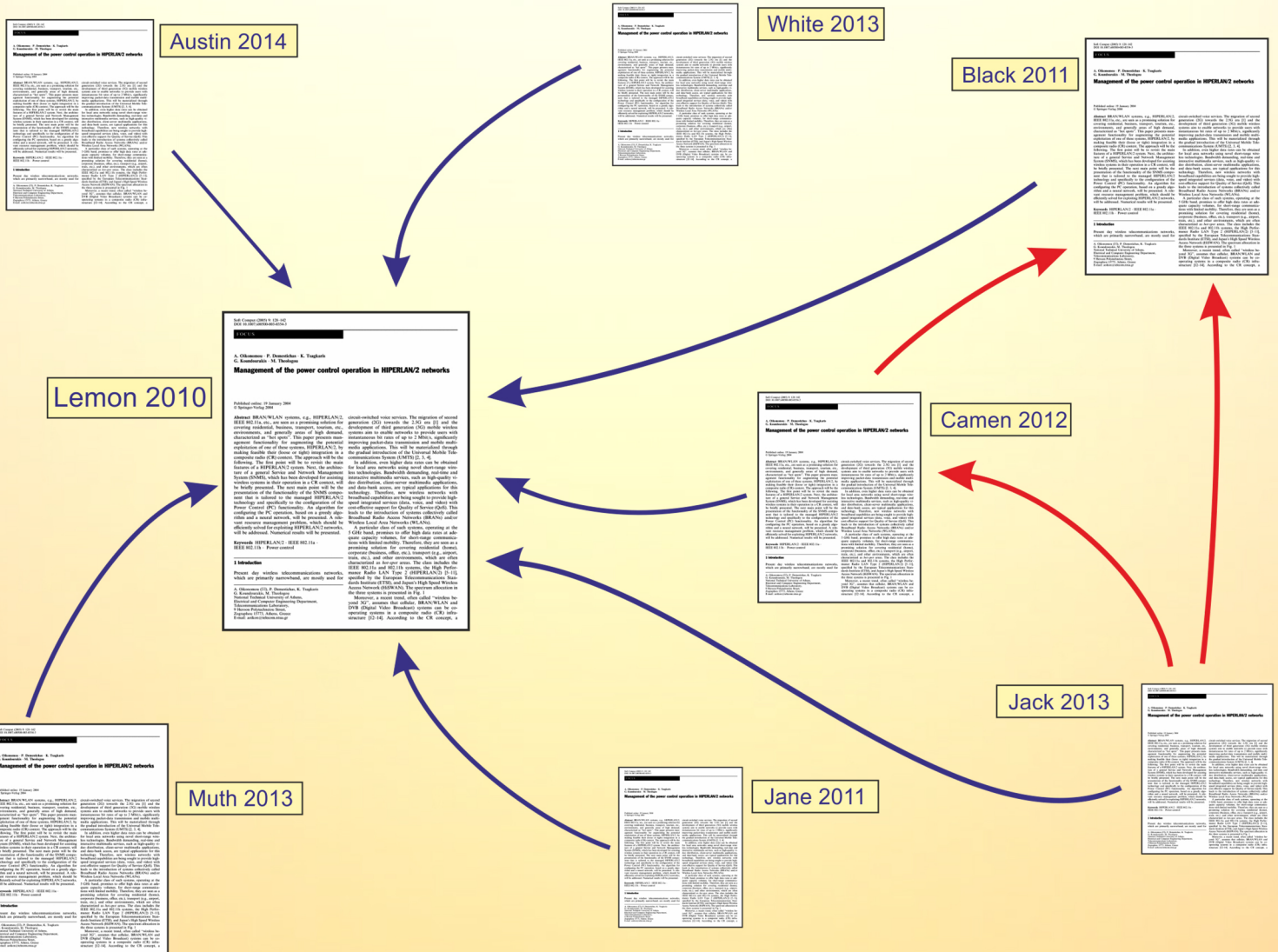


Σώζων Παπαβλασόπουλος

Βιβλιομετρία



Ελληνικά Ακαδημαϊκά Ηλεκτρονικά
Συγγράμματα και Βοηθήματα
www.kallipos.gr

HEALLINK
Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΣΩΖΩΝ ΠΑΠΑΒΛΑΣΟΠΟΥΛΟΣ

Επίκουρος Καθηγητής,
Ιόνιο Πανεπιστήμιο

Βιβλιομετρία



Ελληνικά Ακαδημαϊκά Ηλεκτρονικά
Συγγράμματα και Βοηθήματα
www.kallipos.gr

Βιβλιομετρία

Συγγραφή

Σώζων Παπαβλασόπουλος

Κριτικός αναγνώστης

Μάριος Πούλος

Συντελεστές έκδοσης

Γλωσσική Επιμέλεια: Σώζων Παπαβλασόπουλος

Γραφιστική Επιμέλεια: Σώζων Παπαβλασόπουλος

Τεχνική Επεξεργασία: Ευφροσύνη Αποστόλου

ISBN: 978-960-603-202-8

Copyright © ΣΕΑΒ, 2015



Το παρόν έργο αδειοδοτείται υπό τους όρους της άδειας Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Όχι Παράγωγα Έργα 3.0. Για να δείτε ένα αντίγραφο της άδειας αυτής επισκεφτείτε τον ιστότοπο

<https://creativecommons.org/licenses/by-nc-nd/3.0/gr/>

ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ

Εθνικό Μετσόβιο Πολυτεχνείο

Ηρώων Πολυτεχνείου 9, 15780 Ζωγράφου

www.kallipos.gr

Στον πατέρα μου

Πίνακας περιεχομένων

Πίνακας περιεχομένων.....	5
Πίνακας συντομεύσεων-ακρωνύμια	9
Πρόλογος.....	12
Εισαγωγή	13
Κεφάλαιο 1 Εισαγωγή: Ιστορικές σημειώσεις και θεωρητικές βάσεις της Βιβλιομετρίας	15
1.1 Θεωρητικές Βάσεις.....	15
Εφαρμογές της Βιβλιομετρίας	19
1.2 Οι πρωτοπόροι της Βιβλιομετρίας	19
1.2.1 Ιδρύματα κέντρων έρευνας και παροχής υπηρεσιών	25
1.2.2. Η εκπαιδευτική επικοινωνία.....	26
1.2.3. Η Βιβλιομετρία στη δεκαετία του 90 και στη νέα χιλιετία.....	26
1.3. Οι διάφορες μετρικές.....	29
1.4. Διάκριση Βιβλιομετρίας	34
1.5. Περιορισμοί της Βιβλιομετρίας	35
1.5.1. Βιβλιομετρία και επιστημονικοί κλάδοι	36
1.5.2. Μέσα επιστημονικής επικοινωνίας.....	36
1.5.3. Γλώσσα και τοπικό ενδιαφέρον.	40
1.5.4. Πρακτικές αναφορών	45
1.6. Η Βιβλιομετρία και η μέτρηση της επιστήμης.....	47
1.7 Βιβλιογραφία -βιβλιογραφικές αναφορές και παραπομπές	49
1.7.1 Αναγκαιότητα χρήσης βιβλιογραφίας και αναφορών	49
1.7.2 Διαφορά βιβλιογραφικής αναφοράς και παραπομπής	51
1.8. Διπλώματα ευρεσιτεχνίας (Patents).....	54
Βιβλιογραφία/Αναφορές	55
Κεφάλαιο 2 Εργαλεία καταγραφής επιστημονικών αναφορών και μέτρησης.....	60
2.1. Επιλέγοντας μια μονάδα ανάλυσης.....	60
2.2. Παράγοντες Επηρεασμού της Βιβλιομετρικής Ανάλυσης	61
2.2.1. Επιστημονικά πεδία.....	61
2.2.2. Χρονικό διάστημα ανάλυσης αναφορών	61
2.2.3. Είδος επιστημονικών δημοσιεύσεων	61
2.3. Πώς επηρεάζονται οι αναφορές	62
2.3.1. Η ηλικία της δημοσίευσης και μέσο ετήσιο ποσοστό παραπομπών διαφόρων τομέων	62
2.3.2.Ο τύπος της δημοσίευσης.....	62
2.3.3.Το μήκος της περιόδου παρατήρησης.....	63
2.4. Πού θα βρούμε τις αναφορές.....	63

2.4.1. Web of Science (WoS).....	63
2.4.1.1. Το Ευρετήριο Επιστημονικών Αναφορών (Science Citation Index-SCI)	64
2.4.1.2. Το Ευρετήριο Επιστημονικών Αναφορών και οι περιορισμοί του	66
2.4.1.3. Ο παράγοντας της γλώσσας	66
2.4.1.4. Η κατανομή ανά επιστημονικό κλάδο	67
2.4.2. Scopus	67
2.4.2.1. SCImago Journal and Country Rank.	68
2.4.3. Google Scholar	68
2.4.3.1. Google Scholar Citations.....	68
2.4.3.2. Πώς να βρείτε πόσες αναφορές υπάρχουν σε συγκεκριμένο άρθρο	69
2.4.4. Academic Search.....	69
2.4.5. ACM Digital Library (Association for Computing Machinery)	69
2.4.5.1. Οδηγίες ανάκτησης αναφορών	69
2.4.6. IEEEExplore	69
2.4.6.1.Οδηγίες χρήσης	70
2.4.7. JSTOR (Journal Storage).....	70
2.4.7.1.Οδηγίες χρήσης	70
2.4.8. PubMed / Medline.....	70
2.4.8.1. Οδηγίες χρήσης	71
2.4.9. Academic Search Premier	71
2.4.10. Άλλες βάσεις δεδομένων.....	71
2.4.11. Ειδικές βάσεις δεδομένων	71
2.4.12. Προβλήματα που τίθενται εν γένει από τις βάσεις δεδομένων	72
2.5. Τι μετρούν οι βιβλιομετρικές αναλύσεις;	73
2.6. Χρήση των βιβλιομετρικών δεικτών και οι προφυλάξεις που θα πρέπει να ληφθούν	73
2.6.1.Εισαγωγή	73
2.7. Τι επιφυλάσσει το μέλλον για τους βιβλιομετρικούς δείκτες;	77
2.8. Εργαλεία αναφορών και μέτρησης	77
2.8.1 Εργαλεία	77
2.8.2. Ανάλυση αναφορών	78
2.9. Ανάλυση Περιεχομένου (Content Analysis).....	88
2.10. Βιβλιομετρικοί Χάρτες (Bibliometric Maps).....	89
Βιβλιογραφία/Αναφορές	91
Κεφάλαιο 3 Πρακτικός οδηγός κατασκευής και χρήσης Βιβλιομετρικών Δεικτών και Νόμων	93
Εισαγωγή	93
3.1 Δείκτες μέτρησης και αξιολόγησης περιοδικών	93
3.1.1 Η υπηρεσία Journal Citation Reports.....	93
3.1.1.1 Παράδειγμα εφαρμογής και Σχόλια σχετικά με τον Δείκτη	95

3.1.2. Προβλήματα εφαρμογής του Δείκτη επιρροής.....	95
3.1.3. Αντιμετώπιση των προβλημάτων από την εφαρμογή του Δείκτη	97
3.1.3.1. Υ δείκτης	97
3.1.4. Βαθμολογία Eigenfactor	97
3.1.5. Βαθμολογία Επίδρασης του Άρθρου (Article Influence Score)	97
3.1.6. Ο Αναλυτής Περιοδικού (Journal Analyzer).....	98
3.1.7. Δείκτες Elsevier	98
3.1.8. Μετρήσεις Περιοδικών μέσω της υπηρεσίας Google Scholar	101
3.2. Παραδείγματα χρήσης συνδυασμού δεικτών επιρροής περιοδικών.....	102
3.2.1. Προβλήματα και Λύσεις.....	102
3.3. Οι Βιβλιομετρικοί δείκτες: Ένας πρακτικός οδηγός κατασκευής και χρήσης για ερευνητές.....	108
3.3.1. h-index	108
3.3.2. G -index.....	109
3.4. Χρησιμοποίηση Δεδομένων για Παραγωγή Δεικτών.....	110
3.5. Εξειδικευμένα παραδείγματα:	111
3.6. Σφάλματα Στρέβλωσης και Αδυναμίες Μέτρησης Αναφορών	112
3.7 Βιβλιομετρικοί Νόμοι: εφαρμογή σε εμπειρικά σύνολα δεδομένων	118
3.7.1. Ο Νόμος του Zipf	118
3.7.2. Ο Νόμος του Samuel Clement Bradford 1934.....	119
3.7.3. Νόμος του Lotka	120
Βιβλιογραφία	122
Κεφάλαιο 4.....	124
ΕΙΣΑΓΩΓΙΚΑ.....	124
4.1. Ανοικτή Πρόσβαση (Open Access).....	128
4.2. Βιβλιομετρικοί μέθοδοι στον Ιστό.....	144
4.2.1 Ορολογία.....	145
4.2.1.1 Βασική ορολογία συνδέσμων	145
4.2.1.2 Βασική ορολογία web Κόμβων και διαγράμματα	146
4.3 Εναλλακτικά μοντέλα εγγράφων	148
4.4 Έρευνα και συλλογή δεδομένων του Παγκόσμιου Ιστού	149
4.4.1 Προσωπικοί Web Crawlers.....	149
4.4.2 Εμπορικές μηχανές αναζήτησης	150
4.4.3 Web Log Files.....	151
4.4.4 Χρήση δειγματοληπτικών μεθόδων	152
4.4.5 Ανάκτηση σελίδων δειγματοληπτικά	152
4.4.6 Ανάκτηση δικτυακών τόπων δειγματοληπτικά.....	153
4.4.7 Μελέτες σε μεγάλα τμήματα του παγκοσμίου ιστού.....	153
4.5 Ανάλυση Βιβλιομετρική στο παγκόσμιο ιστό	153
4.5.1 Ανάλυση βιβλιογραφικών αναφορών (citation analysis)	153

4.5.2 Ανάλυση ζευγών λέξεων και αναφορών (co-word and co-citation analysis).....	154
4.5.3 Ανάλυση περιεχομένου (content analysis).....	154
4.5.4 Αξιολογώντας με τη χρησιμοποίηση είτε υπαρχόντων ή νέων μεθόδων	154
4.5.6 Υπολογισμός και επεξήγηση των δεικτών	154
4.6 Μοντέλα βιβλιομετρικής ανάλυσης στο Παγκόσμιο Ιστό	155
4.6.1 Τοπολογικά μοντέλα και εξόρυξη στο Παγκόσμιο Ιστό (web mining).....	155
4.6.2 Μοντέλα σύνδεσης.....	155
4.6.3 Δίκτυα “Small-Worlds”.....	156
4.6.4 Ανάλυση ακολουθούμενης διαδρομής (path analysis).....	159
4.6.5 Τοπολογική μοντελοποίηση και εξόρυξη του Web	159
4.6.6 Μέτρηση του Web 2.0	159
4.6.7 Χρησιμοποιώντας εξυπηρετητές Παγκόσμιου Ιστού.....	160
4.7 Σχεδιάζοντας Βιβλιομετρία	161
4.7.1 Βέλτιστη Διαμόρφωση Εξυπηρετητή Παγκόσμιου Ιστού	161
4.7.2 Διαχείριση αρχείων καταγραφής.....	162
4.7.3 Ο J. S. Downie (Καθηγητής στο Πανεπιστήμιο του Illinois) και η χρήση του Παγκόσμιου Ιστού.....	162
4.7.4 Βέλτιστη παραμετροποίηση περιεχομένου παγκόσμιου ιστού	163
4.7.5 Συλλογή Βιβλιομετρικών δεδομένων	163
4.8 Altmetrics	163
4.8.1 Πηγές Altmetrics.....	164
4.8.2 Επιστημονικά δίκτυα Altmetrics.....	164
4.8.3 Τα Blogs και τα μέσα ενημέρωσης Altmetrics	165
4.8.4 Υποστήριξη της έρευνας Altmetrics.....	165
Βιβλιογραφία	167
Διαδικτυογραφία.....	171
ΕΥΡΕΤΗΡΙΟ	172
Βιβλιογραφία/Αναφορές	174
Δικτυογραφία	183

Πίνακας συντομεύσεων-ακρωνύμια

ACI	Autonomous Citation Indexing Αυτόματο ευρετήριο αναφορών της βιβλιογραφίας
ACM	Association for Computing Machinery Ένωση Μηχανικών Πληροφορικής
ADM	Alternative Document Model Εναλλακτικό Μοντέλο Κειμένου
AHCI	Arts and Humanities Citation Index
AIDS	Σύνδρομο Επίκτητης Ανοσοανεπάρκειας
Altmetric	Alternative metrics Εναλλακτικές μετρικές
ANT	Actor Network Theory Θεωρία του δικτύου των δραστήων
APA	American Psychological Association
BOF	Bijzonder Onderzoeks fonds
CAS	Chemical Abstracts Service
CERN	Ευρωπαϊκό Εργαστήριο Σωματιδιακής Φυσικής
CGI	Common Gateway Interface
CINDOC	Centre of scientific information and Documentation
CNER	Conseil National d'évaluation de la Recherche Εθνικό Συμβούλιο Έρευνας Αξιολόγησης
CORDIS	Community Research and Development Information Service
CPOT	Citation Potential Δυναμικό αναφοράς
CSA	Constrained Spreading Activation
CWTS	Centre for Science and Technology Studies
DOI's	Digital Object Identifier Ψηφιακά αναγνωριστικά αντικειμένου
EBSCO	Elton B. Stephens Co
ECOOM	Expertisecentrum Onderzoek en Ontwikkelingsmonitoring Centre for Research & Development Monitoring
ESSS	European Summer School for Scientometrics
FTP	FILE TRANSFER PROTOCOL
HITS	Hypertext-Induced topic search
HTML	HyperText Markup Language Γλώσσα Σήμανσης Υπερκειμένου
IDIS	Institut für Dokumentation und Information über Sozial medizin und öffentliches Gesundheitswesen
INIST	Ινστιτούτο Επιστημονικών και Τεχνικών Πληροφοριών Digital Bibliography & Library Project
IoP	Institute of Physics
ISBN	International Standard Book Number Διεθνής μοναδικός αριθμός βιβλίου

ISI	Thomson Institute for Scientific Information Ινστιτούτο Thomson για την επιστημονική πληροφορία,
ISSI	International Society for Scientometrics and Informetrics Διεθνής εταιρία της Επιστημομετρίας και της Πληροφοριομετρίας
JASIS	Journal of the American Society for Information Science
JCI	Journal of Clinical Investigation Περιοδικό κλινικών Ερευνών
JCR	Journal Citation Reports
JCR	Journal Citation Reports
JSTOR	Journal Storage
LIS	Library and Information Science
MathSciNet	Mathematical Sciences Network
MeSH	Medical Subject Headings Ιατρικές Θεματικές Επικεφαλίδες
MLA	Modern Language Association
NASA	National Aeronautics and Space Administration Εθνική Υπηρεσία Αεροναυπηγικής και Διαστήματος
NLM	National Library of Medicine
NORSLIS	Nordic Research School in Library and Information Science
NOWT	Ολλανδικό Παρατηρητήριο Επιστήμης και Τεχνολογίας
NSAPI	Network Service Access Point Identifier
NSE	Natural Science and Engineering Φυσικές και Τεχνολογικές Επιστήμες
NSF	Εθνικό Ίδρυμα Επιστημών
OAI	Open Archive Initiative
OST	Observatoire des Sciences et des Techniques
OST	Observatoire des sciences et des technique Παρατηρητήριο Επιστήμης και Τεχνικής
PBRF	Performance-Based Research Fund Απόδοση χρηματοδότηση της Έρευνας
PLoS	Public Library of Science
PoP	Δημοσίευση ή Φθορά Publish or Perish
PRF	Pseudo Random Function
PROLA	Physical Review Online Archive
PsycINFO	American Psychological Association
R&D	Research and Development
RAE	Research Assessment Exercise
REF	Research Excellence Framework
RIP	Row ImPact Μέση επίδραση ανά άρθρο του περιοδικού
RN	Research Networking Ερευνητική δικτύωση

RQF	Research Quality Framework
RSS	Rich Site Summary Σύνοψη Πλούσιας Σελίδας
SAHCI	Social-Arts-and Humanities Science Citations Indexes Δείκτες αναφορών Κοινωνικών, Ανθρωπιστικών και Τεχνών
SCC	Strongly Connected Component
SCI	Science Citations Indexes Δείκτης αναφορών επιστημών
SEI	Science and Engineering Indicators Δείκτες Επιστήμης και Τεχνολογίας
SJR	SCImago Journal Rank
SNIP	Source Normalized Impact per Paper Κανονικοποιημένος Δείκτης Απήχησης
SPIN	Searchable Physics Information Notices
SSCI	Social Science Citation Index Δείκτης αναφορών κοινωνικών επιστημών
SSH	Social Sciences and Humanities Κοινωνικές και Ανθρωπιστικές Επιστήμες
SSRN	Social Science Research Network
URL	Uniform Resource Locator Ενιαίος Εντοπιστής Πόρων
VINITI	All-Soviet Institute for Scientific and Technical Information Ινστιτούτο Επιστημονικής και Τεχνικής Πληροφόρησης
VQR	Valutazione Quinquennale della Ricerca
VTR	Valutazione Triennale della Ricerca
WIF	Web Impact Factor Διαδικτυακός παράγοντας επιρροής
WIF	Web Impact Factor Δείκτης απήχησης Παγκοσμίου Ιστού
WoS	Web of Science
WWW	World Wide Web
Blog	Weblog
BEΠ	Βιβλιοθήκη και Επιστήμη της Πληροφορίας
ΔΑΕ	Διαδικασία αξιολόγησης της έρευνας
E&A	Έρευνα και ανάπτυξη
EXE	Excellence in Research for Australia
IEEE	Institute of Electrical and Electronics Engineers Ινστιτούτο Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών
MIME	Multipurpose Internet Mail Extensions
ΟΟΣΑ OECD	Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης
ΠΕΑ	Πλαίσιο της ερευνητικής αριστείας
ΠΟΥ WHO	Παγκόσμιος Οργανισμός Υγείας
Timeout	Μέθοδος timeout, ορισμός χρονικού ορίου στην λειτουργία
YXT	Υπηρεσία Χημικής Ταυτοποίησης

Πρόλογος

Η εκτίμηση, η μέτρηση, η έρευνα και η αξιολόγηση αποτελούν μέρος των διαδικασιών της επιστήμης, με τα ζητήματα που σχετίζονται με αυτές τις θεματικές ενότητες συχνά να επικαλύπτονται. Η εκτίμηση αναφέρεται στη συλλογή των δεδομένων, για να περιγραφεί ή να κατανοηθεί καλύτερα το υπό εξέταση φαινόμενο. Η μέτρηση είναι η διαδικασία της ποσοτικοποίησης των εκτιμώμενων δεδομένων. Η έρευνα αναφέρεται στη χρήση των δεδομένων για τους σκοπούς της περιγραφής, της πρόβλεψης και του έλεγχου, ως το κατ' εξοχήν μέσο προς την καλύτερη κατανόηση των υπό εξέταση φαινομένων. Η αξιολόγηση, τέλος, αναφέρεται στη σύγκριση των δεδομένων με ένα πρότυπο, με σκοπό τον έλεγχο της αξίας και της ποιότητας τους.

Οι σημειώσεις αυτές έχουν σαν στόχο να δώσουν στους αναγνώστες, με μια απλή και προσιτή γραφή, την παρουσίαση των βασικών εννοιών, τις τεχνικές θεωρητικές θέσεις, τα εργαλεία της έρευνας, τις ιστορικές εξελίξεις, τις πηγές των δεδομένων, καθώς και τις σύγχρονες προεκτάσεις του υποπεδίου της επιστήμης των πληροφοριών, που ασχολείται με την ποσοτική ανάλυση της επιστημονικής και τεχνολογικής βιβλιογραφίας. Προς το παρόν, πριν από τη διασαφήνιση των όρων, ας ονομάσουμε αυτήν την περιοχή "Βιβλιομετρία".

Το παρόν βιβλίο απευθύνεται στους φοιτητές των Τμημάτων των Επιστημών Πληροφορίας και Βιβλιοθηκονομίας. Σκέψεις και ερωτήματα για το τι ακριβώς είναι η Βιβλιομετρία, τι μετράει η βιβλιομετρική ανάλυση, πώς διεξάγουμε μια βιβλιομετρική ανάλυση, για ποιους ακριβώς λόγους είναι σημαντική, ποια η συνεισφορά της, ποια η μετεξέλιξή της κ.ά. θα διερευνηθούν εξονυχιστικά με σκοπό τη χρησιμοποίηση των εργαλείων της, ώστε να δίδουν αξιόπιστα και επιστημονικά αποδεκτά αποτελέσματα, τόσο σε χαμηλό επίπεδο (π.χ. επιστημονικές ή ερευνητικές ομάδες), όσο και σε μεσαίο επίπεδο (π.χ. ερευνητικά ιδρύματα, περιοδικά) ή και υψηλό επίπεδο (υπέρ - εθνικές περιφέρειες, χώρες ή μεγάλα επιστημονικά πεδία).

Το εγχειρίδιο που θα ακολουθήσει θα χρησιμοποιεί, εκτός των άλλων, αρκετά συστηματικά, αν και όχι αποκλειστικά, ιστορικά έγγραφα και στοιχεία, αλλά αυτό δεν αποτελεί σε καμία περίπτωση την πλήρη ή εμβριθή ιστορία της Βιβλιομετρίας.

Κέρκυρα, Σεπτέμβρης 2015

Εισαγωγή

Η ανάλυση του σώματος των τεκμηρίων μπορεί να ταξινομηθεί γενικά ως αποτελούμενη από ποιοτική και ποσοτική ανάλυση. Τα τελευταία χρόνια οι ποιοτικές μέθοδοι για την αποτίμηση του ερευνητικού έργου παρουσιάζουν ύφεση, η βασικότερη μάλιστα από αυτές, η «αξιολόγηση από ομοτίμους» (peer review), έχει θεωρηθεί μη αποτελεσματική. Δεν είναι τυχαίο πως η περίφημη RAE (Research Assessment Exercise), η διαδικασία δηλαδή αξιολόγησης της έρευνας, που πραγματοποιείται περίπου κάθε πέντε (5) χρόνια για λογαριασμό τεσσάρων (4) συμβουλίων χρηματοδότησης της τριτοβάθμιας εκπαίδευσης στο Ηνωμένο Βασίλειο, έπαψε να λειτουργεί. Την αντικατέστησε πρόσφατα, το 2014, το REF (Research Excellence Framework), το πλαίσιο της ερευνητικής αριστείας, το οποίο βασίζεται εφεξής, κυρίως, σε ένα σύστημα ποσοτικών (μετρικών) μεθόδων αξιολόγησης, στηριζόμενο στη βιβλιομετρική ανάλυση. Η βιβλιομετρική ανάλυση ασχολείται με την καταγραφή και επεξεργασία των δεδομένων που σχετίζονται με τις επιστημονικές δημοσιεύσεις και την εξαγωγή των σχετικών δεικτών, όπως οι αναφορές των δημοσιεύσεων, από άλλες δημοσιεύσεις, ο συνδυασμός τους με συγκεκριμένα επιστημονικά πεδία φορείς, κ.λ.π.

Οι μέθοδοι αξιολόγησης εφαρμόζονται με συγκεκριμένες τεχνικές, βασισμένες στα παραπάνω αντικειμενικά δεδομένα δημοσιεύσεων και βιβλιογραφικών αναφορών επιστημονικών άρθρων, μονογραφιών, τεχνικών αναφορών κ.ά. Η αξιολόγηση της ερευνητικής επίδοσης επιτυγχάνεται με τη χρήση των βιβλιομετρικών δεικτών, που παράγονται με τη χρήση των παραπάνω τεχνικών. Η κατάταξη που θα προκύψει από τη χρήση των βιβλιομετρικών δεικτών θα χρησιμοποιηθεί για την κατανομή της χρηματοδότησης της έρευνας κάθε ιδρύματος τριτοβάθμιας εκπαίδευσης, από το αντίστοιχο ανώτατο συμβούλιο τους. Οι μελέτες που στηρίζονται στη βιβλιομετρική ανάλυση αυξάνονται με ταχείς ρυθμούς τα τελευταία χρόνια στο διεθνές, αλλά και στον ελληνικό χώρο. Οι βιβλιομετρικοί δείκτες μέτρησης της ερευνητικής δραστηριότητας συμβάλλουν στη διαμόρφωση μιας, κατά το δυνατόν, αντικειμενικά μετρήσιμης εικόνας των συστημάτων έρευνας, τεχνολογικής ανάπτυξης και καινοτομίας. Εκτός από τη χρησιμότητα τους για την αξιολόγηση ερευνητικών οργανισμών, ομάδων και ερευνητών, καταγράφουν τα ερευνητικά πεδία στα οποία δραστηριοποιείται η επιστημονική κοινότητα, αποτυπώνουν τα νέα ερευνητικά πεδία που αναδύονται, καθώς και τα επιστημονικά δίκτυα που δημιουργούνται για την υλοποίηση ερευνητικών συνεργασιών, επιστημονικών δραστηριοτήτων των ερευνητικών ομάδων με κοινούς ερευνητικούς στόχους.

Με τη χρήση αυτών των δεικτών μπορούν να διερευνηθούν και άλλες πτυχές της έρευνας, όπως οι προσανατολισμοί της έρευνας με χαρακτηριστικά εφαρμοσμένης ή θεωρητικής κ.ά.

Η διερεύνηση των σχέσεων μπορεί να επεκταθεί, επίσης, μεταξύ επιστημονικών περιοδικών, συνεδρίων, δημοσιεύσεων, συγγραφέων, επιστημονικών κλάδων, παραγόμενου ερευνητικού έργου συγκεκριμένων χωρών, καθώς επίσης είναι δυνατός ο προσδιορισμός χαρακτηριστικών και τάσεων της ερευνητικής παραγωγής σε επίπεδο φορέα, χώρας ή ευρύτερου συνόλου χωρών, η εκτίμηση της απήχησης του επιστημονικού έργου, η αξιολόγηση της ερευνητικής δραστηριότητας και ο εντοπισμός εθνικών και πολυεθνικών δικτύων μεταξύ επιστημόνων και επιστημονικών κλάδων. Τόσο η μέθοδος αξιολόγησης από ομότιμους, όσο και οι μέθοδοι της Βιβλιομετρίας (της εφαρμογής δηλαδή των ποσοτικών τεχνικών έρευνας και της στατιστικής στα δημοσιεύματα ενός τομέα ή του συνόλου της βιβλιογραφίας) παρουσιάζουν θετικά και αρνητικά αποτελέσματα και, ως εκ τούτου, έχουν τους υποστηρικτές και τους πολέμιους τους στην ερευνητική κοινότητα.

Προκλήσεις όπως:

- **Η γιγαντιαία ανάπτυξη της επιστημονικής βιβλιογραφίας.**
- **Η βοήθεια που απαιτείται σε περίπτωση υπερφόρτωσης των πληροφοριών.**
- **Η αυξανόμενη πολυπλοκότητα στην επιστημονική επικοινωνία.**
- **Η αλλαγή από τη «μικρή επιστήμη» με την εκθετική ανάπτυξη στη «μεγάλη επιστήμη».**
- **Η ανταγωνιστική πρόσβαση του ανθρώπινου δυναμικού.**
- **Η κρίση στο σύστημα αξιολόγησης.**
- **Η πληρότητα των δεδομένων εναντίον της αντιπροσωπευτικότητας και συνάφειας.**
- **Η ποικιλία κάλυψης των δεδομένων από διαφορετικές βάσεις δεδομένων.**
- **Τα προβλήματα αναγνώρισης των συγγραφέων, ιδρυμάτων, περιοδικών (αναδιοργάνωση, συγχωνεύσεις, τίτλοι κ.λ.π.), όπως και**

- **Τα προβλήματα σε θέματα ταξινόμησης.**

ζητούν άμεση προσέγγιση και λύσεις, αφού η Βιβλιομετρία μπορεί να παρέχει χρήσιμες πληροφορίες σε τομείς εφαρμογής της.

Κεφάλαιο 1

Εισαγωγή: Ιστορικές σημειώσεις και θεωρητικές βάσεις της Βιβλιομετρίας

Σύνοψη

Σε αυτό το κεφάλαιο δίδονται οι θεωρητικές βάσεις, διασαφηνίζονται οι βασικές έννοιες και θεμελιώνεται ο σκοπός της Βιβλιομετρίας καθώς και η χρήση της, ως ένα εργαλείο διαχείρισης της έρευνας με εξελιγμένες τεχνικές, οι οποίες θα συζητηθούν επιπροσθέτως με τους ορισμούς των διάφορων μετρικών: Βιβλιο-Επιστημο-Πληροφοριο-Ιστομετρίας που θα επεξηγηθούν πλήρως.

Η Βιβλιομετρία βασίζεται σε δύο παραδοχές: (Α) Ο στόχος των ερευνητών είναι να προάγουν τη γνώση με τη διάδοση των αποτελεσμάτων της έρευνας και των μελετών τους μέσα από μια ποικιλία μέσων επικοινωνίας, συμπεριλαμβανομένης της γραφής, η οποία βρίσκεται στον πυρήνα της ακαδημαϊκής παράδοσης και (Β) Οι μελετητές πρέπει να δημοσιεύουν, προκειμένου να αποκτήσουν φήμη και να εξελίξουν τη σταδιοδρομία τους.

Η δημοσίευση μιας αναφοράς είναι ένας τρόπος μέτρησης και σύγκρισης της παραγωγής των διαφόρων μακροοικονομικών μεγεθών, όπως άτομα, ιδρύματα, περιφέρειες και χώρες. Μπορεί επίσης να χρησιμοποιηθεί για την αξιολόγηση της παραγωγής σε επιμέρους κλάδους, όπως π.χ. η φιλοσοφία και η οικονομία, καθώς επίσης και για την παρακολούθηση των τάσεων σε τομείς έρευνας, συνεργατικής έρευνας και σε πολλές άλλες πτυχές των αποτελεσμάτων της έρευνας

1.1 Θεωρητικές Βάσεις

Τι είναι η αναφορά

Ο πρωταγωνιστής της ιστορίας μας, κατά μια έννοια το νήμα της Αριάδνης ολόκληρης της επισκόπησης, είναι η βιβλιογραφική αναφορά. Αυτό υπαγορεύεται από την επιλογή μιας αυστηρώς καθορισμένης άποψης για το είδος των πληροφοριών που μετριοούνται και το σκοπό για τον οποίον αυτό γίνεται, δηλαδή τον βαθμό παροχής πολύτιμων πληροφοριών για τον χρήστη, που θα μπορούσε να αναπτύξει ένα σύνολο μεθόδων και αλγορίθμων για τη μέτρηση του βαθμού της ικανότητας πληροφόρησης από ένα τεκμήριο.

Η λέξη αναφορά προέρχεται από τη λέξη *referren* της Μέσης Αγγλικής διαλέκτου, από τη λέξη *referer* της Μέσης Γαλλικής διαλέκτου, από τη λατινική *referre* "να φέρει πίσω", σχηματιζόμενη από το πρόθεμα *re* και τη λέξη *ferre* «να φέρει». Ένας μεγάλος αριθμός από λέξεις που προέρχονται από αυτή τη ρίζα, συμπεριλαμβανομένων των διαιτητής, αναφορά, δημοψήφισμα, διατηρούν το βασικό νόημα της αρχικής Λατινικής ως "το σημείο, τον τόπο ή την πηγή προέλευσης". Σύμφωνα με τη MedLibrary, ο διαιτητής είναι ο πάροχος αυτής της πηγής προέλευσης και το αντικείμενο αναφοράς είναι ο κάτοχος της πηγής προέλευσης, είτε πρόκειται για γνώση, ύλη ή ενέργεια. Λόγω αυτής της σημασίας, η λέξη αναφορά χρησιμοποιείται σε κάθε σφαίρα της ανθρώπινης γνώσης, υιοθετώντας αποχρώσεις της σημασίας, ανάλογα με τα συμφραζόμενα μέσα στα οποία χρησιμοποιείται. Οι αναφορές μπορεί να πάρουν πολλές μορφές, συμπεριλαμβανομένων μια σκέψη, μια αισθητηριακή αντίληψη που μπορεί να είναι ακουστική (ονοματοποιία), οπτική (κείμενο), οσφρητική ή αφή, συναισθηματική κατάσταση, σχέση με άλλους, χωροχρονικές συντεταγμένες, συμβολική ή αλφαριθμητική παράσταση, φυσικό αντικείμενο ή μια προβολή ενέργειας. Ωστόσο, άλλα συγκεκριμένα και αφηρημένα πλαίσια υπάρχουν ως μέθοδοι καθορισμού των αναφορών εντός του πεδίου εφαρμογής διάφορων τομέων που απαιτούν μια προέλευση, το σημείο εκκίνησης ή την αρχική μορφή. Αυτό περιλαμβάνει μεθόδους που σκόπιμα κρύβουν την αναφορά από ορισμένους παρατηρητές, όπως στην κρυπτογραφία. Οι αναφορές μετρώνται για να βρούμε πρότυπα διαφορετικής χρήσης, όπως ο συγγραφέας, η χρονολογία, η γεωγραφία, το αντικείμενο, έντυπα κ.τ.λ. στην Βιβλιοθήκονομία και Επιστήμη της Πληροφορίας (Library Information Science).

Σημαντικό μέρος των ερευνητικών εργασιών, ιδιαίτερα στην επιστήμη, είναι ο κατάλογος των αναφορών που υποδεικνύουν προηγούμενες δημοσιεύσεις. Ο Ziman (1968), ορθώς, ανέφερε «μια επιστημονική εργασία δεν στέκεται μόνη της, είναι ενσωματωμένη στη βιβλιογραφία του αντικειμένου». Ομοίως ο Nann (1976) ορίζει πως «μια αναφορά είναι η αναγνώριση ότι ένα έγγραφο δίνει σε ένα άλλο, η παραπομπή είναι η αναγνώριση του γεγονότος ότι ένα έγγραφο λαμβάνει από ένα άλλο». Ο Malin (1968) σημειώνει πως «μια αναφορά υποδηλώνει μια σχέση μεταξύ ενός μέρους ή του συνόλου του αναφερόμενου

εγγράφου και ενός μέρους ή του συνόλου του παραπεμπόμενου έγγραφου». Από αυτές τις δηλώσεις είναι σαφές ότι η αναφορά είναι σημαντική κατά τη δημοσίευση επιστημονικών ή ερευνητικών ανακοινώσεων. Είναι απαραίτητο να αναφέρεται ο συγγραφέας του οποίου τα δεδομένα χρησιμοποιούνται. Η ερευνητική δραστηριότητα βασίζεται στις επικαλούμενες εργασίες και στη χρήση προηγούμενης γνώσης. Η χρήση αναφορών και η μελέτη τους αποκαλύπτει πολλές έννοιες χρήσιμες για την ορθή ανάπτυξη των βιβλιοθηκών. Η ανάλυση των αναφορών είναι η περιοχή της βιβλιομετρίας που ασχολείται με τη μελέτη των σχέσεων μεταξύ τους, η οποία μπορεί να είναι χρήσιμη για τη γεφύρωση της έρευνας. Τέτοιες μελέτες είναι απαραίτητες για την παρακολούθηση της ακαδημαϊκής ανάπτυξης σε κάθε θεματικό πεδίο.

Κάθε φορά που ένας επιστημονικός τομέας φτάνει σε ένα στάδιο που απαιτεί την υποστήριξη στατιστικών μεθόδων, μία μετρική αναδύεται από τον συγκεκριμένο τομέα. Αυτό ισχύει επίσης και για την επιστήμη της πληροφορίας, όπου υπήρχε η ανάγκη χρήσης στατιστικών μεθόδων για τη μέτρηση και την μοντελοποίηση συγκεκριμένων φαινόμενων των σύγχρονων επιστημονικών ανακοινώσεων.

Εννοιολογικά και θεωρητικά θέματα

Τα εννοιολογικά, καθώς και τα θεωρητικά θεμέλια της επιστήμης της Βιβλιομετρίας εδράζονται στις ακόλουθες υποθέσεις:

- **Η Βιβλιομετρία έχει τις επιστημολογικές ρίζες της στη Βιβλιογραφία.**
- **Η Βιβλιομετρία, ως εφαρμοσμένη Βιβλιογραφία, αποτελεί έναν παράγοντα μέτρησης και αξιολόγησης των πηγών πληροφόρησης.**
- **Η Βιβλιομετρία αποτελείται από μια μέθοδο ή από μια σειρά μεθόδων που μπορούν να στρατευθούν για την αξιολόγηση της έρευνας.**

Αυτές οι τρεις υποθέσεις συναποτελούν τους κύριους πυλώνες, που είναι αναγκαίοι για τη διατύπωση της έννοιας της Βιβλιομετρίας.

Από τις απαρχές, λοιπόν, μέχρι σήμερα η Βιβλιομετρία συνδέεται στενά με τη Βιβλιογραφία και με τις Πηγές Πληροφόρησης. Επιπλέον αυτή η σύνδεση είναι μία σχέση εξάρτησης, εφόσον οι βιβλιομετρικές μελέτες που μετρούν την επιστημονική παραγωγή διεξάγονται με βάση τα αποτελέσματα, τα οποία προέρχονται από τις αναλύσεις των πηγών πληροφόρησης. Η επιστημονική επικοινωνία (scholarly communication)- διαδικασία κατά την οποία οι επιστήμονες και οι ερευνητές μοιράζονται, επικοινωνούν και δημοσιεύουν τα αποτελέσματα της έρευνας, καθιστώντας τα διαθέσιμα στην ακαδημαϊκή κοινότητα και στο ευρύ κοινό- πρέπει να έχει πρόσβαση σε μια ποικιλία καναλιών επικοινωνίας. Ωστόσο, σε γενικές γραμμές, για την κυκλοφορία ενός μηνύματος το μόνο που χρειάζεται είναι ο πομπός, ένας δέκτης και ένα κανάλι επικοινωνίας. Βεβαίως, αυτό το κανάλι είναι το αντικείμενο με το οποίο ασχολείται ο τομέας των Πηγών Πληροφόρησης. Ο πομπός μπορεί να είναι ένας οργανισμός, ένας θεσμός ή ένα άτομο. Αυτοί δίνουν ώθηση στο σχήμα επικοινωνίας, δηλαδή το μήνυμα, το οποίο θα φτάσει στο δέκτη, που είναι ένας άλλος οργανισμός, θεσμός ή άτομο. Αυτό το μήνυμα με τη σειρά του, ίσως, μπορεί να μετατραπεί σε ένα ή περισσότερα άλλα μηνύματα, τα οποία θα κυκλοφορήσουν με τον ίδιο τρόπο και τα οποία θα δημιουργήσουν νέα μηνύματα που πυροδοτούν την επιστημονική ανάπτυξη. Το έγγραφο, το οποίο είναι καρπός της επιστημονικής δημιουργίας του συγγραφέα, περιέχει το μήνυμα, που έστειλε ο συγγραφέας. Οι πηγές πληροφόρησης αποτελούνται από έγγραφα ανεξάρτητα από το μέσο που διοχετεύτηκαν και υπό την προϋπόθεση ότι υπάρχει η πιθανότητα να ερμηνευθούν οι πληροφορίες που περιέχονται. Αυτός είναι ο τρόπος με τον οποίο επιτυγχάνεται η καθοριστική λειτουργία της πληροφόρησης και της επικοινωνίας.

Ένα εμπόδιο που πρέπει να αντιμετωπίσουν αυτοί που διεξάγουν βιβλιομετρικές αναλύσεις είναι ο μεγάλος αριθμός των εγγράφων που μεταφέρουν πληροφορίες, οι οποίες είναι ταυτόχρονα και εξειδικευμένες και απευθύνονται στο ευρύ κοινό. Προτού ανακαλύψουν την αδυναμία πρόσβασης σε όλα τα έργα, πρέπει να καταφύγουν στο έργο αυτών που ασχολούνται με τις Πηγές Πληροφόρησης. Στη συνέχεια, ομαδοποιούν τα έγγραφα, τα οποία, αφότου αναλυθούν, παρέχουν δεδομένα ή βιβλιογραφικά πεδία, που διευκολύνουν την επισκόπηση, την ανάλυση και τη στατιστική επεξεργασία που χρησιμοποιεί η Βιβλιομετρία.

Ο τομέας της Βιβλιομετρίας, αυτός καθαντός, από τη γένεση του, στράφηκε προς τη Βιβλιογραφία ως την κύρια βάση για την ανάπτυξή του. Πράγματι, οι ίδιοι οι ερευνητές δεν δίστασαν να αναφερθούν στον τομέα των Πηγών Πληροφόρησης ως ένα μέσο, στο οποίο εφαρμόζονται οι βιβλιομετρικές τεχνικές. Αυτές οι τεχνικές, με τη σειρά τους, είναι στενά συνδεδεμένες με τις στατιστικές μεθόδους.

Η Βιβλιομετρία, επιπλέον, είναι μία σημαντική περιοχή μελέτης των επιστημών της πληροφορίας, που αναπτύχθηκε τον 20ο αιώνα, από το ενδιαφέρον των ερευνητών για τις τάσεις και τις εξελίξεις των

επιστημών, όπως αυτές αντανακλώνται στη βιβλιογραφική παραγωγή τους. Το ενδιαφέρον αυτό έχει να κάνει με τη διερεύνηση των αλλαγών στα επιστημονικά πεδία μέσα στο χρόνο ή στις διάφορες χώρες, με την υψηλή παραγωγή δημοσιευμάτων πολλών επιστημόνων, μεγάλο μέρος των οποίων έχει ωστόσο μηδενική επιρροή στην ερευνητική διαδικασία. Η εξέλιξη της Βιβλιομετρίας χαρακτηρίστηκε από δύο βασικά στοιχεία: α) τις δευτερογενείς βιβλιογραφικές πηγές και πώς αυτές θα περιέχουν όλα τα στοιχεία που είναι απαραίτητα για την ποσοτική ανάλυση, ενώ παράλληλα θα εξασφαλίζουν οικονομία χρόνου και β) τη διευκόλυνση των επιστημόνων στην αναζήτηση αριθμητικών στοιχείων και στην ανάλυσή τους με στατιστικές ή άλλες μαθηματικές μεθόδους.

Οι πληροφορίες αυτές μπορούν να έχουν διαστάσεις με διάφορους τρόπους και για διάφορους σκοπούς. Από τη βιβλιογραφική περιγραφή των εγγράφων, η Βιβλιομετρία δανείζεται τις πληροφορίες που είναι αναγκαίες για να πραγματοποιήσει τις μελέτες της: συγγραφέας (ή συγγραφείς), γενικός τίτλος του περιοδικού ή της μονογραφίας, έτος έκδοσης, είδος εγγράφου (άρθρο, μονογραφία, βιογραφία κ.τ.λ.), γλώσσα, σύνοψη - εάν η βιβλιογραφία είναι αναλυτική - και λέξεις κλειδιά ή στοιχεία περιγραφής. Συνοπτικά, οι βιβλιογραφικές αναφορές παρέχουν στους ερευνητές μια τεράστια ποσότητα πληροφοριών, χωρίς τις οποίες δεν θα μπορούσαν να κάνουν τις βιβλιομετρικές τους αναλύσεις.

Η Βιβλιομετρία χρησιμοποιείται για την εκτίμηση του αντίκτυπου της επιστημονικής έρευνας μέσω της ποσοτικής ανάλυσης, ενώ η ποιοτική ανάλυση-αξιολόγηση επιτυγχάνεται με τη χρήση των διαφόρων μορφών αξιολόγησης από ομότιμους. Αυτό που αποκαλούμε σήμερα ως Βιβλιομετρία αρχικά ονομαζόταν από τον Hulme (1932), «Στατιστική Βιβλιογραφία». Αυτός ο όρος ανταποκρινόταν στην ανάγκη που δημιουργούσε ο αυξανόμενος όγκος των εκδόσεων, η οποία ήταν τόσο εκτεταμένη που άρχισε να δυσκολεύει τους ερευνητές. Με αυτόν τον τρόπο, η Βιβλιογραφία συμπληρώθηκε από τη χρήση μεθόδων που ανήκαν στη Στατιστική, έτσι ώστε να ανταποκριθεί στις ανάγκες πληροφόρησης των επιστημόνων. Το ίδιο το κριτήριο της Στατιστικής Βιβλιογραφίας διατηρείται από τον Raising έως το 1962.

Το 1969 προτάθηκε, ως πιο αντιπροσωπευτικός, ο όρος «Βιβλιομετρία» από τον Alan Pritchard (1969) και ορίστηκε ως «η εφαρμογή των μαθηματικών και στατιστικών μεθόδων στα βιβλία και τα άλλα μέσα δημοσίευσης» ή αναλυτικότερα, ως «η ποσοτική μελέτη των βιβλιογραφικών αναφορών όπως αυτές εμφανίζονται στις βιβλιογραφίες, με στόχο την παροχή εξελικτικών μοντέλων στις επιστήμες και την τεχνολογία». Εργαλεία για τη βιβλιομετρική ανάλυση του εντοπισμού και της μελέτης των αναφορών είναι η χρησιμοποίηση της πλατφόρμας του Web of Science του οργανισμού της Thomson, της Scopus, της Google Scholar και άλλων βάσεων δεδομένων, τις οποίες θα δούμε αναλυτικότερα παρακάτω. Οι δύο μεγάλες αλλαγές στο χώρο των εκδόσεων, η μηχανογράφηση της εκτυπωτικής διαδικασίας καθώς και η μετατροπή ολόκληρου του κύκλου των εκδόσεων που περιλαμβάνει την υποβολή ενός άρθρου, την αποδοχή και τη δημοσίευση του στο διαδίκτυο, επέτρεψαν την ταχύτερη και φθηνότερη ανταλλαγή γνώσης. Επιπροσθέτως βοήθησαν υιοθετώντας αλγορίθμους και εργαλεία μέτρησης της πληροφορίας από τεκμήρια/ή και δημιουργούς για τους σκοπούς της ποσοτικής ανάλυσης.

Η εφαρμογή, τώρα, της Βιβλιομετρίας απαιτεί τα εξής τέσσερα (4) βήματα:

1) Ποσοτικοποίηση

- i) μείωση δεδομένων (εξαγωγή βασικών πληροφοριών),
- ii) καθαρισμός δεδομένων και αποσαφήνιση δεδομένων,

2) Μέτρηση

- i) τι μπορούμε και τι πρέπει να μετρήσουμε,
- ii) ποιες είναι οι κατάλληλες μετρήσεις για τον σκοπό αυτό,

3) Συγκριτική αξιολόγηση

- i) να βάζουμε δεδομένα σε πλαίσια και να υπολογίζουμε τους περιορισμούς,
- ii) σύγκριση,

4) Παρουσίαση και ερμηνεία

- i) προετοιμασία διαγραμμάτων και χαρτών,
- ii) διάθεση ερμηνείας των παρατηρήσεων,
- iii) σύνδεση βιβλιομετρικών δεικτών έτσι ώστε να έρχονται ως αποτελέσματα άλλοι μέθοδοι.

Τέλος, τα οφέλη που η Βιβλιομετρία μπορεί να προβάλει ως χρήσιμες πληροφορίες σε τομείς εφαρμογής της είναι οι:

επιστημονικές πληροφορίες.

