

MALO ZNANE LASTNOSTI KLASIFIKACIJE KRATEK VPOGLED V TEORIJU KLASIFIKACIJE IN NJENO UPORABO V KNJIŽNIČNIH KATALOGIH

Alenka Šauperl *

UDK: 025.4

Alenka Šauperl: Malo znane lastnosti klasifikacije - Kratak vpogled v teorijo klasifikacije in njeno uporabo v knjižničnih katalogih. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik referatov z dopolnilnega izobraževanja, Maribor 6/2007, str. 83-95.

Izvirnik v slovenščini, izvleček v slovenščini in angleščini, povzetek v angleščini.

Klasična teorija klasifikacije ni dovolj, da bi razumeli in v prid informacijskih sistemov uporabili vse možnosti, ki jih klasifikacija dejansko ponuja. Z vključitvijo novejših teorij, kot je teorija prototipov, in dobrim izkoristkom sodobne informacijske tehnologije lahko klasifikacijski sistemi pripomorejo k obvladovanju velikih količin podatkov, kot se pojavljajo v knjižnicah, arhivih ali na internetu.

UDC: 025.4

Alenka Šauperl: Less Known Characteristics of Classification - A Brief Overview of the Theory of Classification and its Use in Library Catalogues. Technical and Field Related Problems of Traditional and Electronic Archiving. Conference Proceedings, Maribor 6/2007, pp. 83-95.

Original in Slovenian, abstract in Slovenian and English, summary in English.

Classical theory of classification does not offer a sufficient basis to understand the process of classification and to use it to enhance information systems. It would be possible to employ classification for managing large quantity of data (as in libraries, archives or the Internet) by applying modern theories, e.g. theory of prototypes, and with an efficient use of information technology.

UVOD

Klasifikacijski sistemi so bili v 18 in 19. stoletju razmeroma popularni. Ljubiteljski in poklicni znanstveniki so si izmišljali klasifikacijske sisteme za različne namene (Bowler, 1993). Med najpomembnejšimi dosežki tistega časa so prav gotovo botanične in zoološke taksonomije in knjižnični klasifikacijski sistemi. Danes, v 21. stoletju in digitalni dobi, botaniki in zoologi še vedno uporabljajo svoje taksonomije. Bibliotekarji pa smo začeli dvomiti v njihovo uporabnost. Prepričevati so nas začeli predvsem računalniki s svojo vedno večjo kapaciteto in »računalniški guruji«, ki so popolnoma prepričani v njihove zmogljivosti. Toda klasifikacijski sistemi niso le ostanek iz preteklosti predvsem zaradi svoje naravne povezanosti s človekovim mišljenjem. Ker na to velikokrat pozabljamo, želim s svojim prispevkom spomniti na prednosti, ki nam jih klasifikacija nudi pri organiziranju velikih količin podatkov.

* *Izr. prof. dr. Alenka Šauperl, Oddelek za bibliotekarstvo, informacijsko znanost in knjigarstvo Filozofske fakultete Univerze v Ljubljani, Aškerčeva 2, 1000 Ljubljana, Slovenija.*

KLASIFICIRANJE JE V ČLOVEŠKI NARAVI

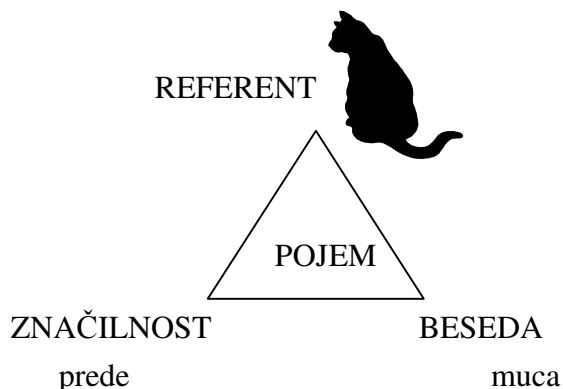
Vse naše življenje je preprejeno z različnimi klasifikacijami. Že kot otroci razlikujemo domače in tuje osebe, nevarne predmete in igrače. Zgodaj začnemo klasificirati pomembne in nepomembne dogodke, zvoke ali objekte v naši okolici. Od tega smo življenjsko odvisni, saj ne moremo nameniti enake pozornosti vsemu, kar se okrog nas dogaja. Na ta način ekonomično uporabljamo svoje sposobnosti. Pozorni smo le na pomembno in temu namenimo pozornost, ki je za naše delo ali preživetje pomembna.

Pravijo, da lahko človek hkrati pomni približno sedem nepovezanih konceptov. Ker je to število razmeroma majhno, smo se naučili koncepte grupirati in na ta način dosegli, da si jih lahko zapomnimo več - zapomnimo si skupine namesto posameznikov. Zanimivo pa je, kako koncepte grupiramo - kako kategoriziramo, oz. kaj so naši klasifikacijski kriteriji. Psiholingvisti, ki se ukvarjajo z raziskovanjem teh kriterijev in njihovim vplivom na jezik, premišljevanje in vedenje, so prišli do zanimivih ugotovitev. Lakoff (1987), npr. opisuje kategorizacijo, ki se pojavlja v jeziku ene od skupin avstralskih domorodcev (Dyrbal). Vsi samostalniki so razvrščeni v štiri skupine. Kadarkoli govorec uporabi nek samostalnik, mora pred njim povedati ime kategorije, nekako tako, kot se v nemščini uporablja spolnik (die, der, das) ali angleščini določni ali nedoločni člen (the, a). Štiri kategorije pa so naslednje:

- »Bayi«: vanjo sodijo moški, kenguruji, oposumi, netopirji, večina kač, žuželk in rib, nekatere ptice, luna, nevihte, mavrice, bumerangi, nekatere sulice itd.;
- »Balan«: v katero sodijo ženske, psi, kljunači, kljunati ježki, nekatere kače in ribe, večina ptic, kresnice, škorpijoni, črički, vse, kar je povezano z vodo ali ognjem, sonce in zvezde, ščiti, nekatere puščice, nekatera drevesa in podobno;
- »Balam«: kar so vsi užitni sadeži in rastline, ki jih obrodijo, gomolji, praproti, med, cigarete, vino in pecivo;
- »Bala«: kamor sodijo deli telesa, meso, čebele, veter, nekatere puščice, večina dreves, trava, blato, kamni, hrup in jezik itd. (Lakoff, 1987, str.92-93).

