiče in politike na področju ravnanja s podatki⁴⁵⁴.

Celovita usmerjenost Avstralije pri skrbi za podatke se kaže tudi v kodeksu *Australian Code for the Responsible Conduct of Research*⁴⁵⁵, ki pokriva vse vidike raziskovanja. Leta 2007 so kodeks prenovili in vanj prvič zapisali tudi pomen ravnanja s podatki. Kodeks vključuje ravnanje z raziskovalnimi podatki in povezanim gradivom, objavo in diseminacijo raziskovalnih ugotovitev, recenzijo, konflikt interesov, sodelovanje ...⁴⁵⁶

V Avstraliji so vzpostavili tudi močno podporo pridobivanju in uporabi trajnih identifikatorjev DOI za raziskovalne podatke.⁴⁵⁷ Ta vključuje tudi celovito promocijo sistema vrednotenja raziskovalnih podatkov, predanih v hrambo in v dostop (*ANDS workshop: Connecting the DOIs – data citation from planning to reward eResearch Australasia*, 2013, Brisbane).

V različnih državah je obveznost priprav načrtov ravnanja s podatki središčna tema diskusij. Temo obravnavajo tudi profesionalna združenja in organizacije v Avstraliji, izdajajo vodiče in orodja, ki podpirajo delo s podatki in njegovo načrtovanje. Npr. *Practical Data Management: A Legal and Policy Guide* (Fitzgerald in dr. 2008)⁴⁵⁸ in vodič *Data Management* pri Australian National University. Podobni vodiči obstajajo tudi drugod.⁴⁵⁹

⁴⁵³ Glej: <http://www.innovation.gov.au/science/ResearchInfrastructure/Pages/2011RoadMap.aspx>, 12. 2. 2014.

⁴⁵⁴ Glej: <http://libguides.anu.edu.au/datamanagement>, 28. 1. 2014

⁴⁵⁵ Glej: <http://www.nhmrc.gov.au/guidelines/publications/r39>, 19. 12. 2013.

⁴⁵⁶ Glej: <http://ands.org.au/guides/code-awareness.html>, 4. 12. 2013 .

⁴⁵⁷ Za več o tem glej: *Implementing DOIs for Research Data*, <http://www.dlib.org/dlib/may12/simons/05simons.html>, 19. 12. 2013.

⁴⁵⁸ Glej: <http://eprints.qut.edu.au/14923/>, 12. 2. 2014.

⁴⁵⁹ UK DA se ponaša z dokaj izčrpnim programom usposabljanj s področja *Managing and Sharing Data: Training Resources* (<http://data-archive.ac.uk/create-manage/training-resources>, 12. 2. 2014), to ima izrazitejšo poudarke pri vprašanih varovanja in zaščite podatkov. ICPSR je pripravil *Guide to Social Science Data Preparation and Archiving: Best Practice the Data Life Cycle* (<http://www.icpsr.umich.edu/icpsrweb/content/deposit/guide/ack.html>, 12. 2. 2014).

RAZISKOVALNI PODATKI IN SPREMLJAJOČA INFRASTRUKTURA ZNANSTVENEGA PUBLICIRANJA

Na že omenjenem lanskem posvetu EK s področja odprtih podatkov je bila izrečena tudi misel, da bi bilo za prihodnji razvoj znanosti verjetno bolje, če bi javni financerji med raziskovalnimi projekti dajali prednost takšnim, ki »zagotavljajo odličen pristop k deljenju podatkov in ne le dobro raziskovanje, pred predlogi, ki obljublajo odlično raziskovanje, toda brez deljenja podatkov« (Ball 2013).⁴⁶⁰ Poleg različnih spodbud, da bi ustvarjalci delili podatke, so se razpravljavci zavzeli tudi za dosledno preverjanje izpolnjevanja obveznosti, da so raziskovalni podatki javno dobro, saj bi s tem preprečili situacije, v katerih imamo uporabnika, ki bi rad dostopal do določenih podatkov, pa jih ne more izkoristiti. Se pravi, da v pristopih k oblikovanju politik na področju odprtih podatkov poudarimo tudi drugo plat zgodbe in izpostavimo interese potencialnih uporabnikov.

Očiten je premik v razmišljanjih. Razpravljavci se ne sprašujejo več, kaj je potrebno storiti, pač pa razpravljajo o podrobnostih, kako to storiti najbolj učinkovito glede na trenutne tehnološke in organizacijske možnosti. Pri tem so financerji znanosti in ministrstva kot predstavniki skupnih interesov javnosti tisti, ki morajo izpostaviti pomen podatkov kot javnega dobrega ter pooblastiti ustanove in posameznike za skrb nad podatki.

VREDNOTENJE RAZISKOVALNIH PODATKOV

V večkrat omenjenem poročilu KE skupine *A Surfboard for Riding the wave* pri vprašanju financiranja navajajo vrsto študij, ki kažejo upravičenost investicij v podatkovno infrastrukturo. To vprašanje so kot glavno temo obravnavali na več skupnih delovnih sestankih, kjer so podrobno proučili tako strukturo stroškov kot modele financiranja v posameznih državah.

Z ocenami stroškov in vpliva je povezana tudi splošnejša tema vrednotenje podatkov, ki obsega znanstveno vrednost in uporabnost podatkov. Pomembna je za raziskovalca ustvarjalca podatkov, da glede na vrednost podatkov določi in prilagodi načrt ravnanja s podatki. Pomembna je tudi za podatkovni arhiv oziroma podatkovno središče, ki presoja ustreznost podatkov za sprejem v hrambo in odprti dostop glede na vsebinsko skladnost s kriteriji arhiva in glede izpolnjevanja drugih formalnih pogojev za zagotavljanje možnosti shranjevanja.⁴⁶¹

Vrednotenje podatkov je lahko izziv tudi v povezavi z običajnimi kriteriji kakovosti in znanstvene vrednosti, ki veljajo pri klasičnih znanstvenih objavah. Ni namreč nujno, da je model klasičnega recenziranja v okviru publikacij zadosten za podatkovne objave. V poročilu, ki je izšlo

⁴⁶⁰ *Making Data Count*, <http://www.dcc.ac.uk/blog/making-data-count>, 7. 8. 2013.

⁴⁶¹ Glej tudi *Stakeholder Values and Ecosystems* (2013), http://recodeproject.eu/wp-content/uploads/2013/10/RECODE_D1-Stakeholder-values-and-ecosystems_Sept2013.pdf, 23. 1. 2013.

pri Knowledge Exchange⁴⁶², avtorji ugotavljajo, da je po analogiji s klasično revijalno objavo pri objavi podatkov, če naj šteje kot znanstvena objava, pomembno zagotoviti kakovost podatkov, oddanih v repozitorij. Na nekaterih znanstvenih področjih še ni ustrezno usposobljenih podatkovnih središč, ki bi imela vzpostavljene protokole za ocenjevanje podatkov. Eden od pomembnih kriterijev kakovosti raziskovalnih podatkov, zlasti če želimo spodbujati tudi medpodročno uporabo podatkov, je primerna opremljenost za ponovno rabo in zagotovilo verodostojnosti podatkov. Še posebej pomembno je to v primerih, ko potencialni uporabniki podatkov niso izvedenci za podatke in sami ne morejo oceniti kakovosti podatkov. V teh primerih je ocena vrednosti zagotovilo kakovosti in je lahko ključnega pomena za ustrezno uporabo in interpretacijo (Costas in dr. 2013, 13).

Vrednotenje podatkov bolj ali manj poglobljeno obravnavajo različni vodiči in študije. Kot zgled so na spletni strani NERC na voljo navodila in obrazci za načrtovanje ravnanja z raziskovalnimi podatki. Gre za krajšo in daljšo različico načrta. Krajša različica načrta obsega tri teme, o katerih morajo razmisliti raziskovalci:

1. Načrt ravnanja s podatki mora biti prisoten skozi ves življenjski krog podatkov (metapodatki, ohranjanje, kakovost podatkov, etična vprašanja in vprašanja dostopa).
2. Ali že obstajajo podatki, ki bi jih lahko uporabili za namene projekta in kako je z njihovo dostopnostjo.
3. Podatki, ki bodo ustvarjeni med raziskovalnim procesom, bodo po izteku financiranja na voljo NERC-u, se pravi, da je opredeljeno tudi priporočeno podatkovno središče, kamor naj bodo podatki oddani, kratek opis podatkov, kdaj bodo podatki oddani, možnosti druge rabe podatkov. ([Guidance Notes for Completion of the Outline Data Management Plan as part of the Case for Support in NERC Proposals](#) 2012)⁴⁶³

Daljša različica predvideva podrobnejše načrtovanje glede ustvarjanja, dokumentiranja, ravnanja s podatki in zagotavljanja njihove kakovosti, zahteva tudi opredelitev vlog in odgovornosti. Raziskovalcem v pomoč so pri NERC pripravili tudi seznam, s katerim je mogoče preveriti vrednost podatkov ([NERC Data Value Checklist](#)).⁴⁶⁴ Namen seznama je pomagati oceniti, ali so podatki primerni za oddajo oz. sprejem v katero od podatkovnih središč NERC. Med kriteriji, ki jih je potrebno upoštevati, so: pravni in zakonski premisleki, politike, znanstvena ali zgodovinska vrednost, izvirnost, pogoji, ohranjanje, dostop in raba ter formati oziroma tehnične omejitve.

Pri NERC ugotavljajo, da uspešno dolgotrajno ravnanje s podatki zahteva veščine, vezane na ohranjanje podatkov, zahteva pa tudi razumevanje znanstvenega ozadja, ki se skriva v podatkih. V svojih politikah je NERC izpostavil naslednje:

- Vzdrževal bo podatkovna središča, katerih naloga sta ravnanje z okoljskimi podatki, ki imajo dolgotrajno vrednost, in njihovo deljenje, ne glede na to, ali so bili ustvarjeni v okviru NERC ali zunaj njega.
- Podatkovna središča se bodo vedla nepristransko do vseh ustvarjalcev podatkov, ne glede na to, ali pripadajo ustanovi znotraj NERC ali zunaj njega, podatki bodo odprti na isti osnovi.

⁴⁶² Costas in drugi (2013): The Value of research Data, Metrics for datasets from a cultural and technical point of view.

⁴⁶³ Glej: <http://www.nerc.ac.uk/research/sites/data/documents/odmp-guidance.pdf>, 12. 2. 2014.

⁴⁶⁴ Glej: <http://www.nerc.ac.uk/research/sites/data/documents/data-value-checklist.pdf>, 12. 2. 2014.

- Pri delu z okoljsko raziskovalno skupnostjo bo NERC ohranjal merila za določanje dolgotrajne vrednosti.
- Informacije o vseh podatkih, ki jih hranijo podatkovna središča, bodo na voljo preko NERC Data Discovery Service.
- Vsi okoljski podatki, dostopni preko NERC, bodo opremljeni s podatkovno licenco.
- Vsi, ki uporabljajo podatke, dostopne preko NERC, so dolžni navajati vir podatkov ([Access to data](#)).⁴⁶⁵

Welcome Trust se v svojih politikah o ravnanju s podatki in deljenju podatkov (*Policy on data management and sharing*) sklicuje na naslednje dokumente: *OECD Principles and Guidelines for Access to Research Data from Public Funding* (2007); *Fort Lauderdale Principles* (2003)⁴⁶⁶ in *Toronto Statement* (2009).⁴⁶⁷ Trust je v politikah izpostavil, da pričakuje od vseh raziskovalcev, ki so financirani z njegove strani, da bodo maksimizirali dostopnost svojih raziskovalnih podatkov, s čim manj omejitvami. Od kandidatov za sredstva pričakuje, da bodo oddali načrt za ravnanje s podatki in načrt za deljenje podatkov. Od uporabnikov podatkov pričakuje, da bodo navajali vir podatkov in upoštevali pogoje, ki spremljajo podatke. ([Policy Statement](#))⁴⁶⁸

Kot motivacijo za uveljavitev politike navajajo povečanje vpliva raziskovalnih rezultatov, prihranek pri ustvarjanju raziskovalnih podatkov ter izboljšanje ponovljivosti in zaupanja v rezultate (kadar jih spremlja objava dostopnih podatkov). Jasno prepoznavajo dodano vrednost podatkov in spoznanj, kadar se ta množijo na osnovi podatkov, dostopnih različnim raziskovalcem.⁴⁶⁹

Primer, kako opredeliti vrednost raziskovalnih podatkov, ponuja Louise Corti z UK DA. Raziskovalca popelje skozi sklope vprašanj, ki se nanašajo na: vsebino, formate, dokumentacijo, potencial za drugo rabo, soglasja, izjave o izročitvi ... V nadaljevanju predstavi obrazec za vrednotenje. V obrazcu so vključeni naslednji kriteriji, glede katerih se raziskovalec opredeli, ali se nanašajo na njegove podatke ali ne:

- primarni podatki,
- širša pokritost (geografska, tematska, področna, časovna),
- metodološka odličnost,
- popularna raziskovalna tema,
- podatki, primerni za poučevanje,
- podatki v sprejemljivem formatu in enostavni za obravnavo,
- podatki dopolnjujejo obstoječe zbirke,
- dokumentacija omogoča seznanjeno drugo rabo,
- ustrezna transkripcija,
- ustrezna velikost vzorca,
- podpisana izjava o izročitvi,
- razrešeno vprašanje avtorskih pravic in intelektualne lastnine,
- soglasja o sodelovanju,

⁴⁶⁵ Glej: <http://www.nerc.ac.uk/research/sites/data/policy2011.asp>, 12. 2. 2014.

⁴⁶⁶ Glej: <http://www.genome.gov/pages/research/welcomereport0303.pdf>, 12. 2. 2014.

⁴⁶⁷ Glej: <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/WTX060361.htm>, 12. 2. 2014.

⁴⁶⁸ Glej: <http://www.wellcome.ac.uk/about-us/policy/policy-and-position-statements/wtx035043.htm>, 12. 2. 2014.

⁴⁶⁹ *Sharing research data to improve public health: full joint statement by funders of health research*, <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/WTDV030690.htm>, 9. 1. 2014.

- vprašanja anonimizacije,
- zahteva za omejen dostop (pogoji/embargo) ([Evaluating Research Data](#) 2012).⁴⁷⁰

V poročilu Royal Society Science Policy Centre z naslovom [Science as an Open Enterprise](#)⁴⁷¹ podajajo oceno, da se bo potreba po FTE-jih za univerzitetne repozitorije povečala za nekajkrat, ko bodo poleg klasičnih publikacij začeli sprejemati tudi raziskovalne podatke, saj ti običajno terjajo dodatne aktivnosti za ustvarjanje dodane vrednosti podatkov, ki pa so z njihovo uporabo večkratno poplačane (*Science as an Open Enterprise* 2012, 67). Tudi članek Piwowar in dr. (2011) [utemeljuje upravičenost večjih vlaganj](#).⁴⁷² Med projekti, ki se ukvarjajo z zbiranjem in presojo različnih modelov stroškov in vrednotenja, omenimo še [projekt 4C](#), *Collaboration to Clarify the Costs of Curation*.⁴⁷³

S tem, kako so ovrednoteni podatki, je tesno povezano tudi vprašanje, kako bomo podatke hranili, torej način skrbništva. Francine Berman v prispevku iz leta 2008 s piramidalno shemo opredeljuje hierarhijo vrednosti podatkov in z njimi povezane organizacijske modele skrbništva. Na dno piramide je umestila podatke, ki imajo individualno vrednost – vrednost za posameznike (npr. družinske fotografije), višje podatke, ki imajo pomen za neko večjo skupnost, npr. zbirke posameznih bolnišnic, fotografije določene regije iz določenega obdobja, njihovi skrbniki so lahko posamezniki, pa tudi agencije. Na vrh piramide je uvrstila podatke, ki imajo širši družbeni pomen, njihovi skrbniki so običajno državne pooblaščenec ali mednarodne ustanove, kot so nacionalni podatkovni centri in arhivi – ti vključujejo podatke, ki jih ni mogoče nadomestiti z drugimi (Berman 2008, 55–56).

Upravičenost vlaganj v odprti dostop do raziskovalnih podatkov se da utemeljiti s sistemom spremljanja stroškov in koristi. Tak sistem se lahko razteza od preprostega sledenja publikacijam, ki so nastale na podlagi dostopnih podatkov kot kazalca uporabnosti podatkov, pa do sistematične razčlenitve nalog in ur, porabljenih za posamezne faze priprave podatkov, ter koristi, vključno s koristmi družbe pri reševanju aktualnih problemov. Merjenje koristi na način stroškov je posebej podrobno razdelala skupina avtorjev v aktualni študiji z naslovom [Economic Impact of the British Atmospheric Data Centre \(BADC\)](#)⁴⁷⁴ o vrednosti in vplivu British Atmospheric Data Centre (BADC). Študija je pokazala, da so raziskovalci, ki uporabljajo BADC, učinkovitejši, kar je v študiji ocenjeno na vrednost najmanj 10 milijonov funtov na leto (Beagrie in Houghton 2013, 6). Neil Beagrie in John Houghton sta podobno raziskavo izvedla tudi na področju družboslovja, preučila sta koristi in stroške, ki nastajajo pri delu Economic and Social Data Service (ESDS), katerega del je sedaj tudi UK DA. Njuna obsežna analiza z naslovom [The Economic Impact Evaluation of the Economic and Social Data Service](#) (2012)⁴⁷⁵ opredeljuje dejavnike, s katerimi je mogoče meriti ekonomski učinek in pomen storitve, ki jo nudi podatkovna organizacija; prišla sta do zaključka o 2,5- do 10-kratnem povračilu stroškov investicije v podatkovno storitveno infrastrukturo (Beagrie in Houghton 2012, 6). Podobno študijo sta avtorja izvedla tudi za področje arheologije, in sicer je na voljo pod *naslovom* [The](#)

⁴⁷⁰ Glej: http://www.data-archive.ac.uk/media/369163/managing_research_data14sept2012b.pdf, 12. 2. 2014.

⁴⁷¹ Glej: <http://royalsociety.org/policy/projects/science-public-enterprise/report/>, 12. 2. 2014.

⁴⁷² *Data archiving is a good investment* (Nature 2011).

⁴⁷³ D3.1–Evaluation of Cost Models and Needs & Gaps Analysis (MS12 Draft), <http://www.4cproject.eu/community-resources/outputs-and-deliverables/d3-1-evaluation-of-cost-models-and-needs-gaps-analysis-ms12-draft>, 19. 12. 2013.

⁴⁷⁴ Glej: http://www.jisc.ac.uk/whatwedo/programmes/di_directions/strategicdirections/badc.aspx, 12. 2. 2014.

⁴⁷⁵ Glej: http://www.esrc.ac.uk/images/ESDS_Economic_Impact_Evaluation_tcm8-22229.pdf, 12. 2. 2014.

[Value and Impact of the Archaeology Data Service](#) (2013) (Beagrie in Houghton 2013, 6–7).⁴⁷⁶ Ocenjevanje podatkov z vidika kakovosti in primernosti za podatkovni center, podatkovno recenziranje z vidika presoje znanstvene vrednosti, preverjanje in zagotavljanje kakovosti podatkov so vse aktualne teme različnih projektov in predmet diskusij na raznih strokovnih srečanjih.⁴⁷⁷

POLITIKE ZNANSTVENIH REVIJ IN UREDNIŠTEV

[Budapest Open Access Initiative](#)⁴⁷⁸ velja za pomembnejši mejnik, kadar govorimo o odprtem dostopu do raziskovalnih publikacij. Šele v zadnjem času pa so glasnejše zahteve, da bo potrebno poskrbeti tudi za odprti dostop do raziskovalnih podatkov, na katerih temeljijo znanstvene objave. Odprti dostop do objav pomeni možnost seznanjanja z znanstvenimi ugotovitvami, odprti dostop do podatkov pa prinese možnost preverjanja teh rezultatov, primerjave postopkov, metod in drugo. Odprti dostop do raziskovalnih podatkov prispeva k preglednosti in preverljivosti znanstvenega dela, ki sodijo med osnovna znanstvena načela. Ponovljivost rezultatov na podlagi istih podatkov se izkaže kot končni test verodostojnosti objavljenih ugotovitev, ki pa ga je pogosto težko izvesti (King 1995).

Za uresničevanje sistema odprtega dostopa do raziskovalnih podatkov so pomembne tudi znanstvene publikacije, in sicer na dva načina. Najprej, ker uredništva od avtorjev zahtevajo dosledno navajanje podatkov, na katerih temelji objava, vključno z zagotovilom odprtega dostopa, in potem še, ker spremljajo drugo rabo podatkov, objavljenih v njihovih publikacijah.

Več mednarodno uveljavljenih znanstvenih revij je že na poti uresničevanja odprtega dostopa do raziskovalnih podatkov. Sprejele so notranje politike, ki jih je največkrat mogoče prepoznati v navodilih avtorjem, s katerimi od avtorjev člankov zahtevajo informacijo o tem, kje in pod kakšnimi pogoji je mogoče dostopati do raziskovalnih podatkov, na katerih temeljijo njihove ugotovitve. Npr. revija *Nature* zahteva, da so podatki, na katerih temelji objava, dostopni že v času recenzentskega postopka, po objavi pa morajo biti dostopni tudi bralcem. Avtorjem priporočajo, da podatke shranijo v katerem od javnih repozitorijev. Pravila glede ravnanja s podatki so objavljena v uredniških politikah na spletni strani revije. V rubriki »izročanje podatkov« je zapisano, da morajo avtorji zagotoviti šifro stalnega mesta dostopa. Ta je obvezna za ustaljene strukture, za katere obstajajo mednarodne banke podatkov, kot so genetski zapisi. »Hramba je močno priporočljiva tudi za številne druge podatkovne datoteke, za katere obstajajo strukturirani javni repozitoriji,« so še zapisali v dokumentu izpolnjevanja obveznosti, namenjenem avtorjem publikacij.⁴⁷⁹

Politikam odprtega dostopa je zavezana tudi založba PLOS⁴⁸⁰, ki s svojimi revijami pokriva različna raziskovalna področja. Npr. [PLOS](#)⁴⁸¹ Biology, PLOS Medicine, PLOS Genetics in drugo. Od avtorjev, ki želijo svoje ugotovitve objaviti v njihovih revijah, zahtevajo informacijo o tem, kje je mogoče dostopati do podatkov, vzorcev ali modelov, na katerih temeljijo njihove ugotovitve.⁴⁸²

⁴⁷⁶ Glej: <http://archaeologydataservice.ac.uk/research/impact>, 12. 2. 2014.

⁴⁷⁷ Npr. *Workshop: Curating for Quality: Ensuring Data Quality to Enable New Science*, <http://datacuration.web.unc.edu/>, 23. 1. 2014.

⁴⁷⁸ Glej: <http://www.budapestopenaccessinitiative.org/>, 12. 2. 2014.

⁴⁷⁹ Glej: <http://www.nature.com/authors/policies/availability.html>, 7. 8. 2013.

⁴⁸⁰ Public Library of Science, <http://www.plos.org/>, 7. 8. 2013.

⁴⁸¹ Glej: <http://www.plos.org/>, 12. 2. 2014.

⁴⁸² Glej: <http://www.plos.org/publish/rcuk-open-access-mandate/>, 7. 8. 2013.

Založba [BioMed Central](#),⁴⁸³ ki izdaja več kot 250 revij s področja znanosti, tehnologije in medicine, je v nekaj svojih revijah že uveljavila zahtevo po odprtem dostopu do podatkov, zavzemanje za odprt dostop do podatkov (»Open Data«) so opredelili tudi v politikah svoje ustanove.⁴⁸⁴ Revija [GigaScience](#)⁴⁸⁵ zahteva od avtorjev, da podatkovne datoteke, na katerih temeljijo njihove ugotovitve, oddajo v javno dostopne repozitorije, kot je npr. [GigaDB](#).^{486, 487}

Založnik [Pensoft](#)⁴⁸⁸ prav tako predvideva odprt dostop do podatkov za skupino svojih revij. Revije *ZooKeys*, *PhytoKeys*, *MycKeys*, *BioRisk*, *NeoBiota*, *Nature Conservation*, *Comparative Cytogenetics* in *International Journal of Myriapodology* so povezane v sistem avtomatizirane oddaje podatkov v repozitorij [Dryad Digital Repository](#).⁴⁸⁹ Pri Pensoftu ponujajo tudi model podatkovne revije, ki je namenjena objavi podatkov oziroma podatkovnih zbirk kot samostojnega raziskovalnega objekta. Lyubomir Penev, direktor in ustanovitelj Pensofta, pravi:

*Podatkovni članek je članek v znanstveni reviji, ki na način »razširjenih metapodatkovnih opisov« opiše podatkovne baze ali zbirke podatkovnih baz. Namen podatkovnih člankov ni predstaviti znanstvene rezultate, ki izvirajo iz podatkov, čeprav lahko vsebujejo nekaj analize in preprostih statistik (npr. porazdelitev enot po skupinah, po regijah). Podatkovni članki zagotovijo trajen znanstveni zapis in ustrezne mehanizme za citiranje, s katerim se ljudem in ustanovam povrne trud, vložen v ustvarjanje in ravnanje s podatki.*⁴⁹⁰

Podatkovni članek gre skozi recenzentski postopek in znanstveno ovrednotenje po načelu drugih znanstvenih objav. Založba Pensoft kot glavni izziv podatkovnih objav vidi raznolikost tipov podatkov, formatov, tehničnih standardov in drugo. Težave, ki izvirajo iz raznolikosti, premoščajo s standardiziranimi rešitvami, ki ustrezajo določenim skupinam podatkov.

Med podatkovnimi revijami velja omeniti še: [CODATA Data Science Journal](#),⁴⁹¹ specializirano za objavo podatkov s področja fizike, geo- in bioznanosti, kmetijskih in medicinskih ved;⁴⁹² *Earth System Science Data*, [Geoscience Data Journal](#)⁴⁹³ in druge.

T. i. podatkovne revije sicer prevladujejo v naravoslovju, mogoče pa jih je najti tudi na drugih področjih. V okviru skupine Ubiquity Press Metajournals izhajajo naslednje revije, ki pokrivajo različna raziskovalna področja: [Journal of Open Archeology Data](#),⁴⁹⁴ *Journal of Open Public Health*

⁴⁸³ Glej: <http://www.biomedcentral.com/about/access>, 12. 2. 2014.

⁴⁸⁴ Glej: <http://www.biomedcentral.com/about/access>, 7. 8. 2013.

⁴⁸⁵ Glej: <http://www.gigasciencejournal.com/authors/instructions/research>, 7. 8. 2013.

⁴⁸⁶ Glej: <http://gigadb.org/>, 7. 8. 2013.

⁴⁸⁷ Še nekaj revij s politikami odprtega dostopa do podatkov je na voljo na seznamu *Journal open-data policies*: <http://oad.simmons.edu/oadwiki/journal-open-data-policies>, 10. 1. 2014 .

⁴⁸⁸ Pensoft Publisher, <http://www.pensoft.net/>, 7. 8. 2013.

⁴⁸⁹ Glej: <http://www.pensoft.net/news.php?n=86>, 7. 8. 2013.

⁴⁹⁰ »Data Paper«: *the data publishing project of Pensoft* (2012), <http://www.openaire.eu/en/component/content/article/76-highlights/399-data-paper-the-data-publishing-project-of-pensoft>, 20. 12. 2013.

⁴⁹¹ Glej: <http://www.codata.org/dsj/>, 12. 2. 2014.

⁴⁹² Glej: <http://www.codata.org/dsj/>, 7. 8. 2013.

⁴⁹³ Glej: <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%292049-6060>, 7. 8. 2013.

⁴⁹⁴ Glej: <http://openarchaeologydata.metajnl.com/about>, 7. 8. 2013.