Ο τι είναι σημαντικό στο δικό μου κλάδο,

- πώς να αποκτήσουμε γρήγορη πρόσβαση σε αυτήν την πληροφορία,
 - τι είναι τα δίκτυα στο δικό μου πεδίο.
- μελέτες δομής και δυναμικές της επιστήμης,**
 - χαρτογράφηση και απεικόνιση της δομής της επιστήμης,
 - εμφάνιση νέων τάσεων, ερευνητικών πεδίων και επίκαιρων θεμάτων,
 - έρευνες αξιολόγησης και χρηματοδότησης.
 - ανάπτυξη μεθόδων για την μέτρηση ή την έρευνα,
 - βιβλιομετρικά συστατικά για τους τύπους χρηματοδότησης,
 - συγκριτικές μελέτες για την αξιολόγηση της έρευνας.

Η Βιβλιομετρία θεωρείται ως διεπιστημονικός και πολυεπιστημονικός τομέας της επιστήμης. Ένα αντικείμενο θεωρείται διεπιστημονικό, όταν δανείζεται στοιχεία από επιστημονικούς τομείς του ίδιου επιστημονικού κλάδου. Πρόκειται, λοιπόν, για ένα πολυεπιστημονικό αντικείμενο, που δανείζει ή δανείζεται γνώσεις ή τεχνικές πέραν του γενικού πεδίου. Η Βιβλιομετρία, ως ένας διεπιστημονικός τομέας, αναλαμβάνει το ρόλο του βοηθήματος ή του μέσου στη μέτρηση διαφόρων πεδίων που αποτελούν το γενικό επιστημονικό κλάδο του, δηλαδή τομείς που εντάσσονται στο πρόγραμμα σπουδών της Βιβλιοθηκονομίας και της Επιστήμης της Πληροφόρησης. Αυτά τα αντικείμενα λειτουργούν ως κρίκοι μιας αλυσίδας, που όταν ένας από αυτούς σπάσει, οι άλλοι δεν μπορούν να λειτουργήσουν ικανοποιητικά. Ας μας επιτραπεί να αποσαφηνίσουμε αυτό το σημείο με ένα παράδειγμα. Η Βιβλιομετρία συνεργάζεται τακτικά με τον τομέα των Πηγών Πληροφόρησης, ώστε να εντοπίσει τα κενά στις βιβλιογραφικές συλλογές, να τις συντηρήσει και να διευκολύνει τις διορθώσεις και την περιστασιακή αφαίρεση στοιχείων που δεν είναι αποδεκτά. Επιπλέον οι αναλύσεις τους βασίζονται στα περιγραφικά πεδία των πρωτογενών εγγράφων, όπως έχουν διατυπωθεί από τον τομέα της ανάλυσης εγγράφων ως: καταλογογράφηση, ταξινόμηση κ.λ.π. Όταν δεν μπορούμε να βρούμε ένα δεδομένο έγγραφο εξαιτίας της αυξανόμενης ποσότητας αυτών, διέξοδος είναι η βιβλιογραφία και οι πηγές πληροφόρησης, είτε σε έντυπη είτε σε μηχανογραφημένη μορφή, έτσι ώστε να αποκτήσουμε τα απαραίτητα έγγραφα που απαιτούνται, είτε για την ανάλυση ενός δεδομένου θέματος είτε για τον προσδιορισμό των ερευνητών του.

Πρόσφατες εξελίξεις στις μεθόδους της Βιβλιοθηκονομίας και της Επιστήμης της Πληροφόρησης έχουν συμβάλει σημαντικά στην εδραίωση της Βιβλιομετρίας. Με τη σειρά της, η Βιβλιομετρία έχει διαδραματίσει ένα σημαντικό ρόλο στην προβολή και την εφαρμογή της Επιστήμης της Πληροφόρησης και, ως εκ τούτου, στη συμβίωση μεταξύ των δύο επιστημών. Η Βιβλιομετρία είναι, επίσης, ένας πολυεπιστημονικός τομέας. Δανείζεται στατιστικές μεθόδους, προκειμένου να διεξάγει τις αναλύσεις της. Η ίδια εξυπηρετεί τις έρευνες και τις δοκιμές των κοινωνιολόγων και, εν τέλει, αξιοποιεί την επιστήμη της πληροφορικής για την επεξεργασία των δεδομένων με τη βοήθεια υπολογιστικών φύλλων, στατιστικών εφαρμογών και βάσεων δεδομένων. Όλα αυτά τα εργαλεία που προέρχονται από άλλους τομείς σπουδών εξυπηρετούν τη Βιβλιομετρία για να πετύχει το στόχο της, δηλαδή την ανάλυση του έργου των επιστημόνων και των ερευνητών στους διάφορους γνωστικούς κλάδους.

Όσον αφορά στο αντικείμενο της Βιβλιομετρίας, όπως υποδείξαμε και παραπάνω, βρίσκει την επιστημολογική βάση του στη Βιβλιογραφία. Λειτουργεί ως παράγοντας μέτρησης των πηγών πληροφόρησης και τελικά εμφανίζεται ως μέθοδος που χρησιμοποιείται από την επιστημονική έρευνα. Ως αποτέλεσμα, η Βιβλιομετρία μελετά την οργάνωση των τεχνολογικών τομέων από την οπτική των πηγών πληροφόρησης. Μετρά την επιστημονική ανάπτυξη με την εφαρμογή στατιστικών μεθόδων στη παραγωγή των επιστημόνων, καθιερώνοντας με αυτόν τον τρόπο τον βαθμό ανάπτυξης των διαφόρων επιστημονικών τομέων, πραγματοποιώντας μελέτες σχετικά με τη κατανάλωση πληροφοριών βάσει εγγράφων, που χρησιμοποιήθηκαν από τους επιστήμονες. Αυτές οι αναλύσεις διεξάγονται με τη χρήση βιβλιογραφικών αναφορών των εκδόσεων που εμπεριέχουν οι βιβλιογραφίες και οι πηγές πληροφόρησης, κατά τη διάρκεια ενός καθορισμένου χρονικού διαστήματος. Αυτά τα βιβλιογραφικά όργανα παρέχουν αρκετά δεδομένα για τα πρωτογενή έγγραφα, που επιτρέπουν τις πολύ σημαντικές βιβλιομετρικές έρευνες.

Ο επιστημονικός τομέας των Πηγών Πληροφόρησης, όπως έχουμε ισχυριστεί, αποτελεί την πρωταρχική βάση των Βιβλιομετρικών μελετών. Με αυτόν τον τρόπο, οι ερευνητές επιλέγουν τις πιο κατάλληλες πηγές για την ανάπτυξη του έργου τους.

Επιπλέον οι κατάλογοι περιοδικών και πηγών που περιέχουν περιλήψεις βιβλίων ή άρθρων που έχουν αναλυθεί, ίσως, χρησιμοποιηθούν για να εντοπίσουν αναφορές που έχουν δημιουργηθεί από το έργο του συγγραφέως. Αυτή είναι η μέθοδος της μέτρησης που προβλέπει η Βιβλιομετρία, δηλαδή η ανάλυση των

αναφορών. Η δυνατότητα ερμηνείας αυτών των ποσοτικών μέτρων ανοίγει νέους ορίζοντες για τη μελέτη των διαφόρων επιστημών. Έχοντας κατά νου αυτές τις θεωρητικές παραδοχές, προτείνουμε τον εξής ορισμό του επιστημονικού τομέα μας: Η Βιβλιομετρία είναι το σύνολο της μεθοδολογικής γνώσης που εξυπηρετεί την εφαρμογή των τεχνικών ποσοτικού προσδιορισμού, προκειμένου να αξιολογήσει τις διαδικασίες παραγωγής, επικοινωνίας και χρήσης των επιστημονικών πληροφοριών. Στόχος της είναι η συμβολή στην ανάλυση και την αξιολόγηση της επιστήμης και της έρευνας.

Εφαρμογές της Βιβλιομετρίας

Σήμερα, η βιβλιομετρική ανάλυση (Βιβλιομετρία) εφαρμόζεται σε ένα ευρύ φάσμα πεδίων:

- Στην ιστορία της επιστήμης, όπου συμβάλλει στη διασαφήνιση της ανάπτυξης και της εξέλιξης των επιστημονικών κλάδων, μέσα από τον εντοπισμό των ιστορικών κινημάτων που αποτυπώνονται στα αποτελέσματα που εξάγουν οι ερευνητές.

- Στις κοινωνικές επιστήμες, όπου, εξετάζοντας την επιστημονική βιβλιογραφία, υποστηρίζει και ενισχύει την ανάλυση της επιστημονικής κοινότητας και της δομής της σε μια δεδομένη κοινωνία, καθώς και τα κίνητρα και τα δίκτυα των ερευνητών.

- Στην τεκμηρίωση, όπου μπορεί να υπολογίζει τον αριθμό των περιοδικών που υπάρχουν ανά βιβλιοθήκη, αλλά και να προσδιορίζει τα περιοδικά που αποτελούν τον πυρήνα, τις δευτερεύουσες πηγές και την περιφέρεια ενός επιστημονικού κλάδου (αναλύοντας την ποσότητα των περιοδικών που απαιτούνται για την κάλυψη του 50 %, 80 % ή 90 % των συνολικών πληροφοριών σε ένα συγκεκριμένο επιστημονικό πεδίο).

- Στην πολιτική της επιστήμης, όπου παρέχει δείκτες μέτρησης της παραγωγικότητας και της επιστημονικής ποιότητας. Ουσιαστικά, δηλαδή, παρέχει μια βάση πάνω στην οποία δομείται η αξιολόγηση και ο προσανατολισμός της έρευνας και ανάπτυξης (εφεξής θα αναφέρεται ως E & A).

Οι βιβλιομετρικές τεχνικές έχουν εξελιχθεί με την πάροδο του χρόνου και συνεχίζουν προς την κατεύθυνση αυτή και είναι οι εξής:

- Η καταμέτρηση των εγγράφων με απόδοση ανά χώρα, ανά ίδρυμα/θεσμό και συγγραφέα.
- Η καταμέτρηση των αναφορών, ώστε να υπολογισθεί ο αντίκτυπος των δημοσιευμένων εργασιών στην επιστημονική κοινότητα.
- Η καταμέτρηση των συν-αναφορών (δηλαδή να υπολογισθεί το πόσες φορές δύο ερευνητικά έγγραφα αναφέρονται και παρατίθενται μαζί σε ένα ενιαίο έγγραφο) κ.τ.λ.

Όλες αυτές οι τεχνικές συνδυάζονται για να δώσουν πιο λεπτομερείς και πιο αποτελεσματικές μετρήσεις. Τα αποτελέσματα παρουσιάζονται με διάφορες μορφές, όπως είναι η χαρτογράφηση, προκειμένου να απεικονίζουν τις σχέσεις που αναπτύσσονται μεταξύ των συμμετεχόντων και να επεκτείνουν τα μέσα που απαιτούνται για την ανάλυση.

1.2 Οι πρωτοπόροι της Βιβλιομετρίας

Ο Shapiro (1992) μας παρουσιάζει τα της Βιβλιομετρίας, ως ένα θέμα που έχει παραμεληθεί από τους ιστορικούς επιστήμονες της επιστήμης της πληροφορίας. Η χρήση των δεικτών αναφοράς έχει παρουσιαστεί εδώ και πολλά χρόνια, ήδη από το 1743, ενώ έχουν δημοσιευτεί πολλά άρθρα, τα οποία βρίσκονται σε νομικά κείμενα, τουλάχιστον από το 1817. Ο Weinberg (1997) σημειώνει πως οι δείκτες αναφοράς Hebrew είναι ακόμη παλαιότεροι και χρονολογούνται περίπου στο 12ο αιώνα. Η ιδέα της διεξαγωγής έρευνας και εξέτασης της βιβλιογραφίας έχει τις ρίζες της στην αρχή του αιώνα. Το 1917, οι επιστήμονες FJ. Cole και Nellie Eales δημοσίευσαν μια στατιστική ανάλυση της ιστορίας της συγκριτικής ανατομίας. Η ημερομηνία αυτή αποτέλεσε ορόσημο στην ιστορία της βιβλιομετρικής ανάλυσης, καθώς ο Cole και η Eales ήταν από τους πρώτους που χρησιμοποίησαν το δημοσιευμένο ερευνητικό έργο για να δημιουργήσουν μια ποσοτική εικόνα της προόδου που συντελείται σε έναν ερευνητικό τομέα. Η εργασία τους περιγράφει τη συνεισφορά της Βιβλιομετρίας καθώς επίσης και τα προβλήματα που αυτή θέτει - μερικά από τα οποία δεν έχουν ακόμη επιλυθεί-.

Περαιτέρω έργο διεξήχθη από τον Hulme (1923), αξιοποιώντας αυτή τη φορά τα διπλώματα ευρεσιτεχνίας. Συσχετίζοντας τα διπλώματα ευρεσιτεχνίας και την επιστημονική βιβλιογραφία, προκειμένου να μετρηθεί η κοινωνική πρόοδος στη Βρετανία, ο Hulme πρωτοστάτησε σε μια σύγχρονη μεθοδολογία για την ιστορία της επιστήμης.

Ακολούθως, ο Cole (1926) έδειξε τις συχνότητες κατανομής της επιστημονικής παραγωγής. Ήταν αναμφισβήτητα ένας από τους πρώτους που συνδέσαν την έννοια της παραγωγικότητας με την καταμέτρηση, χρησιμοποιώντας ανά δεκαετία τους δείκτες που προέρχονταν από την Υπηρεσία Χημικής Ταυτοποίησης (Chemical Abstracts Service - (CAS): πρόκειται για τμήμα της Αμερικάνικης Χημικής Εταιρείας), καθώς και από το εγχειρίδιο του Auerbach, *Geschichts Tafelnder Physik*. Εισήγαγε, επίσης, ένα ποιοτικό μέτρο του επιστημονικού έργου, βάσει στοιχείων και δεδομένων που κατέστησαν δυνατή την επιλογή των συνεισφορών, οι οποίες κρίθηκαν εξέχουσες και διακρίθηκαν.

Εκείνος, όμως, που ανέπτυξε τις στατιστικές σχέσεις με τη δημοσίευση ήταν ο S.W. Fernberger του Πανεπιστημίου της Πενσυλβάνια. Ο Fernberger (1936) μελέτησε την εξέλιξη των ερευνητών και έδωσε αυξανόμενη έμφαση στη δημοσίευση ως κριτήριο για την επιλεξιμότητα. Εξέτασε, επίσης, τα οικονομικά και οργανωτικά θέματα των περιοδικών του Συνδέσμου Ψυχολόγων και των συνεδρίων τους. Επεξεργάστηκε τον αριθμό των εργασιών που παρουσιαζόταν σε κάθε συνέδριο Ψυχολογίας από το 1892. Κατέγραψε την παραγωγικότητα των πανεπιστημίων σε αυτές τις συναντήσεις και περιέγραψε αυτό που αποκάλεσε ως «τη συνοχή της δημοσίευσης και τους τομείς ενδιαφέροντος». Επίσης διαπίστωσε ότι σε 19 πανεπιστήμια παραγόταν το 53% του συνόλου των δημοσιεύσεων. Ο Fernberger ήταν αυτός που επέβαλε τις έννοιες της παραγωγικότητας και του δείκτη για τη μέτρηση της παραγωγικότητας της επιστήμης.

Το 1906, ο Cattell (1906) δρομολόγησε τον βιογραφικό κατάλογο των Αμερικανών επιστημόνων, ο οποίος δημοσιευόταν κάθε πέντε χρόνια και συνέλλεγε πληροφορίες σχετικές με χιλιάδες επιστήμονες που δραστηριοποιούνταν στον τομέα της έρευνας. Επεξεργάστηκε στατιστικές σχέσεις με τον αριθμό των επιστημόνων και τη γεωγραφική κατανομή τους, καθώς επίσης και την κατάταξη των επιστημόνων ανάλογα με την επίδοσή τους. Κατά συνέπεια μπορεί να πιστωθεί στον Cattell ότι ξεκίνησε τη συστηματική μέτρηση της επιστήμης. Στη συνέχεια εισήγαγε δύο διαστάσεις για τη μέτρηση της επιστήμης, την ποιότητα και την ποσότητα. Η ποσότητα ή παραγωγικότητα, όπως την αποκαλούσε, ήταν απλώς η καταμέτρηση του αριθμού των επιστημόνων που ανήκαν σε ένα έθνος, ενώ η ποιότητα ή η επίδοση ορίστηκε ως η συνεισφορά ενός εκάστου στην πρόοδο της επιστήμης και μετρήθηκε από την κατάταξη των συναδέλφων του, σύμφωνα με τον αριθμό των εργασιών τους. Τον Cattell ακολούθησαν και άλλοι ερευνητές όπως ο ψυχολόγος Buchner ο οποίος, στη σειρά των κριτικών σχολίων του για την ψυχολογία, περιλάμβανε τη συζήτηση των πρόσφατων άρθρων, τον αριθμό των ψυχολόγων, τον κατάλογο των νέων περιοδικών, καθώς και στατιστικά στοιχεία σχετικά με τις δημοσιεύσεις. Ο Buchner αναφερόταν επίσης, στην ποσοστιαία κατανομή των άρθρων που περιλαμβάνονταν στο σχετικό δείκτη, καθώς επίσης και στα ενδιαφέροντα των ψυχολόγων.

Το 1971, ένας Ιταλός οικονομολόγος, ο Vilfredo Pareto (1971) παρατήρησε αυτό που στη συνέχεια ονομάστηκε αρχή του Pareto ή κανόνας του 80-20, αφού σύμφωνα μ' αυτήν, στα περισσότερα γεγονότα, περίπου το 80% των αποτελεσμάτων προέρχεται από το 20% των αιτιών. Έτσι, μπορεί να αναμένεται ότι το 80% των αναφορών αφορούν έναν πυρήνα 20% τίτλων σε περιοδικά. Ομοίως, το 80% των άρθρων σε περιοδικά ανήκουν περίπου στο 20% των συγγραφέων. Δηλαδή τα βιβλιομετρικά φαινόμενα έχουν ένα καταμερισμό βαθειά ασύμμετρο, αφού οι δημοσιεύσεις, οι αναφορές κ.λπ. αναπαράγονται από μία μικρή αναλογία πηγών, συγγραφέων, περιοδικών, ιδρυμάτων, σε αντίθεση με τα περισσότερα φαινόμενα που παρατηρούνται στη φύση, τα οποία ακολουθούν την κανονική κατανομή. Το βασικό γνώρισμα της κανονικής κατανομής είναι ο σχηματισμός της "καμπάνας" που την καθιστά, άκρως συμμετρική. Αποτέλεσμα αυτού είναι ότι ο μέσος όρος είναι ίσος με την διάμεσο και την επικρατούσα τιμή (σημείο μέγιστης συχνότητας), δηλαδή όλες οι τιμές των παραμέτρων θέσης συμπίπτουν. Αυτή η κανονικότητα επιτρέπει τη χρήση βασικών στατιστικών τεχνικών όπως συσχέτιση, παλινδρόμηση, καθώς επίσης και δοκιμασμένων στατιστικών τεστ.

Το 1926, ο Alfred.J. Lotka (1926) δημοσίευσε την πρωτοποριακή μελέτη του σχετικά με την κατανομή της συχνότητας της επιστημονικής παραγωγικότητας που καθορίστηκε από τον δεκαετή δείκτη (1907-1916) της Chemical Abstracts. Ο Lotka κατέληξε στο συμπέρασμα ότι «Ο αριθμός των συγγραφέων που παράγουν n δημοσιεύσεις είναι αντιστρόφως ανάλογος με το n^2 - πολλαπλασιασμένος με μια σταθερά που υπολόγισε ο ίδιος ο Lotka- ενώ το ποσοστό εκείνων που παράγουν μία μόνο δημοσίευση είναι περίπου 60 τοις εκατό». Αυτό το αποτέλεσμα μπορεί να θεωρηθεί ως γενικός κανόνας, ακόμη και σήμερα, 90 χρόνια μετά τη δημοσίευσή του. Κατά την ίδια σχεδόν χρονική περίοδο, το 1927, οι Gross και Gross (1927) δημοσίευσαν μία μελέτη εστιασμένη στις αναφορές, προκειμένου να βοηθήσουν την απόφαση, για το ποιά περιοδικά Χημείας θα ήταν καλύτερο να αγοράζονται από μικρές βιβλιοθήκες κολλεγίων. Ειδικότερα, εξετάστηκαν 3633 αναφορές από τον τόμο του περιοδικού της American Chemical Society του 1926. Η μελέτη αυτή θεωρείται ως η πρώτη ανάλυση αναφορών, αν και δεν αποτελούσε ανάλυση αναφοράς με τη σημερινή έννοια.

Οκτώ χρόνια μετά την εμφάνιση του άρθρου του Lotka, ο Bradford, το 1934, δημοσίευσε τη μελέτη του για την κατανομή συχνότητας των δημοσιεύσεων σε περιοδικά. Ο Bradford (1934) ανακάλυψε ότι «Εάν τα επιστημονικά περιοδικά μιας θεματικής ενότητας διαταχθούν κατά φθίνουσα σειρά παραγωγικότητας, μπορούν να χωριστούν σε έναν πυρήνα περιοδικών με αρκετές ομάδες ή ζώνες που περιέχουν τον ίδιο αριθμό άρθρων, όπως αυτόν του πυρήνα, τότε οι αριθμοί των περιοδικών στον πυρήνα και στις διαδοχικές ζώνες αυξάνει με την αναλογία $k^0 : k^1 : k^2$ ». Μια σημαντική συνέπεια του νόμου είναι το γεγονός ότι σε μια αναζήτηση για ένα συγκεκριμένο θέμα, ένας μεγάλος αριθμός των σχετικών άρθρων θα είναι συγκεντρωμένος σε ένα μικρό αριθμό τίτλων περιοδικών (Nordstrom, 2005).

Με τους νόμους αυτούς συνήθως γίνονται οι εκτιμήσεις των δεικτών αναφορών, καθώς και των διαφορών υπηρεσιών των βιβλιοθηκών. Για παράδειγμα, στον Δείκτη Επιστημονικών Παραπομπών (Science Citation Index) παρακολουθείται μόνο το ένα δέκατο πέμπτο (1/15) περίπου των συνολικών περιοδικών, αλλά καταγράφονται περισσότερες από τα τρία τέταρτα (3/4) των συνολικών αναφορών (Price, 1976). Επιπροσθέτως οι βιβλιοθήκες όταν χρειάζεται να περιορίσουν τις δαπάνες τους, χρησιμοποιούν το νόμο του Bradford, για να προσδιορίσουν τις λιγότερο επώδυνες περικοπές περιοδικών.

Ο Otlet στη συνέχεια ήταν αυτός που χρησιμοποίησε τον όρο Βιβλιομετρία (Bibliometrie) για να περιγράψει τη τεχνική, με την οποία επιδίωκε τον ποσοτικό προσδιορισμό της επιστήμης και των επιστημών. Ο Otlet (1920), που ήταν πρωτοπόρος στην επιστήμη της πληροφόρησης και της θεωρίας της, εμμένει στη διαφορά ανάμεσα στη Βιβλιομετρία και τη Στατιστική Βιβλιογραφία υποστηρίζοντας πως η επιστήμη από τη γέννηση της μετρίεται ή προσδιορίζεται ποσοτικά με την εφαρμογή στατιστικών μεθόδων σε πηγές πληροφόρησης.

Η θεώρηση του Otlet είναι ότι η Βιβλιογραφία καθιερώνεται ως μια γενική επιστήμη που συλλέγει και ταξινομεί συστηματικά την ολότητα των δεδομένων, τα οποία αφορούν στην παραγωγή, τη συντήρηση, την κυκλοφορία και τη χρήση των γραπτών και των εγγράφων κάθε είδους. Στην πραγματεία του για την επιστήμη της πληροφόρησης, ο Otlet προβάλλει μια σειρά ιδεών για τη Βιβλιομετρία, μεταξύ των οποίων οι ακόλουθες είναι οι πιο σημαντικές:

α) Σε κάθε μορφή γνώσης, το μέτρο είναι η ανώτερη μορφή που αυτή η γνώση σχηματίζει. Τα μέτρα σχετικά με τα βιβλία και τα έγγραφα μπορούν να συγκροτήσουν ένα σύνολο συντεταγμένων, τη Βιβλιομετρία. Παρόλο που ο Otlet είχε χρησιμοποιήσει τον όρο «bibliometrie» αργότερα, ο Pritchard (1981) εισήγαγε τον όρο Βιβλιομετρία και όρισε ευρέως πως η Βιβλιομετρία είναι «η εφαρμογή μαθηματικών και στατιστικών μεθόδων σε βιβλία και άλλα μέσα επικοινωνίας».

β) Τα μέτρα σχετίζονται με τα αντικείμενα, τα φαινόμενα ή τα γεγονότα, τις σχέσεις ή τους νόμους. Τα μέτρα των κύριων σχέσεων μιας επιστήμης γίνονται δείκτες (για παράδειγμα, όταν οι γεωγράφοι μελετούν τη σχέση του νερού και της βροχής με τη γη, δημιουργούν ένα δείκτη ξηρασίας).

γ) Όταν ασχολούμαστε με τη Βιβλιομετρία, πρέπει να λάβουμε υπόψη μας τα ευρήματα της μετρικής (σε γενικό πλαίσιο) και της μετρικής στις κοινωνικές επιστήμες (σε ειδικό πλαίσιο). Το απόφθεγμα «τα πάντα με μέτρο» έχει γίνει η κατευθυντήρια γραμμή κάθε επιστήμης, η οποία τείνει να προχωρήσει από το ποσοτικό στο ποιοτικό στάδιο. Ο López Yepes (1995) τονίζει την ικανότητα του Otlet να οργανώνει τη γνώση και τη συνεχιζόμενη αναζήτηση του για μια σύνθετη εξήγηση σχετικά με το πώς δημιουργούνται και αναπτύσσονται οι έννοιες. Αναγνωρίζει στον Otlet την ικανότητά του για ορθολογική οργάνωση.

Ο Otlet ανέλυσε τους λόγους για τους οποίους επιστήμες όπως η Αστρονομία, η Βιολογία, η Κοινωνιολογία και άλλες τείνουν να έχουν έναν ποσοτικό χαρακτήρα. Έχουν θεσπίσει μεθόδους μέτρησης, οι οποίες αποδίδουν αποτελέσματα. Όσον αφορά τα βιβλία, τόνισε ότι:

1) Τα αντικείμενα που σχετίζονται με τα βιβλία δεν μπορούν να μετρηθούν εύκολα, είτε υπό την έννοια του υλικού και της λειτουργικότητας τους, είτε υπό την έννοια της υποκειμενικής πραγματικότητας. Οι προσπάθειες, ως εκ τούτου, είναι επιθυμητές προς αυτήν την κατεύθυνση.

2) Οι επιστήμες που σχετίζονται με τα βιβλία πρέπει να εισάγουν την ιδέα της μέτρησης στην έρευνα που προωθούν. Στο μέτρο που τα βιβλία είναι αντικείμενα μελέτης Ψυχολογίας, Κοινωνιολογίας ή Τεχνολογίας τα δεδομένα τους επιδέχονται μέτρησης.

3) Η Βιβλιομετρία θα αποτελεί μέρος της Βιβλιογραφίας, η οποία θα ασχολείται με την εφαρμογή της ποσότητας ή του μέτρου στα βιβλία (αριθμητική ή μαθηματική βιβλιογραφία).

4) Το κάθε στοιχείο με το οποίο ασχολείται η Βιβλιογραφία πρέπει να είναι κατ' αρχήν μετρήσιμο. Είναι σκόπιμο η έρευνα να αντιμετωπίζει τα στοιχεία με ακρίβεια, δηλαδή με τη μορφή αριθμών, έτσι ώστε να περνά από ένα ποιοτικό ή περιγραφικό στάδιο σε ένα ποσοτικό.

Ο Otlet προτείνει μια σειρά βασικών αρχών για τον τομέα της Βιβλιομετρίας, λαμβάνοντας υπόψη έναν αριθμό παραγόντων που επηρεάζουν ή περιβάλλουν το κείμενο. Σε αυτά συγκαταλέγονται η γλώσσα, τα

διαστήματα που περιέχονται και οι συντελεστές που αναφέρονται ανάμεσα σε άλλα, στη μορφή, στη διάταξη και την τιμή της μονάδας καθώς και σε συντελεστές που ανήκουν στη στατιστική, όπως δείκτες σύγκρισης. Επίσης δίνει προσοχή στη συχνότητα με την οποία διαβάζεται ένας δεδομένος συγγραφέας ή το έργο του. Από αυτά τα δεδομένα συνεπάγεται ότι μπορεί να σχεδιαστεί μια καμπύλη «συχνότητας χρήσης», λαμβάνοντας υπόψη τον αριθμό των εκδόσεων ενός κειμένου σε συνάρτηση με τον συγγραφέα και το περιεχόμενό του ή το συγκεκριμένο κοινωνικών προεκτάσεων στο οποίο εμφανίζεται.

Παρόλη τη σημασία του σε σχέση με τα βιβλία, ο Otlet πιστεύει ότι ο τομέας της Στατιστικής εκπληρώνει έναν πολύ διαφορετικό στόχο από αυτόν της Βιβλιομετρίας.

Οι στατιστικές για τα βιβλία, συχνά, συγχέονται με τη Βιβλιομετρία επειδή ίσως μέχρι σήμερα εφαρμόζονταν κυρίως για την απαρίθμηση της ποσότητας των βιβλίων που εκδιδόταν. Ωστόσο η χρήση της Στατιστικής αρχίζει τώρα να εξαπλώνεται στα νούμερα των αντιτύπων που εκτυπώνονται, στην κυκλοφορία βιβλίων, στις βιβλιοθήκες, στα βιβλιοπωλεία, στις τιμές και άλλα. Επιπροσθέτως πολλά έργα έχουν γραφτεί για τη στατιστική για τα βιβλία, τα οποία είχαν να κάνουν με απόλυτες τιμές καθώς και συντελεστές. Βεβαίως, δεν πρέπει να υπερτιμούμε τη σημασία αυτών των αριθμών, καθώς οι λίστες που παρουσιάζονται απέχουν πολύ από το να θεωρηθούν πλήρεις, ακριβείς και συγκρίσιμες. Από την άλλη πλευρά, οι συντελεστές που λαμβάνουμε είναι μόνο μέτρα τα οποία συγκρίνουν κάθε είδους μεταβολές σε συνάρτηση με μια ευρεία ποικιλία μεταβλητών. Ωστόσο, τα υπάρχοντα στοιχεία όπως ισχύουν προσωρινά, πρέπει να μας δείξουν έναν τρόπο για πιο ακριβείς και πλήρεις αριθμούς. Είναι σαφές ότι, σύμφωνα με τον Otlet, ενώ η Βιβλιομετρία μετρά το περιεχόμενο του βιβλίου, η στατιστική ασχολείται με τον κύριο κορμό και τις συνθήκες του.

Στη συνέχεια, το 1949, ο Zipf (1949) διατύπωσε έναν ενδιαφέροντα νόμο της ποσοτικής γλωσσολογίας, που ανακαλύφθηκε από τη μελέτη της συχνότητας των λέξεων σε ένα κείμενο. Σύμφωνα με τον Zipf, ισχύει η σχέση $r \cdot f = C$, όπου r είναι η κατάταξη μιας λέξης σε σχέση με τη συχνότητα εμφάνισης της στο κείμενο, f είναι η συχνότητα εμφάνισης της λέξης στο κείμενο και C είναι μια σταθερά που εξαρτάται από το κείμενο που αναλύεται. Η φιλοσοφία του νόμου εστιάζεται στην αρχή της ελάχιστης προσπάθειας, που σημαίνει ότι ένα άτομο θα προσπαθήσει να λύσει τα προβλήματά του με τέτοιο τρόπο, ώστε να ελαχιστοποιείται το συνολικό έργο που πρέπει να δαπανήσει για την επίλυση τόσο των άμεσων, όσο και των πιθανών μελλοντικών προβλημάτων του. Άρα θα κάνει χρήση της ίδιας λέξης, αντί μιας συνώνυμης, όταν αυτό είναι εφικτό.

Η μέχρι τότε κατάσταση άλλαξε δραματικά στις αρχές της δεκαετίας του 1960, όταν ο ιστορικός της επιστήμης Derek de Solla Price δημοσίευσε τη θεμελιώδη εργασία του στη Βιβλιομετρία που αναλύεται με τα δυο του βιβλία, το πρώτο για την «Επιστήμη από την Εποχή της Βαβυλώνας» (1961) και το δεύτερο για τη «Μικρή Επιστήμη, Μεγάλη Επιστήμη» (1963). Ο Price (1976) εγκαινίασε το ενδιαφέρον για την επιστήμη της επιστήμης, που στηρίζεται σε ακριβείς ποσοτικές αναλύσεις, αφενός των ρυθμών της επιστημονικής παραγωγής, δηλαδή του αριθμού των επιστημονικών βιβλίων και περιοδικών ανά μονάδα χρόνου, κι αφετέρου του πλήθους των ανθρώπων που απασχολούνται με την επιστήμη. Στον Price (Μπουντουρίδης Μ. Α., 1999) οφείλεται η διαπίστωση ότι το μεγαλύτερο μέρος του συνολικού επιστημονικού έργου παράγεται από ένα σχετικά μικρό αριθμό επιστημόνων.

Μία άλλη πεποίθηση του ήταν ότι, ενώ μπορεί να μην είμαστε σε θέση να διαβάσουμε όλα τα επιστημονικά βιβλία, εργασίες κ.λ.π., μπορούμε μόνο μετρώνοντας, να εκμαιεύσουμε πολλά συμπεράσματα σχετικά με την επιστήμη. Έτσι ο Price οδηγήθηκε στο μοντέλο του για τον εκθετικό ρυθμό της χρονικής ανάπτυξης της επιστήμης, σύμφωνα με το οποίο, όπως ισχυριζόταν, η παγκόσμια επιστήμη επεκτεινόταν από τον δέκατο έβδομο αιώνα κι ύστερα με εκθετικό ρυθμό (συν 5% ως 7% κάθε χρόνο), ούτως ώστε η επιστημονική παραγωγή να διπλασιάζεται κάθε δέκα ως δεκαπέντε χρόνια. Επιπροσθέτως ισχυριζόταν ότι ο εκθετικός αυτός ρυθμός που αναπτύσσεται η επιστήμη πρέπει να έχει ένα τέλος. Το 1963 ανέφερε ότι: Διαφαίνεται ότι δεν είναι σωστό να αναρριχηθούμε άλλα δυο επίπεδα ακριβώς όπως αναρριχηθήκαμε τα τελευταία χρόνια. Αν δεν συμμορφωθούμε, θα πρέπει να έχουμε δυο επιστήμονες για κάθε έναν άνδρα, γυναίκα, παιδί και σκύλο του πληθυσμού και θα πρέπει να ξοδεύουμε δυο φορές πιο πολλά χρήματα από αυτά που διαθέτουμε. Επομένως, η ημέρα της κρίσεως για την επιστήμη θα απείχε λιγότερο από έναν αιώνα » (Price, 1963, σ. 19).

Στη συνέχεια της αναλυτικής του έρευνας, παρουσίασε μια σειρά από ποσοτικές τεχνικές αξιολόγησης. Ήταν ο πρώτος που εξέτασε την αυξανόμενη τάση της συνεργασίας μεταξύ των ερευνητών της Χημείας, χρησιμοποιώντας τη Βιβλιομετρία. Από τότε η Βιβλιομετρία αναπτύχθηκε ως ερευνητικό πεδίο από μόνη της, με αποτέλεσμα την εμφάνιση στην επιστημονική κοινότητα των ειδικών που ονομάστηκαν βιβλιομέτρες. Στην επιστημονική ανάπτυξη, ο Price βλέπει από τη ένα μέρος, αυτό που αποκαλεί αρχειακό σώμα της ερευνητικής βιβλιογραφίας κι από το άλλο μέρος, αυτό που λέει ερευνητικό μέτωπο της

βιβλιογραφίας. Το αρχειακό υλικό είναι το τμήμα της βιβλιογραφίας μιας επιστημονικής περιοχής που έχει γραφεί σχετικά παλαιότερα, π.χ. πριν από τα τελευταία έξι επτά χρόνια. Στο μέτωπο της έρευνας υπάρχει ένα μέρος του συνόλου της βιβλιογραφίας, που έχει δημιουργηθεί πρόσφατα και που κατοπτρίζονται σε αυτό οι τελευταίες ερευνητικές εργασίες. Με άλλα λόγια, η επιστήμη φαίνεται να αναπτύσσεται σαν δένδρο που βγάζει συνέχεια νέα κλαδιά προς τα έξω, ενώ μέσα από την αναπτυσσόμενη βλάστησή του κρύβεται μια σταθερή, αλλά λιγότερο ενεργή δομή. Για τον λόγο αυτό χρησιμοποιούσε τον όρο αναπτυσσόμενο δέρμα για το ερευνητικό βιβλιογραφικό μέτωπο και περιέγραφε την ανάπτυξη της επιστήμης ισχυριζόμενος ότι, όσο περισσότερο λεπτό είναι το δέρμα της, τόσο πιο δομημένη και διαυγής είναι η ανάπτυξη και τόσο πιο γρήγορη η όλη διαδικασία. Ο Price ειδικότερα πιστεύει ότι το δέρμα του μετώπου της έρευνας δημιουργείται, κατά μέσο όρο, από πενήντα (50) ερευνητικές εργασίες και στην συνέχεια είναι επιβεβλημένο να ενεργοποιηθεί ένα feedback με την δημιουργία μιας ερευνητικής εργασίας ανασκόπησης που προφανώς θα μεταβάλλεται ανάλογα με το ερευνητικό πεδίο. Πάντως, για τις δυναμικά αναπτυσσόμενες περιοχές, το ότι το ερευνητικό δέρμα είναι λεπτό σημαίνει πως υπάρχει ένας γρήγορος ρυθμός απαρχαίωσης των εργασιών κι ενσωμάτωσης τους στο προηγούμενο υλικό, έτσι ώστε οι αναφορές σε αυτές να θεωρούνται αυτονόητες και άρα, ούτε καν να γίνονται. Αντίθετα, σε αργά αναπτυσσόμενες περιοχές, το ερευνητικό μέτωπο αποτελεί ένα πολύ μεγαλύτερο τμήμα της συνολικής βιβλιογραφίας, με αποτέλεσμα να είναι δύσκολος ο διαχωρισμός του από το παλαιότερο αρχειακό υλικό. Σε αυτές τις περιπτώσεις παρατηρείται μια περιοδική επαναφορά στα άλυτα προβλήματα ή τις ημι-αποδεκτές λύσεις τους, έτσι ώστε η επικαιρότητα της παλαιότερης βιβλιογραφίας να διατηρείται συνεχώς σταθερή.

Ο Price διαμόρφωσε, επίσης, το 1976, τη γενική θεωρία χαρακτηρισμού των μεθόδων της επιστημονικής επικοινωνίας ως την αρχή του σωρευτικού πλεονεκτήματος ως ακολούθως: «Η επιτυχία φαίνεται να αναπαραγάγει την επιτυχία. Ένα έγγραφο που έχει αναφερθεί πολλές φορές είναι πιο πιθανό να αναφέρεται και πάλι περισσότερο από ένα άλλο, το οποίο έχει ελάχιστα αναφερθεί. Ένας συγγραφέας πολλών εργασιών είναι πιο πιθανό να δημοσιεύσει ξανά, από κάποιον που είναι λιγότερο παραγωγικός. Ένα περιοδικό που συχνά αναφέρεται-γνωμοδοτεί για κάποιο θέμα είναι πιο πιθανό να ενεργοποιηθεί και πάλι, από ένα που προηγουμένως είχε σπάνια χρησιμοποιηθεί.» Η παραπάνω αρχή είναι γνωστή και ως «κατά Ματθαίον επίδρασης» -γιατί όποιος έχει, θα του δοθεί και θα εξακολουθεί να έχει σε αφθονία, ενώ όποιος δεν έχει, θα του αφαιρεθεί και αυτό που έχει- (Ευαγγελιστής Ματθαίος 25:29). Η επίδραση κατά Ματθαίον εξηγεί την αύξηση της αναγνώρισης για συγκεκριμένες επιστημονικές προσφορές σε επιστήμονες με αναγνωρισμένη φήμη και τη μη διατήρηση της αναγνώρισης από τους επιστήμονες που δεν έχουν δώσει ακόμα το στίγμα τους.

Ενώ αυτός ο ορισμός επικεντρώνεται στην αναγνώριση, ο κοινωνιολόγος της επιστήμης Merton αναγνώρισε ότι και άλλοι παράγοντες, επίσης, τείνουν να διαφοροποιήσουν τους επιστήμονες (Cole, 2004). Ο Merton κι η σχολή του είχαν έντονα προβληματισθεί στις δεκαετίες του 1960 και 1970 για το αν, όντως, η επιστήμη ζούσε μόνο κάτω από τον ιδανικό κανόνα της καθολικότητας, δηλαδή της ισονομίας και ισοπολιτείας στην επιστημονική κοινότητα ή αν οι ιδιαίτεροι παράγοντες, που διαφοροποιούν την επιστημονική κοινότητα παίζουν κάποιο ρόλο, όπως η ηλικία, το πανεπιστήμιο ή το ίδρυμα που υπηρετεί ένας επιστήμονας. Ο Merton, στη συνέχεια, είχε επιχειρηματολογήσει ότι το φαινόμενο κατά Ματθαίον εμφανίζεται κυρίως σε περιπτώσεις συνεργασίας και ανεξαρτήτων πολλαπλών ανακαλύψεων (Merton, 1973). Επεσήμανε ότι ένα αποτέλεσμα της επίδρασης κατά Ματθαίον είναι ότι, όταν ένας υψηλής στάθμης επιστήμονας κάνει μία επιστημονική προσφορά, είναι πιο πιθανό να εντοπιστεί, από έναν χαμηλότερης στάθμης επιστήμονα. Σχετικά με το Φαινόμενο κατά Ματθαίον υπάρχει άλλος ένας όρος που μερικές φορές εμφανίζεται στην επιστήμη, «το φαινόμενο του φωτοστέφανου» (Crane, 1967). Αυτός ο όρος περιγράφει το πλεονέκτημα που συσσωρεύει ένας επιστήμονας, μέσω της αποτελεσματικότητας του να έχει φοιτήσει σε πανεπιστήμια υψηλής κατάταξης.

Η Margaret Rochiter, σχολιάζοντας το φαινόμενο κατά Ματθαίον, παρατήρησε ότι η μη αναγνώριση ή η δυσκολία που παρουσιάζεται στην αναγνώριση μιας γυναίκας επιστήμονα, όσον αφορά τη σύγκρισή της με τους άνδρες συναδέλφους της, προστίθεται και το θέμα του φύλου της, δηλαδή άλλο ένα επιπλέον εμπόδιο στην καριέρα της, το οποίο στην πραγματικότητα είναι πιο δύσκολο να ξεπεραστεί σε σχέση με την ασημότητα. Αυτή όρισε το παράλληλο, αλλά αντίστροφο φαινόμενο, το “Φαινόμενο Ματίλντα”, που αναφέρεται στην τάση που επικρατεί στις επιστήμες, οι γυναίκες να είναι αγνοημένες ή να μην τους αποδίδεται η αναγνώριση που τους άξιζε στην εποχή τους. Το όνομα Ματίλντα προέρχεται από την Matilda Gage, μια συγγραφέα του δεκάτου ενάτου αιώνα και γνωστή φεμινίστρια, που είχε αντιμετωπίσει την αδιαφορία και τον παραγκωνισμό σε τέτοιο βαθμό, ώστε το όνομά της υιοθετήθηκε από τη θεωρία που δείχνει τη ρατσιστική αντιμετώπιση των γυναικών από τους άνδρες. Ο χρόνος τότε ήταν πλέον ώριμος για την

αποδοχή των παραπάνω ιδεών των σχετικών με την παγκοσμιοποίηση της επικοινωνίας της επιστήμης, την ανάπτυξη των γνώσεων και των δημοσιευμένων αποτελεσμάτων, την αυξανόμενη εξειδίκευση, καθώς και την αυξανόμενη σημασία της διεπιστημονικότητας στην επιστημονική έρευνα. Εκείνη την εποχή, αναπτύχθηκαν επίσης βασικά μοντέλα για την επιστημονική εργασία. Μεταξύ αυτών των μοντέλων είναι τα πρώτα για τις βασικές έννοιες της επιστημονικής ανακοίνωσης, καθώς επίσης για την ανάπτυξη και τη γήρανση των πληροφοριών.

Παρότι λοιπόν η Βιβλιομετρία χρησιμοποιήθηκε ως πρότυπο για τη μέτρηση της παραγωγής των δημοσιεύσεων των επιστημόνων πριν από σχεδόν έναν αιώνα, ο όρος εισήχθη για πρώτη φορά, όπως προαναφέρθηκε, από τον Alan Pritchard στην εργασία του με τίτλο «Στατιστική Βιβλιογραφία ή Βιβλιομετρία;» που δημοσιεύθηκε το 1969. Αυτό όμως που κατέστησε δυνατή σε μεγάλο βαθμό την ποσοτική ανάλυση των επιστημονικών δημοσιεύσεων ήταν το έργο του Eugene Garfield στη δεκαετία του 1960 και των δεικτών, που αυτός εισήγαγε με την ονομασία Social-Arts-and Humanities Science Citations Indexes, μέσω του Ινστιτούτου για την επιστημονική πληροφορία (ISI). Αρχική ιδέα και στόχος του Garfield ήταν να προσφέρει στους ερευνητές έναν γρήγορο και αποτελεσματικό τρόπο εύρεσης δημοσιευμένων άρθρων, που πραγματεύονταν τους διάφορους τομείς της έρευνάς τους (Garfield, 1968). Σύντομα, όμως, επέκτεινε τη μελέτη και το έργο του, καθώς ασχολήθηκε με την αξιολόγηση των αναφορών που συντάσσονται, έτσι: «Το συμπέρασμα που εξάγεται εύλογα είναι ότι, όσο η επιστημονική επιχείρηση γίνεται μεγαλύτερη και πιο σύνθετη και ο ρόλος της στην κοινωνία καθίσταται ολοένα και πιο κρίσιμος, τόσο πιο δύσκολη, δαπανηρή αλλά και απαραίτητη θα καταστεί η αξιολόγηση και ο σαφής προσδιορισμός των μεγαλύτερων και σπουδαιότερων συνεισφορών» (Garfield, 1979b). Ο Garfield προσπάθησε να απεικονίσει την ανάλυση των αναφορών ως ένα θεμιτό και πρακτικό εργαλείο για την αξιολόγηση της επιστημονικής παραγωγής.

Η ύπαρξη του δείκτη αναφορών κοινωνικών επιστημών (SCI) δεν αποτέλεσε μόνο την κινητήριο δύναμη για τη διεξαγωγή ενός μεγάλου αριθμού βιβλιομετρικών μελετών, αλλά ευνόησε παράλληλα και την ανάδυση μιας νέας γενιάς επιστημόνων – ερευνητών της βιβλιομετρικής ανάλυσης, υποστηρίζοντας, χαρακτηρίζοντας και καθιερώνοντας τον επιστημονικό αυτό κλάδο με τον όρο «Επιστήμη της Επιστήμης» (Price, 1965). Ο Derek de Solla Price, υπέρμαχος αυτής της μεθοδολογίας, με μεγάλη σφαιρα επιρροής, προσπάθησε να υιοθετήσει και να ακολουθήσει μια προσέγγιση στην επιστήμη που ήταν ανεξάρτητη από αυτήν που οι επιστήμονες υιοθετούσαν και εφάρμοζαν. Σύμφωνα με τον Price, η επιστήμη θα μπορούσε να μετρηθεί με βάση το δημοσιευμένο υλικό και θα μπορούσε να αναλυθεί ανεξάρτητα από τους επιστήμονες. Σύμφωνα με την άποψη που ο ίδιος έχει εκφράσει, οι επιστήμονες ήταν ειδικοί οι οποίοι, όμως, δεν εξακολουθούσαν να θεωρούνται ειδικοί όταν πραγματεύονταν άλλα αντικείμενα, εκτός των αντίστοιχων πεδίων της έρευνάς τους. Έγραψε: «Ακριβώς όπως η επιστήμη των οικονομικών έχει γίνει ένα πολύτιμο βοήθημα για τη λήψη αποφάσεων στα χέρια της κυβέρνησης και του βιομηχανικού κόσμου, καθώς και ένα ανεξάρτητο αντικείμενο ακαδημαϊκών σπουδών, μπορεί να σημαίνει ότι είμαστε μάρτυρες της γέννησης μιας παρόμοιας επιστημονικής αξιολόγησης και ανάλυσης του κόσμου της επιστήμης» (Price, 1964). Ο Price προέβλεψε ότι, στο εγγύς μέλλον, η ανάλυση των αναφορών θα μπορούσε να χρησιμοποιηθεί στη διαδικασία αξιολόγησης από συναδέλφους.

Σε αυτόν τον τομέα, Ρώσοι ερευνητές -ανατρέχοντας στη δεκαετία του 1930- συνέδεσαν τις επιστημονικές αναλύσεις με τις κοινωνικές επιστήμες, με σκοπό την παροχή μεθοδολογικών περιγραφών των διαφόρων κλάδων. Τα συστήματα μέτρησης που ανέπτυξαν, οδήγησαν στη δημιουργία ενός νέου πεδίου, που ονομάστηκε *Naukometrika* (κυριολεκτικά σημαίνει «η μέτρηση της επιστήμης») και αποτέλεσε πρόδρομο της Βιβλιομετρίας.

Ο Garfield, για να επανέλθουμε στον ιδρυτή του Ινστιτούτου για την επιστημονική πληροφορία (ISI), πιστώθηκε επίσης και την συν-αναφορά (co-citation) ως ένα μέτρο της ομοιότητας μεταξύ δύο ή περισσότερων άρθρων. Αν δύο άρθρα δηλαδή, συχνά εμφανίζονται μαζί στις λίστες αναφοράς (co-cited), είναι πιθανό να έχουν κατά κάποιο τρόπο ομοιότητες. Αυτό, με απλά λόγια, σημαίνει ότι αν συλλογές άρθρων είναι διαταγμένες σύμφωνα με τον αριθμό συν-αναφορών τους, αυτό θα πρέπει να παράγει ένα πρότυπο που αντανακλά τις σχέσεις μεταξύ γνωστικών επιστημονικών περιοχών.

Το 1973, ο Robert Merton θεμελίωσε τη θεωρία της κοινωνιολογίας της επιστήμης, στην οποία αποδεικνύει ότι οι αναφορές είναι ο τρόπος με τον οποίο οι μελετητές αναγνωρίζουν την επιρροή των προηγούμενων εργασιών. Με βάση αυτό, η αναφορά χρησιμοποιείται ως δείκτης της επιστημονικής αξίας της έρευνας.

