

PREGLED SISTEMOV PREDMETNIH OZNAK ZA PODROČJE ZNANOSTI O ŽIVALIH V IZBRANIH BIOTEHNIŠKIH PODATKOVNIH ZBIRKAH

Tomaž BARTOL¹

Delo je prispelo 7. marca 2009, sprejeto 24. septembra 2009.
Received March 7, 2009; accepted September 24, 2009.

Pregled sistemov predmetnih oznak za področje znanosti o živalih v izbranih biotehniških podatkovnih zbirkah

Članek obravnava tri najpomembnejše biotehniške/kmetijske podatkovne zbirke Agricola, Agris, in CAB Abstracts (CABA) glede na predmetne oznake o znanosti o živalih (proizvodnja in varstvo živali, veterina). Podan je obsežen pregled zbirk, indeksiranje in klasifikacijske sheme. Predstavljeni so koncepti, kot so ontologije in metapodatki. Prikazane so predmetne kategorije in razlike med zbirkami pri opisovanju vsebin o živalih, akvakulturi in ribištvu. Različni pojmi se uporabljajo za podobne koncepte. Drevesaste strukture, tezavri, ključne besede oz. deskriptorji so predstavljeni glede na ožje in širše pojme, preferenčne pojme, nedeskriptorje in sorodne pojme oz. gesla. Obstajajo različne hierarhične smeri glede na proizvodne ali taksonomske koncepte. CABA ima najbolj kompleksno hierarhično drevo v smislu taksonomije. V različnih zbirkah se uporabljajo različne relacije med deskriptorji, nedeskriptorji in sorodnimi gesli, zato je kartiranje vsebine odvisno od podatkovne zbirke. Zaslonske slike opisujejo rabo spletnih tezavrov. Prikazane so večjezikovne funkcije tezavra Agrovoc. Portali oz. platforme so obravnavane glede na iskalno sintakso, fraze, boolevo logiko, znake za krajšanje ali maskiranje. Razlike med zbirkami vplivajo na natančnost in odziv oz. priklic ter šum. Za učinkovitejšo rabo bi morali uporabniki pridobiti več znanj o učinkovitih rabah podatkovnih zbirk in informacijskih sistemov.

Ključne besede: predmetne oznake / kartiranje znanosti / deskriptorji / kategorije / klasifikacija / terminologija / ontologije / tezavri / drevesaste strukture / portali / iskalne platforme / sintaksa / podatkovne zbirke / meta podatki / kmetijstvo / biotehnika / znanost o živalih / proizvodnja / varstvo / zootehnika / veterina

Overview of systems of subject headings in the field of animal sciences, production and protection in selected agricultural databases

The article tackles three most important agricultural databases (DB) Agricola, Agris and CAB Abstracts (CABA), produced by FAO, NAL, CABI, with regard to subject headings related to animal sciences, production, protection or health-related veterinary issues. The initial part reviews different approaches to DB and respective indexing and classification schemes. Concepts, such as ontologies and metadata, are presented. Animal, aquatic sciences and fisheries subject categories are shown. Inter-database differences are addressed, e.g. employment of different names for similar concepts. Tree-structures, indexing systems of thesaurus-based keywords-descriptors (DE) are analyzed with emphasis on narrow and broader terms, preferential terms (non-descriptors) and related terms. There exist different tree-structures, depending either on production or taxonomy. CABA exhibits hierarchically the most complex tree with regard to taxonomy. In different DB, keywords are used in relations DE vs. non-DE vs. related terms. Mapping of a concept depends on particular DB. Subject headings are assigned by information specialists, indexers, thus possessing an important degree of subjective choice. Original web-based thesauri screenshots are presented. Emphasis is placed on multilingual functionality of Agrovoc. Portals or search platforms are tackled with regard to retrieval, search syntax, priority, phrases, Boolean logic, wildcards and truncation. Inter-database differences affect retrieval precision, recall, and noise. The complex schemas, subject trees, and headings can sometimes account for a less successful retrieval because they may be too sophisticated and can remain disregarded by users. End-users should acquire better expertise in order to use more effectively the existing information systems and databases.

Key words: subject headings / mapping of science / descriptors / categories / classification / terminology / ontologies / thesauri / tree structures / portals / search platforms / syntax / queries / retrieval / databases / metadata / agriculture / animal sciences / production / protection / zootechny / veterinary sciences

¹ Univ. v Ljubljani, Biotehniška fak., Odd. za agronomijo, Jamnikarjeva 101, SI-1000 Ljubljana, Slovenija, izr.prof., dr., e-mail: tomaz.bartol@bf.uni-lj.si

1 UVOD

Specializirane podatkovne zbirke veljajo v znanosti za najpomembnejši vir referenčnih informacij ne glede na hitri razvoj različnih spletnih iskalnikov, ki prav tako indeksirajo znanstvene informacije. To velja tudi za biotehniko oz. kmetijstvo, kjer se že več kot tri desetletja samostojno razvijajo trije informacijski sistemi in podatkovne zbirke, in sicer Agris, Agricola in CAB Abstracts. Sistemi so bili na začetku razvoja dokaj podobni, kasneje pa so ubrali vsak svojo smer, ki je bila odvisna od usmeritve matičnih organizacij. Vsi ti sistemi še zmeraj nudijo referenčne bibliografske informacije, razvijajo pa tudi številne nove informacijske aplikacije.

Iskanje informacij v teh sistemih temelji na dokaj zapleteno strukturiranih sistemih predmetnih oznak, t.j. predmetnih kategorij in deskriptorjih, urejenih po drevestastih hierarhičnih načelih. Informacijski strokovnjaki, ki so te zbirke gradili, so pred razvojem interneta za končne uporabnike iskali tudi informacije, saj so precej natančno poznali strukturo in posebnosti zbirk. Z razvojem računalnikov in interneta so ti sistemi postali neposredno dostopni posameznim končnim uporabnikom, zlasti raziskovalcem. Končni uporabniki pa pogosto ne poznajo dovolj dobro strukture sistemov predmetnih oznak, saj so ti sistemi dokaj zapleteni in se od zbirke do zbirke precej razlikujejo. V tem prispevku nameravamo izpostaviti nekatere razlike med zbirkami, in sicer na primeru pregleda obstoječih sistemov predmetnih oznak za označevanje pojmov s področja živalske proizvodnje. Poznavanje razlik je pomembno predvsem zato, da bi uporabniki kompetentno prišli do kar najbolj kakovostnih in tehtnih informacij in da bi tako znali sistematično izločiti informacijski šum, torej tiste podatke, ki zanje niso relevantni.

Zbirk ne bomo obravnavali v celoti ali glede na njihove tehnične funkcije, pač pa bomo pregledali in ocenili sisteme za vsebinsko obdelavo. Prispevek sestoji iz začetnega pregleda objavljene literature na temo predmetnih oznak, vsebinske obdelave, tezavrov in značilnosti iskanja informacij v treh osrednjih biotehniških zbirkah ter iz kasnejšega našega pregleda in ocene sistemov predmetnih oznak v teh zbirkah.

2 PODATKOVNE ZBIRKE IN ORODJA ZA VSEBINSKO OBDELAVO

Na začetku za lažje nadaljnje razumevanje povzeto predstavljamo vse tri zbirke in njihova orodja za vsebinsko obdelavo, ki so predmet pregleda literature in kasnejše analize posameznih predmetnih oznak. Citirani viri v tem poglavju se nanašajo predvsem na spletne portale

zbirk in na domače strani ustanov, ki te zbirke gradijo. Prek naslovov URL, navedenih v virih, je možno dostopati do portalov zbirk, do tezavrov in do klasifikacijskih sistemov. Pri aplikacijah zbirke CAB Abstracts je prost javni dostop možen le do osnovnih spletnih informacij, zbirka v celoti, vključno s tezavri, pa je v Sloveniji trenutno dostopna prek portala Biotehniške fakultete, Fakultete za veterino, Fakultete za kmetijstvo v Mariboru, Kmetijskega inštituta Slovenije in prek osebnih gesel raziskovalcev na teh ustanovah.

2.1 AGRICOLA – NAL THESAURUS

Zbirka Agricola nastaja pri ameriški državni kmetijski knjižnici NAL (National Agricultural Library), ki je največja tovrstna knjižnica na svetu (The National, 2009). Za vsebinsko obdelavo in indeksiranje dokumentov uporablja svoj tezaver oz. geslovnik (Agricultural Thesaurus, 2009). Zbirka v zadnjih letih zajema predvsem dokumente iz Severne in Južne Amerike. Dostopna je komercialno prek različnih namenskih iskalnih programov in brezplačno prek portala NAL (Agricola, 2009).

2.2 AGRIS – AGROVOC

Zbirka Agris nastaja pri podorganizaciji Združenih narodov za kmetijstvo FAO (2009). Zbirka se gradi kooperativno, tako da vsaka država članica prek svoje nacionalnega centra samostojno prispeva dokumente po skupni metodologiji zbiranja. Za vsebinsko obdelavo se uporablja tezaver Agrovoc (2009). Zbirka je dostopna komercialno prek namenskih iskalnih programov, predvsem prek programa OVID, brezplačno pa tudi neposredno prek portala FAO (AGRIS/CARIS, 2009).

2.3 CAB ABSTRACTS – CAB THESAURUS

Zbirka CAB Abstracts (2009), v nadaljnjem besedilu CABA, nastaja pri organizaciji CABI (CAB International, 2009). Gradi se predvsem po komercialnih načelih in je dostopna tako prek namenskih iskalnih programov različnih ponudnikov kot prek lastnega komercialnega portala CAB Direct, kjer pa je potrebno prav tako pridobiti plačljivo geslo za dostop. Zbirka CABA velja za najbolj kakovostno biotehniško zbirko na svetu.