Data ([IOPHD](#)),⁴⁹⁵ *Journal of Open Psychology Data* ([JOPD](#)),⁴⁹⁶ kmalu pa bo začel izhajati tudi *Journal of Open Economics Data* ([IOED](#)).⁴⁹⁷

V slovenskem prostoru velja izpostaviti [Acta Chimica Slovenica](#) kot eno redkih znanstvenih revij, ki zahteva oddajo podatkov v ustrezno podatkovno bazo, kadar članek obravnava kristalografske podatke.⁴⁹⁸ V rubriki »navodila avtorjem« priporočajo ustrezen format in repozitorije, npr. Cambridge Crystallographic Data Centre ([CCDC](#))⁴⁹⁹ in Fachinformationszentrum Karlsruhe ([FIZ](#)).⁵⁰⁰

Hiter pregled revij, ki so se že spopadle z vprašanjem odprtega dostopa do podatkov, kaže, da se politike revij razlikujejo po tem, na kakšen način zahtevajo dostop do podatkov. Nekatere zahtevajo oddajo podatkov v javne repozitorije, kot sta Global Biodiversity Information Facility ([GBIF](#))⁵⁰¹ in Dryad⁵⁰², ali javne baze, kot je npr. [GenBank](#).⁵⁰³ Manj pogoste možnosti pa so, da raziskovalci objavijo podatke v repozitorijih svojih ustanov.

V okviru projekta [JoRD](#)⁵⁰⁴ razvijajo model politik o ravnanju z raziskovalnimi podatki za znanstvene revije. Raznolikost tipov podatkov in tudi obseg podatkov lahko revijam povzroča preglavice, ko jih skušajo umestiti v svoje objave. Zato je za revije pogosto bolje, če avtorji podatke shranijo pri ustreznem repozitoriju oziroma podatkovnem središču, skupaj pa poskrbijo za povezavo med objavo in podatki.⁵⁰⁵

Revije čaka še eno pomembno vprašanje, to je uveljavitev sistema doslednega navajanja (citiranja) podatkovnih gradiv po načelu, ki velja za navajanje virov v znanstvenih objavah. To pomeni, da mora biti podatkovno gradivo, ki je uporabljeno v objavi, tudi ustrezno vpisano v seznam literature (referenc)⁵⁰⁶ in ustrezno citirano znotraj objave. Za uresničevanje te zahteve morajo revije predpisati standarde citiranja. S tem bodo revije prispevale k izgradnji sistema nagrajevanja za delo, vloženo v ustvarjanje podatkovnih gradiv.

V zadnjem času je več pobud, ki so promovirale citiranje podatkov. Glej npr. [For Attribution – Developing Data Attribution and Citation Practices and Standards: Summary of an International Workshop](#) (2012).⁵⁰⁷ Iz tega je zrasla potreba po usklajevanju, pri kateri je pomembno vlogo odigrala Research Data Alliance (RDA), pod njenim okriljem je nastalo več interesnih skupin, ki povezujejo strokovnjake iz različnih organizacij. Ena takih je skupina [Data Citation Synthesis](#).⁵⁰⁸ Skupina je objavila osnutek deklaracije – načela citiranja podatkov z naslovom [Declaration of](#)

⁴⁹⁵ Glej: <http://openpublichealthdata.metajnl.com/>, 7. 8. 2013.

⁴⁹⁶ Glej: <http://openpsychologydata.metajnl.com/>, 7. 8. 2013.

⁴⁹⁷ Glej: <http://www.ubiquitypress.com/>, 7. 8. 2013.

⁴⁹⁸ Glej: http://acta.chem-soc.si/depositing_x_ray_data.html, 7. 8. 2013.

⁴⁹⁹ Glej: <http://www.ccdc.cam.ac.uk/pages/Home.aspx>, 7. 8. 2013.

⁵⁰⁰ Glej: <http://www.fiz-karlsruhe.de/index.php?id=15>, 7. 8. 2013.

⁵⁰¹ Glej: <http://www.gbif.org/>, 7. 8. 2013.

⁵⁰² Npr. Pensoft.

⁵⁰³ Glej: <http://www.ncbi.nlm.nih.gov/genbank/>, 7. 8. 2013.

⁵⁰⁴ Glej: <http://jordproject.wordpress.com/>, 12. 2. 2014.

⁵⁰⁵ *Data comes in all sorts of shapes and sizes*, <http://jordproject.wordpress.com/category/data-sharing-policy/>, 23. 1. 2014.

⁵⁰⁶ V rubrikah, v katerih avtorjem dajejo navodila za urejanje člankov pred oddajo in objavo, se navodila redko nanašajo na podatke. Glej npr. navodila avtorjem pri Družboslovnih razpravah: »Dela, navedena v članku, morajo biti v abecednem seznamu navedena na koncu v poglavju z naslovom Literatura. Če je članek napisan na osnovi raziskave ali arhivskih podatkov, naj avtor/avtorica to eksplicitno napiše. Raziskava, raziskovalno poročilo ali baza podatkov se navede kot vir« (<http://druzboslovnirazprave.org/sl/navodila-avtorjem/>, 20. 12. 2013).

⁵⁰⁷ Glej: http://www.nap.edu/catalog.php?record_id=13564, 20. 12. 2013.

⁵⁰⁸ Glej: <http://www.force11.org/node/4432>, 20. 12. 2013.

[Data Citation Principles](#),⁵⁰⁹ in sicer z namenom, da bi spodbudila širšo rabo politik citiranja podatkov med različnimi področji.

Na podlagi informacij, dostopnih preko spletnih strani, ocenjujemo, da se slovenske znanstvene revije vprašanju odprtega dostopa do raziskovalnih podatkov za enkrat še niso veliko posvečale. Prav tako ni mogoče govoriti o politikah, ki bi urejale navajanje podatkov, kakor velja za navajanje drugih virov. Videti je, da v slovenski raziskovalni skupnosti obstaja načelno strinjanje oz. samoumevno pričakovanje, da bodo avtorji publikacij upoštevali pravila citiranja, kadar uporabljajo podatke, ki jih niso ustvarili sami. V praksi pa je mogoče ugotoviti, da so načini citiranja podatkov že znotraj iste publikacije neenotni,⁵¹⁰ večkrat pa je mogoče ugotoviti, da podatki sploh niso citirani. Kar je, predpostavljamo, posledica problema, da je citiranje podatkov prepuščeno avtorjevi osebni presoji in izboru načina citiranja. Za čim večjo izrabo slovenskih podatkovnih gradiv in njihovo prepoznavnost tako v domači kot v tuji raziskovalni skupnosti bo nujno potrebno premostiti stanje brez jasnih navodil.

Financerjevo pričakovanje glede bolj določnih zahtev upravičuje dejstvo, da dostop do podatkov prispeva k uresničevanju znanstvenih načel (preverljivost, veljavnost, zanesljivost), in sicer s tem, da je zagotovljena pregledna sledljivost vseh gradiv, ki so bila uporabljena v publikaciji – tudi podatkovnih! To pomeni, da financer zahtevo po odprtem dostopu do raziskovalnih podatkov vključi v razpisne pogoje financiranja revij skladno s krovno politiko in zahteva od revij, ki kandidirajo za javna sredstva za sofinanciranje, da jo uresničujejo.

Uredništva lahko sama spremenijo svoje notranje politike oziroma navodila avtorjem, ki vključujejo vsaj dve temeljni zahtevi. Prvič, v recenzijo sprejmejo le članke, katerih avtorji izpolnijo zahtevo po odprtem dostopu do podatkov, uporabljenih v objavi. Drugič, podatki, uporabljeni v objavi, morajo biti opremljeni z oznako, iz katere je razvidno, kje so podatki shranjeni in dostopni (iz poročila o delavnici *Politike ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi*).

⁵⁰⁹ Glej: <http://www.force11.org/datacitation>, 12. 2. 2014.

⁵¹⁰ Primerjaj npr. Družboslovne razprave, XXIX (2013) – strani: 36, 59, 95, in Teorija in praksa 3/2012 – str. 460, 483, 496, 518 in 552.

ZAKLJUČKI MEDNARODNEGA PREGLEDA

Večkrat omenjeno poročilo *A Surfboard for Riding the Wave – Towards a four country action programme on research data* (Graaf in Waaijers 2011) se osredotoča na štiri ključne vzvode sprememb: spodbude, usposabljanje, vzpostavljanje infrastrukture ter financiranje njenega vzdržnega delovanja in razvoja. Poročilo navaja zglede dobrih praks in teme, ki so potrebne nadaljnje usklajenega napora, kot so usposabljanje in podpora na področju NRRP.

Poročilo na podlagi ugotovitev podaja predloge za skupino držav za aktivnosti v naslednjih letih, ki zadevajo promocijo in ozaveščanje o pomenu deljenja podatkov, vključevanje načel v profesionalne kodekse, določanje obveznosti priprave načrtov ravnanja s podatki in njihovega uresničevanja, vzpostavljanje infrastrukture za citiranje podatkov, pobude za spremembe politik revij, da zahtevajo dostop do originalnih podatkov (Graaf in Waaijers 2011, 32).

Bistvena ugotovitev, ki sledi iz mednarodnega pregleda, je, da je pomembno tekoče spremljati dogajanje v mednarodni skupnosti in sodelovati pri strokovnih diskusijah, se odzivati ter sproti preizkušati in uvajati novosti. Dogajanje je namreč izredno intenzivno tako na posameznih področjih znanosti, npr. vprašanja t. i. velikih podatkov, kot pri oblikovanju politik in strategij, pri pristopih k poučevanju, tehnologijah in orodjih.⁵¹¹ Ob tem je v nacionalnem okolju nemogoče preizkusiti vse novosti, zato je toliko bolj pomembno mednarodno povezovanje, da se izkoristi kolektivno znanje, sodelovanje pri različnih iniciativah in projektih. Hkrati pa je napak še naprej odlašati pri konkretnih korakih pri načrtovanju in izvajanju lastnih strategij in politik. Kot smo navajali iz različnih poročil, so obrisi prihodnjega sistema odprtih podatkov zdaj že dobro vidni in ni nevarnosti, da bi koraki vodili v slepo ulico. Pravzaprav je povzetek teh različnih prikazov in ugotovitev vsebovan v točkah akcijskega načrta v nadaljevanju. Prišel je čas za ukrepanje; danes je nevarnost v tem, da postaneš zamudnik in s tem izgubljaš pri tekmovanju z znanstveniki od drugod. Pogost pomislek v zvezi z odpiranjem podatkov na nacionalni ravni, ki neizogibno sega tudi v mednarodno okolje, je, da bodo raziskovalci in laboratoriji iz drugih držav izkoriščali naše podatke za znanstvena odkritja, mi pa bomo samo zagotavljali surovine za ostale in bomo s tem na izgubi. Na to se je potrebno pripraviti s tem, da se kot del s podatki povezanih storitev uvaja tudi usposabljanje naših uporabnikov za delo s podatki, ki so v odprtem dostopu, domačimi in od drugod. Deljeni podatki predstavljajo možnost za povezovanje in kombiniranje podatkov z več lokacij, obenem pa so lahko tudi podlaga za sodelovanje pri projektih in pri razvoju metod za njihovo analitično izkoriščanje. Skratka, potrebna je modrost uvida, da je še posebej v znanosti sodelovanje lahko samo koristno. Pri tem pa se morajo majhne in manj razvite države še vedno zavzemati za možnosti izrabe razvojnih sredstev in izrabljati storitve, orodja in možnosti usposabljanja, ki se ponujajo v mednarodnem okolju.

Da v mednarodnem okolju prihaja do poenotenja, dokazuje vedno več konkretnih političnih pobud. Ena takih je nastala na srečanju ministrov za znanost iz G8 ([G8 Science Ministers](#)

⁵¹¹ Glej npr. študijo o razvoju indikatorjev odprtega dostopa z naslovom *Open Data Access Policies and Strategies in the European Research Area and Beyond* (Nicol in dr. 2013), http://www.science-metrix.com/pdf/SM_EC_OA_Data.pdf, 10. 1. 2014.

[Statement](#))⁵¹² junija 2013 v Londonu. Njihovo stališče je, da je potrebno posebno pozornost posvetiti tudi odpiranju raziskovalnih podatkov, saj bo to spodbudilo napredovanje pri znanstvenih odkritjih. Ob tej priložnosti so se poenotili glede načel, ki bodo pomagala pri doseganju tega cilja. Če na kratko povzamemo, so sklenili naslednje:

- raziskovalne podatke je potrebno odpreti (upoštevajoč posebne omejitve),
- odprti podatki naj bodo dostopni, uporabni, interoperabilni ...,
- zagotoviti je potrebno stroškovno učinkovite mehanizme, s katerimi bo mogoče ustvariti največjo možno vrednost podatkov,
- za uresničitev načel odprtih podatkov bo potrebno poskrbeti z ustreznimi politikami.

Tudi delovanje RDA lahko razumemo kot napor v smeri usklajevanja pristopov pri oblikovanju zahtev glede ravnanja z raziskovalnimi podatki. Npr. predlog načel pri citiranju podatkov kaže na široko vključenost različnih iniciativ. Sam predlog je v javni debati in dostopen za dodatne izboljšave.

Tudi uspeh skupine za pripravo aktivnosti COST z naslovom PARD⁵¹³, ki je uspela pritegniti posameznike iz 20 držav, kaže naraščajočo jasno usmerjenost skupnosti v reševanje konkretnih problemov in izmenjavo izkušenj. Raziskovalci, akademske založbe, knjižnice, podatkovna središča, financerji raziskav in drugi bodo morali sodelovati, da bi rešili vprašanja, ki zadevajo diseminacijo, dostop, objave, kakovost ... podatkov, sodelovanje bo potrebno na državni, pa tudi mednarodni ravni.⁵¹⁴ Okvir, ki ga bodo zagotovili pri akciji COST, bo akterjem v Evropi dal skupni glas na globalni sceni.

Odprti podatki postajajo področje z izoblikovanim naborom priporočil, glede katerih je dosežen konsenz. Lep povzetek je denimo [Implementing an Open Data Policy. A Primer for Research Funders](#).⁵¹⁵ Dokument odgovarja na precej pogosta vprašanja, kot npr.: kaj je politika odprtega dostopa, zakaj bi financerji raziskav morali premisliti o politikah odprtega dostopa, zakaj je bolje, da financerji zahtevajo dostop do podatkov, katere dejavnosti in gradiva naj pokrivajo politike odprtega dostopa, kaj mora vsebovati dober načrt za deljenje podatkov ipd.

Izčrpen in koncizen pregled aktivnosti, usmerjen na univerzitetne ustanove, je sveže izdani dokument [LERU Roadmap for research data](#).⁵¹⁶ Se pravi, da je vse več načrtov in vodičev, ki obravnavajo določen segment nosilcev vlog ter njihovih povezav v ekosistemu nalog in odgovornosti v različnih fazah življenjskega kroga podatkov.

Konzultacija o naslednjih korakih pri vzpostavljanju panevropske podatkovne raziskovalne infrastrukture je v programu Obzorje 2020 razvidna v dokumentu [Research Data e-Infrastructures: Framework for Action in H2020](#),⁵¹⁷ ki nakazuje področja, kjer se bodo povezovali obstoječi nacionalni in mednarodni projekti razvoja. Našteta so področja in pa potencialni nosilci. V tem smislu lahko beremo program kot priložnost za tiste, ki vidijo probleme in imajo

⁵¹² Glej: <https://www.gov.uk/government/news/g8-science-ministers-statement>, 12. 2. 2014.

⁵¹³ *Publishing Academic and Research Data*, korespondenca na PARD@JISCMAIL.AC.UK, 15. 11. 2013.

⁵¹⁴ Primerjaj *Opportunities for Data Sharing*, http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/09/8238_ODE_DATA_SHEET_Funders_LR.pdf, 20. 12. 2013.

⁵¹⁵ Glej: <http://sparc.arl.org/sites/default/files/sparc-open-data-primer-final.pdf>, 20. 12. 2013.

⁵¹⁶ Glej: http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf, 12. 2. 2014.

⁵¹⁷ Glej: https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/data_infra_h2020_fiches_on-line_consult.pdf, 12. 2. 2014.

ambicije, da se z njimi spopadejo. Nekatere teme so skupne za več področij, razvoj poteka tudi znotraj področij posameznih disciplin ali skupin, izziv pa ostaja implementacija ponujene transverzalne infrastrukturne storitve (trajni identifikatorji, internetne identitete, standardi za deskriptivne metapodatke). Se pravi, da ni potrebno ponovno razvijati določenih storitev, pač pa je le treba poenotiti njihovo izkoriščanje na način, ki zagotavlja večjo integracijo. Posamezne storitve so že bile predmet obravnave z vidika možnosti uporabe na področju raziskovalnih podatkov, denimo pred kratkim pripravljena študija glede infrastrukture AAA z naslovom [*Advancing Technologies and Federating Communities, A Study on Authentication and Authorisation Platforms For Scientific Resources in Europe*](#).⁵¹⁸ Predstavljena shema izhaja iz spoznanja, da za razliko od klasičnih publikacij, ki so jih v strokovnem smislu pri razvoju podporne infrastrukture pokrivali znanstveni knjižničarji in založniki, področje zagotavljanja trajne skrbi za podatke in njihov dostop zahteva sodelovanje različnih strokovnosti.

Da misli EK resno z vzpostavljanjem podatkovne infrastrukture, kaže tudi zahteva za dostop do raziskovalnih podatkov v okviru novega sklica raziskovalnih razpisov v Obzoru 2020. Kot pilotna aktivnost obveznost zajema posebej opredeljena področja, s tem da je priprava načrtov ravnanja z raziskovalnimi podatki priporočena tudi za vsa ostala področja, prav tako kot je določena tudi obveznost, da se zagotovi dostop do raziskovalnih podatkov, ki spremljajo besedilne raziskovalne objave.

[*Delivering Research Data Management Services*](#) (Pryor, Jones in Whyte 2013)⁵¹⁹ je eden od najnovejših vodičev o sodobnem dogajanju na področju podatkovnih storitev. Avtorji predstavijo korake k vzpostavljanju učinkovitih podatkovnih storitev na primerih iz Velike Britanije, Avstralije in ZDA – s poudarkom na načrtovanju ravnanja z raziskovalnimi podatki. V središče razprave so postavili naslednje teme: možnosti in pristopi k ravnanju s podatki, vloge, odgovornosti in kompetence pri ravnanju s podatki, pot k vzdržnim podatkovnim storitvam, nabor storitev in podpornih infrastruktur s področja ravnanja z raziskovalnimi podatki.

Mednarodni pregled zaključujemo z vprašanjem, ki so ga zastavili v okviru projekta ODE. Vprašanje, ki se nanaša na financerje, se glasi: kaj lahko storite vi in vaša organizacija za izboljšanje deljenja podatkov? Na vprašanje odgovarjajo tako: financerji in tisti, ki oblikujejo politike, naj priskrbijo finančno podporo za skrbno ravnanje z raziskovalnimi podatki. V okviru prijav za projekte je potrebno vzpostaviti zahteve za pripravo načrtov ravnanja z raziskovalnimi podatki, ki vključujejo njihovo predajo in hrambo ter možnost citiranja. Spodbujati je potrebno koordiniran razvoj institucionalnih ali področnih infrastruktur in storitev. Poleg financiranja je potrebno zagotoviti tudi, da podatki postanejo pomembno merilo akademske kakovosti. Nasvet dajejo tudi izobraževalnim ustanovam, ki naj poskrbijo, da bodo v svoje podiplomske programe vključile usposabljanje za delo s podatki ([*Opportunities for Data Sharing*](#)).⁵²⁰

Upoštevanje pobude EU v predlogu akcijskega načrta

V zaključnem predlogu akcijskega načrta za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v Sloveniji smo se zgledovali po številnih zgoraj predstavljenih dobrih praksah. Med najpomembnejšimi smo upoštevali pobude in implementirali posamezne zahteve EU. Tako smo denimo sledili načelu, da mora biti dodatno delo s podatki, ki nastane zaradi

⁵¹⁸ Glej: <https://confluence.terena.org/download/attachments/30474266/2012-AAA-Study-report-final.pdf?version=1&modificationDate=1355503760046&api=v2>, 28. 1. 2014

⁵¹⁹ Glej: <http://www.dcc.ac.uk/node/10114>, 12. 2. 2014.

⁵²⁰ Glej: www.stm-assoc.org/2012_09_05_ODE_DATA_SHEET_Funders.pdf, 20. 12. 2013.

zagotovitev odprtega dostopa, za raziskovalce ustvarjalce podatkov čim manj moteče pri siceršnjem ustvarjalnem delu/raziskovalnem procesu. V ta namen je potrebno izgraditi podporne dejavnosti na institucionalni in disciplinarni ravni, kot: pomoč v obliki svetovanja, priprava orodij in postopkov, vzpostavitev infrastrukture področnih podatkovnih središč kot mest, kamor je mogoče podatke predati. Predvidena je obvezna priprava načrtov ravnanja z raziskovalnimi podatki in sprotne komunikacija s področnimi podatkovnimi središči v vseh fazah načrtovanja, zbiranja in priprave podatkov, kar vključuje zgodnje in sprotne delo na pripravi podatkov za odprti dostop. Tak pristop zagotavlja boljšo organiziranost, s tem pa tudi kakovost in uporabnost podatkov že za izvirnega raziskovalca, ob minimalni dodatni porabi časa. Poleg podpore v obliki storitev je raziskovalce in ustanove potrebno motivirati z nagrajevanjem za njihov prispevek – tako v obliki znanstvenega bibliografskega vrednotenja kakor tudi s priznavanjem dodatnih upravičenih stroškov pri projektnem financiranju.

V zahtevah EK so glede infrastrukturne podpore posebej omenjene storitve prevzema podatkov v digitalni obliki, njihovega beleženja in shranjevanja v ustrezne formate ter skrbi za dolgotrajno hrambo. Razvoj nacionalne podatkovne infrastrukture za razširjanje znanja mora biti usmerjen v podporo vsem fazam v življenjskem ciklu podatkov, pri čemer so posebej izpostavljeni prevzem, skrbništvo, metapodatki, registracija in preverjanje istovetnosti ter podatkovna integriteta. Naloga infrastrukturnih podatkovnih središč je tudi podpora pri razvoju in usposabljanju zaposlenih v podatkovnih poklicih. Posebej je priporočeno, da se zaradi večje učinkovitosti upošteva in izkorišča obstoječe nastavke dejavnosti, obenem pa se spodbuja razvoj inovativnih orodij in pristopov pri izkoriščanju raziskovalnih podatkov. Skladno z zahtevami EU v akcijskem načrtu dajemo poudarek vzpostavitvi vzdržne nacionalne infrastrukture za dostop in hrambo znanstvenih informacij, ki ima tako institucionalna kot finančna zagotovila za svoj obstoj.

Infrastruktura mora po priporočilih EU zagotavljati določeno raven kakovosti in zanesljivosti delovanja, kar – prav tako vključeno v akcijski načrt – je mogoče doseči tako, da se jasno opredelijo in razdelijo odgovornosti, predpišejo zahteve glede delovanja ter da se preverja in certificira skladnost delovanja z mednarodnimi standardi. Poskrbeti je potrebno za usklajeno delovanje infrastruktur, kar se uresničuje s predpisovanjem zahtev glede uporabe mednarodnih metapodatkovnih standardov in uveljavljenih pristopov pri izvajanju storitev, kot so uporaba trajnih identifikatorjev, obveznost citiranja podatkov ipd. Pri tem EU spodbuja izmenjavo dobrih praks in pristopov, ki jih ponujajo posebej usmerjeni zgoraj naštetih razvojni projekti.

V zahtevah EU je poudarjena tudi potreba po izmenjavi tehnologij in dolžnost aktivnega sodelovanja ter spodbujanja dialoga med različnimi deležniki na nacionalni, evropski in mednarodni ravni. Tam so našteje tudi posamične teme, ki naj bodo v teh dialogih deležne večjega poudarka, kot so razprava z založniki glede publikacij in z njimi povezanih podatkov, rešitve za bibliografsko upoštevanje raziskovalnih podatkov, predanih v odprti dostop, in ustrezno nagrajevanje prispevka, kar naj bo predmet obravnave na medsebojnih srečevanjih financerja, predstavnikov raziskovalcev in ponudnikov infrastrukturnih storitev, ter na splošno aktivna promocija in uresničevanje načel odprtega dostopa v dialogu z različnimi javnostmi.

Pomemben korak za zagotovitev uresničevanja naštetih ukrepov na ravni države, ki ga zahtevajo priporočila komisije EU, je tudi imenovanje nacionalnega predstavnika za vprašanja odprtega dostopa. Ta sodeluje pri koordinaciji nacionalnih aktivnosti in v mednarodnih povezavah, njegova naloga pa je tudi poročanje o izvajanju ukrepov. V okviru izvajanja akcijskega načrta je potrebno natančneje opredeliti naloge predstavnika in mu zagotoviti podporo pri njihovem

opravljanju. Akcijski načrt za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v Sloveniji, ki je nastal v okviru CRP Odprti podatki, podajamo v nadaljevanju kot dejanski zaključek in povzetek celotnega projekta.

AKCIJSKI NAČRT

Sledeč Načelom in usmeritvam OECD, so **raziskovalni podatki** na splošno digitalno berljivi zapisi o dejstvih, ki predstavljajo osnovno podlago za znanstveno raziskovanje in ki v okviru znanstvene skupnosti veljajo kot ustrezno sredstvo za preverjanje veljavnosti raziskovalnih spoznanj. Podrobne opredelitve vsebine, pomena in oblike raziskovalnih podatkov so povezane z razumevanjem in znanjem posameznih skupnosti uporabnikov in ustvarjalcev podatkov znotraj znanstvenih področij. Poudarek pričujočega Akcijskega načrta je na dostopu do čim bolj osnovnih podatkov, pripravljenih in dostopnih v takšni obliki, da jih je mogoče uporabiti za različne druge namene. Iz opredelitve so izključene znanstvene objave v besedilni obliki (članki in podobno).

Odprti dostop do raziskovalnih podatkov, podprtih z javnimi sredstvi, pomeni, skladno z mednarodnimi priporočili OECD⁵²¹ in Komisije EU⁵²², po možnosti brezplačen dostop preko interneta za katerega koli uporabnika. Razumne minimalne omejitve dostopa, sorazmerne s tveganjem, se dopustijo kot izjema, ki jo je potrebno posebej utemeljiti in opravičiti. Dopuščene so izjeme zaradi tveganj v primerih ogrožanja zasebnosti, uničenja kulturnih ali naravnih dobrin, ogrožanja državne varnosti, poslovnih skrivnosti ali pravic intelektualne lastnine in konkurenčnosti države.

ODGOVORNOST MIZŠ IN NAJVIŠJIH PREDSTAVNIKOV ZNANSTVENE SKUPNOSTI IN PROFESIONALNIH ZDRUŽENJ

Problem	Cilj	Odgovornost	Zgledi
Deljenje raziskovalnih podatkov v odprtem dostopu je premalo	Opredeliti trajen in kar se da odprt dostop do raziskovalnih podatkov kot razvojno prioriteto na področju	Oblikovalci politik in vplivni predstavniki strokovne in znanstvene	Dokumenti OECD, Evropske komisije, Skupna izjava RCUK ⁵²³ , Deklaracije LERU ⁵²⁴ , LIBER ⁵²⁵ , Science

⁵²¹ OECD, 2007. *Principles and Guidelines for Access to Research Data from Public Funding*. Dostopno prek: <http://www.oecd.org/sti/scienceandtechnologypolicy/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm>, 5. 11. 2013

⁵²² European Commission Research & Innovation. *Policy Initiatives*. Dostopno prek: <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301>, 5. 11. 2013.

⁵²³ Research Councils UK, 2012. *RCUK Common Principles on Data Policy*. Dostopno prek: <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>, 5. 11. 2013.

⁵²⁴ League of European Research Universities. 2012. *Open Research Data*. Dostopno prek: <http://www.leru.org/files/general/Open%20Access%20to%20Research%20Data-FINALdocx.pdf>, 5. 11. 2013.