Στις δεκαετίες του 1970 και του 1980, η Βιβλιομετρία γνώρισε μια απότομη άνοδο και βρήκε ένα νέο προσανατολισμό. Από την αρχή της δεκαετίας του ογδόντα, η Βιβλιομετρία μπόρεσε να εξελιχθεί σε μια

ξεχωριστή επιστημονική οντότητα με ένα συγκεκριμένο προφίλ έρευνας, πολλά υποπεδία και τις αντίστοιχες επιστημονικές δομές επικοινωνίας. Το περιοδικό *Scientometrics* πρωτοδημοσιεύθηκε το 1979 ως το πρώτο περιοδικό που ειδικεύεται στα βιβλιομετρικά θέματα, τα πρώτα σχετικά διεθνή συνέδρια ξεκίνησαν το 1983, το δε περιοδικό Αξιολόγησης της Έρευνας, "Research Evaluation", πρωτοδημοσιεύθηκε το 1991. Τον κύριο λόγο για την εξέλιξη αυτή μπορεί να τον δει κάποιος στη διαθεσιμότητα των μεγάλων βιβλιογραφικών βάσεων δεδομένων σε μορφή αναγνώσιμη από μηχανή, όπως επίσης και στην ταχεία ανάπτυξη της επιστήμης των υπολογιστών και της τεχνολογίας. Αυτό κατέστησε δυνατό το γεγονός οι μετρήσεις της επιστήμης μπόρεσαν να ενεργοποιηθούν και να διαχειριστούν και εκτός της επικράτειας των ΗΠΑ. Καταρχάς, τα τέλη αδειών χρήσης είχαν ως αποτέλεσμα, τουλάχιστον στη δεκαετία του '80, σοβαρούς περιορισμούς, αλλά η τεχνολογία της δεκαετίας του '90 έφερε την επανάσταση. Η «On-line βιβλιομετρία», ωστόσο, παραμένει ένα όνειρο. Η χρηματοδότηση των μεγάλων έργων φαίνεται να έχει γίνει ο κανονικός τρόπος χρηματοδότησης της έρευνας στη Βιβλιομετρία. Από τις "Μικρές...μετρικές" το ερευνητικό πεδίο έχει αλλάξει σε «Μεγάλες ... μετρικές".

Η δημοσίευση πολλών βιβλίων για την ολοκληρωμένη Βιβλιομετρία, μεταξύ άλλων, από τους Haitun (1983), Ravichandra Rao (1983), Bujdoso (1986), van Raan (1988), Courtial (1990), Egghe και Rousseau (1990) αντανakλά αυτή τη διαδικασία. Το γεγονός ότι οι βιβλιομετρικές μέθοδοι είναι ήδη εφαρμοσμένη στον επιστημονικό τομέα «Βιβλιομετρία» από μόνο του δηλώνει, επίσης, την ταχεία ανάπτυξη της επιστημονικής περιοχής. Στις δεκαετίες του 1970 και του 1980, η Βιβλιομετρία βρήκε ένα νέο προσανατολισμό:

- Η Βιβλιομετρία εξελίχθηκε από ένα αόρατο κολλέγιο, από μια υπο-περιοχή της βιβλιοθηκονομίας σε ένα όργανο για την αξιολόγηση και συγκριτική αξιολόγηση. Αυτό μπορεί να θεωρηθεί ως μια «αλλαγή προοπτικής».
- Ως συνέπεια αυτής της αλλαγής προοπτικής, νέα πεδία εφαρμογών και προκλήσεις ανοίγονται για τη Βιβλιομετρία, αλλά σχεδιάζονται ακόμη αρκετά εργαλεία για χρήση τους στην επιστημονική πληροφόρηση, την ανάκτηση πληροφοριών και τις βιβλιοθήκες.

1.2.1 Ιδρύματα κέντρων έρευνας και παροχής υπηρεσιών

Γερμανία:

Ένα από αυτά τα ερευνητικά κέντρα, το «Dokumentationsstelle für Versorgungsmedizin», ιδρύθηκε το 1956 από τον Otto Nacke, που παρέμεινε επικεφαλής του ινστιτούτου μέχρι τη συνταξιοδότησή του, το 1980. Το κέντρο βρισκόταν στο Bielefeld και ήταν επικεντρωμένο στην τεκμηρίωση στο χώρο της ιατρικής. Στη συνέχεια, άλλαξε το όνομά του σε «Institut für Dokumentation und Information über Sozialmedizin und öffentliches Gesundheitswesen» (IDIS) το 1976.

Ο Otto Nacke ήταν αυτός που επινόησε τον όρο *Πληροφοριομετρία*, το 1979. Το δεύτερο κέντρο της Γερμανίας, το «Center for Science Studies», το διηύθυνε ο Peter Weingart και βρισκόταν στο Πανεπιστήμιο του Bielefeld. Το διαδέχθηκε το «Institute for Science & Technology Studies». Ο διευθυντής του, Peter Weingart, επικεντρώθηκε για πρώτη φορά στην Κοινωνιολογία και τη φιλοσοφία της Επιστήμης (κυρίως στη δεκαετία του 1970), αλλά αργότερα επέκτεινε την έρευνά του επίσης σε μελέτες ποσοτικής επιστήμης.

Ουγγαρία:

Το ISSRU στη Βουδαπέστη ιδρύθηκε από τον Tibor Braun στις αρχές της δεκαετίας. Το κέντρο αυτό στεγάζεται στη Βιβλιοθήκη της Ουγγρικής Ακαδημίας Επιστημών. Οι πρώτες διεθνείς επιστημομετρικές δημοσιεύσεις από αυτό το κέντρο χρονολογούνται γύρω στο 1976.

Ολλανδία:

Το ινστιτούτο/κέντρο [Centre for Science and Technology Studies \(CWTS\)](#) στο Πανεπιστήμιο του Leiden διευθύνεται από τον Athony van Raan. Ιδρύθηκε γύρω στο 1982 και αρχικά είχε ονομαστεί Ινστιτούτο LISBON. Το προφίλ του είναι παρόμοιο με εκείνο του ISSRU στη Βουδαπέστη (επικεντρώνεται στην Επιστημομετρία και στη Βιβλιομετρία).

Το δεύτερο κέντρο της Ολλανδίας, το «Dept. Science Dynamics» στο Πανεπιστήμιο του Άμστερνταμ μετονομάστηκε και αναδιαρθρώθηκε πολλές φορές. Επικεφαλής του είναι ο Loet Leydesdorff, ο οποίος έχει δημοσιεύσει τα πρώτα αποτελέσματα σε διεθνή επιστημονικά περιοδικά γύρω στο 1980.

Γαλλία:

Η ομάδα των William Turner, Michel Callon, Jean-Pierre Courtial και οι συνεργάτες τους στην Ecole Mines στο Παρίσι επικεντρώθηκε σε διαρθρωτικά ζητήματα, όπως η χαρτογράφηση και οπτικοποίηση της

επιστήμης (στην πραγματικότητα με βάση την ανάλυση συναφών λέξεων). Η ομάδα ήταν ήδη ενεργή στις αρχές του 1980.

Το δεύτερο κέντρο της Γαλλίας, το «Observatoire des Sciences et des Techniques» (OST) ιδρύθηκε ως μια δια-θεσμική πλατφόρμα, το 1990. Το OST, με επικεφαλής τον Remi Barré, ήταν ένας από τους πρώτους οργανισμούς στην Ευρώπη που εξέδιδε ανά διετία εκθέσεις σχετικά με τους Δείκτες Επιστήμης και Τεχνολογίας.

Ισπανία:

Το «Centre of scientific information and Documentation» CINDOC ξεκίνησε λίγο αργότερα στη δεκαετία του 1980 (διεθνώς ορατό περίπου από το 1985). Πρόσφατα το Ινστιτούτο άλλαξε το όνομά του σε IEDCYT. Η Isabel Gomez είναι πρόεδρος του.

1.2.2. Η εκπαιδευτική επικοινωνία

Η Εκπαιδευτική επικοινωνία επικεντρώνεται σε:

- τακτική δημοσίευση των επιστημονομετρικών άρθρων στο *Czechoslovak Journal of Physics*, απ' το 1970 περίπου.
- διεθνή επιστημονικά περιοδικά.
 - ο - Scientometrics,
 - ο -Research Evaluation,
 - ο -Journal of Informetrics,
- επιμέλεια από τον Otto Nacke των βιβλίων, με τίτλο «Scientometrie und Bibliometrie in Planung und Forschung» (1976) και «Zitatenanalyse und verwandte Verfahren» (1979).
- κυκλοφορία του πρώτου «Handbook of Quantitative Science and Technology Research» που εκδόθηκε από τον A.J.F van Raan εμφανίζεται το 1988.
- οργάνωση διεθνών συνεδρίων
 - ο - International Conference for Informetrics and Scientometrics (ανά διετία από το 1987),
 - ο - International Conference for Science and Technology Indicators (ανά διετία από το 1988),
 - ο - CollNet (από το 1998),
 - ο - Foundation of the International Society for Informetrics and Scientometrics (ISSI, 1993).

Μεθοδολογία

Κατά την περίοδο αυτή, έλαβε χώρα η ανάπτυξη μιας ειδικής επιστημονομετρικής μεθοδολογίας.

1. Η συν-ανάλυση με αναφορές είχε προταθεί για τη δομική χαρτογράφηση της επιστήμης. Το ISI εξέδωσε τη συν-αναφορά βασισμένη στο Atlas of Science.
2. Περίπου μια δεκαετία αργότερα, οι Callon και άλλοι (1983) ανέπτυξαν μια άλλη διαδικασία νοητικής χαρτογράφησης που ονομάζεται Leximappe, η οποία βασίστηκε στην ανάλυση ζευγών λέξεων.
3. Αργότερα, αυτές οι μέθοδοι συμπληρώθηκαν από και σε συνδυασμό με άλλες που βασίζονται σε κείμενο (συχνότητα όρου) και σε τεχνικές αναφορών (βιβλιογραφικές ζεύξεις, άμεση σύνδεση αναφοράς, συγγραφέας συν-αναφορά).
4. Η ανάπτυξη συνεπών συστημάτων επιστημονομετρικών δεικτών για την αξιολόγηση επιδόσεων έρευνας στο ISSRU (Βουδαπέστη, Ουγγαρία) και CWTS (Leiden, Ολλανδία).
5. Η δεκαετία του 1980 είναι χαρακτηριστική για τα σημαντικά βήματα προς ιδρυματοποίηση της Επιστημονομετρικής και Πληροφοριομετρίας με σημαντικές πρωτοβουλίες και ιδρυματοποίηση του πεδίου στην Ευρώπη.

1.2.3. Η Βιβλιομετρία στη δεκαετία του 90 και στη νέα χιλιετία

Η θεαματική εξέλιξη της Βιβλιομετρίας τη δεκαετία του '90 οφείλεται σε μεγάλο βαθμό στην επανάσταση της Τεχνολογίας της Πληροφορίας (Information Technology), της οποίας είμαστε πρόσφατα μάρτυρες.

Οι ακόλουθες εξελίξεις και η συνεργία τους έχουν διευκολύνει καινοτομίες και την δημοφιλία του τομέα ερευνάς μας.

- Διαθεσιμότητα βάσεων δεδομένων
- Τη δεκαετία του 1970 και του 1980, η πρόσβαση στις ηλεκτρονικές εκδόσεις βιβλιογραφικών βάσεων δεδομένων, κατάλληλες για βιβλιομετρική χρήση, ήταν προνόμιο πολύ λίγων ιδρυμάτων σε όλον τον κόσμο. Αυτό αλλάζει τη δεκαετία του 1990, όταν οι εκδόσεις των SCI, Medline και άλλων βάσεων δεδομένων έγιναν διαθέσιμες σε πανεπιστήμια και ιδρυματικές βιβλιοθήκες.
- Ανάπτυξη υλικού
- Η παλαιότερη επεξεργασία πληροφορίας που έτρεχε σε ακριβούς κεντρικούς υπολογιστές μπορούσε σταδιακά να μεταφερθεί σε ιδρυματικούς διακομιστές και PC (φορητοί υπολογιστές βιβλιομετρίας).
- Χρήση λογισμικού
- Άνοιξε τη Βιβλιομετρία σε ένα πιο πλατύ γκρουπ χρηστών μεταξύ επιστημόνων και βιβλιομετρών «ημι-επαγγελματιών»
- Διαδίκτυο και Παγκόσμιος Ιστός
- Διευκόλυνε τη συνεργασία μεταξύ βιβλιομετρών και επιστημόνων («Συνδεδεμένων βιβλιομετρών»)
- Η είσοδος της Βιβλιομετρίας στην Ευρωπαϊκή, εθνική και τοπική Science & Technology
- Το περιοδικό «Science and Engineering Indicators» (SEI) εκδίδεται από τη National Science Board (USA). Η πρώτη έκδοση εμφανίστηκε το 1993, ενώ από το 1996 το SEI εκδίδεται κάθε δύο έτη.
- Η Ευρωπαϊκή Επιτροπή εκδίδει τακτικά τις Ευρωπαϊκές Αναφορές για τους Δείκτες Επιστήμης και Τεχνολογίας από το 1994.
- Η Ετήσια Αναφορά Επίδοσης (OST). Αυτή η αναφορά μετράει τα ιδρυματικά ερευνητικά αποτελέσματα, επιτρέποντας την σύγκριση των επιδόσεών τους με τα γαλλικά ή ευρωπαϊκά ιδρύματα.
- Το «Ολλανδικό Παρατηρητήριο Επιστήμης και Τεχνολογίας» (NOWT) είναι μια συνεργασία μεταξύ του CWTS (Leiden) και του UNU-MERIT (Maastricht). Οι Αναφορές των Δεικτών Επιστήμης και Τεχνολογίας εκδίδονται τακτικά από το 1994.
- Το Flemish Indicator Book on Science, Technology and Innovation καλύπτει ένα συγκριτικά ευρύ φάσμα και εμφανίζεται κάθε διετία από το 1999. Το βιβλίο εκδίδεται από την ECOOM σε συνεργασία με την κυβέρνηση της Φλαμανδίας.
- Η Βιβλιομετρία χρησιμοποιείται στην παρακολούθηση των δημόσιων πόρων και σε διαδικασίες λήψης στρατηγικών αποφάσεων.
- Οι κυβερνήσεις χρησιμοποιούν βιβλιομετρικές πληροφορίες, όχι μόνο με τη μορφή χρηματοδότησης, αλλά και στην παρακολούθηση και πρόβλεψη των διαδικασιών.
- Τα Ιδρύματα χρησιμοποίησαν τη Βιβλιομετρία για την παρακολούθηση των προγραμμάτων και στις στρατηγικές λήψεων αποφάσεων.

Μηχανισμοί χρηματοδότησης και ασκήσεις αξιολόγησης της έρευνας

- Αλλαγές στη Διαδικασία Αξιολόγησης της Έρευνας (RAE) στο Ηνωμένο Βασίλειο έχουν προγραμματιστεί για τη μελλοντική, σταθμισμένη χρηματοδότηση της ποιότητας της έρευνας της τριτοβάθμιας εκπαίδευσης μετά το 2008.
- Ένας από τους κύριους μηχανισμούς χρηματοδότησης για τη βασική επιστήμη στα φλαμανδικά πανεπιστήμια είναι τα Bijzonder Onderzoeks fonds (BOF). Μέρος της

διαδικασίας κατανομής βασίζεται στη δημοσίευση και παράθεση δεδομένων που προέρχονται από το "Web of Science" (Thomson Reuters) από το Centre for Research & Development Monitoring (Expertise centrum Onderzoeken Ontwikkelings monitoring, ECOOM). (Debackere&, 2004).

- Από το 2005, η διαδικασία κατανομής για τη βασική χρηματοδότηση της έρευνας στη Νορβηγία, περιλαμβάνει έναν δείκτη εκροών για επιστημονικές δημοσιεύσεις. Το νορβηγικό μοντέλο εφαρμόζεται τώρα στη Δανία.

Σήμερα χρησιμοποιούνται διαφορετικές μορφές των επιδόσεων Ερευνητικών Πανεπιστημιακών Συστημάτων σε Χρηματοδοτήσεις Pseudo Random Function (PRFS), βασιζόμενες στα βιβλιομετρικά συστατικά.

Περαιτέρω παραδείγματα για PRFS's με ποσοτικά στοιχεία είναι προς το παρόν:

- Ο Γενικός Δείκτης του πανεπιστημίου επιχορηγήσεων, το Research Quality Framework (RQF) και το Excellence in Research for Australia (EXE) στην Αυστραλία.
- Η βασισμένη στην απόδοση χρηματοδότηση της έρευνας Performance-Based Research Fund (PBRF) στη Νέα Ζηλανδία.
- Ο μαθηματικός τύπος χρηματοδότησης για την κατανομή των πόρων των πανεπιστημίων στη Φινλανδία.
- Η Valutazione Triennale della Ricerca(VTR) / Valutazione Quinquennale della Ricerca (VQR).

Κατά τη διάρκεια της τελευταίας δεκαετίας, η Βιβλιομετρία έχει ενταχθεί στο πλαίσιο των μεταπτυχιακών ή διδακτορικών προγραμμάτων πολλών ευρωπαϊκών πανεπιστημίων. Ειδικά μαθήματα προσφέρονται επίσης από τα πανεπιστήμια και τα ερευνητικά κέντρα, όπως:

- «Measuring Science» στο CWTS, στο Πανεπιστήμιο του Leiden.
- «Road show seminars» και διοργάνωση σκανδιναβικών διδακτορικών σπουδών στη Βιβλιομετρία, από το Nordic Research School in Library and Information Science (NORSLIS).
- «European Summer School for Scientometrics» (esss), οργανωμένο από κοινού από το Πανεπιστήμιο της Βιέννης, το Humboldt-Universität Berlin και το KU Leuven.

Ο αντίκτυπος της Βιβλιομετρίας

- Η ανάγκη για μετρήσεις στις υπηρεσίες πληροφόρησης, την πολιτική της επιστήμης και τη διαχείριση της έρευνας έχει γίνει ευρέως αναγνωρίσιμη.
- Η Βιβλιομετρία διαδραματίζει ολοένα και σημαντικότερο ρόλο στην αξιολόγηση της έρευνας και οι ποσοτικές φόρμουλες με βιβλιομετρικά συστατικά χρησιμοποιούνται ολοένα και περισσότερο κατά την κατανομή των χρηματοδοτήσεων.
- Η επιτυχής εφαρμογή των επιστημονομετρικών μεθόδων έχουν συμβάλει σε μεγάλο βαθμό στην αύξηση της δημοτικότητάς τους.
- Η ηλεκτρονική επικοινωνία, ο ιστός και η ανοικτή πρόσβαση άνοιξαν το δρόμο για τον εκδημοκρατισμό της Βιβλιομετρίας.

Όμως, έχουν απαξιώσει τον τομέα:

- γρήγορα και αναληθή στατιστικά στοιχεία και αξιολογήσεις,
- μη ενημερωμένη εφαρμογή και
- κατάχρηση της Βιβλιομετρίας.

Η ανάγκη αξιολόγησης της έρευνας

Οι βιομηχανικές κοινωνίες ήταν ιδιαίτερα θετικές και ευνοϊκές απέναντι στην εξέλιξη της επιστήμης (Bush, 1960). Από τη στιγμή που πρωτοξεκίνησαν το 1957 οι έρευνες που διεξάγονται από το Εθνικό Ίδρυμα Επιστημών (NSF), αποτελεί πεποίθηση της αμερικανικής κοινής γνώμης ότι η συμβολή της επιστήμης και της

τεχνολογίας στην κοινωνική πρόοδο είναι μείζονος σημασίας (Εθνικό Ίδρυμα Επιστημών, 1989, σσ. 170-172).

Ο ανταγωνισμός επιπροσθέτως με τη Σοβιετική Ένωση αποτέλεσε το κίνητρο ώστε να καταβάλουν οι Ηνωμένες Πολιτείες σημαντικές προσπάθειες στον τομέα Ε & Α, κατά τη δεκαετία του '60. Ανάμεσα στις προσπάθειες αυτές συγκαταλέγεται και η δημιουργία ποικίλων υπηρεσιών, γραφείων, οργανισμών και ιδρυμάτων. Ακολούθησαν, εν συνεχεία, παρόμοιες κινήσεις στην Ευρώπη, τη Σοβιετική Ένωση και την Ιαπωνία.

Μια αλλαγή πραγματοποιήθηκε στη δεκαετία του 1970 όταν η επιστήμη παύει να θεωρείται πλέον ως μια επιχείρηση στην οποία η κοινωνία θα μπορούσε να επενδύει απλόχερα και χωρίς όρια. Η πρώτη φάση αυτής της πλήρους μετατόπισης της κοινής γνώμης ήταν απόρροια της επιβράδυνσης της οικονομικής ανάπτυξης. Ταυτόχρονα, προήλθε και από μια πιο κριτική στάση που επικρατούσε πλέον και που λάμβανε, ουσιαστικά, υπόψη της τις αρνητικές συνέπειες της επιστημονικής έρευνας όπως ότι η επιστήμη και η τεχνολογία ήταν δαπανηροί τομείς, ενώ οι επενδύσεις στον τομέα της έρευνας δεν εξασφάλιζαν αυτομάτως την επίλυση των περιβαλλοντικών προβλημάτων ή των κοινωνικών προβλημάτων, όπως το χάσμα μεταξύ των βιομηχανικών χωρών και του Τρίτου Κόσμου.

Αυτό οδήγησε σε έντονη ανησυχία αναφορικά με την αποδοτικότητα της βασικής έρευνας, κατά κύριο λόγο, συνδυαστικά και με το γεγονός ότι ολοένα και περισσότερο κέρδιζε έδαφος η αντίληψη πως οι ερευνητές ήταν «παραγωγοί της επιστήμης», που έπρεπε να λογοδοτήσουν για τους πόρους που έλαβαν. Επιπλέον, η φοιτητική εξέγερση επιδείνωσε την εικόνα των πανεπιστημίων, καθώς την εξουσία και το κύρος των επιστημόνων και των αποφοίτων.

Τέτοιου είδους γεγονότα προκάλεσαν την καχυποψία του ευρύτερου κοινωνικού συνόλου και τη δυσπιστία της κοινής γνώμης απέναντι στην επιστήμη και την τεχνολογία. Ο νέος στόχος ήταν να παραχθεί προστιθέμενη αξία, που συμβάλλει στη διατήρηση των φυσικών πόρων, ο περιορισμός των ποσοστών ρύπανσης και η δημιουργία ενός πιο αποτελεσματικού συστήματος έρευνας, που θα επέτρεπε να αξιοποιηθούν σωστότερα οι ήδη υπάρχουσες επιστημονικές γνώσεις. Μέσα σε αυτό το ιστορικό πλαίσιο, κάνει την εμφάνισή της η αξιολόγηση της επιστημονικής έρευνας και του επιστημονικού έργου εν γένει.

Ως αποτέλεσμα των παραπάνω, οι μέθοδοι των κοινωνικών και ανθρωπιστικών επιστημών (που θεωρούνται ως οι «ήπιες» επιστήμες) χρησιμοποιήθηκαν για την ανάλυση των «σκληρών» επιστημών. Για τον σκοπό αυτό, απαιτούνταν ποσοτικά κριτήρια και μέτρα. Με άλλα λόγια, έπρεπε να βρεθούν μέθοδοι ποσοτικοποίησης, συγκέντρωσης και σύγκρισης των δεικτών. Η καθιέρωση μιας μέτρησης της επιστήμης κατέστη αναπόφευκτη. Η σπουδαία αυτή αλλαγή άνοιξε το δρόμο για την ανάλυση της επιστήμης και της τεχνολογίας και ευνόησε την εισαγωγή της Βιβλιομετρίας στις επιστήμες.

1.3. Οι διάφορες μετρικές

Πλήθος όρων χρησιμοποιείται για να περιγράψουμε τον κλάδο της Επιστήμης της πληροφορίας, η οποία άλλωστε είναι το θέμα συζήτησης. Αυτοί οι όροι έχουν επικαλυπτόμενες, αλλά όχι ταυτόσημες έννοιες και έχουν επίσης βιώσει την αλλαγή της δημοτικότητας. Οι κύριοι όροι εκτός των άλλων που χρησιμοποιούνται είναι Βιβλιομετρία (Bibliometrics), Επιστημομετρία (Scientometrics) και Πληροφοριομετρία (Informetrics). Υπάρχουν επίσης σχετικοί όροι, για παράδειγμα κάποια ουσιαστικά των λέξεων αυτών ή επιρρήματα με την ίδια ρίζα. Άλλοι όροι είναι η Στατιστική Βιβλιογραφία, η Τεχνομετρία κ.ά.

Αυτοί οι ορισμοί υποδεικνύουν αξιοσημείωτη επικάλυψη των όρων, αλλά δεν είναι απαραίτητα συνώνυμοι. Με την πάροδο του χρόνου, η χρήση των όρων έχει αλλάξει, με τον παλαιό όρο Βιβλιομετρία να παραμένει σταθερός, ενώ οι νεότεροι όροι Πληροφοριομετρία και Επιστημομετρία να χρησιμοποιούνται όλο και περισσότερο. Ειδικότερα ερευνητές, εκτός της Επιστήμης της πληροφορίας, θα συνεχίσουν να χρησιμοποιούν τον πιο συνήθη όρο, Βιβλιομετρία.

Ο όρος Πληροφοριομετρία είναι, ίσως, ο πιο γενικός εκ των τριών. Η Πληροφοριομετρία θα μπορούσε να υπάγεται στην Επιστημομετρία και ειδικότερα στη Βιβλιομετρία. Ωστόσο, όσοι εργάζονται στους τρεις αυτούς μετρικούς κλάδους, συνεχίζουν να χρησιμοποιούν την ορολογία που πιστεύουν ότι περιγράφει καλύτερα τη δουλειά τους.

Παρακάτω ορίζονται οι διάφορες μετρικές ως εξής:

Πληροφοριομετρία: Η Πληροφοριομετρία (Informetrics) είναι η μελέτη των ποσοτικών πτυχών της πληροφορίας σε οποιαδήποτε μορφή τεκμηρίου και σε κάθε κοινωνική ομάδα, όχι μόνο στους επιστήμονες.

Αυτό περιλαμβάνει την παραγωγή, διάδοση και χρήση των πληροφοριών. Σύμφωνα με άλλον χρησιμοποιούμενο ορισμό, Πληροφοριομετρία είναι «η εφαρμογή μαθηματικών μεθόδων σε αντικείμενα της Επιστήμης των πληροφοριών για την περιγραφή και την ανάλυση των φαινομένων της, την ανακάλυψη των νόμων της και τη στήριξη των αποφάσεων της» (Nacke, 1979).

Αυτός ο πιο πρόσφατος μετρικός όρος, “Πληροφοριομετρία”, προέρχεται από τον γερμανικό όρο “infometrie” και χρησιμοποιήθηκε πρώτη φορά το 1979 από τον Nacke, για να καλύψει ένα κομμάτι της Επιστήμης της πληροφορίας που είχε να κάνει με τη μέτρηση φαινομένων πληροφοριών και την εφαρμογή μαθηματικών μεθόδων σε προβλήματα ελέγχου στη Βιβλιομετρία και, κατά κάποιο τρόπο, στη θεωρία Ανάκτησης της πληροφορίας και ίσως και σε πιο ευρεία θέματα.

Τον ακόλουθο χρόνο, ο Nacke πρότεινε την Επιστημομετρία ως συγγενικό όρο της Πληροφοριομετρίας. Ο Bonitz (1982) ασχολείται με την εισαγωγή του όρου Πληροφοριομετρία και τη συγκρίνει με τους όρους της Βιβλιομετρίας και της Επιστημομετρίας. Θεωρεί απαραίτητη την εισαγωγή ενός νέου όρου, για να γίνει διάκριση της Επιστημομετρίας, ώστε να μη δημιουργηθεί σύγχυση στον τομέα της Επιστήμης της επιστήμης και της Επιστήμης των βιβλιοθηκών. Το 1984, το Ινστιτούτο Επιστημονικής και Τεχνικής Πληροφόρησης ‘All-Soviet Institute for Scientific and Technical Information’ (VINITI) καθιέρωσε μια διεθνή επιτροπή Πληροφοριομετρίας, υπό την προεδρία του Nacke, όπου ο όρος Πληροφοριομετρία χρησιμοποιήθηκε ως γενικός όρος για τη Βιβλιομετρία και την Επιστημομετρία. Η χρήση αυτή εγκρίθηκε στο VINITI, χάρη στον Gorkova, υπό τη ρωσική ονομασία infometriya (Πληροφοριομετρία).

Στο πρώτο διεθνές συνέδριο για τη Βιβλιομετρία και τις θεωρητικές πτυχές ανάκτησης πληροφοριών το 1988, ο Brookes υποστήριξε πως ο όρος «Πληροφοριομετρία», στον οποίο υπάγεται η Βιβλιομετρία και η Επιστημομετρία, για αρχειακές και ηλεκτρονικές πληροφορίες, ίσως έχει μέλλον. Ο όρος Πληροφοριομετρία υιοθετήθηκε το 87/88. Στο δεύτερο συνέδριο, ο Brookes (1990) ενέκρινε την «Πληροφοριομετρία» ως γενικό όρο για την Επιστημομετρία και τη Βιβλιομετρία, με την Επιστημομετρία να τείνει περισσότερο στις πολιτικές επιστήμες και τη Βιβλιομετρία στις επιστήμες των βιβλιοθηκών. Ο όρος «Πληροφοριομετρία» ενισχύθηκε στα πρακτικά του τρίτου συνεδρίου της τρίτης διεθνούς διάσκεψης για την Πληροφοριομετρία (1992), αλλά αποδυναμώθηκε στο τέταρτο συνέδριο της διεθνούς διάσκεψης της Βιβλιομετρίας, Πληροφοριομετρίας και Επιστημομετρίας. Τα πρακτικά του τέταρτου συνεδρίου δημοσιεύτηκαν σε τέσσερις (4) διαφορετικούς τόμους, οι τρεις (3) εκ των οποίων αποτέλεσαν ολόκληρα θέματα σε περιοδικά. Σε αυτή τη διάσκεψη ιδρύθηκε η διεθνής εταιρία της Επιστημομετρίας και της Πληροφοριομετρίας International Society for Scientometrics and Informetrics (ISSI). Μια ειδική έκδοση σχετική με την Πληροφοριομετρία έκανε την εμφάνιση της στο περιοδικό Information Processing & Management (Tague- Sutcliffe, 1992). Εν ολίγοις, από τις αρχές της δεκαετίας του 1990, ο όρος Πληροφοριομετρία γνώρισε μεγάλη απήχηση.

Κάποιοι άλλοι ενδεικτικοί ορισμοί τη δεκαετία του '90 είναι από τους Egghe&Rousseau (1990) στον υπότιτλο του βιβλίου τους: «Πληροφοριομετρία: Ποσοτικές μέθοδοι της βιβλιοθήκης, της τεκμηρίωσης και της επιστήμης της πληροφόρησης». Η Πληροφοριομετρία καλύπτει τις εμπειρικές μελέτες της βιβλιογραφίας και των εγγράφων, καθώς και τις θεωρητικές επιστήμες των μαθηματικών ιδιοτήτων των νόμων και των κατανομών που έχουν ανακαλυφθεί.

Άλλοι χρησιμοποιούμενοι ορισμοί:

«Πληροφοριομετρία είναι η μελέτη των ποσοτικών πτυχών της πληροφορίας σε οποιαδήποτε μορφή, όχι μόνο τα αρχεία ή οι βιβλιογραφίες, και σε οποιαδήποτε κοινωνική ομάδα, όχι μόνο επιστημονική. Έτσι, εξετάζει τις ποσοτικές πτυχές της άτυπης ή προφορικής επικοινωνίας» (Tague-Sutcliffe, 1992).

«Ο όρος πληροφοριομετρία ορίζει μια πρόσφατη επέκταση των πρόσφατων βιβλιομετρικών αναλύσεων, ώστε να καλύψουν τις μη ακαδημαϊκές κοινότητες στις οποίες η πληροφορία παράγεται, κοινοποιείται και χρησιμοποιείται» (Ingwerse&Christensen, 1997).

«Πληροφοριομετρία είναι η ποσοτική μελέτη των συλλογών των μετρίου μεγέθους μονάδων των πληροφοριακών κειμένων, κατευθυνόμενων προς την επιστημονική κατανόηση της επεξεργασίας σε ένα κοινωνικό επίπεδο» (Wilson, 1999).

Επιστημομετρία (Scientometrics): Το 1969, οι Vassily Nalimov και Mluchenko (1969) επινόησαν τον όρο Επιστημομετρία (naukometriya). Όπως συνεπάγεται η ονομασία, ο όρος αυτός χρησιμοποιείται κυρίως για τη μελέτη όλων των πτυχών της βιβλιογραφίας της επιστήμης και της τεχνολογίας. Ο όρος έχει λάβει ευρεία αναγνώριση απ’ τη στιγμή που άρχισε να γίνεται γνωστό το περιοδικό "Scientometrics" από τον Tibor Braun, στην Ουγγαρία. Σύμφωνα με τον υπότιτλο του, η Επιστημομετρία συμπεριλαμβάνει όλες τις έννοιες της επιστήμης της επιστήμης, της επιστήμης της επικοινωνίας και της πολιτικής επιστήμης. Λίγο μετά την ίδρυση του, ο Nalimov έγινε ο μόνος σύμβουλος-συντάκτης. Μερικά από τα πρώιμα άρθρα του Nalimov, τα οποία

βοήθησαν στην καλλιέργεια του όρου Επιστημομετρία, είναι τα εξής: (Nalimov 1970; Nalimov & Mulchenko 1969; Nalimov et al., 1971). Μεγάλο μέρος της Επιστημομετρίας δεν έχει εμφανή διαφορά από τη Βιβλιομετρία έτσι πολλές βιβλιομετρικές έρευνες έχουν δημοσιευτεί στο περιοδικό “Scientometrics.”

Η Επιστημομετρία τυπικά ορίζεται ως η «ποσοτική μελέτη της επιστήμης και της τεχνολογίας», όπως για παράδειγμα στην πρόσφατη ειδική έκδοση της JASIS (Journal of the American Society for Information Science) σχετική με τους δείκτες της επιστήμης και της τεχνολογίας, που εκδόθηκε από τον VanRaen (1998).

Ακριβέστερα «Επιστημομετρία είναι η μελέτη των ποσοτικών πτυχών της επιστήμης, ως πειθαρχημένη ή οικονομική δραστηριότητα. Είναι μέρος της κοινωνιολογίας της επιστήμης και εφαρμόζεται στην επιστήμη της πολιτικής. Περιλαμβάνει ποσοτικές μελέτες των επιστημονικών δραστηριοτήτων, συμπεριλαμβανομένων, μεταξύ άλλων, δημοσιεύσεων και έτσι επικαλύπτει και τη Βιβλιομετρία σε κάποια σημεία» (Tague & Sutcliffe, 1992).

Η Επιστημομετρία μελετά ποσοτικές πτυχές της επιστήμης ή αφορά την «ποσοτική μελέτη της επιστήμης, της επικοινωνίας στον τομέα της επιστήμης και την πολιτική της επιστήμης.»

Στο Σχήμα 1.1 οι πολιτικό-οικονομικές πτυχές της Επιστημομετρίας καλύπτονται από το μέρος της έλλειψης της, που βρίσκεται έξω από τα όρια της Βιβλιομετρίας.

Η Επιστημομετρία ως σπουδή της επιστήμης με ποσοτικά χαρακτηριστικά εστιάζεται στα τεκμήρια: αρχεία, δημοσιευμένο και επικοινωνιακό υλικό της επιστημονικής παραγωγής και στους ποσοτικούς δείκτες που δημιουργούνται από αυτά. Οι δείκτες αυτοί χρησιμοποιούνται για να περιγράψουν την όλη κατάσταση και τη πορεία της επιστήμης. Η οριοθέτηση της σε σχέση με άλλους κλάδους και μεθοδολογίες οδηγεί συχνά σε σύγχυση, αφού μοιράζονται κοινές περιοχές (Restivo, 1993). Η κοινωνιολογία της επιστήμης όπως την γνωρίζουμε δεν ταυτίζεται με την Επιστημομετρία, γιατί αφενός η πρώτη περιλαμβάνει και ποιοτικές μεθόδους κι αφετέρου η δεύτερη είναι προσανατολισμένη και σε άλλους κλάδους πέραν της Κοινωνιολογίας, όπως Πληροφορική και Βιβλιοθηκονομία. Επίσης η Επιστημομετρία ξεχωρίζει από τη Βιβλιομετρία, η οποία αποσκοπεί στη μέτρηση των μορφωμάτων κάθε γραπτής επικοινωνίας (Broadus, 1987) κι έτσι δεν περιορίζεται στην επιστημονική επικοινωνία, ενώ από τη μεριά της ούτε η Επιστημομετρία περιορίζεται μόνο στα βιβλία. Η Βιβλιομετρία παρέχει τα μέσα για την αξιολόγηση της ερευνητικής παραγωγής (μέσω της ανάλυσης των δημοσιεύσεων), αλλά η Επιστημομετρία, αφενός μεν χρησιμοποιεί τα βιβλιομετρικά εργαλεία για τις μετρήσεις των δημοσιεύσεων, αφετέρου περιλαμβάνει και άλλους δείκτες που αφορούν στην αξιολόγηση της επιστημονικής παραγωγής. Παράδειγμα αυτής της περίπτωσης οι δείκτες της επιστήμης, με τους οποίους περιγράφονται τα διάφορα μέτρα μεταβολών στα επιστημονικά ζητήματα, που γενικώς χρησιμοποιούνται από τους δημιουργούς της επιστημονικής πολιτικής (Elkana et al., 1978).

Η θεμελίωση της Επιστημομετρίας αποδίδεται κυρίως στην πρωτοπόρα δουλειά, στις αρχές της δεκαετίας του 1960, του ιστορικού της επιστήμης, Derek de Solla Price.

Ιστομετρία (Webometrics): Στον απόηχο των εξελίξεων του διαδικτύου ορισμένοι βιβλιομέτρες επέδειξαν αναλογίες μεταξύ των διαδικτυακών βάσεων και των ερευνητικών εγγράφων και γέννησαν την ιδέα ότι το επιστημονικό περιεχόμενο του Παγκόσμιου Ιστού θα μπορούσε να αναλυθεί με τον ίδιο τρόπο, όπως και το σύστημα των επιστημονικών περιοδικών. Υπέρ-συνδέσεις μεταξύ των ιστοσελίδων συγκρίθηκαν άρθρο προς άρθρο, μέσω των αναφορών και αυτό γέννησε μια σειρά από διερευνητικές μελέτες για να διαπιστωθεί κατά πόσον θα μπορούσαν να εφαρμοστούν οι συμβατικές βιβλιομετρικές μέθοδοι σε όλο ή εν μέρει στο διαδίκτυο (Almind and Ingwersen 1997; Ingwersen 1998; Larson 1996; Rousseau 1997). Ορισμένες μέθοδοι που αναπτύσσονται στο πλαίσιο αυτό αναλύονται παρακάτω.

Είναι δυνατόν να μετρηθεί ένας διαδικτυακός παράγοντας επιρροής (WIF). Για παράδειγμα, ο αριθμός των συνδέσεων που αποδίδεται σε έναν τομέα μπορεί να διαιρεθεί με τον αριθμό των σελίδων του δικτυακού τόπου (Bjorneborn and Ingwersen 2001; Li 2003). Μια πιο ενδιαφέρουσα εναλλακτική λύση είναι να μετρήσουμε τον αριθμό των σχετικών υπέρ-συνδέσεων σε ένα ακαδημαϊκό ίδρυμα και στη συνέχεια να διαιρέσουμε τον παραπάνω αριθμό με τον αριθμό των μελετητών που συνδέονται με αυτό (Li 2003 citing Smith and Thelwall 2001; Li et al. 2003; Thelwall 2001). Τα τμηματικά αποτελέσματα των διαδικτυακών παραγόντων επιρροής ταξινομούνται παράλληλα με τις κατατάξεις και καθορίζονται από τη Διαδικασία Αξιολόγησης της Έρευνας (RAE) στο Ηνωμένο Βασίλειο (Li et al. 2003; Thelwall and Harries 2004). Με τον καθορισμό των διαδικτυακών παραγόντων επιρροής είναι δυνατόν να μετρηθεί το κύρος των πανεπιστημίων και όχι η ποιότητα του ερευνητικού τους έργου, όπως ακριβώς και τα μέτρα ανάλυσης αναφορών μετρούν τον αντίκτυπο και όχι την ποιότητα (Li, 2003).

Είναι δυνατό να χτιστούν συστάδες των ιστοσελίδων με βάση το βαθμό της εγγύτητας που παράγεται από την ανάλυση των υπέρ-συνδέσεων. Τα αποτελέσματα είναι παρόμοια με εκείνα που παράγονται από την

ανάλυση συνδυασμού αναφορών και λέξεων. Όπως και στην περίπτωση της Βιβλιομετρίας, η μέτρηση των ιστοσελίδων θα μπορούσε να χρησιμεύσει για να επισημανθούν συγκεκριμένες διασυνδέσεις εντός του επιστημονικού πεδίου. Για παράδειγμα, οι Thelwall και Wilkinson (2004) προσπάθησαν να εφαρμόσουν τις βιβλιογραφικές μεθόδους βιβλιογραφικής ζεύξης και της ανάλυσης υπέρ-συνδέσμων (αντί του συνδυασμού των αναφορών) στο διαδίκτυο. Τα αποτελέσματα δεν οδηγούν σε ασφαλή συμπεράσματα, αλλά δείχνουν προς μια νέα κατεύθυνση για τη δοκιμή των αναδυόμενων πεδίων της Ιστομετρίας.

Ανεξάρτητα από όλο το δυναμικό τους, οι υπέρ-σύνδεσμοι υποφέρουν από ένα σημαντικό μειονέκτημα σε σχέση με τις αναφορές: δεν έχουν σχέση με κυμαινόμενο χρόνο. Υπάρχει πάντα μια ημερομηνία που σχετίζεται με τις αναφορές, κάτι που δε συμβαίνει με τους υπέρ-συνδέσμους (Prime, Zitt and Bassecouard 2002, citing Egghe 2000). Η μεταβλητή του χρόνου είναι πολύ χρήσιμη για την ανάλυση των αναφορών και μπορεί να παράγει πολύ καλά αποτελέσματα, ενώ η ανάλυση της υπέρ-σύνδεσης μπορεί να μην είναι τόσο κατατοπιστική. Ένα δεύτερο, πιο πρακτικό εμπόδιο για την ανάπτυξη των μετρήσεων της ιστοσελίδας βρίσκεται στα πολλά τεχνικά προβλήματα (Li, 2003), ιδίως στο πλαίσιο της συλλογής και διαμόρφωσης των δεδομένων.

Επιπλέον μερικοί συγγραφείς λένε ότι οι υπέρ-συνδέσεις δεν θα πρέπει να αναμένεται να αντικαταστήσουν τις αναφορές (Prime, Zitt and Bassecouard 2002; VanRaen 2001). Κύριο σκεπτικό τους είναι ότι οι υπέρ-συνδέσεις δεν θα δημιουργήσουν αξιόπιστες αξιολογήσεις των ερευνητικών επιπτώσεων (Egghe, 2000).

Με λίγα λόγια, η παρούσα κατάσταση των ιστομετρικών μεθόδων είναι τέτοια, ώστε να μην χρησιμοποιούνται πραγματικά για την αξιολόγηση της έρευνας των βάσεων δεδομένων. Σε αυτό το στάδιο, η εφαρμογή τους στη χαρτογράφηση φαίνεται πιο αξιόπιστη. Ωστόσο, οι εξελίξεις στην ιστομετρία αξίζουν την παρακολούθηση, επειδή μπορεί να οδηγήσουν σε ορισμένες καινοτόμες μεθόδους αξιολόγησης της έρευνας. Ειδικότερα, το έργο της Ομάδας της στατιστικής ιστομετρικής έρευνας στο Πανεπιστήμιο Wolverhampton στο Ηνωμένο Βασίλειο φαίνεται πολύ ελπιδοφόρο.

Η Ιστομετρία, λοιπόν, είναι «η μελέτη των ποσοτικών πτυχών της κατασκευής και χρήσης των πηγών πληροφόρησης, των δομών και τεχνολογιών του World Wide Web, χαράζοντας βιβλιομετρικές και πληροφοριομετρικές προσεγγίσεις». Συνεπώς αυτός ο ορισμός καλύπτει ποσοτικές απόψεις τόσο της πλευράς της κατασκευής αλλά και της χρήσης του web, περιλαμβάνοντας τις τέσσερις κύριες περιοχές της τρέχουσας έρευνας της webometrics: (1) ανάλυση περιεχομένου σελίδας web, (2) ανάλυση δομής συνδέσμων web, (3) ανάλυση χρήσης web (όπως χρήση των Logfiles για τη συμπεριφορά αναζήτησης και ξεφυλλίσματος των χρηστών) και (4) ανάλυση της τεχνολογίας web (περιλαμβάνοντας την απόδοση των μηχανών αναζήτησης). Αυτό περιέχει υβριδικές φόρμες. Για παράδειγμα ο Pirolli και οι συνεργάτες του (1996) εξερεύνησαν τεχνικές ανάλυσης του web για αυτόματη κατηγοριοποίηση, υλοποιώντας τοπολογία γραφήματος συνδέσμων, περιεχόμενο κειμένου και ομοιότητα μεταδεδομένων, καθώς και δεδομένα χρήσης.

Η επιστήμη της webometrics προσπαθεί να μετρήσει το World Wide Web για να πάρει τη γνώση σχετικά με τον αριθμό και το είδος των υπερσυνδέσμων, τη δομή του World Wide Web και τα μοτίβα χρήσης. Σύμφωνα με τους Björneborn και Ingwersen (2004), «Ο όρος webometrics επινοήθηκε για πρώτη φορά από τους Almind και Ingwersen» (1997). Σύμφωνα με έναν άλλο ορισμό, η Webometrics έχει ορισθεί ως «η μελέτη του διαδικτυακού περιεχομένου με κυρίως ποσοτικές μεθόδους για τους κοινωνικούς, επιστημονικούς, ερευνητικούς στόχους, χρησιμοποιώντας τεχνικές που δεν είναι ειδικές για έναν τομέα έρευνας» (Thelwall, 2009). Με αυτόν τον ορισμό δίνεται έμφαση στην ανάπτυξη των εφαρμοζόμενων μεθόδων για χρήση τους στις ευρύτερες κοινωνικές επιστήμες. Σε αυτό το πλαίσιο, το πεδίο Webometrics μπορεί να φανεί ως εξ' ολοκλήρου περικλειόμενο από τη Βιβλιομετρία, γιατί τα έγγραφα του web, είτε κειμένου είτε πολυμέσων, καταγράφονται ως πληροφορία (βλ. τον παραπάνω ορισμό της Βιβλιομετρίας από τους Tague-Sutcliffe) και αποθηκεύονται στους web servers.

Στο διάγραμμα η webometrics είναι μερικώς καλυμμένη από την scientometrics, καθώς στις μέρες μας πολλές ακαδημαϊκές δραστηριότητες είναι βασισμένες στο web. Ακόμη η Webometrics είναι εντελώς περιλαμβανόμενη μέσα στο πεδίο της Cybermetrics, ο ορισμός της οποίας δίδεται παρακάτω.

Η συμπερίληψη των webometrics επεκτείνεται στο πεδίο της Βιβλιομετρίας κι έτσι η webometrics, αναπόφευκτα, θα συνεισφέρει με περαιτέρω μεθοδολογικές εξελίξεις, συγκεκριμένων web προσεγγίσεων. Καθώς οι ιδέες ρίζωσαν στις Bibliometrics, Scientometrics και Informetrics, συνέβαλαν στην εμφάνιση της Webometrics και οι ιδέες στην Webometrics, τώρα, μπορεί να συνεισφέρουν στην ανάπτυξη των περικλειόμενων πεδίων.

Κυβερνομετρία (Cybermetrics): Ο όρος cyber ως πρόθεμα χρησιμοποιήθηκε για πρώτη φορά στη λέξη cybernetics (κυβερνητική), η οποία επινοήθηκε από τον Norbert Wiener στο ομώνυμο βιβλίο του που εκδόθηκε το 1948. Ο Wiener εμπνεύστηκε τη λέξη από την ελληνική που σημαίνει κυβερνήτης, όπου εμπεριέχεται η έννοια του ελέγχου. Σύμφωνα με αυτό ο William Gibson χρησιμοποίησε τη λέξη σε μια ιστορία που δημοσιεύτηκε στο περιοδικό Omni. Έπειτα ξαναχρησιμοποίησε τη λέξη αυτή στο διήγημα Neuromancer, που εκδόθηκε το 1984. Στο διήγημα η λέξη δεν είχε την σημερινή έννοια. Με την πάροδο του χρόνου η έννοια της λέξης εξελίχθηκε και τώρα περιέχει την τεχνολογία πληροφοριών, το διαδίκτυο και την εικονική πραγματικότητα.

Δύο ευδιάκριτα μέρη στη λέξη 'cybermetrics' είναι το 'cyber' και το 'metrics'. Η συνδυαστική μορφή της λέξης metrics έχει προκύψει από τη λέξη 'metre', που με τη σειρά της προκύπτει από τη λατινική λέξη 'metrum' και την ελληνική λέξη 'μέτρον' που όλες σημαίνουν μέτρο. Έχοντας ως βάση τη σημασία των δυο συνδυαζόμενων μορφών, μπορούμε να πάρουμε τη σημασία της Κυβερνομετρικής (cybermetrics) ως της επιστήμης μέτρησης που περιλαμβάνει κυβερνο-αντικείμενα. Η Κυβερνομετρία είναι ο κλάδος της επιστήμης, η οποία χρησιμοποιεί μαθηματικές και στατιστικές τεχνικές για την ποσοτικοποίηση των διαδικτυακών τόπων ή των συστατικών στοιχείων και των εννοιών τους. Παρακολουθεί την ανάπτυξή τους, τη σταθερότητα, τη διάδοση και τη χρήση, εξετάζει την αυθεντικότητα του περιεχομένου, θεσπίζει νόμους που διέπουν αυτούς τους παράγοντες, μελετά την αποτελεσματικότητα των συστημάτων πληροφοριών στον κυβερνοχώρο, τις υπηρεσίες και τα προϊόντα του και αξιολογεί την εξέλιξη του. Η διαφορά μεταξύ των όρων Ιστομετρία και Κυβερνομετρία έγκειται στο γεγονός ότι η Κυβερνομετρία είναι πιο γενική, αναφερόμενη και σε μη-web έρευνα στο διαδίκτυο, όπως τα e-mails ή ασχολείται και με τη μελέτη των newsgroups, επιπλέον της αναζήτησης στον Ιστό. Στην Κυβερνομετρική οι ιστοσελίδες θα παίζουν τον ίδιο ρόλο, όπως τα τεκμήρια στην Βιβλιομετρική.

Στο διάγραμμα, το πεδίο Cybermetrics ξεπερνά τα όρια της Βιβλιομετρίας, γιατί κάποιες δραστηριότητες στον κυβερνοχώρο δεν καταγράφονται, αλλά μεταδίδονται συγχρονισμένα, όπως στα chatrooms. Οι Cybermetric μελέτες τέτοιων δραστηριοτήτων ακόμη εντάσσονται στο γενικό πεδίο Infometrics ως η μελέτη ποσοτικών απόψεων της πληροφορίας σε «κάθε μορφή» και σε «κάθε κοινωνική ομάδα», όπως έχει οριστεί παραπάνω από τους Tague και Sutcliffe (1992).

Βιβλιομετρία (Bibliometrics): Υπάρχουν πολλοί ορισμοί του όρου «Βιβλιομετρία» στη βιβλιογραφία έμμεσοι ή άμεσοι, αλλά μόνο λίγοι θα αναφερθούν. Ένας πρώιμος ορισμός δόθηκε από τον Pritchard (1969): "για να ρίξουμε φως στις διαδικασίες της γραπτής επικοινωνίας και στη φύση και την πορεία της ανάπτυξης μιας επιστήμης, μέσω της μέτρησης και ανάλυσης διαφόρων πτυχών της γραπτής επικοινωνίας, είναι χρήσιμη η εφαρμογή των μαθηματικών και στατιστικών μεθόδων σε βιβλία και άλλα μέσα επικοινωνίας."