Vsak pripadnik zahodne civilizacije se bo vprašal, kaj imajo ženske skupnega s kljunatimi ježki, kresnicami, soncem, ščitom in drevesi. Marsikateri moški, ki je hud na svojo partnerko, si bo mislil, da imajo precej skupnega s psi, kačami in škorpijoni. Tisti, ki je zaljubljen, jo bo morda primerjal s čričkom in soncem. V glavnem pa nas bodo skupine presenetile, kot so presenetile tudi raziskovalce. Ponovno so namreč ugotovili, da vse lastnosti, ki odločajo o razvrstitvi posamezne enote v kategorijo, niso očitne.

Filozofi, psiholingvisti in informacijski strokovnjaki že nekaj časa vedo, da je vsak pojem triada, sestavljena iz referenta, značilnosti in besede (Dahlberg, 1989; Mai, 2001; Peirce, 1991). Referent je stvar (*object*), ki jo poimenujemo z neko besedo (*representamen*), npr. muca (glej sliko 1). Beseda sama pa sploh ne nosi primernih informacij, da bi referenta lahko kategorizirali. Odločilne so referentove značilnosti (*interpretant*). Kaj izberemo za referentove značilnosti, ki odločajo o kategorizaciji, pa je odvisno od nas samih ali, pogosteje, od našega okolja, družbe.



Slika 1: Pojmovni trikotnik

Jeziki in narodi se zelo razlikujejo po kategorizacijskih kriterijih, kar smo videli že v primeru aboridžinske kategorizacije. Razlike pa nastopajo tudi v manjših skupinah, v družini, v podjetju oz. organizaciji, na strokovnem področju. Razlike tudi ne nosijo enakih posledic. Od posameznika je odvisno, kako si bo uredil in organiziral zapiske predavanj, saj so namenjeni le njemu. Uspeh mestnega projekta recikliranja odpadkov pa je vsaj delno odvisen od tega, ali meščani enako razumejo delitev odpadkov na organske odpadke, steklo, papir in druge. Ali je prava lepotica mis Slovenije, Iris Mulej, ali mis Fitness, Brina Romih, je najbrž manj usodno kategorizacijsko vprašanje kot, kdo so domačini in kdo tujci v naši vasi (npr. Ambrus ob božiču 2006). O usodnosti klasifikacije ljudi sta več tragičnih zgodb razkrila Bowker in Star (1999).

NAČINI KLASIFICIRANJA

Da smo presenečeni nad zgornjimi štirimi avstralskimi kategorijami, je torej predvsem krivo nepoznavanje kategorizacijskega načela. Razlog pa je lahko tudi v načinu kategoriziranja. Kot izobraženi ljudje, ki imamo redno opravka z različnimi klasifikacijami, smo vajeni razvrščanja v hierarhične skupine, ki jih družijo skupne lastnosti, torej skladno s klasično teorijo kategorizacije. Ta paradigma prevladuje že od Aristotela dalje in sloni na treh predpostavkah:

- Značilnosti kategorije so vsota značilnosti pojmov, ki sodijo v to kategorijo. To pomeni, da imajo vsi člani razreda določene skupne značilnosti, pa tudi nekaj posebnosti. Npr.: Trikotniki so mnogokotniki, ki so omejeni s tremi stranicami in imajo tri oglišča. Te lastnosti veljajo za vse trikotnike ne glede na to, kakšne oblike so (enakokraki, enakostranični, pravokotni).
- Značilnosti neke kategorije so nujne in zadostne za definiranje te kategorije. Noben drug mnogokotnik, predmet ali pojem nima enakih lastnosti kot trikotnik. Hkrati je v definiciji zajeto vse, kar je za prepoznavanje trikotnika potrebno.
- Značilnosti se “dedujejo” tako, da ima vsaka podkategorija vse značilnosti nadrejene in še nekaj svojih. V našem primeru imajo vsi mnogokotniki stranice

in kote. To so značilnosti, ki se prenašajo z nadrejene kategorije na njene podkategorije, tudi na trikotnik. Trikotnikova posebna značilnost v razredu mnogokotnikov je samo število stranic in kotov (Iyer, 1995; Phares, 1985).

Po klasični teoriji klasifikacije imajo razredi točno določene, jasne in med seboj različne lastnosti. Meje med razredi so torej točno določene in vsak element je lahko član le enega razreda.

Da pa ljudje vedno ne uporabljamo dosledno tega načina klasifikacije, so sumili že filozofi, preden so psihologi v 20. stoletju to tudi dokazali. Filozof Ludwig Wittgenstein (1981) je na svojem slavnem primeru iger pokazal družinsko podobnost kot enega od načinov klasifikacije, ki ga ljudje redno in nezavedno uporabljamo. Med igre spadajo tiste z žogo kot tiste, ki ne potrebujejo nobenega predmeta (npr. mora, ital. »morra«). Med igre z žogo pa navadno ne štejemo tenisa, ki je igra, pri kateri uporabljamo žogo. Med igre spadajo olimpijske in otroške, tiste, ki temeljijo na razumu (šah) in tiste, ki z razumom nimajo veliko (ruleta), tiste, ki jih igramo sami (npr. pasijansa ali solitaire) in tiste, ki jih igramo v družbi kot posamezniki (človek, ne jezi se) ali skupinske (nogomet). Družinska podobnost se od klasične teorije kategorizacije razlikuje v več dimenzijah:

- Značilnosti kategorije so sicer vsota značilnosti pojmov, ki sodijo v to kategorijo, vendar ni nujno, da imajo vsi člani kategorije tudi iste značilnosti. V naši družini smo si ženske zelo podobne, vendar nimamo vse enakega nosa ali barve las. Z mamo imava podobna usta in zobe, s sestro glas, s hčerjo obliko obraza in z nečakinjo kakovost las. Kljub mnogim razlikam, vključno z veliko razliko v starosti, nas vsi, tudi neznanci, brez težav razvrstijo v isto družino.
- Značilnosti neke kategorije posledično niso nujne in zadostne za definiranje te kategorije.
- Vse značilnosti, ki definirajo nadrejeno kategorijo, se pri klasični teoriji prenašajo na podrejeno kategorijo. Pri družinski podobnosti teh skupnih značilnosti ni, zato se ne morejo dedovati.
- Po klasični teoriji klasifikacije imajo razredi točno določene in med seboj različne lastnosti, torej jasne meje. Pri družinski podobnosti meje niso nujno vedno jasne, zato lahko nek pojem sodi v en ali drugi razred. Nogomet in poker sta lahko prijetni družabni igri, lahko pa tudi zelo resna posla.

Toda družinska podobnost ni edini način kategorizacije, ki ga ljudje uporabljamo, in se ne ujema s klasično teorijo klasifikacije. Psihologinja Eleanor Rosch (1973) je z eksperimenti ugotovila, da nas pri kategorizaciji vodi princip prototipa. Naš prototip ptice ima kljun in perje. Četudi neke ptice še nikoli nismo videli, jo bomo uvrstili med ptice zato, ker se podoba nove živali ujema z lastnostmi našega prototipa. Tako kategorizacijo najlažje in redno opazimo pri majhnih otrocih, ki spoznavajo svet in se učijo govoriti. Čeprav ga pojma ptice naučimo na primeru domačega kanarčka, bo kokoš, čeprav jo bo prvič videl, uvrstil med ptice. Tudi z rožami je podobno. Naš prototip rože se najbrž ujema z marjetico. Manj verjetno pa je, da se ujema z ajdo. E. Rosch je ugotovila, da so nekateri pripadniki določenega razreda bolj, drugi manj podobni prototipu. Zato so razredi urejeni od sredine proti robu v bolj ali manj gosti gradaciji. Prototip je lahko konkreten ali abstrakten. Prototip ptice je pri majhnem otroku lahko še povsem konkreten, medtem ko je pri odraslem človeku navadno zelo abstrakten. To pa ni nujno. Pri svojih predavanjih sem nekajkrat ponovila eksperiment, ki ga je H. Iyer (1995) povzela po Eleanor

Rosch. Svoje študente sem prosila, naj najprej naštejejo nekaj pojmov, ki sodijo v določeno kategorijo, in naj jih nato še rangirajo po značilnosti za to kategorijo. Ugotovila sem, da se vrtnica redno uvršča na prvi dve mesti med rožami. Do enakega rezultata je prišla Hemalata Iyer (1995). V kategoriji prevoznih sredstev pa se letalo pri slovenskih študentih vedno uvršča med manj značilne kot pri ameriških, ki jih je testirala Iyer. Vrtnica torej deluje kot prototip rože tako v ameriški kot slovenski kulturi, kar za letalo ne velja. To nas najbrž ne preseneča. Hkrati pa ilustrira zelo pomemben pojav, ki sem ga omenila že v začetku: vse lastnosti, ki odločajo o razvrstitvi posamezne enote v kategorijo, niso očitne in ne izhajajo iz lastnosti pojma. Pogosto izhajajo iz funkcije, ki jo ima pojem v življenju in delu posameznika ali družbene skupine.

KNJIŽNIČNE KLASIFIKACIJE

Lakoff, Wittgenstein in Rosch niso edini raziskovalci, ki so dokazali pomanjkljivosti klasične teorije klasifikacije. Kljub temu je večina klasifikacij, ki jih uporabljamo, od klasifikacij poklicev in bolezni do klasifikacije dokumentarnega in knjižničnega gradiva, zgrajenih po tem principu. Najbolj razširjene knjižnične klasifikacije, kot so Deweyeva decimalna klasifikacija, Klasifikacija Kongresne knjižnice in Univerzalna decimalna klasifikacija, so nastale pred več kot 100 leti. Takrat dokazov o pomanjkljivosti klasične teorije klasifikacije še ni bilo. Vendar se tudi danes, ko jih poznamo, osnovni principi teh klasifikacij ne spreminjajo. Spreminja se kvečjemu vsebina. To se dogaja zaradi pojavljanja novih disciplin ali zaradi kulturnega uravnoteženja. Hope Olson je opozorila na različne teme, ki jih knjižnične klasifikacije ne vključujejo ali stigmatizirajo v skladu z večinskim mnenjem v kulturnem okolju (Olson, 1998: Olson in Schlegl, 2001; Kublik et al., 2003).

Kulturno pristranskost knjižničnih klasifikacij lepo ilustrira skupina verstva v Univerzalni decimalni klasifikaciji (UDK). Izvirno je ta hierarhična klasifikacija, ki je razdeljena po desetiškem principu, v svoji skupini 2 Verstvo (religija). Bogoslovje (teologija) vsebovala naslednje podskupine:

21 Naravna teologija	26 Krščanska cerkev
22 Sveto pismo (Biblija)	27 Obča zgodovina krščanske cerkve
23 Dogmatika	28 Krščanske cerkve in sekte
24 Praktična teologija	29 Nekrščanska verstva in kulti. Mitologije
25 Pastoralna teologija	

Že na prvi pogled je razvidno, da je bilo 90% prostora namenjenega krščanstvu, vsem drugim nekdanjim in sedanjim svetovnim religijam pa le 10%. Če razumemo, da je klasifikacija nastala leta 1895 v Belgiji, potem lahko sprejmemo tedanji pogled na svet in pomembnost posameznih tematskih področij. Seveda pa je taka razporeditev verstev iz kulturnih in praktičnih razlogov zdaj neprimerna. To so ugotovili tudi uredniki klasifikacije in leta 2000 objavili spremembo, ki se je nedavno uveljavila tudi pri nas (Kovač, 2005).

