

Izvirni znanstveni članek ■

Spletni informacijski sistem za podporo projektu Eurethnet

Web Based Information System for the Eurethnet Project Support

Jure Dimec, Branimir Leskošek

Izvleček. V članku predstavljamo podatkovno zbirko in spletno aplikacijo za vnos bibliografskih podatkov, ki smo ju na Inštitutu za biomedicinsko informatiko razvili za partnerje projekta Eurethnet, ki prihajajo iz držav, novih članic EU. Pri razvoju smo se osredotočili predvsem na izmenljivost podatkov z drugimi zbirkami projekta, podporo vsem evropskim naborom znakov, minimalizacijo obremenitev omrežja in zagotavljanje varnosti. Za razvoj sistema je bil na voljo izredno kratek čas, kar smo reševali z lastno programsko opremo za avtomatsko generiranje aplikacij.

Abstract. The paper presents a database and a web-based application for bibliographic data entry, which was developed at the Institute for Biomedical Informatics for Eurethnet project partners from EU New Member States. The development was focused mainly on records compatibility with other Eurethnet databases, support for all European character sets, minimalisation of network traffic, and security issues. The time window available for system development was very small and this problem was solved with our own software for automatic application generation.

■ **Infor Med Slov:** 2006; 11(2): 12-23

Institucija avtorjev: Inštitut za biomedicinsko informatiko
Medicinske fakultete v Ljubljani.

Kontaktna oseba: Jure Dimec, Inštitut za biomedicinsko
informatiko, Medicinska fakulteta, Vrazov trg 2, 1000
Ljubljana. email: jure.dimec@mf.uni-lj.si.

Uvod

Razmislek o etičnih posledicah raziskovalnega in strokovnega dela v vedah o živi naravi postaja vedno pomembnejši, raste pa tudi število dokumentov, prevsem raziskovalnih in zakonodajnih, ki se ukvarjajo s to tematiko. Videti pa je, da informacijska orodja, ki nudijo dostop do teh dokumentov, svoje naloge ne opravljajo dovolj dobro. H. Fangerau je v svoji študiji¹ pokazal, da so članki v revijah s področja medicinske etike zelo razpršeni v različnih bibliografskih zbirkah. Te zbirke je razvrstil z ozirom na pokrivanje medicinske etike in deset najvišje uvrščenih zbirk je skupaj pokrivalo le 45,2% revij s tega področja. Lahko bi ugovarjali, da najpomembnejše zbirke, denimo Medline, vključujejo le najkvalitetnejše revije, vendar pa ima področje etike nekatere posebnosti, ki sicer ne veljajo za večino drugih področij medicinskih in sorodnih ved. Za dobro informiranost in ustrezno odločanje niso pomembni samo raziskovalni članki, ampak, vsaj toliko, tudi drugi viri – članki, ki odsevajo odnos lokalne javnosti in politike do etične problematike, zakonodajna besedila in podobno, skratka viri, ki jih ne najdemo v najodmevnejših znanstvenih revijah. Že dolgo je znana določena pristranskost največjih bibliografskih zbirk v prid severnoameriškim revijam, situacija pa je nedvomno še bolj pereča, če gre za revije, ki izhajajo v državah, novih članicah EU (New Member States, NMS), zlasti če so besedila v ne-angleških jezikih.

Mednarodna mreža Eurethnet² in projekt z istim imenom sta bila ustanovljena z namenom, "da se poveča zavedanje evropske javnosti o bioetičnih problemih in pospešuje mednacionalni in medkulturni dialog" ter poveča vidnost dokumentov, ki opisujejo dosežke bioetičnih prizadevanj v evropskih okvirih.³ Projekt, ki ga je koordiniral Oddelek za medicinsko etiko in zgodovino medicine Univerze v Goetingenu, je, v okviru programa Quality of Life and Management of Living Resources 5. okvirnega programa, financirala Evropska Unija. Prvotni projektni konzorcij je sestavljalo 18 partnerjev iz 9 držav EU.

Delo se je osredotočalo na 6 glavnih sklopov: zbirki Euroethics in Endebit, informacijske produkte z dodano vrednostjo, tezaver TELS, dokumentacijske standarde in internetni portal.