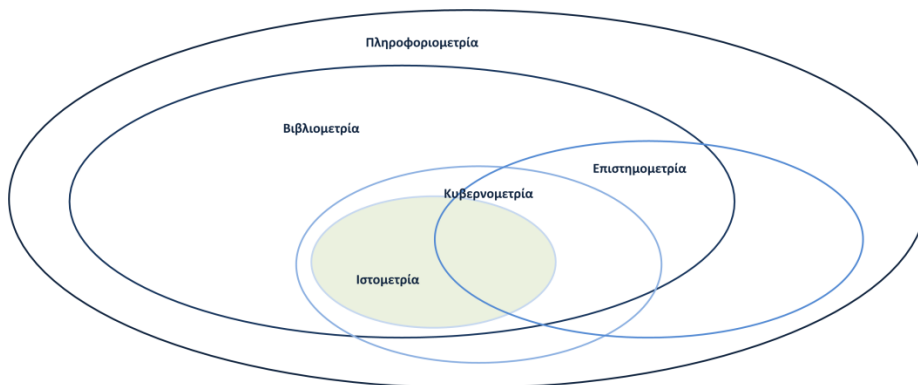
Broadus (1987): «... η ποσοτική μελέτη των δημοσιευμένων μονάδων ή βιβλιογραφικών μονάδων ή των υποκατάστατων του καθενός...»

Brookes (1990): «Δεν έχω καμία αμφιβολία ότι η Βιβλιομετρία, πλέον, πρέπει να γίνεται δεκτή μόνο σε μελέτες πάνω στη βιβλιοθήκη. Το έργο της δεν έχει ολοκληρωθεί, εφόσον οι βιβλιοθήκες συνεχίζουν να προσαρμόζονται στον μεταβαλλόμενο αυτό κόσμο. Και η ίδια η Βιβλιομετρία χρειάζεται το ενδιαφέρον όλων και όχι μόνο των εμπειρογνομόνων και στατιστικολόγων, για να τελειοποιηθεί και να αναπτύξει τις τεχνικές της.»

White και McCain (1989): «Βιβλιομετρία είναι η ποσοτική μελέτη των βιβλιογραφιών, όπως αυτές αντικατοπτρίζονται στις βιβλιογραφίες. Το έργο της, είναι να παρέχει εξελικτικά μοντέλα επιστήμης και τεχνολογίας.»

Η Βιβλιομετρία, όπως προαναφέρθηκε, είναι η εφαρμογή των μαθηματικών και στατιστικών μεθόδων για τις επιστημονικές δημοσιεύσεις (από το Biblos: βιβλίο και metron: μέτρηση). Η Βιβλιομετρία χρησιμοποιείται συχνά για την «αξιολόγηση» της παραγόμενης επιστημονικής έρευνας, μέσω ποσοτικών μελετών των δημοσιεύσεων.

Η Βιβλιομετρία ορίζεται ως «η μελέτη ποσοτικών απόψεων της παραγωγής, διάδοσης και χρήσης καταγεγραμμένων πληροφοριών». Οι μετρήσεις των τεκμηρίων αντιστοιχίζόμενοι με αριθμούς εκφράζονται μέσω των βιβλιομετρικών δεικτών. Οι βιβλιομετρικοί δείκτες είναι μια υποδιαίρεση των επιστημονομετρικών δεικτών, στο βαθμό που σχετίζονται με μια επιστημονική περιοχή και κατασκευάζονται από μετρήσιμες ποσότητες, όπως ο αριθμός των επιστημονικών δημοσιεύσεων, ο οποίος αποτελεί τον απλούστερο δείκτη για την καταγραφή της παραγωγής επιστημονικών εργασιών και, κατά συνέπεια, του ερευνητικού έργου ανά επιστήμονα, οργανισμό, επιστημονικό κλάδο ή χώρα.



Εικόνα 1 1 Απεικόνιση των Πληροφο-Επιστημο-Κυβερνο-Ιστο-μετριών που προκύπτει από τους αντίστοιχους ορισμούς. (Björneborn & Ingwersen, 2004)

Εκτός από τον αριθμό των δημοσιεύσεων, ο συνηθέστερος βιβλιομετρικός δείκτης που χρησιμοποιείται προκειμένου να εκτιμηθεί η απήχηση και η πρωτοτυπία του επιστημονικού έργου, βασίζεται στην ανάλυση των αναφορών σε δημοσιεύσεις από άλλες επιστημονικές δημοσιεύσεις.

1.4. Διάκριση Βιβλιομετρίας

Η βιβλιομετρία προσφέρει τρεις σπουδαίες δυνατότητες: την περιγραφή και την ανάλυση των χαρακτηριστικών της βιβλιογραφίας, την «αξιολόγηση» μέσω των αναφορών των δημοσιεύσεων και την εξέταση των σχέσεων των εμπλεκόμενων μεταβλητών. Ως συνέπεια αυτών των δυνατοτήτων της επιτυγχάνεται και η βασική διάκρισή της σε:

Α) Περιγραφική Βιβλιομετρία: Η περιγραφική Βιβλιομετρία (descriptive Bibliometrics) περιγράφει τα χαρακτηριστικά ή τις δυνατότητες της βιβλιογραφίας και χρησιμοποιείται για τη μέτρηση της παραγωγικότητας των επιστημόνων. Η έρευνα χωρίζεται σε γεωγραφικές περιοχές, σε χρονικές περιόδους, σε πανεπιστημιακά τμήματα και επιστημονικούς κλάδους. Η περιοχή της περιγραφικής Βιβλιομετρίας περιλαμβάνει τη μελέτη του αριθμού των δημοσιεύσεων σε ένα δεδομένο πεδίο ή την παραγωγικότητα της βιβλιογραφίας στο πεδίο, με σκοπό τη σύγκριση της έρευνας σε διάφορα ιδρύματα / χώρες, καθώς και σε διαφορετικές περιόδους.

Β) Αξιολογική Βιβλιομετρία: Η αξιολογική Βιβλιομετρία (evaluative Bibliometrics) χρησιμοποιεί τις αναφορές ως πηγή των δεδομένων της. Οι βιβλιομετρικές αξιολογήσεις βασίζονται στην υπόθεση ότι οι περισσότερες επιστημονικές ανακαλύψεις και τα αποτελέσματα της έρευνας τελικά δημοσιεύονται σε διεθνή επιστημονικά περιοδικά, όπου μπορούν να διαβασθούν και να αναφερθούν από άλλους ερευνητές. Ο αριθμός των αναφορών σε ένα άρθρο περιοδικού μπορεί να θεωρηθεί ότι αντανακλά τον αντίκτυπο του άρθρου στην επιστημονική κοινότητα. Οι περισσότερες αξιολογικές βιβλιομετρικές τεχνικές χρησιμοποιούν αναφορές ως πρωτογενή δεδομένα τους. Η θεωρητική βάση γι' αυτό προέρχεται από την κοινωνιολογία της επιστήμης του Robert Merton, η οποία θέτει ως δεδομένο ότι οι αναφορές είναι ο τρόπος με τον οποίο οι επιστήμονες παραδέχονται προηγούμενη σημαντική δουλειά. Σ' αυτή τη βάση, η καταμέτρηση των αναφορών θα μπορούσε να χρησιμοποιηθεί ως δείκτης επιστημονικής αξίας, γιατί όσο πιο σημαντική είναι μία δουλειά, τόσο πιο συχνά θα τείνει να αναφέρεται. Πραγματικά, ο όρος «επιρροή» είναι τώρα αποδεκτός ως κατάλληλος γι' αυτό που οι αναφορές καταμετρούν ή υποδεικνύουν. Μεταγενέστερη έρευνα έχει αποδείξει ότι η άποψη του Merton είναι μια απλοποίηση της πραγματικότητας, αφού υπάρχουν πολλοί διαφορετικοί λόγοι για να αναφέρει κανείς άρθρα, καθώς και πολλές επιδράσεις στο ποιά άρθρα να επιλέξεις, όταν είναι διαθέσιμες πολλαπλές επιλογές.

Σύμφωνα με τον Merton, τα περιοδικά με υψηλή επιρροή τείνουν να δημοσιεύουν έρευνες με μεγαλύτερη επιρροή και ως εκ τούτου, τείνουν να λαμβάνονται περισσότερο υπόψη. Παρ' όλα αυτά, φαίνεται να είναι κοινώς αποδεκτό ότι, ακόμα και μέσα σε διακριτά θεματικά πεδία, το να κατατάσσει κανείς περιοδικά βασιζόμενος πάνω στην επιρροή τους είναι προβληματικό. Επιπλέον, καθώς η επιρροή έχει κερδίσει σε αξία, φαίνεται να έχουν γίνει προσπάθειες από τους εκδότες των περιοδικών να συστήνουν στους συγγραφείς να αναφέρει κανείς άρθρα στο ίδιο περιοδικό, για να βελτιώσουν την επιρροή τους.

Μία άλλη εφαρμογή, η οποία συνήθως διεξάγεται από ειδικούς βιβλιομέτρους, συγκρίνει ακαδημαϊκά τμήματα μέσω των αναφορών στις δημοσιεύσεις τους. Ακόμη και προσεκτικά κατασκευασμένοι

βιβλιομετρικοί δείκτες, οι οποίοι είναι λογικά αξιόπιστοι εξαιτίας του συνόλου των εκδόσεων ολόκληρων τμημάτων, πρέπει να συνδυάζονται με άλλες αποδεικτικές πηγές (π.χ.τη χρηματοδότηση, τις πηγές εκτίμησης, την αξιολόγηση από ομότιμους), προκειμένου να δώσουν αδιάσειστα στοιχεία για σημαντικές αποφάσεις, όπως αυτές που αφορούν τη χρηματοδότηση.

Γ) Σχεσιακή Βιβλιομετρία: Η Σχεσιακή Βιβλιομετρία (relational Bibliometrics) μία ιδέα που προήλθε από τον γενετιστή Allen το 1960, ο οποίος έστειλε το διάγραμμα αναφορών του στον Garfield, χρησιμοποιείται για να εξετάσει τις σχέσεις στο πλαίσιο της επιστημονικής έρευνας, κάνοντας χρήση των δεδομένων του ISI. Αυτό δεν ήταν δυνατό τον πρώτο καιρό, λόγω της έλλειψης της υπολογιστικής ισχύος και της εμπειρίας στον τομέα της τεχνολογίας. Ακόμα κι έτσι, οι αρχικές σχεσιακές αναλύσεις παρήγαγαν ενδιαφέρουσες γνώσεις για τη δομή της έρευνας της επιστήμης μέσα από απλά μέσα, όπως τα διαγράμματα δικτύου της ροής των αναφορών μεταξύ των κυρίων συνόλων των περιοδικών. Τα διαγράμματα αναφορών περιοδικού ήταν μία άλλη πρώιμη εφεύρεση, καθώς αυτά μπορούν να διευκρινίσουν τις σχέσεις μεταξύ των περιοδικών μέσα σ' έναν τομέα, ν' ανιχνεύσουν περιοδικά, τα οποία ξεπερνούν τα σύνορα του τομέα και να εντοπίσουν κεντρικά ή περιφερειακά περιοδικά. Μία σημαντική σχεσιακή μέθοδος, που αποδίδεται στο Garfield, είναι η συν-αναφορά ως μέτρο ομοιότητας. Η βάση αυτού είναι ότι τα ζεύγη των τεκμηρίων, τα οποία συχνά εμφανίζονται μαζί σε καταλόγους αναφορών (δηλαδή, συν-αναφέρονται) είναι πιθανό να είναι κατά κάποιο τρόπο παρόμοια. Αυτό σημαίνει ότι, εάν οι συλλογές τεκμηρίων τακτοποιούνται σύμφωνα με τη συν-αναφορά τους, τότε αυτό θα πρέπει να παράγει ένα πρότυπο που ν' αντανάκλα τις γνωστικές επιστημονικές σχέσεις. Η Ανάλυση των συν –αναφορών ενός συγγραφέα είναι μια παρόμοια τεχνική στο ότι μετρά την ομοιότητα των ζευγών των συγγραφέων μέσα από τη συχνότητα με την οποία το έργο τους συν – αναφέρεται. Αυτή λειτουργεί σ' ένα αρκετά υψηλό επίπεδο συγκέντρωσης, ώστε να είναι ένα πρακτικό εργαλείο για τη χαρτογράφηση των δομών των τομέων της επιστήμης.

1.5. Περιορισμοί της Βιβλιομετρίας

Ειδικά προβλήματα επηρεάζουν την εγκυρότητα των βιβλιομετρικών αξιολογήσεων. Σύμφωνα με τον γαλλικό οργανισμό, Εθνικό Συμβούλιο Έρευνας και Αξιολόγησης 'Conseil National d'évaluation de la Recherche' (CNER 2002), τα κύρια προβλήματα με τις βάσεις δεδομένων στη δημοσίευση της έρευνας είναι τα ακόλουθα:

- περιορισμένη κάλυψη,
- αποκλεισμός ορισμένων τύπων άρθρων,
- ταξινόμηση των περιοδικών βάσει των επιστημονικών πεδίων,
- αλλαγές στους τίτλους των περιοδικών,
- τα ονόματα δεν γράφονται με τον ίδιο τρόπο,
- ο αριθμός των συγγραφέων (και η κατανομή της εργασίας),
- υπερβολικές, επιλεκτικές, αρνητικές και λανθασμένες αναφορές,
- αυτό -αναφορές και προσωπικές στρατηγικές,
- πολλαπλοί συγγραφείς που μπορούν να μετακινηθούν ή να χαθούν παρόμοια ονόματα συγγραφέων, χρησιμοποιώντας τα αρχικά που μπορεί να αναμειγνύονται με τα ολόκληρα ονόματα,
- ανθρώπινα σφάλματα όπως ορθογραφικά λάθη.

Από τότε που η Βιβλιομετρία γενικά βασίζεται στη χρήση των βιβλιογραφικών βάσεων δεδομένων, όλοι οι παραπάνω παράγοντες μειώνουν την αξιοπιστία και την αξιολόγηση των βιβλιομετρικών ερευνών. Τα προβλήματα είναι κοινά σε όλες τις βάσεις δεδομένων και αντανάκλουν: (1) τους περιορισμούς της κατάρτισης και ευρετηρίασης σε ακαδημαϊκά περιοδικά και άρθρα και (2) τις πρακτικές και ειδικές καταστάσεις του συγκεκριμένου επιστημονικού πεδίου.

Προχωρώντας προς ακόμη μια πιο αυτόματη ανάλυση, γίνονται αντιληπτοί οι τύποι αυτών των σφαλμάτων. Έτσι μπορούμε να αναπτύξουμε το λογισμικό, ώστε να τους εντοπίσουμε μέσω μετρήσεων, όπως αυτές των ευρετηρίων, των κοινών ονομάτων και των ορθογραφικών λαθών ή χτίζοντας τομείς για συγκεκριμένα μεταδεδομένα. Η ανάλυση δεδομένων ευρείας κλίμακας αυτών των απλών μεθόδων μπορεί να

αποκαλύψει τη σχέση μεταξύ όλων των τύπων πληροφοριών. Μπορούμε να ξεκινήσουμε να αναπτύσσουμε προγράμματα και να θέτουμε ενδιαφέρουσες ερωτήσεις όπως:

Οι εργασίες που αναφέρονται στις σημειώσεις και υποσημειώσεις σχετίζονται περισσότερο από ό,τι άλλες αναφορές μέσα σε μία εργασία;

Μήπως η σειρά των αναφορών στο πλαίσιο υπονοεί σχέσεις;

Μήπως ο τύπος της μορφής δεδομένων (διάταξη βιβλιογραφίας ή ο τύπος δημοσίευσης) σημαίνει κάτι;

Έχουν οι αναφορές μεταβατικές σχέσεις (το Α σχετίζεται με το Β, το Β σχετίζεται με το Γ, το Α συγγενεύει με το Γ);

Η ανάλυση του λογισμικού είναι ιδανική για να απαντάει σε τέτοιου τύπου συγκριτικές ερωτήσεις. Γενικά όλες οι τεχνικές βασικής Βιβλιομετρίας λειτουργούν καλά με πολλούς τύπους πληροφοριακών οντοτήτων, όπως συγγραφείς, περιοδικά, οργανισμοί, τμήματα, πανεπιστήμια.

1.5.1. Βιβλιομετρία και επιστημονικοί κλάδοι

Σαφώς, είναι δύσκολο να εκτελέσεις βιβλιομετρική ανάλυση με ένα ενιαίο τρόπο για όλες τις περιπτώσεις, λόγω του ευρέως φάσματος των επιστημονικών κλάδων. Κάθε κλάδος έχει τα δικά του χαρακτηριστικά και τις πρακτικές του, που ασκούν διαφορετικές επιπτώσεις στις συνθήκες για την εφαρμογή της Βιβλιομετρίας. Ως εκ τούτου, ο αναγνώστης θα πρέπει πάντα να έχει κατά νου ότι τα γενικά ευρήματα που παρουσιάζονται εδώ δεν μπορούν να επιτευχθούν σε ορισμένα επιστημονικά πεδία ή τομείς εξειδίκευσης. Θα εξετάσουμε προς αυτήν την κατεύθυνση τις διαφορές δύο τομέων που μας ενδιαφέρουν άμεσα, τις Κοινωνικές και Ανθρωπιστικές Επιστήμες 'Social Sciences and Humanities' (SSH) και τις Φυσικές και Τεχνολογικές Επιστήμες 'Natural Science and Engineering' (NSE).

Οι Social Sciences and Humanities (SSH) επιστήμες διαφέρουν από τις Natural Science and Engineering (NSE), όχι μόνο από την άποψη των υποκειμένων της έρευνας τους, αλλά και λόγω των επιστημονικών πρακτικών και μεθόδων επικοινωνίας τους. Κάθε διαφορά επηρεάζει τις συνθήκες για την εφαρμογή της Βιβλιομετρίας.

Πρώτον, τα μέσα διάδοσης της γνώσης και, κατ' επέκταση, των μέσων επικοινωνίας σε γενικές γραμμές είναι πιο ποικίλα στις SSH από ότι στις NSE. Αυτό αντικατοπτρίζεται στην ενίσχυση του ρόλου που διαδραματίζουν οι μονογραφίες, οι ανακοινώσεις σε επιστημονικά συνέδρια και τα πρακτικά αυτών, καθώς και η μη-επιστημονική βιβλιογραφία στις SSH. Ανάλογα με τον τομέα, τα άρθρα μπορεί να είναι ένα σχετικά μικρό μέσο δημοσίευσης σε σύγκριση με άλλα μέσα, όπως τα βιβλία. Δυστυχώς, δεν υπάρχει βάση δεδομένων που να καλύπτει αυτές τις άλλες μορφές δημοσίευσης, κάτι που συστηματικά και εξαντλητικά η Thomson ISI κάνει για τα άρθρα των περιοδικών.

Ειδικά τα ερευνητικά αντικείμενα των βάσεων στις SSH είναι μερικές φορές περισσότερο τοπικά προσανατολισμένα και, ως εκ τούτου, το αναγνωστικό κοινό περιορίζεται τις περισσότερες φορές σε μια χώρα ή περιοχή (Nederhof et al., 1989; Nederhof and Zwaan 1991; Webster 1998; Winclawska 1996). Μελετητές των βάσεων στις SSH φέρεται να δημοσιεύουν, τις περισσότερες φορές, στη μητρική τους γλώσσα και σε περιοδικά με μία πιο περιορισμένη διανομή (Gingras, 1984).

Επιπλέον, σύμφωνα με τον Hicks (1999), μια σειρά από ερευνητικούς τομείς των SSH έχουν περισσότερα παραδείγματα που ανταγωνίζονται μεταξύ τους σε σχέση με εκείνα των NSE και ως αποτέλεσμα η ερευνητική παραγωγή στις SSH είναι πιο κατακερματισμένη - μια κατάσταση που εμποδίζει το σχηματισμό ενός στερεού "πυρήνα" των επιστημονικών περιοδικών (Hicks 1999)-, καθιστώντας έτσι το άρθρο που είναι η βάση των βιβλιομετρικών αναλύσεων δύσκολο να διεξαχθεί με επιτυχία. Οι κύριες διαφορές μεταξύ των SSH και NSE, όσον αφορά στη χρήση των βιβλιογραφικών μεθόδων, αναλύονται με μεγαλύτερη λεπτομέρεια παρακάτω.

1.5.2. Μέσα επιστημονικής επικοινωνίας

Οι επιστημονικές ανακοινώσεις που αρχειοθετούνται ως βιβλιογραφικές εγγραφές βρίσκονται στον πυρήνα της Βιβλιομετρίας. Για παράδειγμα, οι επιστημονικές τεχνικές μπορούν να χρησιμοποιηθούν για την ανάλυση λεκτικής επικοινωνίας, εφόσον έχουν ευρετηριαστεί συστηματικά. Κατά συνέπεια, ενώ οι μέθοδοι δεν στοχεύουν ειδικότερα σε καμία μορφή επικοινωνίας, η Βιβλιομετρία χρησιμοποιείται κυρίως για

ακαδημαϊκή επίδοση και ανάλυση σε άρθρα. Η Βιβλιομετρία άρχισε πραγματικά να απογειώνεται με την έλευση των Science Citations Indexes (SCI) στη δεκαετία του 1960. Στην πραγματικότητα, είναι πλέον δύσκολο να διαχωριστεί η Βιβλιομετρία από τη μέτρηση άρθρου περιοδικών, αλλά αυτό πρέπει να γίνει, διότι τα επιστημονικά περιοδικά δεν είναι τόσο εκτεταμένα σε ορισμένους κλάδους των SSH, όπως είναι στις NSE.

Για ποιούς λόγους υπάρχει το ευρύ φάσμα μέσων διάδοσης γνώσης των SSH σε σύγκριση με τις NSE; Σύμφωνα με τον Line (1999), οι περισσότερες κοινωνικές επιστήμες μπορεί να περιγραφούν ως «σχετικά νέες και, μόλις και μετά βίας, οργανωμένες ως συνεκτικοί κλάδοι.» Υποστηρίζει ότι οι SSH είναι κατακερματισμένες, επειδή δεν έχουν τα διεθνή πρότυπα για τον επιμελή ορισμό των εννοιών. Οι όροι που χρησιμοποιούνται διαφέρουν σημαντικά μεταξύ των περιφερειών με την πάροδο του χρόνου. Η δημοσίευση άρθρων στα επιστημονικά περιοδικά θεωρείται ως μια ένδειξη της συναίνεσης και ως μια πρακτική που φέρνει μελετητές σε επιστημονικούς κλάδους. Κάποιοι ερευνητές ισχυρίζονται ότι οι μελετητές των SSH ανήκουν σε ένα μεγάλο αριθμό ανταγωνιστικών προσεγγίσεων, η οποία δημιουργεί μια ώθηση να δημοσιευθεί σε μορφή βιβλίου (Hicks, 1999). Από τότε που τα βιβλία δεν υπόκεινται σε αξιολόγηση από ομότιμους στον ίδιο βαθμό όπως τα άρθρα σε περιοδικά, οι μελετητές των SSH θεωρούν τις εκδόσεις βιβλίων δυνητικά ως ένα μέσο ομαδοποίησης, που μπορούν να λάβουν μια θέση έξω από τα επικρατούντα μοντέλα. Σύμφωνα με τον Hicks (1999), η καλύτερη έρευνα των NSE δημοσιεύεται στα άρθρα, ενώ η καλύτερη έρευνα των SSH δημοσιεύεται σε βιβλία.

Ο Hicks στηρίζει αυτά τα συμπεράσματα σε ένα άρθρο των Clemens κ. α. (1995), που συνδύασε τα επιχειρήματα του Merton και Kuhn για να δημιουργήσει μια σύνδεση μεταξύ της επιστημολογικής δυναμικής και της δημοσίευσης των βιβλίων και άρθρων. Ο Clemens και οι συναρθρογράφοι προώθησαν την υπόθεση ότι η εξάπλωση των άρθρων των φυσικών επιστημών αντανάκλα ένα εκδοτικό σύστημα που λειτουργεί ως κοινωνικός φορέας ελέγχου και έτσι δημιουργεί πρωτότυπη, αξιόπιστη, συσσωρευμένη γνώση. Με τη σειρά του, το σύστημα αυτό αντικατοπτρίζει την παραδειγματική κατάσταση των επιμέρους κλάδων, καθώς και μια συναίνεση για τα θέματα και τις μεθόδους που προσδιάζουν σ' αυτήν. Έτσι ο κλάδος, στον οποίο το βιβλίο αποτελεί σημαντικό μέσο δημοσίευσης, είναι πιθανώς ακόμα σε παραδειγματική φάση.

Ο Clemens και οι συναρθρογράφοι αποδίδουν διάφορους ρόλους σε επιστημονικά βιβλία και άρθρα, τουλάχιστον σε σχέση με την κοινωνιολογία της επιστήμης. Τα βιβλία έχουν μεγαλύτερο αντίκτυπο εκτός των κλάδων και διαβάζονται γενικά από τους ανθρώπους. Τα άρθρα καλύπτουν πιο πρόσφατα προβλήματα και είναι ένας καλύτερος τρόπος για την επίτευξη διαρκούς κύρους εντός του κλάδου. Η προτίμηση ενός δεδομένου μέσου δημοσίευσης από ένα άλλο μπορεί να εξαρτάται από το φύλο του ακαδημαϊκού, τη θέση, το ίδρυμα και ακαδημαϊκό του υπόβαθρο.

Σημειώστε ότι ο κίνδυνος να έχεις κλέψει την ιδέα ή την εφεύρεση κάποιου είναι πολύ χαμηλότερη στις SSH, αφού οι μελετητές έχουν περισσότερο χρόνο για να αναπτύξουν και να συμπληρώσουν τα επιχειρήματά τους πριν από τη δημοσίευσή τους (Hicks, 1999).

Τα άρθρα και τα βιβλία είναι οι δύο πιο συχνά χρησιμοποιούμενες επιστημονικές και επικοινωνιακές δημοσίευσης των μέσων επικοινωνίας. Βεβαίως, και πολλά άλλα εκδοτικά μέσα ενημέρωσης διαδραματίζουν σημαντικό ρόλο σε ορισμένες ειδικότητες, όπως π.χ. επίσημες (κυβερνητικές) δημοσιεύσεις και γκρίζα βιβλιογραφία - εσωτερικές εκθέσεις οργανισμών που είναι συνήθως πιο ανεπίσημα και λιγότερο συστηματικά διανεμημένα από τα εμπορικά διαθέσιμα περιοδικά-. Η αξία τους για την παραγωγή της γνώσης αποκτά όλο και μεγαλύτερη δύναμη (Grayson & Gomersall 2003; Lariviere & Godin 2001). Ορισμένες από αυτές τις εκδόσεις παίρνουν ISBN αριθμούς. Επιπλέον αυτές οι δημοσιεύσεις απολαμβάνουν όλο και πιο ευρεία διάδοση, κυρίως λόγω του διαδικτύου. Τα έγγραφα εργασίας έχουν αναρτηθεί στις ιστοσελίδες ορισμένων ερευνητικών κέντρων και διαδραματίζουν ολοένα και σημαντικότερο ρόλο, επειδή είναι μερικές φορές εξουθενωτικά και παρέχουν εξαντλητική πληροφορία, περισσότερη από τις τελικές εκδόσεις που δημοσιεύονται ως άρθρα. Εν ολίγοις, το σύνθημα εδώ πρέπει να είναι προσοχή, επειδή αυτό το είδος παραγωγής γνώσης θα πρέπει να θεωρείται, τουλάχιστον εν μέρει, μέσο διάδοσης των SSH επιστημών, έξω από την ακαδημαϊκή κοινότητα, αντί για επιστημονικό μέσο επικοινωνίας.

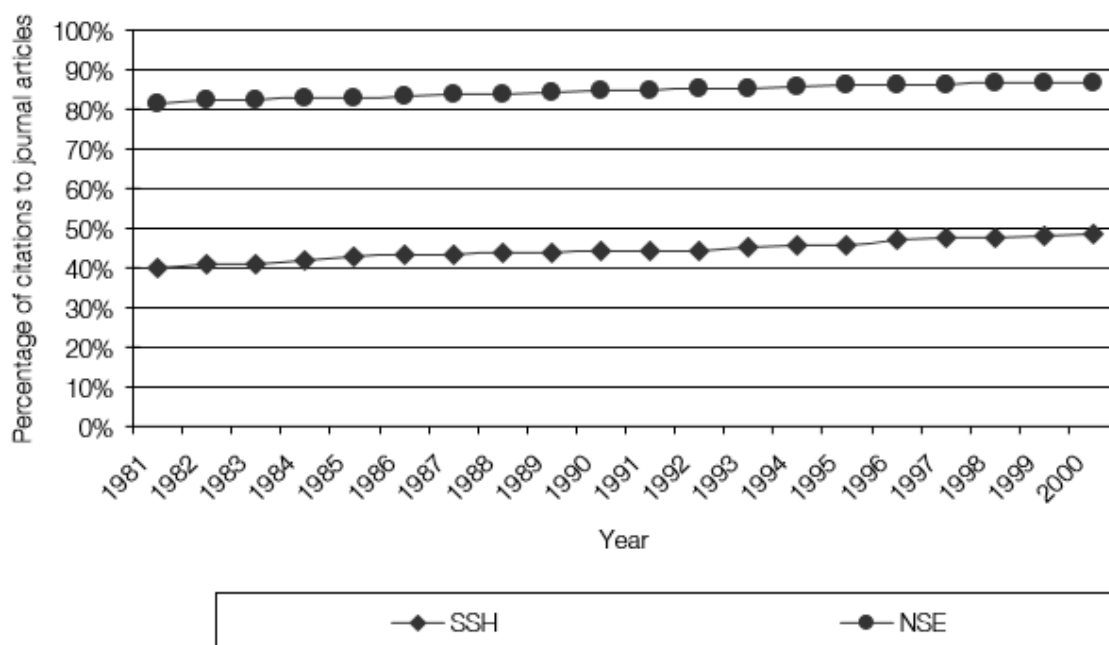
Βιβλία, άρθρα και σε μικρότερο βαθμό, γκρίζα βιβλιογραφία και κυβερνητικές εκθέσεις θεωρούνται όλα ότι είναι επιστημονικά μέσα επικοινωνίας. Όλες οι δημοσιεύσεις είναι μια απάντηση στους παράγοντες (κίνητρα, συμπεριφορές, κριτήρια) των επιστημονικών κλάδων που αναφέρονται. Ωστόσο, οι μελετητές δεν γράφουν μόνο για τους συναδέλφους τους. Αυτοί επίσης δημοσιεύουν άρθρα και βιβλία για το σκοπό της διάδοσης της επιστημονικής γνώσης μέσα στην κοινωνία και στο σύνολό της. Ενώ οι περισσότερες βιβλιομετρικές αναλύσεις εστιάζονται αποκλειστικά στην επιστημονική παραγωγή, μπορεί να αξίζει τον κόπο η μέτρηση των δραστηριοτήτων που αποσκοπούν στη διάδοση της γνώσης στο ευρύ κοινό.

Πολλοί συγγραφείς υποστηρίζουν ότι τα μέσα δημοσίευσης, εκτός από τα άρθρα και τα περιοδικά, διαδραματίζουν σημαντικό ρόλο στις SSH επιστήμες, αλλά το γεγονός είναι ότι δεν υπάρχουν ακριβείς πληροφορίες σχετικά με την εξάπλωση των διάφορων μέσων δημοσίευσης. Ο Nederhof και οι συναρθρογράφοι (1989) μελέτησαν την κατανομή έξι επιστημονικών μέσων επικοινωνίας όπως άρθρα, βιβλία, κεφάλαια βιβλίων, μονογραφίες κ.λπ. σε οκτώ κλάδους των SSH, στην Ολλανδία. Σε όλες τις περιπτώσεις, το άρθρο ήταν το κύριο μέσο, αντιπροσωπεύοντας το 35-57% των δημοσιεύσεων. Ακολουθούν τα κεφάλαια των βιβλίων, με 21-34%. Ακόμα κι αν αυτά τα ευρήματα μπορούν να αντικρούσουν σε ορισμένες περιπτώσεις τις καθιερωμένες απόψεις για την υπεροχή των μονογραφιών στις SSH, δεν θα πρέπει να οδηγήσουν στο συμπέρασμα ότι οι αξιολογήσεις που βασίζονται μόνο σε άρθρα είναι αρκετές. Τα στατιστικά στοιχεία που παρατίθενται από τους Nederhof και τους συναρθρογράφους δείχνουν ότι το 35% των δημοσιεύσεων σε γενική γλωσσολογία είναι τα άρθρα, το 34% είναι τα κεφάλαια βιβλίων, ενώ οι μονογραφίες ακόμη διαδραματίζουν σημαντικό ρόλο στη διάδοση της γνώσης.

Τα ευρήματά τους δεν μπορούν, ωστόσο, να εφαρμοστούν στις κοινωνικές και ανθρωπιστικές επιστήμες σε όλες τις χώρες. Για παράδειγμα, ο Nederhof και οι συναρθρογράφοι σημειώνουν ότι το ποσοστό των άρθρων μπορεί να είναι εξαιρετικά υψηλό λόγω των ειδικών χαρακτηριστικών του επιστημονικού πεδίου στην Ολλανδία, όπου υπάρχει ένας ασυνήθιστα υψηλός αριθμός τοπικής δημοσίευσης στα επιστημονικά περιοδικά. Ως εκ τούτου, είναι σημαντικό να εξετάσουμε τα στατιστικά στοιχεία από άλλες χώρες. Στη Δανία, ο Andersen (2000) διεξήγαγε μια έρευνα που υποδηλώνει ότι μόνο το ένα τέταρτο των δημοσιεύσεων από τους Δανούς μελετητές στις κοινωνικές επιστήμες ήταν άρθρα περιοδικών. Ο Hicks (1999) εκτιμά ότι τα βιβλία αποτελούν το 40-60% της επιστημονικής παραγωγής στις κοινωνικές επιστήμες και ότι έχουν μια πολύ υψηλή επίδραση στην έρευνα, διότι αντιπροσωπεύουν το 40% των αναφορών.

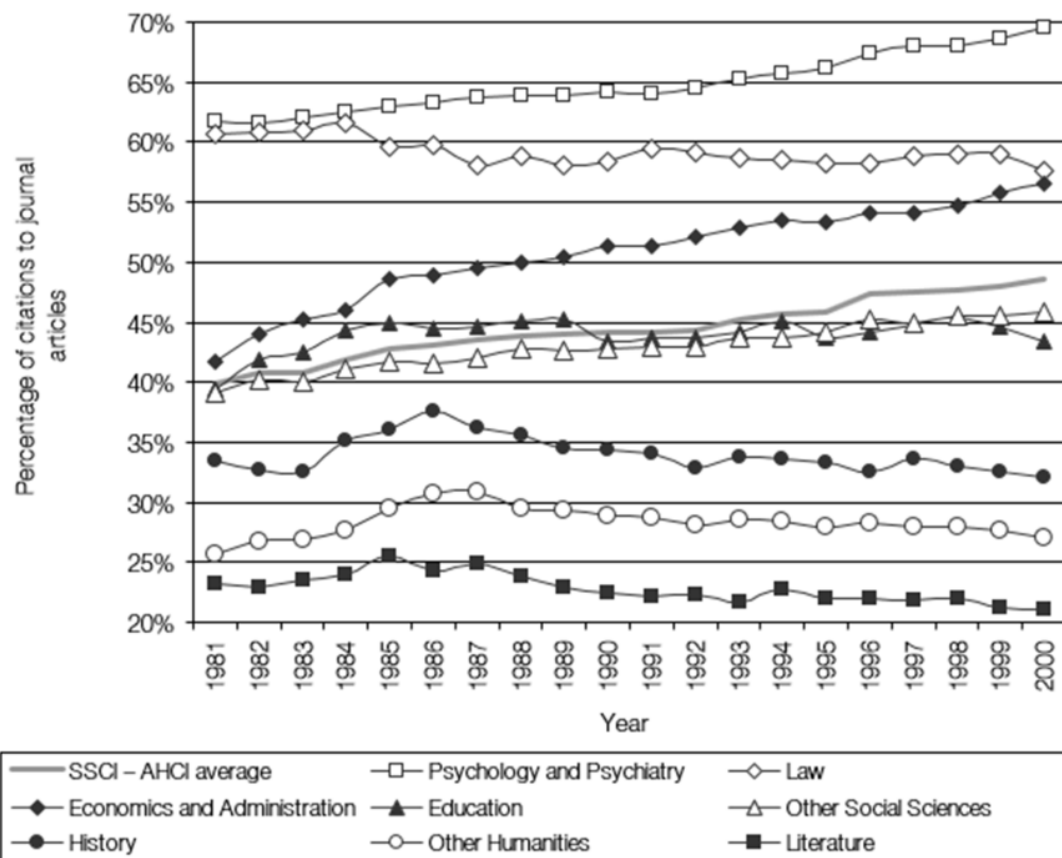
Σημειώστε πως η σημασία που οι ερευνητές δίνουν στη μαζική διανομή διαφέρει ευρέως από κλάδο σε κλάδο. Ο Nederhof και οι συναρθρογράφοι (1989) εξέτασαν το ποσοστό των δημοσιεύσεων στις SSH για την περίοδο 1980-1985. Στην Ολλανδία, στην πειραματική ψυχολογία είναι αφιερωμένο το 3-12 % του εκδοτικού έργου, ενώ στην ολλανδική λογοτεχνία το ποσοστό φτάνει το 30-43 %.

Για να ρίξει φως στη σημασία των άρθρων σε περιοδικά, το Observatoire des Sciences et des Technologies (OST) παρουσίασε μια σειρά από στατιστικά στοιχεία που υπολογίζουν τη σημασία των μέσων επιστημονικής επικοινωνίας στις SSH, εκτός από τα άρθρα. Το σχήμα 1.2 παρέχει στατιστικά στοιχεία σχετικά με τις αναφορές των άρθρων σε περιοδικά που καλύπτουν οι βάσεις δεδομένων SCI, SSCI και AHCI. Δείχνει ότι, ενώ το 85 % των αναφορών στις φυσικές επιστήμες είναι σε άρθρα περιοδικών, λιγότερο από το 50 % των αναφορών στις SSH είναι σε αυτό το μέσο δημοσίευσης (άρθρα περιοδικών).



Εικόνα 1 2 Αναφορές των άρθρων σε περιοδικά που καλύπτουν οι βάσεις δεδομένων SCI, SSCI και AHCI. (Larivière, Archambault, Gingras and Vignola-Gagné, 2004)

Βεβαίως, στο χώρο των SSH τα δεδομένα αλλάζουν γρήγορα, επειδή ο αριθμός των αναφορών σε άρθρα έχει αυξηθεί σχεδόν κατά 22% τα τελευταία 20 χρόνια, σε σύγκριση με το 6% στις φυσικές επιστήμες. Αυτό δεν σημαίνει, ωστόσο, ότι η τάση αυτή συναντάται παντού. Κατά τη διάρκεια των τελευταίων 10 ετών, για παράδειγμα, η Οικονομία και η Διοίκηση, η Ψυχολογία και άλλες κοινωνικές επιστήμες είναι τα πεδία των SSH, όπου ο ρόλος των περιοδικών στην παραγωγή γνώσης έχει αυξηθεί με ρυθμό παρόμοιο με εκείνο των SSH στο σύνολό του (Σχήμα 1.3). Από την άλλη πλευρά, ο ρόλος των περιοδικών στην προώθηση της επιστημονικής γνώσης μειώνεται στην Ιστορία, τη Λογοτεχνία και τις άλλες Ανθρωπιστικές επιστήμες. Με άλλα λόγια, οι επιστημονικοί κλάδοι των SSH δεν έχουν όλοι την ίδια πρόοδο, με τον ίδιο ρυθμό και με τον ίδιο τρόπο.



Εικόνα 1.3 Ποσοστό αναφορών άρθρων περιοδικών στις ανθρωπιστικές επιστήμες. (Larivière, Archambault, Gingras and Vignola-Gagné, 2004)

Παράδειγμα των διαφορών μεταξύ των τύπων της δημοσίευσης

Οι Nederhof και van Raan (1993) συνέκριναν την απόδοση έξι βρετανικών οικονομικών τμημάτων που είχαν λάβει χρηματοδότηση από το Συμβούλιο Οικονομικής και Κοινωνικής Έρευνας (ESRC). Μέτρησαν τις εκδόσεις και τον αριθμό των αναφορών που συνδέονται με αυτές, συμπεριλαμβανομένων των άρθρων, των βιβλίων, των κεφαλαίων βιβλίων και άλλων ειδών της δημοσίευσης. Τα αποτελέσματα στη συνέχεια συνδυάστηκαν με μια ανάλυση βασισμένη σε αξιολόγηση από ομότιμους. Η μελέτη καλύπτει την περίοδο 1980-1988.

Τα έξι τμήματα σε συνδυασμό παρήγαγαν 524 δημοσιεύσεις κατά την συγκεκριμένη περίοδο, συμπεριλαμβανομένων των 140 άρθρων που παρατίθενται από την Thomson ISI, 53 ειδών που δεν αναφέρονται από την Thomson ISI, 58 κεφαλαίων βιβλίων και 13 βιβλίων. Περισσότερα από 260 έγγραφα δεν ανήκουν σε κάποια από αυτές τις κατηγορίες: αυτά περιλαμβάνουν κυβερνητικές εκθέσεις και αδημοσίευτα έγγραφα. Τα 13 βιβλία είχαν κατά μέσο όρο 3,15 αναφορές, σε σύγκριση με το μέσο όρο του 0,95 για τα είδη που καλύπτονται από την Thomson ISI. Έτσι, η επίπτωση των βιβλίων φαίνεται να είναι ο παράγοντας που συμβάλλει περισσότερο στην ισχυρή επίδοση στη μελέτη. Τα αποτελέσματα δείχνουν ότι οι ισχυρές επιδόσεις από μια ομάδα συχνά σχετίζονται με την ισχυρή απόδοση ενός συγγραφέα, του οποίου το ποσοστό της αναφοράς είναι ιδιαίτερα υψηλό.

Το έργο τώρα του Line (1999) και των Cronin, Snyder και Atkins (1997) έχει σίγουρα τα μειονεκτήματά του. Δεν είναι συστηματικό και τα δείγματα είναι σχετικά μικρά. Ωστόσο, εάν τα ευρήματά τους σχετίζονται με εκείνα των Clemens και συναρθρογράφων, η υπόθεση ότι η μέτρηση που βασίζεται σε άρθρα δίνει μια επαρκή εικόνα της έρευνας σε σχέση με όσα διαδίδονται για τις μονογραφίες πρέπει να απορριφθεί. Κάθε αξιολόγηση που κάλυπτε μόνο το ένα από τα δύο μέσα θα μπορούσε πιθανότατα να είναι ελλιπής σε διαφορετικό βαθμό, ανάλογα με τις ειδικότητες. Γι' αυτό είναι σημαντικό να υπογραμμιστεί ότι τα αποτελέσματα των μετρήσεων της έρευνας που βασίζεται σε άρθρα δεν μπορούν να γενικευτούν στη συνολική παραγωγή δημοσιεύσεων.

Οι βιβλιομετρικοί δείκτες που είναι βασισμένοι μόνο σε άρθρα είναι οι λιγότερο αντιπροσωπευτικοί για την έρευνα στις SSH επιστήμες. Ο Line (1999) μένει έκπληκτος με το πόσο απόλυτα μερικές βιβλιομετρικές μελέτες μπορεί να παρουσιάσουν τα αποτελέσματα της αξιολόγησης των κοινωνικών επιστημών, με βάση μόνο την ανάλυση των άρθρων σε περιοδικά, με δεδομένη τη σπουδαιότητα της μονογραφίας σε αυτούς τους κλάδους. Στην ιδανική περίπτωση, οι βιβλιομετρικές έρευνες στις SSH πρέπει να περιλαμβάνουν δεδομένα σχετικά με άρθρα και βιβλία και μάλιστα σε άλλα μέσα επιστημονικής επικοινωνίας, ανάλογα με τον εν λόγω τομέα. Δυστυχώς, οι βάσεις δεδομένων Thomson ISI δεν παρέχουν αυτού του είδους τη κάλυψη και καμία άλλη βάση δεδομένων δεν την παρέχει επαρκώς.

Σε περιπτώσεις όπου τα δεδομένα συγκεντρώθηκαν από τα βιβλία και τα άρθρα συνεδρίων, από τεχνικές εκθέσεις και περιλήψεις, προκύπτει μια άλλη σειρά προβλημάτων. Για να είναι πλήρεις αυτές οι συλλογές θα πρέπει να περιλαμβάνουν αναφορές μεταξύ μονογραφιών, πρακτικά συνεδρίων και άρθρα. Ένα από τα πλεονεκτήματα μιας τέτοιας μεγάλης ανάλυσης θα ήταν μια μεγαλύτερη δεξαμενή αναφορών της ομάδας, αφού οι αναλύσεις περιορίζονται σε ένα μέσο δημοσίευσης και αποκλείουν αναφορές από τα άλλα. Τούτου λεχθέντος, η σύνταξη όλων αυτών των δεδομένων θα απαιτήσει σημαντικές προσπάθειες και σημαντικούς οικονομικούς πόρους. Επιπλέον ο Lewison (2001) αναφέρεται στα προβλήματα της ανάλυσης του βιβλίου στη Βιβλιομετρία, επισημαίνοντας ότι τα βιβλιογραφικά δεδομένα δεν περιλαμβάνουν συστηματικά τις διευθύνσεις όλων των συγγραφέων, ότι το περιεχόμενο του βιβλίου ποικίλλει σε μεγάλο βαθμό και ότι παρατηρείται έλλειψη κριτηρίων για τη διενέργεια συγκρίσεων μεταξύ των βιβλίων. Η Villagra Rubio (1992) έκανε εργασίες για μονογραφίες και βιβλία και εντόπισε παρόμοια προβλήματα.

Καταδεικνύεται λοιπόν ότι στις SSH δεν χρησιμοποιείται μόνο το άρθρο ως μέσο επικοινωνίας, άλλα χρησιμοποιούνται και άλλα είδη της δημοσίευσης. Ωστόσο, αρκετά σοβαρά προβλήματα κάλυψης προκύπτουν ακόμα και στην περίπτωση των ειδών, κυρίως λόγω των τοπικών προσανατολισμών των θεμάτων που εξετάστηκαν και, κατ' επέκταση, της γλώσσας στην οποία είναι γραμμένα τα άρθρα.

1.5.3. Γλώσσα και τοπικό ενδιαφέρον.

Συνοψίζοντας, τα προβλήματα που εντοπίστηκαν στις NSE επιστήμες τείνουν να είναι καθολικά, ενώ εκείνα που αφορούν στις SSH είναι μερικές φορές τοπικά. Η βιβλιογραφία δείχνει ότι το αναγνωστικό κοινό στις φυσικές επιστήμες είναι η διεθνής επιστημονική κοινότητα. Η έρευνα στη Φυσική ή Βιολογία μπορεί να επαναχρησιμοποιηθεί από τυχόν διενέργειες και από ενδιαφερόμενους εμπειρογνώμονες οπουδήποτε στον κόσμο, με τα ίδια αποτελέσματα. Οι ιδιότητες ενός ηλεκτρονίου είναι οι ίδιες παντού στη Γη. Συνέπεια αυτού είναι στις φυσικές επιστήμες η διάδοση των μέσων ενημέρωσης με διεθνή παρουσία, ενώ το προτιμώμενο μέσο είναι το επιστημονικό άρθρο, πολύ συχνά γραμμένο στην αγγλική γλώσσα και δημοσιευμένο σε περιοδικό με διεθνή απήχηση.

Κατά την άποψη των παρακάτω συγγραφέων, η έρευνα στις SSH ασχολείται με τοπικά και περιφερειακά προβλήματα πιο συχνά από ό,τι συμβαίνει στις NSE (Glänzel 1996; Hicks 1999; Hicks 2004; Ingwersen 1997; Nederhof et al., 1989; Nederhof & Zwaan 1991; Webster 1998; Winclawska 1996). Προκύπτει τελικά ότι τα ερευνητικά θέματα στις SSH θα είναι ειδικά για ένα συγκεκριμένο πολιτισμό, πιο συχνά από ότι θα ήταν τα προβλήματα που αντιμετωπίζονται στις NSE. Ο Hicks (1999 και 2004) δείχνει ότι οι θεωρητικές έννοιες στις κοινωνικές επιστήμες είναι λεπτές και δεν μπορούν να εκφραστούν με την παγκόσμια γλώσσα των μαθηματικών τόσο, όσο οι θεωρητικές έννοιες στις φυσικές επιστήμες. Σε πολλές περιπτώσεις, οι έννοιες και τα θέματα που καλύπτονται στις SSH μπορεί να εκφραστούν και να γίνουν κατανοητά μόνο με τη γλώσσα των πολιτισμών που τους διαμορφώνει. Ως εκ τούτου, οι μελετητές στις SSH επιστήμες δημοσιεύουν πιο συχνά στη δική τους γλώσσα και σε περιοδικά με εθνική κατανομή.

Παράδειγμα των διαφορών μεταξύ NSE και SSH σχετικά με τοπικές εκδόσεις.

Ο Πίνακας 1.4 παρουσιάζει στοιχεία σχετικά με τα αποτελέσματα της φινλανδικής έρευνας. Περιλαμβάνει στοιχεία για τον αριθμό των δημοσιευμάτων για τα φινλανδικά άρθρα σε έγκριτα διεθνή περιοδικά με διεθνή διανομή, καθώς επίσης στοιχεία για αυτά που γράφονται στα Φινλανδικά, ανεξάρτητα από το είδος της δημοσίευσης (άρθρο, βιβλίο, πρακτικά συνεδρίων). Τα δεδομένα υποστηρίζουν την άποψη ότι οι SSH έχουν τοπικό προσανατολισμό, δεδομένου ότι ο αριθμός των δημοσιεύσεων στην πρώτη κατηγορία είναι πολύ χαμηλότερος από ότι στη δεύτερη. Υποστηρίζει τη θέση ότι οι SSH γίνονται ολοένα και πιο διεθνοποιημένες, επειδή ο αριθμός των άρθρων που δημοσιεύονται σε περιοδικά διεθνώς καταναμεμένα έχει αυξηθεί σημαντικά από το 1994.

Field	Type of publication	1994	1998	2002
Natural Sciences and Engineering	Local orientation, written in Finnish	3,787	3,032	2,828
	International journal with peer review	6,419	6,702	7,857
	Ratio	0.6	0.5	0.4
Social Sciences and Humanities	Local orientation, written in Finnish	2,871	4,001	3,570
	International journal with peer review	685	984	1,265
		4.2	4.1	2.8

Source: Academy of Finland 2003

Εικόνα 1 4 Αριθμός τεκμηρίων στα Φινλανδικά

Εάν αυτή η επιστημολογική πρόταση είναι σωστή, κάνει τις βάσεις δεδομένων Thomson ISI να παρέχουν επαρκή κάλυψη της θεματολογίας. Έχει λοιπόν μεγαλύτερη συνεισφορά μια πιο τοπική διανομή από αυτήν με εθνική διάσταση. Αρκετοί ερευνητές έχουν προσπαθήσει να απαντήσουν σε αυτήν την ερώτηση και οι σπουδές τους επικεντρώνονται στο βαθμό κάλυψης των SSCI και AHCI βάσεων, ανάλογα με τη χώρα και τη γλώσσα δημοσίευσης.