⁵²⁵ Ligue des Bibliothèques Européennes de Recherche. 2012. *Ten recommendations for libraries to get started with research data management*. Dostopno prek: <http://www.libereurope.eu/sites/default/files/The%20research%20data%20group%202012%20v7%20final.pdf>, 5. 11. 2013.

izkoriščen potencial za napredovanje raziskovanja.	raziskovalnih usmeritev.	javnosti (MIZŠ, Svet za znanost in tehnologijo, ARRS, SAZU, NUK).	Europe ⁵²⁶ , ALLEA ⁵²⁷ itd.
--	--------------------------	---	---

Utemeljitev

Koristi od odprtega dostopa do raziskovalnih podatkov so na splošno zelo jasno izražene, pri tem pa je odgovornost vseh sodelujočih, od ustvarjalcev podatkov v okviru znanstvene skupnosti in uporabnikov do vodstev inštitucij in financerjev skupaj z resornim ministrstvom, da v vsakem konkretnem primeru in znotraj disciplinarnih okvirov presoajo izražene koristi od odprtega dostopa do podatkov v primerjavi z vloženimi stroški in naporji za zagotavljanje dostopa, ki tudi niso zanemarljivi. Pomemben je tudi interes javnosti, da so podatki dostopni in da se čim bolj izkoristi njihov potencial. Nazoren primer je želja bolnikov, ki so sodelovali v kliničnih preizkusih, da so podatki na voljo čim več znanstvenikom, seveda pod pogoji zaščite zasebnosti, da se poveča možnost za odkrivanje načinov zdravljenja in na ta način povrne njihov vložek zaradi tveganj in naporov sodelovanja pri preizkusu. Korist ni izražena nujno samo za konkretnega bolnika, pač pa za celotno človeštvo, če se izboljšajo teoretična in uporabna spoznanja.

Na splošno velja, da so pri presoji proračunskih vlaganj v znanost za financerja prisotne izrazite ekonomske koristi, če lahko do podatkov, katerih potencial za uporabo ni v celoti izkoriščen v okviru trajanja projekta, dostopa še kdo drug. S tem se namreč izognemo podvajanju ustvarjanja podatkov. Upošteva se širjenje interpretativnih okvirov na podlagi istih podatkov, njihove primerjave z drugimi podatki ter možnosti preverjanja izsledkov originalnega projekta na podlagi istih podatkov, za isti denar dosežemo tudi bolj kakovostna odkritja. Našteti je mogoče še mnoge druge koristi.

Obstoječi modeli izmenjave raziskovalnih podatkov so nezadostni za zagotovitev pravičnega in dovolj obsegajočega dostopa. Pravičnega v smislu enakih obveznosti za vse, ki podatke ustvarjajo, in enakih možnosti dostopa za vse uporabnike. Preseči je potrebno občutek, da tisti, ki sodeluje v sistemu odprtega dostopa in prispeva svoje podatke, v zelo konkurenčno usmerjenem okolju znanosti izgublja na račun tistih, ki svoje podatke monopolizirajo. Obstoječi sistem je tudi premalo obsegajoč, saj ne zajema večine najpomembnejših raziskovalnih podatkov z največjim potencialom za drugo rabo, kar bi moral biti cilj prihodnjega izboljšane sistema dostopa do raziskovalnih podatkov. Slovenski financerji raziskav in raziskovalne organizacije še nimajo izdelanih politik odprtih raziskovalnih podatkov, ali pa te niso dovolj jasno artikulirane glede določanja obvez in spodbud zagotavljanja odprtega dostopa.

V mednarodni skupnosti obstaja soglasje glede pomena področja in strateških usmeritev. Večina najpomembnejših mednarodnih organizacij je tako že sprejela deklaracije in pobude za

⁵²⁶ European Science Foundation in EUROHORC. 2009. *EUROHORCs and ESF Vision on a Globally Competitive ERA and their Road Map for actions*. Dostopno prek :

http://www.eurohorcs.org/SiteCollectionDocuments/ESF_Road%20Map_long_0907.pdf, 5. 11. 2013; Inicijativno je prevzela delovna skupina za *Research Data* na Science Europe: <http://www.scienceeurope.org/policy/working-groups>, 5. 11. 2013.

⁵²⁷ European Federation of Academies of Science and Humanities. *Open Science for the 21st Century*. Dostopno prek: <http://www.allea.org/Pages/ALL/31/397.bGFuZz1FTkc.html>, 5. 11. 2013.

vzpostavitev sistema odprtega dostopa do podatkov. K temu se je zavezala tudi Republika Slovenija ob pridruževanju OECD, s tem da je izrazila podporo Načelom in usmeritvam OECD na tem področju, ter s sprejemom obveznosti pri uresničevanju Priporočil Evropske komisije in podpora pri njenih ostalih aktivnostih na področju odprtega dostopa do znanstvenih informacij. Resorno ministrstvo je že naredilo nekatere pomembne korake na področju z Evropo povezanega razvoja raziskovalne infrastrukture, ki vsebuje izboljššan dostop do raziskovalnih podatkov. Sem lahko štejemo tudi podporo projektu Odprti podatki in napore pri usmerjanju sredstev iz strukturnih skladov v spodbujanje vzpostavljanja infrastrukture področnih podatkovnih središč.

Potreben pa je čas, aktivna promocija idej in jasno izražena volja najodgovornejših za razvoj znanosti za njihovo uresničenje, da se že doseženo soglasje o pomenu odprtega dostopa do raziskovalnih podatkov razširi v različna disciplinarna in institucionalna okolja ter da se s postopnimi koraki in z oporo na zametkih dobrih praks vzpostavijo pogoji za izboljšano delovanje sistema. Zato je ključna vloga MIZŠ, da pokaže odločenost za uvajanje sprememb, ki pa morajo potekati s stalnim prilagajanjem na konkretne razmere za doseganje stroškovne učinkovitosti in ob podpori zlasti predstavnikov znanstvene skupnosti, kot so društva, sveti, akademija znanosti, vodstva inštitutov in ugledni posamezniki. Spremembe je potrebno uvajati tako, da bodo finančno učinkovite, pravične z vidika vseh prizadetih in da bodo prinesle očitno izboljšanje kakovosti raziskovanja. Pomembno pa je tudi zagotavljanje zaupanja splošne javnosti, ko gre za organiziranje znanstvenega sistema na pregleden način, vključno z možnostjo dostopa do raziskovalnih podatkov s strani laikov.

Spremembe je treba ponuditi kot priložnost za vse vključene, da na inovativne načine prevzamejo svoje vloge pri njihovem uvajanju na najbolj učinkovit in smiseln način v okviru svojih pristojnosti. Upoštevati je potrebno tudi interes raziskovalcev in raziskovalnih organizacij, ki so prispevali k izvedbi raziskav. Ti imajo pravico za določeno razumno omejeno obdobje pridržati uporabo raziskovalnih podatkov za namene v okviru svojih projektov. Upoštevati in ustrezno spodbujati je potrebno dodatno delo, vloženo v pripravo raziskovalnih podatkov za odprti dostop. Hkrati pa je potrebno na zgledih pokazati, da ustrezno podprt dostop do raziskovalnih podatkov tudi kot samostojen produkt s svojimi povezavami dodatno prispeva k prepoznavnosti raziskovalne odličnosti celotnega raziskovalnega izplena raziskovalcev in raziskovalnih organizacij ter s tem na njihovo odmevnost.

Zadolžitve	Koraki
<p><i>Strateško načrtovanje</i></p> <p>Področje odprtega dostopa do raziskovalnih podatkov vključiti v strateške in delovne dokumente ter razprave o prednostnih razvojnih usmeritvah znanosti in znanstvene infrastrukture.</p>	<p>Pregled, preverjanje uresničevanja ter periodično posodabljanje strateških dokumentov države na področju znanosti in kulture, kot so:</p> <ul style="list-style-type: none"> - RISS, - NRRI – Načrt razvoja raziskovalnih infrastruktur in v tem okviru poudarek na vzdrževanju nacionalne infrastrukture odprtega dostopa, ki je integrirana z mednarodnimi organizacijami s seznama ESFRI (European Strategy Forum on Research Infrastructures),

	<ul style="list-style-type: none"> - Informacijska družba, - Sodelovanje in medsebojna koordinacija pristopov k zagotavljanju odprtega dostopa do znanstvenih in kulturnih digitalnih vsebin, npr. Digitalna agenda v kulturi⁵²⁸. <p>(2014–2019, ob periodičnem posodabljanju strateških dokumentov)</p>
<p><i>Sprejem krovne nacionalne politike odprtih raziskovalnih podatkov.</i></p> <p>Jasno določene in med sabo usklajene vloge in odgovornosti raziskovalcev in raziskovalnih organizacij, uporabnikov podatkov in nosilcev podatkovnih storitev.</p>	<ol style="list-style-type: none"> a. Sopodpis Priporočil z dne 5. 12. 2012 (različni nosilci vlog, leto 2014). b. Deliberacija glede določil politik med predstavniki raziskovalnih organizacij, uporabniki podatkov in nosilci podatkovnih storitev (prva polovica 2014). c. Potrditev politike v ustreznih odločevalskih telesih (konec 2014).
<p><i>Odgovornost za izvajanje nalog nacionalnega predstavnika</i></p> <p>Zagotoviti delovanje nacionalnega predstavnika za vprašanja odprtega dostopa pri koordinaciji nacionalnih aktivnosti in v mednarodnih povezavah, z jasnimi usmeritvami glede odprtega dostopa do raziskovalnih podatkov.</p>	<ol style="list-style-type: none"> a. Oblikovanje letnega načrta delovnih nalog nacionalnega predstavnika na področju uresničevanja akcijskega načrta. Delovne naloge vključujejo: <ul style="list-style-type: none"> - sodelovanje pri koordinaciji priprave in uresničevanja strategij in politik posameznih področnih izvajalcev storitev, povezanih z raziskovalnimi podatki, - organiziranje medsektorskih posvetovanj z različnimi predstavniki znanstvene skupnosti, - deliberacija glede nadaljnjih korakov razvoja, - vključevanje v mednarodna posvetovanja in poročanje o sprejetih odločitvah doma in po svetu. <p>(Rok konec 2013, letno)</p> b. Poročilo o izvajanju, predano ministru za znanost, javno objavljeno in obravnavano na Koordinaciji predstavnikov izvajalcev (glej podrobneje poglavje Razvoj podatkovne infrastrukture, Koordinacija raziskovalnih podatkovnih storitev) (začetek 2014, letno).
<p><i>Promocija in vodenje dialoga s predstavniki</i></p>	<ol style="list-style-type: none"> a. Na spletnih straneh MIZŠ objavljene mednarodne zaveze

⁵²⁸ Glej: http://evsebine.ljudmila.org/Main_Page, 5. 11. 2013.

<p><i>različnih javnosti</i></p> <p>Seznanjanje javnosti z usmeritvami ter aktivno pridobivanje podpore pri najvišjih predstavnikih znanstvene skupnosti in raziskovalnih ustanov. To vključuje pripravljenost sprejemanja usklajenih strategij, politik in drugih ukrepov, pri katerih kot splošno pravilo velja dolžnost, da morajo biti raziskovalni podatki ponujeni v odprti dostop, morebitne izjeme pa posebej utemeljene. Raziskovalcem z zagotavljanjem podpornih storitev sistem omogoča, da pripravijo in predajo ustvarjene podatke na primerno mesto, jih pri tem spodbuja in nagraduje.</p>	<p>države (Načela in usmeritve OECD, Priporočila Evropske komisije) in ostalih predstavnikov deležnikov v znanosti (LERU, Science Europe in njena delovna skupina Research Data, LIBER, ALLEA ...) (2013).</p> <p>b. Dostop do raziskovalnih podatkov se uvrsti na dnevni red sestankov odločevalskih teles, predstavnikov znanstvene skupnosti in strokovnih združenj, na katerih nacionalni predstavnik za področje odprtega dostopa poroča o usmeritvah in dosežkih in prevzema predloge (leto 2014, na dve leti).</p> <p>c. Organiziranje medsebojnih srečanj predstavnikov različnih zainteresiranih skupin za delo s podatki (ustvarjalci, uporabniki, podatkovna središča, oblikovalci politik, znanstvena in strokovna združenja, založniki, skrbniki podatkov javnega sektorja) (MIZŠ, začetek 2014, letno).</p> <p>d. Pregled, strokovna posvetovanja in vlaganje pobud za spremembe nacionalne in EU zakonodaje ter prakse izvajanja glede varovanja zasebnosti, IPR, avtorskih pravic, varovanja nacionalne dediščine ipd.: legitimni cilj uporabe podatkov za raziskovalne namene se kot upoštevanja vreden postavi ob bok splošnih določil. (MIZŠ, tekoče)</p>
<p><i>Odgovornost zagotoviti izvedljivost</i></p> <p>Zagotoviti pogoje izvedljivosti akcijskega načrta.</p>	<p>a. Zagotovitev ustreznih finančnih sredstev za izvajanje aktivnosti. (MIZŠ, 2013, letno)</p> <p>b. Identifikacija morebitnih ovir pri izvajanju akcijskega načrta in aktivno prilagajanje. (MIZŠ, 2013, letno)</p> <p>c. Merjenje učinkov akcijskega načrta in sprotna skrb za izpolnjevanje obveznosti z obveščanjem, opozarjanjem in uveljavljanjem. (MIZŠ, 2013, letno)</p>
<p><i>Poslovno sodelovanje</i></p> <p>Aktivno iskanje povezav s poslovnim okoljem.</p>	<p>a. Vključevanje iniciativ v programe razvoja infrastrukture ter druge oblike sodelovanja in sofinanciranja med javnimi zavodi in podjetji. (MIZŠ, univerze, inštituti, predstavniki gospodarstva, 2014, periodično na dve leti)</p> <p>b. Iskanje modelov in odpravljanje ovir za poslovno izkoriščanje dostopa do javnih podatkov, vključno z raziskovalnimi podatki, in obratno, javnega dostopa za raziskovalne namene do podatkov, nastalih v poslovnih okoljih. (Isti deležniki, dajo skupno pobudo za organiziranje konferenc, 2015)</p>

RAZVOJ PODATKOVNE INFRASTRUKTURE

Problem	Cilj	Odgovornost	Zgledi
Nizka stopnja razvitosti podatkovne infrastrukture.	Zagotoviti enakomerno razvitost podatkovne infrastrukture po področjih znanosti.	MIZŠ, izvajalci nacionalnih področnih podatkovnih storitev z vzpostavljenimi povezavami z mednarodnimi podatkovnimi središči, raziskovalne ustanove, specialne in visokošolske knjižnice (znanstvene knjižnice) v raziskovalnih organizacijah in univerzah.	UK Data Service ⁵²⁹ iz VB; ADP ⁵³⁰ v Sloveniji; delovanje JISC ⁵³¹ v VB kot zgled financiranja razvojnih projektov; izsledki projekta DigCurV ⁵³² kot primer izboljševanja informacijske pismenosti v segmentu dela z raziskovalnimi podatki.

Utemeljitev

Razvita podatkovna infrastruktura na nacionalni ravni zagotavlja celovit pregled nad ustvarjenimi raziskovalnimi podatki in njihovim mestom dostopa. To nalogo evidentiranja in ovrednotenja glede dolgotrajnega pomena raziskovalnih podatkov za znanost opravljajo nacionalna področna podatkovna središča. Pri tem sledijo načelom stroškovne upravičenosti: v večini primerov zelo intenzivna obravnava podatkov, ki vključuje tudi postopke digitalnega skrbništva za zagotavljanje dolgotrajnega dostopa, mora biti v sorazmerju s pomenom in trajno vrednostjo podatkov. Raziskovalci kot ustvarjalci podatkov morajo imeti možnost ponuditi ustvarjene raziskovalne podatke v pooblaščenem področno podatkovno središče za zagotavljanje enakega dostopa do podatkov za vsakogar.

Podatkovno središče podatke skupaj s spremljajočo dokumentacijo najprej pregleda in ovrednoti po vnaprej opredeljenih kriterijih pomena za znanost in skladno s poslanstvom podatkovnega središča. V primeru, da podatki zadostijo merilom izbora, jih podatkovno središče kot primarno priporočeno mesto predaje prevzame v digitalno skrbništvo in zagotovi ustrezen trajen dostop, obenem pa se ustvarjalcu podatkov dodeli točke v bibliografiji za podatkovno objavo (podrobneje o notranji ureditvi področnih podatkovnih središč in njihovih podpornih storitvah v 3. in 4. poglavju).

Prototip razvite podatkovne infrastrukture je lahko delovanje podatkovnega središča za družboslovje, Arhiva družboslovnih podatkov, povezanega v mednarodno enoto CESSDA, za katero opravlja nacionalne podatkovne storitve. Druga znanstvena področja imajo svoje vsebinske posebnosti, ki jih je pri vzpostavitvi potrebno upoštevati, obenem pa slediti nekaterim skupnim organizacijskim pristopom za bolj učinkovito upravljanje. Nekatera področja že imajo posamezne zametke dejavnosti, ki pa jih je potrebno povezati in spodbuditi, da se vzpostavijo kot polno delujoča podatkovna središča, ki so integrirana v mednarodno okolje. Akcijski načrt

⁵²⁹ UK Data Service. Dostopno prek: <http://ukdataservice.ac.uk/>, 5. 11. 2013.

⁵³⁰ Arhiv družboslovnih podatkov. Dostopno prek: <http://www.adp.fdv.uni-lj.si/>, 5. 11. 2013.

⁵³¹ JISC Data Services & Collections. Dostopno prek: <http://www.jisc.ac.uk/whatwedo/topics/dataservices.aspx>, 5. 11. 2013.

⁵³² Digital Curator Vocational Education Europe. Dostopno prek: <http://www.digcur-education.org/>, 5. 11. 2013.

predvideva seznam ukrepov za prehodno pripravljalo obdobje, preden lahko začno področna podatkovna središča delovati v polnem obsegu. Namen je, da se spodbudi začetne aktivnosti pri vzpostavljanju področnih podatkovnih središč, in sicer najprej z identificiranjem potencialnih nosilcev in določanjem pravil njihovega delovanja, pilotnim poskusnim delovanjem, šele nato pa s spremljanjem in nadzorom delovanja polno delujoče infrastrukturne enote. Pomembno je stalno spodbujati inovativno dejavnost na različnih področjih opravljanja storitev v povezavi z raziskovalnimi podatki, čemur so tudi namenjeni pilotni razvojni projekti. Pri načrtovanju celovite nacionalne podatkovne infrastrukture pa je s koordinacijo potrebno zagotoviti dopolnjevanje in preprečiti prekrivanje storitev v okviru različnih institucij in v različnih oblikah delovanja. Podatkovne storitve, če so pazljivo načrtovane skozi različna obdobja v življenjskem krogu podatkov, se dopolnjujejo z različnimi nosilci najprej v raziskovalnih ustanovah v času trajanja projekta v okviru raziskovalnih skupin, nato kot možnost v vmesnem obdobju v okviru institucionalnih repozitorijev in podprte kot razširjena dejavnost raziskovalnih knjižnic ter dolgoročno kot storitev digitalnega skrbništva in organiziranja dostopa do podatkov kot dejavnost med sabo povezanih področnih podatkovnih središčih.

Zadolžitve	Koraki
<p><i>Odgovornost za trajen obstoj mreže področnih podatkovnih središč</i></p> <p>Ustanovitev in skrb za vzdržno delovanje področnih podatkovnih središč.</p>	<p>a. Področna podatkovna središča se ustrezno umestijo v sistem financiranja raziskovalnih infrastrukturnih programov (MIZŠ, obdobje 2014–2019, odvisno od dinamike vzpostavljanja storitvenih kapacitet področja).</p> <p>b. Storitve nacionalne infrastrukture s področja hrambe in dostopa do raziskovalnih podatkov imajo dolgoročno zakonsko opredeljeno zagotovilo delovanja v nacionalnih strategijah in načrtih (tekoče).</p> <p>c. Področna podatkovna središča imajo zagotovilo nasledstva. (Potrditev točke v predlogu DOLOČIL POLITIKE ODPRTEGA DOSTOPA: MIZŠ zagotavlja kontinuiteto dejavnosti skrbi za podatke; naloga je izpolnjena s sprejemom dokumenta, leto 2014)</p> <p>d. Pokrivanje glavnih raziskovalnih področij glede na dolgoročne potrebe po hrambi in dostopu do raziskovalnih podatkov, koordinacija in skrb, da se storitve ne podvajajo in da se povezujejo znotraj področij s ciljem, da se izkoristi obstoječe znanje in omogoči širjenje znanja ter storitev znotraj področij ali medpodročno v povezavi z določenimi tipi podatkov. Nujno je poskrbeti za kontinuiteto dejavnosti, zlasti upoštevati že delujoče infrastrukturne enote s seznama NRRI in ESFRI, ki opravljajo z raziskovalnimi podatki povezane storitve, ter z dodeljevanjem dodatnih pristojnosti graditi na obstoječem, da se izkoristi že dosežena raven storitev in ohrani strokovno znanje. Širitev obstoječih storitev ali oblikovanje mreže storitev sledi značilnostim sorodnih tipov podatkov in skupnosti znanja, ki so jim podatkih namenjeni. (MIZŠ oz.</p>

	<p>ARRS izda poziv za zbiranje predlogov za projekte vzpostavitve oz. nadgradnje področnih in/ali omrežnih podatkovnih središč; koordinira pobude in izbere določene izvajalce oz. njihove konzorcije – leto 2015)</p> <p>e. Področna podatkovna središča dobijo pooblastilo za storitve za izboljšanje dostopa do podatkov javnega sektorja za raziskovalne namene, v prvi vrsti registrskih podatkov, evidenc in javnih meritev. Skrbniki podatkov s področja javnega sektorja so dolžni pri tem sodelovati s ciljem, da se zmanjšajo ovire za dostop do podatkov in da se zagotovi skupno izvajanje storitev dodane vrednosti, kot je delo na urejanju in dokumentaciji podatkov, uporabi standardnih metapodatkov, registraciji, digitalni hrambi, iskanju, povezovanju podatkov in vključevanju v kataloge. Področje dostopa do raziskovalnih podatkov se usklajuje z drugimi pobudami odprtih podatkov javnega sektorja npr. za ureditev dostopa in hrambe registrskih podatkov za znanstvene namene (leta 2014–2019).</p>
<p><i>Financiranje razvoja podatkovnih storitev</i></p> <p>Spodbuditi inovativne pristope pri zagotavljanju podatkovne infrastrukture v obliki pilotnih projektov, npr. ponudba usposabljanja in svetovanja uporabnikom, posebne storitve orodij za integracijo, analizo, iskanje, povezane s posameznimi podatkovnimi tipi.</p>	<p>a. Zbiranje predlogov za izvajanje dodatnih storitev na področju dela z raziskovalnimi podatki s strani raziskovalnih organizacij in skupin, ki so zaradi svojih izkušenj in poznavanja posameznih tipov raziskovalnih podatkov primerni za izvajanje teh storitev, kot so svetovanje in usposabljanje za delo s podatki. (Instrumenti, usmerjeni k izboljšanju izkoriščanja odprto dostopnih raziskovalnih podatkov: 2014, na dve leti)</p> <p>b. Razvoj dejavnosti storitev, povezanih s podatki, se posebej izpostavi v okviru razpisov projektov s področja informacijske družbe. (Direktorat za informacijsko družbo, 2014, ob razpisih)</p> <p>c. Na področjih, kjer trenutno ni nosilcev dejavnosti področnih podatkovnih središč, ki bi lahko prevzeli predvidene naloge, skupine kandidirajo za projekte priprave načrtov za vzpostavitev izvajanja področnih podatkovnih storitev. (MIZŠ oz. ARRS koordinira poenotenje predlogov in izbere nosilce, 2014, na dve leti)</p>
<p><i>Podpora znanstvenim knjižnicam za delo z raziskovalnimi podatki</i></p> <p>Ciljno usmerjanje delovanja znanstvenih knjižnic v podporo</p>	<p>Zagotoviti pogoje za:</p> <p>Vključevanje v podporo dela s podatki na domačih ustanovah. Delo obsega tudi svetovanje ustvarjalcem in uporabnikom podatkov ter sodelovanje v raziskovalnih projektih pri pripravi raziskovalnih podatkov za odprti dostop, npr. pri pisanju načrtov ravnanja z raziskovalnimi podatki, pripravi dokumentacije raziskovalnega projekta in metapodatkov v</p>

<p>izvajanja podatkovnih storitev.</p>	<p>povezavi z raziskovalnimi podatki. (Ustanove opredelijo vloge in odgovornosti v svojih strategijah in politikah za ravnanje z raziskovalnimi podatki, 2014–2019, glej točko v DOLOČILIH POLITIKE ODPRTEGA DOSTOPA: Raziskovalne ustanove sprejmejo notranje politike; potrebno določiti vire financiranja dodatne dejavnosti)</p>
<p><i>Koordinacija raziskovalnih podatkovnih storitev</i></p> <p>Koordinacija delovanja podatkovnih središč in ostalih ponudnikov storitev.</p>	<p>a. V ta namen se ustanovi posvetovalno telo Koordinacija raziskovalnih podatkovnih storitev, ki ga sestavljajo predstavniki ponudnikov storitev v celotnem spektru disciplin in oblik delovanja (ustanovitev prva polovica leta 2014, sestanki najmanj dvakrat letno, oblikovane delovne skupine glede na potrebe).</p> <p>b. Zbiranje predlogov za vzpostavitev medsektorskih in medinstitucionalnih storitev z namenom povezovanja tehnoloških pristopov, standardov in organizacijskih pristopov. (Oblikuje se ločen tematski sklop znotraj razpisov CRP ipd., leto 2015)</p>
<p><i>Usposabljanje za delo s podatki</i></p> <p>Izobraževanje za podatkovne poklice in veščine.</p>	<p>a. Spodbujanje vključevanja vsebin podatkovnih veščin v univerzitetne programe usposabljanja prihodnjih področnih raziskovalcev in izvajalcev storitev (informatiki, knjižničarji ipd.). MIZŠ da pobudo za dopolnitev Meril in kazalnikov informacijske pismenosti (leto 2014) – medpodročna delovna skupina pripravi dopolnila in načine izvajanja usposabljanja (leto 2015).</p> <p>b. Dodatno usposabljanje knjižničarjev znanstvenih knjižnic. (Podatkovna središča, ustvarjalci podatkov, NUK pripravijo seminarje in jih vključijo v redne programe usposabljanja; MIZŠ razpiše poziv za sodelovanje in financira izvajanje programa, leto 2014, kontinuirano)</p> <p>c. Program usposabljanja podatkovnih skrbnikov v podatkovnih središčih in na ustanovah. (Program pripravi Koordinacija podatkovnih središč, leto 2015)</p> <p>d. Vključevanje v pobudo projekta Digital Curator Vocational Education Europe. (Delovna skupina implementira zaključke projekta na nacionalni ravni – leto 2015)</p>

MOTIVIRANJE RAZISKOVALCEV IN RAZISKOVALNIH USTANOV ZA IZPOLNJEVANJE NALOG ODPRTEGA DOSTOPA DO PODATKOV

Problem	Cilj	Odgovornost	Zgledi
Nizka kultura deljenja raziskovalnih podatkov, različna stopnja zavedanja pomena po disciplinah, odsotnost pravil, nezadostno nagrajevanje, druge ovire.	Vzpostaviti ekosistem odprtega dostopa do podatkov, ki vključuje kulturne, pravne, organizacijske, tehnološke vidike in vidike človeških virov.	Ustvarjalci podatkov, raziskovalne ustanove; izvajalci podatkovnih storitev; predstavniki znanstvene skupnosti, javni financerji znanstvenih projektov.	Primeri politik financerjev glede obveznosti priprave načrtov ravnanja z raziskovanimi podatki (VB, ZDA); poročilo o pripravah politik v DCC; vodila etičnih kodeksov ESF in ALLEA ⁵³³ .

Utemeljitev

Raziskovalci večinoma slabo poznajo prednosti in koristi odprtih raziskovalnih podatkov. Oviro predstavljata prevladujoča odsotnost kulture deljenja podatkov in z vidika odprtega dostopa nestimulativen sistem vrednotenja znanstvenih dosežkov.⁵³⁴ Med raziskovalci je razširjena težnja po monopolizaciji podatkov. Različne študije sicer kažejo, da obstajajo posamezni primeri in področja, kjer je izmenjava podatkov bolj razširjena, vendar daleč od tega, da bi bila to splošna praksa in da bi bilo podatke lahko najti in pridobiti takrat, ko jih nekdo potrebuje.

Hramba in deljenje raziskovalnih podatkov sta pogosto prepuščena raziskovalčevi osebni presoji, znanju, izkušnjam in pripravljenosti. Raziskovalci pogosto nimajo časa, sredstev in znanja za dodatno delo z raziskovalnimi podatki, poročajo, da so preobremenjeni z različnimi administrativnimi zahtevami, in imajo občutek, da bo delo s podatki samo še ena obveznost na papirju.