3 RAZVOJ STANDARDIZIRANIH KMETIJSKIH BESEDNJAKOV

Orodja za vsebinsko obdelavo oz. tezavri so že kmalu postala predmet samostojnih scientometričnih raziskav. Tezavri kot specializirani strokovni geslovniki predvidevajo precej natančna pravila za določanje ključnih besed dokumentom, ki jih zbirke zajamejo. Tezaver Agrovoc zbirke Agris in tezaver CAB (CAB Thesaurus) zbirke CABA sta nastala najprej. Zbirka Agricola ameriške knjižnice NAL je na začetku uporabljala tezaver CAB, in sicer tudi v dogovoru z nekaterimi drugimi združenji, npr. American Society for Animal Science, kjer pa je bilo potrebno razreševati tudi rabo t.i. amerikanizmov (Thomas, 1985, saj je tezaver CAB temeljil predvsem na britanski rabi). Pri tem so potekali tudi postopki v zvezi s harmonizacijo aplikacij tezavra CAB s tezavrom Agrovoc. Že takrat se je namreč pokazalo, da razlike med tremi globalnimi sistemi za kmetijske informacije povzročajo nepotrebno izgubo časa pri organizaciji in iskanju informacij (Deselaers, 1986). Kljub podobnosti zbirk pa so bile razlike v zajemanju dokumentov dovolj velike, da so avtorji za učinkovito pridobivanje informacij že takrat svetovali rabo vseh treh zbirk (Chen, 1989).

Da bi zmanjšali stroške vzdrževanja tezavrov in bibliografske obdelave je prišlo do pobude, ki naj bi pripomogla k mednarodnemu standardiziranju (angleškega) kmetijskega besednjaka oz. terminologije (Thomas, 1990). To se je še kot posebej potrebno kazalo pri knjižnici NAL, ki je dokumente, sicer indeksirane s tezavrom CAB, pošiljala tudi v zbirko Agris, ki pa je imela nekoliko drugačen sistem indeksiranja (Hood, 1990). Takrat so se pri vseh treh organizacijah, torej CABI, FAO in NAL, začeli resni postopki za poenotenje kmetijskega tezavra oziroma za gradnjo skupnega kmetijskega tezavra UAT (Unified Agricultural Thesaurus) (Andre, 1992), ki pa v praksi kasneje niso zaživel. Kljub načelni pripravljenosti za sodelovanje so vse tri zbirke dolgoročno ohranile vsaka svoj sistem vsebinske obdelave, saj se je pri potencialnem poenotenju pokazalo kar nekaj težko rešljivih tehničnih problemov (Dextre Clarke, 1996). Od sredine 90-tih let dalje o poenotenju ni bilo več dosti govora. Razvoj internetnih oz. spletnih tehnologij pa je še pospešil razlike pri načinih obdelave podatkov pri teh informacijskih službah, ki so se začele razvijati vedno bolj vsaka v svojo smer.

4 VSEBINSKA OBDELAVA IN BIOTEHNIŠKE ZBIRKE KOT PREDMET RAZISKAV

Tri najpomembnejše biotehniške podatkovne zbirke so že zelo dolgo predmet bibliometričnih oziroma sci-

entometričnih raziskav, in sicer tako glede vsebinskega zajemanja posameznih tematik kot tudi značilnosti vsebinske obdelave pri posameznih zbirkah. Začetne raziskave so pogosto obravnavale prekrivanje in zajemanje revij, ki so jih te zbirke indeksirale. Avtorji so npr. ugotavljali slabo zastopanost dokumentov iz manj razvitih delov sveta (Longo in Machado, 1981). Posamezni avtorji so obravnavali tudi prekrivanje specializiranih kmetijskih tematik znotraj teh zbirk, npr. področje pašništva (Farget, 1984). Pri teh tematskih raziskavah se je zbirka CABA že kmalu izkazala kot najbolj kakovostna (Longo in Machado, 1981). Pri tem pa je treba povedati, da se načela zbiranja in indeksiranja dokumentov pri omenjenih zbirkah kar precej razlikujejo, namen služb, ki te zbirke gradijo, pa ni tekmovalnost, pač pa komplementarnost. Oide and Moriwaki (1990) sta primerjala predmetne kategorije in deskriptorje zbirk CABA in Agris na modelu sekcije zbirke CAB Dairy Science Abstracts. Weintraub (1992) je analiziral terminologijo ekološkega kmetijstva (alternative agriculture) in potrdil, da se pojavljajo problemi zaradi razlik med tezavri posameznih zbirk. Področje zelišč sta analizirala Chadwick in Craker (1992) in identificirala CAB kot sicer najkvalitetnejši vir, vendar pa je Agris takrat bolje zajemal hortikulture teme. Element obravnave so bile tudi posamezne sekcije na kompaktnih diskih (CD-ROM), ki so bile v rabi pred nastankom spletnih aplikacij, kot npr. sekciji za veterino (VETCD) in živinorejo (BEASTCD) pri zbirki CABA (Lee in Bredderman, 1993). Posamezni avtorji so tudi opazili, da veliki splošni kmetijski tezavri ne kažejo najboljše potreb po indeksiranju na nekem bolj specifičnem področju. Raupp (1994) se je poskusno lotil tudi gradnje še bolj specializiranega tezavra, ki bi dopolnjeval tezaver AGROVOC na področju ekološkega kmetijstva. Bartol (2001) je tezaver Agrovoc uporabil pri vsebinski analizi nacionalnih publikacij s področja zootehnike, Juvan in sod. (2005a, 2005b) pa so poskusno sestavili geslovnik za področje funkcionalnih živil, ki je prav tako temeljil na dopolnitvah obstoječih biotehniških tezavrov.

5 RAZVOJ NOVEJŠIH SPLETNIH APLIKACIJ

Kot smo že omenili (Dextre Clarke, 1996), od sredine 90-tih let dalje ni bilo več sistematičnih poskusov poenotenja mednarodnega angleškega biotehniškega besednjaka v smeri enotnega strukturiranega tezavra. Z množično rabo zgoščenk in kasneje s prehodom baz na svetovni splet so se začele razvijati nove aplikacije. Pojavile so se različne platforme, portali oz. namenski iskalni programi, kot npr. SPIRS, WinSpirs in nato OVID (Fremer in Larsson, 1997). Nove tehnične možnosti so

omogočile hiter nadaljnji razvoj, kar pa je povzročilo medsebojno tehnično oddaljevanje teh treh informacijskih sistemov, ki so imeli še nekaj let pred tem precej podobno strukturo. Tudi zaradi varčevanja so posamezni ponudniki skrčili obseg dokumentov, ki so jih zajemali oz. indeksirali za vnos v podatkovne zbirke, ali pa so skrčili obseg dokumentov, ki so jih pošiljali v nadaljnjo obdelavo. NAL je pri zbirki *Agricola* začel razvijati nove aplikacije, zbirka pa se je vedno bolj osredotočala zgolj še na ameriške objave in dokumente. *Agricola* je tako prej uporabljala tezaver sistema CAB, okrog 2002 pa je NAL začel razvijati svoj lastni tezaver (Lauser in sod., 2008). V podobnem času so nove aplikacije začeli razvijati tako pri FAO oz. informacijskem sistemu *Agris/Waicent* kot pri CABI. Leta 2002 so na svetovnem srečanju COAIM v Rimu (*Consultations on Agricultural Information Management*) predstavili nove standarde XML za zbiranje podatkov v sistemu *Agris* (*Agris Application*, 2004). Leta 2003 so tudi pri CABI vzpostavili lastno platformo za iskanje informacij, imenovano *CAB Direct* (*Our history*, 2009). Organizaciji FAO in NAL sta omogočili dostop do zbirk *Agris* in *Agricola* po brezplačnih načelih, oziroma prek javno dostopnih spletnih portalov in namenskih iskalnih programov. Portal *CAB Direct* organizacije CABI pa je še naprej plačljiva storitev, kar je razumljivo, saj je organizacija CABI komercialni založnik, storitev pa je temu primerno kakovostna in dokaj celovito zajema svetovne biotehniške informacije, zato CABA šteje kot najbolj referenčna zbirka. Poseben pomen novih programskih aplikacij znotraj sistemov *Agris* oz. *WebAgris* ni več v tem, da bi zbirka še naprej tekmovala z zbirko CABA in želela celovito zajeti globalne biotehniške informacije, temveč se *WebAgris* lahko lokalno uporablja tudi kot sistem za upravljanje z bazami podatkov (DBMS – Database Management System) oz. kot prikladno orodje za gradnjo LISAGR (*Integrated Library System for Agricultural Libraries*), in sicer predvsem v deželah, ki si ne morejo privoščiti komercialnih integriranih knjižnično informacijskih sistemov (Kaloyanova, 2007).

6 PREHOD ELEKTRONSKIH GESLOVNI-KOV V NOVO FAZO

V zadnjih letih je eksplodirala raba pojmov, kot so metapodatki, ontologije ipd. To so izrazi za pojme, ki so jih računalničarji začeli uvajati še ne tako dolgo nazaj, v tradicionalni informatiki oz. dokumentalistiki in knjižničarstvu pa so z drugimi besedami poznani že zelo dolgo (Šauperl, 2005). Ti pojmi se v smislu podatkovnih zbirk, rečeno nekoliko poenostavljeno, v bistvu nanašajo na že dolgo obstoječe bibliografske podatke (metapodatki) in strukturirane tezavre, drevesaste strukture deskriptorjev

in klasičnejše sisteme (ontologije). V smislu sedanje rabe v podatkovnih zbirkah in spletnih informacijskih sistemih pa so izrazi, kot so metapodatki in ontologije, pojmovani bolj kompleksno in se nanašajo na vsebinske koncepte in njihove medsebojne odnose, ki so definirani bolj natančno kot v tezavrih (Beck in Pinto, 2002), zajemajo pa tudi še vrsto drugih aplikacij.