21 Predzgodovinska in primarna verstva	26 Judovstvo
22 Verstva Daljnega vzhoda	27 Krščanstvo
23 Verstva indijskega podkontinenta	28 Islam
24 Budizem	29 Moderna duhovna gibanja
25 Verstva starega veka. Manjši kulti in verstva	

Nova razvrstitev omogoča knjižnicam predvsem preglednejšo razvrstitev gradiva o različnih verstvih, ki so na našem ozemlju prisotna in pomembna, niso pa kulturno dominantna.

Pregledna fizična razvrstitev gradiva na knjižničnih policah je vloga, ki jo knjižničarji pretežno pripisujemo knjižničnim klasifikacijam. Predvsem Klasifikacijo Kongresne knjižnice Američani pogosto hvalijo kot odlično sredstvo za označevanje in postavitve (»mark and park«). Mnogi pa se tudi ne zavedajo, da je UDK prva knjižnična klasifikacija, ki ni bila izdelana za signiranje gradiva (torej označevanje »poštne naslova knjige«), temveč je bila pravo orodje za iskanje informacij. Razvita je bila namreč kot pripomoček za iskanje gradiva po tematiki v veliki svetovni bibliografiji (McIlwaine, 1999; Rayward, 1997; Rieusset-Lemarié, 1997; Day, 1997). To bibliografijo sta si zamislila Paul Otlet in Henry La Fontaine. Njuna zbirka ni vsebovala knjig oz. dokumentov kot takih, temveč le njihov opis. V listkovnih katalogih je Mednarodni bibliografski inštitut imel na milijone zapisov o gradivu z vsega sveta. UDK pa je bila tisto sredstvo, ki je omogočalo opis po vsebini in tudi iskanje gradiva na določeno tematiko. (Bila je Google svojega časa.) Sprva so vsakemu dokumentu dodelili toliko klasifikacijskih oznak, kolikor tematik je pokrivalo. Ko pa so ugotovili, da se listkovni katalogi naglo povečujejo in postajajo zaradi obsega neobvladljivi, so se odločili za zelo pragmatično rešitev: vsak dokument dobi le eno klasifikacijsko oznako, in to za dominantno temo. Ta odločitev se je tudi pod vplivom drugih dveh dominantnih knjižničnih klasifikacijskih sistemov, ki dovoljujeta le eno klasifikacijsko oznako, med knjižničarji obdržala do danes. Žal pa je danes, v času računalniških katalogov, ta odločitev kontraproduktivna: za oblikovanje klasifikacijske oznake potrebujemo veliko znanja in časa, ko je izdelana pa je ne izkoristimo v celoti. Namesto, da bi jo uporabili za iskanje informacij, jo uporabljamo le za označevanje fizične lokacije gradiva v prostoru.

Klasifikacijska oznaka, ali vrstilec UDK ima veliko informativno vrednost ali izrazno moč. Z njim lahko izrazimo več lastnosti nekega dokumenta:

- temo (dogodek, osebo, pojav itd.),
- kraj, kjer se je tema dogajala ali jo dokument opisuje,
- čas, ko se je tema dogajala ali ga dokument opisuje,
- obliko, v kateri se dokument fizično (knjiga, CD) ali vsebinsko (enciklopedija, učbenik) pojavi in
- jezik, v katerem je dokument napisan.

Kljub temu, da vse lastnosti izrazimo v istem vrstilcu, npr.

061.3:930.25(497.4)"2007"(082), kar pomeni zbornik del z zborovanja slovenskih arhivistov leta 2007. Posamezni deli vrstilca pa imajo naslednji pomen:

061.3 zborovanje

930.25 arhivistika

(497.4) Slovenija

"2007" leto

(082) zbrana dela več avtorjev

Čeprav se tako zdi in bodo vsi bibliotekarji zagotovili, da dokument sodi v en sam razred, ki ima tako oblikovano oznako, to ne drži v celoti. Lahko bi rekli, da dokument sodi v kar pet različnih razredov. Bibliotekarji take sestavine posameznega razreda imenujemo fasete. Lahko jih razumemo kot različne vidike, s katerih opisujemo določen predmet in nam le v celoti dajejo popolno sliko tega predmeta. Slika kovanca je npr. brez prikaza obeh strani in roba nepopolna. V tem primeru bi »cifra« pomenila eno, »glava« drugo in »rob« tretjo faseto opisa tega predmeta.

Fasete si je zamislil že indijski bibliotekar in matematik Shiyali Ramrita Ranganathan v svoji »Colon classification« leta 1933 (Ranganathan, 1963; Iyer, 1995; Broughton, 2004). Vendar je bil zaradi enodimenzionalnosti listkovnega kataloga to neuporaben princip opisovanja in iskanja vsebine dokumentov in se ni uveljavil. Idejo so ponovno odkrili računalničarji, saj računalniški katalogi in evidence naravno podpirajo večdimenzionalne opise, kot je zgornji petdimenzionalni opis zbornika posvetovanja arhivistov.

Kljub temu, da je UDK nastala leta 1895, z uporabo faset dejansko krši pravila klasične teorije klasifikacije. Vsak dokument sodi v več razredov, moral pa bi le v enega. Če mi bodo bibliotekarji ugovarjali, da je zgornji vrstilec za zbornik posvetovanja arhivistov en sam razred, potem klasifikacija krši drugo pravilo: da moramo za delitev v podrazred uporabiti le eno lastnost. V našem primeru imamo kar pet lastnosti. Četudi sprejmemo tezo, da sta vrstilca za kongrese in arhivistiko istovrstna, torej tematska, nam še vedno ostanejo štiri različne lastnosti. Ravno večdimenzionalnost pa je tista lastnost, ki UDK naredi zanimiv tudi za sodobne računalniške kataloge in informacijske sisteme ter pokaže genialnost njegovih kreatorjev.