Raziskovalci so v diskusijah podpirali predloge, da se pri raziskovalnih projektih uvede dodatek, ki bi bil namenjen pokrivanju predvidenih upravičenih dodatnih stroškov, ki bi nastali s posebno skrbjo za podatke ter z urejanjem metapodatkov in dokumentacije v skladu z navodili področnih podatkovnih središč.⁵³⁵ Osrednja spodbuda za večjo pripravljenost na sodelovanje ustvarjalcev podatkov je vrednotenje podatkov, predanih v področno podatkovno središče, kot znanstvenih prispevkov v osebnih bibliografijah.⁵³⁶ Sprejetje politik pripomore, da so vloge in odgovornosti posameznih nosilcev dejavnosti zagotavljanja dostopa do raziskovalnih podatkov jasno in za vse

⁵³³ ESF in ALLE, 2011. *The European Code of Conduct for Research Integrity*. Dostopno prek: http://www.esf.org/fileadmin/Public_documents/Publications/Code_Conduct_ResearchIntegrity.pdf, 23. 1. 2014.

⁵³⁴ »Dokler morajo raziskovalci razmišljati predvsem o tem, koliko točk jim bo prinesla neka objava, in ne o njeni vsebini in kvaliteti, bo njihov glavni cilj izigravanje sistema in ne kvalitetno delo.« (Iz korespondence z L. Š., pripomba na osnutek Akcijskega načrta, 14. 10. 2013).

⁵³⁵ Diskusijske pripombe na *Določila politike odprtega dostopa do raziskovalnih podatkov v Sloveniji* – verzija 2, 14. 3. 2013, z delavnice Politike ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi (Dostopno prek: <http://www.adp.fdv.uni-lj.si/odpp10D2/>, 5. 11. 2013).

⁵³⁶ Poročilo *Opis stanja*, str. 70. Dostopno prek: http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odpp10_opis_stanja/, 5. 11. 2013.

enako opredeljene, vključno z dolžnostmi raziskovalcev in njihovih organizacij, da sodelujejo s svojim prispevkom raziskovalnih podatkov v odprti dostop in da so za to pravično nagrajeni.

Dejavnosti promocije, podpore in svetovanja glede postopkov zagotavljanja odprto dostopnih podatkov niso ustrezno razvite, zaradi tega se raziskovalci počutijo negotovo, kako izpolnjevati svoje obveznosti. Zahteve raziskovalcem kot ustvarjalcem podatkov, da sodelujejo v sistemu odprtega dostopa, je potrebno podpreti z razvojem podporne storitvene dejavnosti, kot je opisano v prejšnjem razdelku. Vloga področnih podatkovnih središč pri tem je v jasnem oblikovanju zahtev za prevzem podatkov in oblikovanju storitev vrednotenja ter registracije podatkov, ki omogoča vključevanje v osebne bibliografije ustvarjalcev in bibliografsko navajanje pri drugi rabi. Znanstvene knjižnice in pisarne na ravni ustanov se morajo usposobiti za podporo pri pripravi načrtov ravnanja z raziskovalnimi podatki in izvajanja dejavnosti vmesne skrbi za raziskovalne podatke v obdobju pred predajo v področno podatkovno središče. Prvi korak v tej smeri na ravni ustanov je priprava notranjih politik ravnanja z raziskovalnimi podatki in vzpostavitev storitev za podporo delu s podatki za svoje raziskovalce.

Ukrep za zagotovitev preglednosti in odgovornosti pri izpolnjevanju vlog v celovitem ekosistemu zagotavljanja dostopa do podatkov je uvajanje obveznosti priprave načrtov ravnanja z raziskovalnimi podatki (NRRP). NRRP gre skozi postopek recenzije predlogov raziskovalnih projektov. Sistem zagotavlja strokovno sodelovanje ustvarjalcev podatkov in izvajalcev podatkovnih storitev že pri načrtovanju in nato skozi vse faze življenjskega cikla raziskovalnih podatkov. Tako se lahko v večji meri izognemo pomanjkljivostim na poti do odprtega dostopa do podatkov, saj je zagotovljena skrb za formalne in pravne vidike priprave podatkov ter na koncu večja kakovost in širša uporabnost ustvarjenega, če je to ustvarjeno z mislijo na prihodnjo dostopnost. NRRP ponuja tudi možnost opredelitve utemeljenih razlogov, da se podatkov ne nudi v odprtem dostopu, npr. ko gre za varovanje strateških interesov in zaščito znanja, ki ga je financirala RS.⁵³⁷

Zadolžitve	Koraki
<p><i>Sprejem jasno določenih obveznosti izvajalcev raziskovalnih projektov</i></p> <p>Sprejem zavezujočih določil v obliki javno objavljenih politik glede obveznosti izvajalcev raziskovalnih projektov z javnim financiranjem, ki – ob upoštevanju pravnih in etičnih obveznosti – vsem zagotavljajo dostop in uporabo podatkov, ki so nastali v okviru javno financiranih raziskav.</p>	<p>Ustvarjalci raziskovalnih podatkov so za popolnitev in vzdržno delovanje področnih podatkovnih središč dolžni ponuditi podatke ob nastanku, najkasneje pa ob zaključku projekta, v področno podatkovno središča na način, da ne kršijo zakonodaje s področja varstva osebnih podatkov in intelektualne lastnine. (Sprejem in javna objava krovnega dokumenta DOLOČILA POLITIKE ODPRTEGA DOSTOPA DO RAZISKOVALNIH PODATKOV V SLOVENIJI; določitev obveznosti v Pravilnikih ARRS; načrtovanje obveščanja prejemnikov projektov in usposabljanja za izvajanje obveznosti) (jeseni leta 2014, periodično posodabljanje na 2 leti)</p>

⁵³⁷ Glej poročila delavnic Politike ravnanja z raziskovalnimi podatki, financiranimi z javnimi sredstvi (dostopno prek: <http://www.adp.fdv.uni-lj.si/odpp10D2/>, 5. 11. 2013), in *Priporočila udeležencev delavnice Problemi in rešitve na področju podatkovnih storitev v Sloveniji*, Fakulteta za družbene vede, 5. 12. 2012. Dostopno prek: <http://www.adp.fdv.uni-lj.si/odpp10D1/priporocila.html>, 5. 11. 2013.

<p><i>Podpora za prispevek raziskovalnih podatkov</i></p> <p>Motiviranje ustvarjalcev podatkov za pripravo in izročanje raziskovalnih podatkov, pripravljenih po merilih stroke tako, da so razumljivi, pregledni in uporabni za druge uporabnike, v sistem odprtega dostopa, in sicer – če podatki ustrezajo – v pooblaščen področna podatkovna središča. Naloga področnih podatkovnih središč je ovrednotenje in izbor najpomembnejših podatkov za dolgotrajno hrambo in dostop ter napotitev preostanka podatkov na druga za njihov status primerna mesta.</p>	<p>a. Sprejem določila o financiranju raziskovalnih projektov:</p> <ul style="list-style-type: none"> - delež do 10 % upravičenih stroškov pri izvajanju raziskovalnih projektov je lahko namenjen za dodatno delo pri načrtovanju, ustvarjanju in pripravi raziskovalnih podatkov ter dokumentacije po zahtevah za predajo v okolje odprtega dostopa. Predvideni stroški se opredelijo v NRRP na posebnem obrazcu, upoštevaje posebnosti projekta. (Izvesti reformo sistema financiranja raziskovalnih projektov, 2015) <p>b. Zagotoviti izvajanje storitev podpore skrbništva nad podatki v raziskovalnih ustanovah za čas trajanja projekta in kot pripravo za predajo v področno podatkovno središče. (Odvisnost od sklopa Ustanove prevzamejo odgovornost v nadaljevanju. Nosilci: mešane infrastrukturne in raziskovalne skupine, inštitutske pisarne, znanstvene knjižnice ustanov v povezavi s področnimi podatkovnimi središči.)</p> <p>c. Zagotovi se pooblaščenemu mestu predaje podatkov in presoje ustreznosti podatkov za uvrstitev v točkovanje. (Pooblaščen podatkovna središča usposobljena za prevzem in izbor, obdobje 2014–2019, začetek po področjih odvisno od zaključka vzpostavitve področnih storitev)</p> <p>d. Objava podatkov visokega znanstvenega pomena in uporabne vrednosti za nadaljnjo rabo na za to pooblaščenem mestu se vrednoti s točkami v osebni bibliografiji⁵³⁸ in nagrajuje pri sledenju v Indeksu citiranosti podatkov⁵³⁹. (Sprejem DOLOČIL POLITIKE ODPRTEGA DOSTOPA; dopolnitev Pravilnika za vrednotenje znanstvenih prispevkov z upoštevanjem do tedaj nastalih pooblaščenih nacionalnih področnih podatkovnih znanstvenih središč, leto 2015)</p>
<p><i>Obveznost priprave NRRP</i></p> <p>Obveznost priprave načrta ravnanja z raziskovalnimi podatki (NRRP) v okviru projektne dokumentacije</p>	<p>Sprejem določila o obvezni pripravi NRRP, kjer je opredeljen način zagotavljanja kakovosti ter dolgoročne uporabnosti in dostopnosti podatkov v pooblaščenem podatkovnem središču oziroma na drugem ustreznem mestu, specializiranem za določen</p>

⁵³⁸ Glej *Tipologija dokumentov/del za vodenje bibliografij v sistemu COBISS* na http://home.izum.si/COBISS/bibliografije/Tipologija_slv.pdf, kategorija 2.20.

⁵³⁹ Npr. http://wokinfo.com/products_tools/multidisciplinary/dci/, 5. 11. 2013.

<p>novih raziskovalnih projektov.</p>	<p>tip podatkov. Obvezno je posebej utemeljiti, zakaj je potrebno ustvarjanje novih podatkov in zakaj za namene projekta ne zadoščajo že obstoječi odprto dostopni podatki. Posebej morajo biti utemeljene izjeme od sicer splošno veljavnega pravila, da morajo biti podatki obvezno ponujeni v odprti dostop. Financer presoja upravičenost izjeme v vsakem primeru posebej, npr. kadar bi bila povzročena škoda ali kadar koristi zaradi odsotnosti povpraševanja ne opravičujejo dodatnih stroškov oz. še ni vzpostavljene ustrezne infrastrukture za podporo zagotavljanja dostopnosti.</p> <p>(Določilo se uvede v razpisno dokumentacijo novih razpisov raziskovalnih projektov; uveljavi se presoja ustreznosti NRRP v okviru ocenjevanja predlogov raziskovalnih projektov in se temu ustrezno dodatno informira in usposobi recenzente; naloga ARRS, leto 2015; odvisnost: sprejem DOLOČIL POLITIKE ODPRTEGA DOSTOPA)</p>
<p><i>Registracija odprto dostopnih raziskovalnih podatkov za upoštevanje prispevka v znanosti</i></p> <p>Evidentiranje NRRP in v njih opredeljenih raziskovalnih podatkov, ustvarjenih v okviru raziskovalnih projektov, v informacijskem sistemu o raziskovalni dejavnosti v Sloveniji (SICRIS, COBISS).</p>	<p>a. Dopolnitev SICRIS za sprejem in informacijsko podporo osnovnih polj obrazcev NRRP v povezavi s financiranimi projekti. Minimalna določila glede načrtovanih raziskovalnih podatkov zadevajo predvideno pooblaščenno mesto dostopa, predvideni datum nastanka, in posebej pomembno, predvideni datum predaje v odprti dostop ter izbor ustrezne licence, priporočene za odprte podatke.</p> <p>b. Določanje pogojev in pravil za vnos v bibliografijo raziskovalca, kot so obvezna polja modela za citiranje, trajni označevalnik lokacije verzije podatkov, presoja znanstvenega pomena in izpolnjevanja formalnih zahtev področnih podatkovnih središč. Področni OSIC (osrednji specializirani informacijski center) pri točkovanju v okviru bibliografij upošteva samo pri področnem podatkovnem središču registrirano objavo raziskovalnih podatkov, ki je prestala ustrezno presojo. (Odvisnost: sprejeta in ustrezno zastavljena določila krovnih DOLOČIL POLITIKE ODPRTEGA DOSTOPA in politik področnih podatkovnih središč; naloga ARRS: dopolnjevanje Pravilnika o kazalnikih in merilih znanstvene in strokovne uspešnosti z novimi pooblaščenimi podatkovnimi središči)</p> <p>c. Zagotovi se možnost navajanja podatkov v originalnih znanstvenih publikacijah s pomočjo modela za</p>

	<p>navajanje, ki vključuje trajni označevalnik mesta dostopa do raziskovalnih podatkov. Omogoči se sledenje navedkov raziskovalnih podatkov v citatnih indeksih.</p> <p>(Vse navedene naloge zahtevajo sodelovanje IZUM-a, ARRS, OSIC in področnih podatkovnih središč, začetek delovanja leto 2015)</p>
<p><i>Vzpostavitev storitev za podporo pripravi NRRP</i></p> <p>Zagotovi se izvajanje usposabljanja in podpore pri pisanju NRRP. Posebna skrb je namenjena pripravi navodil in svetovalne dejavnosti glede pogojev etično in legalno dostopnih raziskovalnih podatkov za odpravljanje ovir pravne narave.</p>	<p>a. Področna podatkovna središča skupaj z drugimi izvajalci storitev izvajajo svetovanje in podporo pri pisanju NRRP. Določijo tudi modele glede izjem odprtega dostopa do podatkov.</p> <p>b. Dopolnila etičnih kodeksov in pravil delovanja etičnih komisij. (Delovne skupine strokovnih združenj poročajo Nacionalnemu predstavniku za vprašanja odprtega dostopa.)</p> <p>c. Dejavnost strokovnega svetovanja glede avtorskih pravic in varovanja zasebnosti za premagovanje ovir pri omejitvah dostopnosti. Oblikovanje dobrih praks bolj fleksibilnega pristopa pri izpolnjevanju zahtev zadevne zakonodaje. (Opravi se zbiranje prijav kandidatov za izvajanje dejavnosti svetovanja, usposabljanja in usmerjanja z vodiči in izda pooblastila; dejavnost lahko izvajajo tudi področna podatkovna središča in drugi usposobljeni izvajalci storitev, povezanih s podatki.)</p> <p>d. Pregledovanje zakonskih omejitev in dajanje pobud za spremembo zakonodaje glede avtorskega prava in prava glede varovanja zasebnosti. (Nacionalni predstavnik za vprašanja odprtega dostopa, Koordinacija raziskovalnih podatkovnih storitev, tekoča dejavnost)</p> <p>(Začetek leto 2014, kontinuirano)</p>
<p><i>Ustanove prevzamejo odgovornost</i></p> <p>Raziskovalne ustanove sprejmejo politike glede vmesne skrbi za podatke in vzpostavijo storitve skrbništva do točke predaje v pooblaščenno podatkovno središče.</p>	<p>a. MIZŠ zahteva skrb za vmesno (ob nastanku in v času trajanja projekta do trenutka predaje v ustrezno področno podatkovno središče) vzdrževanje in varno hrambo raziskovalnih podatkov ter pripravo za odprti dostop na ravni ustanov. Ustanove so dolžne razvijati dejavnost podpore priprave kakovostnih raziskovalnih podatkov s tem, da nudijo skupne storitve in orodja na ravni ustanov ter razbremenijo posamezne raziskovalne skupine. (Normativno urejena zahteva glede skrbi za podatke, realizacija</p>

	<p>projektov vzpostavitve podporne dejavnosti najkasneje leta 2019)</p> <p>b. Ustanove pripravijo notranje politike in strategije ravnanja z raziskovalnimi podatki in načrte delovanja podpornih podatkovnih storitev, na katere se opirajo in sklicujejo raziskovalci pri pripravi in izvajanju svojih raziskovalnih projektov. Politike naj bodo skladne z načeli odprtega dostopa do raziskovalnih podatkov. Na ravni ustanov naj vključujejo spodbude v obliki kariernih možnosti za tiste, ki opravljajo podatkovne poklice (podatkovni znanstvenik, podatkovni knjižničar, skrbnik podatkov) in nagrad za tiste, ki prispevajo raziskovalne podatke v odprti dostop (obveznost sprejeta v krovnih nacionalnih DOLOČILIH POLITIKE ODPRTEGA DOSTOPA, ustanove so dolžne svoje politike in strategije uskladiti najkasneje do leta 2019)</p>
<p><i>Vzpostavitev sistema nadzora na ravni ustanov</i></p> <p>Vzpostavi in opravlja se nadzor izpolnjevanja obveznosti raziskovalcev in ustanov.</p>	<p>a. Vzpostavi se delovanje svetov in drugih nadzornih organov raziskovalnih projektov in raziskovalnih ustanov za nadzor nad izpolnjevanjem obveznosti iz odprtega dostopa (normativna ureditev statutom, področna koordinacija notranjih politik ustanov in načrtov delovanja podpornih podatkovnih storitev; obveznost poročanja – leto 2014, periodično vsako leto).</p> <p>b. Preverjanje izpolnjevanja obveznosti predaje raziskovalnih podatkov na pooblaščenih mestih in izvajanje določil sankcij za neizpolnjevanje. V primeru neizpolnjevanja obveznosti umakniti plačilo zaključnega obroka financiranja ipd. (Področna podatkovna središča na podlagi evidenc NRRP v SICRIS poročajo o izvajanju, financer predvidi in izvaja sankcije; od leta 2015 dalje, po posameznih področjih najkasneje do leta 2019, odvisno od vzpostavljenih infrastrukture za prevzem raziskovalnih podatkov; potrebno ustrezno spremeniti zakonske podlage)</p>

VZDRŽEVANJE STORITEV ZA PODPORO PRI DELU Z RAZISKOVALNIMI PODATKI, VREDNOTENJE, IZBOR IN PREVZEM, DIGITALNO SKRBNISŤTVO IN ZAGOTAVLJANJE DOSTOPA

Problem	Cilj	Odgovornost	Zgledi
Raziskovalci različno opredeljujejo raziskovalne podatke; razlike so pri upoštevanju metapodatkovnih standardov; ni izoblikovanih meril in postopkov za zagotavljanje izbora, digitalne hrambe ter dostopa do raziskovalnih podatkov.	Vzpostaviti odgovornost izvajalcev podatkovnih storitev v obliki raziskovalne infrastrukture za izvajanje nalog za podporo pri pripravi NRRP in njihovem izvajanju, evidentiranju in podpori pri prevzemu raziskovalnih podatkov od ustvarjalcev, nadaljnji obravnavi podatkov, pregledu in dopolnitvi metapodatkov skladno s standardi digitalnega skrbništva ter zagotavljanju odprtega dostopa za uporabnike.	Področna podatkovna središča.	Vodiči in politike področnih podatkovnih središč (npr. ICPSR ⁵⁴⁰) in centrov za svetovanje, kot je DCC ⁵⁴¹ .

Utemeljitev

Celovit mozaik ekosistema odprtega dostopa sloni na jasno opredeljenih nalogah področnih podatkovnih središč, ki glede na uveljavljene pristope posameznih disciplin in skladno s povpraševanjem uporabnikov razvijajo svoje storitve na pregleden način in z jasno opredeljeno odgovornostjo. Delovanje mora biti usklajeno na nacionalni in mednarodni ravni za zagotovitev integracije storitev. Za področja znanosti, kjer že obstaja mednarodno uveljavljena infrastruktura za odprte raziskovalne podatke, se na nacionalni ravni izvede evidentiranje raziskovalnih podatkov, podatke se shrani v mednarodne digitalne arhive in dodeli točkovanje.

Oblikovati je potrebno kriterije in merila za kakovostno delovanje področnih in integriranih podatkovnih storitev, zagotoviti razvoj in svetovanje pri vzpostavljanju novih storitev ter preverjati skladnost delovanja podatkovnih storitev z merili/kriteriji.

Z vidika enakomerne zastopanosti vseh znanstvenih področij, katerih raziskovalci imajo obveznost v okviru javno financiranih projektov ustvarjene raziskovalne podatke predati na za odprti dostop najprimernejše mesto, je pomembno na nacionalni ravni medpodročno usklajeno razvijati storitve prevzema (*»ingest«*) in vrednotenja raziskovalnih podatkov kot dejavnosti pooblaščenih področnih podatkovnih središč. Predpogoj za tovrstno dejavnost je vzpostavitev področne infrastrukture v obliki storitev področnih podatkovnih središč, ki se med sabo koordinirajo pri minimalnih zahtevah za usklajeno izvajanje osrednjih skupnih storitev znotraj življenjskega kroga digitalnih podatkov po modelu OAI⁵⁴². Takšne osrednje storitve so storitve

⁵⁴⁰ ICPSR Data Management & Curation. Dostopno prek:

<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/>, 5. 11. 2013.

⁵⁴¹ Digital curation centre. Vodiči dostopni prek: <http://www.dcc.ac.uk/resources>, 5. 11. 2013.

⁵⁴² ISO 14721: 2012. *Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model* <http://public.ccsds.org/publications/archive/650x0m2.pdf>, 5. 11. 2013.

prevzema, vrednotenja, digitalne hrambe in dostopa. Obenem pa te storitve prilagajajo svoji ciljni skupnosti uporabnikov (»designated community«) in, denimo, razvijajo znotraj skupnosti razumljive zahteve in v stroki sprejeta merila glede kakovostnih znanstveno pomembnih raziskovalnih podatkov, pripravljenih za odprti dostop.

Za podatke, ki zahtevajo posebne pogoje dostopa, je ponekod uveljavljena rešitev dostop na obstoječi lokaciji, kot je na primer varna soba na Statističnem uradu RS. Takšen način se največkrat uporablja v primeru izredno občutljivih podatkov (npr. osebni podatki), za katere je potrebna najvišja stopnja zaščite in kadar področna podatkovna središča nimajo vzpostavljene tovrstne storitve. Ker je zagotavljanje takšne storitve precej zahtevno, bi jo bilo smiselno razvijati medpodročno, tako da bi zajela vse tipe podatkov, ki zahtevajo podobne posebne pogoje nadzorovanega dostopa, npr. v tem pogledu so si sorodni podatki s področja družboslovja in medicine. Tudi v teh primerih imajo področna podatkovna središča pomembno posredniško in svetovalno vlogo pri poenotenju priprave podatkov in pri upravljanju pogojev dostopa.⁵⁴³

Koordinirano je treba razvijati nacionalno povezane storitve trajnega digitalnega ohranjanja kot eno od področij, kjer se je pokazala potreba po skupnem strokovnem središču, katerega krovna naloga je uveljavljanje zahtev izpolnjevanja priznanih meril in standardov zaupanja vrednega organizacijskega delovanja. Razvoj področja trajnega digitalnega ohranjanja zajema tudi robustno med seboj povezano infrastrukturo podatkovnih središč ter ustanov nacionalnega pomena, kot so NUK in Arhiv Slovenije. Področje je primerno tudi za medsektorsko sodelovanje med ustanovami javnega sektorja s področij javne uprave, kulture in znanosti.

Zadolžitve področnih podatkovnih središč	Koraki
<p><i>Podatkovna središča prevzamejo odgovornost</i></p> <p>Vzpostaviti odgovornost področnih podatkovnih središč.</p>	<p>a. Področna podatkovna središča uskladijo svoje delovanje z določili nacionalnih politik s sprejetjem ustreznih notranjih strategij in politik (leto 2015 oz. ob ustanovitvi). (Odvisnost: Sprejem dokumenta DOLOČILA POLITIKE ODPRTEGA DOSTOPA DO RAZISKOVALNIH PODATKOV V SLOVENIJI⁵⁴⁴, glej poglavje Motiviranje raziskovalcev in raziskovalnih ustanov za izpolnjevanje nalog odprtega dostopa do podatkov, sklop Podpora za prispevek raziskovalnih podatkov, c)</p> <p>b. Področna podatkovna središča med seboj uskladijo minimalne kriterije kakovostne ravni opravljenih storitev. Zagotovijo spremljanje delovanja in poročanje. (Naloga koordinacije izvajalcev podatkovnih storitev, leto 2015)</p>
<p><i>Opravljanje področnih podatkovnih storitev za podporo raziskovalcem</i></p> <p>Razvoj storitev identifikacije raziskovalnih podatkov in</p>	<p>a. Področna podatkovna središča so v svojih politikah delovanja dolžna določiti kriterije in postopke izbora raziskovalnih podatkov, ki jih obravnavajo. Pri tem upoštevajo pomen ter dolgoročno znanstveno in uporabno vrednost podatkov, pa tudi stroške in bremena,</p>

⁵⁴³ Iz poročila prve delavnice, str. 8. Dostopno prek: <http://www.adp.fdv.uni-lj.si/odpp10D1/>, 5. 11. 2013.

⁵⁴⁴ Glej Prilogo 6: Določila politike odprtega dostopa do raziskovalnih podatkov v Sloveniji.

<p>njihovega vrednotenja ter integracija sistemov nadzorovanega dostopa.</p>	<p>povezane z zagotavljanjem trajne dostopnosti in uporabnosti.</p> <p>b. Zagotoviti vrednotenje in zajem metapodatkov za raziskovalne podatke iz mednarodnih arhivov, pri ustvarjanju katerih so sodelovali slovenski raziskovalci ali imajo poseben pomen za slovensko znanost in družbo.</p> <p>c. Zagotoviti prevzem raziskovalnih podatkov, kadar ni drugega ustreznega mesta za predajo raziskovalnih podatkov, povezanih z objavo. Pripisati obliko za primerno citiranje raziskovalnih podatkov v seznamu virov znanstvenih besedilnih objav.</p> <p>d. Zagotoviti izvajanje storitev avtentikacije in avtorizacije uporabnikov po protokolu AAI in zagotavljati odprti dostop do raziskovalnih podatkov z uporabo primernih licenc.</p> <p>(Leto 2014 oz. ob vzpostavitvi delovanja)</p>
<p><i>Zagotavljanje trajnega dostopa za prevzete raziskovalne podatke</i></p> <p>Razvoj in vzdrževanje storitev dolgotrajne digitalne hrambe.</p>	<p>a. Razviti način obravnave raziskovalnih podatkov skladno s standardi na področju digitalne hrambe za zagotavljanje trajnega skrbništva in dostopa (OAIS, DSA⁵⁴⁵, ISO 16363).</p> <p>b. Sodelovati v nacionalnem in mednarodnem omrežju usklajenega delovanja strojne infrastrukture in storitev digitalne hrambe.</p> <p>(Leto 2014 oz. ob vzpostavitvi delovanja)</p>
<p><i>Vzdrževanje in širjenje znanja s področja ravnanja s podatki</i></p> <p>Razvoj in promocija kakovostnih podatkovnih storitev.</p>	<p>a. Razvoj in promocija področnih dobrih praks in standardov pri načrtovanju ravnanja s podatki za posamezne tipe in vsebine podatkov. (Podatkovna središča organizirajo področna in medpodročna srečanja, seminarje, pripravijo vodiče in pripomočke, letno)</p> <p>b. Področna podatkovna središča glede na uveljavljene področne dobre prakse določijo formalne obveznosti glede predaje raziskovalnih podatkov (formati, metapodatki, standardi, licence, ostala dokumentacija) za zagotovitev preglednosti raziskovalnih podatkov pri iskanju in kot podlago za kakovostno drugo rabo s strani uporabnikov, vključno z naprednimi možnostmi hrambe, dostopa in medsebojnega povezovanja podatkov. Namen je maksimizirati potencial za drugo rabo, ki ga vsebujejo podatki. Spremljajo področje in vključujejo nove koncepte. (Način: formalne zahteve, ki jim morajo raziskovalni</p>

⁵⁴⁵ Glej: <http://datasealofapproval.org/en/>, 5. 11. 2013.

	podatki zadostiti za prevzem, vključene v javno objavljena merila za izbor podatkov s strani področnih podatkovnih središč; leto 2014 oz. ob ustanovitvi; periodično posodabljanje na dve leti)
<p><i>Sodelovanje pri koordinaciji izvajanja podatkovnih storitev</i></p> <p>Usklajeno z drugimi mednarodnimi in nacionalnimi infrastrukturami vrednotijo tehnološke, organizacijske in pravne pristope, prevzemajo dobre prakse in skrbijo za stroškovno učinkovit pristop pri razvijanju in izvajanju storitev.</p>	<p>a. Sodelovanje z nacionalnimi splošnimi repozitoriji, kot je dLib.si, repozitoriji ustanov⁵⁴⁶ ter bibliografskimi storitvami sistema COBISS, poteka na način zajema metapodatkov iz podatkovnih objav področnega podatkovnega središča po protokolu OAI-PMH. (Obveznost vzpostavitve storitve OAI-PMH za raziskovalne podatke in drugih posodobljenih protokolov – DataCite⁵⁴⁷ ipd. – na strani področnih podatkovnih središč; leto 2014 oz. ob ustanovitvi; periodično posodabljanje)</p> <p>b. Splošni repozitoriji in repozitoriji z objavami povezanih podatkov (Dryad⁵⁴⁸ ipd.) lahko prevzamejo raziskovalne podatke v svoj zajem v primeru, da se raziskovalni podatki zaradi premajhnega pomena glede na postavljene kriterije ne uvrstijo v zajem podatkovnega središča. Vrednotenje in usmerjanje ob upoštevanju stroškovne upravičenosti glede na korist opravi pooblaščen področno podatkovno središče, ki ima prednost pri prevzemu skrbništva nad najkakovostnejšimi podatki. Samo podatki, vključeni v področni podatkovni center, štejejo kot znanstvena objava pri točkovanju v bibliografiji. (Pogoj vključen v krovni dokument DOLOČILA POLITIKE ODPRTEGA DOSTOPA, ob sprejemu)</p> <p>c. Spodbuja se nacionalna in mednarodna izmenjava podatkov, metapodatkov, znanja, storitev. (Vključevanje področnih podatkovnih središč v evidenco SICRIS; Podatkovna središča so dolžna objaviti skupno dogovorjeni strojno berljivi profil, npr. re3data.org Vocabulary; sodelovanje z Digitalno agendo na področju kulture; predstavniki izvajalcev so člani Koordinacije delovanja podatkovnih središč, leto 2014)</p> <p>d. MIZŠ in področna podatkovna središča skrbijo za mednarodne integracije podatkovnih storitev s seznama</p>

⁵⁴⁶ Glej Nacionalni portal odprte znanosti na: <http://www.openscience.si/>, 5. 11. 2013.

