

JABLANOV ŠKRLUP: BIBLIOMETRIČNA ANALIZA - PREKRIVANJE ZBIROK IN OSREDNJE REVIE

Karmen STOPAR¹

¹Biotehniška fakulteta, Oddelek za agronomijo, Ljubljana

IZVLEČEK

Jablanov škrlup, ki ga povzroča gliva *Venturia inaequalis*, velja za najpomembnejšo bolezen jablan, varstvo pa zahteva intenzivno rabo fungicidov. V prispevku je analizirana literatura s tega področja, ki je bibliometrično ovrednotena na podlagi obsega, rasti in vključevanja relevantnih podatkov v revije in bibliografske podatkovne zbirke. V analizo prekrivanja je vključenih pet bibliografskih podatkovnih zbirk: AGRIS, AGRICOLA, CAB, SCISEARCH in Biological Abstracts (BIOSIS), ki so bile izbrane glede na vključenost relevantnih dokumentov. Rezultati kažejo, da je med zbirkami precejšnje prekrivanje, večino relevantnih dokumentov pa najdemo že v zbirkah AGRIS in CAB. S pomočjo Jaccardovega in ekvivalenčnega koeficienta smo ugotavljali tudi trend prekrivanja med zbirkama. Članki so po revijah razporejeni v skladu z Bradfordovim zakonom, po katerem je glavnino relevantnih informacij mogoče najti že v nekaj virih. Tretjino vseh relevantnih člankov vključuje že 5% analiziranih revij.

Ključne besede: bibliometrija, literatura, podatkovne zbirke, revije, jablanov škrlup

ABSTRACT

APPLE SCAB: A BIBLIOMETRIC ANALYSIS - DATABASE OVERLAP AND CORE JOURNALS

Apple scab, caused by the fungal pathogen *Venturia inaequalis*, is considered to be the most important disease of apple worldwide. Its control requires intensive use of fungicides. In the paper apple scab-related publishing was bibliometrically evaluated with the view of growth, quantity and coverage of relevant information in periodicals and bibliographic databases. Five bibliographic databases (AGRIS, AGRICOLA, CAB, SCISEARCH and Biological Abstracts (BIOSIS)) were examined with regard to inclusion of scab-related documents. The results show high overlapping, however, the majority of scab-related documents can be found in two databases: AGRIS and CAB. Trend of overlapping was analyzed with Jaccard's and Equivalence index. Scatter of journal-articles is in accordance with Bradford's bibliometric law, and shows that 5% of all journals account for a third of all relevant articles.

Key words: bibliometrics, publishing, Bradford's law, databases, journals, apple scab

1 UVOD

Živimo v času, ko se na vseh področjih število informacij pospešeno povečuje, kar še posebej velja za znanstvene in strokovne informacije. Hkrati z naraščajočim številom publikacij, pa je zaznati tudi izjemno hiter razvoj in razmah tehnologij za sistematično obdelovanje in hranjenje vseh teh podatkov. Dostop do informacij se je z bibliografskimi in drugimi podatkovnimi zbirkami spremenil, saj omogočajo shranjevanje velike količine podatkov in, kar je še pomembneje, standardiziran dostop do njih. Z naglim razvojem komunikacijskih tehnologij in interneta pa postajajo informacije še dostopnejše in za uporabnika tudi neposredno dosegljive.