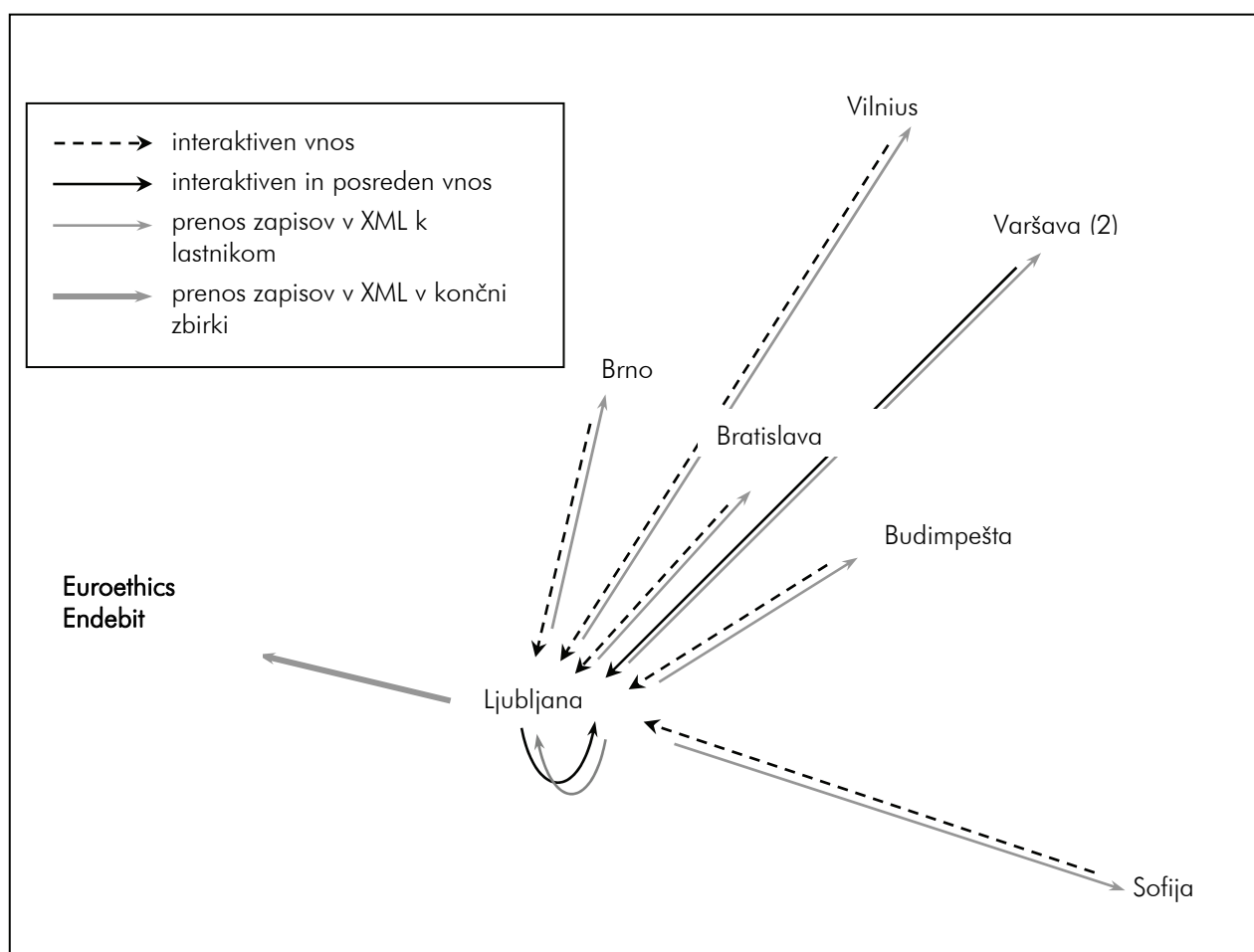
Euroethics in Endebit sta bibliografski zbirki, ki pokrivata evropsko produkcijo dokumentov s področja bioetike. Gradnja zbirke Euroethics,⁴ ki vključuje podatke o dokumentih s področja medicinske etike, se navezuje na rezultate projekta Euroethics Health Network, ki ga je financirala EU v okviru programa BIOMED-2 (1996 – 1999). Euroethics gradi in vzdržuje Univerzitetna knjižnica švedskega inštituta Karolinska (KIB), bibliografski zapisi pa izvirajo iz različnih nacionalnih zbirk partnerjev projekta Eurethnet. Zbirka Endebit⁵ vključuje opise dokumentov s področja biotehnologije, nastaja na podoben način, gradi pa jo Medoddelčni center za etiko v naravoslovju in humanistiki Univerze v Tuebingenu (IZEW), Nemčija. Obstaja tudi iskalnik⁶, ki omogoča poizvedovanje po obeh zbirkah hkrati. Informacijske produkte z dodano vrednostjo, naprimer izbrane bibliografije o aktualnih temah in sezname spletnih kazalcev na dokumente s področja etike, prispevajo različni projektne partnerji v okviru Eurethnet. Pomemben rezultat tega projekta je tudi tezaver TELS⁷ (Thesaurus Ethics in Life Sciences), ki je v rabi za opisovanje vsebine dokumentov v Eurethnet in Endebit, njegov razvoj pa vodi Nemški referenčni center za etiko v znanostih o živi naravi.

Leta 2003 se je pridružilo projektному konzorciju Eurethnet osem partnerjev iz držav NMS ter Bolgarije. Za sodelovanje teh institucij pri gradnji zbirk Euroethics in Endebit je bila predvidena ista metodologija, kot je veljala za stare članice – identifikacija nacionalnih bibliografskih zbirk oziroma knjižničnih katalogov z bioetično vsebino, izvažanje relevantnih zapisov skupaj z opisi njihove strukture, pošiljanje skrbnikom zbirk Euroethics in Endebit, preformatiranje v formate teh zbirk in nalaganje vanje. Vendar pa je sčasoma postalo jasno, da pri partnerjih iz NMS (z izjemo Slovenije in deloma Poljske) ni na voljo ustreznih zbirk ali katalogov. Časa je zmanjkovalo, ker je bil zaključek projekta sprva predviden za konec l.

2004, brez podatkov iz NMS pa njegov cilj ne bi bil dosežen in financiranje ustavljeno. Analizirane so bile različne možnosti za izhod iz te krize. Partnerji iz NMS so bili pripravljene na gradnjo novih lokalnih zbirk, a niso imeli na voljo ustrezne programske opreme, sredstva za nakup pa v projektu niso bila predvidena. Opravljene so bile analize ustreznosti cenovno dostopne programske opreme za gradnjo osebnih bibliografskih zbirk, vendar na tržišču ni take, ki bi ustrezala vsem zahtevam projekta (prilagodljive strukture,

podpora različnim naborom znakov, možnost vključevanja tezavrov...).

Spomladi 2004 je bila sprejeta pobuda Inštituta za biomedicinsko informatiko Medicinske fakultete v Ljubljani (IBMI) in slovenske Komisije za medicinsko etiko o razvoju centralizirane zbirke in ustrezne aplikacije, ki bi omogočala partnerjem z NMS omrežni vnos podatkov. Odločeno je bilo, da se pri tem uporabi odprtokodna programska oprema in spletna tehnologija. Razvoj zbirke in aplikacije je bil zaupan strokovnjakom IBMI.



Slika 1 Povezanost partnerjev v sistemu. Vsi centri vnašajo podatke online, dva (ljubljski in eden od poljskih) pa tudi z izvažanjem iz svojih zbirk. Dokončani zapisi se pretakajo v zbirki Euroethics in Endebit ter, na željo, v centre, ki so jih vnesli.