UDK IN KLASIFIKACIJE V RAČUNALNIŠKEM OKOLJU

Ena od mojih diplomantk je ugotovila, da računalniški katalogi različnih evropskih knjižnic razmeroma skromno uporabljajo klasifikacijske sisteme kot pomoč pri iskanju dokumentov po vsebini (Smrekar, 2003). V njeni raziskavi se je izkazalo, da COBISS/OPAC edini med pregledanimi računalniškimi sistemi omogoča iskanje po klasifikacijski oznaki. Celo več, omogoča tudi veliko različnih možnosti za iskanje po različnih delih sestavljenega vrstilca (npr. po posameznih delih vrstilca za naš zbornik), kar je med pregledanimi računalniškimi katalogi izjema.

Žal pa raziskave in izkušnje kažejo, da klasifikacijo uporablja le malo ljudi, tako bibliotekarjev, še manj pa uporabnikov knjižničnih katalogov. Številke si je težko zapomniti in še težje razumeti. Zato se tudi v Sloveniji mnogi zavzemajo za opustitev klasifikacije. V celoti naj bi jo nadomestilo iskanje po ključnih besedah. Toda tudi iskanje s ključnimi besedami ima v COBISS/OPAC z vidika iskanja po vsebini hude pomanjkljivosti. Predvsem to, da pri iskanju »Radencev« s ključnimi besedami najdemo tako tiste dokumente, ki govore o Radencih, kot tudi tiste, ki so v Radencih izšli ali bili tam natisnjeni, medtem ko pri iskanju po vsebini najdemo le tiste, ki

govore o Radencih. Razlika je v velikih katalogih lahko občutna in usodna. Raziskave so namreč pokazale, da uporabniki informacijskih sistemov nismo vztrajni in pogledamo le zaslon ali dva, redkeje tri zaslone rezultatov (Šaupperl in Saye, 2005).

Klasifikacijski sistemi v računalniškem okolju pa lahko podpirajo ravno tako vrsto človeške nepotrpežljivosti. Če v COBISS/OPAC v okence »ključne besede« vpišem izraz »klasifikacija«, me sistem obvesti, da je na voljo kar 3518 dokumentov s to oznako (rezultat 8. februarja 2007, ko je katalog obsegal nekaj nad 3 milijone zapisov). Na prvih treh zaslonih s po 10 najdenimi zapisi so bila dela z naslednjimi naslovi: *Astronomical spectroscopy*, *Ravnanje z odpadki znotraj meja pridobivalnega prostora Premogovnika Velenje*, *Enzyme nomenclature*, *Manual on monetary and financial statistics*, *Uporaba evropske osrednje klasifikacije proizvodov v slovenski statistiki in Predlog klasifikacije izkoristkov pri obločnem varjenju*. Toliko zaslonov bi pregledal vztrajen povprečen uporabnik. Sama sem seveda vztrajala, vendar med prvimi 100 zadetki ni bilo nobenega dokumenta, ki bi obravnaval knjižnično ali arhivsko klasifikacijo. »Pameten« informacijski sistem bi v takem primeru analiziral najdenih 3518 dokumentov in jih grupiral po skupnih lastnostih, npr. po UDK. Namesto, da bi mi odgovoril »Izpis je omejen na 2000 zadetkov! število zadetkov 3518«, bi lahko ponudil izbiro: »Izraz klasifikacija se pojavlja na naslednjih področjih: bibliotekarstvo, ekonomija in finance, tehnika, medicina, umetnost, šport, arhivistika«. Še število zapisov bi mi lahko povedal ob vsakem strokovnem področju, ki bi bilo opremljeno s hiperpovezavo. Za grupiranje pa bi uporabil vrstilec UDK, ki ga vsak zapis nujno vsebuje. Če zgornje iskanje ponovim z omejitvijo na vrstilec za bibliotekarstvo (02), mi sistem ponudi veliko manj - le 101 zapis. V tem primeru bi lahko sistem ponovil grupiranje, tokrat po podrejenih vrstilih in mi ponudil v pregled še manjše skupine. Če bi bile skupine premajhne, na arhivistiko (930.25) omejeno iskanje npr. obrodi le 2 dokumenta, bi sistem lahko izkoristil hierarhično strukturo klasifikacije in me obvestil, da so koristni dokumenti morda razvrščeni v širše skupine. V primeru arhivistike bi te bile:

930.2 Metodologija zgodovine, pomožne zgodovinske vede

930 Zgodovinska znanost. Zgodovinopisje

93/99 Zgodovina

9 Geografija. Biografija. Zgodovina (Univerzalna, 2007)

Pri vseh omenjenih transakcijah pa uporabnik sploh vedel ne bi, da je uporabil klasifikacijski sistem in se mu klasifikacijske oznake niti ne bi prikazale na zaslonu.

Neopazno bi klasifikacijski sistem lahko uporabili še s pretvorbo v besedilo. Iskali bi arhivistiko, računalniški sistem pa bi besedo avtomatsko pretvoril v klasifikacijsko oznako in iskanje v katalogu izvedel z njo. V odgovor bi ponudil dokumente, ki ustrezajo iskalni besedi, ne da bi klasifikacijsko oznako kakorkoli omenil. S tako vrsto pretvorbe sta se ukvarjala Riesthuis in Bliedung (1991) in ugotovila, da bi bilo UDK sicer mogoče uporabiti na ta način, da pa bi ga bilo treba pregledati in mestoma preurediti, ker nekaterih vrstilcev ne moremo enoznačno prevesti v besede, kot tudi ni mogoče vseh besed neposredno prevesti v en sam vrstilec. Toda v veliko primerih bi ta način pretvarjanja med oznako in besedo iz naravnega jezika deloval in četudi sistem ni popoln, bi lahko pomagal pri »iskanju igle v kopici sena«.