¹ mag., univ. dipl. inž. agr., Jamnikarjeva 101, SI-1111, Ljubljana

Ogromna količina sistematično shranjenih podatkov, pa po mnenju mnogih, še vedno ni učinkovito uporabljena. Pri iskanju relevantnih informacij se uporabnik pogosto izgublja v veliki množici podatkov. Garfield (1994) in Kuhlthau (1997) pišeta o občutkih tesnobe, ki jih doživlja ob tem. Kljub razvoju na področju avtomatiziranega iskanja informacij je iz množice podatkov namreč zelo težko dobiti pravo število relevantnih dokumentov za posameznega uporabnika. Brier (1996) ugotavlja, da nastaja potreba po zmanjšanju števila informacij v število, s katerim je še mogoče upravljati. Rezultat iskanja je namreč pogosto preveliko število najdenih dokumentov in med njimi številni nerelevantni ali manj relevantni dokumenti (t. i. šum). Poleg tega, da je preciznost iskanja nizka, pa je zelo težko tudi oceniti koliko pomembnih dokumentov pa pri iskanju nismo zajeli (t. i. odziv). Relevantnim informacijam je pogosto težko slediti tudi zato, ker so razpršene po številnih zbirkah oziroma informacijskih virih. Bartol (1998) ugotavlja, da je za uspešno iskanje zelo pomembno, da podatkovno zbirko vnaprej poznamo, predvsem kako je sestavljena, strukturirana in indeksirana. Poudarja tudi vlogo informacijskega specialista, ki lahko na podlagi poznavanja informacijskih virov uporabnika pravilno usmerja. Laine-Cruzel in sod. (1996) sklepajo, da je to pomembno še zlasti zato, ker zaradi velikega števila informacij uporabnik pogosto kar sam zmanjšuje število zbirk, po katerih išče.

Pospešen razvoj in uporaba informacijske tehnologije ter razširjenost računalniško dostopnih podatkovnih zbirk pa so vzpodbudile tudi razvoj bibliometrije - znanstvenega področja, ki se ukvarja s študijem kvantitativnih vidikov proizvodnje, širjenja in uporabe zapisanih informacij. Pri tem se, razen štetja publikacij, citatov in drugih informativnih enot, poslužuje tudi različnih matematičnih, verjetnostnih in statističnih metod. Predmet proučevanja so različne vrste bibliografskih informacij: deskriptorji, avtorji, faktor vpliva ipd. (Ungern-Sternberg, 1998). Diodato (1994, cit. po Osareh, 1996) deli bibliometrične analize na tri večja področja: bibliometrične zakone in razporeditve po teh zakonih, spremljanje citiranosti oziroma odmevnosti objav in analizo raziskovalne učinkovitosti. Burton (1988) zelo na splošno sklepa, da lahko rezultati bibliometričnih analiz pomagajo knjižničarjem in informacijskim specialistom pri boljšem upravljanju z lastnimi viri, raziskovalcem pa pri boljšem razumevanju in pravilnem napovedovanju sprememb o informacijskih virih na področju, kjer delujejo.

V prispevku je bibliometrično ovrednotena literatura s področja jablanovega škrlopa, ki je analizirana na podlagi obsega, rasti in vključevanja relevantnih podatkov v revije in bibliografske podatkovne zbirke. S prekrivanjem treh kmetijskih zbirk (AGRIS, AGRICOLA in CAB) se je v preteklem obdobju ukvarjal že Reid (1992) in ugotovil, da je prekrivanje med zbirkami precejšnje. Dejansko število novih različnih dokumentov, ki izidejo v nekem letu, naj bi bilo le okrog 58% bibliografskih zapisov, ki jih je moč najti v vseh treh omenjenih zbirkah skupaj.

2 MATERIAL IN METODE DE LA

V analizo smo zajeli bibliografske zbirke, ki vključujejo največ zapisov za dokumente s področja jablanovega škrlopa; poleg treh t. i. kmetijskih AGRICOLA, AGRIS in CAB, še zbirki Biological Abstracts (BIOSIS) in SCISEARCH (SCI). Iskalni profil, na podlagi katerega smo iskali najprej najrelevantnejše zbirke za analizirano področje, kasneje pa bibliografske zapise za relevantne dokumente, smo zapisali kot presek med izrazi za jablano in glivo oz. bolezen, pri tem pa smo upoštevali angleško in latinsko poimenovanje. V bibliometričnih analizah operiramo z velikim številom podatkov, zato smo izbrali zbirke na CD-ROM-u. Ta tehnologija omogoča časovno neomejeno analiziranje zbirk in pretakanje (downloading) podatkov. Za potrebe analize prekrivanja zbirk smo glede na čas zajemanja podatkov in število obravnavanih zbirk ločeno analizirali dva seta podatkov. Za daljše časovno obdobje, publikacije z letom izida 1986 do 1997, so bili posebej analizirani bibliografski zapisi iz treh kmetijskih zbirk, za katere se predpostavlja, da skupaj pokrivajo večino kmetijske literature. Vzporedno smo za potrebe analize prekrivanja zbirk