Σύμφωνα με τους Royle και Over (Royle and Over, 1994), το 73% των αυστραλιανών άρθρων στις φυσικές επιστήμες καλύπτονται από την SCI, αλλά μόνο το 27% των αυστραλιανών άρθρων στις κοινωνικές επιστήμες καλύπτονται από την SSCI. Οι Nederhof και Zwaan (1991) σημείωσαν ότι η κάλυψη που παρέχεται από τις δύο βάσεις δεδομένων διαφέρει σημαντικά από τον τομέα, τη σημασία και τη γλώσσα του περιοδικού. Για παράδειγμα μόνο το 3% των ολλανδικών άρθρων στη δημόσια διοίκηση καλύφθηκαν από τη SSCI, σε σύγκριση με το 58% των τεκμηρίων στην πειραματική ψυχολογία. Όσον αφορά στις ανθρωπιστικές επιστήμες, η κάλυψη κυμαίνεται από το 10% των τεκμηρίων σε μελέτες της ολλανδικής γλώσσας για το 39% των ειδών στη γενική Βιβλιογραφία.

Σύμφωνα με τον Kyvik (1988), μόνο το ένα τρίτο των νορβηγικών δημοσιευμάτων στην SSH, κατά την περίοδο 1979-1981, γράφτηκαν σε γλώσσα άλλη από τη νορβηγική, σε σύγκριση με το 74% των δημοσιεύσεων στις φυσικές επιστήμες. Είναι σαφές ότι αυτό δείχνει ότι οι μεταβλητές της γλώσσας είναι σημαντικές. Υπάρχει μια καλά εδραιωμένη φήμη στη Γαλλία, Ισπανία και Γερμανία οι οποίες υπό-εκπροσωπούνται στο SSCI (Ingwersen, 2000). Ακόμα κι αν και οι τρεις χώρες απολαμβάνουν μια ισχυρή παράδοση στις SSH, το γεγονός αυτό δεν αντικατοπτρίζεται στο SSCI. Για παράδειγμα μια μελέτη που καλύπτει τις περιόδους 1989-1993 και 1994-1998 αναφέρει ότι η Γερμανία είναι η τελευταία από τις 17 χώρες οι οποίες κατατάσσονται ανάλογα με την επίδραση της έρευνας στις κοινωνικές επιστήμες (Ingwersen, 2000). Πιο συγκεκριμένα, ο Schoepflin (1992) δίνει τα αποτελέσματα των ερευνών των γερμανών επιστημόνων, οι

οποίοι κλήθηκαν να προσδιορίσουν τα ακαδημαϊκά περιοδικά με το υψηλότερο προφίλ και τη μεγαλύτερη αξία σε σχέση με την ειδικότητά τους. Με βάση τις απαντήσεις τους, το SSCI καλύπτει το 94% των γερμανικών περιοδικών στην αναπτυξιακή Ψυχολογία, αλλά μόνο το 26% στον τομέα της Κοινωνιολογίας και το 8% στην Εκπαίδευση. Με δεδομένη την ισχυρή γερμανική παράδοση στις κοινωνικές επιστήμες, τα στοιχεία αυτά είναι πολύ περίεργα.

Από τότε που η Thomson ISI επιλέγει περιοδικά ανάλογα με τον αριθμό των αναφορών που λαμβάνουν, οι συνηθισμένες αναφορές των διαφόρων γλωσσικών κοινοτήτων διαδραματίζουν σημαντικό ρόλο στην πραγματική συλλογή των βάσεων δεδομένων Thomson ISI. Για παράδειγμα οι κοινωνιολόγοι των Η.Π.Α. και του Η.Β. (Ηνωμένου Βασιλείου) αναφέρουν άρθρα που έχουν γραφτεί στην αγγλική γλώσσα στο 99% των περιπτώσεων. Την ίδια στιγμή, τα άρθρα αυτά αντιπροσωπεύουν περίπου το 70% της διεθνούς βιβλιογραφίας στην Κοινωνιολογία (Yitzhaki, 1998). Αυτό εξηγεί σε μεγάλο βαθμό την αγγλοσαξονική υπέρ-εκπροσώπηση στις βάσεις δεδομένων Thomson ISI.

Μια σύγκριση του καταλόγου των κοινωνικών επιστημονικών περιοδικών της UNESCO με εκείνη της ISI αποκαλύπτει κάποιες πολύ σημαντικές διαφορές (Schopflin, 1992). Η λίστα της UNESCO περιέχει περίπου 2,5 φορές περισσότερα ακαδημαϊκά περιοδικά από τη λίστα της SSCI. Τα περιοδικά των ΗΠΑ αντιπροσωπεύουν περίπου το 60% της κάλυψης της SSCI, αλλά αντιπροσωπεύουν μόνο το 17% του συνόλου των περιοδικών ανάλογα με τη λειτουργική λίστα της UNESCO. Η σύγκριση δείχνει ότι η SSCI περιλαμβάνει περισσότερα περιοδικά από τις ΗΠΑ, από τον αριθμό που περιλαμβάνεται στον κατάλογο της UNESCO.

Υπό το φως αυτών των στατιστικών στοιχείων, ένας αριθμός βιβλιομετρητών ισχυρίζεται ότι οι βάσεις δεδομένων SSCI και AHCI έχουν μια προκατάληψη υπέρ της αγγλικής γλώσσας και των περιοδικών που προέρχονται από τις αγγλοσαξονικές χώρες.

Η Εικόνα 1.5 παρουσιάζει στατιστικά στοιχεία αναλογικά με τον εκδότη των περιοδικών ανά χώρα. Δίνει, με τη σειρά, τα στοιχεία για περιοδικά στις NSE και SSH που καλύπτονται σε βάσεις δεδομένων της Thomson ISI και στη βάση δεδομένων του περιοδικού Ulrich, στην οποία αναγνωρίζεται πως παρέχει έναν από τους πιο εξαντλητικούς διαθέσιμους καταλόγους. Ο πίνακας δείχνει ότι τα περιοδικά με συντάκτες του Ηνωμένου Βασιλείου σε πολύ μεγάλο βαθμό υπέρ-εκπροσωπούνται στις βάσεις δεδομένων Thomson ISI, ειδικά στις SSH. Σύμφωνα με την Ulrich, το 18% των περιοδικών έχουν συντάκτες του Ηνωμένου Βασιλείου, παρόλο που οι δείκτες της Thomson ISI δείχνουν το 27 % των αξιολογήσεων με έναν συντάκτη σε αυτή τη χώρα με μια υπέρ-εκπροσώπηση με συντελεστή 55%. Τα περιοδικά στις SSH δείχνουν ότι οι συντάκτες που βρίσκονται στη Ρωσική Ομοσπονδία, στις ΗΠΑ, την Ελβετία και την Ολλανδία υπέρ-εκπροσωπούνται, ενώ σχεδόν όλες οι άλλες χώρες υπο-εκπροσωπούνται.

Country	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
United Kingdom	23%	17%	36%	27%	18%	55%
Russian Federation	1.6%	1.4%	12%	0.3%	0.3%	36%
United States	36%	31%	19%	50%	37%	35%
Switzerland	2.7%	2.1%	26%	0.6%	0.5%	8%
Netherlands	9.4%	8.3%	14%	7.7%	7.4%	5%
Canada	1.3%	1.3%	1%	2.5%	3.2%	-21%
France	2.4%	2.6%	-6%	1.0%	1.4%	-24%
Germany	7.7%	6.2%	25%	3.9%	5.9%	-34%
Japan	2.3%	3.7%	-39%	0.5%	1.0%	-55%
Australia	1.2%	2.1%	-42%	1.1%	3.6%	-71%
Spain	0.4%	1.3%	-72%	0.3%	1.0%	-75%
Belgium	0.2%	0.4%	-52%	0.5%	2.1%	-75%
India	0.9%	2.2%	-61%	0.2%	1.6%	-86%
Poland	0.7%	1.6%	-58%	0.2%	1.3%	-87%
Italy	1.1%	1.7%	-38%	0.1%	1.2%	-89%
China	0.9%	2.9%	-69%	0.1%	0.9%	-91%
Brazil	0.3%	1.1%	-72%	0.04%	1.0%	-96%
Other	7.5%	14%	-45%	3.5%	13%	-73%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

Εικόνα 1 5 Παρουσιάζει στατιστικά στοιχεία αναλογικά με τον εκδότη των περιοδικών ανά χώρα

Ο πίνακας 1.5 δίνει την κατανομή της κάλυψης της Thomson ISI και της Ulrich με κύρια γλώσσα της χώρας του συντάκτη του περιοδικού. Αυτό δείχνει ότι μόνο περιοδικά με συντάκτες σε χώρες όπου ομιλείται η γλώσσα Αγγλικά ή Ρωσικά υπέρ-εκπροσωπούνται. Για παράδειγμα, τα περιοδικά στις SSH με τους συντάκτες να βρίσκονται σε γαλλόφωνες χώρες υπό-εκπροσωπούνται κατά 27 % στις βάσεις δεδομένων Thomson ISI.

Language	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
Russian	1.6%	1.4%	12%	0.3%	0.3%	36%
English	64%	55%	17%	82%	65%	26%
Dutch	10%	8.6%	11%	8.2%	9.4%	-13%
Czech	0.3%	0.7%	-55%	0.3%	0.4%	-15%
French	2.4%	2.6%	-6%	1.0%	1.4%	-27%
German	11%	9.0%	19%	4.6%	7.0%	-34%
Japanese	2.3%	3.7%	-39%	0.5%	1.0%	-55%
Danish	1.2%	0.9%	40%	0.3%	0.7%	-61%
Afrikaans	0.3%	0.6%	-57%	0.4%	1.1%	-66%
Swedish	0.3%	0.4%	-10%	0.2%	0.7%	-67%
Chinese	1.7%	3.9%	-56%	0.4%	1.7%	-76%
Spanish	1.0%	3.2%	-69%	0.6%	3.0%	-82%
Polish	0.7%	1.6%	-58%	0.2%	1.3%	-87%
Italian	1.1%	1.7%	-38%	0.1%	1.2%	-89%
Portuguese	0.3%	1.1%	-74%	0.1%	1.1%	-92%
Arabic	0.1%	0.6%	-90%	0.0%	0.5%	-100%
Other	2.3%	5%	-57%	1.1%	4%	-76%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

Εικόνα 1 6 Παρουσιάζει στατιστικά στοιχεία αναλογικά με τον εκδότη των περιοδικών ανά χώρα

Για να καθοριστεί ο ρόλος των γλωσσικών παραγόντων στα ποσοστά κάλυψης των περιοδικών στις βάσεις δεδομένων Thomson ISI, η Science-Metrix εξέτασε επίσης την πραγματική γλώσσα των περιοδικών. Ο Πίνακας 1.6 δείχνει μια σαφή προκατάληψη επιλογής υπέρ των περιοδικών που τα άρθρα τους έχουν γραφτεί στην αγγλική γλώσσα. Ενώ το 75 % των έγκριτων περιοδικών που αποδελτιώνονται στην Ulrich είναι στα Αγγλικά, σχεδόν το 90% αυτών που επιλέγονται από την Thomson ISI είναι στα Αγγλικά, αποδίδοντας ένα ποσοστό υπέρ-επιλογής περίπου 20%. Η μόνη άλλη γλώσσα που υπέρ-εκπροσωπείται στις βάσεις δεδομένων Thomson ISI είναι η τσεχική, μια γλώσσα που παίζει περιθωριακό ρόλο στο παγκόσμιο σύστημα επιστήμης. Τα Γαλλικά υπό-εκπροσωπούνται κατά 26%.

Language	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
English	89%	78%	13%	90%	75%	20%
Czech	0.04%	0.3%	-85%	0.2%	0.2%	8%
Russian	0.5%	0.9%	-48%	0.3%	0.4%	-24%
French	3.3%	3.4%	-3%	3.2%	4.4%	-26%
Multiple languages	0.2%	0.2%	-14%	0.3%	0.5%	-45%
Dutch	2.2%	2.2%	0%	1.3%	2.6%	-48%
German	3.2%	3.9%	-18%	3.0%	5.8%	-50%
Japanese	0.4%	1.7%	-74%	0.2%	0.6%	-64%
Swedish	-	0.1%	-100%	0.1%	0.4%	-69%
Spanish	0.6%	2.6%	-75%	0.9%	3.0%	-69%
Italian	0.1%	0.8%	-83%	0.2%	1.1%	-80%
Danish	0.04%	0.1%	-50%	0.1%	0.3%	-83%
Portuguese	0.1%	0.7%	-85%	0.1%	1.0%	-86%
Chinese	0.3%	2.4%	-88%	0.04%	1.2%	-96%
Polish	0.05%	0.7%	-92%	-	0.9%	-100%
Arabic	-	0.1%	-100%	-	0.3%	-100%
Turkish	0.01%	0.2%	-95%	-	0.1%	-100%
Other	0.3%	1.6%	-80%	0.3%	2.4%	-87%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

Εικόνα 1 7 Ο ρόλος των γλωσσικών παραγόντων στα ποσοστά κάλυψης των περιοδικών

Τα δεδομένα δείχνουν σαφώς ότι η επιλογή περιοδικών της Thomson ISI στις SSH γίνεται υπέρ των Αγγλικών. Αυτό μπορεί να οφείλεται σε διαφορές στην ποιότητα των αποτελεσμάτων της έρευνας σύμφωνα με τη γλώσσα των άρθρων. Σύμφωνα με τους Hodgson και Rothman (1999), 388 ή 84%, από τους 463 συντάκτες των 30 πιο σημαντικών οικονομικών περιοδικών είναι συνδεδεμένοι με ιδρύματα των ΗΠΑ. Τούτου λεχθέντος, είναι αμφίβολο κατά πόσο τα ερευνητικά άρθρα γραμμένα σε όλες τις γλώσσες, εκτός από τα Αγγλικά, είναι κατώτερης ποιότητας σε όλες τις περιπτώσεις. Αντιθέτως, η κατάσταση αυτή οφείλεται στην αδυναμία της Thomson ISI να αναλύσει το περιεχόμενο των περιοδικών σε άλλες γλώσσες, πλην των Αγγλικών, κάτι που αναφέρεται στην ιστοσελίδα της Thomson ISI: Οι τίτλοι των άρθρων, οι περιλήψεις και οι λέξεις-κλειδιά στα Αγγλικά. Οι αναφορές που παρατίθενται στα Αγγλικά επίσης. Παρά το γεγονός ότι σημαντικές επιστημονικές πληροφορίες δημοσιεύονται σε όλες τις γλώσσες, οι συγγραφείς πρέπει να παρέχουν μεταφράσεις των τίτλων των άρθρων, των λέξεων-κλειδίων, καθώς και αποσπάσματα στα Αγγλικά, αν ελπίζουν να συγκεντρώσουν το ευρύτερο δυνατό κοινό.

Είναι σημαντικό να υπολογίσουμε αυτήν την προκατάληψη σε οποιαδήποτε διεθνή συγκριτική ανάλυση. Πράγματι, κάθε αξιολόγηση με βάση τον 'Social Science Citation Index' (SSCI) και 'Arts and Humanities Citation Index' (AHCI) μπορεί κάλλιστα να υπερτιμήσει τις ΗΠΑ, το Ηνωμένο Βασίλειο και τον Καναδά και να υποτιμήσει τη Γερμανία, την Ισπανία, τη Γαλλία και άλλες χώρες που δεν μιλούν Αγγλικά και αυτή η προκατάληψη θα επηρεάσει τόσο τις μετρήσεις της δημοσίευσης, όσο και τις αναλύσεις αναφορών.

Σύμφωνα με ορισμένους συγγραφείς, το πρόβλημα προκύπτει μόνο όταν η βιβλιομετρική αξιολόγηση πρέπει να καλύπτει τοπικά προσανατολισμένο αποτέλεσμα έρευνας. Κατά τη γνώμη των Moed, Luwel και Nederhof (2002), η πραγματική ακαδημαϊκή έρευνα θα πρέπει να είναι σχετικά διεθνής και ο τοπικός προσανατολισμός δεν πρέπει να αποτελεί παράγοντα στις βιβλιομετρικές αξιολογήσεις. Από την άποψη αυτή, η έρευνα που δεν καλύπτεται από τις βάσεις δεδομένων Thomson ISI απλώς αποτυγχάνει να φθάσει στο όριο ενδιαφέροντος που θα δικαιολογούσε τη στενότερη αξιολόγηση. Αυτό είναι ένα αδύναμο επιχείρημα, επειδή μετατρέπει την Thomson ISI σε ένα αμερόληπτο δικαστή του τι είναι και τι δεν είναι «ποιοτικό» αποτέλεσμα της έρευνας. Στην πραγματικότητα, επειδή τα κριτήρια επιλογής της απαιτούν οι πληροφορίες βιβλιογραφίας περιοδικών να είναι στα Αγγλικά, η Thomson ISI μπορεί να αποτύχει στην ευρετηρίαση του περιεχομένου

ενός εξαιρετικού περιοδικού φιλοσοφίας για παράδειγμα, επειδή το συνολικό περιεχόμενο του είναι στα Γερμανικά.

Υπό το πρίσμα των παραπάνω στατιστικών στοιχείων, υπάρχει λόγος ανησυχίας σχετικά με την επίδραση παραγόντων της γλώσσας στη βιβλιομετρική μέτρηση. Παρόλο που η Thomson ISI φαίνεται προκατειλημμένη υπέρ των άρθρων στα αγγλικά, η ερευνητική παραγωγή στη γλώσσα του Κεμπέκ και στα Γαλλικά θα μπορούσε να μην υποτιμηθεί σε σύγκριση με τα αγγλόφωνα άρθρα και περιοδικά. Η Science-Metrix εξέτασε το ζήτημα αυτό στο πλαίσιο της συγκριτικής ανάλυσης των βάσεων δεδομένων Ulrich και Thomson ISI. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 1.7, ο οποίος δείχνει ότι το Κεμπέκ είναι, στην πραγματικότητα, κάπως υπέρ-εκπροσωπημένο στις SSH στον Καναδά (17% των καναδικών συντακτών στις βάσεις δεδομένων Thomson ISI και 15% στην Ulrich). Με άλλα λόγια, με βάση τον τόπο καταγωγής των συντακτών στις SSH, το Κεμπέκ δεν υπό-εκπροσωπείται στις βάσεις δεδομένων Thomson ISI.

	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
Quebec	11%	13%	-18%	17%	15%	12%
Other provinces	89%	87%	3%	83%	85%	-2%

Source: Compiled by Science – Metrix from Thomson ISI's and Ulrich's data.

Εικόνα 1 8 Συγκριτική ανάλυση των βάσεων δεδομένων Ulrich και Thomson ISI.

Τελικά, εναπόκειται στους επαγγελματίες των βιβλιομετρικών αξιολογήσεων να καθορίσουν εάν η πιο τοπικά προσανατολισμένη έρευνα μπορεί να ληφθεί υπόψη επαρκώς ή αν μπορούν να υπολογισθούν μόνο οι διεθνείς συνεισφορές. Είναι επίσης στο χέρι τους να προειδοποιούν τον αναγνώστη τους – τους αναλυτές της επιστημονικής πολιτικής και τους λήπτες αποφάσεων - ότι τα εργαλεία τους έχουν σοβαρές ελλείψεις, ιδιαίτερα στις ανθρωπιστικές επιστήμες. Προς το παρόν, η κάλυψη των SSCI και AHCI είναι κατάλληλη μόνο στην περίπτωση των περιοδικών διεθνούς εμβέλειας που δημοσιεύονται στα Αγγλικά. Αυτή η κατάσταση θα μπορούσε να αλλάξει σύντομα, όμως μια πρόσφατη ερευνητική εργασία δείχνει ότι η ισχύς των δεδομένων SSCI για τη συγκριτική αξιολόγηση των επιδόσεων στις SSH των μη αγγλοσαξονικών χωρών έχει αυξηθεί τα τελευταία χρόνια (Ingwersen, 2000). Αυτό το συμπέρασμα υποστηρίζεται από τον Hicks (1999 και 2004), ο οποίος ισχυρίζεται, μεταξύ άλλων, ότι οι κοινωνικές επιστήμες γίνονται όλο και πιο διεθνείς και λιγότερο κατακερματισμένες, καθώς οι μελετητές υιοθετούν όλο και περισσότερο κοινές έννοιες και παραδείγματα. Ως αποτέλεσμα, μια σειρά ερευνητικών πεδίων αναπτύσσουν μια βασική ομάδα των ακαδημαϊκών περιοδικών που μπορεί να αποτελέσει μια σταθερή βάση για αποκλειστική βιβλιομετρική ανάλυση.

Το ευρύτερο φάσμα των μέσων διάδοσης της γνώσης στις SSH από ότι στις NSE και το γεγονός ότι τα αποτελέσματα της έρευνας στις SSH είναι μερικές φορές περισσότερο τοπικού προσανατολισμού δεν είναι οι μόνοι παράγοντες που καθιστούν δύσκολο να χρησιμοποιηθεί η Βιβλιομετρία για τη μέτρηση της έρευνας στις SSH. Πρέπει επίσης να δοθεί προσοχή και στις ξεχωριστές πρακτικές αναφοράς των τμημάτων στις SSH.

1.5.4. Πρακτικές αναφορών

Η στέρεα, έγκυρη βιβλιομετρική ανάλυση απαιτεί μεγάλες ποσότητες δεδομένων (Van Raan, 2003) αφού πρέπει να είναι δυνατή η εφαρμογή του νόμου των μεγάλων αριθμών. Λαμβάνοντας υπόψη αυτή την απαίτηση, μερικά στατιστικά στοιχεία για τις βάσεις δεδομένων Thomson ISI χρησιμεύουν για να εξηγήσουν, γιατί οι μετρήσεις των άρθρων και η ανάλυση αναφορών μπορεί να είναι προβληματικές στις SSH. Κάθε εβδομάδα, περίπου 17.000 νέα άρθρα και 300.000 νέες αναφορές προστίθενται στη βάση δεδομένων SCI, 2800 άρθρα και 50.000 αναφορές στην SSCI, και 2200 άρθρα και 15.000 αναφορές στην AHCI (CNER, 2002). Τα SSCI άρθρα παραπέμπουν σε περίπου 72.000 διακριτά αντικείμενα. Εκείνα στην SCI, σε περίπου 195.000. Το 1999, το SSCI περιείχε 2,3 εκατομμύρια αναφορές, ενώ η SCI περιείχε πάνω από 20 εκατομμύρια (Leydesdorff, 2003).

Τα Journal Citation Reports (JCR) είναι μια έκδοση της Thomson ISI που δίνει στοιχεία για τον παράγοντα των επιπτώσεων στα περιοδικά. Το 1999, το JCR για τις κοινωνικές επιστήμες κάλυπτε περίπου 1700 περιοδικά παρόλο που το JCR για τις NSE καλύπτονταν από 5500. Στην περίπτωση του AHCI, δεν υπάρχει δημοσίευση που να δίνει στοιχεία στις αναφορές. Έτσι υπάρχει μια σημαντική διαφορά μεταξύ των SCI, SSCI και AHCI σε σχέση με τον αριθμό των αναφορών. Οι παράγοντες αυτοί που παρουσιάζονται στις κοινωνικές επιστήμες έχουν επιπτώσεις στις JCR και είναι λιγότερο αξιόπιστοι από τους παράγοντες των NSE, λόγω της σταθερής αναφοράς των δύο ετών που χρησιμοποιεί η Thomson ISI για τον υπολογισμό του συντελεστή επιπτώσεων. Στις SSH, ένα άρθρο δεν είχε αρχίσει να έχει πραγματικό αντίκτυπο μέχρι πριν από λίγα χρόνια μετά τη δημοσίευση. Απόδειξη αυτού έγκειται στη φήμη ότι μόνο 32 από τους 100 πιο γνωστούς συγγραφείς που παρατίθενται στις ανθρωπιστικές επιστήμες γεννήθηκαν κατά τον εικοστό αιώνα (Nederhof Luwel & Moed 2001 αναφέροντας Garfield, 1979).

Discipline	No. of articles	Percentage of journals	Mean No. of references	Mean reference age
Immunology	23,396	94.3	29.6	6.9
Research Medicine	24,369	92.1	25.9	7.9
Solid State Physics	28,466	85.2	23.6	10.1
Analytical Chemistry	9,605	83.8	20.9	9.4
Mathematics	11,987	64.4	16.2	11.3
Psychology and Psychiatry	11,886	64.0	31.0	11.4
Electronic Engineering	19,222	62.2	15.0	8.6
Business	3,663	56.0	20.8	10.9
Economics	7,959	48.7	21.6	10.6
Library and Information Science	2,128	47.6	14.9	9.1
Sociology	3,675	40.4	32.7	12.5
History and Philosophy of Science and Social Sciences	658	34.7	48.7	38.8

Εικόνα 1 9 Χαρακτηριστικά των SSH αναφορών που μπορούν να περιορίσουν την εφαρμογή της βιβλιομετρίας. *Science metrix.com*

Υπάρχουν και άλλα χαρακτηριστικά των αναφορών στις SSH που μπορούν να περιορίσουν την εφαρμογή της Βιβλιομετρίας. Οι Glänzel και Schoepflin (1999) περιέγραψαν τις ειδικές βιβλιομετρικές ιδιότητες των κλάδων των κοινωνικών και φυσικών επιστημών (βλ. εικόνα 1.9). Το ποσοστό των ερευνητικών άρθρων χωρίς αναφορές είναι μόνο 3% στις φυσικές επιστήμες, αλλά μεταξύ 3% και 30% σε διάφορους κλάδους των κοινωνικών επιστημών. Οι αριθμοί είναι σχετικά χαμηλοί στην ιστορία της επιστήμης, της ψυχολογίας, της κοινωνιολογίας και της οικονομίας, αλλά το ποσοστό είναι 26% στην πληροφορική και στην επιστήμη της Βιβλιοθηκονομίας και 28% στις επιχειρήσεις.

Το 1993, στην SCI αναφέρονται περίπου 28.500 άρθρα σε Φυσική στερεάς κατάστασης, ενώ στην SSCI αναφέρονται περίπου 3700 είδη περιοδικών στην Κοινωνιολογία. Ωστόσο η Κοινωνιολογία έχει υψηλότερο μέσο αριθμό αναφορών ανά άρθρο (περίπου 33) από τη Φυσική (περίπου 24), ενώ αυτό σημαίνει μόνο ότι η προτιμώμενη ηλικία είναι ελαφρώς υψηλότερη (12 και 10 ετών αντίστοιχα) (Glänzel and Schoepflin, 1999). Με βάση αυτές τις στατιστικές, ο αριθμός των άρθρων και αναφορών που μπορούν να αναλυθούν στις SSH είναι πολύ μικρός σε σύγκριση με την κατάσταση στις NSE.

Όπως αναφέρθηκε προηγουμένως, υπάρχει μια σημαντική διαφορά μεταξύ του ποσοστού των αναφορών των ακαδημαϊκών άρθρων σε περιοδικά και του ποσοστού των αναφορών σε άλλα μέσα διάδοσης της γνώσης. Ο δείκτης για τα άρθρα είναι μόνο 40% στην Κοινωνιολογία, αλλά είναι διπλάσιος (85%) στη Φυσική. Άλλα στατιστικά στοιχεία μπορούν να προσκομιστούν προς στήριξη της ανάγκης να δοθεί μεγαλύτερη σημασία στα τεκμήρια τα οποία καλύπτονται ελάχιστα, αν όχι καθόλου, από τις SSCI και AHCI. Το 1999, περίπου 2,4 εκατομμύρια βιβλιογραφικές αναφορές συγκεντρώθηκαν στο SSCI. Ωστόσο, μόνο περίπου ένα εκατομμύριο από τις αναφερόμενες παραπομπές ή 45% περιλαμβάνονται στο SSCI. Το ποσοστό αυτό είναι 79% για το SCI (Leydesdorff, 2003).

Ο χρόνος που απαιτείται για να συσσωρευτούν οι αναφορές στις SSH καθιστά δύσκολη την ανάλυση, ιδιαίτερα όταν ο στόχος είναι να βοηθήσει στη λήψη αποφάσεων και τη χάραξη πολιτικής. Ο κλάδος της αναφοράς ποικίλλει ανάλογα με τον τομέα. Η ανάλυση αυτή μπορεί να γίνει σε ορισμένους τομείς, αλλά θα πρέπει να αποφεύγεται για λόγους μεθοδολογίας σε άλλους, όπου ο νόμος των μεγάλων αριθμών δεν εφαρμόζεται και οι αναφορές δεν συντάχθηκαν επαρκώς και αντανακλούν το κυρίαρχο μέσο διάδοσης. Οι μονογραφίες και οι σχετικές αναφορές απλά δεν μπορούν να αγνοηθούν.

Η ανάλυση συν-αναφοράς είναι μια μέθοδος που μπορεί να χρησιμοποιηθεί για να χαρακτηρίσει τη δομή της επιστήμης. Είναι ένα μέσο δυνητικά προσδιορισμένο των αναδυόμενων ερευνητικών θεμάτων και πεδίων. Ο Hicks (1987) καθόρισε ένα μεγάλο πρόβλημα για τη χρήση της ανάλυσης συν-αναφοράς για τους σκοπούς της πολιτικής της επιστήμης. Η μέθοδος απαιτεί ένα σχετικά μεγάλο αριθμό άρθρων και αναφορών, έτσι ώστε οι ερευνητές να έχουν μια αντιπροσωπευτική μάζα για να εργαστούν. Ωστόσο ένα νέο πεδίο έρευνας μπορεί να προκύψει πολύ πριν από την αντιπροσωπευτική μάζα που απαιτείται για την ανάλυση της συν-αναφοράς, που έχει ήδη δημιουργηθεί. Ο Hicks δίνει το παράδειγμα της έρευνας όπου η αντιπροσωπευτική μάζα αναπτύχθηκε πριν από δέκα χρόνια πριν από το πεδίο που άρχισε να διαφαίνεται. Το χρονικό διάστημα που μεσολαβεί μεταξύ της εμφάνισης και της ανάπτυξης του τομέα της κρίσιμης μάζας, προφανώς, ποικίλλει και επηρεάζεται εν μέρει από εξωτερικούς παράγοντες (κοινωνιολογικούς παράγοντες, κ.λπ.). Γι' αυτό είναι δύσκολο να εφαρμοστεί ειδικά η ανάλυση της συν-αναφοράς στο πλαίσιο της πολιτικής της επιστήμης και η δυσκολία επιτείνεται από τη φήμη των βιβλιομετρικών προσδιορισμών των αναδυόμενων ερευνητικών πεδίων στις SSH, όπου ο αριθμός των αναφορών χτίζεται με βραδύτερους ρυθμούς.

Ο Hicks προσδιορίζει ένα άλλο εμπόδιο για την ανάλυση της αναφοράς για σκοπούς πολιτικής της επιστήμης - η διαδικασία της οριοθέτησης στα ερευνητικά πεδία-. Αυτό το βήμα είναι απαραίτητο για όλες τις βιβλιομετρικές αναλύσεις. Σύμφωνα με τον Hicks, η διαδικασία οριοθέτησης θα πρέπει να καλύπτει πάντα τις ακόλουθες απαιτήσεις: Θα πρέπει να είναι σαφής, αυτό πρέπει να φαίνεται σε έναν εύλογο αριθμό των μελετητών στο πεδίο (αν και μάλλον δεν συμφωνούν σχετικά με τα όρια της περιοχής του πεδίου λάθους τους) και θα πρέπει να είναι σταθερή διαχρονικά, έτσι ώστε οι τάσεις να μπορούν να καθοριστούν σε επαρκή βαθμό αξιοπιστίας. Δυστυχώς, ο αριθμός των από κοινού αναφορών που αναλύει δεν πληροί αυτές τις απαιτήσεις.

Ο Hicks χαρακτηρίζει ακόμη προβλήματα τα λάθη αναφοράς και την υπό-εκπροσώπηση ορισμένων τύπων τεκμηρίων. Τέλος, ο Hicks τονίζει το απαγορευτικό κόστος της ανάπτυξης της, σε μια στερεά βάση ανάλυσης συν-αναφοράς. Σε πολλές περιπτώσεις, οι απαιτούμενοι πόροι περιλαμβάνουν εμπειρογνώμονες στη Βιβλιομετρία και την Επιστήμη της πολιτικής ή της διαχείρισης της έρευνας. Ο Van Raan (1998) αμφισβητεί το συμπέρασμα του Hicks, ισχυριζόμενος ότι οι βιβλιομετρικές αναλύσεις είναι μια ανέξοδη με πληρότητα αξιολόγηση από ομότιμους. Στην πραγματικότητα, οι δύο συγγραφείς αναφέρονται σε δύο διαφορετικές βιβλιομετρικές μεθόδους. Η πρώτη συζήτηση από τον van Raan είναι απλούστερη και μπορεί να εξηγήσει τη διαφορά απόψεων. Σημειώστε ότι η ανάλυση του Hicks χρονολογείται από το 1987. Επίσης με την ανάπτυξη νέου λογισμικού, το κόστος μπορεί να μειωθεί.

1.6. Η Βιβλιομετρία και η μέτρηση της επιστήμης

Οι κυβερνήσεις σε όλες τις χώρες έχουν αντιληφθεί σταδιακά την ανάγκη για κριτική ανάλυση των επιστημονικών και τεχνολογικών πολιτικών τους. Μερικοί έκριναν ότι αρκεί να δημιουργηθούν διοικητικές μονάδες στο πλαίσιο των υπουργείων έρευνας των χωρών τους (υπουργεία Παιδείας, Βιομηχανίας, κ.λπ.). Άλλοι προτίμησαν την εκπαίδευση και κατάρτιση ειδικών, καθώς και την ανάπτυξη δεικτών σε ένα ακαδημαϊκό περιβάλλον που ενθαρρύνει την αλληλεπίδραση των ιδεών.

Στις Ηνωμένες Πολιτείες, το Εθνικό Ίδρυμα Επιστημών δημοσίευσε τους πρώτους του Δείκτες Επιστήμης & Τεχνολογίας, το 1972. Κατά την παρουσίαση του έργου τους, οι αρμόδιοι υπάλληλοι εξήγησαν ότι «ο απώτερος στόχος αυτής της προσπάθειας είναι ένα σύνολο δεικτών που θα αποκαλύψει τα πλεονεκτήματα και τις αδυναμίες της επιστήμης και της τεχνολογίας των ΗΠΑ, από την άποψη της ικανότητας και των επιδόσεων της επιχείρησης να συμβάλει στην επίτευξη των εθνικών στόχων. Εάν δεν μπορούν να αναπτυχθούν οι δείκτες αυτοί κατά τα επόμενα χρόνια, θα πρέπει να συνδράμουν στη βελτίωση της κατανομής και της διαχείρισης των πόρων που διατίθενται για την επιστήμη και την τεχνολογία, καθώς και στην καθοδήγηση της έρευνας και της ανάπτυξης του αμερικανικού έθνους σε μονοπάτια που θα είναι απόλυτα εποικοδομητικά και ωφέλιμα για την κοινωνία μας» (NSF, 1972). Στις επόμενες εκθέσεις, ο ρόλος των βιβλιομετρικών δεικτών επεκτάθηκε σημαντικά.

Έκτοτε η Βιβλιομετρία είναι όλο και περισσότερο προσανατολισμένη προς την πολιτική της επιστήμης. Ομάδες βιβλιομετρητών από διαφορετικές σχολές έχουν προτείνει διάφορες μεθόδους μέτρησης της επιστημονικής εξέλιξης και διασύνδεσης των μεθοδολογιών τους με τις αξιολογήσεις. Κατά συνέπεια η Βιβλιομετρία εισέρχεται τώρα σε μία δύσκολη φάση, κατά την οποία προσπαθεί να συμβάλει στην αξιολόγηση. Πολλές από αυτές τις μεθόδους έχουν παρουσιαστεί σε εξειδικευμένα σεμινάρια ανά τον κόσμο.

Μέσα σε μισό αιώνα, η Βιβλιομετρία έχει συνεπώς καταφέρει να κερδίσει τη θέση της ως βασικό όργανο μέτρησης της επιστήμης, τόσο στις βιομηχανικές χώρες του δυτικού κόσμου όσο και στις εκβιομηχανισμένες χώρες της Ανατολικής Ευρώπης. Το 1989, ο ΟΟΣΑ αφιέρωσε ένα κεφάλαιο του εγχειριδίου Frascati στον τομέα της τριτοβάθμιας εκπαίδευσης, επισφραγίζοντας με τον τρόπο αυτό τη θέση της Βιβλιομετρίας στην ανάλυση της επιστήμης (ΟΟΣΑ, 1989, σσ. 49-53).

Οι Κάτω Χώρες και το Ηνωμένο Βασίλειο ήταν μεταξύ των πρώτων που δημοσίευσαν τακτικές μελέτες αναφορικά με την επιστήμη που εφαρμόζει βιβλιομετρικούς δείκτες. Οι ερευνητικές ομάδες σε αυτές τις χώρες ήταν πρωτοπόρες στη θεωρητική κατασκευή και την πρακτική εφαρμογή στον τομέα αυτό (Irvine & Martin 1980; Martin & Irvine 1984; Leven 1982; Moed κ.ά., 1983), που ενσωματώνει κατά τη δεκαετία του 1980, τα μέτρα της Βιβλιομετρίας στην ανάλυση της πολιτικής της επιστήμης. Το 1987, το Ιαπωνικό Υπουργείο Παιδείας, Επιστημών και Πολιτισμού ανέθεσε σε μια ομάδα βιβλιομετρητών τη διεξαγωγή μιας συγκριτικής μελέτης του αριθμού των επιστημονικών άρθρων που έχουν δημοσιευθεί σε επτά (7) μεγάλες χώρες, προκειμένου να «δημιουργηθεί ένας δείκτης που θα επιτρέπει την βαθύτερη κατανόηση των ερευνητικών δραστηριοτήτων της Ιαπωνίας στη διεθνή κοινότητα».

Ωστόσο, σε όλες τις χώρες χρειάστηκε μεγάλο χρονικό διάστημα έως ότου η βιβλιομετρική προσέγγιση να κερδίσει την αποδοχή των πολιτικών και των επιστημονικών κύκλων ως τρόπος μέτρησης της επιστήμης. Ορισμένοι επιστήμονες συνεχίζουν να είναι εχθρικοί απέναντι στο ρόλο που διαδραματίζει η Βιβλιομετρία. Πολλοί από αυτούς δεν έχουν εξοικειωθεί με τη μεθοδολογία και αισθάνονται άβολα απέναντι σε μια επικείμενη ανάλυση (ή ακόμη και αξιολόγηση) με ποσοτική μέτρηση του επιπέδου της ερευνητικής τους δραστηριότητας. Ενδεικτική είναι η ακόλουθη φράση: «Πρόκειται για ένα παντελώς άκομψο μέσο μέτρησης του επιπέδου της ερευνητικής δραστηριότητας, τόσο άκομψο, μάλιστα, σε βαθμό που κάποιος μπορεί να πουν ότι είναι απλώς καταμέτρηση χαρτιού» (Rappa, 1989, σ. 28). Δεν είναι εύκολο ψυχολογικά ή διανοητικά να μεταπηδήσει κανείς από τη μέτρηση σε επίπεδο χώρας ή επιστημονικού κλάδου στην αξιολόγηση του κάθε ερευνητή ξεχωριστά. Η βιβλιομετρική προσέγγιση έχει χαρακτηριστεί ακόμη και ως «ανυπόφορο σκάνδαλο» (Chauvin, 1991, σ. 782). Μερικοί επιστήμονες έχουν προτείνει μια εναλλακτική μέθοδο αξιολόγησης, σύμφωνα με την οποία οι αιτούντες εργασίας θα κρίνονται όχι βάσει ενός καταλόγου δημοσιευμένων έργων, που θα προέρχεται από κάποια ανώνυμη και απόρρητη βάση δεδομένων, αλλά βάσει επιλεγμένων άρθρων – μελετών, που οι ίδιοι θεωρούν ως αντιπροσωπευτικό δείγμα του έργου τους. Οι ερευνητές θεωρούν την αξιολόγηση που προέρχεται από συναδέλφους ως μοναδικό τρόπο άσκησης κριτικής. Την άποψη αυτή συμμαρτίζει μεγάλο μέρος της επιστημονικής κοινότητας (Ourisson, 1991).

Σε ατομικό επίπεδο, η Βιβλιομετρία καταμετρά την παραγωγικότητα της έρευνας, αλλά δεν διευκρινίζει απαραίτητα κάτι αναφορικά με την ποιότητα ή την επάρκεια των ερευνητών ως εκπαιδευτικών. Οι αντιδράσεις των επιστημόνων είναι απολύτως φυσικό να υπάρξουν, ενώ οι ίδιοι υπογραμμίζουν την ανάγκη για αλληλεπίδραση μεταξύ των ανθρώπων που αξιολογούνται και αυτών που αξιολογούν.

Κατά τη διαδικασία αξιολόγησης, ο διάλογος μεταξύ των «δημιουργών της επιστήμης» και των «βιβλιομετρο-αναλυτών» μπορεί να είναι εποικοδομητικός. Μπορεί να βοηθήσει στην τροποποίηση των δεδομένων και των μεθόδων που χρησιμοποιούνται, αλλά πάνω απ' όλα μπορεί να επηρεάσει την ερμηνεία των αποτελεσμάτων. Πολλοί ειδικοί στην «αξιολογική Βιβλιομετρία» επιβεβαιώνουν ότι η συζήτηση των αποτελεσμάτων πρέπει να αποτελεί πάντα μέρος της συνολικής διαδικασίας αξιολόγησης ενός ερευνητή (Moed et al., 1983). Στο διάλογο μεταξύ του αναλυτή και του προσώπου που αναλύεται, κρίνεται απαραίτητος ο έλεγχος των αριθμητικών στοιχείων. Σαφώς, η χρήση των βιβλιομετρικών δεικτών απαιτεί πολύ μεγαλύτερη εγρήγορση όταν προσφέρεται για την ατομική αξιολόγηση, από ότι για μια γενική περιγραφή της επιστήμης σε εθνικό επίπεδο.

Με την πάροδο του χρόνου, οι εν λόγω μέθοδοι έχουν γίνει ευρύτερα γνωστές, αλλά ίσως αυτό να ισχύει περισσότερο στους έξω-επιστημονικούς κύκλους από ότι μέσα στην ίδια την επιστημονική κοινότητα (Chelimsky, 1991). Πολλές είναι οι χώρες εκείνες που δημοσιεύουν στατιστικά στοιχεία στους Δείκτες της Επιστήμης & της Μηχανικής του Εθνικού Ιδρύματος Επιστημών (National Science Foundation - NSF):

- αυστραλιανή Επιστήμη & Τεχνολογία με μια ματιά, 1990 (Αυστραλία),
- συλλογή Επιστημονικών Δεικτών, 1991 (Καναδάς),

- δείκτες της Επιστήμης & της Τεχνολογίας, 1991 (Ιαπωνία),
- δείκτες της E & T, 1992 (Γαλλία),
- έκθεση Δεικτών E & T, 1994 (Κάτω Χώρες),
- πολιτική της E & T - Απολογισμός και προοπτικές (ΟΟΣΑ),
- ευρωπαϊκή έκθεση για τους δείκτες της επιστήμης και της τεχνολογίας.

1.7 Βιβλιογραφία -βιβλιογραφικές αναφορές και παραπομπές

1.7.1 Αναγκαιότητα χρήσης βιβλιογραφίας και αναφορών

Για τη συγγραφή μιας εργασίας ή μιας δημοσίευσης είναι αναγκαίο να αξιοποιούνται πληροφορίες και να αναφέρονται αποσπάσματα έργων άλλων δημιουργών, προκειμένου το πρόσωπο που γράφει την εργασία να υποστηρίξει τις απόψεις του, αλλά και να τεκμηριώσει τα γραφόμενά του. Σε αυτήν την περίπτωση πρέπει να στηρίζεται σε αυτό που λέμε βιβλιογραφική έρευνα (Bibliographic research), αναφέροντας τις πηγές που μελέτησε προκειμένου να αντλήσει πληροφορίες. Το αποτέλεσμα της βιβλιογραφικής έρευνας αποτυπώνεται σε ειδικό κεφάλαιο στο τέλος της εργασίας ως βιβλιογραφία (Bibliography), αποτελώντας έναν από τους πιο καθοριστικούς παράγοντες αξιολόγησης του περιεχομένου της. Με τον τρόπο αυτό αποφεύγεται η λογοκλοπή (plagiarism), η ιδιοποίηση δηλαδή της ξένης πνευματικής ιδιοκτησίας και επιπλέον επιτρέπεται στον αναγνώστη να ανατρέξει στις αρχικές πηγές, προκειμένου να επαληθεύσει ή να αντιπαραβάλει τα όσα γράφτηκαν. Ανακεφαλαιώνοντας, ως βιβλιογραφία ορίζεται το σύνολο των τεκμηρίων τις οποίες μελέτησε ο συγγραφέας για τη συγγραφή της εργασίας του, ανεξάρτητα από το αν χρησιμοποίησε ή όχι στοιχεία από αυτές. Η αναφορά (citation) είναι η διαδικασία αναγνώρισης της συνεισφοράς μιας επιστημονικής εργασίας σε κάποιο άλλο ερευνητικό έργο. Η αναφορά στο έργο ενός συγγραφέα πραγματοποιείται όταν γίνεται χρήση κάποιας ιδέας, μεθόδου ή αυτούσιου αποσπάσματος από το δημοσιευμένο έργο του.

Η χρήση αναφορών / βιβλιογραφίας βοηθά τον αναγνώστη να καταλάβει το ευρύτερο θεματικό πλαίσιο μέσα στο οποίο κινήθηκε ο συγγραφέας προκειμένου να ολοκληρώσει την έρευνα του. Έτσι, με τον τρόπο αυτόν, μπορεί να συμβουλευτεί άλλες πηγές σχετικές με το θέμα, ώστε να προχωρήσει, εάν επιθυμεί, στην περαιτέρω διερεύνηση του.

Λόγοι μελέτης των αναφορών συνοδευόμενες με τις απαραίτητες βιβλιογραφικές πληροφορίες μπορεί να αναφερθούν οι παρακάτω:

- Η αντίκρουση ή αμφισβήτηση έργων ή ιδεών διαφόρων συγγραφέων (αρνητικές απόψεις).
- Η κριτική δημοσιευμένων έργων και η τεκμηρίωση απόψεων.
- Η δημιουργία στοιχείων επικύρωσης των αποτελεσμάτων των ερευνών και η εγγύηση της διασφάλισης του επιστημονικού κύρους.
- Ο προσδιορισμός των πρωτογενών δημοσιεύσεων που περιγράφουν τα χαρακτηριστικά γνωρίσματα μιας ιδέας ή ενός όρου, καθώς βεβαιώνει την αυθεντικότητα και τη νομιμότητα των ιδεών και αντιλήψεων των δημιουργών τους.

Λόγοι μελέτης της βιβλιογραφίας μπορεί να αναφερθούν οι παρακάτω:

1. Ο εντοπισμός σχετιζόμενων έργων, σχέσεων μεταξύ των συγγραφέων και των έργων τους.
2. Η εύρεση εργαλείων, μεθοδολογιών, κ.τ.λ..
3. Η παροχή ώθησης σε έργα που δεν είναι πολύ γνωστά ή ευρετηριασμένα ή δεν έχουν χρησιμοποιηθεί ως αναφορές.
4. Η επεξήγηση των αρχικών τεκμηρίων που περιέχουν νέες μεθόδους ιδέες και εργαλεία.
5. Η απονομή φόρου τιμής στους πρωταγωνιστές των επιστημονικών κλάδων.

Οι βιβλιογραφικές αναφορές προέρχονται από το ευρετήριο αναφορών. Αυτό με τη σειρά του βασίζεται στο Αγγλικό Νομικό Σύστημα. Το αγγλικό δίκαιο λειτουργεί υπό το δόγμα της εκδίκασης βάσει παλαιών. Αυτό σημαίνει ότι όλα τα δικαστήρια πρέπει να ακολουθούν τις προηγούμενες αποφάσεις τους, καθώς κι εκείνες που καθορίζονται από ανώτερα δικαστήρια. Για να διερευνήσει ένας δικαστής μια υπόθεση σύμφωνα με την αρχή του δεδικασμένου, θα πρέπει να βασιστεί σε προηγούμενες αποφάσεις που αφορούν

μία παρόμοια νομική περίπτωση, λαμβάνοντας υπόψη απόρριψη ενστάσεων, ανατροπές δικαστικών αποφάσεων κι άλλους περιορισμούς.

Εκμεταλλευόμενος την αρχή αυτή, ο Frank Shepard επινόησε έναν κατάλογο που παρουσίαζε περιπτώσεις όπου συγκεκριμένες αποφάσεις παρέπεμπαν σε αντίστοιχες μεταγενέστερες υποθέσεις. Οι καταχωρήσεις έδειχναν επίσης ποιες καταστάσεις και περιοδικά ανέφεραν την αρχική απόφαση.

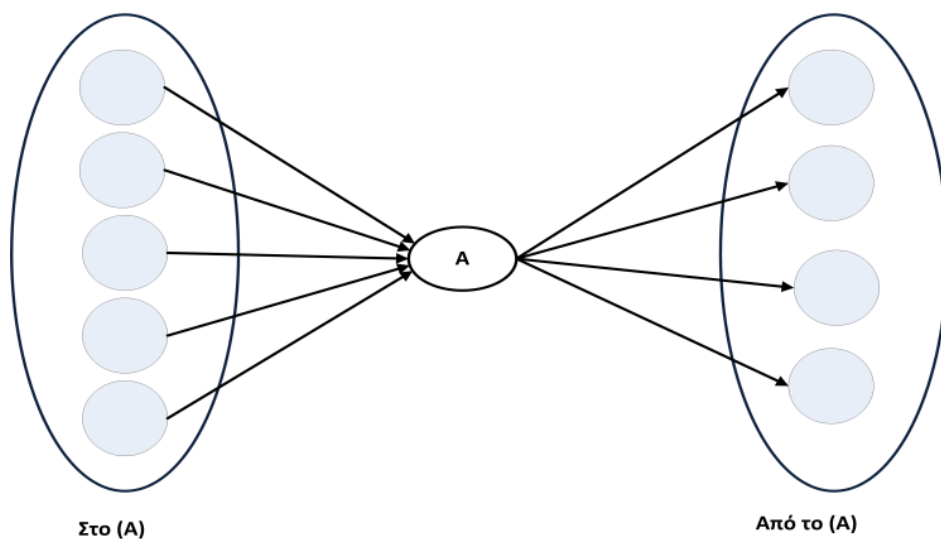
Όπως το δόγμα της εκδίκασης βάσει παλαιών αποφάσεων αποτέλεσε τη λογική και το ερέθισμα για τα συμπεράσματα του Shepard, έτσι και η αντίστοιχη σύμβαση πρόσφερε λογική εξήγηση για επιστημονικές αναφορές. Αυτή η παράδοση προϋποθέτει ότι όταν ένας συγγραφέας δημοσιεύει ένα άρθρο, θα πρέπει να παραθέτει αντίστοιχα προγενέστερα έργα τα οποία αποδείχθηκαν χρήσιμα για την προετοιμασία του δικού του έργου. Οι Price, Ziman, και Merton δίνουν έμφαση σε αυτή τη σύμβαση των βιβλιογραφικών αναφορών στο πλαίσιο της ακαδημαϊκής ανοικοδόμησης, ο Cronin παρομοίασε τις αναφορές με παγωμένα αποτυπώματα στην άμμο της ακαδημαϊκής επιτυχίας. Όπως τα αποτυπώματα μπορούν να οδηγήσουν στη σωστή κατεύθυνση, έτσι κι οι αναφορές αποτελούν ένα χρήσιμο εργαλείο για την εκτίμηση της ποιότητας και του αντίκτυπου των συγγραφέων και των τίτλων των περιοδικών εντός και μεταξύ των κλάδων. Η άποψη αυτή εκφράζεται ακόμη πιο δυναμικά μέσω της ανάλυσης του Kuhn σχετικά με την ανάπτυξη τη κανονικής επιστήμης, αλλά και μέσω του σχολίου του Newton πως «αν βλέπω μακρύτερα, είναι γιατί στέκομαι στους ώμους των γιγάντων», το οποίο έχει γίνει και τίτλος ενός βιβλίου του Merton.