Med obravnavanimi tremi biotehničnimi informacijskimi sistemi je razvoj sistemov za vsebinsko obdelavo in rabe formata XML verjetno šel še najdlje pri FAO, saj so tam močno razširili razvoj aplikacij tezavra *Agrovoc* in vsebinskih drevesastih struktur v smer ontologij in t.i. semantičnega spleta (Gangemi in sod., 2002). Tezaver *Agrovoc* je tako začel postajati samostojni terminološki informacijski sistem in kompleksna podatkovna zbirka, in naj ne bi bil namenjen več zgolj potrebam bibliografske zbirke *Agris* ampak številnim drugim rabam. Na temo raziskav novih semantičnih možnosti na področju biotehnike oz. aplikacij tezavra *Agrovoc* je izšlo kar nekaj prispevkov (Medelyan in Witten, 2005). Predmet raziskav o razmerjih med ontologijami in tezavri je postal tudi tezaver zbirke *Agricola* (Kim and Beck, 2006). Ta tezaver je postal orodje za indeksiranje metapodatkov tudi zunaj knjižnice NAL, npr. na nekaterih ameriških univerzah (Straus in Gardner, 2006). Močno pa prevladujejo prispevki na temo ontologij sistemov FAO oz. tezavra *Agrovoc*, saj je ta tezaver postal pravo globalno orodje za indeksiranje biotehniških dokumentov (Lauser in sod., 2006), katerega možnosti raziskujejo v vedno več deželah. Ta tezaver prevajajo v vedno več jezikov, uporabljajo ga zlasti v Aziji, tudi na Japonskem ter Kitajskem, kjer je že v rabi dvojezični CAT (*Chinese Agricultural Thesaurus*). Liang in Sini (2006) sta raziskala možnosti harmonizacije obeh tezavrov in razvoj tezavra *Agrovoc* v smeri globalnega standardiziranega orodja za izmenjavo biotehniških podatkov. Končana je verzija v japonščini, ki zajema skoraj 40.000 konceptov (Takezake, 2008).

Indeksiranje dokumentov oz. vsebinska obdelava in kartiranje vsebine (mapping) s pomočjo tezavrov in strokovnjakov je relativno drag proces, zato se na primeru tezavra *Agrovoc* odvijajo tudi raziskave, ki primerjajo učinkovitost avtomatiziranih računalniških procesov kartiranja vsebine in učinkovitost tistih procesov, ki jih usmerjajo ljudje (Lauser, 2008). *Agrovoc* preizkušajo za rabo v sistemih ontologij na različnih specializiranih področjih, npr. pri metapodatkih s področja ekološkega kmetijstva (Sanchez Alonso in Sicilia, 2009). Pri najnovejših eksperimentih in analizah je tezaver CAB udeležen v nekoliko omejenem obsegu. Znova je treba omeniti, da je zbirka CABA dostopna le prek naročnine in gesla, kar velja tudi za tezaver te zbirke, ki je prav tako licenčni izdelek z omejeno javno rabo.

7 PREGLED PREDMETNIH OZNAK

Vse tri obravnavane zbirke in večina drugih kvalitetnejših in bolj kompleksno strukturiranih informacijskih sistemov za vsebinsko obdelavo dokumentov oz. objav (in posledično tudi za kasnejše iskanje dokumentov) predvidevajo dve ločeni skupini predmetnih oznak, in sicer *predmetne kategorije* za označevanje hierarhično višjih in bolj splošnih tematik ter *deskriptorje* za bolj natančno oz. bolj specifično označevanje bolj podrobnih pojmov in tematik, ponavadi na hierarhično nižjih ravneh. Te pojme dokumentom priredijo informacijski strokovnjaki, zato izbor zavisi od njihove osebne presoje.

7.1 PREDMETNE KATEGORIJE

Predmetne kategorije so namenjene za opremljanje dokumentov s hierarhično "višjimi" oz. bolj splošnimi oznakami. Njihovo vsebinsko območje je odvisno od informacijskega sistema oz. zbirke. V preglednicah 1, 2 in 3 predstavljamo le tiste kategorije, ki so namenjene označevanju dokumentov z živalsko oz. zootehniško tematiko. Za prihranek prostora v preglednicah smo ponavljajoči-se pojem *animal* povsod okrajšali s črko *a*.

Preglednica 1: Predmetne kategorije zbirke Agris
Table 1: Agris Subject Categories – Category Codes

<i>L – animal production</i>	
L01	a. husbandry
L02	a. feeding
L10	a. genetics and breeding
L20	a. ecology
L40	a. structure
L50	a. physiology and biochemistry
L51	a. physiology – nutrition
L52	a. physiology – growth and development
L53	a. physiology – reproduction
L60	a. taxonomy and geography
L70	veterinary science and hygiene
L72	pests of a–s
L73	a. diseases
L74	miscellaneous a. disorders
<i>M – aquatic sciences and fisheries</i>	
M01	fisheries and aquaculture – general aspects
M11	fisheries production
M12	aquaculture production and management
M40	aquatic ecology

Zbirke, ki jih obravnava ta pregled, zajemajo dokaj podobne vsebinske sklope znotraj kmetijstva oz. biotehnike, zato so kategorije podobne. Same kategorije so v vsaki od zbirk imenovane nekoliko specifično. Opazimo lahko, da je kategorij pri zbirki Agris nekoliko manj (pregl. 1), največ pa jih je pri sistemu CABI (pregl. 2), kjer so tudi bolj podrobno razdeljene na ožje sklope. To pomeni, da lahko pri zbirki CABA iščemo bolj natančno, vendar pa je več možnosti, da se naš koncept nekega pojma ne bo skladal s tistim, ki ga je določil informacijski strokovnjak, zato bo informacijski šum (noise) nekoliko višji, priklic oz. natančnost iskanja (precision) pa nižji. Večje število kategorij torej ne pomeni nujno boljših iskalnih rezultatov.

Če bi raziskovali natančnejše, bi tudi ugotovili, da zbirka Agris pri vsakem dokumentu predvideva opis z največ tremi kategorijami, medtem ko pri zbirki CABA take omejitve ni. Nek dokument oz. bibliografski zapis ima lahko pri CABA pet ali še več kategorij. Zaradi tega je scientometrično težje primerjati zadetke, pridobljene pri eni ali drugi zbirki, saj so že začetna načela klasificiranja nekoliko različna. Očitno CABA kategorije prireja tudi manj osrednjim oz. manj bistvenim konceptom v nekem dokumentu.

Že malo bolj natančna primerjava pokaže, kako pomembna bi bila boljša iskalna usposobljenost uporabnikov teh informacij in njihovo boljše poznavanje nekaterih vsebinskih značilnosti teh informacijskih sistemov. Pri zbirki Agris (pregl. 1) imamo npr. ločeni kategoriji za krmljenje (*Feeding*) in prehrano živali (*Animal Physiology – Nutrition*), kjer prva označuje tehnološke druga pa fiziološke procese. Pri zbirki CABA (pregl. 2) te razlike ni, hkrati pa ta zbirka predvideva kar tri različne fiziološke kategorije prehranjevanja. Zbirka Agris dokumente o anatomskih vidikih označi kot *Animal Structure*, zbirka CABA pa kot *Animal Anatomy and Morphology*. Vidike obnašanja in varstva živali oz. živalskih interakcij z okoljem Agris označi kot *Animal Ecology*, CABA predvideva dve kategoriji, in sicer *Animal Behaviour* in *Animal Welfare*, prav tako Agricola: *Animal Welfare* ter *Animal Ecology and Behavior*.

Pri zbirki Agricola lahko izpostavimo relativno zelo visoko število kategorij, namenjenih označevanju pojmov v povezavi z veterino oz. zdravjem živali (pregl. 3). Teh kategorij je 14 oz. več kot polovica, medtem ko so pri zbirki Agris le štiri. Pri zbirki Agricola imajo vse glavne skupine organizmov, ki povzročajo bolezni, svoje lastne kategorije, pri zbirki CABA pa so te skupine združene, vendar pa so pri CABA izrecno navedeni tudi prioni, torej *Prion*, *Viral*, *Bacterial and Fungal Pathogens of Animals* kot skupna kategorija, ti pa pri zbirki Agricola sploh niso posebej omenjeni in lahko le ugibamo, katera kategorija bi tako zajela tudi prione.