⁵⁴⁷ Dostopno prek: <http://www.datacite.org/>, 5. 11. 2013.

⁵⁴⁸ Dostopno prek: <http://datadryad.org/>, 5. 11. 2013.

	<p>ESFRI in širše (2013 in dalje).</p> <p>e. Podpora sodelovanju v mednarodnih projektih razvoja storitev in kooperativnih iniciativah (zlasti KE⁵⁴⁹, RDA⁵⁵⁰, DataCite⁵⁵¹, OpenAIRE⁵⁵²). (MIZŠ in strokovna združenja dajo pobudo za pridružitve, imenujejo predstavnike, zagotovijo sredstva, leto 2015)</p> <p>f. Poudarki integriranih storitev na nacionalni ali mednarodni ravni, s katerimi se povezujejo področna podatkovna središča, so storitve trajnega ohranjanja digitalnih virov (sodelovanje z NUK, Arhivom Slovenije, javno upravo), uporaba tehnologij akademskega omrežja ARNES (GRID, AAI, ipd.), pravno in etično svetovanje ter pobude (organiziranje informacijskih točk, ki jih lahko vodijo pooblaščen izvajalci, aktivnosti SAZU, večja vloga etičnih komisij pri ocenjevanju in usmerjanju projektov za izpolnjevanje pogojev odprtega dostopa); skupna storitev nadzorovanega dostopa do občutljivih podatkov (področna podatkovna središča za medicino, družboslovje in humanistiko skupaj s SURS in IVZ). (Naloga področnih središč, Koordinacije in MIZŠ, da poskrbijo za dialog in prenos rešitev med okolji, 2013 in dalje)</p>
<p><i>Opravljanje podpornih storitev za raziskovalce</i></p> <p>Področna podatkovna središča skupaj z drugimi izvajalci (glej poglavji Vzpostavitev storitev za podporo pripravi NRRP in Ustanove prevzamejo odgovornost) nudijo osnovne storitve za podporo raziskovalcem pri izvajanju obveznosti akcijskega načrta in DOLOČIL POLITIKE ODPRTEGA DOSTOPA.</p>	<p>a. Organiziranje in vodenje seminarjev ter usposabljanj za izdelavo in izvajanje NRRP.</p> <p>b. Storitve podpore pri NRRP ustvarjalcem raziskovalnih podatkov.</p> <p>c. Priprava metapodatkov v obliki, primerni za evidentiranje v nacionalnem registru raziskovalnih podatkov (priprava oblike bibliografskega zapisa objavljenih raziskovalnih podatkov, pripis trajnega označevalnika posamezni verziji, skladnost z uveljavljenimi splošnimi standardi za dostopne metapodatke, kot sta DC⁵⁵³ in DataCite).</p> <p>d. Razvoj storitev iskanja in dostopa do raziskovalnih podatkov, prilagojenih zahtevam podatkov glede tehničnih karakteristik in pravnih omejitev.</p> <p>e. Usposabljanje uporabnikov in promocija druge rabe z organizacijo seminarjev, poletnih šol, vključevanjem podatkov v izvajanje rednega poučevanja, pripravo učnih</p>

⁵⁴⁹ Knowledge Exchange. Dostopno prek: <http://www.knowledge-exchange.info/>, 5. 11. 2013.

⁵⁵⁰ Research Data Alliance. Dostopno prek: <https://rd-alliance.org/>, 5. 11. 2013.

⁵⁵¹ DataCite. Dostopno prek: <http://www.datacite.org/>, 5. 11. 2013.

⁵⁵² OpenAIRE. Dostopno prek: <http://www.openaire.eu/>, 5. 11. 2013.

⁵⁵³ DublinCore. Dostopno prek: <http://dublincore.org/>, 5. 11. 2013.

	<p>gradiv na podlagi podatkov, uporabniških konferenc ipd. Izvajalci so področna podatkovna središča skupaj s pomembnimi ustvarjalci in uporabniki.</p> <p>(Začetek 2014, prehodno obdobje in vzpostavitev najkasneje do leta 2019)</p>
--	---

PRISPEVEK IZDAJATELJEV ZNANSTVENIH PUBLIKACIJ ZA ODGOVORNO RAVNANJE Z RAZISKOVALNIMI PODATKI

Problem	Cilj	Odgovornost	Zgledi
Izdajatelji znanstvenih publikacij nimajo določenih pravil glede obveznosti dostopa, recenziranja in citiranja raziskovalnih podatkov.	Izboljšati integriteto znanstvenih publikacij z možnostjo dostopa do originalnih raziskovalnih podatkov.	MIZŠ, predstavniki znanstvene skupnosti pri določanju politik revij v okviru uredništev, izdajatelji.	Priporočila STM ⁵⁵⁴ ; The Royal Society ⁵⁵⁵ .

Utemeljitev

Slovenski izdajatelji znanstvenih revij in monografij se lahko pridružijo najuglednejšim svetovnim založnikom z obvezo odprtega dostopa do raziskovalnih podatkov, ki so povezani s publikacijo, že pri recenziranju prispevkov in pri objavi, saj se tako poveča integriteta objavljenih izsledkov.

Prav tako je tudi citiranje podatkov v primeru druge rabe del akademske kulture, ki ga revije gojijo zato, da se moralno nagradi avtorje, ki so drugim ponudili ustvarjene raziskovalne podatke. Z vzpostavitvijo dobro podprtih trajnih označevalnikov publikacij, raziskovalnih podatkov in avtorjev se odpira prostor za sledenje uporabi in boljše izkoriščanje znanja.⁵⁵⁶

Zadolžitve	Koraki
<p><i>Uveljaviti odgovornost izdajatelja</i></p> <p>Javno financirani izdajatelji znanstvenih publikacij morajo aktivno prispevati pri uveljavljanju sistema odprtega dostopa do raziskovalnih</p>	<p>a. Zahtevati od kandidatov za sofinanciranje izdajanja domačih znanstvenih periodičnih publikacij, da uveljavijo notranje uredniške politike ravnanja z raziskovalnimi podatki, povezanimi z objavo. Pogoji za objavo članka je odprti dostop do raziskovalnih podatkov, ki so bili podlaga izpeljavam v članku, v pooblaščenem podatkovnem središču ali na drugem ustreznem mestu.</p>

⁵⁵⁴ International Association of Scientific, Technical & Medical Publishers (STM), Research Data Group. Dostopno prek: <http://www.stm-assoc.org/research-data-group/>, 5. 11. 2013.

⁵⁵⁵ Science as an open enterprise: The Royal Society Science Policy Centre report 02/12. Dostopno prek: <http://royalsociety.org/policy/projects/science-public-enterprise/report/>, 5. 11. 2013.

⁵⁵⁶ Primerjaj z izsledki druge delavnice projekta Odprti podatki. Dostopno prek: <http://www.adp.fdv.uni-lj.si/odpp10D2/>, 5. 11. 2013.

<p>podatkov z jasno določenimi pravili glede obveznosti avtorjev.</p>	<p>(2014 –2019, odvisno od vzpostavljenih storitev prevzema raziskovalnih podatkov; vključeno v pogodbe o subvencioniranju financiranja publikacij v letu 2015)</p> <p>b. Spremeniti recenzentske postopke tako, da vključujejo obravnavo podatkov (najkasneje leto 2017).</p>
<p><i>Dolžnost navajanja raziskovalnih podatkov</i></p> <p>Opredeliti navodila avtorjem glede obveznosti in načina citiranja raziskovalnih podatkov in upoštevanje tehnološke infrastrukture za zagotavljanje sledenja uporabe podatkov s strani založnika (kot je skladnost delovanja s storitvami sledenja citiranosti podatkov).</p>	<p>a. Podatki, uporabljeni v objavi, morajo biti opremljeni z oznako, iz katere je razvidno, kje so shranjeni in dostopni (dopolnjena navodila, v katerih je opredeljena celotna bibliografska referenca na raziskovalne podatke z vključenim stalnim identifikatorjem lokacije odprtega dostopa).</p> <p>b. Razvoj in implementacija tehnične infrastrukture za povezovanje objav in podatkov, sledenja objavam in uveljavljanja mednarodnih iniciativ (upoštevanje pobud, orodij in standardov DataCite, OpenAIREplus ipd.).</p> <p>c. Uveljaviti podporo in sodelovanje s pobudami za globalno istovetnost raziskovalcev (npr. ORCID) za zagotavljanje povezljivosti in ustrezno sledenje citatov.</p>

LITERATURA

- A Vision for Global Research Data Infrastructures*. Dostopno prek: <http://www.grdi2020.eu/>, 23. 1. 2014.
- Alliance for Permanent Access*. Dostopno prek: <http://www.alliancepermanentaccess.org/>, 26. 11. 2013.
- Alliance of German Science Organisations. *Principles for the Handling of Research Data*. Dostopno prek: http://www.allianzinitiative.de/en/core_activities/research_data/principles/, 23. 1. 2014.
- ANDS. 2011. What is research data? Dostopno prek: <http://ands.org.au/guides/what-is-research-data.html>, 7. 8. 2012.
- ANDS Technical Working Group. 2007. *Towards the Australian Data Commons A proposal for an Australian National Data Service*. Canberra. Dostopno prek: <http://ands.org.au/towardstheaustraliandatacommons.pdf%20>, 14. 1. 2014.
- Anderson, Chris. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *WIRED MAGAZINE*. Dostopno prek: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory. 7. 8. 2012.
- Arzberger Peter, Peter Schröder, Anne Beaulieu, Geoffrey C. Bowker, Kathleen Casey, Leif Laaksonen, David Moorman, Paul F. Uhler, Paul Wouters. 2004. Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal* 3: 135–152.
- Association of Research Libraries (ARL). 2007. *Agenda for Developing E-Science in Research Libraries*. Dostopno prek: http://www.arl.org/bm~doc/ARL_EScience_final.pdf, 7. 8. 2012.
- Ball, Alex. 2013. Making Data Count. Dostopno prek: <http://www.dcc.ac.uk/blog/making-data-count>, 7. 8. 2013.
- Beagrie, Neil in John Houghton. 2013. *The Value and Impact of the British Atmospheric Data Centre*. Dostopno prek: http://repository.jisc.ac.uk/5382/1/BADCReport_Final.pdf, 23. 1. 2014.
- Berman, Francine. 2010. We Need a Research Data Census. *Communications of the ACM*. Dostopno prek: http://www.cs.rpi.edu/~bermaf/Data_Census_CACM_10.pdf, 7. 8. 2012.
- Berman, Francine. 2008. Got Data? A Guide to Data Preservation in the Information Age: a Guide to Data Preservation in the Information Age. *Communications of the ACM* 51 (12): 50–56. doi:10.1145/1409360.1409376. Dostopno prek: <http://cacm.acm.org/magazines/2008/12/3360-got-data-a-guide-to-data-preservation-in-the-information-age/abstract>, 23. 1. 2014.
- Biemer, Paul P. in Lars Lyberg. 2003. *Introduction to survey quality*. Hoboken, New Jersey: J. Wiley.

- Board on Research Data and Information. Dostopno prek: <http://sites.nationalacademies.org/PGA/brdi/index.htm>, 26. 11. 2013.
- Borgman, Christine L. 2010. Research Data: Who will share what, with whom, when, and why? V: Beijing: China-North America Library Conference. Dostopno prek: <http://works.bepress.com/borgman/238/>, 7. 8. 2012.
- Borgman, Christine L. 2011. The Conundrum of Sharing Research Data. *Journal of the American Society for Information Science and Technology*: 1–40. Dostopno prek: <http://ssrn.com/abstract=1869155> ali <http://dx.doi.org/10.2139/ssrn.1869155>, 7. 8. 2012.
- BRTF-SDPA. 2010. *Sustainable Economics for a Digital Planet: Ensuring Long Term Access to Digital Information*. Dostopno prek: <http://brtf.sdsc.edu/>, 7. 8. 2012.
- Budapest Open Access Initiative. Dostopno prek: <http://www.budapestopenaccessinitiative.org/>, 23. 1. 2014.
- Büttner, Stephan, Hans-Christoph Hobohm in Lars Müller. 2011. *Handbuch Forschungsdatenmanagement*. Dostopno prek: http://www.forschungsdatenmanagement.de/?page_id=2, 23. 1. 2014.
- Colledge, Michael in Jörgen Dalén. 2009. *ESS Handbook for Quality Reports*. Luxembourg. Dostopno prek: http://epp.eurostat.ec.europa.eu/portal/page/portal/ver-1/quality/documents/EHOR_FINAL.pdf, 7. 8. 2012.
- Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age. 2009. Dostopno prek: <http://www8.nationalacademies.org/cp/projectview.aspx?key=48721>, 7. 8. 2012.
- CORDIS. e-Infrastructure. Dostopno prek: http://cordis.europa.eu/fp7/ict/e-infrastructure/publications_en.html, 7. 8. 2012.
- Corti, Louise. 2012. Evaluating Research Data. Dostopno prek: http://www.data-archive.ac.uk/media/369163/managing_research_data14sept2012b.pdf, 23. 1. 2014.
- Corti, Louise, Veerle Van den Eynden, Libby Bishop in Bethany Morgan-Brett. 2011. *Managing and Sharing Data*. Essex: UK Data Archive. Dostopno prek: <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>, 7. 8. 2012.
- Costas, Rodrigo, Ingeborg Meijer, Zohreh Zahedi, and Paul Wouters. 2013. The Value of Research Data, Metrics for Datasets from a Cultural and Technical Point of View. *Knowledge Exchange*. Dostopno prek: <http://www.knowledge-exchange.info/datametrics>, 23. 1. 2014.
- Cragin, Melissa H. and Jacob R. Carlson in Michael Witt Carole L. Palmer. 2010. Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society* 368, št. 1926: 4023–4038. Dostopno prek: <http://rsta.royalsocietypublishing.org/content/368/1926/4023>, 7. 8. 2012.
- Diane Harley, Sophia Krzys in C. Judson King. 2010. *Interim Report: Assessing the Future Landscape of Scholarly Communication: An In-depth Study of Faculty Needs and Ways of Meeting them*. Los Angeles, California. Dostopno prek: <http://escholarship.org/uc/item/15x7385g#page-3>, 7. 8. 2012.

- Digging into data*. Dostopno prek: <http://www.diggingintodata.org/>, 7. 8. 2012.
- Downs, Robert R. in Robert S. Chen. 2012. Improving the Trustworthiness of an Interdisciplinary Scientific Data Archive. In: Washington, DC, ZDA. Dostopno prek: <http://www.iasist2012.org/indexfolder/program/index.php?show=session:L>, 7. 8. 2012.
- CESSDA - Consortium of European Social Science Data Archives. Dostopno prek: <http://www.cessda.org/>, 7. 8. 2012.
- CODATA-ICSTI Task Group on Data Citation Standards and Practices. 2013. "Out of Cite, Out of Mind: Citation of Data: The Current State of Practice, Policy, and Technology." *Data Science Journal* 12: 76. doi:<http://dx.doi.org/10.2481/dsj.OSOM13-043>. Dostopno prek: https://www.jstage.jst.go.jp/article/dsj/12/0/12_OSOM13-043/article, 28. 1. 2014.
- Council of the European Union. 2013. 3223rd COMPETITIVENESS (Internal Market, Industry, Research and Space) Council Meeting – Bruselj, 18. in 19. februar. 56(6399/13): 20. Dostopno prek: http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/intm/135548.pdf, 22. 1. 2014.
- Dallmeier-Tiessen, Sunje, Robert Darby, Kathrin Gitmans, Patricia Herterich, Simon Lambert, Salvatore Mele, Josefine Nordling, Hans Pfeiffenberger, Sergio Ruiz in Eefke Smit. 2012. *D6.1 Summary of the Studies, Thematic Publications and Recommendations*. Dostopno prek: http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/11/ODE-WP6-DEL-0001-1_0.pdf, 23. 1. 2014.
- Danish Agency for Science in Innovation and Technology. 2011. Danish Research Infrastructure Roadmap. 57. Dostopno prek: <http://fivu.dk/en/publications/2011/danish-roadmap-for-research-infrastructure-2011/>, 23. 1. 2014.
- Data Seal of Approval*. Dostopno prek: <http://datasealofapproval.org/en/>, 23. 1. 2014.
- DDI - Data Documentation Initiative. Dostopno prek: <http://www.ddialliance.org/>, 7. 8. 2012.
- Deutsche Forschungsgemeinschaft Foundation. 2009. *Recommendations for Secure Storage and Availability of Digital Primary Research Data*. Dostopno prek: http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf, 23. 1. 2014.
- Developing Data Attribution and Citation Practices and Standards. V Berkeleyju, Kalifornija. Dostopno prek: http://sites.nationalacademies.org/PGA/brdi/PGA_064019, 23. 1. 2014.
- DFG Committee on Scientific Library Services and Information Systems. 2012. Taking Digital Transformation to the Next Level, The Contribution of the DFG to an Innovative Information Infrastructure for Research. Dostopno prek: http://www.dfg.de/download/pdf/foerderung/programme/lis/strategy_paper_digital_transformation.pdf, 23. 1. 2014.
- Dietrich, Dianne, Trisha Adamus, Alison Miner in Gail Steinhart. 2012. De-Mystifying the Data Management Requirements of Research Funders. *Issues in Science and Technology Librarianship*. Dostopno prek: <http://www.istl.org/12-summer/refereed1.html>, 23. 1. 2014.

- E-IRG. 2009. *E-IRG Report on data management. Data Management Task Force*. Dostopno prek: http://www.e-irg.eu/images/stories/publ/task_force_reports/dmtfjointreport.pdf, 7. 8. 2012.
- e-IRG White Paper 2013. 2013. Dostopno prek: <https://europe.rd-alliance.org/Repository/document/Publications and Reports/e-irg white paper 2013 - final version.pdf>, 23. 1. 2014.
- Editorial. 2008. Community cleverness required. *Nature* 455, št 7209: 1. Dostopno prek: <http://www.nature.com/nature/journal/v455/n7209/full/455001a.html>, 7. 8. 2012.
- Editorial. 2009. Data's shameful neglect. *Nature* 461, št. 7261: 145. Dostopno prek: <http://www.nature.com/nature/journal/v461/n7261/full/461145a.html>, 7. 8. 2012.
- Educopia Institute. 2011. *Aligning National Approaches to Digital Preservation*. Dostopno prek: <http://www.educopia.org/events/ANADP>, 23. 1. 2014.
- ESF. 2008. *Shared Responsibilities in Sharing Research Data: Policies and Partnerships*. Dostopno prek: http://www.dfg.de/download/pdf/foerderung/programme/lis/sharing_research_data_esf_dfg_0709.pdf, 7. 8. 2012.
- ESRC. 2010. *ESRC Research data policy*. Dostopno prek: <http://www.esrc.ac.uk/about-esrc/information/data-policy.aspx>, 7. 8. 2012.
- European Commission. 2007. *Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee. On Scientific Information in the Digital Age: Access, Dissemination and Preservation*. Dostopno prek: http://ec.europa.eu/research/science-society/document_library/pdf_06/communication-022007_en.pdf, 22. 1. 2014.
- European Commission. 2010. *Riding the wave: How Europe can gain from the rising tide of scientific data*. Dostopno prek: <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>, 7. 8. 2012.
- European Commission. 2012a. *Advancing Technologies and Federating Communities, A Study on Authentication and Authorisation Platforms For Scientific Resources in Europe*. Dostopno prek: <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/aaa-study-final-report.pdf>, 23. 1. 2014.
- European Commission. 2012b. *Recommendation on Access to and Preservation of Scientific Information*. Dostopno prek: http://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf, 14. 8. 2013.
- European Commission. 2012c. *Towards Better Access to Scientific Information: Boosting the Benefits of Public Investments in Research - COM(2012) 401 Final*. Dostopno prek: http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf, 14. 8. 2013.
- European Commission. *Data Protection in the European Union*. Dostopno prek: http://ec.europa.eu/justice/policies/privacy/docs/guide/guide-ukingdom_en.pdf.

- European Commission Research & Innovation. Policy Initiatives. Dostopno prek: <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301>, 7. 8. 2012.
- European Federation of Academies of Science and Humanities. 2012. *Joint Declaration: Open Science for the 21st Century*. Dostopno prek: <http://www.allea.org/Pages/ALL/31/397.bGFuZz1FTkc.html>, 25. 11. 2013.
- European Science Foundation in All European Academies. 2011. *The European Code of Conduct for Research Integrity*. Dostopno prek: http://www.esf.org/fileadmin/Public_documents/Publications/Code_Conduct_ResearchIntegrity.pdf, 23. 1. 2014.
- European Science Foundation in EUROHORC. 2009. *EUROHORCs and ESF Vision on a Globally Competitive ERA and Their Road Map for Actions*. Dostopno prek: <http://www.kowi.de/Portaldata/2/Resources/fp/EUROHORCs-ESF-RoadMap.pdf>, 22. 1. 2014.
- European Social Survey. 2009. *Round 5 Specification for Participating Countries*. London. Dostopno prek: http://europeansocialsurvey.org/index.php?option=com_content&view=article&id=248:r5-project-specs-release&catid=22:news&Itemid=48, 7. 8. 2012.
- European Union. 2011. *National Open Access and Preservation Policies in Europe*. Dostopno prek: http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-report-2011_en.pdf, 22. 1. 2014.
- Fitzgerald, Anne M., Kylie M. Pappalardo in Anthony C. Austin. 2008. *Practical Data Management: A Legal and Policy Guide*. Dostopno prek: <http://eprints.qut.edu.au/14923/>, 23. 1. 2014.
- Feijen, Martin. 2011. *What Researchers want*. Dostopno prek: http://www.surf.nl/nl/publicaties/documents/what_researchers_want.pdf, 7. 8. 2012.
- Graaf, Maurits van der in Leo Waaijers. 2011. *A Surfboard for Riding the Wave*. Dostopno prek: <http://www.knowledge-exchange.info/Default.aspx?ID=469>, 7. 8. 2012.
- GRDI2020 Consortium. 2012. *Towards a 10 Year Vision for Global Data Infrastructure*. Dostopno prek: <http://www.grdi2020.eu/Repository/FileScaricati/e2b03611-e58f-4242-946a-5b21f17d2947.pdf>, 23. 1. 2014.
- Green, Ann in Myron Gutmann. 2006. *Building Partnerships Among Social Science Researchers, Institution-based Repositories and Domain Specific Data Archives*. Dostopno prek: <http://deepblue.lib.umich.edu/handle/2027.42/41214>, 7. 8. 2012.
- Griffin, Tom, Brian Matthews, Alistair Mills, Sri Nagella, Arif Shaon, Michael Wilson in Erica Yang. 2012. *Moving from a scientific data collection system to an open data repository*. V: *The 7th International Conference on Open Repositories*. Dostopno prek: https://www.conftool.net/or2012/index.php?page=browseSessions&form_session=2https://www.conftool.net/or2012/index.php?page=browseSessions&form_session=2, 7. 8. 2012.

- Heather A Piwowar. 2011. Who Shares? Who Doesn't? Factors Associated with Openly Archiving Raw Research Data. *PLoS ONE* 6(7), št. e18657. Dostopno prek: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0018657>, 7. 8. 2012.
- Henty, Margaret, Belinda Weaver, Stephanie Bradbury in Simon Porter. 2008. *Investigating Data Management Practices in Australian Universities*. Dostopno prek: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0018657>, 7. 8. 2012.
- Hey, Tony, Stewart Tansley in Kristin Tolle, ed. 2009. *The Fourth Paradigm. Data-Intensive Scientific Discovery*. Redmond, Washington: Microsoft Research. Dostopno prek: http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf, 7. 8. 2012.
- ICPSR. Deposit Data. Dostopno prek: <http://www.icpsr.umich.edu/icpsrweb/deposit/> 7. 8. 2012.
- Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council. 2009. *Harnessing the Power of Digital Data for Science and Society*. Dostopno prek: http://www.nitrd.gov/About/Harnessing_Power_Web.pdf, 7. 8. 2012.
- International Association of Scientific Technical & Medical Publishers. 2007. *Brussels Declaration*. Dostopno prek: <http://www.stm-assoc.org/brussels-declaration/>, 20. 12. 2013.
- International Association of Scientific, Technical & Medical Publishers*. Dostopno prek: <http://www.stm-assoc.org/about-the-association/>, 20. 12. 2013.
- Interuniversity Consortium for Political and Social Research. 2013. Sustaining Domain Repositories for Digital Data: A Call for Change from an Interdisciplinary Working Group of Domain Repositories. Dostopno prek: <http://www.icpsr.umich.edu/icpsrweb/ICPSR/support/announcements/2013/09/sustaining-domain-repositories-for>, 9. 1. 2014.
- Jones, Sarah. 2009. *A report on the range of policies required for and related to digital curation*. Glasgow, Velika Britanija. Dostopno prek: <http://hdl.handle.net/1842/3371>, 7. 8. 2012.
- Kaul, Inge, Isabelle Grunberg in Marc A. Stern, ur. 1999. *Defining Global Public Goods. Global Public Goods: International Cooperation in the 21st Century*. New York: Oxford University Press.
- Kevin Ashley. 2011. *Data Quality and Curation*. Dostopno prek: http://www.grdi2020.eu/Pages/SelectedDocument.aspx?id_documento=201287b9-9f1c-4626-8b62-44cfe418707a, 7. 8. 2012.
- King, Gary. 1995. Replication, Replication. *PS: Political Science & Politics* 28: 443–99. Dostopno prek: <http://j.mp/jCyff1>, 23. 1. 2014.
- Koski, Kimmo, Claudio Gheller, Stefan Heinzl, Alison Kennedy, Achim Streit in Peter Wittenburg. 2009. *Strategy for a European Data Infrastructure*. Dostopno prek: <http://www.csc.fi/english/pages/parade>, 23. 1. 2014.
- Kuula Arja in Sami Borg. 2008. *Open Access to and Reuse of Research Data – The State of the Art in Finland*. Dostopno prek: http://www.fsd.uta.fi/fi/julkaisut/julkaisusarja/FSDjs07_OECD_en.pdf, 7. 8. 2012.