Υπάρχουν αμφιβολίες, ωστόσο, για το κατά πόσο οι βιβλιογραφικές αναφορές γίνονται αποκλειστικά για τους σκοπούς της ακαδημαϊκής οικοδόμησης. Οι Weinstock και Frost έχουν παραθέσει πολλούς λόγους για τους οποίους οι συγγραφείς αναφέρονται σε προγενέστερα έργα, όπως το να αποτίσουν φόρο τιμής στους πρωτοπόρους έως και να αμφισβητήσουν ισχυρισμούς που διατυπώθηκαν πρώτοι από άλλους. Οι Kaplan, May και Davies υπαινίσσονται ότι οι συγγραφείς εκούσια αποφεύγουν τις αναφορές σε άλλες εκδόσεις και η μελέτη του Broadus έρχεται να επαληθεύσει αυτές τις υποψίες. Ο Smith πιστεύει ότι παρότι υπάρχουν λόγοι για να γίνει μία βιβλιογραφική αναφορά, υπάρχουν παράλληλα πολλοί άλλοι για τους οποίους ένας συγγραφέας, κατά τις αναφορές του, δεν παρέχει σκόπιμα τη σύνδεση με άλλα έγγραφα. Παρότι ο πιο προφανής λόγος είναι ότι ένα προγενέστερο έγγραφο δεν σχετίζεται με την παρούσα εργασία, είναι πιθανό να ισχύει κι ότι ο συγγραφέας δεν είναι γνώστης του εγγράφου ή δεν μπορεί να έχει πρόσβαση σε αυτό ή δεν μπορεί να διαβάσει τη γλώσσα στην οποία έχει γραφτεί.

Για τον λόγο αυτό, οι βιβλιογραφικές αναφορές είναι σήμερα η πηγή δεδομένων για τη διεξαγωγή πολλών ειδών βιβλιομετρικών μετρήσεων. Οι αναφορές είναι διακριτικά δεδομένα που δεν απαιτούν τη συνεργασία του ερωτώμενου και δεν προκαλούν αντίδραση.

1.7.2 Διαφορά βιβλιογραφικής αναφοράς και παραπομπής

Η αναφορά (bibliographic citation) είναι μια παράθεση σε ένα άρθρο που εκδίδεται από ένα πιο πρόσφατα δημοσιευμένο άρθρο που την δημοσιοποιεί το άρθρο που επικαλείται την αναφορά είναι το αναφερόμενο άρθρο (citing) και το άρθρο που λαμβάνει την αναφορά είναι το παραπεμπόμενο (cited) άρθρο (bibliographic reference).

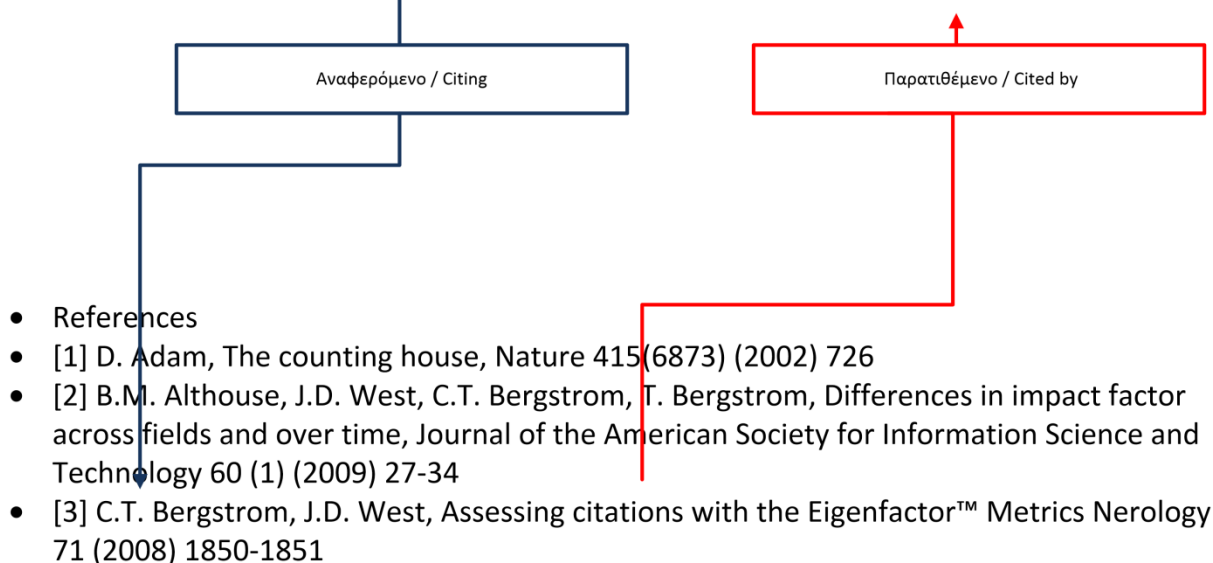


Εικόνα 1 10 Το άρθρο A δέχεται αναφορές και κάνει παραπομπές

InformationSciences

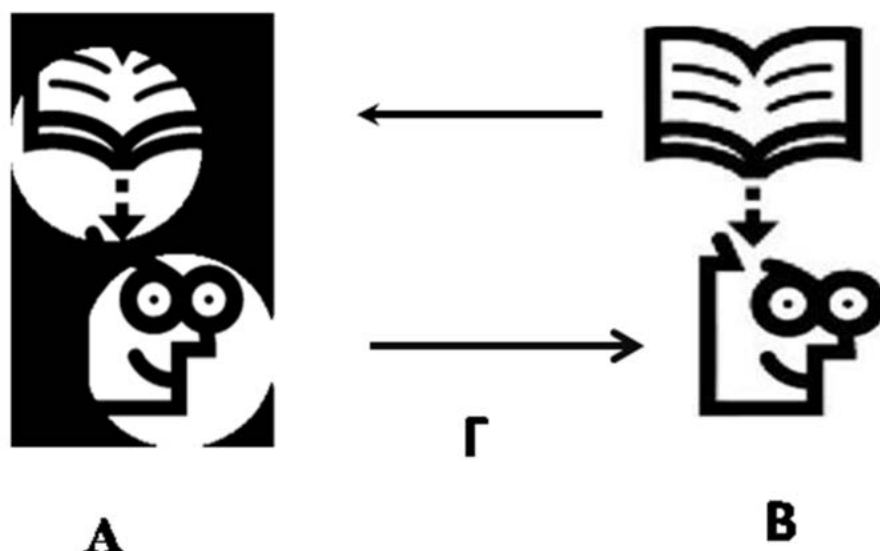
journalhomepage:www.elsevier.com/locate/ins

- A non-linear index to evaluate a journal's scientific impact
- Sozon Papavasopoulos a, Marios Poulos a, Nikolaos Korfiatis b, *, George Bokosa



Εικόνα 1 11 Σχηματική απεικόνιση άρθρου και της λίστας βιβλιογραφικών αναφορών του

Με άλλα λόγια, αν ο Α αναφέρει τον Β σε μια βιβλιογραφική αναφορά Γ, από την οπτική μεριά του Α η σχέση Γ είναι μία αναφορά ενώ από την οπτική του Β η σχέση Γ είναι μία παραπομπή.



Εικόνα 1 12 Σχηματική απεικόνιση αναφοράς-παραπομπής

Παράθεση των πηγών σε μια εργασία

Τρόποι παράθεσης

Υπάρχουν δύο τρόποι παράθεσης των πηγών σε μια εργασία:

A. Αναφορές εντός κειμένου (In-Text citations)

B. Λίστα βιβλιογραφίας / βιβλιογραφικών αναφορών

Στην πρώτη περίπτωση που χρησιμοποιείται κάποια ιδέα ή αυτούσιο απόσπασμα από δημοσιευμένο έργο οι αναφορές γίνονται:

α) Με τη μορφή υποσημειώσεων στο κάτω μέρος της σελίδας ή σημειώσεων στο τέλος της ενότητας ή του εγγράφου, με την προσθήκη ενός αύξοντα αριθμού υπό μορφή εκθέτη στο σημείο, όπου επιθυμεί ο συγγραφέας να γίνει η παραπομπή. Εμφανίζονται, όπως είπαμε, στο υποσέλιδο με μικρότερη γραμματοσειρά και ξεχωρίζουν από το υπόλοιπο κείμενο με μια διαχωριστική ευθεία γραμμή στο κάτω μέρος της σελίδας, ενώ οι σημειώσεις εμφανίζονται στο τέλος του κάθε κεφαλαίου ή στο τέλος όλου του έργου. Η αρίθμηση μπορεί να είναι συνεχής για όλα τα κεφάλαια ή να ξεκινά από την αρχή για κάθε κεφάλαιο ξεχωριστά.

β) Σε συνοπτική μορφή εντός του κειμένου, με παρένθεση στο αντίστοιχο σημείο, μέσα στην οποία αναφέρεται ο συγγραφέας ή/και η χρονολογία δημοσίευσης ή/και η σελιδαρίθμηση, ανάλογα με το πρότυπο που ακολουθείται. Οι βιβλιογραφικές πληροφορίες πρέπει να βρίσκονται με τη μορφή *συγγραφέας/χρονολογία*, προκειμένου να επιτρέπουν στον αναγνώστη να βρει τα πλήρη στοιχεία της πηγής που χρησιμοποιήθηκαν μέσα από τη λίστα της βιβλιογραφίας.

Παραδείγματα αναφορών εντός κειμένου

1. Υποσημειώσεις

Για περισσότερες πληροφορίες σχετικά με τους νόμους της Βιβλιομετρίας βλ. Σ. Παπαβλασόπουλος, Βιβλιομετρικές Εφαρμογές, Παραδόσεις, Τμήμα Αρχαιονομίας- Βιβλιοθηκονομίας- Μουσειολογίας, Ιόνιο Πανεπιστήμιο, Γ' ανατύπωση επαυξημένη και βελτιωμένη, 2013 σσ. 5-9.

2. Σημειώσεις

Στην περίπτωση ενός συγγραφέα (Παπαβλασόπουλος, 2006). Εάν χρησιμοποιήθηκαν δύο αναφορές από τον ίδιο συγγραφέα και έχουν και οι δύο δημοσιευτεί το 2008, τότε η μία θα φέρει ημερομηνία 2008a και η άλλη 2008b και θα εμφανίζονται ομοίως και στην αναφορά. Σε ελληνική βιβλιογραφία οι αναφορές θα φέρουν ημερομηνίες 1998a και 1998b αντίστοιχα.

Εάν υπάρχουν δύο συγγραφείς για ένα συγκεκριμένο τεκμήριο, τότε αναφέρονται τα ονόματα με τη σειρά με την οποία εμφανίζονται. Παράδειγμα: (Papavlasopoulos & Poulos, 2012).

Εάν στην χρησιμοποιούμενη πηγή υπάρχουν περισσότεροι από δύο συγγραφείς, τότε χρησιμοποιείται η συντομογραφία et al./κ.ά. Παράδειγμα: (Papavlasopoulos et al., 2007) ή (Παπαβλασόπουλος κ.ά., 2007)

Οι αριθμοί των σελίδων μπορούν ή και όχι να συμπεριληφθούν ανάλογα με την ιδιομορφία της αναφοράς στην πηγή για να υποδείξουν μια συγκεκριμένη σελίδα ή μια ακολουθία σελίδων. Παράδειγμα: (Papavlasopoulos et al., 2013, p.55) ή (Παπαβλασόπουλος κ.ά. 2004, σσ.16-19).

Σε ηλεκτρονική πηγή, στην οποία δεν υπάρχει αρίθμηση σελίδων, μπορεί να χρησιμοποιηθεί ο αριθμός μιας παραγράφου προκειμένου να υποδεικνύεται το μέρος του κειμένου στο οποίο αναφέρεται ο συγγραφέας.

Όταν αναφέρονται δύο ή περισσότερα κείμενα με διαφορετικούς συγγραφείς ξεχωρίζονται με ερωτηματικό(;). Παράδειγμα: (Παπαβλασόπουλος 2004; Πούλος 2001)

3. Άμεσης αναφοράς

Χρησιμοποιούνται διπλά εισαγωγικά σημεία («») για να εσωκλείνονται επιπλέον λόγια ενός συγγραφέα. Η θέση του χρησιμοποιούμενου τεκμηρίου (αριθμοί σελίδας ή αριθμοί παραγράφου) πρέπει οπωσδήποτε να αναφερθεί. Εάν η άμεση παραπομπή στην πηγή είναι περισσότερες από 40 λέξεις, παρατίθεται το κείμενο της παραπομπής σε εσοχή παραγράφου, χωρίς τη χρήση των εισαγωγικών σημείων («»). Παράδειγμα: Σύμφωνα με τον Παπαβλασόπουλο (2003), «υπάρχει και εναλλακτική μορφή του ανωτέρω νόμου» (σ.17).

4. Έμμεσης αναφοράς/παράφρασης:

Εάν παραφράζονται πορίσματα της έρευνας ή ιδέες ενός άλλου συγγραφέα, τότε καλό θα ήταν να εντάσσονται σαν μέρος του κειμένου αλλά με λόγια του προσώπου που γράφει την εργασία. Όταν παραφράζεται ή αναφέρεται μια ιδέα η οποία εμπεριέχεται σε μια άλλη εργασία, δεν είναι υποχρεωτικό πάντα να αναφέρεται η θέση της πηγής (αριθμός σελίδας), αν και θεωρείται χρήσιμο. Θα πρέπει ωστόσο να φαίνονται ξεκάθαρα στο κείμενο οι ιδέες του συγγραφέα και οι ιδέες άλλων προσώπων.

Παράδειγμα:

Το «λογισμικό των windows» δεν είναι το καταλληλότερο σε αυτήν την περίπτωση (Πούλος, 2009).

Παρά το γεγονός ότι ο Παπαβλασόπουλος (Παπαβλασόπουλος, 2003) ορίζει τον νόμο του Zipf με την εξίσωση 2, πιστεύω πως ...

5. Αναφοράς σε έργο χωρίς συγγραφέα ή ανώνυμο συγγραφέα:

Όταν ένα έργο δεν έχει όνομα συγγραφέα ή πρόκειται για ανώνυμο συγγραφέα, η αναφορά (in-text citation) αποτελείται από τις πρώτες λέξεις του τίτλου (σε πλάγια γραμματοσειρά), ακολουθούμενες από τη χρονολογία και την αρίθμηση της σελίδας.

Παράδειγμα:

Η λογοκλοπή είναι μια πράξη η οποία προσβάλλει το δικαίωμα της πνευματικής ιδιοκτησίας (Ελληνικά Γράμματα 1998, σελ.74).

Στη δεύτερη περίπτωση της λίστας οι αναφορές καταγράφονται και στο τέλος της εργασίας στην αναλυτική τους μορφή με κεφαλίδα Βιβλιογραφικές Αναφορές ή Βιβλιογραφία (References/Bibliography). Στη λίστα βιβλιογραφικών αναφορών παρατίθενται με αλφαβητική σειρά (κατά όνομα συγγραφέα) μόνο οι πηγές που χρησιμοποιήθηκαν στη συγγραφή της εργασίας, ενώ στη βιβλιογραφία δεν παρατίθενται μόνο οι χρησιμοποιούμενες πηγές, αλλά και οι πηγές οι οποίες, αν και μελετήθηκαν κατά τη διάρκεια της συγγραφής της εργασίας, τελικά δεν χρησιμοποιήθηκαν. Για τη σύνταξη βιβλιογραφικών αναφορών στην αναλυτική τους μορφή μπορείτε να χρησιμοποιήσετε κάποιο στυλ (πρότυπο) αναφοράς από τα κατά καιρούς αναπτυσσόμενα διάφορα διεθνή πρότυπα, τα οποία διαφέρουν μεταξύ τους ως προς τον τρόπο που δομείται και εμφανίζεται η βιβλιογραφική πληροφορία.

Στυλ βιβλιογραφικών αναφορών

Έχουν αναπτυχθεί διάφορα στυλ μορφοποίησης βιβλιογραφικών αναφορών, τα οποία περιλαμβάνουν τα ίδια βιβλιογραφικά στοιχεία, όπως όνομα συγγραφέα, τίτλος, τόπος έκδοσης, έτος δημοσίευσης και εκδότης. Ανάλογα, τώρα, με το τεκμήριο χρησιμοποιούνται και επιπρόσθετα στοιχεία όπως:

- Για βιβλία, δίνεται το όνομα του εκδότη και ο τόπος δημοσίευσης. Εάν δίνονται δύο ή περισσότεροι τόποι δημοσίευσης, προτείνεται αυτός που εμφανίζεται πρώτος στο βιβλίο.
- Για περιοδικά, δίνεται ο αριθμός του τόμου, του τεύχους και ο αριθμός της σελίδας.
- Για ιστοσελίδες, δίνεται η πλήρης ηλεκτρονική διεύθυνση (URL, αν και καλό είναι να αποφεύγεται η χρήση ιστοσελίδων ως κύριων πηγών, γιατί το περιεχόμενο τους αλλάζει

σχετικά γρήγορα. Οι ιστοσελίδες χρησιμοποιούνται κυρίως για εύρεση αναφορών για μια πρώτη σχετική βιβλιογραφία.

Παρότι τα στυλ μορφοποίησης περιλαμβάνουν τα ίδια στοιχεία, διαφέρουν μεταξύ τους ως προς τον τρόπο που δομείται και εμφανίζεται η πληροφορία, γιατί το καθένα έχει τις δικές του ξεχωριστές απαιτήσεις. Για τον λόγο αυτό, πρέπει να γίνεται χρήση ενός μόνο στυλ βιβλιογραφικών αναφορών σε ένα τεκμήριο τόσο για τη λίστα βιβλιογραφικών αναφορών, όσο και για τη βιβλιογραφία. Για τη σύνταξη βιβλιογραφικών αναφορών στην αναλυτική τους μορφή μπορείτε να χρησιμοποιήσετε τα πρότυπα σύνταξης βιβλιογραφικών αναφορών. Τα πλέον διαδεδομένα είναι: APA (American Psychological Association), Numeric, Harvard, MLA (Modern Language Association).

1.8. Διπλώματα ευρεσιτεχνίας (Patents)

Ένα δίπλωμα ευρεσιτεχνίας είναι ένα σύνολο χρονικά περιορισμένων αποκλειστικών δικαιωμάτων σε μια εφεύρεση, που συνήθως χορηγείται από το γραφείο διπλωμάτων ευρεσιτεχνίας της κυβέρνησης. Ο όρος ευρεσιτεχνία χρησιμοποιείται, επίσης, για τις επίσημα εγγεγραμμένες περιγραφές εφευρέσεων. Αυτά τα έγγραφα είναι παρόμοια από ορισμένες απόψεις με τα επιστημονικά άρθρα για παράδειγμα, μέσω της συμπερίληψης ενός συνόλου αναφορών. Η αξία των διπλωμάτων ευρεσιτεχνίας ως δείκτης της επιστημονικής αξίας πηγάζει από την αναγνώριση ότι οι ακαδημαϊκοί ερευνητές μπορεί να είναι, και ίσως μερικές φορές θα πρέπει να είναι, άμεσα εμπλεκόμενοι στην ανάπτυξη χρήσιμων τεχνολογιών.

Οι δείκτες ευρεσιτεχνίας μπορεί να είναι άμεσοι, με την έννοια της καταμέτρησης των βραβείων ευρεσιτεχνίας των ερευνητών, που ανταμείβουν εκείνους που προωθούν νέα έρευνα με δυνητική εμπορική αξία. Οι δείκτες ευρεσιτεχνίας μπορεί επίσης να είναι έμμεσοι, χρησιμοποιώντας τις αναφορές διπλωμάτων ευρεσιτεχνίας, για να εντοπίσουν την παραπεφθείσα ακαδημαϊκή δουλειά, η οποία ως εκ τούτου αναγνωρίζεται πως έχει εφαρμόσιμη αξία.

Οι αναλύσεις ευρεσιτεχνίας έχουν επίσης χρησιμοποιηθεί για να αξιολογήσουν την απόδοση της τεχνολογίας μιας χώρας και για να εντοπίσουν ροές της μεταφοράς γνώσεων ανάμεσα στην επιστήμη και την τεχνολογία. Για παράδειγμα μία εμπειρική μελέτη των διπλωμάτων ευρεσιτεχνίας σχετική με την Ολλανδία κατέληξε στο συμπέρασμα ότι τα αποτελέσματα δεν ταίριαζαν με τα υπάρχοντα / ισχύοντα θεωρητικά μοντέλα των σχέσεων μεταξύ πανεπιστημίων – βιομηχανίας τα οποία, ως εκ τούτου, έπρεπε να αναθεωρηθούν.

Βιβλιογραφία/Αναφορές

- Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the World Wide Web: Methodological approaches to 'webometrics'. *Journal of Documentation*, 53(4), 404–426.
- Andersen, H. (2000). Influence and reputation in the social sciences-how much do researchers agree? *Journal of Documentation*, 56(6), 674–692.
- Archambault, É., & Gagné, É. V. (2004). The use of bibliometrics in the social sciences and humanities. Montreal: Science-Metrix.
- Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.
- Bonitz, M. (1982). *Scientometrie, Bibliometrie, Infor-metrie*. Zentralblatt für Bibliothekswesen, 96(1), 19-24.
- Broadus, R. N. (1987). Early approaches to bibliometrics. *Journal of the American Society for Information Science (1986-1998)*, 38(2), 127.
- Brookes, B. C. (1990). Biblio-, sciento-, infor-metrics?? what are we talking about? Retrieved from <https://uhdspace.uhasselt.be/dspace/handle/1942/857>
- Cattell, J. M. (1906). A statistical study of american men of science. III. The distribution of american men of science. *Science*, 732–742.
- Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis.
- Clemens, E. S., Powell, W. W., McIlwaine, K., & Okamoto, D. (1995). Careers in print: Books, journals, and scholarly reputations. *American Journal of Sociology*, 433–494.
- Cole, S. (2004). Merton's Contribution to the Sociology of Science. *Social Studies of Science*, 34(6), 829–844.
- Comité national d'évaluation de la recherche (CNER). (2002). *Évaluation de la recherche publique dans les établissements publics français*. Paris.
- Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. *Journal of Documentation*, 53(3), 263–273.
- Crane, D. (1967). The gatekeepers of science: Some factors affecting the selection of articles for scientific journals. *The American Sociologist*, 195–201.
- Egghe, L. (2000). New informetric aspects of the Internet: some reflections-many problems. *Journal of Information Science*, 26(5), 329–335.
- Egghe, L., & Rousseau, R. (1990). Introduction to informetrics: Quantitative methods in library, documentation and information science. Retrieved from <http://eprints.rclis.org/handle/10760/6011>
- Elkana, Y., & others. (1978). *Toward a metric of science*. Wiley. Retrieved from <http://agris.fao.org/agris-search/search.do?recordID=US201300538654>
- Garfield, E. (1986). THE 250 MOST-CITED AUTHORS IN THE ARTS-AND-HUMANITIES CITATION INDEX, 1976-1983. *Current Contents*, (48), 3–10.
- Glänzel, W. (1996). A bibliometric approach to social sciences. National research performances in 6 selected social science areas, 1990–1992. *Scientometrics*, 35(3), 291–307.
- Glänzel, W., & Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing & Management*, 35(1), 31–44.
- Gingras, Y. (1984). La valeur d'une langue dans un champ scientifique. *Recherches Sociographiques*, 25(2), 286–296.

- Grayson, L., & Gomersall, A. (2003). A difficult business: finding the evidence for social science reviews. *ESRC UK Centre for Evidence Based Policy and Practice. Working Paper, 19*. Retrieved from <http://www.kcl.ac.uk/sspp/departments/politiceconomy/research/cep/pubs/papers/assets/wp19.pdf>
- Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44(2), 193–215.
- Hicks, D. (2005). The four literatures of social science. In *Handbook of quantitative science and technology research* (pp. 473–496). Springer. Retrieved from http://link.springer.com/chapter/10.1007/1-4020-2755-9_22
- Hodgson, G. M., & Rothman, H. (1999). The editors and authors of economics journals: A case of institutional oligopoly? *The Economic Journal*, 109(453), 165–186.
- Ingwersen, P., & Christensen, F. H. (1997). Data Set Isolation for Bibliometric Online Analyses of Research Publications: Fundamental Methodological Issues. *JASIS*, 48(3), 205–217.
- Ingwersen, P. E. R. (1997). *The central international visibility of Danish and Scandinavian research 1988-1996*. Centre for Informetric Studies, Royal School of Librarianship.
- Ingwersen, P. E. R. (1997). *The central international visibility of Danish and Scandinavian research 1988-1996*. Centre for Informetric Studies, Royal School of Librarianship.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236–243.
- Larson, R. R. (1996). Bibliometrics of the World Wide Web: An exploratory analysis of the intellectual structure of cyberspace. In *Proceedings of the Annual Meeting-American Society for Information Science* (Vol. 33, pp. 71–78). Retrieved from <http://free-journal.umm.ac.id/files/file/asis96.ps>
- Larivière V. and Godin B. (2001). *Profil de la production scientifique du Service canadien des forêts*. Report prepared for Canadian Forest Service.
- Lewison, G. (2001). Evaluation of books as research outputs in history of medicine. *Research Evaluation*, 10(2), 89–95.
- Leydesdorff, L. (2003). Can networks of journal-journal citations be used as indicators of change in the social sciences? *Journal of Documentation*, 59(1), 84–104.
- López Yepes, J. (1995). La documentación como disciplina: teoría e historia. actual. y ampli. *Panplona: EUNSA*.
- Li, X. (2003). A review of the development and application of the Web impact factor. *Online Information Review*, 27(6), 407–417.
- Li, X., Thelwall, M., Musgrove, P., & Wilkinson, D. (2003). The relationship between the WIFs or inlinks of Computer Science Departments in UK and their RAE ratings or research productivities in 2001. *Scientometrics*, 57(2), 239–255.
- Line, M. B. (1981). The structure of social science literature as shown by a large-scale citation analysis. *Social Science Information Studies*, 1(2), 67–87.
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of Washington Academy Sciences*. Retrieved from <http://doi.apa.org/?uid=1926-10141-003>
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press. Retrieved from <https://www.google.com/books?hl=el&lr=&id=zPvcHuUMEMwC&oi=fnd&pg=PR9&dq=The+Sociology+of+Science:+Theoretical+and+Empirical+Investigations&ots=x5QIQmbavM&sig=hfWufcJoat-4WNmuCGe6Dh2hDz0>
- Moed, H. F., Luwel, M., & Nederhof, A. J. (2002). Towards research performance in the humanities. *Library Trends*, 50(3), 498–520.
- Nacke, O. (1979). Informetrie: Ein neuer Name für eine neue Disziplin. *Nachrichten Für Dokumentation*, 30(6), 219–226.

- Nalimov, V. V. (1970). Influence of mathematic statistics and cybernetics on the methodology of scientific investigations. *Zavodskaya Laboratoriya*. 36 (10). *English Translation in Industrial Laboratory*, 36(10), 1549–1558.
- NALIMOV, V. V., I. V. KORDON, A. YA. KORNEEVA (1971), *Geograficheskoe Raspredelenie Nauchnoi Informatsii*. [Geographic Distribution of Scientific Information.] Informatsionnye Materialy. Moscow: an SSSR Nauchnyi Sovet po Kompleksnoi Probleme Kibernetiki. [Informational Papers. Moscow: Soviet Academy of Science, Scientific Council on Cybernetics. 2 : 3–37.
- Nalimov, V. V., & Mulczenko, Z. M. (1969). *Naoukometria [Scientometrics. The study of science as an information process]*. Moscow: Nauka
- Nederhof, A. J., & Zwaan, R. A. (1991). Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. *Journal of the American Society for Information Science*, 42(5), 332–340.
- Nederhof, A. J., Zwaan, R. A., De Bruin, R. E., & Dekker, P. J. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioral sciences: A comparative study. *Scientometrics*, 15(5-6), 423–435.
- Nederhof, A., Luwel, M., & Moed, H. (2001). Assessing the quality of scholarly journals in linguistics: An alternative to citation-based journal impact factors. *Scientometrics*, 51(1), 241–265.
- Nederhof, A. J., & Van Raan, A. F. (1993). A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy*, 22(4), 353–368.
- Fernberger, S. W. (1917). On the number of articles of psychological interest published in the different languages. *The American Journal of Psychology*, 141–150.
- Gross, P. L., & Gross, E. M. (2007). College libraries and chemical education Science. 1927; 66 (1713): 385-389. *Smith DR. Historical Development of the Journal Impact Factor and Its Relevance for Occupational Health. Ind Health*, 45(6), 730–42.
- Nederhof, A. J., & Noyons, E. C. (1992). International comparison of departments' research performance in the humanities. *Journal of the American Society for Information Science*, 43(3), 249–256.
- Nederhof, A. J., & Zwaan, R. A. (1991). Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. *Journal of the American Society for Information Science*, 42(5), 332–340.
- Nordstrom, L. (1990). 'Bradford's law' and the relationship between ecology and biogeography. *Scientometrics*, 18(3-4), 193–Kostagiolas, P., Papadaki, E., Kanlis, G., & Papavasopoulou, S. (2013). Responding to crises with alliances: evidence from an academic library survey in Greece. *Advances in Librarianship*, 36, 247–279.
- Bokos, M. P. G., Papavasopoulou, N. K. S., & Avlonitis, M. (2007). Specific selection of FFT amplitudes from audio sports and news broadcasting for classification purposes. Retrieved from <http://www.emis.ams.org/journals/JGAA/accepted/2007/Poulos+2007.11.1.pdf>
- Papavasopoulou, S., & Poulos, M. (2012). Neural network design and evaluation for classifying library indicators using personal opinion of expert. *Library Management*, 33(4/5), 261–271.
- Pareto, V., & Page, A. N. (1971). Translation of *Manuale di economia politica* ('Manual of political economy'). *AM Kelley*.
- Pirolli, P., Pitkow, J., & Rao, R. (1996). Silk from a sow's ear: extracting usable structures from the Web. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 118–125). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=238450>

- Price, D.J. de Solla (1976). *A general theory of bibliometric and other cumulative advantage processes*. Journal of the American Society for Information Science 27 (5), 292–306. doi:10.1002/asi.4630270505
- Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, (25), 348–349.
- Restivo, S. P. (1994). *Science, society, and values: Toward a sociology of objectivity*. Lehigh University Press. Retrieved from https://www.google.com/books?hl=el&lr=&id=h0z5zdPAR0gC&oi=fnd&pg=PR9&dq=Toward+a+Sociology+of+Objectivity&ots=AGBJKUR--t&sig=ek_adOiOf0XtgGKw-uPS65F850c
- Rousseau, R. (1997). *Sitations: an Exploratory Study*. *Cybermetrics*, [Online], 1 (1).
- Royle, P., & Over, R. (1994). The use of bibliometric indicators to measure the research productivity of Australian academics. *Australian Academic & Research Libraries*, 25(2), 77–88.
- Shapiro, F. R. (1992). Origins of bibliometrics, citation indexing, and citation analysis: the neglected legal literature. *Journal of the American Society for Information Science*, 43(5), 337–339.
- Schoepflin, U. (1992). Problems of representativity in the social sciences citation index. *Representations of Science and Technology*. Leiden: DSWO, 177–188.
- Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing & Management*, 28(1), 1–3.
- Thelwall, M. (2001). Extracting macroscopic information from web links. *Journal of the American Society for Information Science and Technology*, 52(13), 1157–1168.
- Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1–116.
- Thelwall, M., & Harries, G. (2004). Do the web sites of higher rated scholars have significantly more online impact? *Journal of the American Society for Information Science and Technology*, 55(2), 149–159.
- van Raan, A. F. (1998). *Special Topic Issue: Science and Technology Indicators*. American Society for Information Science (ASIS).
- van Raan, A. F. (2001). Bibliometrics and Internet: Some observations and expectations. *Scientometrics*, 50(1), 59–63.
- Rubio, A. V. (1992). Scientific production of Spanish universities in the fields of social sciences and language. *Scientometrics*, 24(1), 3–19.
- Webster, B. M. (1998). Polish sociology citation index as an example of usage of national citation indexes in scientometric analysis of social sciences. *Journal of Information Science*, 24(1), 19–32.
- Winclawska, B. M. (1996). Polish sociology citation index (principles for creation and the first results). *Scientometrics*, 35(3), 387–391.
- Weinberg, B. H. (1997). The earliest Hebrew citation indexes. *Journal of the American Society for Information Science*, 48(4), 318–330.
- White, H. D., & McCain, K. W. (1989). Bibliometrics. *Annual Review of Information Science and Technology*, 24, 119–186.
- Wilson, C. S. (1999). Informetrics. *Annual Review of Information Science and Technology (ARIST)*, 34, 107–247.
- Yitzhaki, M. (1998). The ‘language preference’ in sociology: Measures of ‘language self-citation’, ‘relative own-language preference indicator’, and ‘mutual use of languages’. *Scientometrics*, 41(1-2), 243–254.
- Zipf, G. K. (1949). Human behavior and the principle of least effort. Retrieved from <http://psycnet.apa.org/psycinfo/1950-00412-000>
- Μπουντουρίδης, Μουσιός Α. (1999). <http://hyperion.math.upatras.gr/courses/sts/>, Πανεπιστημίου Πατρών

Σίτας, Α. Κώνστα, Ο. Ντούλια, Π. (2007). *Βιβλιομετρική-Πληροφοριομετρική-Ιστομετρική*. Δημοσιευμένο στο: Μιχάλης Τζεκάκης: ο οραματιστής. Ένωση Ελλήνων Βιβλιοθηκονόμων και Επιστημόνων της Πληροφόρησης, Χεροχ, 44-63.

Κεφάλαιο 2

Εργαλεία καταγραφής επιστημονικών αναφορών και μέτρησης

Σύνοψη

Η Βιβλιομετρία είναι ένα σύνολο μεθόδων για την ποσοτική ανάλυση των επιστημονικών και τεχνολογικών βιβλιογραφιών. Οι, συνήθως, χρησιμοποιούμενες βιβλιομετρικές μέθοδοι είναι η ανάλυση αναφορών και η ανάλυση περιεχομένου. Σε αυτό το κεφάλαιο αναλύεται η έννοια της βιβλιομετρικής ανάλυσης και της ανάλυσης περιεχομένου και απαντώνται ερωτήματα για το πώς και πού αποκτούμε τις αναφορές, καθώς και για το ποια δεδομένα χρησιμοποιούμε για την παραγωγή δεικτών. Επίσης γίνεται προσπάθεια να αναδειχθεί ότι, μέσω των βιβλιομετρικών αναλύσεων, οδηγούμεθα σε δείκτες ποσοτικής έρευνας και απόδοσης. Στόχος είναι η παράθεση των πιο σημαντικών βιβλιομετρικών δεικτών με τους ορισμούς τους, ορισμένες παρατηρήσεις σχετικές με τα πλεονεκτήματα και τις αδυναμίες τους, καθώς και την κατασκευή και χρήση τους. Επίσης η επιλογή της κατάλληλης βάσης δεδομένων και μέσω αυτής η εκμάθηση, βήμα προς βήμα, της ανάκτησης, αποθήκευσης και διαχείρισης των βιβλιογραφικών δεδομένων για βιβλιομετρικές αναλύσεις. Θα γίνει εκτενής αναφορά με παραδείγματα και μελέτες περίπτωσης από τις βάσεις: Web of Science, Scopus, Google Scholar. Δίδονται, δηλαδή, τα εργαλεία αναφορών και μέτρησης και τα τεχνικά ζητήματα που προκύπτουν με τα δεδομένα. Το ερευνητικό έργο των ακαδημαϊκών μελών, αλλά και των ιδρυμάτων και των χωρών κατ'επέκταση μπορεί εύκολα να ποσοτικοποιηθεί και να μετατραπεί σε μετρήσιμες μονάδες, μέσω βιβλιομετρικής ανάλυσης. Όπου βιβλιομετρική ανάλυση νοείται η καταγραφή και επεξεργασία δεδομένων που σχετίζονται με τις επιστημονικές δημοσιεύσεις και την εξαγωγή των αντίστοιχων βιβλιομετρικών δεικτών.

2.1. Επιλέγοντας μια μονάδα ανάλυσης

Το σημείο εκκίνησης σε μια βιβλιομετρική ανάλυση είναι να επιλέξουμε μια ομάδα δημοσιεύσεων, συνήθως βάσει των πληροφοριών που διατίθενται στους δείκτες αναφορών της Thomson ή άλλων βάσεων. Αυτή η επιλογή των δημοσιεύσεων αποτελεί τη μονάδα ανάλυσης. Οι δημοσιεύσεις μπορούν, για παράδειγμα, να επιλεγούν με βάση τις οργανωτικές, ιδρυματικές συνεργασίες των συγγραφέων και μπορεί θεωρητικά να είναι:

- συγγραφέας,
- ερευνητική Ομάδα,
- τμήμα,
- ερευνητικά Κέντρα / Δίκτυα,
- πανεπιστήμια,
- χώρες.

Ένα σημαντικό μέρος προετοιμασίας και επαλήθευσης των τοπικών δεδομένων είναι αναγκαίο, προκειμένου να δημιουργηθεί μια μονάδα ανάλυσης που βασίζεται σε ερευνητική ομάδα, πανεπιστημιακό τμήμα, ερευνητικό κέντρο ή ερευνητικό δίκτυο. Αυτές οι πληροφορίες είναι πολύ δύσκολο να εντοπιστούν στα δεδομένα από τους δείκτες αναφορών της Thomson και μπορεί, σε πολλές περιπτώσεις, να μην υπάρχουν καθόλου. Είναι ακόμα δύσκολο να αποδώσουμε τις δημοσιεύσεις σε ένα συγκεκριμένο πανεπιστήμιο, αφού και τα ονόματα συμμετοχής και οι διευθύνσεις τους μπορεί να γραφτούν με πολλούς διαφορετικούς τρόπους και δύο διαφορετικά πανεπιστήμια κατά καιρούς μοιράζονται ένα κοινό όνομα.

Μια μονάδα ανάλυσης μπορεί, επίσης, να επιλέγεται με βάση τις ιδιότητες των επιμέρους άρθρων (αντί των συγγραφέων ή το ίδρυμα του συγγραφέα). Για μεμονωμένες δημοσιεύσεις π.χ.:

περιοδικό,

θέμα το οποίο συχνά βασίζεται στη θεματική ταξινόμηση του περιοδικού,

τύπος εγγράφου: άρθρο, επισκόπηση, σημείωμα, επιστολή, πρακτικά συνεδρίων, κ.λ.π.,

έτος έκδοσης.

Τα αποτελέσματα των στατιστικών μεθόδων που χρησιμοποιούνται σε βιβλιομετρική έρευνα βελτιώνονται με μεγαλύτερες μονάδες ανάλυσης. Αυτό συμβαίνει εν μέρει, επειδή τα μεμονωμένα φαινόμενα - όπως είναι οι αρνητικές αναφορές - ακυρώνονται από τη μεγάλη ποσότητα των αντικειμένων (Moed, H. F. 2005, σ. 80). Οι βιβλιομετρικοί δείκτες που βασίζονται σε οποιαδήποτε μονάδα ανάλυσης, η οποία περιέχει

λιγότερο από 10 άρθρα (δηλαδή αφορά σε ένα μεμονωμένο ερευνητή ή αντικείμενο), δεν πρέπει να συνιστάται. Είναι, επίσης, αναγκαίο να ληφθεί υπόψη κάθε δυνατότητα μιας συστηματικής μεροληψίας. Αυτό θα μπορούσε, για παράδειγμα, να είναι διαφορετικοί λόγοι αναφορών των εργασιών λόγω «παραδόσεων» ή των υφισταμένων συμβάσεων για την ένταξη και την κατάταξη συγγραφέων που διαφέρουν σημαντικά μεταξύ των διαφόρων ερευνητικών περιοχών (Moed, H. F. 2005, σ. 223).

2.2. Παράγοντες Επηρεασμού της Βιβλιομετρικής Ανάλυσης

Η βιβλιομετρική ανάλυση όπως έχουμε προείπει ασχολείται με την επεξεργασία των δεδομένων και την καταγραφή αυτών που έχουν σχέση με τις δημοσιεύσεις, όπως ο αριθμός των αναφορών σε αυτές από άλλες δημοσιεύσεις ο αριθμός των δημοσιεύσεων, η κατανομή τους ανά χώρα, επιστημονικό πεδίο, συγγραφέα, φορέα, κ.ο.κ.. Από τα στοιχεία των δημοσιεύσεων είναι εφικτή η εύρεση των τάσεων της ερευνητικής παραγωγής καθώς και των χαρακτηριστικών της σε επίπεδο οργανισμού, χώρας ή χωρών, καθώς επίσης ο προσδιορισμός της απήχησης και της αξιολόγησης του ερευνητικού έργου, και ο προσδιορισμός κοινωνικών δικτύων μεταξύ των ερευνητών και των επιστημονικών πεδίων. Οι παράγοντες επηρεασμού των βιβλιομετρικών αναλύσεων είναι οι κάτωθι:

2.2.1. Επιστημονικά πεδία

Στα επιστημονικά πεδία παρατηρούνται πάρα πολλές διαφορές, σε τέτοιο βαθμό που η σύγκριση δεικτών είναι προβληματική και όχι μόνο σε αυτούς αλλά και μεταξύ διαφορετικών ερευνητικών τομέων, ακόμα δε και μεταξύ επιμέρους περιοχών του ίδιου ερευνητικού πεδίου όσον αφορά στα μοτίβα αναφορών των δημοσιεύσεων στο χρόνο απαξίωσης των ερευνητικών αποτελεσμάτων κ.λ.π.. Για παράδειγμα, στην έρευνα για τον καρκίνο και το Aids παρατηρείται υψηλότερη παραγωγικότητα ως προς την ποσότητα των επιστημονικών άρθρων που δημοσιεύονται ανά έτος.

Κατά κανόνα, ο αριθμός των αναφορών αυτών των δημοσιεύσεων φτάνει στη μέγιστη τιμή του σε σχετικά μικρό χρονικό διάστημα από την ημερομηνία δημοσίευσής τους. Σε αντίθεση, στις κοινωνικές επιστήμες οι ρυθμοί δημοσίευσης είναι πολύ πιο αργοί, ενώ καταγράφονται αναφορές για μεγάλο χρονικό διάστημα μετά τη δημοσίευσή τους. Στις ανθρωπιστικές επιστήμες, όπως έχουμε αναφέρει εκτενέστατα στο προηγούμενο κεφάλαιο, το μεγαλύτερο μέρος των επιστημονικών δημοσιεύσεων είναι βιβλία, μονογραφίες και άρθρα που δημοσιεύονται σε επιστημονικά περιοδικά εθνικής εμβέλειας. Σε άλλα επιστημονικά πεδία - όπως ορισμένοι κλάδοι της πληροφορικής- είναι συχνότερη η δημοσίευση των σημαντικών ερευνητικών αποτελεσμάτων σε συνέδρια, χωρίς να ακολουθεί απαραίτητα αντίστοιχη δημοσίευση σε επιστημονικά περιοδικά. Και στις δύο περιπτώσεις, δημοσιεύσεις υψηλής επιστημονικής αξίας, όπως και μέρος των αναφορών τους, συνήθως δεν υπολογίζονται στις βάσεις δεδομένων, έτσι ώστε οι βιβλιομετρικοί δείκτες των αντίστοιχων ερευνητικών περιοχών να μην αποτυπώνουν πλήρως την πραγματική κατάσταση.

2.2.2. Χρονικό διάστημα ανάλυσης αναφορών

Είναι προφανές ότι το πλήθος των αναφορών σε μια ερευνητική εργασία θα σχετίζεται με το χρονικό διάστημα που έχει διανυθεί μετά την αρχική δημοσίευσή της. Είναι επίσης προφανές ότι οι παλαιότερες δημοσιεύσεις δυνητικά έχουν τις περισσότερες αναφορές, χωρίς τούτο να συνοδεύεται πάντα με την απήχηση που αυτές έχουν στην επιστημονική κοινότητα. Για την εξομάλυνση των διαφορών που οφείλονται αποκλειστικά στη φυσιολογική αύξηση του αριθμού των αναφορών σε παλαιότερες δημοσιεύσεις, ορίζονται συγκεκριμένα χρονικά διαστήματα για τη μέτρηση των αναφορών σε μία δημοσίευση (Simkin M.V. & Roychowdhury V.P., 2012).

2.2.3. Είδος επιστημονικών δημοσιεύσεων

Το πλήθος των αναφορών σχετίζεται, επίσης, και από το είδος της δημοσίευσης. Όπως ήδη έχει αναφερθεί τα άρθρα ανασκόπησης έχουν αρκετά περισσότερες αναφορές από τις απλές ερευνητικές εργασίες. Ως εκ τούτου είναι αναγκαίος ο προσδιορισμός εναλλακτικών δεικτών, ανάλογα με το είδος της δημοσίευσης, σε περιπτώσεις αξιολόγησης μεμονωμένων ερευνητών ή ερευνητικών ομάδων.

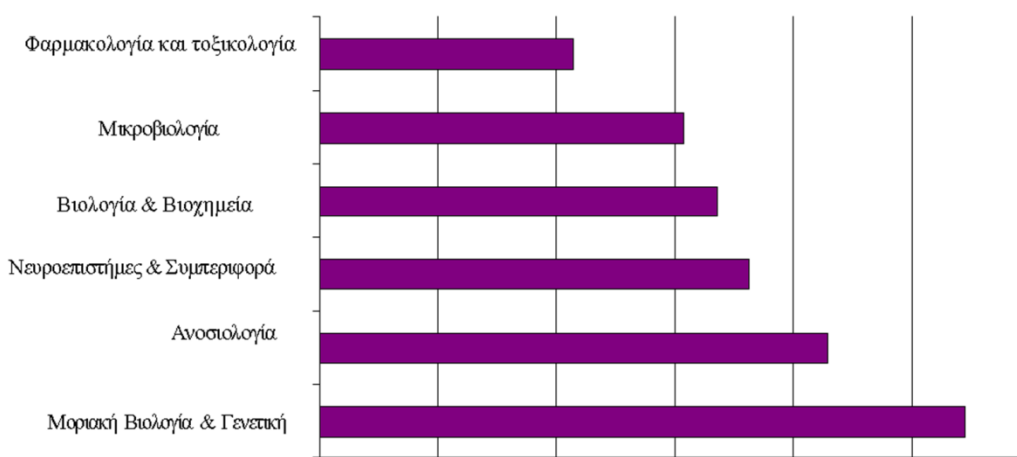
Για τους παραπάνω λόγους είναι φανερό ότι η αξιολόγηση του επιστημονικού έργου δεν πρέπει να βασίζεται μόνο στη χρησιμοποίηση των βιβλιομετρικών δεικτών. Δεν πρέπει να παραγνωρίζεται το γεγονός όμως πως οι βιβλιομετρικοί δείκτες δίδουν, σημαντικές πληροφορίες για πλείστα θέματα όπως την καταγραφή των επιστημονικών ομάδων ενός οργανισμού ή χώρας και του ερευνητικού έργου που παρέχουν, τα επιστημονικά πεδία που αυτές δραστηριοποιούνται, τον βαθμό συνεργασίας που αναπτύσσεται και την απήχηση που έχουν διεθνώς στην επιστημονική κοινότητα.

2.3. Πώς επηρεάζονται οι αναφορές

Μια περιεκτική περίληψη της έρευνας σχετικά με το γιατί οι συγγραφείς αναφέρουν ο ένας τον άλλον έχει γραφτεί από τον Henk Moed (2005). Έχει βρεθεί και είναι λογικό να υποθέσουμε ότι οι περισσότερες αναφορές είναι "θετικά προσανατολισμένες", δηλαδή είναι ένα σημάδι του γεγονότος ότι ο συγγραφέας που αναφέρει βρίσκει κάτι χρήσιμο στο υλικό που παραθέτει. Αποκλίνοντα μοτίβα αναφορών, όπως αρνητικές αναφορές, μπορεί να επηρεάσουν την ανάλυση ενός μεμονωμένου άρθρου ή συγγραφέα, αλλά αυτή η αρνητική επίδραση τείνει να εξαφανιστεί σε μια ανάλυση των μεγάλων ομαδοποιήσεων των συγγραφέων, όπως αυτές συμβαίνουν στα τμήματα, πανεπιστήμια ή χώρες. Το πλήθος των αναφορών σε ένα άρθρο επηρεάζεται από τους εξής παράγοντες σύμφωνα με τον Glanzel (2003, σ. 61.): το θεματικό περιεχόμενο, και στο πλαίσιο αυτού το «επίπεδο αφαίρεσης». Η εκδοτική τώρα δραστηριότητα στα θεωρητικά πεδία όπως λ.χ. τα Μαθηματικά είναι χαμηλότερη από αυτήν που παρατηρείται στις επιστήμες της ζωής (Δες σχήμα 2. 1). Δημοσιεύσεις σε ένα ερευνητικό τομέα με επίδοση στο πλήθος των αναφορών λαμβάνουν επίσης, κατά μέσο όρο, περισσότερες αναφορές όπως έχουμε ήδη προείπει.

2.3.1. Η ηλικία της δημοσίευσης και μέσο ετήσιο ποσοστό παραπομπών διαφόρων τομέων

Είναι προφανές ότι όσο παλαιότερη μια δημοσίευση τόσο μεγαλύτερη η χρονική περίοδος κατά την οποία μπορεί να πάρει αναφορές. Άρα, παλαιότερα άρθρα, λογικά θα αναφέρονται περισσότερο από ό,τι τα νεότερα.



Εικόνα 2 1 Το μέσο ποσοστό παραπομπής το 2000 σε πέντε διαφορετικούς τομείς της έρευνας.

2.3.2. Ο τύπος της δημοσίευσης

Ορισμένοι τύποι δημοσιεύσεων, όπως ανασκοπήσεις, όπως προείπαμε, τείνουν να αναφέρονται περισσότερες φορές από ό,τι τα πρωτότυπα άρθρα. Στις ποιο πολλές βιβλιομετρικές έρευνες περιλαμβάνονται μόνο συγκεκριμένα είδη δημοσιεύσεων. Μόνο εκείνα τα είδη δημοσιεύσεων που είναι φορείς των σχετικών ερευνητικών δεδομένων, που χρησιμοποιούνται για τη συγγραφή της εργασίας. Αυτές οι εκδόσεις είναι: δημοσιεύσεις σε περιοδικά όπως: επιστημονικά άρθρα, ανασκοπήσεις κ.λ.π.. «Κριτικές βιβλίων περιλήψεις,

παροράματα, και άλλες μορφές δημοσιεύσεων είναι αντικείμενα ειδικών βιβλιομετρικών εργασιών». (Glanzel, 2003).