primerjalno analizirali še podatke iz vseh petih zbirk, za publikacije z letom izida 1992-1997, saj so bili za to obdobje razpoložljivi podatki na CD-ROM-u za vse omenjene zbirke. Zbirke sledijo različnim pravilom v okrajševanju, črkovanju, transliteraciji, zato je bilo potrebno nehomogen set podatkov, ki je nastal kot posledica vključevanja različnih zbirk v analizo, najprej poenotiti. Podatke smo obdelovali s pomočjo programa MS Excel. S koeficienti sopojavljanja smo spremljali trende v podvajanju oziroma hkratnem pojavljanju bibliografskih zapisov za iste dokumente v zbirkah AGRIS in CAB. Uporabili smo Jaccardov in ekvivalenčni koeficient (DataView, 2000). Razporejanje člankov po različnih revijah in glavnino revij za področje jablanovega škrlupa smo analizirali z Bradfordovim zakonom. Revije, ki smo jih predhodno razvrstili v zaporedje po padajočem številu člankov na revijo, smo razdelili na enakovredne skupine, s približno enakim številom člankov po skupini. Prvo skupino - jedro predstavljajo revije, ki so vsebinsko najbolj namenjene izbranemu področju. Teoretično naj bi po tem modelu število revij v jedru in sledečih skupinah sledilo geometričnemu zaporedju $1:k:k^2...$

3 REZULTATI IN RAZPRAVA

3.1 Prekrivanje zbirk

V literaturi je hkratno pojavljanje zapisov za iste dokumente v različnih zbirkah opisano kot prekrivanje zbirk, sopojavljanje ali tudi podvajanje zapisov. Uporabnika, ki želi imeti celovit pregled nad literaturo za izbrano področje in išče podatke po različnih zbirkah, dvojniki ne zanimajo, če jih je veliko, je to gotovo celo moteče. Preglednica 1 številčno prikazuje podatke iz obeh setov, vključenih v analizo prekrivanja zbirk. V zbirkah AGRIS, AGRICOLA, CAB smo našli skupaj 2020 bibliografskih zapisov za dokumente s področja jablanovega škrlupa z letom izida od 1986 do 1997 (Pregl. 1). Imenujemo jih podvojeni zapisi, saj je med njimi dejansko le 1441 (71%) različnih dokumentov. Nekaj dokumentov je indeksiranih namreč v več zbirkah hkrati.

V vseh petih zbirkah smo našli 1412 bibliografskih zapisov za relevantne dokumente z letom izida 1992-1997, med katerimi je dejansko 800 (57%) različnih dokumentov. Iz preglednice 1 je razvidno, da vključuje, med tremi kmetijskimi zbirkami, AGRICOLA najmanj bibliografskih zapisov za relevantne dokumente z letom izida 1986-1997 - le 284 oziroma 20% (284 od 1441), zbirka AGRIS 750 oziroma 52% (750 od 1441) in CAB 986, kar je kar 68% (986 od 1441) vseh relevantnih dokumentov.

Preglednica 1: Število bibliografskih zapisov za dokumente s področja jablanovega škrlupa v treh kmetijskih zbirkah, skupno število podvojenih zapisov in različnih dokumentov po posameznih letih v obdobju 1986-1997 ter število relevantnih zapisov v BIOSIS in SCI v obdobju 1992-1997.