Preglednica 2: Predmetne kategorije CAB International
Table 2: CABI Subject Categories – CABI Codes

<i>LL000 – animal science (general)</i>	
LL010	apiculture
LL020	sericulture
LL030	other invertebr. culture (not aquac.)
LL040	laboratory a. science
LL050	game a–s
LL060	draught a–s
LL070	pets and companion a–s
LL075	sport a–s
LL080	zoo a–s
LL110	dairy a–s
LL120	meat producing a–s
LL130	egg producing a–s
LL145	wool producing a–s
LL148	fur-bearing a–s
LL180	a. husbandry and production
LL190	a. slaughter
LL240	a. genetics and breeding
LL250	a. reproduction and embryology
LL300	a. behaviour
LL400	a. anatomy and morphology
LL500	a. nutrition (general)
LL510	a. nutrition (physiology)
LL520	a. nutrition (production responses)
LL600	a. physiology and biochemistry (excluding nutrition)
LL650	a. immunology
LL700	a. tissue and cell culture
LL800	a. health and hygiene (general)
LL810	a. welfare
LL821	prion, viral, bacterial and fungal pathogens of a–s
LL822	protoz, helminth, mollusc and arthropod paras. of a–s
LL823	veterinary pests, vectors and intermediate hosts
LL860	non-communicable diseases and injuries of a–s
LL884	a. surgery and non-drug therapy
LL886	diagnosis of a. diseases
LL950	toxicology and poisoning of a–s
<i>MM000 – aquatic sciences (general)</i>	
MM110	fisheries
MM120	aquaculture (a–s)
MM130	aquaculture (plants)
MM300	aquatic biology and ecology

Preglednica 3: Predmetne kategorije zbirke Agricola
Table 3: Agricola Subject Categories – Category Codes

<i>L000 – animal science</i>	
L001	entomology related
L002	apiculture related
L003	sericulture related
L100	a. production
L105	a. welfare
L110	laboratory and experimental a–s
L200	a. breeding and genetics
L210	a. reproduction
L300	a. ecology and behavior
L400	a. structure
L500	a. nutrition
L600	a. physiology and biochemistry
L700	a. taxonomy and geography
L800	veterinary science
L810	veterinary pharmacology, toxicology and immunology
L820	Pests of a–s (general)
L821	Pests of a–s (insects and other arthropods)
L822	Pests of a–s (helminths)
L823	Pests of a–s (protozoa)
L830	a. diseases (general)
L831	a. diseases (fungal)
L832	a. diseases (bacterial)
L833	a. diseases (viral)
L840	a. diseases (physiological)
L841	a. disorders and injuries
L850	protection of a. products (general)
L851	protection of a. products (insects and other arthropods)
<i>M000 – aquatic sciences</i>	
M001	aquatic sciences related
M110	fisheries
M120	aquaculture (a–s)
M130	aquaculture (plants)
M210	fisheries management
M220	aquaculture management
M300	aquatic biology and ecology (general)
M310	aquatic biology and ecology (a–s)
M320	aquatic biology and ecology (plants)
M400	oceanography
M500	limnology

Tudi za specifično področje akvatičnih oz. vodnih organizmov predvideva zbirka Agricola največ kategorij. Agricola in CABA tu zajameta tako živali kot rastline, medtem ko pri zbirki Agris ni posebne kategorije za vodne rastline, ampak to zajame kar splošna kategorija *Aquatic Biology and Ecology*.

Že nekaj primerov pokaže, da je sistematske klasifikacijske sisteme zbirk možno učinkovito uporabljati le, če poznamo vsaj osnovna načela metodike klasifikacije v vsaki posamezni zbirki. Če tega ne poznamo dovolj dobro, bomo seveda kljub temu priklicali neke dokumente, ne bomo pa se zavedali, česa vsega sploh nismo dobili prav zaradi naše šibke iskalne strategije.

Pri tem pa ne gre pozabiti še nekaterih drugih zakonitosti vsebinske obdelave na ravni klasifikacijskih kategorij. Poleg specifičnih kategorij, namenjenih označevanju vsebin o živalih, obstajajo še nekatere druge kategorije, ki se na živali nanašajo posredno, a jih ne gre zanemariti. To so npr. kategorije, namenjene označevanju pojmov v zvezi s transportom kmetijskih produktov, kar se nanaša tudi na produkte živalskega izvora ter predvsem kategorije z vsebinskega področja živilstva, torej kategorije, namenjene živilom in prehrani, kjer pa pogosto ni razlike med živili rastlinskega ali živalskega izvora in se vsebine dokumentov tako označijo s splošno kategorijo za živila.

7.2 DESKRIPTORJI

Deskriptorji so namenjeni opremljanju dokumentov s hierarhično "nižjimi" oz. bolj specifičnimi oznakami. Zbrani so v specializiranih geslovnikih, imenovanih tezavri, ki lahko zajemajo več deset tisoč deskriptorjev, urejenih v različnih hierarhičnih in asociativnih relacijah. Tudi tu je njihova raba odvisna od informacijskega sistema oz. zbirke. Za razliko od širših predmetnih kategorij, kjer se razhajanja v poimenovanju nanašajo bolj

na širino vsebinskega obsega posamezne kategorije, pa se pri deskriptorjih pojavljajo tudi terminološka razhajanja, torej tista, ki so med drugim povzročila, do so zamrli poskusi gradnje enotnega svetovnega angleškega tezavra oz. geslovnika UAT (Unified Agricultural Thesaurus). Poleg tega se pri različnih tezavrih oz. zbirkah dokaj razlikujejo tudi hierarhična drevesa in globina oz. višina hierarhij pri drevesastih strukturah (tree structures). V pregledu izpostavljamo nekaj primerov, in sicer posebej hierarhijo v smeri navzgor, torej širše (*Broader Terms*) in navzdol oz. ožje (*Narrower Terms*).

7.3 HIERARHIJA NAVZGOR IN ŠIRŠI POJMI

Za primer vzemimo nek pojem, npr. race oz. *ducks* (pregl. 4). Pri zbirki Agris vidimo, da obstajata dve različni hierarhični skupini, in sicer skupina, ki jo imenujemo "tehnološka", in ki izhaja iz deskriptorja *poultry* ter "taksonomska" skupina, ki se začne z deskriptorjem *Anseriformes*. Na peti stopnji se obe skupini zopet združita. Pri zbirki Agricola je hierarhično drevo nekoliko bolj razvejano. Že na začetku se razdeli na skupini *poultry* ter *waterfowl*. Prva je tehnološka, na četrti stopnji pa se deli še bolj natančno, ko se skupina *livestock* razdeli na dve enakovredni skupini, ki se potem združita, in sicer *domestic animals* ter *farmed animal species*. Pri zbirki CABA se deskriptor *ducks* hierarhično nadgrajuje v eni sami taksonomski skupini, ki pa je zelo podrobna in sega najvišje. Kot zanimivost lahko še omenimo, da se pri zbirki CABA sicer obstoječi deskriptor *livestock* ne uporablja za ptiče in ni del hierarhije za *ducks*, medtem ko se *livestock* uporablja v zvezi z racami tako pri zbirkah Agricola kot Agris. Pri zbirki CABA se npr. izraz *poultry* uporablja le v zvezi s kokošmi. Pri zbirki Agricola se upošteva tudi skupina *monogastric livestock*, pri zbirki Agris pa za skupino *domestic animals* obstaja še nadrejena skupina *useful animals*.

Preglednica 4: Hierarhično širši izrazi glede na deskriptor *ducks* v posameznih zbirkah

Table 4: Hierarchically broader terms with regard to descriptor *ducks* in respective databases

Agricola			Agris/Agrovoc		CAB
Ducks			ducks		ducks
	poultry	waterfowl	Anseriformes	poultry	Anatidae
	monogastric livestock	water birds	birds	livestock	Anseriformes
	livestock	birds	Vertebrates	domestic animals	birds
domestic animals	farmed animal species	vertebrates	Chordata	useful animals	vertebrates
	animals			animals	Chordata
	organisms			organisms	animals
	biological sciences				eukaryotes
					organisms

7.4 HIERARHIJA NAVZDOL IN OŽJI POJMI, SO-RODSTVENE RELACIJE, NEDESKRIPTORJI

Kot pri hierarhični smeri navzgor se glede na deskriptor *ducks* razlike kažejo tudi pri smeri navzdol oz. k ožjim pojmom (pregl. 5). Pri tezavru Agrovoc oz. zbirki Agris ožji deskriptor posebej označuje race *muscovy ducks*, torej vrsto *Cairina moschata*. Tudi pri nedeskriptorjih, označenih z UF (Used For), vidimo, da je deskriptor namenjen za označevanje rodov *Anas* in *Cairina*. Hkrati je deskriptor *ducks* rabljen tudi za pojme označevanje mladih živali (*ducklings*) ter označevanje samcev (*drakes*). Pri sorodnih pojmi, označenih z RT, predvideva Agrovoc pojem *waterfowl* kot poseben pojem, ki ni del hierarhije pri *ducks*. Sorodna skupina so tudi lovni ptiči, *game birds*, ter meso rac, *duck meat*.

Pri zbirki Agricola se kot ožji pojem pojavijo tako *drake* kot *ducklings*, nedeskriptorji (UF) pa se nanašajo le na tehnološke pojme v povezavi s prirajo. Sorodni pojmi so tako jajca kot meso, pa tudi družina *Anatidae*, ki je pri CABA širši deskriptor in kot tak del iste hierarhične skupine.

Pri zbirki CABA ožji deskriptor označuje le mlade živali, pojem za samčke (*drakes*) pa se sicer sploh ne pojavlja v tezavru niti kot nedeskriptor (UF), kot je primer pri tezavru Agrovoc. Tudi pri CABA se nedeskriptorji nanašajo predvsem na tehnološke pojme. Deskriptor *poultry* se tokrat pojavlja kot sorodni pojem, ki se pri drugih dveh zbirkah sicer uvršča med širše pojme (pregl. 4). Edino pri CABA se kot sorodstveni pojem pojavljajo bolezni rac (*duck diseases*), kar pa seveda ne pomeni, da se pri drugih dveh zbirkah bolezni rac ne morejo označiti z deskriptorjem. To je seveda možno storiti že s kombinacijo dveh osnovnih deskriptorjev: *ducks AND diseases*.