- League of European Research Universities. 2012. Open Research Data. Dostopno prek: http://www.leru.org/files/publications/Open_Access_to_Research_Data-FINALdocx.pdf, 23. 1. 2014.
- LeClere, Felicia B. 2006. Data Sharing for Demographic Research at ICPSR. *ICPSR Bulletin* 6, št. 2: 3–8. Dostopno prek: <http://www.icpsr.umich.edu/files/ICPSR/org/publications/bulletin/2006-Q1.pdf>, 7. 8. 2012.
- LERU Research Data Working Group. 2013. LERU Roadmap for Research Data. *ADVICE PAPER* 14: 36. Dostopno prek: http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf, 22. 1. 2014.
- Ligue des Bibliothèques Européennes de Recherche. 2012. *Ten Recommendations for Libraries to Get Started with Research Data Management, Final Report of the LIBER Working Group on E-Science / Research Data Management*. Dostopno prek: http://www.libereurope.eu/sites/default/files/The_research_data_group_2012_v7_final.pdf, 22. 1. 2014.
- Lossau, Norbert. 2012. An Overview of Research Infrastructures in Europe - and Recommendations to LIBER. *LIBER Quarterly* 21, št. 3/4: 313–329. Dostopno prek: <http://liber.library.uu.nl/index.php/lq/article/view/8028>, 7. 8. 2012.
- Lyon, Liz. 2007. *Dealing with Data: Roles, Rights, Responsibilities and Relationships*. Dostopno prek: http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing_with_data_report-final.pdf, 7. 8. 2012.
- Lyon, Liz. 2012. Incremental Change or Revolution? Libraries and the Informatics Transform. V *LIDA Conference*, Zadar, 49. Dostopno prek: <http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/liz-lyon-lida-keynote-croatia.pdf>, 26. 11. 2013.
- Marx, Vivien. 2013. Biology: The Big Challenges of Big Data. *Nature* 498: 255–60. Dostopno prek: http://www.nature.com/nature/journal/v498/n7453/full/498255a.html?WT.ec_id=NATURE-20130613, 23. 1. 2014.
- Mathae, Kathie Bailey in Paul F. Uhler. 2012. *The Case for International Sharing of Scientific Data: A Focus on Developing Countries: Proceedings of a Symposium*. The National Academies Press. Dostopno prek: http://www.nap.edu/openbook.php?record_id=17019, 23. 1. 2014.
- McGovern, Nancy Y., ur. 2012. *Allining national approach to digital preservation*. Atlanta: Educopia Institut. Dostopno prek: http://www.educofia.org/sites/default/files/ANADP_Educopia_2012.pdf, 7. 8. 2012.
- Memorandum for the Heads of Executive Departments and Agencies*. 2013. <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-16.pdf>, 23. 1. 2014.
- Milia, Nicola, Alessandra Congiu, Paolo Anagnostou, Francesco Montinaro, Marco Capocasa, Emanuele Sanna in Giovanni Destro Bisol. 2012. Mine, Yours, Ours? Sharing Data on Human

- Genetic Variation. *PloS one* 7 (6): e37552. Dostopno prek: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0037552>, 14. 8. 2013.
- Ministrstvo za visoko šolstvo, znanost in tehnologijo. 2011. Načrt razvoja raziskovalnih infrastruktur 2011–2020. Dostopno prek: <http://www.arhiv.mvzt.gov.si/fileadmin/mvzt.gov.si/pageuploads/pdf/znanost/nacrt-RI.pdf>, 7. 8. 2012.
- Mossink, Wilma, Magchiel Bijsterbosch in Joeri Nortier. 2013. *European Landscape Study of Research Data Management*. Dostopno prek: http://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM_landscape_report_vs1_4_14.08.13.pdf, 11. 12. 2013.
- Muñoz Palma, Patricia. 2011. *Developing a research data access policy in Chile*. Washington, DC. Dostopno prek: http://sites.nationalacademies.org/PGA/biso/PGA_061353, 7. 8. 2012.
- National Academy of Sciences Board on Research Data and Information. BRDI DATA AND INFORMATION CHALLENGE, Using Data for the Public Good. Dostopno prek: http://sites.nationalacademies.org/PGA/brdi/PGA_085255, 26. 11. 2013.
- National Research Council. 1999. *A Question of Balance: Private Rights and the Public Interest in Scientific and Technical Databases*. Washington, DC: National Academy Press.
- National Science Board. 2005. *Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century*. Dostopno prek: <http://www.nsf.gov/pubs/2005/nsb0540/>, 7. 8. 2012.
- National Science Board. 2011. Task Force on Data Policies Statement of Principles. Dostopno prek: <http://www.nsf.gov/nsb/committees/dp/principles.pdf>, 23. 1. 2014.
- National Science Foundation. 2007. *Cyberinfrastructure Vision for 21st Century Discovery*. Dostopno prek: <http://www.nsf.gov/pubs/2007/nsf0728/>, 23. 1. 2014.
- National Science Foundation. Dostopno prek: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>, 26. 11. 2013.
- Natural Environment Research Council. Dostopno prek: <http://www.nerc.ac.uk/>.
- OECD. 2007. *Principles and Guidelines for Access to Research Data from Public Funding*. OECD. Dostopno prek: <http://www.oecd.org/sti/scienceandtechnologypolicy/oecdprinciplesandguidelinesforaccessstoresearchdatafrompublicfunding.htm>, 7. 8. 2012.
- OECD. 2004. *Declaration on Access to Research Data from Public Funding*. Dostopno prek: <http://webnet.oecd.org/OECDACTS/Instruments/ShowInstrumentView.aspx?InstrumentID=157&InstrumentPID=153&Lang=en&Book>, 23. 1. 2014.
- Open Access Committee. Recommendations for Implementation of Open Access in Denmark. Dostopno prek: http://www.kulturstyrelsen.dk/fileadmin/publikationer/publikationer_engelske/open_access_2010/html/chapter03.htm, 23. 1. 2014.

- Pampel, Heinz. 2012. Quality Assurance of Research Data: Perspectives of Scientist, Infrastructure Providers and Publishers. V *Quality Assurance of Research Data: a common endeavour*, Dunaj, Avstrija, 36. Dostopno prek: <http://de.slideshare.net/heinzpampel/quality-assurance-of-research-data>, 23. 1. 2014.
- Pampel, Heinz, Hans Pfeiffenberger, Angela Schäfer, Eefke Smit, Stefan Pröll in Christoph Bruch. 2012. *Report on Peer Review of Research Data in Scholarly Communication*. Dostopno prek: http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/05/APARSEN-DEL-D33_1A-01-1_0.pdf, 23. 1. 2014.
- Pampel, Heinz, Paul Vierkant, Frank Scholze, Roland Bertelmann, Maxi Kindling, Jens Klump, Hans-Jürgen Goebelbecker, Jens Gundlach, Peter Schirmbacher in Uwe Dierolf. 2013. Making Research Data Repositories Visible: The Re3data.org Registry. *PeerJ PrePrints*: 1:e21v1. Dostopno prek: <http://dx.doi.org/10.7287/peerj.preprints.21v1>, 23. 1. 2014.
- PARSE.Insight. Dostopno prek: <http://www.parse-insight.eu/project.php>, 7. 8. 2013.
- PARSE.Insight. 2009. *INSIGHT into Issues of Permanent Access to the Records of Science in Europe*. Dostopno prek: http://www.surf.nl/nl/publicaties/documents/what_researchers_want.pdf, 7. 8. 2012.
- Parsons, M. A., O. Godoy, E. LeDrew, T. F. de Bruin, B. Danis, S. Tomlinson in D. Carlson. 2011. A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science. *Journal of Information Science* 37(6): 555–69. Dostopno prek: <http://jis.sagepub.com/cgi/content/long/37/6/555>, 23. 1. 2014.
- Parsons, Mark. GRDI2020 Expert Report on Data Policy And Data Access. Dostopno prek: <http://www.grdi2020.eu/Repository/FileScaricati/e31a1aab-b01e-4e7e-9b10-0fd93d4b710f.pdf>, 23. 1. 2014.
- Parsons, Mark. 2010. Towards a 10-year vision for Global Data Infrastructures Workshop. Dostopno prek: <http://www.grdi2020.eu/Repository/FileScaricati/f43bfb7-3dac-4cf6-a506-03b07d01b19f.pdf>, 7. 8. 2012.
- Piwowar, Heather A., Todd J. Vision in Michael C. Whitlock. 2011. Data Archiving Is a Good Investment. *Nature* 473. Dostopno prek: <http://www.nature.com/nature/journal/v473/n7347/full/473285a.html>, 23. 1. 2014.
- Pravilnik o kazalcih in merilih znanstvene in strokovne uspešnosti. ARRS. Dostopno prek: <http://www.rrs.gov.si/sl/akti/prav-znan-strok-uspesn-maj09.asp>, 7. 8. 2012.
- Pryor, Graham, Sarah Jones in Angus Whyte. 2013. Delivering Research Data Management Services. Dostopno prek: <http://www.dcc.ac.uk/node/10114>.
- Reilly, Susan. 2013. European Commission Public Consultation on Open Research Data: Some Impressions from the Day. Dostopno prek: <http://www.libereurope.eu/blog/european-commission-public-consultation-on-open-research-data-some-impressions-from-the-day>, 23. 1. 2014.
- Research Information Network. 2008. *Stewardship of digital research data - principles and guidelines*. London. Dostopno prek: <http://www.rin.ac.uk/data-principles>, 7. 8. 2012.

- Research Councils UK. 2012. RCUK Common Principles on Data Policy. Dostopno prek: <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>, 23. 1. 2014.
- Research Councils UK. <http://www.rcuk.ac.uk/>, 7. 8. 2013.
- Research Data e - Infrastructures: Framework for Action in H2020. Dostopno prek: https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/data_infra_h2020_fiches_on-line_consult.pdf, 23. 1. 2014.
- Research Information Network in British Library. 2010. *Patterns of information use and exchange: case studies of researchers in the life sciences*. Dostopno prek: <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/patterns-information-use-and-exchange-case-studie>, 3. 1. 2014.
- Research Information Network. 2008. *Stewardship of Digital Research Data: a Framework of Principles and Guidelines*. London. Dostopno prek: <http://rinarchive.jisc-collections.ac.uk/our-work/data-management-and-curation/stewardship-digital-research-data-principles-and-guidelines>, 23. 1. 2014.
- Robert S. Chen. 2013. CODATA Strategic Plan 2013-2018. Dostopno prek: http://www.codata.org/CODATA_Strategic_Plan_2013-2018.pdf, 23. 1. 2014.
- Ruusalepp, Raivo. 2008. *A comparative study of international approaches to enabling the sharing of research data*. Dostopno prek: <http://www.dcc.ac.uk/sites/default/files/documents/publications/reports/Data-Sharing-Report.pdf>, 7. 8. 2012.
- Sandberg, Marie. 2012. *Reinforced Nordic Collaboration on Data Resources – Nordisk Samarbejde*. Dostopno prek: <http://www.norden.org/da/publikationer/publikationer/2012-514>, 21. 1. 2014.
- Sesink, Laurents, René van Horik in Henk Harmsen. 2008. *Quality guidelines for digital research data in the Netherlands*. Dostopno prek: http://www.data-archive.ac.uk/media/57319/dsa_booklet.pdf, 7. 8. 2012.
- Soehner, Catherine, Catherine Steeves in Jennifer Ward. 2010. *E-Science and Data Support Services: A Study of ARL Member Institutions*. Dostopno prek: http://www.arl.org/bm~doc/escience_report2010.pdf, 7. 8. 2012.
- Support for Establishment of National/Regional Social Sciences Data Archives*. Dostopno prek: <http://www.serscida.eu/en/>, 23. 1. 2014.
- Sveinsdottir, Bridgette, Thordis Wessels, Rod Smallwood, Peter Linde, Vasso Kala in Victoria Tsoukala. 2013. *Stakeholder Values and Ecosystems*. Dostopno prek: http://recodeproject.eu/wp-content/uploads/2013/10/RECODE_D1-Stakeholder-values-and-ecosystems_Sept2013.pdf, 23. 1. 2014.
- Swan Alma in Sheridan Brown. 2008. *To Share or not to Share: Publication and Quality Assurance of Research Data Outputs Report commissioned by the Research Information Network (RIN)*. Dostopno prek: <http://www.rin.ac.uk/system/files/attachments/To-share-data-outputs-report.pdf>, 7. 8. 2012.

- Swan Alma in Sheridan Brown. 2008. *The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs*. Dostopno prek: <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareerfinalreport.pdf>, 7. 8. 2012.
- Štebe, Janez. 2013. Aktualnost mednarodne primerjalne raziskave volilnih sistemov v časih družbene krize (CSES). V Niko Toš (ur.): *Vrednote v prehodu V.: Slovenija v evropskih primerjavah: evropska družboslovna raziskava 2002–2010*. Wien: Echoraum; Ljubljana: Fakulteta za družbene vede, IDV-CJMMK.
- Štebe, Janez. 2010. Arhiv družboslovnih podatkov in storitve prostega dostopa do znanstvenih podatkov v republiki Sloveniji. *Prost dostop do dosežkov slovenskih znanstvenikov*: 64–78. Dostopno prek: <http://www.dlib.si/v2/Details.aspx?URN=URN:NBN:SI:doc-JGYFFYW7>, 7. 8. 2012.
- Štebe, Janez, Sonja Bezjak, Sanja Lužar in Irena Vipavc Brvar. 2012. Opis stanja na področju raziskovalnih podatkov v Sloveniji, Poročilo projekta Odprti Podatki – priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji. URN:NBN:SI:DOC-F9KIZ8B0. Dostopno prek: http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odpp10_opis_stanja/, 17. 1. 2014.
- Štebe, Janez in Irena Vipavc Brvar. 2008. Vodič za pripravo družboslovnih podatkov za arhiviranje. Dostopno prek: http://adp.fdv.uni-lj.si/za_uporabnike/delavnice/#vodic, 7. 8. 2012.
- Štebe, Janez in Irena Vipavc Brvar. 2011. Analiza stanja in perspektiva digitalne hrambe v Arhivu družboslovnih podatkov (ADP). *Knjižnica* 55, št. 1: 57–85. Dostopno prek: http://revija-knjiznica.zbds-zveza.si/lzvodi/K1101/Stebe_Vipavc-Brvar.pdf, 7. 8. 2012.
- Štebe, Janez, Jože Hudales in Boris Kragelj. 2011. Archiving and Re-using Qualitative and Qualitative Longitudinal Data in Slovenia. *IASSIST Quarterly* 2010, št. Spring/Summer/Fall/Winter: 50–59. Dostopno prek: http://iassistdata.org/downloads/iqvol34_35_stebe.pdf, 7. 8. 2012.
- Tenopir, Carol, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, Lei Wu, Eleanor Read, Maribeth Manoff in Mike Frame. 2011. Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6, št. 6. Dostopno prek: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0021101>, 7. 8. 2012.
- The Consultative Committee for Space Data Systems. 2012. *Reference Model for an Open Archival Information System (OAIS)*. Washington, DC, ZDA. Dostopno prek: <http://public.ccsds.org/publications/archive/650x0m2.pdf>, 7. 8. 2012.
- The Economist. 2010. Data, data everywhere. *The Economist*. Dostopno prek: <http://www.economist.com/node/15557443>, 7. 8. 2012.
- The Royal Society. 2012. Science as an Open Enterprise, Final Report. Dostopno prek: <http://royalsociety.org/policy/projects/science-public-enterprise/report/>, 26. 11. 2013.
- The Royal Society Science Policy Centre. 2012. *Science as an open enterprise*. Dostopno prek: http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf, 7. 8. 2012.

- Thessen Anne in David Patterson. 2011. Data issues in the life sciences. *ZooKeys* 150: 15–51. Dostopno prek: <http://www.pensoft.net/journals/zookeys/issue/150/>, 7. 8. 2012.
- Toronto International Data Release Workshop Authors. 2009. Prepublication Data Sharing, Toronto Statement. *Nature* 461: 168–170. doi:10.1038/461168a. Dostopno prek: <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/WTX060361.htm>, 23. 1. 2014.
- Treloar, Andrew, William Michener in G. Sayeed Choudhury. 2012. Contrasting National Research Data Strategies: Australia and the USA. V *Managing Research Data*, London: Facet Publishing, 173–203.
- Uhlir, Paul F. 2012. For Attribution– Data Attribution and Citation Practices and Standards. *The National Academies Press*. Dostopno prek: http://www.nap.edu/openbook.php?record_id=13564, 23. 1. 2014.
- UK Data Archive. Dostopno prek: <http://data-archive.ac.uk/>, 7. 8. 2012.
- UK Research Data Service. 2010. *Proposal and Business Plan for the Initial Pathfinder Development Phase*. Dostopno prek: <http://www.ukrds.ac.uk/resources/download/id/47>, 7. 8. 2012.
- Univerza v Oxfordu. Data Quality and Data Quality Assurance Policy. Dostopno prek: http://www.admin.ox.ac.uk/pras/aboutus/data_quality/, 7. 8. 2012.
- Vipavc Brvar, Irena in Jožica Klep. 2004. Primerjava metapodatkovnega standarda Arhiva družboslovnih podatkov (DDI) in standarda Statističnega urada RS (METIS). *Organizacija: revija za management, informatiko in kadre* 37, št. 5: 296–304.
- Wellcome Trust. 2003. *Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility*. Dostopno prek: <http://www.genome.gov/pages/research/wellcomereport0303.pdf>, 23. 1. 2014.
- Wicherts, Jelte M., Marjan Bakker, in Dylan Molenaar. 2011. Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results. *PLoS ONE* 6 (11) (2. november): e26828. Dostopno prek: <http://dx.plos.org/10.1371/journal.pone.0026828>, 7. 8. 2012.
- Williford, Christa in Charles Henry. 2012. *One Culture. Computationally Intensive Research in the Humanities and Social Sciences*. Dostopno prek: <http://www.clir.org/pubs/reports/pub151/pub151.pdf>, 7. 8. 2012.
- Wynholds, Laura A., Jillian C. Wallis, Christine L. Borgman, Ashley Sands in Sharon Traweek. 2012. Data, data use, and scientific inquiry: Two case studies of data practices. *ACM Press*. Dostopno prek: <http://works.bepress.com/cgi/viewcontent.cgi?article=1282&context=borgman>, 7. 8. 2012.
- Yoon, Ayoung in Helen Tibbo. 2011. Examination of Data Deposit Practices in Repositories with the OAIS Model. *IASSIST Quarterly WINTER*: 6–13. Dostopno prek: http://www.iassistdata.org/downloads/iqv0135_tibbo.pdf, 7. 8. 2012.

PRILOGE

PRILOGA 1: Prošnja za sodelovanje v projektu Odprti podatki



Arhiv družboslovnih podatkov
Univerza v Ljubljani, Fakulteta za
družbene vede
Kardeljeva ploščad 5
1000 Ljubljana
tel: 01/5805-292
arhiv.podatkov@fdv.uni-lj.si

Ljubljana, 25. 1. 2011

doc. dr. Janez Štebe, predstojnik ADP
Kardeljeva ploščad 5
1000 Ljubljana

Tel. 01/5805-292
<http://www.adp.fdv.uni-lj.si/blog>
http://twitter.com/#!/Janez_Stebe
E-pošta: janez.stebe@fdv.uni-lj.si

Zadeva: Prošnja za sodelovanje v projektu Odprti podatki

Spoštovani!

Arhiv družboslovnih podatkov za Ministrstvo za visoko šolstvo, znanost in tehnologijo v letih 2010-13 izvaja projekt CRP s šifro V5-1018 »Odprti podatki - Priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji«. Prosimo vas za sodelovanje pri aktivnostih, ki jih bomo organizirali v okviru projekta za namenom zbiranja informacij o stanju na področju ustvarjanja in dostopa do raziskovalnih podatkov doma in v svetu. Z vašim sodelovanjem bomo v nadaljevanju iskali in preizkušali najbolj ustrezne rešitve, ki bodo ob zaključku projekta kot rezultat podane v predlogu akcijskega načrta. S svojim sodelovanjem nam boste pomagali priti do relevantnih rezultatov - takih, ki bodo ustrezali vašim željam in potrebam ter odsevali posebnosti vaše znanstvene discipline in institucionalnega okolja.

Odprt dostop do raziskovalnih podatkov je aktualen zaradi možnosti, ki jih ponujajo sodobne digitalne tehnologije in omrežja. Kot je navedeno v »Načelih in usmeritvah OECD za dostop do javno financiranih raziskovalnih podatkov«¹ (v nadaljevanju »Načela OECD«), bolj odprt dostop do podatkov povečuje raznolikost analiz in presoj, spodbuja nova raziskovanja, omogoča preizkušanje novih ali alternativnih hipotez in metod raziskovanja, proučevanje metod zbiranja podatkov in meritev, spodbuja izobraževanje novih raziskovalcev, omogoča raziskovanje prvotno nepredvidenih tem ter oblikovanje novih podatkovnih serij s kombiniranjem podatkov iz več virov. Odstranjene ovire pri dostopu do podatkov bodo torej zagotovile njihovo širšo uporabo, s tem izboljšale znanstveno odličnost rezultatov in njihovo uporabnost ter povečale donos javnih naložb v raziskovalno dejavnost.

K izpolnjevanju Načel OECD, ki predstavljajo izhodišče projekta, se je zavezala tudi Slovenija z nedavno včlanitvijo v OECD. Odprt dostop do podatkov po Načelih OECD pomeni preglednost nad viri raziskovalnih podatkov, ki so nastali v skladu z uveljavljenimi znanstvenimi merili kakovosti in uporabnosti za nadaljnje analize. Raziskovalne podatke je potrebno urediti in opremiti z vsemi potrebnimi dodatnimi informacijami (metapodatki), ki opredeljujejo avtorske, patentne, metodološke in vsebinske podrobnosti o nastanku podatkovnega vira in možnostih nadaljnje uporabe. Odstranjene ovire pri dostopu do podatkov bodo torej zagotovile njihovo širšo uporabo, s tem pa se bo izboljšala znanstvena odličnost in uporabnost rezultatov.

¹ OECD Principles and Guidelines for Access to Research Data from Public Funding,
<http://www.oecd.org/dataoecd/9/61/38500813.pdf>

Vabimo Vas, da kot poznavalci svojega področja prispevate informacije, predloge in komentarje na različna izhodišča akcijskega načrta ter sodelujete v aktivnostih, ki jih bomo s tem namenom organizirali: v intervjujih, fokusnih skupinah in drugih dogodkih ter diskusijah, ki jih bomo vodili s pomočjo različnih spletnih orodij. Le tako bomo lahko zagotovili, da bo končni rezultat projekta – akcijski načrt z ukrepi za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji - realistično zasnovan, na podlagi obstoječih razmer na področju ustvarjanja podatkov, storitev njihovega shranjevanja in dostopa, ter dejanskih potreb po njihovem nadaljnjem izkoriščanju pri nas in na tujem.

Podrobnosti o poteku izvajanja projekta, pa tudi druge informacije iz področij, povezanih z vsebino projekta, so predstavljene na našem spletnem naslovu. Vabljeni ste, da čim bolj tekoče spremljate, komentirate in dodajate svoje vsebine tudi s pomočjo spletnih tehnologij (bloga, Twitterja, Facebooka, Connotee).

Za dodatna pojasnila smo vselej na voljo!

Za Vaše sodelovanje se Vam že vnaprej zahvaljujemo in Vas lepo pozdravljamo!

Vodja projekta Odprti podatki
doc. dr. Janez Štebe

PRILOGE:

- Priporočilno pismo MVZT

PRILOGA 2: Priporočilno pismo za sodelovanje v projektu Odprti podatki

REPUBLIKA SLOVENIJA
**MINISTRSTVO ZA VISOKO ŠOLSTVO,
 ZNANOST IN TEHNOLOGIJO**

Kotnikova 38, 1000 Ljubljana

T: 01 478 46 00
 F: 01 478 47 19
 E: gp.mvzt@gov.si
 www.mvzt.gov.si

Številka: 631-101/2010/2
 Datum: 17.12.2010

Priporočilno pismo za sodelovanje pri projektu »Odprti podatki«

Arhiv družboslovnih podatkov (ADP) na Fakulteti za družbene vede, Univerza v Ljubljani, v okviru razpisa za CRP v letih 2010-13 po naročilu Ministrstva za visoko šolstvo, znanost in tehnologijo (MVZT) izvaja projekt »Odprti podatki - Priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji« (Šifra projekta: V5-1018). Republika Slovenija se je k ureditvi tega področja zavezala ob včlanitvi v organizacijo OECD. V ta namen vas prosimo, da sodelujete pri omenjenem projektu in na ta način prispevate k oblikovanju rezultatov, ki bodo upoštevali širše vidike učinkovite, kakovostne in sodobno zasnovane znanstvene dejavnosti.

Nosilci projekta, sodelavci ADP, vas bodo povabili k posameznim aktivnostim v različnih fazah poteka projekta, hkrati pa je predvideno tudi vključevanje predstavnikov MVZT. Vabimo vas torej, da skupaj z nami prispevate k uspehu projekta in ustvarjalno sodelujete v različnih fazah izvajanja projekta v svojih vlogah kot direktorji in vodje raziskovalnih ustanov, kot ustvarjalci podatkov, kot izvajalci podatkovnih storitev ali kot njihovi uporabniki.

S spoštovanjem,

Direktorat za znanost in tehnologijo
 Generalna direktorica

dr. Jana Kolar



Pripravil/a:
 Petra Tramte

PRILOGA 3: Izjavo o strinjanju s pogoji sodelovanja



Odprti podatki

IZJAVA O STRINJANJU S POGOJI SODELOVANJA

Vaše sodelovanje pri intervjuju je prostovoljno. Intervju bo predvidoma trajal eno uro. Pridobljene podatke bomo uporabili za raziskovalne cilje projekta 'ODPRTI PODATKI' in jih shranili v podatkovnih arhivih, kjer bodo dostopni uporabnikom za druge raziskovalne ali izobraževalne namene. Osebni kontaktni podatki bodo iz tega izključeni in jih bomo uporabili samo za korespondenco v okviru tega projekta. Razen v kolikor izrecno ne želite, vaše ime ne bo navedeno v poročilih in drugih projektnih gradivih v povezavi z izrečenim v okviru intervjuja.

Potrdite, ali izrecno želite, da je vaše ime navedeno v poročilih in drugih gradivih projekta, ki se navezujejo na izrečeno v okviru intervjuja. DA NE

Prosimo, da s podpisom potrdite, da se strinjate z zgoraj navedenimi pogoji sodelovanja.

Ime in priimek intervjuvanca

Podpis

Datum

Doc. dr. Janez Štebe
Vodja projekta Odprti podatki

Podpis

Datum

ŠT. INT. _____

PRILOGA 4: Vprašalnik



Odprti podatki

Datum: _____

ŠT. INT. _____

Kraj: _____

OSNOVNI PODATKI O INTERVJUVCANU

Ime in priimek: _____

Organizacija, ki jo zastopa: _____

Znanstveno področje: _____

Položaj: _____

Izobrazba: _____

Leta dela: _____

Leto rojstva: _____

Opreделите svojo sedanjo ali pričakovano vlogo v zvezi z raziskovalnimi podatki; poimenujte jo tako kot ste navajeni.

Vloge (akterji) (nosilci pravic, dolžnosti, izvajajo naloge...)	Vloge v kateri nastopa	Lastno poimenovanje vloge
Ustvarjalci , raziskovalna enota, raziskovalni inštitut, podatkovni znanstvenik	<input type="checkbox"/>	
Storitve (skrb za podatke, katalogiziranje, opremljanje z metapodatki za uporabo, spravilo, povezovanje, specializirani področni podatkovni arhivi, mednarodni arhivi, digitalni institucionalni ali nacionalni repozitoriji, knjižnice)	<input type="checkbox"/>	
Uporabniki (drugi raziskovalci kot uporabniki, študentje, poučevanje)	<input type="checkbox"/>	
Vodstva in oblikovalci politik (ministrstvo, drugi financerji, vodstva institucij)	<input type="checkbox"/>	

V katerih fazah življenjskega kroga podatkov nastopate ali se povezujete z drugimi; poimenujte jih tako kot ste navajeni.