Leto	1986-1997					1992-1997	
	AGRICOLA	AGRIS	CAB	Podvojeni zapisi	Različni dokumenti	BIOSIS	SCI
1986	49	81	83	213	143		
1987	16	55	84	155	117		
1988	18	60	74	152	116		
1989	11	65	60	136	109		
1990	24	51	82	157	109		
1991	15	54	53	122	85		
1992	24	44	87	155	105	28	16
1993	24	52	69	145	104	21	16
1994	37	114	106	257	171	62	25
1995	15	44	77	136	99	23	18
1996	28	77	89	194	138	35	28

1997	23	53	122	198	145	31	24
Skupaj	284	750	986	2020	1441	200	127

Analiza prekrivanja posameznih dvojic zbirk je pokazala, da so dokumenti, ki jih najdemo v zbirki AGRICOLA, v velikem obsegu vključeni tudi v zbirki AGRIS in CAB. Kar 74% (211 od 284) vseh zapisov iz te zbirke se podvoji v zbirki AGRIS in 66% (189 od 284) v zbirki CAB. V primeru zbirke AGRIS je prekrivanje manjše. 28% (211 od 750) zapisov iz te zbirke najdemo tudi v zbirki AGRICOLA in 42% (318 od 750) v zbirki CAB. Zbirka CAB izstopa po številu vključenih dokumentov za analizirano področje, nižji pa je tudi delež prekrivanja z ostalima zbirkama. 19% (189 od 986) zapisov iz te zbirke se podvoji v zbirki AGRICOLA in 32% (318 od 986) v zbirki AGRIS (Pregl. 2). Zanimiv je tudi podatek, koliko dokumentov je indeksiranih samo v določeni zbirki. Od skupno 1441 različnih dokumentov (tri kmetijske zbirke, obdobje 1986-1997) jih samo AGRICOLA in nobena druga zbirka, vključuje le 22, kar predstavlja okrog 1,5%. V zbirki CAB je takšnih dokumentov, ki jih drugi dve zbirki ne zajemata 625 (43%), v zbirki AGRIS pa 364 (25%). Skupno predstavljajo dokumenti, ki so indeksirani samo v zbirki AGRIS in samo v zbirki CAB skoraj 69% vseh 1441 relevantnih dokumentov.

Preglednica 2: Število relevantnih dokumentov/zapisov v posameznih zbirkah in število dokumentov, ki se pojavljajo hkrati v posameznih parih zbirk (obdobje 1986-1997)

Zbirka	AGRICOLA	AGRIS	CAB
AGRICOLA	284	211	189
AGRIS	211	750	318
CAB	189	318	986

Za krajše časovno obdobje smo dodatno analizirali še vseh pet zbirk skupaj. Zbirki Biological Abstracts (BIOSIS) in SCISEARCH (SCI) vključujeta za obdobje 1992-1997, poleg treh kmetijskih, največ zapisov za dokumente s področja jablanovega škrlopa.

Preglednica 3: Število relevantnih zapisov/dokumentov v posameznih zbirkah in število dokumentov, ki se pojavijo hkrati v posameznih parih zbirk (obdobje 1992-1997)

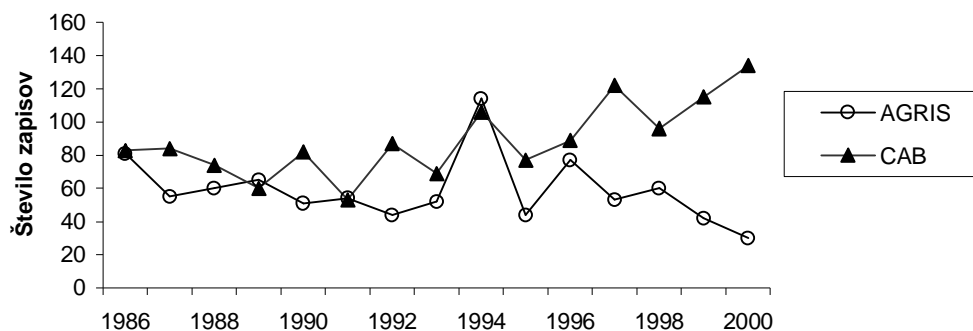
Zbirka	AGRICOLA	AGRIS	BIOSIS	CAB	SCI
AGRICOLA	151	119	80	114	76
AGRIS	119	384	126	186	75
BIOSIS	80	126	200	151	99
CAB	114	186	151	550	99
SCI	76	75	99	99	127