2.3.3. Το μήκος της περιόδου παρατήρησης

Το μέγιστο πλήθος των αναφορών σε διαφορετικές χρονικές περιόδους μετά τη δημοσίευση εξαρτάται από το κάθε ερευνητικό πεδίο. Έτσι είναι ουσιαστικό να αποφασίσουμε για το χρονικό παράθυρο -τα έτη δηλαδή δημοσίευσης των άρθρων που μας ενδιαφέρουν- καθώς επίσης για το παράθυρο αναφοράς -για ποια έτη δηλαδή μετά τη δημοσίευση θέλουμε να μετρήσουμε τις αναφορές-. Υπάρχουν δύο εξαιρέσεις όσον αφορά το μοτίβο της αναφοράς όπως αναφέρθηκε πιο πάνω η:

Α. Ωραία κοιμωμένη ή Φαινόμενο του Μέντελ. Σύμφωνα με το οποίο ένα άρθρο δεν έχει καμία αναφορά για μερικά χρόνια, έως τη στιγμή που η ερευνητική κοινότητα θα ανακαλύψει την αξία του και θα ξεκινήσει η αναγνώριση του δια των αναφορών.

Β. Εξάλειψη με ενσωμάτωση. Που υφίσταται όταν οι πληροφορίες που ενσωματώνει ένα άρθρο έχουν χρησιμοποιηθεί τόσο πολλές φορές, σε σημείο που να εκλαμβάνονται πλέον ως γνωστές και, να μην αναφέρεται δι' αυτού του λόγου το άρθρο που τις παρουσίασε αρχικά. Βέβαια η εξάλειψη με ενσωμάτωση χρειάζεται πολύ χρόνο για να εμφανιστεί, ενώ σχεδόν όλες οι βιβλιομετρικές αναλύσεις αναφέρονται στα τελευταία 5 ή 10 χρόνια. Πάντως, και οι δύο περιπτώσεις αυτές είναι αρκετά σπάνιες και δεν διαφοροποιούν την ανάλυση η οποία πραγματοποιείται σε ένα μεγάλο αριθμό άρθρων.

Ο καταμερισμός των αναφορών σε μια ομάδα δημοσιεύσεων είναι σχεδόν πάντα ασύμμετρος, έστω κι αν λάβουμε υπόψη τους παραπάνω παράγοντες. Οποσδήποτε κάποιες από τις δημοσιεύσεις σε ένα σύνολο είναι αναφέρονται πολύ περισσότερο από κάποιες άλλες. Η κατανομή εντός της ίδιας ομάδας μπορεί να ποικίλει από μη αναφερόμενα άρθρα έως άρθρα μεταξύ του 1% των περισσότερων αναφερομένων συγκριτικά με τον παγκόσμιο μέσο όρο. Αυτό μπορεί να το δει κανείς τόσο σε κορυφαίες ερευνητικές ομάδες, όσο και σε λιγότερο υποσχόμενες και είναι επίσης αληθές για εργασίες σε υψηλής, καθώς και για εκείνες σε χαμηλής επιρροής περιοδικά. Ο Moed λέει ότι αυτό μπορεί να οφείλεται στο ότι ορισμένες εργασίες μπορούν να θεωρηθούν ως «σημαίες» και άλλες ως «τούβλα». Τα τούβλα θέτουν τα θεμέλια που χρειάζονται οι σημαίες να κυματίσουν με προορισμό να είναι υψηλής ποιότητας δημοσιεύσεις, αλλά μόνο τα άρθρα σημαίες παίρνουν αναφορές (Moed, H. F. 2005, σ. 86).

2.4. Πού θα βρούμε τις αναφορές

Στην επιστήμη της Βιβλιομετρίας, ως πηγή πληροφοριών θεωρούμε πάντοτε μία βάση δεδομένων. Οι πρωτογενείς πηγές βάσεων δεδομένων που μπορούν να υποστηρίξουν τη βιβλιομετρική ανάλυση, εκτός από βιβλιογραφικές εγγραφές επιστημονικών δημοσιεύσεων, περιέχουν στοιχεία για τις αναφορές μεταξύ των δημοσιεύσεων. Οι βάσεις δεδομένων είναι απαραίτητες, διότι περιέχουν στοιχεία τα οποία διαφορετικά θα πρέπει να συλλέγονται με το χέρι. Ας σημειωθεί, ωστόσο, ότι έχουν κατασκευαστεί περισσότερα για την εύρεση δεδομένων από ότι για την εκτέλεση πολύπλοκων μαθηματικών πράξεων ή, με άλλα λόγια, έχουν σχεδιαστεί για βιβλιογραφικούς παρά για βιβλιομετρικούς σκοπούς. Οι διάφορες βάσεις που ιδρύθηκαν είτε από επιχειρήσεις, είτε από δημόσιους ή ιδιωτικούς φορείς χρησιμοποιούνται για να δείχνουν τα αποτελέσματα της επιστημονικής και τεχνολογικής δραστηριότητας. Με ειδική επεξεργασία μπορούν να χρησιμοποιηθούν στην καθιέρωση βιβλιομετρικών δεικτών. Οι περισσότερες βάσεις δεδομένων είναι εξειδικευμένες. Ωστόσο, μόνο λίγες εξ αυτών δύναται να έχουν γενικό πεδίο εφαρμογής. Οι, πλέον, καθιερωμένες βάσεις με παγκόσμια, διεπιστημονική κάλυψη είναι οι παρακάτω:

2.4.1. Web of Science (WoS)

Εργαλεία για τη βιβλιομετρική ανάλυση του εντοπισμού και της μελέτης των αναφορών είναι η χρησιμοποίηση του Web of Science, μιας ευρέως γνωστής υπηρεσίας εντοπισμού βιβλιογραφικών αναφορών, αλλά και αρθρογραφίας, που αποτελεί μέρος της πλατφόρμας Web of Knowledge της Thomson Reuters, η οποία καλύπτει περισσότερες από 250 ειδικότητες και χωρίζεται σε 3 βάσεις δεδομένων:

- Science Citation Index με δεδομένα από 6.500 περιοδικά σε θέματα Επιστημών Υγείας και Φυσικών & Εφαρμοσμένων Επιστημών, από το 1970 έως σήμερα.

- Social Sciences Citation Index, με δεδομένα από 1.960 περιοδικά στο πεδίο των Κοινωνικών Επιστημών, από το 1970 .
- Arts & Humanities Citation Index, με δεδομένα από 1.180 περιοδικά στο πεδίο των Ανθρωπιστικών Επιστημών και Τεχνών, από το 1975 .

Υστερούν κυρίως περιοχές των επιστημονικών πεδίων “Engineering Technology”, “Social Sciences” και “Humanities”, ενώ υπερτερούν άλλες, όπως του πεδίου “Natural Sciences”.

Η Web of Science είναι η αρχική πηγή της έρευνας αναφορών, που μαζί με τη Google Scholar, αποτελούν τις πιο διεπιστημονικές και πιο ολοκληρωμένες πηγές αναφορών στη διάθεση της επιστημονικής κοινότητας. Η Web of Science παρέχει δεδομένα για τις αναφορές άρθρων, σε 12.500 περιοδικά και πάνω, τα οποία υπόκεινται σε αξιολόγηση κριτών, σχεδόν σε κάθε επιστήμη από το 1990, και εξής εξάγοντας πληροφορίες για τις αναφορές σε περισσότερα από 151.000 πρακτικά συνεδρίων. Οι βιβλιογραφικές εγγραφές και οι αναφορές που περιέχονται στη βάση χρονολογούνται, για ορισμένους επιστημονικούς κλάδους, από το 1900 και περιλαμβάνουν αναλυτικά μεταδεδομένα για άρθρα, συγγραφείς και ερευνητικούς οργανισμούς.

Σημαντικό πλεονέκτημα του Web of Science θεωρείται η αξιοπιστία του, αποτέλεσμα της αυστηρής αξιολόγησης των εκδόσεων και κυρίως των περιοδικών που εισάγονται στο σύστημα βάσει συγκεκριμένων κριτηρίων, μεταξύ των οποίων και η επιστημονική τους απήχηση. Στα μειονεκτήματα αναφέρονται κυρίως η ανισομερής κάλυψη των επιστημονικών δημοσιεύσεων, τόσο γεωγραφικά καθώς η συντριπτική πλειοψηφία του υλικού προέρχεται από αγγλόφωνες χώρες και κυρίως τις ΗΠΑ, όσο και θεματικά καθώς μια αναζήτηση αναφορών στη Web of Science έχει τις δυσκολίες της, αφού δεν περιλαμβάνει στη βάση δεδομένων όλα τα περιοδικά, αναφορές από βιβλία, διατριβές, διπλώματα ευρεσιτεχνίας και τεχνικές εκθέσεις, ενώ επίσης δεν περιλαμβάνονται ισομερώς όλοι οι κλάδοι, αφού οι θετικές επιστήμες είναι καλύτερα καλυμμένες από άλλες, όπως οι θεωρητικές. Επίσης δεν καλύπτονται ομοιόμορφα όλες οι επιστήμες από χρονολογικής άποψης, τα επιστημονικά περιοδικά που καλύπτουν τις θετικές επιστήμες πάνε πολύ πιο πίσω στο χρόνο από ότι τα περιοδικά που καλύπτουν τις τέχνες, τις ανθρωπιστικές και κοινωνικές επιστήμες. Ορισμένες δε θεματικές περιοχές καλύπτονται ελάχιστα, όπως η οικονομία και η εκπαίδευση. Η χρήση των υπηρεσιών της βάσης μπορεί να γίνει είτε στο σύνολό της ή μεμονωμένα για κάποιο κλάδο που περιέχει, απαιτεί όμως ετήσια συνδρομή για κάθε έναν από αυτούς ξεχωριστά.

Ο H.F.Moed(2005) υπολόγισε το βαθμό κάλυψης της WoS ανά επιστημονικό πεδίο ως εξής. Επειδή οι αναφορές των τεκμηρίων που περιλαμβάνονται στη βάση αφορούν και τεκμήρια που δεν περιλαμβάνονται σε αυτήν, θεώρησε ως σύνολο της δημοσιευμένης βιβλιογραφίας όλα τα έργα στα οποία αναφέρουν όλες οι δημοσιεύσεις που περιλαμβάνονται στη βάση δεδομένων WoS. Στη συνέχεια υπολόγισε το ποσοστό των αναφορών που αφορούν δημοσιεύσεις που περιλαμβάνονται στη WoS προς το σύνολο των αναφορών των εργασιών αυτών (ποσοστό συνολικής κάλυψης) και προς το σύνολο των αναφορών σε άρθρα περιοδικών (ποσοστό κάλυψης περιοδικών).

Επίσης υπολόγισε την επιροή που έχουν τα περιοδικά για κάθε επιστημονικό τομέα ως ποσοστό των αναφορών στα περιοδικά προς το συνολικό αριθμό των αναφορών.

2.4.1.1. Το Ευρετήριο Επιστημονικών Αναφορών (Science Citation Index-SCI)

Δομή και δυνατότητες του SCI

Εξέχουσας σημασίας είναι η χρήση δεδομένων ομοιόμορφης ποιότητας, ώστε να καταμετρηθεί η ποσότητα του επιστημονικού υλικού. Αυτή η βάση δεδομένων καθορίζεται με συγκεκριμένα κριτήρια έτσι ώστε να προσδιορίζεται επακριβώς το υπό εξέταση σύνολο.

Μέσα σε αυτό το πλαίσιο κινείται το Science Citation Index (SCI) το Ευρετήριο Επιστημονικών Αναφορών που κατασκευάστηκε από το Αμερικανικό Ινστιτούτο στο πλαίσιο της Επιστήμης της Πληροφόρησης.

Ο δημιουργός του ευρετηρίου, ο Eugene Garfield, παραδέχθηκε ότι δεν ήταν εφικτό να καλύψει και να αξιοποιήσει όλο το φάσμα που καλύπτουν τα υπάρχοντα επιστημονικά περιοδικά. «Ένας λόγος που εξηγεί την παραδοχή αυτή είναι ότι κανείς δεν ξέρει ακριβώς πόσα περιοδικά δημοσιεύονται, επειδή δεν έχει συμφωνηθεί και καθοριστεί το πότε ένα γραπτό συνιστά περιοδικό» (1972).

Ο Garfield υπολόγισε πρώτα τον αριθμό των επιστημονικών περιοδικών που απαιτούνται για να αξιοποιήσει με τον καλύτερο δυνατό τρόπο το μέγιστο ποσό των επιστημονικών πληροφοριών. Έτσι αξιοποίησε τον νόμο του Bradford. Σύμφωνα με αυτόν τον νόμο χρειάστηκε ένας πυρήνας αποτελούμενος

από 450 έως 970 ερευνητικά περιοδικά ώστε να καλυφθεί το 93% του ποιο αξιόλογου τμήματος της επιστημονικής παραγωγής που υπάρχει σε έναν συγκεκριμένο ερευνητικό πεδίο.

Εν συνεχεία, ο Garfield προέβη σε συνδυασμό του νόμου περί διασποράς, που εισήχθη από τον Bradford, με τον νόμο περί συγκέντρωσης, που είχε ο ίδιος αναπτύξει (1972). Ο νόμος του Bradford ορίζει και οριοθετεί ένα επιστημονικό πεδίο, αλλά αν μια βάση δεδομένων κληθεί να καλύψει πάνω από ένα επιστημονικό πεδίο, τότε αυτό συνεπάγεται ότι και ο αριθμός των επιστημονικών περιοδικών θα πρέπει να πολλαπλασιαστεί, ώστε να είναι ποσοτικά ανάλογος με τον αριθμό των πεδίων. Σύμφωνα με τον Garfield, λόγω του ότι ένα σημαντικό μέρος των επιστημονικών κλάδων επικαλύπτονται, ο πυρήνας της επιστημονικού παραγόμενου έργου για όλους αυτούς τους κλάδους μπορεί επίσης να καλυφθεί από περίπου 500 έως 1.000 περιοδικά.

Πρωταρχικό ήταν να δημιουργηθεί μια μέθοδος που να προσδιορίζει αυτά τα 450 με 970 επιστημονικά περιοδικά. Προς την κατεύθυνση αυτή, ο Garfield θεώρησε ότι ο αριθμός των αναφορών ήταν ένα κριτήριο που κάνει μία έρευνα σημαντική και ιδιαίτερη. Από τη στιγμή που οι δημιουργοί αναφέρουν εργασίες προηγούμενες έτσι ώστε να υποστηρίξουν, επεξηγήσουν, περιγράψουν ή αναλύσουν ένα τμήμα του έργου τους, η αναφορά σε ένα ερευνητικό άρθρο προσδίδει τη σημασία που αποδίδεται από την ερευνητική κοινότητα στην έρευνα. Άρα οι αναφορές πιθανόν να εκληφθούν ως κριτήριο για την επιλογή των ερευνητικών περιοδικών που έχουν υψηλό αντίκτυπο, από τα άρθρα που δημοσιεύουν.

Στην αρχή έγινε καταμέτρηση του αριθμού των αναφορών σε ένα άρθρο, σε ένα δεδομένο επιστημονικό περιοδικό. Στη συνέχεια υπολογίστηκε ο παράγοντας του αντίκτυπου, διαιρώντας τον αριθμό των αναφορών με τον αριθμό των άρθρων που περιέχονται σε ένα δεδομένο επιστημονικό περιοδικό. Επετεύχθη, με τον τρόπο αυτό, η εξάλειψη κάθε προκατάληψης που απορρέει από το μέγεθος ενός περιοδικού, καθιστώντας την αναφορά ανάλογη με τον αριθμό των άρθρων.

Εύλογα εξάγεται το συμπέρασμα ότι η βάση δεδομένων SCI καλύπτει τα επιστημονικά περιοδικά που ασκούν μεγαλύτερη επιρροή και που είναι ευρύτερα χρησιμοποιούμενα και αναγνωρισμένα παγκοσμίως, όπως προκύπτει από τη μέτρηση των «δεικτών αναφοράς» τους. Περιορίζει το πεδίο που καλύπτουν τα παγκοσμίου κλάσης επιστημονικά περιοδικά, που αντιπροσωπεύουν τον «πυρήνα» της επιστημονικής παραγωγής σε συγκεκριμένους τομείς και που εξαλείφουν την έρευνα που δε θεωρείται καθιερωμένη και η οποία περιορίζεται σε μια συγκεκριμένη ομάδα περιοδικών. Το πρόβλημα για μια βάση δεδομένων έγκειται στο να χαρακτηί ένα όριο ανάμεσα στο «ασχυρό» και στο «αδύναμο». Ποιος παράγοντας κάνει τη διαφορά εν τέλει; Η ποιότητα, η ποσότητα ή κάποιο άλλο κριτήριο; Εν πάση περιπτώσει, αυτό που ισχύει για τους ερευνητές είναι η διεθνής αναγνώριση του παραγόμενου έργου τους, που καθορίζει την ιεραρχία και την κατάταξη βάσει της αξίας του καθενός. Περί το 1981, η βάση δεδομένων SCI κάλυπτε περίπου 3.100 επιστημονικά περιοδικά. Εκείνη την εποχή, υπήρχαν περίπου 70.000 ανάλογες δημοσιεύσεις σε όλον τον κόσμο! Αυτός ήταν ο λόγος για τον οποίο το περιβάλλον της SCI είχε βαπτιστεί ως το «καθιερωμένο» (Frame, 1977). Η επιλογή των περιοδικών που καλύπτονται από την SCI έδωσε έτσι μια ποιοτική διάσταση και πνοή στην καταμέτρηση του παραγόμενου επιστημονικού ερευνητικού έργου.

Εκτός από την, σχεδόν, μονοπωλιακή της θέση, ο λόγος που η βάση δεδομένων SCI χρησιμοποιείται σήμερα τόσο έντονα, δεν μπορεί να αποδοθεί αποκλειστικά και μόνο στη χρήση των αναφορών. Η, εν λόγω, βάση δεδομένων καλύπτει μια τεράστια έκταση της επιστήμης. Είναι διεπιστημονική, σε αντίθεση με τις περισσότερες βάσεις δεδομένων, που είναι ιδιαίτερα εξειδικευμένες. Με την SCI, είναι δυνατόν να αναλάβει κανείς μια ευρεία μελέτη της επιστήμης, χάρη στην ομοιόμορφη μεταχείριση των θετικών και φυσικών επιστημών. Παρόλα αυτά, το Ινστιτούτο για την Επιστήμη της Πληροφόρησης (ISI) έχει επίσης επεκτείνει τη συλλογή των δεδομένων και είναι πλέον δυνατό να βρεθούν επαρκείς πληροφορίες σχετικά με τις κοινωνικές και τις ανθρωπιστικές επιστήμες, αξιοποιώντας το Ευρετήριο Παραπομπών για τις Κοινωνικές Επιστήμες (Social Science Citation Index - SSCI) και το Ευρετήριο Παραπομπών για τις Τέχνες & τις Ανθρωπιστικές Επιστήμες (Arts & Humanities Citation Index - AHCI), όπως ήδη έχουμε αναφέρει στο προηγούμενο κεφάλαιο.

Η SCI καταγράφει επίσης την υπαγωγή (διευθύνσεις, θεσμικές συνδέσεις κ.λ.π.) όλων των συγγραφέων του κάθε άρθρου, σε αντίθεση με τις περισσότερες βάσεις δεδομένων, οι οποίες καταγράφουν μόνο τον πρώτο συγγραφέα των συν-υπογεγραμμένων άρθρων (η αρχική λίστα ενδέχεται ορισμένες φορές να αποτελεί «φόρο τιμής» στους γνωστότερους ερευνητές, οι οποίοι δεν έχουν κατ' ανάγκη συμβάλει ιδιαίτερα στον τομέα). Προς το παρόν, οι βάσεις δεδομένων SCI και Physic Brief συγκαταλέγονται στις πιο σπάνιες, που ακολουθούν μια πολιτική πολλαπλής καταχώρησης. Και όμως, η μέθοδος αυτή προσφέρει μια σειρά από πλεονεκτήματα εξέχουσας σημασίας, ιδίως σε μια εποχή όπου οι μελέτες της διεθνοποίησης της επιστήμης είναι προηγμένες κατά πολύ. Επειδή όλες οι συνεργασίες των συγγραφέων που συνυπογράφουν τα διάφορα

επιστημονικά άρθρα- καταγράφονται, ένα πρόγραμμα σε υπολογιστή έχει τη δυνατότητα να επιλέξει ένα άρθρο που γράφτηκε από ερευνητές σε διαφορετικά εργαστήρια και σε διαφορετικές χώρες. Αυτό προσδίδει στην SCI διάρκεια και μόνιμο χαρακτήρα στο πλαίσιο της βιβλιομετρικής έρευνας, κάτι που οι βιβλιομετρητές δεν μπορούν να αγνοήσουν.

2.4.1.2. Το Ευρετήριο Επιστημονικών Αναφορών και οι περιορισμοί του

Ωστόσο, η βάση δεδομένων SCI έχει και τους περιορισμούς της, οι οποίοι είναι κοινοί σε πολλές άλλες βιβλιομετρικές βάσεις δεδομένων.

Γιατί οι συγγραφείς χρησιμοποιούν αναφορές;

Οι λόγοι που παρακινούν έναν συντάκτη επιστημονικού άρθρου να κάνει αναφορά σε άλλο έργο επιστημονικού περιεχομένου είναι πολύπλοκοι. Οι κοινωνιολόγοι της Επιστήμης ανέλυσαν επί σειρά ετών τη σπουδαιότητα που ενέχει η κάθε αναφορά και έχουν επισημάνει ότι η αναφορά στο έργο άλλων ερευνητών δεν σχετίζεται πάντα με την αυθεντικότητα, τη σημασία ή ακόμα και την «ποιότητα» του, εν λόγω, έργου.

Μια αναφορά δύναται, επίσης, να εξαρτάται από την ικανότητα ενός άρθρου να απευθύνεται και να αφορά ένα μεγάλο ακροατήριο. Διάσημοι επιστήμονες επιβλέπουν συχνά πολλούς από τους φοιτητές τους και τα άρθρα τους είναι πολύ πιο πιθανό να αποτελέσουν αντικείμενο αναφοράς σε σύγκριση με εκείνα των συναδέλφων τους που παρουσιάζουν μικρότερη δύναμη επιρροής (αυτό ισχύει, ιδίως, αν το όνομα του «επιβλέποντα» εμφανίζεται, καλώς ή κακώς, μεταξύ των συγγραφέων που συνυπογράφουν ένα άρθρο). Η βαρύτητα που φέρει η κοινωνική δομή στο πλαίσιο ενός επιστημονικού κλάδου δεν συνδέεται κατ' ανάγκη άμεσα με την ποιότητα της έρευνας. Οι συγγραφείς επιστημονικών άρθρων μπορεί να αναφέρονται σε διακεκριμένους επιστήμονες, είτε λόγω τακτικής, είτε ως αφιέρωμα και όχι, απαραιτήτως, επειδή αναγνωρίζουν και θαυμάζουν το κομμάτι του έργου που παραθέτουν ως αναφορά. Το έργο του ερευνητή που διαθέτει μια πειραματική τεχνική ή μεθοδολογία, όχι απαραίτητα υψηλής ποιότητας, αλλά απλώς χρήσιμη, θα αναφέρεται ως αναφορά κάθε φορά που χρησιμοποιείται.

Οι αναφορές αποτελούν μονάδα μέτρησης της συνολικής επίδρασης που ασκεί ένα άρθρο ή οι δημιουργοί του στην επιστημονική κοινότητα. Πρόκειται για μια σύνθετη κοινωνικό-επιστημολογική παράμετρο, που προκαλεί πιθανώς έναν παράγοντα ποιότητας, αλλά αυτός ο παράγοντας δεν ισοδυναμεί, ούτε συσχετίζεται κατηγορηματικά με την επιστημονική ποιότητα (Seglen, 1992).

Αρνητικές αναφορές

Οι αναφορές μπορεί επίσης να είναι αρνητικές. Ειδικότερα, ένας συγγραφέας μπορεί να αναφερθεί σε κάποια αναφορά, εξαιτίας μίας έρευνάς του με αμφιλεγόμενο χαρακτήρα ή λόγω λανθασμένης μεθοδολογίας του. Και στην περίπτωση αυτή, η αναφορά δεν μετρά και δεν υπολογίζει πάντοτε την ποιότητα της έρευνας, αλλά την επίδραση ενός συγκεκριμένου έργου ή ενός μεμονωμένου επιστήμονα.

Επιστήμονες & άρθρα που μένουν εκτός αναφορών

Επιπρόσθετα ο αριθμός των επιστημόνων που αναφέρονται σε αναφορές είναι εξαιρετικά περιορισμένος. Περισσότερο από το ήμισυ (δηλαδή περίπου ένα 55%) των άρθρων που δημοσιεύονται στα επιστημονικά περιοδικά και που καλύπτονται από τη βάση δεδομένων SCI δεν αναφέρθηκαν ούτε μια φορά σε αναφορές, μέσα στα πέντε χρόνια που ακολούθησαν από τη δημοσίευσή τους. Το ποσοστό των άρθρων που μένουν εκτός αναφορών κυμαίνεται σε διαφορετικά επίπεδα ανά επιστημονικό κλάδο: για παράδειγμα στην κατηγορία της Μηχανικής - Μηχανολογίας, το ποσοστό υπερβαίνει το 72% (Pendlebury, 1991).

2.4.1.3. Ο παράγοντας της γλώσσας

Ένα άλλο σημείο που έχει αναφερθεί συχνά, που επικρίθηκε και σε ορισμένες περιπτώσεις αναλύθηκε, είναι το γεγονός ότι η συγκεκριμένη βάση δεδομένων ευνοεί σαφώς τους αγγλόφωνους επιστήμονες (όπως αναλύσαμε στο προηγούμενο κεφάλαιο). Σύμφωνα με τον Garfield, δεν είναι ούτε εύκολο, ούτε οικονομικά αποδοτικό για τη βάση δεδομένων SCI να περιλαμβάνει περιοδικά που δεν χρησιμοποιούν το λατινικό αλφάβητο (1975). Αυτή η παρατήρηση δείχνει ότι, προκειμένου να γίνει δεκτή η βάση δεδομένων στην κατηγορία των πιο πρωτοποριακών, τα άρθρα πρέπει να συντάσσονται στη γλώσσα που έχει επικρατήσει. Είναι απολύτως σαφές ότι η σημερινή προσφορά γίνεται στην αγγλική γλώσσα και το σύστημα είναι αυτό-διδιωνιζόμενο (Garfield, 1988). Επιπλέον οι ερευνητές σε μη αγγλόφωνες χώρες, που όμως δημοσιεύουν στα Αγγλικά, απολαμβάνουν μία συγκριτικά πιο διευρυμένη παρουσία στο χώρο. Ενδεικτικό είναι το παράδειγμα της Σκανδιναβίας (Sivertsen, 1991).

2.4.1.4. Η κατανομή ανά επιστημονικό κλάδο

Ένα από τα βασικά πλεονεκτήματα των βιβλιομετρικών βάσεων δεδομένων είναι η δυνατότητα χρήσης πολύ λεπτομερών και αναλυτικών ταξινομήσεων των επιστημονικών κλάδων. Κατά κανόνα, οι αναφορές αναφέρονται σε άρθρα κ.λ.π. που εμπίπτουν στην ίδια κατηγορία ή υποκατηγορία ενός συγκεκριμένου πεδίου ή κλάδου.

Ένα πρόβλημα προκύπτει, ωστόσο, όταν η SCI χρησιμοποιείται για να μελετήσει την κατάσταση της επιστήμης, μέσα από μια διεπιστημονική βάση δεδομένων, η οποία καλύπτει ένα ευρύ πεδίο και δεν μπορεί να αντανakλά τις διαφορές δυναμικές και τις ιδιαιτερότητες των αναφορών που πηγάζουν από διαφορετικούς επιστημονικούς κλάδους. Θεωρητικά, τα κριτήρια επιλογής είναι τα ίδια για κάθε ένα από τα περιοδικά που καλύπτονται από τη βάση δεδομένων. Παρόλα αυτά, οι παραδόσεις και οι συνήθειες της εκάστοτε δημοσίευσης και αναφοράς διαφέρουν ανάλογα με τον επιστημονικό τομέα, γεγονός που επηρεάζει την αντιπροσωπευτικότητα των δεδομένων. Για παράδειγμα, το μεγαλύτερο ποσοστό του επιστημονικού υλικού, που έχει δημοσιευθεί παγκοσμίως και αφορά στη Φυσική, τη Χημεία και τη Βιοϊατρική, εκπροσωπείται επαρκώς από την SCI. Αντίθετα, προβλήματα παρουσιάζει η κάλυψη του πεδίου των Γεωεπιστημών, της Βιολογικής έρευνας, της Μηχανικής και της Τεχνολογίας, των Μαθηματικών και, ως έναν βαθμό, της κλινικής Ιατρικής. Ποια τα αίτια; Η απάντηση βρίσκεται στο γεγονός ότι σε ορισμένους επιστημονικούς κλάδους η επικοινωνία είναι συγκεντρωμένη στα κυριότερα και εξειδικευμένα, διεθνή, επιστημονικά περιοδικά, καθώς και στο γεγονός ότι ορισμένα περιοδικά ασκούν πιο περιορισμένη «επιρροή». Το πρόβλημα της γλώσσας είναι επίσης εντονότερο σε ορισμένους τομείς της επιστήμης.

Όλα αυτά ισχύουν εξίσου για τις αναφορές. Προκειμένου να καθοριστεί, για παράδειγμα, η ικανότητα μιας χώρας από την άποψη των αναφορών είναι απαραίτητο να διεισδύσουμε στις ιδιαιτερότητες του κάθε επιστημονικού πεδίου, οι οποίες επηρεάζουν τα ευρετήρια αναφορών. Για παράδειγμα, προέκυψε ότι, κατά μέσο όρο, σε σύντομο χρονικό διάστημα, τα άρθρα Βιοϊατρικής αναφέρονται συχνότερα, συγκριτικά με τα άρθρα που πραγματεύονται τα Μαθηματικά ή την κλινική Ιατρική. Τέτοιοι συσχετισμοί πρέπει να ληφθούν υπόψη, ειδικά όταν τα στοιχεία αποτελούν αντικείμενο ερμηνείας. Χρειάζεται να γίνουν διορθώσεις, αλλά κάτι τέτοιο απαιτεί πληρέστερη κατανόηση της δομής των επιστημονικών πεδίων που αναλύονται. Βασική προϋπόθεση είναι να τηρείται -ως ένα βαθμό- επιφυλακτικότητα κατά την ερμηνεία των αριθμών. Η συνεργασία μεταξύ των επαγγελματιών, που δραστηριοποιούνται στους διάφορους επιστημονικούς κλάδους και των βιβλιομετρικών διαδραματίζει καίριο ρόλο σε τέτοιες περιπτώσεις.

2.4.2. Scopus

Η Scopus είναι μια σχετικά νέα βάση δεδομένων που παράγεται από τις εκδόσεις Elsevier και παρέχεται μέσα από την πλατφόρμα SciVerse. Θεωρείται η μεγαλύτερη στον κόσμο βάση δεδομένων αναφορών και περιλήψεων έγκριτης, διεθνούς βιβλιογραφίας, με παροχή εργαλείων βοήθειας για αναζήτηση, ανάκτηση, ανάλυση και οπτικοποίηση τεκμηρίων. Τα μεταδεδομένα περιέχουν αναλυτικά στοιχεία για άρθρα, συγγραφείς και οργανισμούς, ενώ η εισαγωγή των εκδόσεων πραγματοποιείται μετά από ποιοτική αξιολόγηση βάσει κριτηρίων. Η Scopus υστερεί, σε σχέση με την WoS, χρονικά αφού οι παλαιότερες βιβλιογραφικές εγγραφές ξεκινούν το 1966, ενώ οι αναφορές σε δημοσιεύσεις καταγράφονται μετά το 1995. Καλύπτει πάνω από 20.500 τίτλους περιοδικών με κριτές από 5.000 διεθνείς εκδοτικούς οίκους σε όλον τον κόσμο, 49 εκατομμύρια εγγραφές (το 78% με περιλήψεις), πάνω από 5,3 εκατομμύρια άρθρα συνεδρίων και το 100% της κάλυψης της βάσης Medline (βλ. παρακάτω). Οι τίτλοι είναι από όλες τις γεωγραφικές περιοχές, συμπεριλαμβανομένων και μη-αγγλικών τίτλων με την προϋπόθεση της ύπαρξης περίληψης της δημοσίευσης στα Αγγλικά.

Η Scopus καλύπτει πολλές διαφορετικές επιστήμες, όπως Χημεία, Φυσική, Μαθηματικά, Μηχανική, Επιστήμες Ζωής και Υγείας (συμπεριλαμβανομένων των αναφορών που ανακτώνται από το Medline), Κοινωνικές Επιστήμες, Ψυχολογία, Οικονομικές Επιστήμες, Βιολογικές, Γεωργικές και Περιβαλλοντικές Επιστήμες. Οι πρώτες έξι (6) καλύπτονται πληρέστερα από τις επόμενες. Στη Scopus, οι λίστες αναφοράς περιλαμβάνονται από το 1996 και αυτό οδηγεί αυτόματα σε αντίστοιχες εγγραφές δημοσιεύσεων που δύνανται να αναφέρονται.

Η Scopus προσφέρει παρόμοια μετρικά άρθρων, συγγραφέων, και επιπέδου περιοδικών με το Web of Knowledge, αλλά χρησιμοποιεί ελαφρώς διαφορετικούς αλγόριθμους για τον υπολογισμό τους. Τα μετρικά περιλαμβάνουν τις τυποποιημένες επιλογές, όπως τιμές αναφορών και h-index, καθώς και πρωτότυπες

προσφορές, όπως οι: *SCImago Journal Rank* (SJR) και *Source Normalized Impact per Paper* (SNIP) από το SCImago. Το Scopus ξεκίνησε πρόσφατα τις Εναλλακτικές μετρικές "Altmetrics για το Scopus", μια τριμερή εφαρμογή που τρέχει μέσα στο sidebar των σελίδων του Scopus, για να εντοπίζει αναφορές εγγράφων σε τοποθεσίες κοινωνικών μέσων ενημέρωσης, ή επιστημονικά blogs.

Πρόσβαση: <http://www.scopus.com>.

2.4.2.1. SCImago Journal and Country Rank.

Η SCImago είναι μια δωρεάν ιστοσελίδα που λειτουργεί με δεδομένα Scopus για να υπολογίσει δύο μετρικές: το SCImago Journal Rank (SJR) και το Source Normalized Impact per Paper (SNIP), οι οποίες τα συγκρίνουν άμεσα με το Impact Factor του Web of Knowledge. Το SJR βασίζεται στις φορές που ένα άρθρο αναφέρεται, αλλά επίσης χρησιμοποιεί έναν αλγόριθμο που είναι παρόμοιος με το PageRank της Google για να υπολογίσει τα άρθρα επιρροής, τα οποία χρησιμοποιεί για να δημιουργήσει τις ταξινομήσεις. Με τη χρήση της ηλεκτρονικής διασύνδεσης SCImago, οι χρήστες μπορούν να συγκρίνουν τις ταξινομήσεις έως δέκα περιοδικών τη φορά, να εμφανίζουν κορυφαία περιοδικά, ακόμα και να εμφανίζουν τις χώρες με περιοδικά επιρροής σε μια επιστήμη. Πρόσβαση: <http://www.SCImagojr.com>.

2.4.3. Google Scholar

Η Google Scholar η οποία δημιουργήθηκε από την GoogleInc παρέχει, με έναν απλό τρόπο αναζήτησης, πληροφορίες αναφερομένων άρθρων. Παρέχει επίσης ελεύθερη πρόσβαση σε βιβλιογραφικά στοιχεία και αναφορές γενικότερα επιστημονικών τεκμηρίων (άρθρα, διατριβές, βιβλία, περιλήψεις και αποφάσεις δικαστηρίων) στο σύνολο των επιστημονικών πεδίων από πηγές πολλών ακαδημαϊκών εκδοτών, επαγγελματικών ενώσεων, ιδρυματικών και θεματικών αποθετηρίων, πανεπιστημίων και άλλων δικτυακών τόπων.

Η Google Scholar μας βοηθά να αναζητήσουμε όλη την ακαδημαϊκή βιβλιογραφία, να εξερευνήσουμε συναφή έργα, αναφορές, συγγραφείς και δημοσιεύσεις, να εντοπίσουμε το πλήρες έγγραφο μιας εργασίας, μέσω του παγκόσμιου ιστού, να ελέγξουμε ποιος αναφέρεται στις δημοσιεύσεις μας και τέλος, να δημιουργήσουμε ένα δημόσιο προφίλ συγγραφέα.

Παρά το μεγάλο εύρος σε περιεχόμενο, τα μεταδεδομένα των εγγράφων είναι περιορισμένα και απουσιάζουν κρίσιμα στοιχεία για τη γεωγραφική προέλευση ή τη διεύθυνση των οργανισμών στους οποίους ανήκουν οι συγγραφείς. Επιπλέον τα κριτήρια εισαγωγής των πηγών του περιεχομένου είναι περιορισμένα, ενώ δεν διατίθενται ακριβή στοιχεία για τη γεωγραφική ή τη θεματική κάλυψη που παρέχει το σύστημα.

Επίσης οι πληροφορίες, οι σχετικές με την παράθεση αναφορών, είναι μάλλον αναξιόπιστες, γιατί αυτές μπορεί να περιέχουν διπλοεγγραφές ή και αναντιστοιχίες. Επιπλέον ο μελετητής Google δεν ευρετηριάζει όλα τα επιστημονικά άρθρα, ενώ ευρετηριάζει αρχεία γραμμένα ως εφαρμογές powerpoint και word. Επειδή η αναζήτηση με το όνομα του συγγραφέα είναι προβληματική, ενδείκνυται να γίνεται μόνο με το επώνυμο του συγγραφέα, συνδυάζοντας το με τον κύριο τίτλο του άρθρου μέσα σε εισαγωγικά. Ο όρος «αναφορά» σε αγκύλες δηλαδή [citation], στην αρχή της αναφοράς του άρθρου, δείχνει ότι το πλήρες κείμενο του τεκμηρίου δεν είναι προσβάσιμο μέσω του Google Scholar.

2.4.3.1. Google Scholar Citations

Αυτή η δωρεάν υπηρεσία της Google επιτρέπει στους συγγραφείς να δημιουργήσουν προφίλ, βάσει του οποίου διαχειρίζονται, υπολογίζονται και εντοπίζονται βιβλιογραφικά δεδομένα, όπως το h-index και το i10-index (δηλαδή, ο αριθμός των άρθρων με τουλάχιστον δέκα αναφορές). Χρησιμοποιώντας ένα στατιστικό μοντέλο που βασίζεται στο συγγραφέα και στα μεταδεδομένα των άρθρων του για τον εντοπισμό των σχετικών αναφορών, η υπηρεσία προσφέρει τη δυνατότητα της αυτόματης προσθήκης νέων άρθρων σε δημόσια ή ιδιωτικά προφίλ των χρηστών. Η Google, επίσης, ξεκίνησε πρόσφατα μια σχετική υπηρεσία, το Google Scholar Metrics, που μετρά την «ορατότητα και επιρροή» των άρθρων και δημοσιεύσεων από το 2007, με βάση τα βιβλιογραφικά δεδομένα του Google Scholar.

Πρόσβαση: <http://scholar.google.com/intl/en/scholar/citations.html>.

2.4.3.2. Πώς να βρείτε πόσες αναφορές υπάρχουν σε συγκεκριμένο άρθρο

ή "Ποιος επικαλείται (αναφέρει) Ποιον"

- Πηγαίνετε στο Google Scholar.
- Επιλέξτε Σύνθετη Αναζήτηση στον Μελετητή (κάτω βέλος στο πλαίσιο αναζήτησης).
- Εισάγετε τους κατάλληλους όρους αναζήτησης.
- Κάντε κλικ στο κουμπί αναζήτησης του Μελετητή.
- Εντοπίστε το σωστό άρθρο στη λίστα αποτελεσμάτων αναζήτησης.
- Εάν το άρθρο αυτό αναφέρεται από τους άλλους, θα δείτε ένα σύνδεσμο "Αναφορά από" στο κάτω μέρος της οθόνης.
- Κάντε κλικ σε αυτόν το σύνδεσμο για να δείτε ποιοι ανέφεραν αυτό το άρθρο.

2.4.4. Academic Search

Η Academic Search είναι μια δωρεάν υπηρεσία που αναπτύχθηκε από τη Microsoft για να βοηθήσει ερευνητές, καθηγητές και φοιτητές να βρουν εύκολα και γρήγορα πληροφορίες ακαδημαϊκού περιεχόμενου. Η Academic Search ευρετηριάζει όχι μόνο εκατομμύρια επιστημονικές εργασίες, αλλά αναδεικνύει τις βασικές σχέσεις μεταξύ των εργασιών και μεταξύ των συγγραφέων με κατάλληλο οπτικό τρόπο, που υπογραμμίζει τις κρίσιμες συνδέσεις οι οποίες συμβάλλουν στον καθορισμό της επιστημονικής έρευνας.

2.4.5. ACM Digital Library (Association for Computing Machinery)

Η ACM καλύπτει τον χώρο της Πληροφορικής από την Ένωση Μηχανικών Πληροφορικής και άλλους εκδότες. Περιλαμβάνει βιβλία, άρθρα περιοδικών, πρακτικά συνεδρίων, διδακτορικές διατριβές, διπλωματικές, καθώς και τεχνικές εκθέσεις.

2.4.5.1. Οδηγίες ανάκτησης αναφορών

- Συνδεθείτε με την ACM (Association for Computing Machinery) και κάντε κλικ στο "ψηφιακή βιβλιοθήκη" στην αριστερή πλευρά της σελίδας.
- Κάντε κλικ στο "Σύνθετη Αναζήτηση" στην κορυφή της δεξιάς στήλης.
- Στην αναζήτηση για συγγραφέα ή τίτλο θα πρέπει να αναφερθούν: Κατά την αναζήτηση για τους Συγγραφείς στην ενότητα Ονόματα - αν το όνομα είναι γνωστό - εισάγετε όνομα σε εισαγωγικά για να βελτιώσετε την αναζήτηση. Εάν το πλήρες όνομα δεν είναι γνωστό, χρησιμοποιήστε το "Refine by People " σύνδεσμο, στην αριστερή στήλη που προέκυψε, στη σελίδα αποτελεσμάτων για να επιλέξετε τον συγγραφέα που ψάχνετε.
- Σιγουρευτείτε ότι η "expanded form» έχει επιλεγεί για την "Sort by", στο πάνω δεξιά μέρος των αποτελεσμάτων.
- Επιλέξτε ένα αποτέλεσμα, στη συνέχεια κάντε κλικ στη "Cited by" καρτέλα στην ενότητα "Additional Information " στο κάτω μέρος των επιθυμητών εγγραφών του αποτελέσματος αναζήτησης, για να δείτε αν, τυχόν, άλλες εκδόσεις στη βάση δεδομένων τους αναφέρουν.

2.4.6. IEEEXplore

Η IEEEXplore περιέχει σχεδόν 2 εκατομμύρια άρθρα και μελέτες από τα περιοδικά και τα πρακτικά συνεδρίων του Ινστιτούτου Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών (IEEE- Institute of Electrical and Electronics Engineers) συν τις δημοσιεύσεις στα πρότυπα δημοσιεύσεων IEEE. Εδώ ορίζονται ευρέως τα θέματα των «Ηλεκτρολόγων Μηχανικών» και τα αντικείμενα ευρίσκονται από όλους τους τομείς της μηχανικής, στα οποία εφαρμόζονται τα ηλεκτρονικά.

2.4.6.1.Οδηγίες χρήσης

- Βρείτε το άρθρο που σας ενδιαφέρει σε (αναζήτηση με συγγραφέα, τίτλο κ.λ.π.).
- Κάντε κλικ στο σύνδεσμο «Abstract Plus» κάτω από την αναφορά του άρθρου, στη λίστα των αποτελεσμάτων.
- Στο αριστερό παράθυρο με τίτλο "On This Page" επιλέξτε το σύνδεσμο "Cited By IEEE" – ή κυλίστε προς τα κάτω στη σελίδα με τις ενότητες «Cited By IEEE " και "Cited By Other Publishers." Σημείωση: Μπορείτε επίσης να αλλάξετε το "Sort By ", στη σελίδα αποτελεσμάτων "Most Cited ", που θα παραθέσει τους αριθμούς των αναφορών, από την υψηλότερη στη χαμηλότερη, στη σελίδα αποτελεσμάτων.

2.4.7. JSTOR (Journal Storage)

Η JSTOR είναι μία online συλλογή ακαδημαϊκών περιοδικών σε όλους τους τομείς. Η βάση δεδομένων JSTOR καλύπτεται τα τελευταία χρόνια από επιστημονικά περιοδικά σε όλους τους κλάδους. Η JSTOR παρέχει τον αριθμό αναφορών ενός άρθρου στο JSTOR.

2.4.7.1.Οδηγίες χρήσης

Το JSTOR, της ITHAKA, έχει δύο επιλογές αναζήτησης: α) τον Εντοπισμό σημείου αναφοράς και β) τη σύνθετη αναζήτηση. Τα εργαλεία αυτά χρησιμεύουν ως καλές επιλογές για τους ερευνητές που γνωρίζουν την προσβασιμότητα ενός περιοδικού στο JSTOR και έχουν στη διάθεσή τους τις πληροφορίες των αναφορών. Ο Εντοπισμός σημείου αναφοράς μας μεταφέρει απευθείας στο άρθρο. Μόλις ο χρήστης επιλέγει ένα έγγραφο από μια αναζήτηση, αυτός ή αυτή έχει την επιλογή να χρησιμοποιήσει στο Google Μελετητή τη λειτουργία Αναφερόμενα Αντικείμενα σχετικά με την αναφορά του στοιχείου αναζήτησης. Με την επιλογή αυτής της λειτουργίας, η αναζήτηση θα τρέξει στον Μελετητή Google για τα αναφερόμενα άρθρα στο αρχικό συγκεκριμένο άρθρο. Οι χρήστες έχουν και ένα άλλο χαρακτηριστικό του JSTOR στη διάθεσή τους, που ονομάζεται Ανίχνευση Αναφορών. Επιλέγοντας αυτό το χαρακτηριστικό, οι χρήστες λαμβάνουν ένα e-mail, όταν ένα άλλο άρθρο στο JSTOR παραθέτει το αρχικό συγκεκριμένο άρθρο. Πρόσβαση: <http://about.jstor.org>.

Η αναζήτηση πραγματοποιείται με τη χρήση του Citation Locator, μια επιλογή κάτω από την επικεφαλίδα αναζήτησης.

Η αναζήτηση πραγματοποιείται με τη χρήση πεδίων του συγγραφέα ή του τίτλου.

Εάν χρειάζεται, περικλείουμε τις λέξεις του τίτλου σε εισαγωγικά για να αναγκάσουμε την αναζήτηση να ψάξει ακριβώς τη φράση. Κάνουμε κλικ στον τίτλο του άρθρου ή της περίληψης του. Ο αριθμός των τεκμηρίων που επικαλούνται το άρθρο αυτό στο JSTOR θα είναι στη δεξιά στήλη και στο κάτω μέρος της περίληψης. Υπάρχει επίσης μια επιλογή για να δούμε τα αναφερόμενα άρθρα στο Google Scholar. Αν δεν υπάρχουν αναφερόμενα άρθρα στο JSTOR, δεν θα εμφανιστεί καταμέτρηση.

2.4.8. PubMed / Medline

Η Medline είναι η μεγαλύτερη ιατρική βάση δεδομένων στον κόσμο (PubMed Overview 2006). Προέρχεται από την Εθνική Βιβλιοθήκη Ιατρικής (NLM-National Library of Medicine) των ΗΠΑ και καλύπτει τους τομείς της Ιατρικής, Νοσηλευτικής, Οδοντιατρικής, την Κτηνιατρική, το σύστημα υγείας, καθώς και τις προκλινικές επιστήμες. Η Medline περιλαμβάνει βιβλιογραφικές αναφορές και περιλήψεις από περισσότερα από 4.800 περιοδικά βιοϊατρικής. Η βάση δεδομένων, επί του παρόντος, περιέχει πάνω από 14 εκατομμύρια αναφορές, που χρονολογούνται από τα μέσα του 1960. Η κάλυψη είναι παγκόσμια, αλλά τα περισσότερα αρχεία είναι από πηγές στην αγγλική γλώσσα ή έχουν μια αγγλική περίληψη. Πολλή δουλειά έχει γίνει επίσης στην ευρετηρίαση των αναφορών του άρθρου με το ελεγχόμενο λεξιλόγιο MeSH (Medical Subject Headings). Το PubMed έχει τη δική του διεπαφή αναζήτησης για την Medline. Εκτός από την Medline αυτό ανακτά:

"OLDMEDLINE" αναφορές για προ του 1966 εγγραφές

Αναφορές σε άρθρα που είναι εκτός του ερευνητικού του ενδιαφέροντος (π.χ. Αστροφυσική) από ορισμένα περιοδικά MEDLINE, επίσης περιοδικά κυρίως γενικά περί Επιστήμης και Χημείας, για τα οποία οι δημοσιεύσεις τους στις επιστήμες της ζωής ευρετηριοποιούνται στη βάση MEDLINE.

Αναφορές σε επεξεργασία που καταγράφονται για μία δημοσίευση πριν από την ευρετηρίαση της στην MeSH και προστέθηκαν στη Medline ή μεταφέρθηκαν σε κατάσταση εκτός ερευνητικού ενδιαφέροντος.

Αναφορές που προηγούνται της ημερομηνίας που ένα περιοδικό επιλέχθηκε για τη βάση ευρετηρίασης MEDLINE (όταν παρέχονται ηλεκτρονικά από τον εκδότη).

Ορισμένα περιοδικά των επιστήμων της ζωής που υποβάλλουν πλήρες κείμενο σε PubMed Central και μπορεί να μην έχουν προταθεί για ένταξη στο MEDLINE, αν και έχουν υποστεί κριτική από την NLM και μερικά περιοδικά Φυσικής που ήταν μέρος ενός πρωτοτύπου PubMed στις αρχές έως τα μέσα της δεκαετίας του 1990. Η διαφορά μεταξύ MEDLINE και PubMed έγκειται στο γεγονός ότι η Medline δεν περιλαμβάνει καμία πληροφορία σχετική με αναφορές.

2.4.8.1. Οδηγίες χρήσης

Αναζήτηση αναφερόμενου τεκμηρίου με βάση τον συγγραφέα ή τον τίτλο του άρθρου.

Κάντε κλικ στον τίτλο του τεκμηρίου από τη λίστα αποτελεσμάτων.

Στη δεξιά στήλη, αναζητήστε το κουτί με την ένδειξη "Cited by X PubMed Central articles", όπου X είναι ο αριθμός των φορών που παρατίθενται.

Σημείωση: Εάν η δεξιά στήλη δεν έχει το κουτί "Cited by", υποθέτουμε ότι δεν υπάρχουν στην PubMed άρθρα που έχουν επικαλεστεί το συγκεκριμένο τεκμήριο.