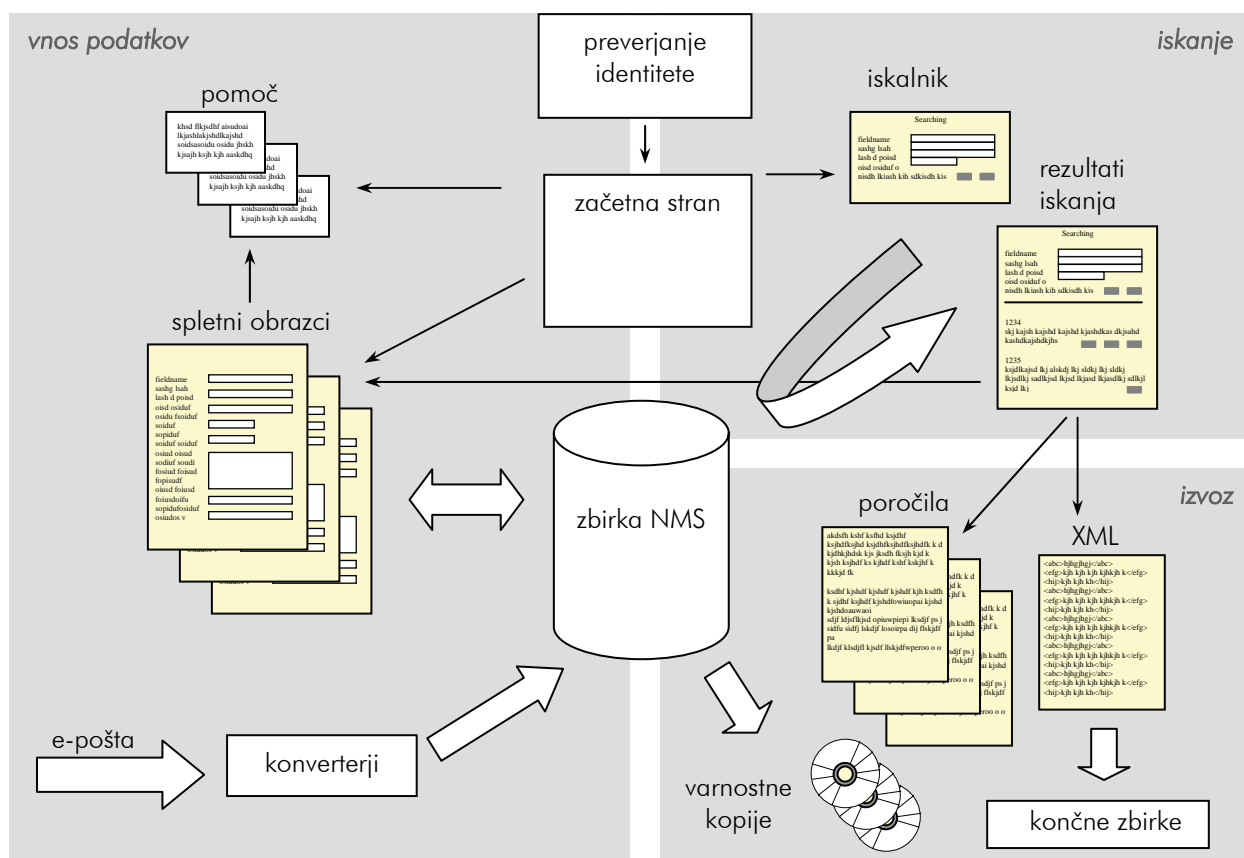
Metodologija in opis sistema

Sistem, ki ga opisujemo, sodi v mrežo zbirk projekta Eurethnet, ki predstavljajo vseevropsko shrambo bibliografskih zapisov o dokumentih s področij medicinske in biotehnoške etike.

Gradnja zbirk Euroethics in Endebit poteka v dobro definiranem in urejenem okolju, ki vključuje zbiranje zapisov iz obstoječih nacionalnih zbirk, do neke mere drugačna situacija v NMS pa zahteva drugačen pristop. Zbirka in aplikacija za vnos iz NMS služi kot vmesna postaja med avtorji zapisov, ki so običajno zdravniki, vpleteni v odločanje o etičnih problemih in končnima zbirkama Euroethics in Endebit.

Slika 1 prikazuje splošno strukturo sistema za zbiranje podatkov. Vpletenih je devet institucij: IBMI, Ljubljana, kot avtor aplikacije in gostitelj

zbirke (v nadaljevanju imenovane zbirka NMS), KIB in IZEW kot gostitelja zbirk Euroethics in Endebit in šest centrov, aktivnih na različnih področjih bioetike, ki prispevajo bibliografske zapise. Štirje od šestih centrov vnašajo podatke le z uporabo interaktivnih funkcij (vnos "online"), dva centra (Ljubljana in Varšava) pa pretakata tudi podatke iz njihovih lokalnih zbirk. V tem primeru se izvoženi podatki preformatirajo z uporabo posrednih funkcij aplikacije in vključijo v zbirko NMS. Občasno se zapisi iz zbirke NMS izvozijo v formatu XML in pošljejo v Euroethics ali v Endebit ali v obe zbirki. Pošiljajo se le zapisi, ki imajo izpolnjena vsa obvezna polja. Če pride do spremembe v že poslanem zapisu, se ta ponovno pošlje v ciljno zbirko, kjer nadomesti prejšnjo verzijo. Tudi partnerji iz NMS si lahko pretočijo iz zbirke zapise, ki so jih prispevali, v formatu XML.



Slika 2 Shematski prikaz funkcij sistema. Enostavne puščice predstavljajo možna dejanja uporabnikov, široke puščice pa tok podatkov.

Odločitev o gradnji zbirke, centralizirane z vidika partnerjev iz NMS, ima nekatere dobre posledice. Zapisi so, neglede na izvor, že ob nastanku oblikovani v skladu s standardnim formatom, ki je v rabi v projektu Eurethnet, s čemer smo se izognili izgubi informacij, ki je zelo pogosta pri pretvorbah med formati. Težave s prilagoditvijo na različne nabore znakov, ki so v rabi v Evropi, morajo biti rešene le enkrat, da ne omenjamo enostavnejšega zagotavljanja varnosti, vključno z varnostnimi kopijami podatkov. Posledice uporabe spleta kot infrastrukture so dvojne – dobre in malo manj dobre. Ker so vsi programi, ki izvajajo logiko aplikacije, nameščeni na strežniku, ali pa se sproti prenašajo s spletnimi obrazci, odpade potreba po nameščanju odjemalske programske opreme, partnerji iz NMS pa so osvobojeni skrbi za tehnične detalje. Njihove računalniške potrebe so skromne – ustreza vsak PC s spletnim brskalnikom. Zbiranje besedil z različnimi nabori znakov ali celo v različnih pisavah je (teoretično) enostavno vsaj v fazi vnašanja podatkov, ker vsi najbolj razširjeni brskalniki podpirajo standard Unicode. Po drugi strani pa pride do izraza dokajšna okornost spletnega uporabniškega vmesnika, ta pa je posledica omejitev, ki jih postavljata HTTP in HTML. Aplikacija, ki komunicira z uporabnikom preko spleta, ima omejen dostop do njegovih računalniških virov, kot so pomnilnik, procesorska moč in prostor na disku. Vsa programska koda, potrebna za delovanje aplikacije, mora biti vključena v spletno stran, zato je potreben za vsako uporabnikovo intervencijo obsežen prenos podatkov in programov, ki manipulirajo s podatki.

Po izčrpni analizi situacije smo identificirali nekaj osnovnih zahtev, ki jim mora zadostiti sistem:

1. Biti mora uporaben tudi pri počasnih povezavah in z uporabo katerekoli računalniške platforme, sposobne spletne komunikacije.
2. Zapisi morajo biti neposredno združljivi z ostalimi relevantnimi bibliografskimi zbirkami, ki nastajajo v okviru projekta Eurethnet.
3. Uporabniški vmesnik mora biti enostaven in intuitivno jasen.
4. Vse funkcije, ki imajo opravka z besedili, morajo podpirati vse nabore znakov različnih evropskih jezikov.
5. Varnostna politika mora urejati dostop do zbirke, prenos podatkov in preprečevati nehotene poškodbe podatkov.
6. Institucije, ki prispevajo podatke, ostajajo njihovi lastniki in morajo imeti možnost njihovega izvoza v standardnem formatu.
7. Uvažanje zapisov v drugih, dobro definiranih formatih, mora biti omogočeno.
8. S istem mora biti dovolj enostaven, da bo operativen v zelo kratkem času.