Zbirka BIOSIS, ki sicer po opisu zajema vsa področja žive narave, vendar bolj s poudarkom na biologiji in biomedicini, vključuje relativno veliko dokumentov s področja jablanovega škrlopa. Od skupno 800 različnih dokumentov jih najdemo kar 200 tudi v tej zbirki, kar predstavlja četrtino vseh relevantnih dokumentov. V zbirki SCI jih najdemo 127 oziroma 16%. To navede na misel, da bi bilo morda potrebno pri poizvedbah oziroma iskanju dokumentov s kmetijsko vsebino, poleg tako imenovanih "treh velikih" - kmetijskih zbirk, bolj upoštevati tudi omenjeni dve. Analiza prekrivanja vendarle pokaže, da se 76% (151 od 200) zapisov iz zbirke BIOSIS podvoji v zbirki CAB, 63% (126 od 200) v zbirki AGRIS, 50% (99 od 200) v SCI in 40% (80 od 200) v zbirki AGRICOLA (Pregl. 3). Hkrati je zelo malo tistih dokumentov, ki jih najdemo izključno v BIOSIS oziroma SCI in nobeni drugi

zbirki. V zbirki BIOSIS je takšnih od 800 različnih dokumentov z letom izida 1992-1997 le 12, v SCI pa le 20.

3.2 Trend podvajanja zbirk AGRIS in CAB

V zbirkah AGRIS in CAB je torej mogoče najti večji del literature za področje jablanovega šklrupa.



Slika 1: Letna rast bibliografskih zapisov s področja jablanovega šklrupa v zbirkah CAB in AGRIS v obdobju 1986-2000

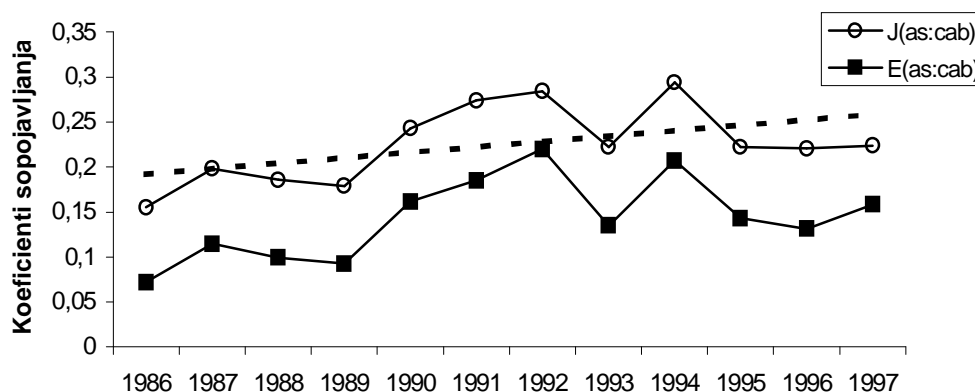
Z analizo sopojavljanja smo ugotavljali, ali se hkratio pojavljanje zapisov za dokumente v omenjenih zbirkah z leti spreminja. Iz vrste funkcij, ki se uporabljajo za merjenje sopojavnosti, smo za primerjavo izbrali dva - Jaccardov (**J**) in ekvivalenčni (**E**) koeficient. Koeficienti so izračunani s spodaj navedenimi formulami, pri čemer **i** in **j** imenujemo zbirki; **A** - število zapisov, ki jih najdemo tako v zbirki **i**, kot v zbirki **j**; **B** - število zapisov, ki jih najdemo samo v zbirki **i** in **C** - število zapisov, ki jih najdemo samo v zbirki **j**.