Ob vseh teh razlikah se pojavlja vprašanje, ali povprečni uporabnik dovolj dobro pozna specifične konota-

cije in medsebojne povezave teh pojmov v posameznih zbirkah. Kot lahko vidimo, obstaja več enakovrednih logik gradnje tezavrov, ki pa imajo vse večje ali manjše nedoslednosti in druge pomanjkljivosti. Zdi pa se kar nekako nerealno pričakovati, da bi povprečni končni uporabnik natančneje preučili te posebnosti, zato lahko prezapletene strukture pogosto privedejo do slabšega iskalnega odziva v konkretnih iskalnih situacijah. Pri tem lahko omenimo, da pojem *ducks* glede na nekatere druge pojme ni preveč zapleten. Nismo se natančneje poglobili v celotno strukturo tezavrov, temelječo na angleški in hkrati še latinski taksonomiji živali, saj bi to zahtevalo samostojno raziskavo. Vseh deskriptorjev je namreč več deset tisoč, k temu pa lahko dodamo še deset tisoč nedeskriptorjev in iz tega izhajajoče permutacije. Ni tudi odveč komentirati, da se ne zdi več tako nenavadno, da so poskusi gradnje poenotnega globalnega kmetijskega tezavra pred več kot desetletjem zamrli.

7.5 ZASLONSKE SLIKE TEZAVROV

V nadaljevanju prikazujemo nekaj izvernih zaslon-skih slik iz računalniških aplikacij prej omenjenih tezavrov, izhajajočih iz deskriptorja *ducks*. Slika 1 prikazuje le majhen zasloni izsek iz tezavra zbirke CABA, podatke v prejšnjih dveh preglednicah pa smo uredili ročno, na podlagi izbranih možnih hiperpovezav navzgor (BT – Broader Term), navzdol (NT – Narrower Term) ali k sorodnim pojmom. Pri tezavru CAB lahko izberemo več deskriptorjev hkrati in iščemo z vsemi temi gesli naenkrat po načelu booolove iskalne unije OR, kjer zadetke prikliče vsaj eno izbrano deskriptorsko geslo. Če želimo preveriti celotno hierarhijo oz. celotno hierarhično drevo, moramo klikniti na vsak posamezen pojem, ki potem na novih zaslonih prikaže svoje lastne hierarhične

Preglednica 5: Hierarhično ožji (NT), sorodni (RT) ter nedeskriptorski izrazi (UF) glede na deskriptor *ducks* v posameznih zbirkah

Table 5: Hierarchically narrower terms (NT), related terms (RT), and non-descriptors (UF) with regard to descriptor *ducks* in respective databases

	Agricola	Agris/Agrovoc	CAB
	<i>Ducks</i>	<i>Ducks</i>	<i>Ducks</i>
NT (Narrower Terms)	drakes ducklings	muscovy ducks	ducklings
RT (Related Terms)	Anatidae duck eggs duck meat	duck meat game birds waterfowl	duck diseases duck feeding duck meat poultry
UF (Used For)	duck fattening duck feeding duck finishing	Anas Cairina drakes ducklings	duck duck eggs duck fattening



Slika 1: Izsek iz tezavra CAB zbirke CABA – primer deskriptorja ducks s hierarhičnimi povezavami.

Figure 1: Detail from CAB Thesaurus – a case of descriptor ducks with hierarchical relations.

relacije. Možne so torej številne permutacije, manj večji uporabnik pa se lahko hitro “izgubi”.

Kot smo že omenili, lahko do tezavra CAB dostopamo le s pridobljeno licenco za rabo same zbirke. Predstavljena zaslonska slika 1 prikazuje tisto obliko tezavra CAB, kot jo ponuja namenski iskalni program OVID, na katerega je vezana raba zbirke. Isti licenčni iskalni program nudi dostop tudi do tezavra Agrovoc, ki pa je sicer dostopen tudi popolnoma brezplačno in javno prek spleta oz. strežnika FAO, kjer je ta tezaver strukturiran kot prava ontološka podatkovna zbirka. To aplikacijo obravnavajo številni znanstveni prispevki, kar smo predstavili v enem od uvodnih poglavij. Izsek zaslonske slike 2 prikazuje tudi večjezične aplikacije, jeziki pa so označeni z dvočrkovnimi kodami, npr. AR (arabsko), CS (češko), ZH (kitajsko), JA (japonsko). V tezaver se vključuje vedno več jezikov, zajema pa npr. tudi slovaški, madžarski in tajski jezik. Celotna hierarhija, predstavljena v preglednicah 4 in 5, je permutirano dostopna v vseh jezikih tezavra. Gre torej za izjemno veliko in zmogljivo relacijsko zbirko, ki v več kot deset jezikih omogoča dostop do več

deset tisoč biotehniških ali kmetijskih gesel javno prek svetovnega spleta brez kakršnihkoli omejitev.

Brezplačno je dostopen tudi tezaver zbirke Agricola (slika 3), ki je najmlajši med omenjenimi tezavri, saj je Agricola prej uporabljala tezaver CAB. Tudi ta tezaver je hierarhično izjemno široko razvejan, kar smo pokazali že v preglednicah 4 in 5. Pred nekaj leti je tezaver prešel na dvojezično platformo in je zaradi naraščajočega pomena španščine v Ameriki vpeljal tudi ta jezik; zbirka namreč indeksira strokovne in znanstvene dokumente iz Severne in Južne Amerike.

Predmetne kategorije so pri vseh treh zbirkah poimenovane podobno, in sicer *codes* ali *categories*, za deskriptorje pa so imena nekoliko različna in določena še bolj podrobno, in sicer po različnih načelih.

Pri zbirki Agricola obstajajo splošni oz. vsebinski deskriptorji *NAL Subject(s)*. Večina deskriptorjev te zbirke se uvršča v ta razred. Poleg tega obstajajo tudi posebni deskriptorji, namenjeni označevanju geografskih pojmov *NAL Geographic(s)*, vezanih zlasti na pokrajine, države in ozemlja.

Pri zbirki Agris (Agrovoc) obstajajo specifični deskriptorji *Indexer-Assigned Descriptors*, to so tisti, ki jih dokumentom določi informacijski strokovnjak ter hierarhično širši deskriptorji *Computer-Assigned Descriptors*; to so tisti, ki se dokumentom priredijo avtomatsko glede na nadrejene pojme v tezavru. Ti dve skupini se pri javno dostopni verziji te zbirke na portalu FAO imenujeta *AGROVOC Terms* oz. *Other subjects* (slika 4). Če torej strokovnjak dokumentu priredi pojem *ducks*, se temu dokumentu nadrejeno prilepijo (up-posting) vsi širši pojmi, predstavljeni v preglednici 4. Pri tem se računalniško priredijo oz. prevedejo še deskriptorji v francoščini in španščini (slika 4). To sta poleg angleščine dva uradna jezika zbirke Agris.

Zbirka CAB ima nekoliko bolj zapleten sistem. Tu obstajajo *Broad Terms* in *Descriptors*, kjer gre za osnovne ter širše pojme. Poleg tega ima tako kot Agricola CABA



Slika 2: Izsek iz tezavra Agrovoc zbirke Agris – primer deskriptorja ducks s hierarhičnimi in večjezičnimi povezavami.

Figure 2: Detail from Agris database Agrovoc Thesaurus – a case of descriptor ducks with hierarchical and multilingual relations.



Slika 3: Izsek iz tezavra zbirke Agricola – primer deskriptorja ducks s hierarhičnimi povezavami.

Figure 3: Detail from Agricola's Agricultural Thesaurus – a case of descriptor ducks with hierarchical and multilingual relations.

tudi deskriptor za označevanje geografskih pojmov *Geographic Location* ter še poseben deskriptor za t.i. organizme, *Organism Descriptors*, ki so v bistvu latinska rodovna in vrstna imena ter angleška imena za rastline in živali. Ker je v podatkovni zbirki CABA možno iskati usmerjeno tudi s posameznimi polji, je potrebna posebna pazljivost, saj specifično polje vsebuje zgolj deskriptorje ene skupine pojmov.

Poleg teh dveh skupin predmetnih oznak se pri vseh treh zbirkah pojavljajo še "nekontrolirane" ali prosto oblikovane predmetne oznake, imenovane *identifiers* (*other subjects* pri zbirki Agricola). To so tisti pojmi, ki jih ni najti ne med kategorijami ne med deskriptorji. Navadno gre za zelo redko rabljene pojme in jih kontrolirani deskriptorski tezavri ne zajemajo. Če pa raba nekega pojma začne naraščati, bo verjetno prej ali slej postal deskriptor in ga bo tako zajel tudi tezaver. Tu pa ni splošnih pravil. Tezaver CAB je npr. strukturiran zelo podrobno. Marsikateri deskriptor iz CAB pa se pri zbirki Agris pojavlja

zgolj kot identifikator in ga tezaver Agrovoc ne zajema. Standardiziranje nekega pojma določi predvsem daljša raba.

8 ISKALNE SINTAKSE (S PREDMETNIMI OZNAKAMI) / PORTALI / ZBIRKE

Prej predstavljene razlike močno vplivajo tudi na iskanje informacij v teh zbirkah. Končni uporabniki, ki teh posebnosti ne poznajo dovolj dobro, lahko izgubijo informacijo o marsikaterem zanimivem dokumentu. Poleg tega lahko na iskalni uspeh močno vpliva tudi dokaj stroga struktura sintakse, ki jo predvidevajo posamezni portali, platforme oz. namenski iskalni programi. Pri tem je potrebno dobro poznati načela oblikovanja skupin in prioritet pri boolovi sintaksi.