V članku opisujemo, kako smo zadostili navedenim zahtevam.

Na sliki 2 so prikazane tri skupine funkcij, ki omogočajo v sistemu vnos in popraviljanje, iskanje in izvoz podatkov. Običajen način vnosa bibliografskih zapisov je s pomočjo spletnih obrazcev, dodatno možnost pa nudi fleksibilen program za konverzije, ki zapise, poslane na IBMI v nekem drugem formatu, preoblikuje v format zbirke NMS in jih vključi vanjo. Obrazci za iskanje pomagajo pri izbiranju zapisov ali skupin zapisov, ki zahtevajo uporabnikovo pozornost, najsibo to popraviljanje, izdelavo izpisov ali izvažanje. Moduli za izvažanje, poleg različnih oblik izpisov, nudijo izdelavo varnostnih kopij in izvažanje podatkov v formatu XML, tako podmnožice, ki jo je kreiral posamezen partner in jo izvaža za svojo lokalno rabo, ali podmnožice, ki se prenaša v končni zbirki.

Zbirke, ki nastajajo v konzorciju Eurethnet, so v osnovi bibliografske, vendar zaradi namena projekta, da informira o etičnih temah, ki se pojavljajo v vseh plasteh javnosti, vključujejo tudi podatke, ki se pojavljajo v ne-znanstvenih medijih. Posebno zanimive informacije lahko izvirajo iz dnevnega časopisja ali v zakonodajnih aktih. S tem

v mislih standardi projekta definirajo (naša aplikacija pa implementira) širši nabor tipov dokumentov, kot je to običajno za bibliografske zbirke. Med t.i. sekundarnimi tipi, ki natančneje definirajo tip dokumenta, so, med drugim:

odločitev sodišča, predpis, zakon in parlamentarna publikacija. Dostopnost takih dokumentov, če niso v e-obliki, je pogosto otežkočena, zato v takem primeru zapisi vsebujejo tudi podatke o lokalni dostopnosti.

https://unimed.mf.uni-lj.si - Book chapter, conference paper (analytic) - Mozilla Firefox

Book chapter, conference paper (analytic)

creator: tone NMEC-RS
creation date: 25-10-2004
document ID: 511
updated 18-07-2005

Author *	<input type="button" value="delete"/>	Trontelj JV
Author's type *		author
Author's affiliation		SPS Nevrološka klinika, Klinični oddelek za nevrofiziologijo, Klinični center Ljubljana
Author's email		jaze.trontelj@
Ordinal position of author *		1

Author *	<input type="button" value="delete"/>	Rees D
Author's type *		author
Author's affiliation		- choose one -
Author's email		corpor. author
Ordinal position of author *		2

Author *		
Author's type *		- choose one -
Author's affiliation		
Author's email		
Ordinal position of author *		

Done unimed.mf.uni-lj.si

Slika 3 Spletni obrazec za zbiranje podatkov o avtorstvu. Isti obrazec je v rabi za različne tipe avtorstva.

Delovanje sistema v neugodnih pogojih

Razvoj sistema je temeljil na pričakovanju, da ga uporabniki ne bodo uporabljali v širokopasovno povezanih računalniških centrih, temveč na svojih delovnih mestih. Ker so ta večinoma v okoljih,

kjer stanje računalniških povezav še vedno zaostaja za zahodnoevropskimi standardi, je bila naša osnovna skrb namenjena minimalizaciji količine podatkov, ki se prenašajo po omrežju med vnosom ali popravljanjem podatkov. Odločili smo se, da na uporabnikovi strani uporabimo le standardne, računalniško nezahtevne rešitve iz

arzenala spletnega programiranja, brez slik, animacij ali apletov.

Zbiranje podatkov s spletnimi obrazci lahko zelo obremeni omrežno infrastrukturo. Vsaka izmenjava spletnega obrazca, neglede na to, ali obrazec pošljamo, ali sprejemamo, povzroči prenos vseh opisov polj, programskih vključkov in morebitnih podatkov, ki so že vnešeni v ta polja. Struktura bibliografskega zapisa zbirke NMS je kompleksna, z velikim številom polj in številna med njimi je treba kontrolirati glede pravilnosti vnosa. Kompletna struktura z vsemi polji vključenimi v en sam obrazec, ki bi se prenašala ob vnosu ali pregledovanju enote podatkov, bi verjetno povzročala nesprijetno obremenitev omrežja, zato jo je bilo treba razdeliti na večje število obrazcev, prijaznejših do omrežja. Osnovi za delitev sta bili dve – dokumentacijska pravila, določena v projektu in tendenca nekaterih polj, da nastopajo skupaj. Dokumentacijska pravila razlikujejo med sedmimi tipi dokumentov, zanimivimi za etično srenjo. Obrazci za vnos so sledili posebnostim teh tipov, kar je že občutno zmanjšalo število polj v vsakem od njih (v zapisu o članku v reviji, naprimer, nista potrebni polji naslov knjige in ISBN).

Podatki, ki sodijo v nekatere skupine polj, se pojavljajo v vsakem dokumentu, neglede na njegov tip, in jih je, v skladu z njihovo naravo, najlažje organizirati kot samostojne entitete v zbirki. Taki podatki opisujejo avtorstvo in opis vsebine. V skladu z zmanjševanjem omrežnega prometa smo te podatke zbirali v ločenih obrazcih. Slika 3 prikazuje obrazec za zbiranje podatkov o avtorstvu dokumentov. O vsakem avtorju smo zbirali več, kot samo njihova imena, zato je bil potreben spletni obrazec s ponovljivimi polji. Programi, ki generirajo obrazce, ne poznajo vnaprej števila avtorjev in s tem števila zapisov, ki jih morajo pripraviti za nek dokument, zato je bil izhod v iterativnem oblikovanju obrazca. Vsakič, ko se tak obrazec pošlje spletnemu brskalniku, je v njem prostor za podatke o novem avtorju. Ko so ti podatki vnešeni in poslani strežniku, se obrazec oblikuje na novo, spet s praznimi polji za

naslednjega avtorja. Različne tipe avtorstva, ki jih predvidevajo dokumentacijski standardi, zbiramo z istim obrazcem, ob vključevanju v zbirko pa se podatki shranijo v različna polja.