Jaccardov koeficient: $J(i,j)=A/(A+B+C)$ in

Ekvivalenčni koeficient: $E(i,j)=A*A/((A+B)*(A+C))$

Preglednica 4: Podatki za izračun koeficientov sopojavljanja ter vrednosti posameznih koeficientov (Jaccardov- J, ekvivalenčni- E), za zbirki AGRIS in CAB (1986-1997)

Leto	AGRIS	CAB	Skupni zapisi	Koeficienti sopojavljanja	
				Jaccardov	Ekvivalenčni
1986	81	83	22	0,15	0,07
1987	55	84	23	0,20	0,11
1988	60	74	21	0,19	0,10
1989	65	60	19	0,18	0,09
1990	51	82	26	0,24	0,16
1991	54	53	23	0,27	0,18
1992	44	87	29	0,28	0,22
1993	52	69	22	0,22	0,13
1994	114	106	50	0,29	0,21
1995	44	77	22	0,22	0,14
1996	77	89	30	0,22	0,13
1997	53	122	32	0,22	0,16



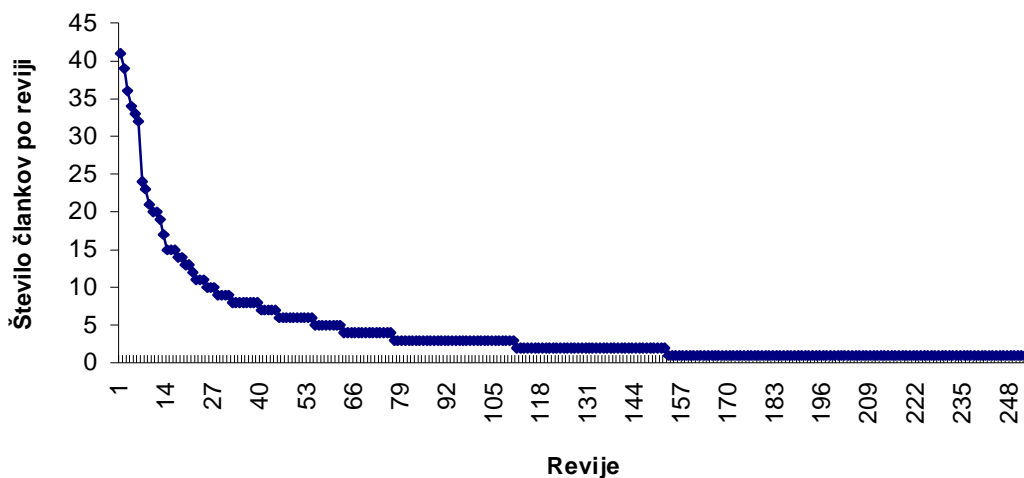
Slika 2: Gibanje koeficientov sopojavljanja (Jaccardov- J, ekvivalenčni- E) za zbirki AGRIS in CAB (1986-1997)

Ugotovimo lahko, da sta krivulji obeh koeficientov podobnih oblik, le vrednosti Jaccardovih koeficientov so nekoliko višje. Iz slike 1 je tudi razvidno, da izrazitejšega naraščanja v podvajanju omenjenih zbirk ni, govorimo lahko le o večjih ali manjših nihanjih med leti.

3.3 Razporejanje člankov po revijah in osrednje revije za področje jablanovega škrlupa

Članki se po revijah razporejajo v skladu z bibliometričnim načelom, po katerem zajame večji del člankov, ki se navezujejo na neko področje raziskovanja, že manjše število revij (Slika 3). S pomočjo Bradfordovega zakona, ki je eden temeljni indikatorjev razporejanja znanstvenih informacij, so bile določene tudi osrednje revije za področje jablanovega škrlupa. To so tiste revije, ki se v procesu znanstvenega komuniciranja na nekem področju raziskovanja največ uporabljajo, pri čemer pa razvrščanje v skupine ni odvisno od kakovosti prispevkov.

Slika 3: Razporeditev člankov po različnih naslovih revij



V analizo je bilo vključenih 1112 člankov (leto izida 1986-1997, podatki iz zbirk AGRIS, AGRICOLA, CAB), ki so objavljeni v 254 različnih revijah. Glede na število vključenih relevantnih člankov smo jih razvrstili po padajočem redu in jih razdelili na tri enakovredne skupine s približno enakim številom člankov po skupini.