V preglednici 7 predstavljamo primer iskalne zahteve, kjer želi končni uporabnik pridobiti splošne informa-

Preglednica 6: Izvirno poimenovanje ožjih (deskriptorji), širših (predmetne kategorije) in prosto oblikovanih (identifikatorji) predmetnih oznak

Table 6: Original names for narrow (descriptors), broader (Categories), and free (identifiers) subject headings

Zbirka	Deskriptorji/Descriptors	Identifikatorji/Identifiers	Kategorije/Categories
Agricola	NAL Geographic(s)	Other Subject(s)	Subject Codes
	NAL Subject(s)		
Agris	1. Indexer-Assigned Descriptors	Identifiers	Subject Category Codes
	2. AGROVOC Terms		
	1. Computer-Assigned Descriptors		
CAB	2. Other subjects		
	Broad Terms	Identifiers	CabiCodes
	Descriptors		
CAB	Geographic Location		
	Organism Descriptors		

cije bodisi o ovčjem bodisi o kozjem mleku. Ista zbirka je lahko dostopna prek različnih portalov. Vsak portal predvideva posebna pravila glede sintakse. Ta pravila se npr. nanašajo na predpone, torej specifična polja, po katerih želimo iskati, na boolove operatorje in na znake za desno krajšanje ali maskiranje, za kar se lahko uporabljajo različni neabecedni znaki.

Izpostavimo specifikko zbirke Agris na brezplačnem portalu FAO (pregl. 7), kjer se za boolovo unijo predvideva operator OR, obvezno zapisan z velikimi črkami, za boolov presek oz. operator in/and pa se zapiše zgolj presledek, podobno kot bi iskali z Googlom oz. njegovim učenjakom (Google Scholar). Primerjajmo tudi različne znake za desno krajšanje in različne predpone za iskanje po deskriptorskem polju (*hw, descriptor*) pri zbirki CABA (CAB), ki so prav tako odvisne od zahtev portalov OVID ali Direct. Zgornji del preglednice predstavlja iskanje s pomočjo deskriptorjev, spodnji pa hkratno iskanje po vseh vsebinskih poljih, to so naslov (title), izvleček (abstract) ter predmetne oznake, in sicer predvsem deskriptorji. Pri enotavnem iskanju (Basic Search) OVID ne išče dobro brez določila za polje. Uporabiti moramo določilo oz. pripono *mp*, ki zajame vsa prej omenjena "kartirana" (mapped) vsebinska polja, to so naslov, izvleček in predmetne oznake. Zbirke in portali lahko omogočajo več načinov iskanja, kar pomeni tudi to, da se uporabniki poleg vseh potencialnih vsebinskih zagat spopadajo še z različnimi tehničnimi aplikacijami programov za iskanje, poleg tega se platforme nadgrajujejo in spreminjajo.

9 OZNAČEVANJE DOKUMENTOV S PREDMETNIMI OZNAKAMI

Za konec pogledjmo, kako se pri pridobljenih dokumentih prikažejo predmetne oznake v samem računalniškem izpisu podatkov. Oglejmo si primer istega do-

kumenta z naslovom "The effect of a lactic acid bacterial inoculant with enzymes on the fermentation dynamics, intake and digestibility of *Digitaria eriantha* silage", objavljenega v reviji *Animal Feed Science and Technology*. Za ta isti članek prikazujemo primere predmetnih oznak v zbirkah Agris (slika 4), Agricola (slika 5) in CABA (slika 6). Kot že prej tudi tukaj predstavljamo le izseke z zaslon-skih slik.

Opazimo lahko razlike v tehnični izvedbi, pa tudi vsebinske razlike, ki so seveda pomembnejše, saj vplivajo na dejanski priklic podatkov. Iščemo lahko namreč le po tistih geslih, ki so v podatkovni zbirki tudi dejansko prisotna.

Pri zbirki Agris (slika 4) opazimo, da se deskriptorji izpišejo v treh jezikih, med katerimi pa ni optične ločnice. Naj omenimo, da gre za javno dostopno različico zbirke Agris (*Agris New Search Engine*), ki je še nekoliko nova in je še vedno v testni fazi. Ne glede na to pa je dokument s tem naslovom možno na spletu priklicati s katerimkoli od pojmov, ki so zajeti na tej sliki, kar predstavlja precejšnje vsebinsko obogatitev iskalnih možnosti.

Če smo pozorni, lahko pri zbirki Agricola pri tem dokumentu takoj opazimo, da se nikjer ne pojavi objekt *sheep* ali kakšne druge širše oznake za živali, pač pa le specifičen pojem, vezan na ovne oz. osebke moškega spola, torej *rams*. Ta pojem je sicer prisoten tudi pri zbirki Agris, vendar pa je tam isti dokument označen še s številnimi drugimi hierarhično nadrejenimi pojmi v logični povezavi s pojmom *rams*, torej *sheep, ruminants, caprinae, domestic animals* itd. Pri zbirki Agris je tako možno iskati in najti ta dokument z vsemi temi gesli, pri zbirki Agricola pa le z geslom *rams*. Predmetne kategorije (*Subject Code(s)*) so predstavljene le s šiframi, za kar je potrebno imeti šifrant, ki je sicer prosto dostopen na spletu.

Slika 6 kaže različne predmetne oznake zbirke CABA. Lahko bi se pojavljali še celo posebni geografski deskriptorji. Prav tako kot pri zbirki Agris so tudi tu na

Preglednica 7: Iskalna sintaksa za iskanje po deskriptorskih in vseh vsebinskih poljih

Table 7: Search syntaxes for retrieval with descriptors and other mapped terms

Zbirka	Iskanje po deskriptorskih poljih (Tezaver)
Agricola/Ebsco	<i>su</i> ((<i>sheep or goat* or caprine* or ovine*</i>) and <i>milk</i>)
Agricola/NAL	(<i>skey sheep or skey goat? or skey caprine? or skey ovine?</i>) and <i>skey milk</i>)
Agris/FAO	<i>subject</i> ::((<i>sheep OR goat* OR caprine* OR ovine*</i>) <i>milk</i>)
CAB/OVID	((<i>sheep or goat\$ or caprine\$ or ovine\$</i>) and <i>milk</i>). <i>hw</i>
CAB Direct	((<i>sheep or goat* or caprine* or ovine*</i>) and <i>milk</i>): <i>descriptor</i>
	Hkratno iskanje po vseh vsebinskih poljih
Agricola/NAL	("dairy cow?" or "dairy cattle") and (feed? or nutri?)
Agris/FAO	((("dairy cow" OR "dairy cows" OR "dairy cattle") (feed* OR nutri*))
CAB/OVID	((("dairy cow\$" or "dairy cattle") and (feed\$ or nutri\$)). <i>mp</i>

AGRIS Subj. Cat.	Feed additives;Animal physiology - Nutrition
AGROVOC Terms	RAMS, DIGITARIA ERIANTHA, SILAGE, LACTIC ACID BACTERIA, DIGESTIBILITY, FEED INTAKE, RUMEN DIGESTION, BELIER, DIGITARIA ERIANTHA, ENSILAGE (PRODUIT), BACTERIE LACTIQUE, DIGESTIBILITE, PRISE ALIMENTAIRE (ANIMAUX), DIGESTION DU RUMEN, MORUECO, DIGITARIA ERIANTHA, ENSILADO, BACTERIAS ACIDOLACTICAS, DIGESTIBILIDAD, INGESTION DE PIENSOS, DIGESTION RUMINAL
Other subjects	BEHAVIOUR, BOVIDAE, CAPRINAE, DIGESTION, DIGITARIA, DOMESTIC ANIMALS, FEEDING HABITS, FEEDS, FERMENTED PRODUCTS, GRAMINEAE, LIVESTOCK, MAMMALS, MICROBIAL FLORA, PHYSIOLOGICAL FUNCTIONS, PROCESSED PRODUCTS, RUMINANTS, SHEEP, USEFUL ANIMALS, ALIMENT POUR ANIMAUX, ANIMAL DOMESTIQUE, ANIMAL UTILE, BETAIL, BOVIDAE, CAPRINAE, COMPORTEMENT, COMPORTEMENT ALIMENTAIRE, DIGESTION, DIGITARIA, FLORE MICROBIENNE, FONCTION PHYSIOLOGIQUE, GRAMINEAE, MAMMIFERE, OVIN, PRODUIT FERMENTE, PRODUIT TRANSFORME, RUMINANT, ANIMALES DOMESTICOS, ANIMALES UTILES, BOVIDAE, CAPRINAE, COMPORTAMIENTO, DIGESTION, DIGITARIA, FLORA MICROBIANA, FUNCION FISIOLÓGICA, GANADO, GRAMINEAE, HABITOS ALIMENTARIOS, MAMIFEROS, OVINOS, PIENSOS, PRODUCTOS FERMENTADOS, PRODUCTOS PROCESADOS, RUMIANTE

Slika 4: Vse predmetne oznake, deskriptorji in kategorije, v zbirki Agris, portal FAO.

Figure 4: All subject headings, descriptors and categories, in Agris database, search platform FAO.

voljo številni širši in bolj splošni deskriptorji, ki označujejo rod *Ovis*, tu predstavljen kot širši izraz ali Broader Term, s to razliko, da pa pri zbirki Agris ni rodov ampak družine. Tako imenovani organizmi so tu razvrščeni v posebnem deskriptorskem polju. Vidimo lahko, da za te "organizme" veljajo neka specifična interna pravila sistema CAB. Kot organizem se pojavi le geslo *sheep*. Gesla *rams* za samce tu sploh ni, čeprav lahko iz deskriptorskih podatkov zbirk Agricola in Agris sklepamo, da gre za nek eksperiment, kjer so raziskovalci izpostavili prav ovne.