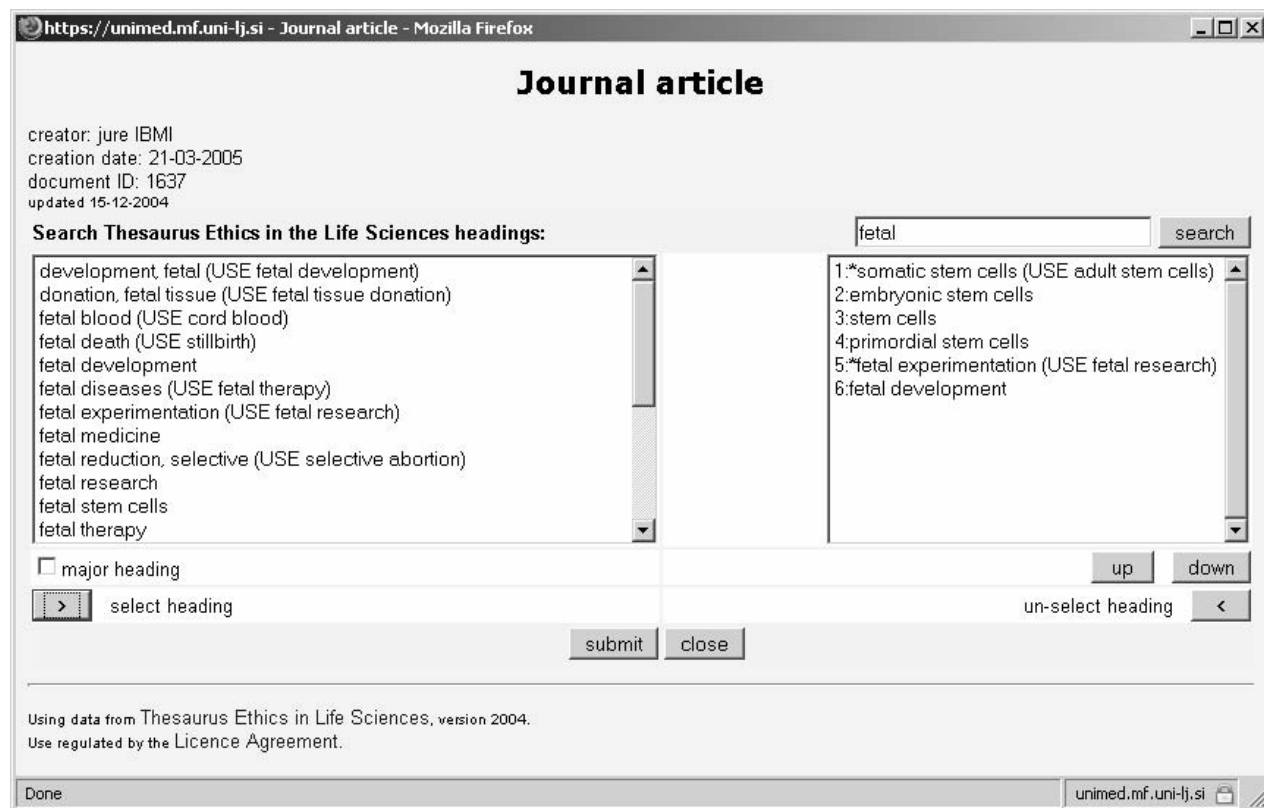
Podobno načelo samostojnih obrazcev smo uporabili za vse tri tipe polj vsebinskega opisa dokumentov. Dokumentacijski standardi definirajo 7 tipov nekontroliranih izrazov, ki pa jih je, na srečo, mogoče zbirati z istim spletnim obrazcem in razpršiti med polji zapisa ob vključevanju v zbirko. Način opisovanja vsebine dokumentov s kontroliranimi izrazi, deskriptorji, je opisan drugje v tem članku.

Vrstni red, v katerem so za neki dokument vnešeni zapisi o avtorjih, je pomemben in ga je treba ohraniti. Načela relacijskega modela ne zagotavljajo, da bodo vsakič priklicani v istem zaporedju, zato ima vsak zapis o avtorju dodatno polje, ki beleži mesto tega avtorja v dokumentu.

Tezavri

Dokumentacijski standardi projekta Eurethnet zahtevajo tri tipe polj vsebinskega opisa – t.i. nekontrolirane izraze, deskriptorje iz tezavra MeSH⁸ in deskriptorje iz tezavra Ethics in Life Sciences (TELS), ki so tudi obvezni za vsak zapis. Oba tezavra, MeSH in TELS, sta v rabi v zbirkah projekta, vključno z zbirko NMS.

Želeli smo olajšati izbiranje deskriptorjev in istočasno zmanjšati možnost vnašanja napak v zbirko, zato se nam je zdelo zelo pomembno vključiti e-obliki obeh tezavrov v aplikacijo za vnos bibliografskih podatkov. V skladu z načelom popolne prenosljivosti in neodvisnosti od uporabnikove računalniške opreme smo opustili možnost nameščanja tezavrov na uporabniški strani. Spletna tehnologija ne omogoča začasnih namestitev, zato je bil edini izhod iz zagate namestitev in uporaba tezavrov na strežniku aplikacije. Oba spletna obrazca, namenjena izbiranju deskriptorjev, sta podobna, slika 4 pa prikazuje tistega, ki je namenjen rabi tezavra TELS.



Slika 4 Po iskanju z nizom 'cells' so bili izbrani štirje deskriptorji iz tezavra TELS, dva pa po naslednjem iskanju z nizom 'fetal'. Izraza #1 in #5 sta sinonima ustreznih deskriptorjev, v katere bosta avtomatsko prevedena ob vključevanju v zbirko.

Obrazec sestavljajo iskalno okno, dve okni za izbiranje deskriptorjev in nekaj gumbov. Prenašanje celotnega tezavra TELS ob vsakem odpiranju obrazca za indeksiranje bi bilo zelo nepraktično in potratno. Uporabnik najprej vnese v iskalno okno niz, za katerega meni, da je vsebovan v deskriptorju ali deskriptorjih, ki dobro opisujejo vsebino indeksiranega dokumenta, sproži iskanje in le podmnžica teh deskriptorjev se prenese s strežnika v levo okno. Uporabnik v njem izbere tiste, ki ustrezajo in jih z gumbom [>] prenese v desno okno. Iskanje poteka tudi po sinonimih deskriptorjev. Če uporabnik v levem oknu izbere sinonim, se ta prevede v prednostni izraz. Tako bo izraz #1 na sliki 4 – 'somatic stem cells (USE adult stem cells)' shranjen v zbirki kot 'adult stem cells'.