Faze življenjskega kroga podatkov	Faze v katerih sodeluje	Lastno poimenovanje aktivnosti
Načrtovanje	<input type="checkbox"/>	
Financiranje	<input type="checkbox"/>	
Ustvarjanje (zagotavljanje kakovosti, uporabnosti, evidentiranje nastanka)	<input type="checkbox"/>	
Dodana vrednost (metapodatki, transformacije, čiščenje)	<input type="checkbox"/>	
Izbor, vrednotenje kakovosti, uporabnosti	<input type="checkbox"/>	
Spravilo	<input type="checkbox"/>	
Analiza , objava rezultatov v raziskovalnem poročilu	<input type="checkbox"/>	
Dostop, iskanje (lociranje), ponovna uporaba podatkov drugih	<input type="checkbox"/>	



Odprti podatki

Faza: **USTVARJANJE PODATKOV**

Opis stanja nastajanja podatkov, z njimi povezana orodja, postopki in navodila.

(Opis nastanka podatkov in metapodatkov; kakšni podatki, kaj se z njimi dogaja; katera orodja, metapodatki, standardi, postopki so uporabljeni.)

Problemi povezani z drugimi fazami življenjskega kroga (upoštevani ali niso upoštevani)

(načrtovanje, kaj se bo zbiralo za kakšen namen; upoštevanje vrednotenja, izbora in zagotavljanje načrtovane kakovosti; upoštevanje zahtev hrambe; razreševanje problemov za zagotavljanje dostopa; vprašaj zakaj)

Odgovornost (lastna in drugih) v zvezi z reševanjem problemov (primeri pomanjkljivosti in razlogi za nereševanje problemov)

(kdo nastopa s kakšnimi dolžnostmi, pravicam, medsebojna pričakovanja in odvisnost nosilcev vlog; pomen politik, vodil, standardov pri opredeljevanju odgovornosti; ovire, sankcije in spodbude, nagrade);

Identifikacija nosilcev dobrih praks, zgledov reševanja problemov.

(Kdo so favoriti, zgledi, kje so že vzpostavljene za prenos znanja uporabne povezave doma in v tujini; uporaba orodij, dobrih praks in standardov, katerih?)

Zap. Št.	Kdo (poimenuj nosilce, njihove vloge, če to ni sam)	Kaj (kateri problem rešuje, katera sredstva uporablja, na katero področje se nanaša) *	Zunanja referenca

*Sredstva, orodja, vodila

- Orodja (za ustvarjanje, za prikaz in analizo podatkov/ metapodatkov)
- Standardi, dobre prakse, navodila (glede kakovosti, podatkov, metapodatkov, hrambe)
- Politike, strategije, zakonodaja, načrti, projekti

(Opomba za analizo: namen je identificirati nosilce problemov ter nosilce zgledov njihovega reševanja; V nadaljnjih fazah iščemo rešitve na način preučevanja možnosti in pogojev za uveljavljanje zgledov rešitev).

Primeri vprašanj: USTVARJANJE PODATKOV

Opis stanja

Ali imate podatke: Kaj se s podatki zgodi, dogaja;

Kakšne podatke producirate in uporabljate pri svojem delu. (cf. Finsko poročilo)

Lahko pokažete konkretno datoteko in kako izgleda? Opišite.

Z vidika hrambe: Format (nekateri formati so pokriti, drugi ne). Uvrstite v razred tipa podatkov. Velikost (količina podatkov), kompleksnost.

ali so kakšna pravila glede tega? (poznavanje standardov, postopkov za zagotavljanje kakovosti;

Ugotavljanje podatkovnih potreb skladno z poslanstvom, ciljem projekta (Harnesing...)

Pregled uporabnosti obstoječih podatkov preden zbiraš nove (Harnesing...)

evidentiranje podatkov ob nastanku, dokumentiranje (cf. Parse: spremljajoči podatki o nastanku)

pregledovanje podatkov, kontrola kakovosti ob nastanku

pomen podatkov (izbor, vrednotenje za cilje projekta);

Povezave z objekti (metapodatki, standardi, publikacije)

Problemi

Kateri od problemov, postopkov so pomembni, s kakšnim ciljem (Faze: shranjevanje, ponovna uporaba, objava, načrtovanje)

Pravni vidiki (avtorske pravice, zasebnost, etični vidiki)

Odgovornost, pripravljenost prevzeti vlogo, odvisnost od pogojev (nagrada, sankcij), odgovornost drugih.

Ali je ustvarjanje podatkov ustrezno vrednoteno (sodelovanje kot avtor pri objavi, posebej namenjena sredstva za skrb za podatke?)

Ustrezna kariera za podatkovnega znanstvenika, nagrade;

Ali so upoštevana pravila/zahteve sistema preverjanja kakovosti podatkov ob objavi: sankcija je zavrnjena objava; ali je zagotovljena kakovost ob nastanku

Preverjanje kakovosti ob shranjevanju: zavrjen vstop v sistem hrambe;

Kdo skrbi za zagotavljanje pravnih vidikov (navodila, kdo jih je dolžan priskrbeti; etični kodeksi, etične komisije; skrb za izvajanje navodil).

Povezave z nosilci vlog (uporabnik, skrbnik); fazami

Faza: **DODANA VREDNOST (METAPODATKI, TRANSFORMACIJE, ČIŠČENJE)**

Opis stanja na področju dodane vrednosti podatkov, z njimi povezana orodja, postopki in navodila.

(Opis spreminjanja podatkov in metapodatkov v fazi dodajanja vrednosti; kaj se s podatki in metapodatki dogaja; katera orodja, metapodatki, standardi, postopki so uporabljeni.)

Problemi povezani z drugimi fazami življenjskega kroga (upoštevani ali niso upoštevani)

(Za kakšne namene služijo postopki dodane vrednosti; upoštevanje načrtovane kakovosti; upoštevanje zahtev hrambe, ali je zagotovljena sledljivost postopkov in transformacij; razreševanje problemov za zagotavljanje dostopa, npr. lažja dostopnost, odstranjene ovire zaupnosti ipd.)

Odgovornost (lastna in drugih) v zvezi z reševanjem problemov (primeri pomanjkljivosti in razlogi za nereševanje problemov)

(kdo nastopa s kakšnimi dolžnostmi, pravicam, medsebojna pričakovanja in odvisnost nosilcev vlog; pomen politik, vodil, standardov pri opredeljevanju odgovornosti; ovire, sankcije in spodbude, nagrade);

Identifikacija nosilcev dobrih praks, zgledov reševanja problemov.

(Kdo so favoriti, zgledi, kje so že vzpostavljene za prenos znanja uporabne povezave doma in v tujini; uporaba orodij, dobrih praks in standardov, katerih?)

Zap. št.	Kdo (poimenuj nosilce, njihove vloge, če to ni sam)	Kaj (kateri problem rešuje, katera sredstva uporablja, na katero področje se nanaša) *	Zunanja referenca

*Sredstva, orodja, vodila

- Orodja (za ustvarjanje, za prikaz in analizo podatkov/ metapodatkov)
- Standardi, dobre prakse, navodila (glede kakovosti, podatkov, metapodatkov, hrambe)
- Politike, strategije, zakonodaja, načrti, projekti

Primeri vprašanj: DODANA VREDNOST (METAPODATKI, TRANSFORMACIJE, ČIŠČENJE)

Opis stanja

Ali surove podatke obdelate, da jim dodate vrednost?

Vrednotenje, uporabnost: ali so surovi podatki sploh uporabni (zase, za druge) preden jih opremimo, dopolnimo; kakšna znanja in orodja so potrebna za uporabo; katere dodatne informacije so potrebne (metapodatki, dokumentacija) o nastanku in dodajanju vrednosti;

Ali je možna standardizacija na tem področju, navodila, politike, metapodatki (jih pozna, so pomembne, ovire pri implementaciji)?

Kaj se dogaja s podatki od nastanka do točke spravila z vidika dodane vrednosti?

Problemi

Čemu dodana vrednost koristi? (za namen spravila, iskanja, dokumentiranja, analize)

Odgovornost Pripravljenost prevzeti vlogo, odgovornost, odvisnost od pogojev (nagrada, sankcij), odvisnost od objektov (obstoje načrtov, politik, strategij), odgovornost drugih (vloge, kdo bi lahko prevzel).

Kdo opravlja, sam ali pooblaščen podatkovni center, repozitorij, ali specialist v okviru organizacije; centralni nacionalni digitalni arhiv; Mednarodni centri...

Ali je dodana vrednost nagrajena, izvajalec dobi povrnitev stroškov;

Ali so opredeljene obveznosti glede dodane vrednosti, metapodatkov (na nivoju organizacije, discipline, nacionalno); kdo bi moral opredeliti obveznosti;

Ali je obveznosti potrebno opredeliti za doseganje ciljev – enotna raziskovalna politika?

Identifikacija nosilcev dobrih praks

Kje bi lahko dobil navodila, napotke, usposabljanje za delo na tem področju (področni centri za strokovno svetovanje, kdo je vodilni na tem področju, nastavki dejavnosti);

Faza: **SPRAVILO, DOLGOROČNA RAZPOLOŽLJIVOST PODATKOV**

Opis stanja na področju spravila podatkov, s skrbjo za podatke povezana orodja, postopki in navodila.

(Opis korakov v fazi spravila; kaj se s podatki dogaja, kakšne oblike, formati podatkov nastopajo; kje se podatki hranijo; kakšna so znanja glede hrambe; katera orodja, metapodatki, standardi, postopki so uporabljeni – ali imajo načrte hrambe, poznajo OAIS.)

Problemi povezani z drugimi fazami življenjskega kroga (upoštevani ali niso upoštevani)

(Ali je shranjevanje za lastne potrebe ali tudi za potrebe drugih po koncu projekta; Prednosti hrambe na obstoječi lokaciji ali centralizirano po disciplinah; kaj je ciljna vsebina, ki se hrani; ali je opravljeno strokovno vrednotenje pomena podatkov – dolgotrajna vrednost za projekt in za druge, kaj se zgodi s podatki male vrednosti, ali so podatki 'enkratni' ali pa jih z lahkoto reproduciramo; upoštevanje načrtovane kakovosti – ali je zagotovljeno ohranjanje vrednosti podatkov – ali je razpoložljivost podatkov ogrožena zaradi odsotnosti hrambe; ali je v okviru sistema hrambe zagotovljena sledljivost izvora podatkov in postopkov v predhodnih fazah; razreševanje problemov povezave podatkov in metapodatkov, dokumentacije v paketu povezanih gradiv skozi proces hrambe)

Odgovornost (lastna in drugih) v zvezi z reševanjem problemov (primeri pomanjkljivosti in razlogi za nereševanje problemov)

(kdo nastopa s kakšnimi dolžnostmi, pravicam, medsebojna pričakovanja in odvisnost nosilcev vlog – kje bi bilo primerno mesto za odlaganje podatkov; pomen politik, vodil, standardov pri opredeljevanju odgovornosti – ali imamo merila in vrednotenje izpolnjevanja zahtev kakovostne hrambe; ovire, sankcije in spodbude, nagrade – kdo za hrambo plača, kdo skrbi za usposabljanje za hrambo; pravni problemi hrambe – prenos pravic, zaščita podatkov);

Identifikacija nosilcev dobrih praks, zgledov reševanja problemov.

(Kdo so favoriti, zgledi, kje so že vzpostavljene za prenos znanja uporabne povezave doma in v tujini; uporaba orodij, dobrih praks in standardov, katerih?)

Zap. Št.	Kdo (poimenuj nosilce, njihove vloge, če to ni sam)	Kaj (kateri problem rešuje, katera sredstva uporablja, na katero področje se nanaša)*	Zunanja referenca

Primeri vprašanj: SPRAVILO, DOLGOROČNA RAZPOLOŽLJIVOST PODATKOV

Opis stanja

Ali hranite podatke: Kaj se s podatki zgodi, dogaja;

Ali ima načrt in strategijo skrbi za podatke.

Kje hranijo proizvedeni svoje podatke (v mednarodnem/državnem/institucionalnem repozitoriju/na strežniku znotraj ustanove, kjer delajo/na svojem osebem računalniku...)?

Kakšna je skrb za podatke: kje, v kakšni obliki (formatu) se hranijo, verzije, načrt ravnanja s podatki

Kakšna je časovna potreba po hranjenju

Z vidika hrambe: Format (nekateri formati so pokriti, drugi ne). Uvrstite v razred tipa podatkov. Velikost (količina podatkov), kompleksnost.

Ali so kakšna pravila glede spravila in dolgotrajne hrambe tega? (poznavanje standardov, postopkov za zagotavljanje kakovosti;

Kakšna orodja uporablja ali načrtuje;

Problemi

Evidentiranje in dokumentiranje spremljajočih podatkov o nastanku pomembni z vidika hrambe?

Pregledovanje podatkov, kontrola kakovosti ob vstopu v sistem hrambe

Pomen podatkov (izbor, vrednotenje za dolgotrajno razpoložljivost);

Nepomembni podatkovni viri: kaj z njimi? Bi tovrstne podatkovne vire dali v depo majhne vrednosti ali zavrgli? V povezavi z vrednotenjem;

Prenos podatkov v druge sisteme, povezani sistemi hrambe

Kaj so prednosti hranjenja podatkov na obstoječi lokaciji (priročnost/varnost/backup/enostavnost dostopa oz. prilagodljivost uporabe/nizka cena/kapaciteta hranjenja/tisto, kar je na razpolago/deljenje podatkov/samostojnost pri upravljanju oz. nadzoru) in slabosti (omejena kapaciteta/omejena varnost/omejena možnost backup/zunanji težko dostopajo do podatkov/predraga cena storitev/ne ustreza potrebam/za backup je potreben čas/omejena publiciteta)? Načrti za oblikovanje /pridruževanje / uporabo teh storitev?

Uničenje podatkov, primeri ravnanja? Koliko so podatki enkratni oz. se jih da ponovno pridobiti? Koliko dosega prag 'ogroženosti' uničenja;

Kje so med podatkovnimi viri praznine oz. do katerih podatkov ne morete dostopati (pomanjkljivosti kot je npr. obstoj pomembnega podatkovnega vira, ki je ogrožen, nedostopen, slabo dokumentiran...)? Kaj so grožnje shranjevanju: tehnologija, katastrofe, diskontinuiteta?

Kateri od problemov, postopkov so pomembni, s kakšnim ciljem (Faze: vrednotenje, ponovna uporaba za lastne namene, za namene drugih, objava, načrtovanje)

Tehnološki vidiki (orodja, postopki, standardi) - pomen

Uporaba specialnih metapodatkov za shranjevanje, standardi, navodila, ki jih upošteva.

Odgovornost

Kdo je odgovoren? Kaj bi bilo treba narediti?

Kje bi jih lahko objavili oz shranili? (institucionalni repozitorij, svetovni repozitorij, revije, etc.) Kam odlagati – kaj bi bilo primerno mesto...?.



Odprti podatki

Ali je shranjevanje podatkov ustrezno vrednoteno (posebej namenjena finančna sredstva?)

Ustrezna kariera za podatkovnega skrbnik, usposabljanje; kdo bi lahko za to nalogo usposabljal;

Preverjanje kakovosti ob shranjevanju, ali je opredeljeno, ali je nekdo pooblaščen?

Kdo skrbi za zagotavljanje pravnih vidikov (pravice glede spravila, prenos pravic, obveznosti, nadzor dostopa).

Faza: **ISKANJE, DOSTOP (OBJAVA), PONOVNA UPORABA**

Opis stanja na področju objave in dostopa do podatkov, z iskanjem, dostopom in ponovno uporabo povezana orodja, postopki in navodila.

(Opišite komu, pod kakšnimi pogoji, kakšne oblike, formati in verzije podatkov so dostopni - kakšna je praksa glede omejevanja dostopa; kje se podatki hranijo; katera orodja, metapodatki, standardi, postopki so uporabljeni za iskanje in dostop; ali so podatki evidentirani, kako podrobno so opisani v katalogu.)

Problemi povezani z drugimi fazami življenjskega kroga (upoštevani ali niso upoštevani)

(Ali se dostopnost podatkov upošteva in zagotavlja v predhodnih fazah priprave podatkov; ali imajo podatki vrednost za druge (preverjanje objav, raziskovalni potencial za nova odkritja in objave, kombiniranje več podatkovnih virov med sabo); ali se za vrednost za ponovno uporabo vrednoti, kakšen je postopek, merila; ali ima uporabnik možnost oceniti vrednost in uporabnost podatkov; ali ima interes za podatke drugih, za kakšen namen; zagotovljena sledljivost izvora podatkov in postopkov v predhodnih fazah za zagotavljanja uporabnosti podatkov za druge, dostop do metapodatkov, dokumentacije; podatki in publikacije;)

Odgovornost (lastna in drugih) v zvezi z reševanjem problemov (primeri pomanjkljivosti in razlogi za nereševanje problemov)

(kdo nastopa s kakšnimi dolžnostmi, pravicam, medsebojna pričakovanja in odvisnost nosilcev vlog – kje bi bilo primerno mesto za dostop do podatkov; kdo bi lahko opravljal storitve zagotavljanja dostopa (knjižnice), katere funkcije bi moral izpolnjevati; pomen politik, vodil, standardov pri opredeljevanju odgovornosti – ali imamo merila kakovostne podatkovne objave, pozna in podpira strategije in politike za zagotavljanje dostopa; ovire, sankcije – nezmožnost lociranja podatkov, finančne zahteve; in spodbude, nagrade – kaj kot avtor pričakuje, citiranje, zabeležena objava; pravni problemi dostopa, avtorske pravice, etična občutljivost podatkov); pogled na obveznost dostopa do podatkov s strani naročnika (javno financirani projekti, zahteve EU), revije ipd.; kultura izmenjave, pogled glede omejevanja in pogojev dostopa; ali imaš kot interesent za podatke opredeljene pravice glede dostopa; odnos do OECD vodil (kdo in kako zagotoviti izpolnjevanje – nacionalno, ali vidi sebe, druge; samoumevnost ali neizvršljivost); model načrtovanja upravljanja podatkov;)

Identifikacija nosilcev dobrih praks, zgledov reševanja problemov.

(Kdo so favoriti, zgledi, kje so že vzpostavljene za prenos znanja uporabne povezave doma in v tujini; uporaba orodij, dobrih praks in standardov, katerih?)

Zap. št.	Kdo (poimenuj nosilce, njihove vloge, če to ni sam)	Kaj (kateri problem rešuje, katera sredstva uporablja, na katero področje se nanaša) *	Zunanja referenca

Primeri vprašanj: ISKANJE, DOSTOP (OBJAVA), PONOVNA UPORABA

Opis stanja

Ali je omogočeno lociranje, dostop in ponovna uporaba podatkov?

Ali omogoča dostop do podatkov. Komu, pod kakšnimi pogoji, kje, v kakšni obliki;

Podatki – publikacije (da objaviš v reviji in je zraven še povezava na podatke)

Koliko ljudi uporablja podatke (izbranega raziskovalca)? Jaz/ostali člani projektne skupine/omejeno število zunanjih uporabnikov (do 20)/veliko zunanjih uporabnikov (več kot 20)/drugi – meri pripravljenost, da daš na razpolago drugim/da uporabiš podatke drugim/da iščeš

Orodja in tehnologije za analizo, dostop

Ali predstavljajo oviro z vidika dostopnosti podatkov?

Kje bi jih lahko objavili? (institucionalni repozitorij, svetovni repozitorij, revije, etc.)

Kam odlagati – kaj bi bilo primerno mesto spravila in dostopa...?

Ali ima načrt in strategijo objavljanja podatkov.

Kako hitro se producent/repozitorij/revija odzove na prošnjo za podatke? Ali se sploh odzovejo?

Problemi

Ka so ovire dostopa (npr. nezmožnost lociranja podatkov; avtorske pravice; zasebnost; patent)?

Kateri od problemov, postopkov so pomembni, s kakšnim ciljem (Preizkusiti pomen posamezne faze za dostop: vrednotenje (recenzija, selekcija, zagotovilo znanstvene odličnosti, uporabnosti, ponovna uporaba za lastne namene, za namene drugih, načrtovanje objave...)

Tehnološki vidiki (orodja, postopki, standardi) – pomen za dostop

Odgovornost

Kdo skrbi za zagotavljanje pravnih vidikov (pravice glede spravila, prenos pravic, obveznosti, nadzor dostopa).Povezave z nosilci vlog (uporabnik, skrbnik); fazami

Do katerih podatkov ne morete dostopati zaradi nerazpoložljivosti; kaj so ovire za dostop? Kdo je odgovoren? Kaj bi bilo treba narediti?

Ali je zagotavljanje dostopa nagrajeno/kaznovano (pravice uporabe do objave, konkurenca/ citiranje, priznanje avtorstva.

Ali je obvezno zagotoviti dostop (zahteva naročnika); Ali podpira obveznost zagotavljanja dostopa?

Kakšen pogled imate na obveznost izročanja podatkov – npr. v revijah;

Kultura? Splošna ocena koristi dostopa oz. kulture podatkovne izmenjave. Interdisciplinarne izmenjave (interes, navade)?

Ali so opredeljene pravice, jih uporabnik pozna, glede dostopa.

Pomembno z vidika pravic in dolžnosti; ali so stvari poljubne, določene...; Kakšna pa je dejanska praksa? Kakšen je prag dostopnosti (to je ločen vidik problema, tudi če so podatki shranjeni, so lahko dostopni ožjemu krogu);

Interes, da se citira...

Stališča glede odprtega dostopa (koristi, ali velja za njegovo delo, področje...= , zavedanje nevarnosti uničenja Percipirane ovire glede dostopa do podatkov...?

Koristi in spodbude

Politike (institucionalne, mednarodne, nacionalne) – raziskovalni projekti – podatki: ali se zahteva v okviru kakšnih slovenskih/tujih/mednarodnih projektov, da se poleg poročila shrani v javno dostopnem depozitoriju tudi podatke? Npr. FP7

Odnos imate možnost/želite objavljati/imate koristi?

Ali se vam šteje v dobro pri napredovanju, če objavljate podatke?

Potrebe po ustreznih kadrih, finančnih in drugih sredstvih (Odgovornost):

Ali mislite, da bi za to moral dobiti denar? Koliko?

Javen/plačljiv dostop/kombinacija? (glej DRIVER)

Ali knjižnice/revije vidijo v hranjenju/dostopu do podatkov nišo za svojo dejavnost? V kolikšni meri bi jo izvajale na račun dosedanje?

OECD deklaracijo pozna, se mu zdi izvedljivo, kakšna sredstva, čigava odgovornost (etično, univerze same, financerji...) kdo določiti? Odvisnost od drugih (ocena moči, odzivnosti drugih akterjev na področju): kdo mislite, da mora rešiti problem odprtih podatkov? Sami/država/... Ali mora priti pobuda od spodaj ali od zgoraj?

Pomanjkanje poznavanja ustrezne tehnologije in znanja za njeno uporabo? (problemi, ovire pri dostopu; ali je potrebno poznavanje orodij; ali je treba kaj kompenzirati (npr. z obljubo soobjave...)

Ali je ideja dostopa po OECD samoumevna: kakšne so težave, bariere, pomisleki glede dostopa? Tu gre tudi za vidik izmenjave podatkov z vidika kulture.

Podpora nacionalni ureditvi tega področja? Poznavanje, ocena težav itd., izkušnje (ali je sam odklonil, dostop bil onemogočen)...: kdo računa na free riderje, kje so individualni interesi pretehtali nad skupno koristjo itd.

Faza objavljanja podatkov (motivacije, ovir; primerjaj pull push) v širšem smislu kultura izmenjave; problem citiranja in povezovanja podatkov in publikacij.

Lastništvo podatkov:

Ali so avtorske pravice, intelektualna lastnina, patenti ipd. ovira za uveljavitev odprtih podatkov? Kako je s problemom zaupnosti podatkov? (meri pripravljenost na objavlanje z vidika ovir;,) PRAVICE

Katere politike, strategije poznate oz. delujejo kot spodbuda za objavlanje podatkov?

Vprašanje glede sprejemljivosti različnih politik: ali bi bilo sprejemljivo obvezno predajanje podatkov; Kaj pa model data management planning (DMP),

Zopet glede politik: centralizirano/ distribuirano: Kakšne so izhodiščne predstave, kako bi se lahko to reševalo? Vloga knjižnic, založnikov, medsebojna pričakovanja...?

Ali raziskovalci vidijo knjižnice kot partnerje pri hranjenju in dostopanju do podatkov?

Od odkritja do dostopa

Katera znanja mora imeti uporabnik, da podatke lahko uporablja? Kako morajo biti shranjeni? V kakšnem formatu?

To vprašanje je povezano tudi s profilom podatkov: nekateri podatki so razumljivi sami po sebi, pri drugih rabiš razlago. Običajno potrebuješ za njihovo odpiranje določen program (npr. SPSS). Opremljeni morajo biti z meta podatki.



Odprti podatki

Faza: IZBOR, VREDNOTENJE KAKOVOSTI, UPORABNOSTI

Pojasnilo: Vrednotenje je vključeno v večino predhodnih faz. Tukaj na enem mestu izpostavimo ta vprašanja z namenom, da preverimo, če smo izpustili kakšen pomemben vidik vrednotenja.

Opis stanja na področju vrednotenja podatkov, povezana orodja, postopki in navodila.

(Opišite pomen podatkov, njihovo enkratnost, uporabnost za lastno uporabo, za ponovno uporabo; katera orodja, metapodatki, standardi, postopki so uporabljeni za vrednotenje; ali je ocena vrednotenja kje označena, vidna.)

Problemi povezani z drugimi fazami življenjskega kroga (upoštevani ali niso upoštevani)

(Ali se vrednotenje podatkov upošteva in zagotavlja v predhodnih fazah priprave podatkov, ali se pri vrednotenju upošteva pomen za druge ; ali se ob objavi)

Odgovornost (lastna in drugih) v zvezi z reševanjem problemov (primeri pomanjkljivosti in razlogi za nereševanje problemov)

(kdo nastopa s kakšnimi dolžnostmi, pravicam, medsebojna pričakovanja in odvisnost nosilcev vlog – kje bi bilo primerno mesto za vrednotenje (ob nastanku, ob shranjevanju, ob dostopu, ob objavi); pomen politik, vodil, standardov – ali imamo merila vrednotenja, ovire, sankcije, finančne zahteve; in spodbude, nagrade – ali se kakovost podatkov izplača, je nagrajena?)

Identifikacija nosilcev dobrih praks, zgledov reševanja problemov.

(Kdo so favoriti, zgledi, kje so že vzpostavljene za prenos znanja uporabne povezave doma in v tujini; uporaba orodij, dobrih praks in standardov, katerih?)

Zap. Št.	Kdo (poimenuj nosilce, njihove vloge, če to ni sam)	Kaj (kateri problem rešuje, katera sredstva uporablja, na katero področje se nanaša) *	Zunanja referenca

Dodatna vprašanja: IZBOR, VREDNOTENJE KAKOVOSTI, UPORABNOSTI

Opis stanja

*Ali se podatki ločijo na bolj pomembne in manj pomembne;
Kako lahko jih reproduciramo, če se uničijo, izgubijo? (Ali imate opredeljene postopke za vrednotenje podatkov ob nastanku, ob sprejemu v hrambo, ob dajanju v dostop; Kakšni so postopki, kriteriji (recenziranje, ocenjevanje na podlagi česa?)*

Kaj se dogaja s podatki od nastanka do točke spravila z vidika dodane vrednosti, ali podatki pridobivajo na vrednosti?

Ali so podatki dragoceni? (ponovljivi ali ne itd.; ali obstaja povpraševanje po podatkih, interes in možnosti ponovne uporabe – dostop do tujih podatkov)

Ali imate izoblikovana merila za ocenjevanje podatkov/bi morala biti? Koliko so disciplinarno specifična, koliko splošna?

Problemi

Čemu vrednotenje koristi? (za namen spravila – selekcija najpomembnejših podatkov zaradi stroškov hrambe) ; iskanja, dokumentiranja, ponovne uporabe (rangiranje podatkov po pomenu za zagotovilo kakovosti pri ponovni uporabi; za ekvivalent znanstvene objave)

Ali vrednotenje obsega različne komponente, vključno s fazami in izpolnjevanje (opredeljenih) zahtev v posameznih fazah (nastanek, načrtovanje, uporaba standardov in meril; ustreznost dodane vrednosti, primernosti za hrambo, ustrezne opremljenosti z metapodatki...); ustreznost za ponovno uporabo (vrednost za uporabnike, za poučevanje, edinstvenost podatkov – praznina v znanju);

Odgovornost

Kdo opravlja vrednotenje, sam ali pooblaščen podatkovni center, repozitorij, ali specialist v okviru organizacije; centralni nacionalni digitalni arhiv; Mednarodni centri...; znanstvena skupnost z opredeljenim postopkom, revija ob publikaciji.