10 SKLEPI

V prispevku smo želeli s pregledom obstoječih sistemov predmetnih oznak o živalski proizvodnji izpostaviti nekatere razlike med zbirkami, ki vplivajo na rezultate pri računalniškem iskanju. Pri tem je pomembno, da se uporabniki zbirk in podobnih informacijskih sistemov teh razlik zavedajo, saj lahko slabo poznavanje negativno

vpliva na iskalni odziv oz. priklic podatkov, ki jih informacijski sistemi zajemajo.

Predmetne oznake, deskriptorje, kategorije oz. gesla dokumentom prirejajo informacijski strokovnjaki, zato je izbor potrjen precej močni subjektivni presoji, ki se ne sklada nujno s pričakovanji in idejami uporabnikov teh informacijskih sistemov. Poleg tega se pojavljajo različne že pri sami tehnični izvedbi iskalnih procesov in sistemski strukturi klasifikacijskih sistemov in tezavrov. Pri predmetnih kategorijah smo tako npr. videli, da Agricola predvideva specifične širše kategorije za vsako različno skupino patogenov, kjer pa nikjer ni prionov. K sreči so ti razvrščeni med deskriptorje. Za podobne pojme, npr. voli, imajo zbirke precej različne angleške deskriptorje, ki se razlikujejo med britansko in ameriško rabo. Pri deskriptorskih predmetnih oznakah zbirke CAB se v razmerah samostojnega uporabniškega iskanja postavlja vprašanje, ali je tako podrobno razdelan sistem sploh potreben. Že samo na pojem govedo se navezuje skoraj trideset različnih ožjih gesel. Poleg tega pa se kategori-

NAL Subject(s): Digitaria eriantha	
rams	Enterococcus faecium
grass silage	Pediococcus acidilactici
fermentation	duration
feed intake	pH
digestibility	protein degradation
lactic acid	hay
enzymes	cellulases
silage making	Subject Code(s): R300
O-glycoside hydrolases	L500
silage additives	L600
Lactobacillus plantarum	

Slika 5: Vse predmetne oznake, deskriptorji in kategorije, v zbirki Agricola, portal NAL.

Figure 5: All subject headings, descriptors and categories, in Agricola database, search platform NAL.

CABICODES	Feed Storage and Preservation [RR110]. Feed Composition and Quality [RR300]. Animal Nutrition (General) [LL500].
Subject Headings	cellulase . composition . digestibility . feeds . fermentation . hay . intake . lactic acid . lactic acid bacteria . preservation . processing . silage . silage additives . silage fermentation . silage making . silage quality .
Organism Descriptors	Digitaria . Digitaria eriantha . Enterococcus faecium . Lactobacillus . sheep .
Broad Terms	Ovis . Bovidae . ruminants . Artiodactyla . mammals . vertebrates . Chordata . animals . ungulates . eukaryotes . Poaceae . Cyperales . monocotyledons . angiosperms . Spermatophyta . plants . Digitaria . Enterococcus . Enterococcaceae . Firmicutes . bacteria . prokaryotes . Lactobacillaceae .
Identifiers	ensilage . ensiling . lactate .

Slika 6: Vse predmetne oznake, deskriptorji in kategorije, v zbirki CAB Abstracts, portal OVID.

Figure 6: All subject headings, descriptors and categories, in CAB Abstracts, search platform OVID.

je navzgor med zbirkami močno razlikujejo, tako da bi uporabnik v povezavi z nekim geslom oz. pojmom dejansko moral natančno terminološko preučiti tudi vse povezane pojme. Pri tem se zdi, da so te zbirke oz. njihovi sistemi za predmetne oznake morda kar preveč ambiciozni. Včasih bi bilo bolj uporabno, da bi bili klasifikacijski koncepti malo širši oziroma bolj preprosti.

Naš namen ni bil, da bi zbirke vrednotili po kakovosti. Želeli smo predvsem opozoriti na razlike, ki jih je potrebno upoštevati, če želimo iskati s kar najboljšim iskalnim odzivom. Že dalj časa med strokovnjaki ni dvoma, da velja zbirka CABA za najpomembnejšo biotehniško zbirko, ki se najbolj osredotoča na kakovost pri izboru informacij in dokumentov. Za raznovrstne scientometrične ali tehniške eksperimente pa je ta zbirka nekoliko manj dostopna, saj je raba vezana na naročnino in geslo, uporabniki pa pričakujejo nek stalni sistem brez pretiranih menjav iskalne platforme. Eksperimenti so lažje izvedljivi pri brezplačnih javno dostopnih portalih, kot je npr. zbirka Agris, saj nastaja kot kooperativna kompilacija izbranih nacionalnih biotehniških bibliografij dežel članic FAO oz. OZN. Plačljiv dostop, npr. prek portala OVID, omogoča več iskalnih možnosti, ki so tudi bolj natančne, pri tem pa se od uporabnikov pričakuje tudi nekoliko višja raven informacijske pismenosti. Svoje posebnosti pa imajo tudi brezplačni portali, tako da se pogosto pojavljajo celo nasprotna iskalna načela pri izvedbi iskalnih korakov, kar uporabnike pogosto zmede.

Pravzaprav v zadnjih letih primerjave in poenotenja med temi zbirkami niso več smiselna, saj je šel razvoj sistemov v različne smeri. Zbirka Agricola knjižnice NAL je hkrati tudi OPAC katalog (Online Public Access Catalog) te ustanove. Z uvedbo španščine je zbirka dejansko postala zbirka obeh Amerik. Ker je zbirka izdelek predvsem ene ustanove, to je knjižnice NAL, so možnosti za oblikovanje in usmerjanje razvoja precej večje. Zbirki Agris je v

zadnjih letih nekoliko pošla sapa, vsaj kar se tiče celokupnega letnega števila novih dokumentov, vendar pa je ta zbirka, ali bolje, informacijski sistem, v zadnjih letih prešla na popolnoma novo kakovostno raven obdelave dokumentov po najnovejših načelih zbiranja metapodatkov in aplikacij XML. To še zlasti velja za tezaver Agrovoc, ki je postal večjezikovna relacijska zbirka in pravo javno dostopno globalno orodje za indeksiranje z deset tisoči vzporednih gesel v različnih jezikih. Pri tem pa moramo omeniti, da se morajo drugojezični prevodi držati hierarhije tega tezavra, ki pa ni absolutna glede na obstoječe razlike med tezavri, ki smo jih izpostavili.

Ob iskalnem obnašanju povprečnega uporabnika se je težko znebiti občutka, da uporabniki pogosto iščejo brez prave predstave o tem, kaj je pri nekem informacijskem sistemu sploh možno pričakovati, še posebej upoštevajoč precejšnje razlike med sistemi. Zelo verjetno je, da povprečni uporabnik ni dovolj usposobljen, da bi do informacij prišel dovolj sistematično in celovito. Uporabniki pogosto iščejo naključno, brez poglobljene analize problematike glede na možne taksonomske oznake, pojme in gesla. Tako uporabniki sicer vedno pridobijo določene dokumente, vprašanje pa je, česa ne dobijo, ne da bi se tega zavedali. Morda pa so, kot smo omenili že prej, ti sistemi uporabniško nekoliko preveč ambiciozni. Ta pregled je zato namenjen tudi temu, da bi pri uporabnikih vzbudil potrebo, da bi bolje raziskali vsebinske sisteme in različna možna terminološka drevesa na področju, ki ga raziskujejo in pri svojem iskanju bolj upoštevali posebnosti informacijskih sistemov in zbir. Tako bodo pridobili ustrežnejše publikacije in bo manj možnosti, da bi pri uporabi virov spregledali kak resnično pomemben dokument, ki bi lahko odločilno vplival na potek neke raziskave.

11 SUMMARY

The article tackles three most important agricultural databases Agris, Agricola and CAB Abstracts (CABA), maintained respectively by National Agricultural Library (NAL), FAO and CABI (CAB International), with regard to subject headings related to animal sciences, production, protection or health-related veterinary issues. The initial part reviews different approaches to these databases and respective indexing and classification schemes. Several concepts are presented, such as ontologies and metadata. Animal science and production and aquatic science and fisheries subject headings are presented, such as Agris Category Codes, Agricola Subject Categories, and CABI Codes. Differences among databases with regard to categories are addressed, e.g.: Agris employs two concepts related to animal nutrition: Feeding as a technological category and Animal Nutrition category as a physiological category. CABA employs three separate physiological categories. Agris uses the name Animal Structure, whereas this concept is expressed as Animal Anatomy and Morphology in CABA. Animal environmental interactions and well-being are addressed as Animal Ecology in Agris. Agricola and CABA define two separate categories, related either to welfare or ecology/behavior. Agricola dedicates the most important share to animal-health and veterinary-related issues. Fisheries and Aquatic Sciences-related categories separately classify also plants in Agricola and CABA. There is no such distinction in Agris. The tree-structures or indexing systems of thesaurus-based keywords, or descriptors, are presented, with special emphasis on narrow and broader terms, preferential terms or non-descriptors and related terms. There exist different classification-tree branching directions of broader descriptors, related to production or taxonomy. CABA exhibits hierarchically the most complex tree with regard to taxonomy. Differences among indexing systems are sometimes important. In different databases, different descriptors or keywords are used in relations descriptors vs. non-descriptors vs. related terms, so there is frequently no internationally standardized single descriptor to map a concept. Descriptors can take up different role relative to a database. Different screenshots are presented to illustrate the original uses of thesauri. Emphasis is placed on multilingual functionality of Agrovoc. All subject heading names in different databases are presented, such as Descriptors, Subjects, Identifiers, Subject Codes or Category Codes. Access privileges and password-based licenses are addressed. Portals or search platforms are also tackled with regard to retrieval, search syntax, priority, phrase-searching, Boolean logic, wildcards and truncation. Subject headings are assigned by information specialists, indexers,

thus possessing an important degree of subjective choice. On top of that, the above inter-database differences strongly affect retrieval precision, recall, and noise. The existing classification schemas, subject headings, thesauri, subject trees, and categories contribute to a more precise indexing. They can, however, sometimes account for a less successful retrieval owing to a structure that is maybe too sophisticated. The improvement for search precision and better focus possibilities can thus remain disregarded by users. The end-users should thus acquire better awareness and acquaintance with the principles of subject headings in order to use more effectively the existing information systems and databases - what is an objective of this review article.