Vključevanje besedil v nacionalnih naborih znakov

Uporabnost Eurethnetovih zbirk temelji tudi na njihovi večnacionalni naravi, ki pa prinaša zaplete, povezane s shranjevanjem podatkov, zapisanih z različnimi nacionalnimi nabori znakov v rabi v partnerskih institucijah. Težave z nabori znakov niso pereče v zbirkah, kakršna je Medline, ki organizirajo podatke o člankih z angleščino kot prevladujočim jezikom znanstvene komunikacije, v našem primeru pa so bile. Da bi omogočili doseganje osrednjega cilja bibliografskih zbirk projekta Eurethnet – nudenje karseda širokega vpogleda v etični diskurz – je bilo neizbežno vključevanje pomembnih delov bibliografskih zapisov, zapisanih v ne-ASCII naborih znakov. Poleg polstandardnih Microsoftovih kodnih strani obstajajo različni standardni nabori, v rabi v zahodnoevropskih, srednjeevropskih in baltskih

državah, vpletenih v Eurethnet. Običajni nabori znakov v teh okoljih sodijo v družino ISO 8859 (npr. ISO 8859-1 v rabi v Zahodni Evropi, ISO 8859-2 v Srednji Evropi), vendar pa se opisi teh naborov prekrivajo – iste identifikacije znakov predstavljajo različne znake pri različnih predstavnikih družine ISO 8859. Kmalu je postalo jasno, da mora aplikacija za vnos in zbirka NMS podpirati tak nabor znakov, ki velja za vse evropske države in istočasno ne predstavlja nepremagljive ovire za skrbnike zbirk Euroethics in Endebit, ki so do sedaj vključevale le podatke, zapisane v ASCII. Očiten izbor je standard Unicode,⁹ natančneje verzija UTF-8. Lastnost UTF-8, da pokriva vse nacionalne nabore, a istočasno ohranja identifikacije znakov, ki sodijo v standard ASCII ter je zato neposredno združljiva z njim, je bila še posebno dobrodošla. To dejstvo je precej olajšalo nadgradnjo zbirk Euroethics in Endebit, da so lahko začele vključevati tudi zapise iz NMS.

UTF-8 podpirajo novejšje verzije številnih programov, namenjenih procesiranju besedil, vključno s spletnimi brskalniki, spletnimi strežniki, gonilniki podatkovnih zbirk in sistemi za upravljanje s podatki. Programska okolja, ki smo jih uporabljali pri razvoju aplikacije, so bila izbrana s to lastnostjo v mislih.

Kontrola podatkov in zagotavljanje varnosti

Za zagotavljanje minimalizacije omrežnega prometa je zelo pomembno odkrivanje čimvečjega deleža napak že na odjemalski strani, s čemer se izognemo nepotrebnim izmenjavam okvarjenih podatkov in diagnoz o napakah. Obrazci za vnos so že oblikovani tako, da zmanjšujejo možnost vnašanja napak. Na višjem nivoju aplikacijska logika zmanjšuje možnost hotenih ali nehotenih destruktivnih dejanj na podatkih, ki so že vnešeni v zbirko in možnost razkritja podatkov.

Podatke, vnešene v spletne obrazce, programsko preverjamo večinoma na uporabnikovi strani, takoj po uporabnikovi zahtevi za shranjevanje in pred dejanskim prenosom. Bilo bi bolj v skladu z

varnostnimi priporočili, če bi vpeljali še drugo kontrolo na strežniški strani pred vključevanjem podatkov v zbirko. Z njo bi preprečili posledice morebitnega prestrezanja in namernega poškodovanja podatkov, vendar pa smo prepričani, da je zelo malo verjetno, da bi bile, ob uporabi protokola HTTPS in kodiranju podatkov, take poškodbe neopažene in podatkov ne bi zavrnili postopki preverjanja, ki jih opravlja sama podatkovna zbirka.

Preverjanje podatkov opravljajo programski vključki v jeziku JavaScript, ki so del spletnih obrazcev. Preverjajo prisotnost podatkov v obveznih poljih, znake v poljih z omejenim naborom (npr. numeričnih) in obliko podatkov v poljih, za katera je predviden standarden format (npr. datumi, številke ISSN, e-poštni naslovi, URL-ji...). Kadarkoli so vnaprej znane vse možne vrednosti nekega polja, se potreba po kontroli zmanjša z uporabo standardnih elementov, kot so izbirni gumbi (radio buttons), potrdilna polja (checkboxes) in padajoči sezname (pull-down lists). Potrdilna polja smo uporabili celo v primeru, ko je imelo polje 45 možnih vrednosti (sekundarni tip dokumenta), s čemer je nastalo obsežno področje z izbirami, ki je delno vidno na sliki 5. Izkušnje so nam pokazale, da je tak način, še posebno, če je možnih več izbir, zelo pregleden, in časovni prihranek pri uporabi odtehta večjo porabo prostora. Za polja z mnogimi možnimi vrednostmi, denimo deskriptorji iz tezavra, smo razvili posebne rešitve, ki so predstavljene v ustreznem poglavju.

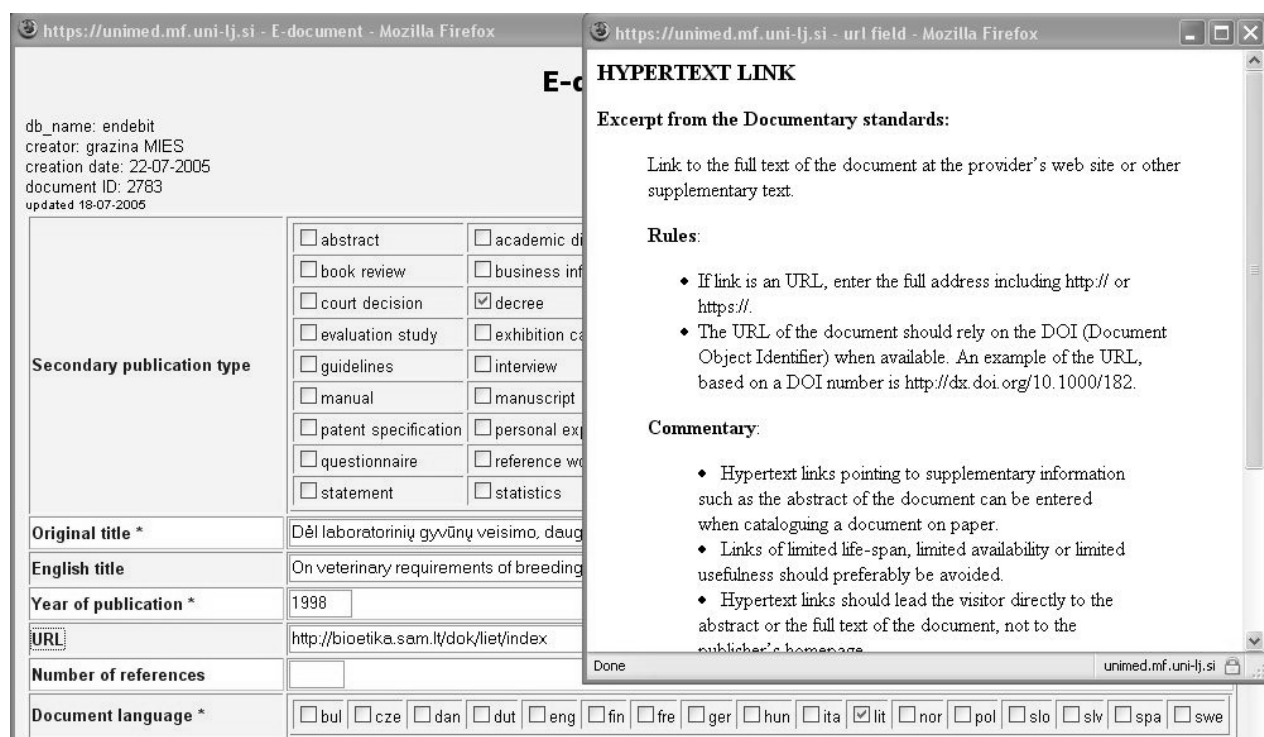
Dostop do vseh funkcij sistema, razen do uporabe priročnika, je dovoljen le registriranim uporabnikom, promet med odjemalci in strežnikom pa je kodiran z 128-bitnim protokolom SSL. Vsak uporabnik se identificira z uporabniškim imenom in geslom, vsa njegova dejanja pa se beležijo. Bibliografski podatki v zbirki so vidni vsem uporabnikom, dovoljenja za popraviljanje in druge potencialno destruktivne funkcije pa imajo le uporabniki iz partnerskega centra, ki je prispeval zapis. Isto načelo velja za izvažanje zapisov – uporabniki si lahko pretočijo samo tiste zapise,