Zelo velik katalog je že COBIS/OPAC. Mnogo večji pa je WorldCat (dostopen na URL: <http://www.worldcatlibraries.org/>). To je skupni katalog nad 10.000 knjižnic in knjižničnih sistemov vsega sveta, v katerem sodelujemo tudi Slovenci z vzajemnim katalogom. To je elektronska oblika svetovne bibliografije, kot sta si jo konec 19. stoletja zamislila Paul Otlet in Henri La Fontaine. Ustvaril pa ga je še eden od velikih mož svetovnega bibliotekarstva, Fred Kilgour (*The best of...*, 2006). Zaradi izrednega obsega svetovnega kataloga, vsebuje namreč nad milijardo dokumentov, je iskanje po tematiki precej težavno. Večina vidi njegovo prednost predvsem v tem, da je mogoče ugotoviti, katera knjižnica ima določeno gradivo in si z medknjižnično izposajo to gradivo tudi zagotoviti. Toda struktura kataložnega zapisa v WorldCatu je standardizirana in vsak podatek ima svoje mesto. Tako UNIMARC kot osnovni standard za strukturo kataložnega zapisa in posledično vsi nacionalni standardi, ki iz njega izhajajo, vsebujejo posebno polje za vpis vsake od klasifikacijskih oznak. Ker imajo mnogi dokumenti vsebino opisano z različnimi klasifikacijskimi sistemi, jih lahko uporabimo za »prevajanje« klasifikacijskih oznak med različnimi klasifikacijskimi sistemi.

Pretvarjanje klasifikacijskih oznak iz enega v drug sistem je v bibliotekarstvu že dolgo prisotno in razširjeno predvsem v ZDA. Tam večje knjižnice, ki večinoma vpisujejo gradivo v vzajemni katalog, uporabljajo Klasifikacijo Kongresne knjižnice, manjše, ki večinoma le kopirajo kataložne zapise iz vzajemnega v lokalni katalog, pa Deweyevo decimalno klasifikacijo. Spretni založniki so zaznali tržno nišo in izdali pretvorbene tabele med eno in drugo klasifikacijo (npr. Scott, 1999), knjižničarji pa jih z veseljem uporabljajo, saj jim prihranijo veliko truda in časa.

Tako pretvorbo pa nameravajo s pridom izkoristiti tudi evropske nacionalne knjižnice v skupnem projektu Evropska knjižnica (*The European Library*). Zamišljeno je, da bi uporabnik lahko iskanje gradiva začel v nacionalni knjižnici svoje države in v svojem jeziku. Potem pa bi sistem sprožil iskanje tudi v katalogih drugih knjižnic, vendar z avtomatsko pretvorbo iskalnega izraza iz enega v druge evropske jezike, ne da bi s tem obremenjeval uporabnika. Uspešni poskusi so že bili opravljeni (MacEwan, 2000).

Prevajanje klasifikacijskih sistemov, geslovnikov in tezavrov sicer prinaša nekaj težav. Mnogo pojmov lahko prevedemo v enoznačen pojem v drugem jeziku. Vedno pa ni tako. Nekateri banalni pojmi nimajo enoznačnega prevoda, npr. prst:

slovensko	prst (na roki, nogi)
angleško	finger / toe
nemško	Finger / Zeh
italijansko	dito (del mano, del piede)

Toda zaradi takih primerov ne bi smeli obupati. Proučiti bi morali še druge možnosti, ki nam bi omogočile avtomatsko razločevanje in združevanje podobnih pojmov. To pa je že področje avtomatske klasifikacije, ki je obsežno in plodno raziskovalno in strokovno področje že vrsto desetletij. Konkreten rezultat raziskav je vizualizacija skupin dokumentov pri iskanju v sistemu EBSCO Host.

Avtomatska klasifikacija je v bibliografskih podatkovnih zbirkah že prisotna, v knjižničnih katalogih pa ne. V sistemih za vzdrževanje velikih količin podatkov (*content management systems*) pa je ta proces zelo pomemben. Čeprav večina takih sistemov omogoča in zahteva, da dokumente opišemo z gesli ali predmetnimi oznakami, lahko avtomatski postopki klasifikacije, ki za delovanje uporabljajo

celotno besedilo dokumentov, pripomorejo k temu, da lažje najdemo dokumente. Ni namreč nujno, da si bo geslilec v času zapisa gesel zamislil vse pomembne vidike za opis določenega dokumenta.