Izvajalec vrednotenja dobi povrnitev stroškov;

Interes za vrednotenje (ali je ustrezno nagrajena izvedba, ali ima vrednotenje posledice za ugled, citiranost, in zato koristno),

Ali so opredeljeni kriteriji, politike vrednotenja (na nivoju organizacije, discipline, nacionalno);

Ali bi vrednotenje moralo biti preverjeno s statistikami uporabe (citiranje).

Identifikacija nosilcev dobrih praks, zgledov reševanja problemov.

Ali so kakšne poznane prakse, zgledi; vključevanje v podatkovne revije; kdo bi ocenjeval podatke...

Kje bi lahko dobil navodila, napotke, usposabljanje za doseganje ciljev na tem področju (etični kodeksi in dobre prakse disciplin, organizacij, izdajateljev, podatkovnih centrov, repozitorijev);



Odprti podatki

Sodelovanje v fokusnih skupinah

Na katerih področjih iz Vodil OECD bi bil pripravljen sodelovati v razpravah (tehnologija, standardi; digitalna hramba; organizacija, odgovornost, politike; avtorske in druge pravne omejitve, zasebnost, patenti; kultura izmenjave, etika; interoperabilnost, učinkovitost; kakovost podatkov; ocenjevanje storitev in nadzor nad izvajanjem; obveznosti načrtov ravnanja s podatki in njihov nadzor ter vrednotenje izpolnjevanja v okviru projektov).

Spodnja tabela je namenjena identifikaciji nosilcev dobrih praks in zglede reševanja problemov. Če veste za kak takšen primer, ga prosimo navedite (nanaša se na vse faze življenjskega kroga podatkov)

Zap. Št.	Faza	Kdo (poimenuj nosilce, njihove vloge, če to ni sam)	Kaj (kateri problem rešuje, katera sredstva uporablja, na katero področje se nanaša)*	Zunanja referenca

ŠT. INT. _____

PRILOGA 5: Priporočila udeležencev delavnice *Problemi in rešitve na področju podatkovnih storitev v Sloveniji*



Projekt Odprti podatki

Arhiv družboslovnih podatkov

VZPOSTAVLJANJE NACIONALNEGA OKOLJA ODPRTEGA DOSTOPA DO PODATKOV IZ JAVNO FINANCIRANIH RAZISKAV

Priporočila udeležencev delavnice *Problemi in rešitve na področju podatkovnih storitev v Sloveniji*, Fakulteta za družbene vede, 5. 12. 2012

Izhodišča

Odprti raziskovalni podatki omogočajo odprto znanost, spodbujajo raznolikost analiz, hipotez in podatkov ter meddisciplinarno raziskovanje, omogočajo boljši izkoristek javnega financiranja raziskovalne dejavnosti in spodbujajo poslovno izrabo rezultatov raziskav. S povezovanjem odprtih podatkov raziskav in odprtih objav nastanejo obogatene objave.

Odprto dostopnost podatkov raziskav priporočata Organizacija za gospodarsko sodelovanje in razvoj (OECD) ter Evropska komisija. Po besedah Neelie Kroes, evropske komisarke za Digitalno agendo, je deljenje podatkov raziskav in možnost njihove ponovne uporabe bistvenega pomena za razvoj znanosti.

Odprti dostop do podatkov raziskav je del nacionalnega okolja odprtega dostopa do rezultatov javno financiranih raziskav, katerega pričakovana zasnova v državah članicah EU je opisana v dokumentu *Priporočila Komisije z dne 17. julija 2012 o dostopu do znanstvenih informacij in njihovem arhiviranju* (2012/417/EU).

Ugotovitve

Udeleženci delavnice glede stanja odprtih podatkov v Sloveniji podpirajo zgornja izhodišča in pozivajo k izpolnjevanju vseh obveznosti, ki iz njih sledijo. Pri tem ugotavljajo, da:

- Raziskovalci večinoma slabo poznajo prednosti in koristi odprtih podatkov, raven njihove seznanjenosti je odvisna od znanstvenega področja, na katerem delujejo. Zadržki raziskovalcev glede odprtega objavljanja podatkov raziskav se nanašajo v prvi vrsti na vprašanje varovanja intelektualne lastnine in varovanja osebnih podatkov, oviro pa

predstavljata tudi odsotnost kulture deljenja podatkov in nestimulativen sistem vrednotenja znanstvenih dosežkov.

- Slovenski financerji raziskav in raziskovalne organizacije še nimajo izdelanih politik odprtih podatkov, ali pa te niso dovolj jasno artikulirane glede določanja obvez in spodbud zagotavljanja odprtega dostopa, celovitejša nacionalna infrastruktura odprtih podatkov še ni zgrajena.
- Dejavnosti promocije, podpore in svetovanja glede zagotavljanja odprto dostopnih podatkov, razen redkih izjem (kot so npr. pojasnila in novice na nacionalnem informacijskem portalu openaccess.si), na državni ravni še niso razvite.

Zato priporočajo:

- Slovenski financerji raziskav in raziskovalne organizacije morajo v skladu s priporočili OECD ter Evropske komisije opredeliti politike odprtih raziskovalnih podatkov, ki bodo, ob upoštevanju pravnih in etičnih obveznosti, zagotovile za najširši krog dostopne in uporabne podatke javno financiranih raziskav.
- Politike morajo zagotoviti ustrezno motivacijo in nagrado za ustvarjalce podatkov na osnovi njihovega prispevka pri pripravi raziskovalnih podatkov, skladno z zahtevami pooblaščenih področnih podatkovnih centrov (obveznost oddaje načrtov ravnanja s podatki, plačilo dodatnega dela pri pripravi podatkov za shranjevanje in dostop, vrednotenje prispevka raziskovalnih podatkov v osebnih bibliografijah).
- Potrebno je nadaljevati z izgradnjo nacionalne infrastrukture odprtih podatkov, vključno z vzpostavitvijo nacionalnega portala/agregatorja in področnih podatkovnih centrov ter mehanizmov za zajemanje metapodatkov iz mednarodne podatkovne infrastrukture (uvedba sistema trajnih identifikatorjev, etična obveznost citiranja, sledenje drugi rabi podatkov), in zagotoviti članstvo države v evropski podatkovni raziskovalni infrastrukturi (zlasti CESSDA, CLARIN in DARIAH, ki so že vključene v nacionalni Načrt razvoja raziskovalnih infrastruktur 2011-2020). Zagotovljena mora obstojna infrastruktura za trajno digitalno skrbništvo in hranjenje odprto dostopnih podatkov raziskav.
- Oblikovati je potrebno kriterije za delovanje področnih in integriranih podatkovnih storitev, zagotoviti razvoj in svetovanje pri vzpostavljanju novih storitev ter preverjati skladnost delovanja podatkovnih storitev z merili/kriteriji. Zlasti je pomembno na nacionalnem nivoju usklajeno razvijati storitve vrednotenja raziskovalnih podatkov, in sicer kot dejavnost pooblaščenih področnih podatkovnih centrov, ter koordinirano razvijati nacionalno povezane storitve digitalne hrambe.
- Pomoč raziskovalcem pri pripravi načrtov ravnanja s podatki raziskave lahko v času prijave projekta nudijo ustrezno usposobljeni knjižničarji raziskovalnih organizacij. Skupaj z ustvarjalci podatkov lahko sodelujejo pri izvajanju priprave podatkov, nudijo pomoč in posredujejo pri storitvah shranjevanja, priprave metapodatkov in promocije ter svetovanja pri drugi rabi. V sodelovanju z Osrednjimi specializiranimi informacijskimi centri (OSIC) in na podlagi sprejema podatkov raziskav v pooblaščenih podatkovnih centrih knjižničarji poskrbijo za evidentiranje podatkovne objave v bibliografskem informacijskem sistemu.

Seznam udeležencev delavnice *Problemi in rešitve na področju podatkovnih storitev v Sloveniji*

Zap. št.	Ime in priimek, kontaktni podatki
1	Ambrožič Melita, Narodna in univerzitetna knjižnica, melita.ambrozic@nuk.uni-lj.si
2	Bezjak Sonja, Univerza v Ljubljani, Fakulteta za družbene vede, Arhiv družboslovnih podatkov, sonja.bezjak@fdv.uni-lj.si
3	Černač Barbara, NIB in BF, Oddelek za biologijo, knjižnica, barbara.cernac@nib.si
4	Frelih Martina, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za geografijo, knjižnica, martina.frelih@ff.uni-lj.si
5	Hajtnik Tatjana, Arhiv Republike Slovenije, tatjana.hajtnik@gov.si
6	Hudelja Mihaela, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za etnologijo in kulturno antropologijo, knjižnica, mihaela.hudelja@ff.uni-lj.si
7	Intihar Ferjan Jana, Moderna galerija, jana.ferjan@mg-lj.si
8	Jug Janez, Univerza v Ljubljani, Fakulteta za družbene vede, Osrednja družboslovna knjižnica Jožeta Goričarja, janez.jug@fdv.uni-lj.si
9	Kanič Ivan, Ekonomska fakulteta, Centralna ekonomska knjižnica, ivan.kanic@ef.uni-lj.si
10	Kavčič-Čolić Alenka, Narodna in univerzitetna knjižnica, alenka.kavcic@nuk.uni-lj.si
11	Kotar Mirjam, Univerza v Ljubljani, Fakulteta za družbene vede, Osrednja družboslovna knjižnica Jožeta Goričarja, mirjam.kotar@fdv.uni-lj.si
12	Kotar Mojca, Univerza v Ljubljani, Univerzitetna služba za knjižnično dejavnost, mojca.kotar@uni-lj.si
13	Kragelj Matjaž, Narodna in univerzitetna knjižnica, matjaz.kragelj@nuk.uni-lj.si
14	Krek Simon, Institut Jožef Stefan, Laboratorij za umetno inteligenco, simon.krek@guest.arnes.si
15	Krstulović Zoran, Narodna in univerzitetna knjižnica, zkrstulovic@gmail.com
16	Kunej Drago, ZRC SAZU, Glasbenonarodopisni inštitut, drago.kunej@zrc-sazu.si
17	Logar Nataša, Univerza v Ljubljani, Fakulteta za družbene vede, natasa.logar@fdv.uni-lj.si
18	Lužar Sanja, Univerza v Ljubljani, Fakulteta za družbene vede, Arhiv družboslovnih

	podatkov, sanja.luzar@fdv.uni-lj.si
19	Marinček Lili, Kmetijski inštitut, knjižnica, lili.marincek@kis.si
20	Mesarič Judita, Arhiv Republike Slovenije, judita.mesaric@gov.si
21	Peče Miha, ZRC SAZU, Inštitut za slovensko narodopisje, miha.pece@zrc-sazu.si
22	Rivo Ksenija, Univerza v Ljubljani, Pedagoška fakulteta, knjižnica, ksenija.rivo@pef.uni-lj.si
23	Slabe Vesna, Urbanistični inštitut Republike Slovenije, knjižnica, vesna.slabe@uirs.si
24	Smrekar Tomaž, Statistični urad RS, tomaz.smrekar@gov.si
25	Smukavec Andreja, Statistični urad RS, andreja.smukavec@gov.si
26	Šoštarič Davor, IZUM, Maribor, davor.sostaric@izum.si
27	Štebe Janez, Univerza v Ljubljani, Fakulteta za družbene vede, Arhiv družboslovnih podatkov, janez.stebe@fdv.uni-lj.si
28	Tramte Petra, Ministrstvo za izobraževanje, znanost, kulturo in šport, petra.tramte@gov.si
29	Vipavc Brvar Irena, Univerza v Ljubljani, Fakulteta za družbene vede, Arhiv družboslovnih podatkov, irena.vipavc@fdv.uni-lj.si
30	Zaletel Metka, IVZ, Zdravstveno podatkovni center, metka.zaletel@ivz-rs.si
31	Zidar Nataša, Statistični urad RS, natasa.zidar@gov.si

PRILOGA 6: Določila politike odprtega dostopa do raziskovalnih podatkov v Sloveniji. Priloga k akcijskemu načrtu

1. Uvod

a. Pomen in motivacija za pripravo politik

Raziskovalni podatki predstavljajo pomemben strošek posameznih raziskovalnih projektov, zato je potrebno povečati stopnjo izkoriščenosti podatkov in tako prihraniti pri stroških, ki nastanejo pri drugih raziskovalnih projektih zaradi podvajanja ustvarjanja podatkov. Dostop do raziskovalnih podatkov mora biti kolikor mogoče odprt, saj se s tem poveča zaupanje v znanstvene objave, ki jih je mogoče bolj podrobno preveriti, krepi se sodelovanje v znanosti, z različnimi analitičnimi pristopi se poveča pluralizem in dialog o možnih interpretacijah podatkov in omogočajo se nova odkritja. Izhodišče za pripravo politik na področju dostopa do odprtih podatkov v Sloveniji predstavljajo načela in usmeritve OECD, priporočil Evropske komisije in pobud UNESCO-a za dostop do javno financiranih raziskovalnih podatkov. Po njih je temeljni cilj izboljšati učinkovitost in izkoristek nacionalnih in globalnega znanstvenega sistema. Boljši dostop do podatkov zagotavlja večji izkoristek javnih sredstev, vloženih v raziskovanje. Pomeni pa tudi potencialni vir znanja, ki je potrebno pri soočanju z izzivi našega časa. Ministrstva ter javne ustanove, ki pripravljajo in izvajajo politike na področju raziskovanja in ki nastopajo kot financerji raziskovalnih projektov, v celoti podpirajo ta izhodišča in s pripravo pričujoče zavezujoče politike pozivajo vse deležnike k njihovem uresničevanju.

Glavni načeli odprtega dostopa, katerima se je s pridružitvijo OECD zavezala Republika Slovenija, sta:

- raziskovalni podatki iz raziskav, ki se izvajajo ob podpori iz javnih virov so javno dobro, namenjeno za izkoriščanje v javne namene;
- potrebno je zagotoviti kolikor je mogoče odprt in prost dostop do raziskovalnih podatkov ob ustrezni skrbi in v čim krajšem času po nastanku.

b. Vloge in odgovornost

Kaj so raziskovalni podatki: Pri opredelitvi raziskovalnih podatkov izhajamo iz zavedanja, da je razumevanje vezano na posamezne znanstvene discipline in da so podrobne opredelitve vsebine, pomena in oblike raziskovalnih podatkov povezane z razumevanjem na podlagi znanja v določeni skupnosti uporabnikov podatkov. Poudarek je na osnovnih podatkih, ki jih v različnih disciplinah najdemo v različnih oblikah in jih je mogoče uporabiti za nadaljnje analize. Tu podajamo splošno opredelitev. Sledeč Načelom in usmeritvam OECD, so "raziskovalni podatki" digitalno berljivi zapisi o dejstvih, ki predstavljajo osnovno podlago za znanstveno raziskovanje in ki v okviru znanstvene skupnosti veljajo kot ustrezno sredstvo za preverjanje veljavnosti raziskovalnih spoznanj.

Kdo sodeluje: Odgovornost v okviru dostopa do podatkov je porazdeljena med različne deležnike. Pomemben poudarek pričujoče politike je, da jasno določa obveznosti ustvarjalcev podatkov, da zagotovijo pravočasen in kolikor je mogoče odprt dostop do podatkov. Obenem pa omogoča

izpolnjevanje teh obveznosti z gradnjo in vzdrževanjem infrastrukture podpornih storitev, ki jih na različnih ravneh in med seboj usklajeno izvajajo naslednji nosilci:

- področna podatkovna središča,
- podatkovni znanstveniki ter knjižničarji na raziskovalnih ustanovah;
- nacionalni ter mednarodni centri storitev za digitalno shranjevanje in posredovanje.

Posebno vlogo in odgovornost imajo odločevalci na vseh nivojih odločanja (ministrstva in agencije ter vodstva inštitucij), da zagotovijo optimalne pogoje za izvajanje te politike.

Uresničevanje pričujoče politike sloni na naslednjih elementih, ki skupaj zagotavljajo vzdržen in za ustvarjalce podatkov spodbuden okvir zagotavljanja priprave in dostopa do raziskovalnih podatkov:

- Obvezna priprava Načrtov ravnanja z raziskovalnimi podatki (NRRP) pri prijavi raziskovalnih projektov. Namenjena je zagotavljanju izpolnjevanja tehničnih in pravnih pogojev za izročanje podatkov v pooblaščenim področnim podatkovnim središčem.
- Vzpostavitev storitev podpore in usposabljanja za pripravo in izvajanje NRRP, ki jih nudijo pooblaščenim podatkovnim središčem in knjižnično informacijski servisi ustanov, ki delujejo na podlagi javno objavljenih notranjih politik ustanov, pripravljenih skladno z ustreznimi standardi in uveljavljenimi dobrimi praksami.
- Vzpostavitev storitve izbora in vrednotenja raziskovalnih podatkov glede na znanstveni pomen s strani pooblaščenih podatkovnih središč, podpora storitev za uveljavitev in izvajanje citiranja raziskovalnih podatkov pri uporabi v članku, ter podpora storitev za priznanje znanstvenega dosežka v obliki bibliografskega zapisa, kadar so raziskovalni podatki oddani v odprti dostop.
- Spodbujanje druge rabe podatkov, ki so brez dodatnih stroškov na razpolago pri pooblaščenih področnih podatkovnih središčih, pri novih raziskovalnih projektih.
- Zagotavljanje izvajanja NRRP ter vzdržnega delovanja infrastrukturnih raziskovalnih podpornih storitev za trajno hrambo in dostop do raziskovalnih podatkov.

2. Določila politike

a. Dolžnosti raziskovalcev

Raziskovalni podatki morajo biti kmalu po nastanku opremljeni z ustreznimi metapodatki, skladno s priporočili in standardi posameznih raziskovalnih področij, glede česar se je potrebno uskladiti s pristojnimi področnimi podatkovnimi središči. Minimalno morajo metapodatki vsebovati podatke o nastanku in dokumentacijo o predelovanju, vsebinske oznake in ostalo dokumentacijo, ki omogoča iskanje, dokumentacijo, ki omogoča neodvisno presojo kakovosti

podatkov in njihove uporabnosti za določen namen, ter informacijo o mestu, kjer so dostopni, in morebitnih potrebnih omejitvah dostopa.

Prejemniki javnih sredstev v okviru instrumentov financiranja raziskovalnih aktivnosti so pri prijavi dolžni priložiti Načrt ravnanja z raziskovalnimi podatki (NRRP). NRRP vsebuje podrobnosti glede načina uresničevanja obveznosti določenih v pričujoči politiki. V NRRP se poda predlog ocene dolgotrajnega pomena podatkov za znanost, enkratnosti, ter široke uporabnosti za več namenov in na tej podlagi predlog glede režima digitalnega skrbništva za zagotavljanje bodoče obstojnosti in uporabnosti podatkov. NRRP lahko vsebuje tudi izjavo, da projekt ne predvideva ustvarjanja raziskovalnih podatkov. Glede na pomen in vrednost raziskovalnih podatkov se lahko prilagodi režim ravnanja z njimi, kadar premajhna uporabnost podatkov ne opravičuje dodatnih vlaganj. V okviru NRRP mora biti obvezno utemeljena potreba po ustvarjanju novih podatkov, iz katere je razvidno, zakaj obstoječi podatki ne zadoščajo za namene projekta. Priporočeno je, da raziskovalne ustanove oblikujejo notranje politike ravnanja z raziskovalnimi podatki, na katere se lahko sklicuje posamezni projekt.

Raziskovalci morajo podatke najkasneje leto dni po nastanku pripraviti skladno z navodili in področnimi standardi ter jih tako oskrbljene izročiti pristojnemu podatkovnemu središču, ki bo izvedel postopek oceno izpolnjevanja formalnih pogojev za zagotovilo odprtega dostopa (format podatkov, opremljenost z minimalnimi metapodatki, ustrezne licence) ter presodil, kakšen je njihov dolgotrajni znanstveni pomen. Natančno morajo biti opredeljene in utemeljene izjeme odstopanja od praviloma popolnoma odprtega dostopa, npr. zaradi interesov državne varnosti, varovanja dediščine, pravic intelektualne lastnine privatnih sofinancerjev. Razčiščeno mora biti izpolnjevanje pravnih in etičnih pogojev za zagotovilo odprtega dostopa. V kolikor je dostop do podatkov zaradi utemeljenih izjem omejen, morajo biti za katalog področnega podatkovnega središča pripravljene vsaj prosto dostopni metapodatki.

Raziskovalne podatke, ki so objavljeni v uveljavljenih mednarodnih podatkovnih zbirkah, raziskovalci registrirajo v nacionalnem podatkovnem registru in zagotovijo metapodatkovni zapis v področnem podatkovnem središču, ki vsebuje povezavo na mesto dostopa.

Uporabniki raziskovalnih podatkov so dolžni te uporabljati skrbno in skladno s pogoji in določili glede uporabe ter z upoštevanjem prispevka avtorjev raziskovalnih podatkov. Znanstvene publikacije, ki uporabljajo raziskovalne podatke, morajo vsebovati bibliografsko referenco na raziskovalne podatke z vključenim stalnim identifikatorjem lokacije dostopa.

b. Odgovornost ustanov

Ustanove, nosilke raziskovalnih aktivnosti, pripravijo svoje notranje politike in postopke, na katere se lahko oprejo raziskovalni projekti pri pripravi in izvajanju NRRP z jasno opredeljenimi vlogami in odgovornostjo raziskovalcev, raziskovalno podpornih dejavnosti in vodstev ustanov. Pri tem upoštevajo ustrezne krovne politike, zakonske obveznosti in etična načela.

Vpelje se strokovno usposabljanje glede ravnanja s podatki (tečajji, delavnice, poletne šole, študijski programi), z vizijo, da postane obvezni del strokovnega usposabljanja raziskovalcev in podpornega osebja.

Vzpostavi se infrastruktura za stroškovno učinkovito izvajanje skupnih storitev v podporo izvajanju NRRP kot so tehnološke in svetovalne storitve za zagotovitev digitalnega skrbništva

nad podatki skozi celotni življenjski cikel. V politikah se je potrebno maksimalno nasloniti na obstoječo storitveno infrastrukturo za podporo projektov kot so raziskovalne pisarne in knjižnice projektov, OSIC, SICRIS, in izkoristiti delujoče področna podatkovna središča in storitvena omrežja.

c. Določila glede delovanja področnih podatkovnih središč

Področna podatkovna središča v sodelovanju z drugimi izvajalci knjižničnih in informacijskih storitev s področja znanosti izvajajo storitve svetovanja, usposabljanja in podpore pri pripravi in izvajanju NRRP, storitve prevzema podatkov, obravnave in priprave podatkov za dolgotrajno hrambo, medsebojno povezanega dostopa do podatkov pod ustreznimi režimi, ter skupaj z ustvarjalci podatkov storitve promocije, svetovanja, usposabljanja in podpore pri uporabi.

Pristojno podatkovno središče na podlagi registriranih NRRP potrdi načrtovane aktivnostih in ponudi ustrezno svetovalno podporo v teku celotnega življenjskega cikla izvajanja projekta, s čimer se zagotovi odstranjanje morebitnih ovir odprte dostopnosti raziskovalnih podatkov še pred začetkom zbiranja podatkov.

Pristojna podatkovna središča na podlagi vnaprej objavljenih kriterijev in postopkov ocenijo dolgotrajni pomen za znanost njim ponujenih podatkov. Sprejem podatkov je omejen na primere, pri katerih je jasna korist za znanost in zagotovljena stroškovna učinkovitost dolgotrajnega digitalnega skrbništva. Dokumentacijo postopka presoje shranijo v spremljajočih metapodatkih v registru raziskovalnih podatkov. Podatki, ki so kot znanstveno pomembni sprejeti v objavo v odprtem dostopu, se upoštevajo pri znanstvenih bibliografijah.

Podatkovna središča sodelujejo pri razvoju in uvajanju metapodatkovnih standardov, pri dodeljevanju stalnih identifikatorjev raziskovalnih podatkov in storitvah podpore navajanja podatkov v publikacijah in sledenja bibliografskih navajanj.

Pristojna podatkovna središča delujejo na podlagi javno objavljenih lastnih notranjih politik, ki vključujejo določila glede znanstvenega področja in vrste podatkov, ki ga pokrivajo, strategij ter postopkov dolgotrajne hrambe in strokovno ustreznih zagotovil glede njihovega izpolnjevanja ter opredelitve ciljne skupine uporabnikov in pogojev dostopa. Vsebuje naj tudi določila glede sodelovanja z drugimi področnimi ali ustreznimi mednarodnimi ustanovami pri ponudbi dostopa do raziskovalnih podatkov in določila glede nalog strokovnega razvoja, vzgoje kadrov in svetovanja uporabnikom.

d. Obveznosti in pravice javnih financerjev

Financer raziskovalnih projektov in raziskovalnih inštitucij lahko v primeru neizpolnjevanja določb umakne plačilo nadaljnjih obrokov financiranja.

Financer del sredstev raziskovalnega projekta nameni za izvedbo nalog opredeljenih v NRRP.

Financer zagotavlja upoštevanje prispevka na ustrezen način objavljenih raziskovalnih podatkov v okviru državnega sistema znanstvenega vrednotenja. Ustrezna objava pomeni objavo v odprtem dostopu, ki je ovrednotena z vidika dolgotrajnega znanstvenega pomena pri pristojnem podatkovnem središču in registrirana v podatkovnem registru skupaj z metapodatki in stalnimi identifikatorji.

Financer zahteva od kandidatov za sofinanciranje izdajanja domačih znanstvenih periodičnih publikacij, da uveljavijo notranje politike ravnanja z raziskovalnimi podatki, ki zadostijo zahtevam v tem dokumentu predstavljenih krovnih politik in izpolnjujejo vsaj naslednje: (1) odprt dostop do raziskovalnih podatkov, uporabljenih v publikacijah, (2) citiranje podatkovnih gradiv po načelu, ki je uveljavljeno za znanstvene publikacije, ter (3) pripoznanje podatkovnega gradiva kot sestavnega dela publikacije, ki je po uveljavljenih kriterijih znanstvene kakovosti podvrženo recenziji.

NRRP predstavlja sestavni del dokumentacije prijav raziskovalnih projektov, ki je ob sprejemu projekta na voljo pristojnemu podatkovnemu središču za potrebe izvajanja nadzora nad uresničevanjem prijavljenega načrta. Register za podporo javne evidence NRRP in register na podlagi teh načrtov ustvarjenih raziskovalnih podatkov, se vodita v povezavi z obstoječimi znanstveno informacijskimi sistemi o raziskovalni dejavnosti.

NRRP gre skozi postopek recenzije predlogov raziskovalnih projektov in njegova ocena predstavlja sestavni del celovite ocene vrednosti projekta.

Pristojna področna podatkovna središča so potrjeni s strani ARRS v okviru infrastrukturnih raziskovalnih programov, pri čimer se zagotovi organizacijsko in finančno vzdržno delovanje za opravljanje nacionalnih storitev in za sodelovanje v mednarodnih povezavah. Poskrbi se za nasledstvo odgovornosti skrbništva nad podatki v primeru prenehanja delovanja.

Financer zagotovi razpise za projekte študij izvedljivosti in demonstratorskih projektov vzpostavljanja podpornih storitev delovanja področnih podatkovnih središč in podatkovnih storitev v okviru raziskovalnih ustanov, priprave politik in njihovega izvajanja s strani ustanov, ter za razvojne projekte medinstitucionalnega nacionalnega in mednarodnega mrežnega sodelovanja akterjev s področja podatkovne infrastrukture.

Financer aktivno spodbuja drugo rabo obstoječih raziskovalnih podatkov v financiranih raziskovalnih projektih.

3. Uveljavitev določil

Določila te politike stopijo v veljavo ob naslednjih razpisih raziskovalnih projektov. Na področjih in ustanovah, kjer še ni ustreznih politik in potrebnih podatkovnih storitev, bo potrebno pripraviti akcijske načrte za izpolnitev vrzeli do konca leta 2015 in jih izpeljati do leta 2019.

Dokument politike se bo pregledoval in posodabljal enkrat na leto glede na zahteve izvajalcev politike in potrebe po usklajevanju s politikami držav EU.

Univerza v Ljubljani
Fakulteta *za družbene vede*