označene v jeziku XML, ki jih je prispeval nekdo iz njihovega centra.

Iskanje

Pomemben del aplikacije je vmesnik, ki omogoča iskanje in izdelovanje izpisov. Iskalni vmesnik ne poskuša biti klasični iskalnik po bibliografski zbirki, ampak orodje, ki pomaga uporabnikom pri izboru zapisov, ki jih je treba na nek način spremeniti. Iskalna polja, prisotna v vmesniku, so bila izbrana s to funkcijo v mislih. Uporabnik lahko opravi selekcijo na osnovi intervala identifikacij, imena zbirke, ki ji je namenjen zapis, primarnega tipa

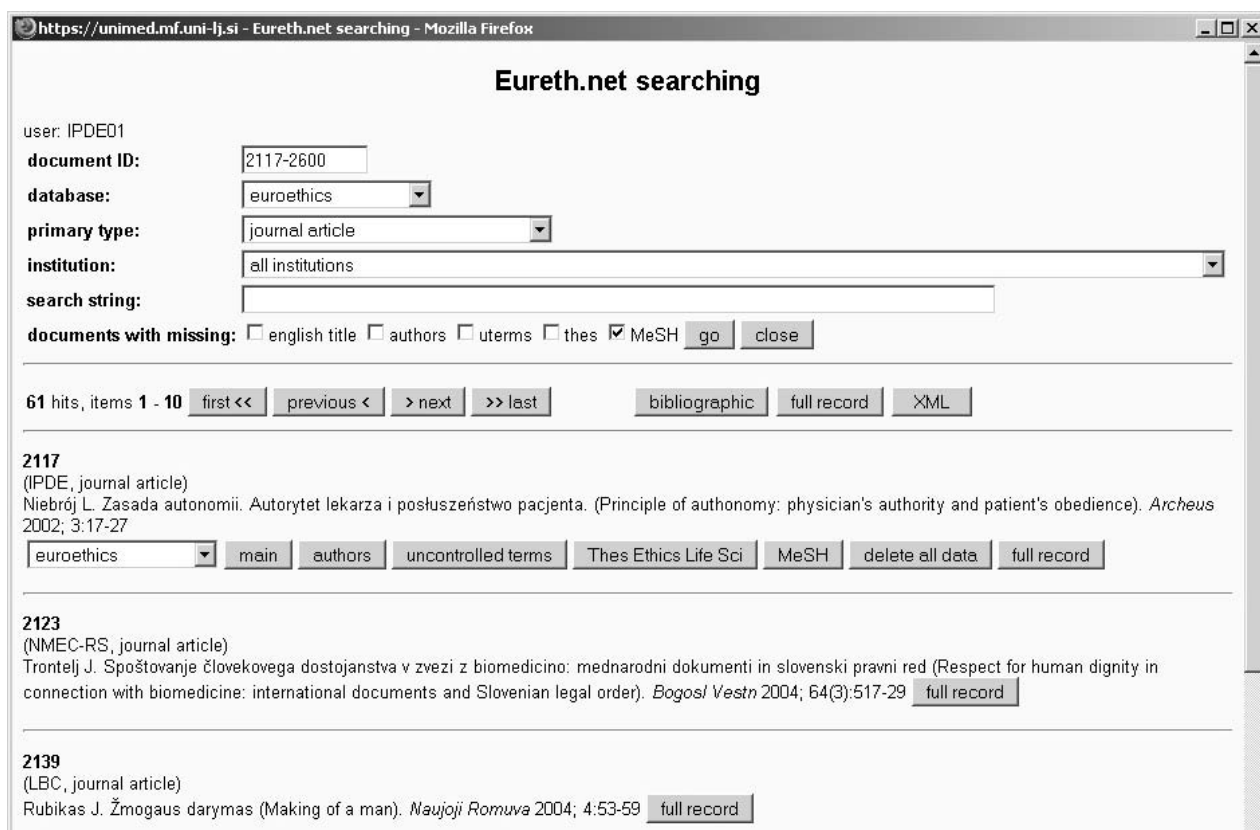
dokumenta, imena institucije, ki je zapise prispevala, ali z iskanjem po besedah v vseh besedilnih poljih. Vrstica potrdilnih polj, naslovljena 'documents with missing:' ima vlogo bližnjice do zapisov, ki jim manjkajo vrednosti v enem ali več pomembnih polj. Ta možnost je praktična, ker lajša delitev nalog pri vnosu podatkov. Ena oseba lahko vnese osnovne podatke o dokumentih, nekdo drug pa enostavno izbere ustrezno podmnožico zapisov in jih indeksira. Slika 6 prikazuje iskalni vmesnik, v katerem je uporabnik kombiniral iskalne kriterije, s katerimi je izbral zapise o člankih z ID-ji med 2117 in 2600, namenjene v zbirko Euroethics in še ne indeksirane z deskriptorji MeSH.



Slika 5 Okno za pomoč uporabnikom z navodili za vnos hipertekstnih kazalcev. Navodila so na voljo za vsako polje vsakega spletnega obrazca, okno na sliki pa se je odprlo po kliku na ime polja URL.

Med poiskanimi zapisi na sliki 6 imajo nekateri gumbje, ki priključijo obrazce za editiranje, drugi pa ne, odločitev za to pa je osnovana na pripadnosti iskalca. Oseba, ki je prispevala zapis #2117 in

iskalec prihajata iz iste institucije, zato ima iskalec pravico spreminjati podatke v tem zapisu, ne pa v zapisih #2123 in #2139, ki si jih lahko le ogleda s pritiskom na gumb [full record].