Da geslilci ne bi pozabili pomembnih lastnosti dokumentov, ki jih je treba opisati, so zelo primerni obrazci, ki geslilca vodijo od ene do druge fasete (beri lastnosti dokumenta) in tako preprečujejo površnost ali pozabljenost, dve zelo človeški lastnosti. Ker smo ljudje zmotljivi ne glede na to, kako skrbni in vestni smo, nam računalniki lahko pomagajo še na dva načina: s preverjanjem črkovanja in s šifranti. Preverjanja črkovanja nam v programih za pisanje besedil popravljajo zatipkane besede. Tudi v računalniških katalogih bi nam ta možnost zelo prav prišla, tako knjižničarjem kot bralcem. Šifrante, drugi pomemben pripomoček, v katalogizaciji dobro poznamo. To so navadno končni sezname pojmov, ki jih smemo za opis določene lastnosti uporabiti. Mednarodna šifra za Slovenijo je npr. SI, za slovenščino pa SVN (SIST, 2003). Med take šifrante bi lahko šteli tudi normativne datoteke. Te pa nimajo le vloge naštevanja dovoljenih šifer, temveč pomenijo bistveno več. Poleg seznama imen ali gesel vsebujejo tudi vse psevdonime in alternativne izraze ter povezave med imeni in psevdonimi ter medsebojne povezave pojmov. Tako npr. normativna datoteka osebnih in korporativnih imen CONOR¹, ki v COBISS-u deluje od leta 2003, vsebuje vseh 15 oblik Prešernovega imena, kot se je pojavilo v publikacijah, ki jih hranijo slovenske knjižnice. Čar normativne datoteke je v tem, da lahko uporabnik išče katerokoli obliko imena, pa bo v vsakem primeru prišel do vseh Prešernovih del, ne glede na to, s kakšno obliko zapisa avtorjevega imena je izšla. Iskalec tematiko, ki ga zanima, navadno poimenuje z izrazi, ki so mu blizu ali jih je spoznal. Ni pa nujno, da se ti izrazi ujemajo z izrazi, ki so v katalogu uporabljeni za vsebinski opis. Normativna datoteka lahko s svojimi povezavami med pojmi kot s pajčevino ujame takega iskalca in ga usmeri na izraze, ki so uporabljeni v katalogu. Recimo, da je v katalogu za opis dokumentov uporabljen izraz »književnost«, vi pa iščete z izrazom »literatura«. Če katalog nima normativne datoteke, ne boste našli ničesar. Ne zato, ker v katalogu ni nič primernega, temveč zato, ker se iskalni izraz ne ujema z opisom v katalogu. Katalog z normativno datoteko pa bo uporabniku ponudil vse dokumente, ki obravnavajo »književnost«, ne da bi ga obremenjeval z razliko. Take mrežne povezave pojmov so vzpostavljene tako v klasifikacijskih sistemih kot v geslovnikih in tezavrih, ki jih v knjižnicah uporabljamo za vsebinski opis dokumentov. Že zelo dolgo imamo torej intelektualno osnovo za normativno datoteko vsebinskih oznak, v COBISS-u pa še nimamo tehnologije, ki bi to dohajala.

ZAKLJUČEK

Čeprav pa knjižnične kataloge internet in različni drugi informacijski sistemi (kot npr. EBSCOhost) pri izkoriščanju avtomatske klasifikacije preHITEVajo, večinoma še ne izkoriščajo drugih omenjenih prednosti, ki jih klasifikacije lahko ponudijo. Sistemi, ki so nastali v času, ko je bila racionalizacija osrednji pojem, se zdijo zdaj, ko smo z računalniškim prostorom in zmogljivostmi lahko razsipni, popolnoma odveč in zastareli. Sredstvo, ki je bilo učinkovito v listkovnih katalogih, se zdi v času zmogljivih informacijskih sistemov nepotrebno. Ostaja pa naša ranljivost in naše omejitve, omejitve, s katerimi se sooča tudi najbriljantnejši um človeštva. Uporabljamo lahko le en procesor in računalniški sistemi z več procesorji, nam pri

¹ URL: <http://cobiss4.izum.si/scripts/cobiss?id=1553133220231872&ukaz=RPAX&rpa=26&bno=++509>

tem lahko delo olajšajo ali otežijo. Med procesi, ki pripomorejo k olajšanju, je tudi klasifikacija s svojo lastnostjo grupiranja podobnih elementov.

Seveda pa ne bi bilo smiselno klasifikacij v novem okolju uporabljati na povsem tradicionalen način. Modro je združiti prednosti, ki jih nudi sodobna informacijska tehnologija, in znanje, ki so ga prispevali raziskovalci z različnih področij, tako glede človekovega mišljenja kot tudi filozofskih osnov klasifikacije in jezika ter značilnosti in algoritmov za avtomatsko obdelavo podatkov. Z moderno zasnovano klasifikacijo, ki je prožna, da lahko hitro vključuje nove znanosti in ki se je sposobna prilagajati terminologiji uporabnika, bi nam informacijski sistemi morda celo omogočili obvladati plaz informacij, ki se na nas vali iz knjižničnega kataloga, interneta ali drugih virov.

Podobno velik vir so tudi arhivski fondi z ustreznimi dokumentacijami. Pomen njihove ureditve je vam, kolegom iz arhivov, veliko bolj znan kot meni, bibliotekarki. Vendar so med upravljanjem evidence velike knjižnice in arhiva mnoge podobnosti. Četudi se mnogi atributi za opis knjižničnega in arhivskega gradiva razlikujejo, imajo skupnih nekaj kategorij atributov: identifikacijo dokumenta in njegov vsebinski opis. Nekateri elementi za identifikacijo se med obema ustanovama najbrž razlikujejo. Pri vsebinskem opisu pa se najbrž tudi vi ubadate s konceptom faset, ki ni skladen z npr. klasifikacijskim načrtom za potrebe samoupravnih lokalnih skupnosti (Klasinc in Verbič, 1999). Prav tako se najbrž ukvarjate z nekaterimi elementi opisa, s katerimi se ukvarjamo tudi bibliotekarji, npr. imeni oseb, korporacij in krajev. Predmet sodelovanja se torej ponuja kar sam. Prijazno povabilo organizatorjev na posvetovanje pa je tudi dejanje v tej smeri. Hvala.