12 VIRI

- Agricola Database - NAL Catalog. 2009.
<http://agricola.nal.usda.gov/> (29. dec. 2008)
- Agricultural Thesaurus and Glossary Home Page. 2009.
<http://agclass.nal.usda.gov/> (7. jan. 2009)
- AGRIS Application Profile: Methodology. 2004. FAO, Rome.
ftp://ftp.fao.org/gi/gil/gilws/aims/publications/papers/2004_7.pdf (3. feb. 2009)
- Agris Subject Categories. 2009.
<http://www.fao.org/scripts/agris/c-categ.htm> (7. jan. 2009)
- AGRIS/CARIS Information Centre Homepage. 2009.
<http://www.fao.org/Agris/> (29. dec. 2008)
- Agrovoc. 2009.
http://www.fao.org/aims/ag_intro.htm (7. jan. 2009)
- Andre P. 1992. Toward a unified agricultural Thesaurus. Quarterly Bulletin of IAALD, 37, 4: 224–226
- Bartol T. 2001. Mapping of contents in "Zbornik Biotehniške fakultete Univerze v Ljubljani. Kmetijstvo. Zootehnika": an analysis of narrower and broader subject areas. Research Reports Biotechnical Faculty University of Ljubljana. Agriculture. Zootechny, 78, 2: 85–103
- Beck H., Pinto H.S. 2002. Overview of Approach, Methodologies, Standards, and Tools for Ontologies. V: Third Agricultural Ontology Service (AOS) Workshop, Gainesville, USA, 09–10 maj 2002. University of Florida: 58 str.
- CAB Abstracts. 2009.
<http://www.cabi.org/datapage.asp?iDocID=165> (7. jan. 2009)
- CAB International. 2009.
<http://www.cabi.org/home.asp> (7. jan. 2009)
- CABICODE Alphabetical List. 2009.
<http://www.cabi.org/guides/cabicodes.htm> (7. jan. 2009)
- Chadwick A.F., Craker L.E. 1992. Herbal literature in selected databases. Acta Horticulturae, 306: 340–345
- Chang Chun, Lu Wenlin. 2002. The Translation of Agricultural Multilingual Thesaurus. V: Proceedings of the Third Asian Conference for Information Technology in Agriculture, Beijing, China, 26–28 okt. 2002. Chinese Academy of Agricultural Sciences: 526–528
- Chen Q.B. 1989. A check on overlapping between AGRICOLA,

- AGRIS and CAB for tropical agricultural records. Quarterly Bulletin of IAALD, 34, 2: 67–72
- Deselaers N. 1986. The necessity for closer cooperation among secondary agricultural information services: an analysis of AGRICOLA, AGRIS and CAB. Quarterly Bulletin of IAALD, 31, 1: 19–26
- Dextre Clarke S.G. Integrating Thesauri in the Agricultural Sciences. Compatibility and integration of order systems. Research Seminar Proceedings of the TIP/ISKO Meeting, Warsaw, Poland, 13–15 sep. 1996. 111–122
- FAO. 2009. Food and Agriculture Organization of the United Nations. <http://www.fao.org/> (3. feb. 2009)
- Farget M.A. 1984. A performance test on various bibliographic databases relevant to agricultural economics. *Economie Rurale*, 160: 37–39
- Fremer E., Larsson B. 1997. SPIRS, WinSPIRS, and OVID: a question of free-text versus Thesaurus retrieval? Bulletin of the Medical Library Association, 84, 1: 63–70
- Gangemi A., Fisseha F., Pettman I., Keizer J. 2002. Building an Integrated Formal Ontology for Semantic Interoperability in the Fishery Domain. V: *Agricultural Information and Knowledge Management Papers*. Rome, Knowledge and Communication Department, FAO: 13 str. <ftp://ftp.fao.org/docrep/fao/008/af242e/af242e00.pdf> (2002).
- Hood M.W. 1990. Reconciling the CAB Thesaurus and AGROVOC. Quarterly Bulletin of IAALD, 35, 4: 181–185
- Juvan S., Bartol T., Boh B. 2005. Design and development of a relational database for functional foods. *Acta agriculturae Slovenica*, 86, 1: 3–15
- Kaloyanova S., Onyancha I., Salokhe G., Ward F.L.H., Keizer J. 2007. Information technologies and standards for agricultural information resources management: AGRIS application profile, AGROVOC and LISAGR. Quarterly Bulletin of IAALD, 52, 1–2: 17–21
- Kim S.H., Beck H.W. 2006. A practical comparison between Thesaurus and ontology techniques as a basis for search improvement. *Journal of Agricultural & Food Information*, 7, 4: 23–42
- Lauser B., Sini M. From AGROVOC to the agricultural ontology service/concept server: an OWL model for creating ontologies in the agricultural domain. V: *Proceedings of the 2006 international Conference on Dublin Core and Metadata Applications: Metadata For Knowledge and Learning*, Manzanillo, Mexico, 03–06 okt. 2006. 76–88
- Lauser B., Sini M., Salokhe G., Keizer J., Katz S. 2006. AGROVOC web services: improved, real-time access to an agricultural Thesaurus. Quarterly Bulletin of IAALD, 51, 2: 79–81
- Lauser B., Johannsen G., Caracciolo C., van Hage W.R., Keizer J., Mayr P. Comparing human and automatic Thesaurus mapping approaches in the agricultural domain. *Papers and Project Reports for International Conference on Dublin Core and Metadata Applications*, Berlin, Germany, 22–26 sep. 2008. 43–53
- Lebowitz A., Portegies-Zwart R., Schmid H. 1991. Multilingual indexing and retrieval in bibliographic systems: the AGRIS experience. Quarterly Bulletin of IAALD, 36, 3: 187–192
- Lebowitz A.I., Portegies-Zwart R.P., Schmid H. 1991. Retrieval in bibliographic systems: the AGRIS experience. Quarterly Bulletin of IAALD, 36, 3: 187–192
- Lee T.P., Bredderman P.J. The literature of veterinary and animal sciences on the CAB ABSTRACTS databases: a description and evaluative appraisal of CAB International's VETCD and BEASTCD CD-ROM products. V: *Animal health information: planning for the 21st century*. The First International Conference of Animal Health Information Specialists, Reading, United Kingdom, 16–19 jul. 1993. Reading, University of Reading: 141–177
- Liang A.C., Sini M. 2006. Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, tools, procedures. *New Review of Hypermedia and Multimedia*, 12, 1: 51–62
- Longo R.M.J., Machado U.D. 1981. Characterization of databases in the agricultural sciences. *Journal of the American Society for Information Science*, 32, 2: 83–91
- Medelyan O., Witten I.H. Thesaurus-Based Index Term Extraction for Agricultural Documents. V: *Proceedings of the 6th Agricultural Ontology Service (AOS) workshop at EFITA/WCCA*, Vila Real, Portugal, 25–28 jul. 2005. 1–8
- National Agricultural Library: NAL Catalog (AGRICOLA) - Category Codes. 2009. <http://agricola.nal.usda.gov/help/categorycodes.html> (7. jan. 2009).
- Oide N., Moriwaki N. 1990. Comparisons of indexing words used in CAB ABSTRACTS and AGRIS. *Bulletin of the Japanese Association of Agricultural Librarians and Documentalists*, 78: 1–10
- Our history. 2009. CAB International. <http://www.cabi.org/datapage.asp?iDocID=235> (3. feb. 2009)
- Raupp J. 1994. Concept for a standard Thesaurus in the field of organic farming. *Nachrichten fur Dokumentation*, 45, 6: 343–348
- Sanchez-Alonso S., Sicilia M.A. 2009. Using an AGROVOC-based ontology for the description of learning resources on organic agriculture. V: *Metadata and Semantics*. Sicilia M.A., Lytras, M.D. (eds.). US, Springer: 481–492
- Straus C., Gardner M. Implementing the NAL Thesaurus in a Knowledge Repository. *USAIN Conference*, Ithaca, N.Y., 08–11 okt. 2006. 1–13
- Šaupperl A. 2005. Knjižnica. Izobraževanje za katalogizacijo in organizacijo informacij, 49, 3: 95–111
- Takezake A. Development of Agricultural Vocabularies for Efficient Searching. V: *The APAN 25th Meeting*, Bangkok, Thailand, 23–25 feb. 2008. 24 str.
- The National Agricultural Library (NAL). 2009. <http://www.nal.usda.gov/> (29. dec. 2008)
- Thomas S.E. 1990. Bibliographic control and agriculture. *Library Trends*, 38, 3: 542–561
- Thomas S.E. 1985. Use of the CAB thesauri at the National Agricultural Library. Quarterly Bulletin of IAALD, 30, 3: 61–65
- Weintraub I. 1992. The terminology of alternative agriculture searching AGRICOLA, CAB and AGRIS. Quarterly Bulletin of IAALD, 37, 4: 209–213