Slika 6 Iskalni vmesnik, namenjen izbiri zapisov, ki zahtevajo spreminjanje. Od poiskanih zapisov sme uporabnik IPDE01 spreminjati le zapis #2117, ki ga je vnesel nekdo iz njegovega centra.

Dodatna naloga iskalnega vmesnika je izdelava izpisov, neobremenjenih z gumbi za editiranje. Za uporabnike je posebno pomembna možnost izvažanja zapisov, označenih z XML. Medtem, ko izpisi referenc ali polnih zapisov (dostopni z gumboma [bibliographic] in [full record]) delujejo na vsej množici zadetkov iskanja (ali celi zbirki, če je bilo iskanje opravljeno brez omejujočih kriterijev), sproži gumb [XML] pretakanje samo tistih zapisov, ki jih je vnesel nekdo iz iste institucije, v katero sodi oseba, ki jih pretaka. Ta možnost izvažanja zapisov je popolnoma v skladu z eno od osnovnih idej projekta Eurethnet – poleg skupnih tudi pospeševanje razvoja lokalnih zbirk in informacijske podpore službam s področja bioetike. Z vidika posameznega partnerja deluje zbirka NMS tudi kot njegova lokalna shramba za zbiranje in shranjevanje zapisov, iz katere si jih bo pretočil, ko bo imel možnost namestitve in vzdrževanja ustrezne programske opreme za zbirke podatkov.

Pretakanje zapisov, opisanih z XML, bo prav gotovo olajšalo ta postopek.

Pomoč uporabnikom

Zaradi časovnih omejitev in geografske razpršenosti partnerjev iz NMS ni bilo mogoče organizirati delavnic za usposabljanje uporabnikov sistema. Vsa navodila, ki so jih potrebovali za vnašanje podatkov, je bilo treba vključiti v samo aplikacijo. Na vstopni strani je na voljo izčrpen uporabniški priročnik.¹⁰ Poleg njega lahko uporabnik priključ za vsako polje vsakega spletnega obrazca stran z navodili, na kateri so deli dokumentacijskega standarda, ki razlagajo vsebino in obliko podatkov ter deli uporabniškega priročnika, ki razlagajo način dela s tem poljem. Primer dela strani z informacijami o vnosu hipertekstnih kazalcev, ki jo je uporabnik priklical s klikom na ime polja URL, je prikazan na sliki 5.

Reševanje časovne stiske

Pomembna značilnost razvoja sistema je bilo pomanjkanje časa. Zaradi pozne vključitve partnerjev iz NMS v projektno skupino je obstajala resna nevarnost, da njihov prispevek v skupne zbirke ne bo pripravljen pred zaključkom projekta Eurethnet. Bilo nam je jasno, da bo, kljub natančnemu načrtovanju in jasnih predstavah o funkcionalnosti sistema ter pričakovanih uporabnikov, zelo težko končati razvoj v danih časovnih okvirih. Namesto običajnih in zamudnih iteracij programiranja in testiranja aplikacije, smo razvili programsko opremo za generiranje aplikacij, ki je, po našem mnenju, občutno skrajšala čas, potreben za razvoj.

Jedro programske opreme za generiranje aplikacij je formalizem, ki smo ga imenovali *jezik za definiranje aplikacij* (Application Definition Language – ADL), eklektična mešanica SQL-ovega Data Definition Language (DDL) in oznak v duhu HTML. Z ADL smo opisali zbirko, uporabniški vmesnik in logiko aplikacije, ki ju uporablja. Z vidika zbirke je ADL definiral vse entitete, njihove atribute in povezanost entitet. Z vidika spletnih obrazcev pa smo z ADL opisali preslikavo elementov zbirke v obrazce, soodvisnost in ponovljivost obrazcev, pa tudi načine, na katere so bili podatki zbirani (padajoče sezname, izbirne gumbke, besedilna polja...), skupaj z morebitnimi dovoljenimi in prednastavljenimi vrednostmi. Konstrukte v ADL smo vpisovali v MS Excel-ove preglednice, ki predstavljajo priročen vmesnik za sestavljanje definicij na enem mestu. V naslednjem koraku se iz preglednic zgradijo definicijske datoteke za gradnjo zbirke in spletnih obrazcev. Te datoteke t. i. *transformacijske skripte*, napisane v jeziku Perl, pretvorijo v CGI skripte, module za administracijo uporabniških podatkov in DDL skripte, ki zgradijo tabele zbirke. Vsi deli aplikacije uporabljajo tudi dve knjižnici – t. i. *validacijsko knjižnico* v JavaScript-u in knjižnico CGI.

Tako relativno zapleteno razvojno okolje nam je omogočilo hitre posege, kadarkoli je bila potrebna sprememba aplikacije, bodisi da je šlo za boljšo prilagoditev dokumentacijskim standardom, ali za spremembo, ki so jo za lažje delo zahtevali uporabniki. Vsako spremembo smo vnesli v opis v ADL in v roku nekaj minut ponovno avtomatsko generirali celo aplikacijo. V primerih, ko zahtevane spremembe niso bile povezane s spreminjanjem tabel zbirke, sploh ni bilo treba prekinjati vnosa podatkov.

Zaradi takega pristopa so bile prve verzije programov pripravljene v treh mesecih od začetka razvoja, do desetega meseca, ko je bil razvoj uradno zaključen, pa je bila potrebna ena sama obsežna revizija aplikacije.

Literatura

1. Fangerau H. Finding European bioethical literature: an evaluation of the leading abstracting and indexing services. *J Med Ethics* 2004;30(3):299-303.
2. Eureth.net. European Information Network - Ethics in Medicine & Biotechnology. <http://www.eureth.net/>
3. Fangerau H, Simon A, Wiesemann C. Improving information systems in Europe: EURETHNET. *Med Health Care Philos* 2003;6(1):67-9.
4. Domača stran zbirke Euroethics: <http://www.idem.uni-goettingen.de/euroethics/database.htm>
5. Domača stran zbirke Endebit: <http://www.eureth.net/endebit/index.htm>
6. Skupni iskalnik po zbirkah Euroethics in Endebit: <http://www.eureth.net/multisearch.htm>
7. Thesaurus Ethics in the Life Sciences, 3rd edition. <http://www.drze.de/BELIT/thesaurus?la=en>
8. Medical Subject Headings. <http://www.nlm.nih.gov/mesh/meshhome.html>
9. Unicode Home Page. <http://unicode.org/>
10. Dimec J, Leskošek B. Data entry application for NMS partners. User manual. Ljubljana: Medicinska fakulteta, 2005: 24. <https://unimed.mf.uni-lj.si/javascript/manual-v2.pdf>.