VIRI

- *The best of Fred Kilgour*. 2006. *The OCLC Newsletter* [online]. Št. 3. [Uporabljeno 2007-02-08]. Dostopno na <http://www.oclc.org/nextspace/003/2.htm>.
- Bowker, G.C. in Star, S.L. 1999. *Sorting things out : classification and its consequences*. London : MIT Press.
- Bowler, P.J. 1993. *The Norton history of the environmental sciences*. 1st American Ed. New York : W.W. Norton.
- Broughton, V. 2004. *Essential classification*. London : Facet.
- Dahlberg, I. 1989. *Concept and definition theory*. V *Classification theory in the computer age : conversations across disciplines. Proceedings from the conference, November 18-19, 1988, Albany, New York* (str. 12-24). Albany, New York : Rockefeller College Press, University at Albany, State University of New York.
- Day, R. 1997. *Otlet's book and the writing of social space*. *Journal of the American Society for Information Science*, letn. 48, št. 4, str. 310-317.
- *The European Library*. [online]. 2005-2007. [Uporabljeno 2007-02-08]. URL: <http://www.theeuropeanlibrary.org/portal/index.html>.
- Iyer, H. 1995. *Classificatory structures : concepts, relations and representation*. Frankfurt a/M : Indeks Verlag.
- Klasinc, P.P. in Verbič, D. 1999. *Dokumentarno in arhivsko gradivo v samoupravnih lokalnih skupnostih*. Maribor : Arhivsko društvo Maribor.
- Kovač, T. 2005. *Novi UDK vrstilci za teologijo in verstva*. *Knjižnica*, letn. 49, št. 3, str.81-94.
- Kublik, A., Clevette, V., Ward, D. in Olson, H.A. 2003. *Adapting dominant classifications to particular contexts*. *Cataloging and Classification Quarterly*, letn. 37, št. ½, str. 13-31.
- Lakoff, G. 1987. *Women, fire, and dangerous things : what categories reveal about the mind*. Chicago : University of Chicago Press.
- MacEwan, A. 2000. *Crossing language barriers in Europe: linking LCSH to other subject heading languages*. *Cataloging and Classification Quarterly*, letn. 29, št. ½, str. 199-207.

- Mai, J.E. 2001. *Semiotics and indexing: An analysis of the subject indexing process*. *Journal of Documentation*, 57, 591-622.
- Mcllwaine, I. 2000. *The universal decimal classification : a guide to its use*. Hague : UDC Consortium.
- Olson, H.A. 1998. *Mapping beyond Dewey's boundaries : constructing classificatory space for marginalized knowledge domains*. *Library Trends*, letn. 47, št. 2, str. 233-255.
- Olson, H.A. in Schlegl, R. 2001. *Standardization, objectivity and user focus : a meta-analysis of subject access critiques*. *Cataloging and Classification Quarterly*, letn. 32, št. 2, str. 61-80.
- Peirce, C.S. 1991. *Peirce on signs : writings on semiotics*. Chapel Hill : University of North Carolina Press.
- Phares, A.D. 1985. *Charles S. Peirce and the linguistic sign*. Amsterdam : John Benjamins Publishing.
- Ranganathan, S.R. 1963. *Colon classification : basic classification*. 6th ed. Bombay : Asia Publishing House.
- Rayward, W.B. 1997. *The origins of information science and the International Institute of Bibliography/International Federation for Information and Documentation (IFID)*. *Journal of the American Society for Information Science*, letn. 48, št. 4, str. 285-288.
- Riesthuis, G.J.A. in Bliedung, S. 1991. *Thesaurification of the UDC*. V *Tools for knowledge organization and the human interface. Proceedings of the First International ISKO conference, Darmstadt, 14-17 August 1990. Part 2*. Frankfurt : Indeks, str. 109-117.
- Rieusset-Lemarié, I. P. 1997. *Otlet's Mundaneum and the international perspective in the history of documentation and information science*. *Journal of the American Society for Information Science*, letn. 48, št. 4, str. 301-309
- Rosch, E. 1973. *Classification of real world objects : origin and representations in cognition*. V Moore, T. (ur.). *Cognitive development and acquisition of knowledge*. New York : Academic Press.
- Scott, M.L. 1999. *Conversion tables*. 2nd ed. Englewood, Co. : Libraries Unlimited.
- SIST EN ISO 3166-1 :2003 - *Kode za predstavljanje imen držav in njihovih podrejenih enot*. Prvi del : *Kode držav*.
- Smrekar, U. 2003. *Možnosti za boljše izkoriščanje Univerzalne decimalne klasifikacije v računalniškem katalogu : diplomsko delo*. Ljubljana : Univerza v Ljubljani, Filozofska fakulteta, Oddelek za bibliotekarstvo, informacijsko znanost in knjigarstvo.
- Šauperl & Saye. 2005. *Opis dokumentov in možnosti za njihov priklic v svetovnem spletu*. V *Informacijski viri in storitve knjižnic v elektronskem okolju, strokovno posvetovanje Zveze bibliotekarskih društev Slovenije, Portorož, 24.-26. oktober 2005*. Ljubljana : Zveza bibliotekarskih društev Slovenije. Str. 95-111.
- *Univerzalna decimalna klasifikacija*. [online]. 2007. Ljubljana : Narodna in univerzitetna knjižnica. [Uporabljeno 2007-02-08]. Dostopno na <http://www.nuk.uni-lj.si/udk/>.
- Wittgenstein, L. 1981. *Philosophical investigations*. Oxford : Blackwell.

SUMMARY

LESS KNOWN CHARACTERISTICS OF CLASSIFICATION

A BRIEF OVERVIEW OF THE THEORY OF CLASSIFICATION AND ITS USE IN LIBRARY CATALOGUES

Classification systems, which were very popular in the past, seem quite redundant today. However being a part of human cognition, they are not only a relict. Besides powerful computers we also need intellectual tools, which will help us manage large quantities of data. Classification systems are such tools. Traditional library classification systems need to be adapted for contemporary use. They need to implement newer theories of classification, e.g., the prototype theory. Contemporary classification systems also need to take advantage of all the qualities of information technology. Traditional card catalogues confined the description and

search to one element of dimension, e.g. the author of the work. Digital world is multidimensional, which means that we can describe and search by several elements simultaneously, e.g. author, publication year, and subject. Faceted classification systems are ideal for this environment and also more similar to human natural processing than hierarchical systems. In a digital environment classification systems can be used in other beneficial ways as well. In the European Library project it will be possible to automatically expand the search in the language of one of the national libraries into the catalogues of other national libraries that use other languages. Classification systems could also be employed for guiding users during information search. Relations among the terms, built into the classification systems would only then be used to the full extent.