Preglednica 5: Osrednje revije za področje jablanovega škrlupa (AGRIS, AGRICOLA, CAB; obdobje 1986-1997)

Število člankov	Naslov revije
41	Obstbau Weinbau
39	Acta Horticulturae
36	Obstbau
34	Bulletin OILB SROP
33	Phytopathology
32	Norwegian Journal of Agricultural Sciences
24	Plant Disease
23	Fruit Varieties Journal
21	Sadovodstvo i Vinogradarstvo
20	Erwerbsobstbau
20	Fungicide and nematocides tests: results (Am. Phytopath.Soc.)
19	Zashchita Rastenii
17	Mededelingen van de Faculteit Landbouw. Rijksuniversiteit Gent

Osrednje revije, ki v analiziranem obdobju vključujejo največ člankov za področje jablanovega škrlupa, so prikazane v preglednici 5. 13 prikazanih revij predstavlja 5% med vsemi 254 različnimi naslovi revij, vključuje pa tretjino (359) vseh relevantnih člankov. Druga skupina obsega 41 revij, za zadnjo tretjino člankov pa bi morali pregledati še dodatnih 200 različnih naslovov revij.

4 SKLEP

Prikazani rezultati bodo lahko v pomoč pri pridobivanju relevantnih informacij v okviru obsežne znanstvene in strokovne literature. Večji del dokumentov za področje jablanovega škrlupa je mogoče najti v dveh podatkovnih zbirkah - AGRIS in CAB. Prekrivanje omenjenih zbirk z leti bistveno ne narašča. Članki se razporejajo v skladu z bibliometričnim načelom, po katerem je glavnino relevantnih člankov mogoče najti že manjšem številu revij. Določitev osrednjih revij za določeno področje pa je lahko v pomoč tako uporabnikom, kot tudi knjižnicam pri smotrnejši nabavi gradiva.

5 LITERATURA

- Bartol, T. 1998. Vrednotenje biotehniških informacij o rastlinskih drogah v dostopnih virih v Sloveniji: doktorska disertacija. Ljubljana, Biotehniška fakulteta: 202 str.
- Bradford, S. C. 1934. Sources of information specific subjects. *Engineering*, 137, 85 (Reproduced: *Journal of information science*, 1985, 10: 173).
- Brier, S. 1996. Cybermiotics: a new interdisciplinary development applied to the problems of knowledge organisation and document retrieval in information science. *Journal of documentation*, 52, 3: 296-344.
- Burton, H. D. 1988. Use of virtual information system for bibliometric analysis. *Information processing & management*, 24, 1: 39-44.
- DataView: Bibliometric software for analysis of downloaded data. 2000.

- <http://crrm.u-3mrs.fr/commercial/software/dataview/dataview.html> (13.04. 2000)
- Garfield, E. 1994. Citation comments: Using the Impact Factor. *Current contents*, 29: 3-5.
- Kuhlthau, C. C. 1997. Inside the search process: information seeking from the user's perspective. *Journal of American Society for Information Science*, 42, 6: 232
- Laine-Cruzel, S., Lafouge, T., Lardy, J. P., Ben Abdallah, N. 1996. Improving information retrieval by combining user profile and document segmentation. *Information Processing & Management*, 32, 3: 305-315.
- Osareh, F. 1996. Bibliometrics, citation analysis and co-citation analysis. *Libri*, 46: 149-158.
- Reid, J. N. A. 1992. Agriculture, food and biosciences information from Silver Platter. *Aslib proceedings*, 44, 6: 245-249.
- Ungern-Sternberg, S. 1998. Application in teaching bibliometrics. 1st IFLA General Conference. *Proceedings*, 20-25 august, 1995. 1998.
<http://ifla.init.fr/iv/ifla61/61-ungs.htm>
- Zbirke na CD-romu: AGRIS, AGRICOLA, CAB, Biological Abstracts (BIOSIS), SCISEARCH.