2.4.9. Academic Search Premier

Αυτή η βάση δεδομένων της εταιρείας του EltonB. Stephens Co. (EBSCO) έχει μερικά εργαλεία που επιτρέπουν στους ερευνητές να αναζητήσουν αποτελεσματικά αναφορές. Τα εργαλεία αυτά περιλαμβάνουν τα: Βρείτε παρόμοια αποτελέσματα, Αναφορές που παρατίθενται και Αντιστοίχιση αναφορών. Το Βρείτε παρόμοια αποτελέσματα χρησιμοποιεί για τη λειτουργία του Αναζήτηση Έξυπνου Κειμένου προκειμένου να βρει άλλα έγγραφα, που σχετίζονται με το αρχικό άρθρο. Με τη διεξαγωγή αναζήτησης στο τμήμα Αναφορές που παρατίθενται, ο χρήστης λαμβάνει μια λίστα των άρθρων που παραπέμπουν στο αρχικό άρθρο. Τοποθετημένο κάτω από την επιλογή Περισσότερα χαρακτηριστικά στο Academic Search Premier (πάνω στη γραμμή μενού), η λειτουργία Αντιστοίχιση παραπομπών επιτρέπει στους ερευνητές να ψάξουν για συγκεκριμένα έργα.

Πρόσβαση: <http://www.ebscohost.com/academic/academic-search-premier>.

2.4.10. Άλλες βάσεις δεδομένων

Με περιορισμένες δυνατότητες

Citeseer (1997). Η CiteSeer είναι μια εξελισσόμενη επιστημονική ψηφιακή βιβλιογραφική βιβλιοθήκη και μηχανή αναζήτησης που εστιάζει, κατά κύριο λόγο, στη βιβλιογραφία με επίκεντρο τον υπολογιστή και την επιστήμη της πληροφορίας.

2.4.11. Ειδικές βάσεις δεδομένων

PsycINFO.

Παράγεται από την American Psychological Association, το PsycINFO έχει την ένδειξη Βρείτε Παραπομπές ως μία από τις επιλογές αναζήτησης στην κεντρική σελίδα αναζήτησης. Αυτό το εργαλείο λειτουργεί αποτελεσματικά, όταν ο ερευνητής μπορεί να εισάγει στην αναζήτηση λεπτομερείς πληροφορίες της ζητούμενης αναφοράς. Πρόσβαση: <http://www.apa.org/psycinfo>.

Chemical Abstracts:

Μια ειδική βάση δεδομένων της Φυσικής και της Χημείας που δημιουργήθηκε από την Αμερικανική εταιρεία, την Chemical Abstracts Services, για την Αμερικανική Χημική Κοινότητα και καταγράφει κατά μέσο όρο περίπου 500.000 αναφορές το χρόνο, που λαμβάνονται από περίπου 10.000 περιοδικά.

Compendex:

Μια ειδική βάση δεδομένων για τη μηχανική και την τεχνολογία που δημιουργήθηκε από μια αμερικανική εταιρεία, την Engineering Information και που στη συνέχεια πουλήθηκε από αυτήν στην Elsevier το 1998. Καταγράφει κατά μέσο όρο περίπου 150.000 αναφορές το χρόνο, που λαμβάνονται από περίπου 4.500 επιστημονικά περιοδικά.

Embase:

Μια ειδική βάση δεδομένων της Ιατρικής επιστήμης που παράγεται από μια Ολλανδική εταιρεία, την Excerpta Medica και καταγράφει κατά μέσο όρο περίπου 250.000 αναφορές το χρόνο, που λαμβάνονται από περίπου 3.500 περιοδικά.

Inspec:

Μια ειδική βάση δεδομένων φυσικών επιστημών που παράγεται από το Ινστιτούτο Ηλεκτρολόγων Μηχανικών του Ηνωμένου Βασιλείου (Institution of Electrical Engineers IEE) και καταγράφει κατά μέσο όρο περίπου 200.000 αναφορές το χρόνο, που λαμβάνονται από περίπου 2.200 περιοδικά.

Pascal:

Μια γενική βάση δεδομένων που καλύπτει διάφορους τομείς και παράγεται από το Ινστιτούτο Επιστημονικών και Τεχνικών Πληροφοριών (INIST) στο Εθνικό Κέντρο Επιστημονικής Έρευνας Γαλλίας (CNRS) και καταγράφει κατά μέσο όρο περίπου 450.000 αναφορές σε ένα χρόνο, που λαμβάνονται από περίπου 6.000 περιοδικά.

Digital Bibliography & Library Project DBLP:

Μια βιβλιογραφική βάση δεδομένων σχετική με την επιστήμη των υπολογιστών που κατασκευάστηκε και φιλοξενείται στο Πανεπιστήμιο του Trier, στη Γερμανία.

2.4.12. Προβλήματα που τίθενται εν γένει από τις βάσεις δεδομένων

Η επιλογή μιας βάσης δεδομένων για τη συγκέντρωση βιβλιομετρικών δεικτών εξαρτάται άμεσα από τους επιδιωκόμενους στόχους και τα ερωτήματα στα οποία η βάση καλείται να απαντήσει. Κάθε βάση δεδομένων έχει το δικό της περιεχόμενο και τα δικά της κριτήρια εισαγωγής όρων σε αυτή. Δεν δύναται, συνεπώς, να υπάρχουν δύο βάσεις πανομοιότυπες. Σε οποιοδήποτε δεδομένο θέμα, η ποσότητα των άρθρων (ή άλλων μονάδων μέτρησης) θα ποικίλει, ανάλογα με τη βάση δεδομένων που χρησιμοποιείται. Η ποιότητα (όπως π.χ. η ακριβής ανάλυση ανά επιστημονικό κλάδο) θα διαφέρει επίσης, μεταξύ άλλων, σύμφωνα με τα περιοδικά από τα οποία αντλούνται δεδομένα. Για την ίδια μελέτη, η ποικιλομορφία αυτή μπορεί να δώσει διαφορετικά αποτελέσματα, καθιστώντας επιτακτική την ανάγκη αναζήτησης των στοιχείων εκείνων που είναι, όσο το δυνατόν, συνεκτικότερα. Όταν τα αποτελέσματα διαφέρουν ανάλογα με τις πηγές που χρησιμοποιούνται, δεν υπάρχει αντικειμενικός τρόπος διαχωρισμού τους και εξακρίβωσης του ποια από αυτά απεικονίζουν με μεγαλύτερη ακρίβεια την πραγματικότητα γύρω από την επιστημονική παραγωγή. Όλοι οι χρήστες των βιβλιομετρικών δεικτών, ως εκ τούτου, πρέπει να ξεκινούν από την επιλογή των βάσεων δεδομένων που ανταποκρίνονται καλύτερα στις ιδιαίτερες ανάγκες τους. Για να το επιτύχουν αυτό, θα πρέπει πρώτα να αναλύσουν τα πλεονεκτήματα, τις αδυναμίες και τους περιορισμούς των διάφορων βάσεων δεδομένων.

Για βιβλιομετρικές μελέτες επίπεδου “macro”, οι βάσεις που επιλέγονται πρέπει να είναι αντιπροσωπευτικές, αλλά δεν είναι απαραίτητο να καλύπτουν όλα τα δεδομένα. Με συνδυασμό ή παραγοντοποίηση πληροφοριών από διάφορες πηγές (βάσεις δεδομένων) στη διαδικασία ανάλυσης, ο κίνδυνος του να μην είναι αντιπροσωπευτικό ή πλήρες το τελικό αποτέλεσμα μπορεί πάντα να ελαχιστοποιηθεί.

Επιπλέον τα βιβλιογραφικά αρχεία, που αποτελούν τις διάφορες βάσεις δεδομένων, δημιουργήθηκαν για την παροχή πληροφοριών μίας, πρωτίτως, ποιοτικής φύσεως και όχι για να λειτουργούν ως βάση για κάθε είδος καταγραφή δημοσίευσης ή άρθρου. Για το λόγο αυτό κρίνεται αναγκαίο, αφού εξαχθούν τα δεδομένα, να υφίστανται ειδική επεξεργασία, ώστε τελικά να μπορούν να χρησιμοποιηθούν κατά την παραγωγή των δεικτών. Για το σκοπό αυτό, οι βιβλιομετρητές έχουν αναπτύξει μια ποικιλία μεθόδων επεξεργασίας και εφαρμογής τους κατά την αναλυτική εργασία τους.

Προκειμένου να μελετηθεί η ανάπτυξη της επιστήμης σε έναν συγκεκριμένο επιστημονικό κλάδο, είναι απαραίτητο να προταθεί ή να πραγματοποιηθεί η συνάθροιση του επιστημονικού πεδίου, επειδή οι βάσεις δεδομένων δεν κατατάσσονται αυτόματα ανά ειδικότητα. Εάν το πεδίο που θα αναλυθεί είναι π.χ. η «Χημεία», οι αναλυτές θα μπορούσαν να κατατάξουν την αντίστοιχη βάση δεδομένων με βάση τα ακόλουθα κριτήρια:

- 1) τους τίτλους των εξειδικευμένων περιοδικών στον τομέα (όπως προσδιορίζονται από τους ίδιους τους αναλυτές).
- 2) τις λέξεις κλειδιά που συνδέονται με την έρευνα τους (η επιλογή επαφίεται και πάλι στους ίδιους τους αναλυτές).
- 3) μια επίλεκτη λίστα των περιοδικών που χρησιμοποιούνται από τους ειδικούς.

Ανάλογα με το ρυθμό κατάταξης, τα αποτελέσματα ποικίλουν. Το να προσδίδεις μια έρευνα σε ένα συγκεκριμένο επιστημονικό πεδίο ή σε έναν επιστημονικό κλάδο είναι μία λεπτή και ιδιαίτερα απαιτητική αποστολή, στο βαθμό που ένα πρόγραμμα αναζήτησης μπορεί κάλλιστα να συγκεντρώνει μια ποικιλία επιστημονικών κλάδων.

2.5. Τι μετρούν οι βιβλιομετρικές αναλύσεις;

Οι Βιβλιομετρικές αναλύσεις οδηγούν σε ερευνητικούς δείκτες ποσότητας και απόδοσης. Μπορούν, επίσης, να παρέχουν μετρήσεις των συνδέσεων μεταξύ των ερευνητών και των ερευνητικών περιοχών, μέσω της στατιστικής ανάλυσης των συν-δημοσιεύσεων και των αναφορών. Η Βιβλιομετρία χρησιμοποιεί τέσσερα κύρια είδη δεικτών:

Α) Δείκτης Παραγωγικότητας (productivity).

Αριθμός δημοσιεύσεων: Ο αριθμός των άρθρων που δημοσιεύονται στα επιστημονικά περιοδικά κατά τη διάρκεια ενός συγκεκριμένου χρονικού πλαισίου αποτελεί δείκτη της παραγωγής σε σύνολο ή υποσύνολο εντός του συστήματος της επιστήμης. Είναι επίσης δυνατό να συγκρίνουμε τους αριθμούς, προκειμένου να εκτιμηθεί η ένταση της παραγωγής σε συγκεκριμένους τομείς.

Β) Δείκτης Απήχησης (impact).

Αναφορές και παράγοντες επιρροής: Ο Αριθμός των αναφορών μπορεί να χρησιμοποιηθεί για να αξιολογηθούν οι επιστημονικές επιπτώσεις της έρευνας. Ο αριθμός των αναφορών που λαμβάνονται από γνωστά περιοδικά καταρτίζεται συστηματικά από την Thomson ISI και πωλείται με την εμπορική ονομασία Journal Citation Reports (JCR). Αυτό το προϊόν περιλαμβάνει μια σειρά από δείκτες που συνδέονται με αναφορές που λαμβάνονται από τα περιοδικά, καθώς αποτελεί τον καθοριστικό παράγοντα που εφαρμόζεται πιο συχνά.

Γ) Δείκτης Αποδοτικότητας (efficiency).

Αναφορές ανά δημοσίευση: Υπολογίζεται ως το ποσοστό (%) του αριθμού των αναφορών που έλαβαν οι επιστημονικές δημοσιεύσεις.

Δ) Δείκτες Υβριδικοί (hybrid: productivity+ impact).

Συνεργασία αναφορών και ανάλυση συνδυασμού λέξεων, h-index. Πολλοί δείκτες που βασίζονται σε αναφορές χρησιμοποιούνται για τη χαρτογράφηση της ερευνητικής δραστηριότητας: ανάλυση συν-αναφορών, ανάλυση ζευγών λέξεων και βιβλιογραφική σύζευξη. Η χαρτογράφηση είναι ένα μέσο για τη μελέτη της ανάπτυξης στους αναδυόμενους τομείς που χρησιμοποιούν το χρόνο ως μεταβλητή. Οι δείκτες συν-αναφοράς μπορούν να συνδυαστούν με τη δημοσίευση και την αναφορά για την κατασκευή πολύπλευρης αναπαράστασης των ερευνητικών πεδίων, της ανάπτυξης δεσμών μεταξύ τους και των συμμετεχόντων που τους διαμορφώνουν.

2.6. Χρήση των βιβλιομετρικών δεικτών και οι προφυλάξεις που θα πρέπει να ληφθούν

2.6.1.Εισαγωγή

Η βιβλιομετρική ανάλυση χρησιμοποιεί πολλές παραμέτρους, όπως είναι η επιστημονική παραγωγή (άρθρα κ.λ.π.), το σύνολο των συντακτών που συνυπογράφουν τα άρθρα, τα διπλώματα ευρεσιτεχνίας, οι αναφορές, οι συν-αναφορές και τα ζεύγη λέξεων. Αυτές οι παράμετροι αποτελούν έμμεσα μέτρα υπολογισμού της

επιστημονικής κοινότητας, της δομής και της παραγωγής της. Παραδείγματα για το πώς χρησιμοποιούνται οι δείκτες αυτοί μπορούν να βρεθούν στο Κεφάλαιο 3.

Τα βιβλιομετρικά δεδομένα και η ανάλυση αυτών παρέχουν πληροφορίες αναφορικά με τον επιστημονικό προσανατολισμό και τη δυναμική μιας χώρας (ή κάποια άλλης μονάδας / συνόλου), καθώς και τη συμμετοχή της στην επιστήμη και την τεχνολογία σε παγκόσμιο επίπεδο, με άλλα λόγια, παρέχει πληροφορίες για τη συμβολή και την επιρροή που ασκεί μια χώρα, τόσο σε εθνικό επίπεδο όσο και στο επίπεδο της διεθνούς κοινότητας. Η ανάλυση που βασίζεται στη συνεργασία καθιστά δυνατό τον προσδιορισμό και την αντιπροσώπευση των επιστημονικών δικτύων και, παράλληλα, επισημαίνει τους δεσμούς μεταξύ των χωρών, των θεσμών και των ερευνητών, καθώς και τις επιπτώσεις των μεγάλων προγραμμάτων [CERN (Ευρωπαϊκό Εργαστήριο Σωματιδιακής Φυσικής), WHO (Παγκόσμιος Οργανισμός Υγείας - ΠΟΥ), κ.λ.π.]. Η Βιβλιομετρία επισημαίνει επίσης, τη δομή των επιστημονικών κλάδων και των μεταξύ τους δεσμών. Τα βιβλιομετρικά δεδομένα και οι βιβλιομετρικοί δείκτες μπορούν να λειτουργήσουν ως εργαλεία ή τουλάχιστον ως συμπληρωματική βοήθεια, για την περιγραφή και τη διατύπωση ερωτημάτων που τίθενται στον κόσμο της επιστήμης.

Όπως άλλωστε συμβαίνει και σε άλλους τομείς, είναι, και εν προκειμένω, σημαντικό να τονίσουμε ότι οι δείκτες που λαμβάνονται από βιβλιομετρικές βάσεις δεδομένων θα πρέπει να παρατίθενται κατ' αναλογία. Οι δείκτες βασίζονται σε μια συγκριτική προσέγγιση, απόλυτες τιμές δεν είναι εξ ορισμού ενδεικτικές, αλλά προσλαμβάνουν την πλήρη σημασία τους μόνο σε σύγκριση με εκείνες των άλλων ομάδων.

Επιπλέον η ανάλυση πρέπει να ενσωματώνει όσο το δυνατόν μεγαλύτερο όγκο δεδομένων, έτσι ώστε να επιτρέπει τη στατιστική αντιστάθμιση για οποιαδήποτε προκατάληψη, η οποία είναι σε θέση να επηρεάσει κάθε μικρή επιχείρηση που λαμβάνεται ξεχωριστά.

Οι περιορισμοί των δεδομένων που χρησιμοποιούνται στη Βιβλιομετρία προέρχονται κυρίως από τα διάφορα μέσα επικοινωνίας που χρησιμοποιούν οι επιστήμονες για να μεταφέρουν πληροφορίες μεταξύ τους, πέραν του συνήθη δίαυλου των επιστημονικών περιοδικών. Επιπλέον η προφορική επικοινωνία μεταξύ των επιστημόνων δεν περιορίζεται στα στατιστικά στοιχεία, ούτε οι εσωτερικές εκθέσεις μεταξύ πανεπιστημίων, εργαστηρίων και ερευνητικών ομάδων και οι εκθέσεις μεταξύ των χωρών που συνεργάζονται μέσω επιτροπών, προγραμμάτων ή εργαστηρίων. Υπάρχει, επίσης, και η ηλεκτρονική επικοινωνία μεταξύ των ερευνητών, η οποία αναπτύσσεται με ταχύτατους ρυθμούς. Όλες οι μορφές επικοινωνίας, που καλύπτονται από τις «παραδοσιακές» βιβλιομετρικές μεθόδους, αποτελούνται συνεπώς από ανταλλαγές που έχουν «επισημοποιηθεί». Κάποια ανεπίσημη ή άτυπη ανακοίνωση δεν έχει ενσωματωθεί και πιθανότατα δεν θα ενσωματωθεί ποτέ.

Η παραδοσιακή προσέγγιση είναι ακόμη πιο περιοριστική σε σχέση με οτιδήποτε περιλαμβάνει βιομηχανική έρευνα ή έρευνα σχετιζόμενη με την άμυνα. Υπάρχουν μεγάλες χρονικές υστερήσεις στην επικοινωνία μεταξύ επιστήμης (πρωτίστως ακαδημαϊκά) και βιομηχανίας, λόγω της επιθυμίας, από πλευράς βιομηχανίας να προστατεύει τις ανακαλύψεις της και λόγω του γεγονότος ότι τα δημοσιευμένα πορίσματα παρουσιάζουν γενικά συντομευμένη μορφή. Τα άρθρα που δημοσιεύθηκαν από βιομηχανικά εργαστήρια παρέχουν σκόπιμα μια περιορισμένη εικόνα των στόχων της έρευνας, οι οποίοι είναι, κατά κανόνα, η δημιουργία προϊόντων ή διαδικασιών που υπόκεινται σε εμπορικό ανταγωνισμό.

Επιπλέον ένα μεγάλο μέρος των έρευνας που σχετίζεται με την άμυνα (η οποία συχνά συνδέεται με τη βιομηχανική έρευνα) δεν συμπεριλαμβάνεται ποτέ στη συνήθη επιστημονική επικοινωνία, παρά την τεχνολογική σημασία του και το γεγονός ότι τείνει να είναι στην κορυφή της βασικής έρευνας.

Το πρόβλημα της «συνολικής καταμέτρησης» ή της «κλασματικής καταμέτρησης»

Μία από τις ασάφειες που παρουσιάζει η βιβλιομετρική μέθοδος είναι η πολυμορφία των μεθόδων μέτρησης. Για παράδειγμα η ταξινόμηση των άρθρων που έχουν συνταχθεί από περισσότερους από έναν συγγραφείς αποτελεί εδώ και καιρό ένα σημαντικό θέμα συζήτησης στον κύκλο των βιβλιομετρητών (Martin, 1991; Braun et al., 1991; Leydesdorff, 1991; Kealey, 1991). Πώς είναι δυνατόν να μετράται και να υπολογίζεται η συμμετοχή των δημιουργών σε ένα επιστημονικό έργο, τη στιγμή που πρόκειται για μία διαδικασία που, εκ φύσεως, βασίζεται στη συνεργασία; Μπορούν όλα τα άτομα που έχουν συμβάλει, να λάβουν «πλήρως τα εύσημα» για τη συμμετοχή τους ή μήπως, αν ένα άρθρο είναι γραμμένο με εννέα άλλους συγγραφείς, για παράδειγμα, θα δικαιούται ο καθένας μόνο το ένα δέκατο της συνολικής αναγνώρισης του έργου; Άραγε, μια χώρα που έχει δεσμευτεί να αποτελεί μέλος μιας συνεργασίας τριών χωρών, συνιστά «σύνδεσμο» («σύνολο») ή μόνο ένα ποσοστό, εν προκειμένω το «ένα τρίτο του συνδέσμου»;

Στην πράξη, όταν ένα άρθρο συνοπογράφεται από ερευνητές που προέρχονται από διαφορετικές χώρες, οι βιβλιομετρητές εφαρμόζουν, κατά βάση, δύο τρόπους αναγνώρισης των ενδιαφερόμενων χωρών:

- Κάποιοι δίνουν την πλήρη αναγνώρισή τους (βαθμολόγηση/εύσημα), δηλαδή υπολογίζουν πως κάθε χώρα που έχει συμμετάσχει στη συγγραφή ενός άρθρου παίρνει «1» βαθμό (/πόντο). Πρόκειται για τη συνολική μέθοδο καταμέτρησης.

- Άλλοι χωρίζουν σε τμήματα τη συν-δημιουργία, βάση του αριθμού των χωρών προέλευσης των συγγραφέων που έχουν συμβάλει σε αυτή και δίνεται κλασματικά σε κάθε χώρα ο βαθμός (/ο πόντος) αναγνώρισής της (κλασματική μέθοδος καταμέτρησης). Η μέθοδος αυτή βασίζεται στη «μαθηματική λογική»: προκειμένου να προκύψει ένας τελικός αριθμός επί τοις εκατό, οι βαθμοί (τα εύσημα ουσιαστικά) κάθε χώρας πρέπει να μοιράζονται και το «1» πρέπει να διαιρείται με τον αριθμό των χωρών προέλευσης των συμβαλλόμενων συγγραφέων που συναντούμε σε κάθε διεθνή δημοσίευση, προερχόμενη από τη συνεργασία πολλών συντελεστών.

Κάθε μέθοδος καταμέτρησης έχει σίγουρα τη δική της λογική, αλλά ο τρόπος με τον οποίο κατανέμεται η βαθμολογία θα πρέπει να υπόκειται στον πλήρη έλεγχο των βιβλιομετρητών και να είναι κατανοητός από τους ανθρώπους που χρησιμοποιούν το προϊόν τους.

Για τους επιστήμονες και τους πολιτικούς που χρησιμοποιούν τους επιστημονικούς δείκτες, η συνολική μέθοδος καταμέτρησης είναι πολύ πιο κατανοητή και ερμηνεύεται πολύ πιο εύκολα. «Ένα ποσοστό της τάξεως του 10% (μιας χώρας) σημαίνει, υπό αυτή την έννοια, ότι για κάθε 100 εργασίες δημοσιευμένες ανά τον κόσμο, οι 10 έχουν τουλάχιστον από έναν συντελεστή συμβολής στη συγγραφή τους (που προέρχεται από τη χώρα αυτή). Είναι δύσκολο να εξηγηθεί, ωστόσο, αυτή η ‘έννοια’ του μεριδίου της τάξεως του 10% στην κλασματική κλίμακα, η οποία μπορεί να είναι το αποτέλεσμα της προσθήκης δέκα ή περισσότερων εργασιών» (Braun et al., 1991).

Ακόμη πιο σημαντικό είναι το γεγονός ότι η κλασματική καταμέτρηση αποδίδει μικρότερη αξία στη διεθνή συνεργασία για τη σύνταξη ενός άρθρου σε σύγκριση με τη μονομερή σύνταξη ενός εθνικού άρθρου, όταν υπολογίζεται η απόδοση σε εθνικό επίπεδο. Όσο περισσότερους διεθνείς εταίρους έχει ένα άρθρο, τόσο λιγότερη βαθμολογία λαμβάνει κάθε συμβαλλόμενη στη συγγραφή του χώρα. Γιατί θα πρέπει, στα στατιστικά στοιχεία της βιβλιομετρικής ανάλυσης, να αξιολογηθεί μια χώρα υψηλότερα για ένα άρθρο από κάποιον εθνικό της συγγραφέα, σε σύγκριση με ένα άρθρο που έχει συνταχθεί από συγγραφείς διαφορετικών εθνικοτήτων, στο πλαίσιο της διεθνούς συνεργασίας (Leydesdorff, 1991); Για αυτούς ακριβώς τους λόγους, κάποιοι βιβλιομετρητές ισχυρίζονται ότι η κλασματική καταμέτρηση είναι μια υποδεέστερη διαδικασία. Ιδίως όταν ο όγκος των δεδομένων είναι μεγάλος, υποστηρίζουν ότι «η ίση καταμέτρηση όλων των δημιουργών - συγγραφέων είναι, στις περισσότερες περιπτώσεις, η καλύτερη λύση» (VanRaaij and Tijssen, 1990).

Είναι δύσκολο να επιλέξει κανείς μία εκ των δύο μεθόδων. Ωστόσο, εφ’ όσον εφαρμόζονται και οι δύο, η ανάλυση της επιστήμης και της τεχνολογίας σε εθνικό ή εργαστηριακό επίπεδο ενδέχεται να ποικίλλει και, ενίοτε, να παράγει αντιφατικά αποτελέσματα. Η προκατάληψη ότι οι μετρήσεις που συσσωρεύονται είναι άλλοτε υπέρ και άλλοτε εις βάρος ορισμένων χωρών μπορεί να κάνει τις διεθνείς συγκρίσεις προβληματικές.

Το πρόβλημα της κάλυψης των βάσεων δεδομένων

Άλλες συζητήσεις έχουν ως επίκεντρό τους μια θεμελιώδη πτυχή των βιβλιομετρικών μετρήσεων, πώς είναι δυνατόν να μετρηθούν οι τάσεις δημοσίευσης επιστημονικού υλικού, επί σειρά ετών, με τη χρήση βάσεων δεδομένων που εξελίσσονται από το ένα έτος στο επόμενο; Εκ των υστέρων, αυτό το πρόβλημα, το οποίο συνδέεται με την εξέλιξη της Βιβλιομετρίας, φαίνεται να είναι το αποτέλεσμα των προσπαθειών που έχουν ως απώτερο σκοπό το να καταστούν τα βιβλιομετρικά εργαλεία περισσότερο αντιπροσωπευτικά. Κάποια επιστημονικά περιοδικά εξαφανίζονται, άλλα αλλάζουν ονομασία ή συγχωνεύονται, και, πάνω απ’ όλα, νέα περιοδικά κάνουν την εμφάνισή τους. Το Ινστιτούτο για την Επιστήμη της Πληροφόρησης (ISI) παρακολουθεί όλες αυτές τις μεταβολές και ανανεώνει τακτικά τον κατάλογο των περιοδικών που καλύπτονται από την SCI, δημιουργώντας έναν ετήσιο κύκλο εργασιών της τάξεως περίπου του 7% (Garfield, 1979a). Το 1964, το Ινστιτούτο είχε συμπεριλάβει περίπου 610 περιοδικά στη βάση δεδομένων SCI, ενώ μέχρι το 1981, ο αριθμός αυτός είχε ανέλθει στα 3.600 περιοδικά. Έχει σημειωθεί επίσης μια παράλληλη αύξηση του αριθμού των άρθρων. Από τα 100.000 το 1964, φτάσαμε στις 500.000 το 1981.

Ωστόσο, η αύξηση αυτή μπορεί να αποτελέσει πρόβλημα για κάποιους, όταν γίνονται προσπάθειες παρακολούθησης των εθνικών επιστημονικών επιδόσεων διαχρονικά. Για την καταμέτρηση της δημοσίευσης του επιστημονικού υλικού, μπορεί να φαίνεται προτιμότερο να παραμένουν σταθερά ο αριθμός και η προετοιμασία των περιοδικών, προκειμένου η μέτρηση και ο υπολογισμός της κατάστασης της επιστήμης σε μια δεδομένη χώρα να είναι συγκρίσιμα από το ένα έτος στο άλλο. Η αύξηση του αριθμού των αναφορών σε μια χώρα θα μπορούσε απλώς να είναι το αποτέλεσμα της προσθήκης νέων περιοδικών και όχι ένα μέτρο μέτρησης και απεικόνισης της πραγματικής αύξησης της παραγωγικότητας. Τα δεδομένα πρέπει να

υποβάλλονται σε επεξεργασία κατά τρόπο που να καθίστανται συγκρίσιμα από το ένα έτος στο άλλο (Anderson et al., 1989). Για να καταστεί ευκολότερη η ερμηνεία των δεδομένων, αποφασίστηκε να παρακολουθείται ένας σταθερός αριθμός επιστημονικών περιοδικών, που συνιστά το «καθορισμένο (σταθερό) σύνολο επιστημονικών περιοδικών» και αντιπροσωπεύει περίπου 2.100 δημοσιεύσεις, που παρακολουθούνταν και ελέγχονταν από το Ινστιτούτο για την Επιστήμη της Πληροφόρησης (ISI), στη διάρκεια του 1973. Επομένως ο κατάλογος των περιοδικών παρέμενε «σταθερός» μεταξύ του 1973 και του 1980. Ο κατάλογος των επιλεγμένων επιστημονικών περιοδικών αναθεωρήθηκε εκ νέου το 1981, ώστε να αντικατοπτρίζει τις νέες εξελίξεις στην επιστήμη. Νέοι τίτλοι, που ενσωματώθηκαν στη βάση δεδομένων SCI, δεν προστέθηκαν στο «καθορισμένο (σταθερό) σύνολο επιστημονικών περιοδικών», κατά τη διάρκεια των «παγωμένων» -όπως αποκαλούνται- περιόδων (1973-1980 και 1981-1986). Ωστόσο, τα περιοδικά που διαγράφηκαν από τη βάση δεδομένων SCI, αποκλείστηκαν και από τη βάση.

Είναι αναπόφευκτο ότι ο περιορισμός των αναφορών σε συγκεκριμένο αριθμό περιοδικών θα εκπροσωπούσε ελλιπώς τη φυσική δυναμική της γνώσης και αυτό το «πάγωμα» θα μπορούσε να δημιουργήσει έναν «τεχνητό» επιστημονικό κόσμο. Ακόμη χειρότερα, ενδέχεται νέα πεδία έρευνας, όπως η υπεραγωγιμότητα ή το AIDS, τα οποία αποτελούν θέμα συζήτησης σε περιοδικά που δημιουργούνται ως αποτέλεσμα των σχετικών ευρημάτων, να μη μελετώνται άμεσα. Η απόδοση των εκβιομηχανισμένων χωρών αντιπροσωπεύεται καλύτερα στο «μη παγωμένο σύνολο περιοδικών», επειδή οι επιστήμονες σε αυτές τις χώρες έχουν την τάση να επιλέγουν νέους τίτλους και όχι τα παραδοσιακά επιστημονικά περιοδικά, για την παρουσίαση του έργου τους. Ακριβώς αυτά τα καινούργια περιοδικά δημιουργούν ένα πρόσφορο έδαφος, όπου τα νέα πεδία έρευνας κάνουν το ντεμπούτο τους (Kealey, 1991). Συνεπώς «η αναπαράσταση της έρευνας (μέσω της σταθερής βάσης περιοδικών) θεωρείται εξαιρετικά συντηρητική και στατική» (Callon and Leydesdorff, 1987).

Και εδώ, λοιπόν, υπάρχουν δύο διαφορετικοί τρόποι επεξεργασίας των δεδομένων της SCI, προκειμένου να μετρηθούν οι επιστημονικές επιδόσεις:

- i) με βάση ένα μη σταθερό σύνολο περιοδικών, με ποσοτικές διαφορές ή
- ii) με βάση ένα σταθερό σύνολο περιοδικών, που μειώνεται διαχρονικά.

Μόνο τα σημεία αφετηρίας (δηλαδή τα έτη 1973 και 1981) παρουσιάζουν σταθερότητα. Τα μεταγενέστερα αποτελέσματα δεν εξελίσσονται κατά τον ίδιο τρόπο. Αυτό που είναι σημαντικό, είναι να γνωρίζουμε ποια από τις υποκατηγορίες είναι η πιο αντιπροσωπευτική ή ποια παράγει τους πιο αξιόπιστους δείκτες για τη μέτρηση των εθνικών επιδόσεων στον τομέα της επιστήμης (Martin, 1991). Σε ορισμένες ειδικότητες, οι αποκλίσεις είναι τόσο μεγάλες, ώστε οι τάσεις μεταβάλλονται με την πάροδο του χρόνου.

Αυτό που είναι πραγματικά περίπλοκο είναι το να μπορέσει να κατανοήσει κανείς τα αριθμητικά στοιχεία, ώστε να βγάλει κάποιο ξεκάθαρο νόημα από αυτά. Οι αριθμοί δεν «μιλούν» από μόνοι τους. Αντιθέτως, χρειάζονται ερμηνεία και περαιτέρω εξήγηση, λαμβάνοντας υπόψη τα πραγματικά και τα τεχνητά σφάλματα, που διαπιστώνονται στα δεδομένα και στη μέθοδο μέτρησης που χρησιμοποιείται η ερμηνεία δεν είναι δύσκολη.

Η σχέση του μάρκετινγκ με τα βιβλιομετρικά δεδομένα και την ανάλυση

Μια άλλη δυσκολία που αντιμετωπίζουν οι βιβλιομετρητές συνδέεται άμεσα με τις εμπορικές πιέσεις που έχουν προκύψει και που ασκούνται στον τομέα.

Μία ολόκληρη αγορά έχει στηθεί γύρω από αυτό το είδος μελέτης. Τα βιβλιομετρικά δεδομένα και η ανάλυσή τους πουλάνε. Δεν είναι, άλλωστε, σπάνιο φαινόμενο τα συμβόλαια για την εκμίσθωση, επεξεργασία και ανάλυση βιβλιομετρικών δεδομένων να περιλαμβάνουν διόλου ευκαταφρόνητα χρηματικά ποσά. Κάθε μία από τις διάφορες βιβλιομετρικές σχολές έχει αναπτύξει τη δική της μέθοδο και θεωρία. Ως εκ τούτου, τα αποτελέσματα ποικίλουν. Η αγοραστική δύναμη έχει αυξήσει τον ανταγωνισμό και, ταυτόχρονα, έχει ενθαρρύνει και τονώσει τον επαγγελματισμό που διέπει τον τομέα.

Επί του παρόντος, οι βιβλιομετρητές επιδίδονται σε έντονες συζητήσεις αναφορικά με τη μεθοδολογία. Οι διάφορες μέθοδοι είναι συχνά παρόμοιες, αλλά δεν υπάρχουν ούτε ενιαίες προδιαγραφές (Glanzel and Schoepflin, 1994) ούτε συναίνεση ως προς το ποιες συνιστούν τις καλύτερες μεθόδους ή εφαρμογές της Βιβλιομετρίας.

Το πρόβλημα της αυτοαναφοράς

Δες: 2.8.4., 3.2.1.

2.7. Τι επιφυλάσσει το μέλλον για τους βιβλιομετρικούς δείκτες;

Οι βιβλιομετρητές, λόγω του ότι βρίσκονται περιορισμένοι στην αξιολόγηση και την ανάλυση της ακαδημαϊκής έρευνας και των, μείζονος σημασίας, δημόσιων προγραμμάτων (τόσο σε εθνικό, όσο και σε διεθνές επίπεδο), κερδίζουν έδαφος και σε άλλους τομείς, ιδίως χάρη στην εξέλιξη των δεικτών για την παρακολούθηση των διαφόρων τύπων συνεργασίας (Nederlands Observatorium van Wetenschappen Technologie, 1994; MERIT, 1994; Katz and Hicks, 1996).

Οι βασικοί βιβλιομετρικοί δείκτες έχουν ακόμη πολύ δρόμο να διανύσουν. Οι βιβλιομετρητές συνεχίζουν τις προσπάθειές τους για εφαρμογή και βελτίωση των υφιστάμενων δεικτών. Ένα από τα πιο ενδιαφέροντα πεδία μελέτης περιλαμβάνει την επιστημονική και την τεχνολογική πρόβλεψη - τομέας για τον οποίο κάποιοι βιβλιομετρικοί δείκτες έχουν ήδη αναπτυχθεί (Leydesdorff, 1995; Noyon and VanRaaij, 1995). Καταβάλλονται, επίσης, προσπάθειες για την ανάπτυξη «ενοποιημένων δεικτών», δηλαδή μιας ένωσης στην οποία θα συγκεντρώνονται διάφοροι δείκτες, που αντιπροσωπεύουν επιστημονικές και τεχνολογικές δραστηριότητες (Niwa and Romizawa, 1995). Εκτός από αυτή τη μακροσκοπική προσέγγιση, προσπάθειες «μικρό - επιπέδου» επιδιώκουν να αντιπροσωπεύουν τις μετατοπίσεις της επιστήμης προς την τεχνολογία, εξετάζοντας την ανάπτυξη των δικτύων των ερευνητικών ομάδων (Hirasawa, 1995). Ένας άλλος στόχος των εντατικών αυτών προσπαθειών είναι η επέκταση της ενίσχυσης προς τις μεθόδους αξιολόγησης, με βάση την ποιότητα των επιστημονικών περιοδικών. Οι προσπάθειες αυτές θα πρέπει να αξιοποιούνται με σκοπό να επιτευχθεί ένας συγκεκριμένος στόχος: η ανάλυση των πολιτικών της επιστήμης.

Κάθε δείκτης παρουσιάζει πλεονεκτήματα και περιορισμούς. Χρειάζεται προσοχή, καθώς δε θα πρέπει να θεωρούνται ως «απόλυτοι» δείκτες, αλλά συμπληρωματικοί. Οι διάφορες βιβλιομετρικές διαδικασίες και μέθοδοι πρέπει να χρησιμοποιούνται συνδυαστικά, ακόμη και αν τα αποτελέσματά τους είναι ενίοτε αντιφατικά, για όσο χρονικό διάστημα θα προσφέρουν χρήσιμες πληροφορίες και θα συμμορφώνονται με τις επιστημονικές και τις επαγγελματικές προδιαγραφές. Παρά τους περιορισμούς της, η Βιβλιομετρία παρέχει μία ποσοτική μέτρηση της επιστημονικής παραγωγής, που είναι κατ' ουσία αντικειμενική.

2.8. Εργαλεία αναφορών και μέτρησης

2.8.1 Εργαλεία

Πρότυπα δημοσίευσης συνδημοσίευσης (publication and co-publication pattern)

Εάν τα στοιχεία δημοσίευσης που θα χρησιμοποιηθούν για την ανάλυση περιέχουν τα στοιχεία όλων των συγγραφέων (όπως τα δεδομένα από τους δείκτες αναφοράς Thomson), μπορούν να χρησιμοποιηθούν για να βρούμε τα πρότυπα της συνδημοσίευσης μεταξύ των διαφόρων συγγραφέων και μονάδων. Αυτά μπορούν να απεικονιστούν σε πίνακες συνδημοσίευσης και χάρτες που δείχνουν για παράδειγμα τους συγγραφείς, τα πανεπιστήμια και τις χώρες που δημοσιεύουν από κοινού και σε ποιο βαθμό καθώς και σε ποιες ερευνητικές περιοχές υποβάλλει εργασίες ο ερευνητής, διευκρινίζοντας πώς αυτές τον συνδέουν με άλλους συν-συγγραφείς. Επίσης εάν οι ερευνητικές περιοχές που εμφανίζονται συχνά στις εκδόσεις ενός πανεπιστημιακού τμήματος διοργανώθηκαν από ένα μεγάλο δίκτυο έρευνας του τμήματος ή πολλά μικρότερα.

Χάρτης Αναφορών (Citation Map)

Η πλατφόρμα Web of Science δημιούργησε αυτό το εργαλείο που απεικονίζει τα πριν και τα μετά της αναφοράς μιας δημοσίευσης, χρησιμοποιώντας μια μορφή χάρτη. Το πριν χαρακτηριστικό αντιπροσωπεύει τις αναφορές στο τρέχον επιλεγμένο έγγραφο και το μετά χαρακτηριστικό αντιπροσωπεύει τα έγγραφα που ανέφεραν το τρέχον επιλεγμένο έγγραφο. Το εργαλείο αυτό δίνει δυναμική αναπαράσταση των επιρροών που έχει ένα έγγραφο σε ένα πεδίο, μια θεματική περιοχή ή στην τάση. Πρόσβαση: <http://wokinfo.com>.

Εκθέσεις αναφορών (Citation Reports)

Αυτό το εργαλείο της Web of Science παρέχει μία γραφική παρουσίαση και ένα στιγμιότυπο από μία σειρά άρθρων. Κατά τη διεξαγωγή μιας αναζήτησης (ανά θέμα, συγγραφέα ή τίτλο δημοσίευσης) και επιλέγοντας μια ομάδα εγγράφων, οι χρήστες θα βρουν μια σύνδεση στο Web of Science για να δημιουργήσουν εκθέσεις αναφορών. Η έκθεση περιλαμβάνει γραφήματα για δημοσιευμένα τεκμήρια για κάθε έτος και αναφορές για κάθε έτος. Το τμήμα των αποτελεσμάτων της έκθεσης που παράγεται περιγράφει λεπτομερώς το σύνολο των αναφορών, το σύνολο των αναφορών χωρίς αυτοαναφορές, τα αναφερόμενα άρθρα, το μέσο όρο των αναφορών ανά δημοσίευση και τον h-index (βλ. παρακάτω). Η έκθεση αναφορών αποτελεί ένα χρήσιμο εργαλείο που βοηθά τους χρήστες να κατανοήσουν διάφορες τάσεις του θέματος, πώς

ένας συγγραφέας έχει δημοσιεύσει διαχρονικά και τα πιο αναφερόμενα άρθρα που σχετίζονται με το θέμα της αναζήτησης. Πρόσβαση από την ιστοσελίδα: <http://wokinfo.com>

Ιχνηλάτης αναφορών (Citation Tracker)

Αυτό το προϊόν της Scopus παρέχει λεπτομερείς πληροφορίες για την εύρεση, τον έλεγχο και την παρακολούθηση των αναφορών. Κατά τη διεξαγωγή μιας αναζήτησης και επιλέγοντας μια ομάδα εγγράφων στην Scopus, ο ιχνηλάτης αναφορών, εμφανίζεται κάτω από την σελίδα προβολής της αναφοράς, δίνοντας στους χρήστες μια επισκόπηση για το πόσες φορές ένα επιλεγμένο έγγραφο έχει αναφερθεί από την αρχική αναφορά του και τον αριθμό των εγγράφων που έχουν αναφέρει το επιλεγμένο έγγραφο από το 1996. Αυτές οι πληροφορίες επισκόπησης αναφορών βοηθούν τους ερευνητές στο να εκτιμήσουν την αξία και τον αντίκτυπο της έρευνας διαχρονικά. Πρόσβαση από την ιστοσελίδα: <http://www.info.sciverse.com/scopus>.

h-γράφημα (h-graph)

Το h-γράφημα της Scopus απεικονίζει οπτικά τον αντίκτυπο της έρευνας του συγγραφέα ή ενός συνόλου άρθρων. Με την επιλογή h-γράφημα, οι χρήστες έχουν τη δυνατότητα να δουν ή να επιλέξουν μεταξύ τριών διαφορετικών γραφημάτων: του h-index, άρθρων που δημοσιεύθηκαν και αναφορών των άρθρων αυτών. Οι χρήστες μπορούν να δουν, επίσης, αυτά τα γραφήματα είτε με το όνομα του συγγραφέα ή με τον τίτλο του άρθρου. Πρόσβαση από την ιστοσελίδα: <http://www.info.sciverse.com/scopus>.

h-δείκτης (h-index)

Πρόσβαση: Web of Knowledge σε <http://www.wokinfo.com> και Scopus στο <http://www.info.sciverse.com/scopus>

Περισσότερα για τον h-δείκτη στο 3^ο κεφάλαιο, ενότητα Βιβλιομετρικοί δείκτες και συνδυασμοί τους με παραδείγματα εφαρμογής.

2.8.2. Ανάλυση αναφορών

Όταν ένας συγγραφέας παραθέτει έναν άλλο συγγραφέα, δημιουργείται μια σχέση. Η ανάλυση αναφορών, χρησιμοποιεί τις αναφορές σε δημοσιευμένα ερευνητικά έργα από άλλα ερευνητικά έργα για να αντιστοιχίσει, όπως ήδη έχουμε αναφέρει, διασυνδέσεις μεταξύ επιστημονικών εργασιών, επιστημόνων, επιστημονικών κλάδων, επιστημονικών περιοδικών καθώς επίσης και μεταξύ του παραγόμενου επιστημονικού έργου συγκεκριμένων χωρών. Η συνήθης χρησιμοποίηση της ανάλυσης αναφορών είναι να προσδιορίσει την απήχηση ενός εκάστου επιστήμονα σε ένα δεδομένο επιστημονικό τομέα με τη μέτρηση του αριθμού των περιπτώσεων που ο συγκεκριμένος επιστήμονας έχει αναφερθεί από άλλους.

Η ανάλυση αναφορών είναι ένα σημαντικό συστατικό της βιβλιομετρίας. Η ανάλυση των βιβλιογραφικών αναφορών επηρέασε σε μεγάλο βαθμό το σύγχρονο πεδίο της ανάκτησης πληροφοριών. Ως λόγοι μελέτης των αναφορών μπορεί να αναφερθούν παράλληλα με αυτούς που αναφέρθηκαν στο Κεφάλαιο 1 οι παρακάτω:

- Η απόδοση φόρου τιμής στους πρωτοπόρους ενός επιστημονικού κλάδου.
- Ο εντοπισμός σχετιζόμενων έργων, σχέσεων μεταξύ των συγγραφέων και των έργων τους.
- Ο εντοπισμός μεθοδολογιών, εργαλείων κ.τ.λ.
- Η παροχή ώθησης σε έργα που δεν είναι πολύ γνωστά, ευρητηριασμένα ή δεν έχουν χρησιμοποιηθεί ως αναφορές.

Μπορούν να αναφερθούν αρκετά πλεονεκτήματα και μειονεκτήματα της μεθόδου ανάλυσης αναφορών. Ως πλεονεκτήματα θεωρούνται η αντικειμενικότητα των μετρήσεων, η ευκολία και το χαμηλό κόστος, η δυνατότητα μετρήσεων για έναν απεριόριστο αριθμό δημοσιεύσεων και η διαβαθμισιμότητα από το μικρο-επίπεδο (άτομα) στο μακρο-επίπεδο (παγκόσμιες μετρήσεις) (Wallin, 2005). Οι υποστηρικτές των βιβλιομετρικών μεθόδων σε σχετικές έρευνες τάσσονται υπέρ της εγκυρότητας και της αξιοπιστίας της ανάλυσης αναφορών και ισχυρίζονται πως, όπου υπάρχουν στοιχεία για τις δύο μεθόδους για το ίδιο corpus δημοσιεύσεων, διαπιστώνεται θετική συσχέτιση μεταξύ της αξιολόγησης από ομότιμους και των βιβλιομετρικών μετρήσεων. (Holmes & Oppenheim, 2001).

Ως μειονεκτήματα της μεθόδου ανάλυσης αναφορών θεωρούνται η αδυναμία ποιοτικής διαφοροποίησης των αναφορών (αρνητικές και θεωρητικές αναφορές εξομοιώνονται, συμπεριλαμβάνονται και αυτοαναφορές), η ανυπαρξία ενός συνεκτικού θεωρητικού υπόβαθρου για την ανάλυση αναφορών (MacRoberts & MacRoberts, 1996), η έλλειψη σταθερών και τυποποιημένων δεικτών (Wallin, 2005), η δυσκολία σύγκρισης μεταξύ διαφορετικών επιστημονικών πεδίων, επειδή το κάθε επιστημονικό πεδίο έχει

ιδιαίτερη δομή δημοσιεύσεων και πρότυπα αναφορών (publication and citation pattern) (Seglen, 1998). Θα πρέπει εδώ να σημειωθεί ότι επειδή οι βάσεις δεδομένων αναφορών αφορούν κυρίως δημοσιεύσεις σε περιοδικά (όχι βιβλία, κεφάλαια βιβλίων, τεχνικές αναφορές κ.λ.π.) κάθε αποτίμηση του ερευνητικού έργου με βιβλιομετρικές μεθόδους στηρίζεται στην εξής βασική αρχική υπόθεση: «κάθε επιστήμονας που έχει να πει κάτι σημαντικό, δημοσιεύει τα ευρήματά του σε επιστημονικά περιοδικά με διεθνή απήχηση» (Van Raan, 1998). Όμως στις ανθρωπιστικές επιστήμες και στις κοινωνικές, τα βιβλία και τα κεφάλαια των βιβλίων αποτελούν μια αξιολογη πρακτική δημοσίευσης επιστημονικών εργασιών. Διαφορετικά δε απ' ότι συνήθως συμβαίνει στις θετικές επιστήμες, οι κοινωνικές επιστήμες είναι συγκριτικά περισσότερο ενταγμένες στο ιδιαίτερο, εθνικό περίγραμμα μιας συγκεκριμένης κοινωνίας. Το ερευνητικό ημερολόγιο μιας κοινωνικής επιστήμης ακολουθεί εν συγκρίσει περισσότερο τις εθνικές τάσεις και επηρεάζεται περισσότερο από τα τεκταινόμενα στο εθνικό, παρά στο υπερεθνικό περίγραμμα. Επιπρόσθετα τα απορρέοντα ερευνητικά αποτελέσματα ενδιαφέρουν περισσότερο ένα εθνικό, παρά ένα υπερεθνικό κοινό. Το τελικό αποτέλεσμα, εν προκειμένω, είναι η συγγραφή των εργασιών στην εθνική γλώσσα και η δημοσίευση σε εθνικά περιοδικά να εμφανίζονται συχνότερα απ' ότι σε άλλες επιστήμες (Hicks, 1995). Κατ'αυτόν τον τρόπο δημοσιεύσεις υψηλής επιστημονικής αξίας, όπως και μέρος των αναφορών τους, κατά κανόνα δεν καταμετρώνται στις βάσεις δεδομένων με αποτέλεσμα οι βιβλιομετρικοί δείκτες στις αντίστοιχες επιστημονικές περιοχές να μην αποτυπώνουν πλήρως την πραγματικότητα.

Συμπερασματικά τα γεγονότα αυτά καθιστούν αδόκιμη την άμεση σύγκριση δεικτών μεταξύ διαφορετικών ερευνητικών πεδίων, συχνά και μεταξύ επιμέρους περιοχών του ίδιου ευρύτερου επιστημονικού πεδίου. Αφού π.χ., στην ιατρική και στη μοριακή βιολογία, παρατηρείται υψηλότερη παραγωγικότητα ως προς την ποσότητα των επιστημονικών άρθρων που δημοσιεύονται ανά έτος. Συνήθως, ο αριθμός αναφορών των δημοσιεύσεων αυτών αυξάνεται και φτάνει στο μέγιστο σε μικρό χρονικό διάστημα από τη δημοσίευσή τους. Σε αντίθεση, στις κοινωνικές επιστήμες οι ρυθμοί δημοσίευσης είναι πολύ πιο αργοί, ενώ καταγράφονται αναφορές για μεγάλο χρονικό διάστημα μετά τη δημοσίευσή τους.

Τέλος, οι MacRoberts (1996) φθάνουν στο σημείο να ισχυριστούν πως οι θεμελιώδεις προϋποθέσεις της ανάλυσης αναφορών έχουν αποδειχθεί ψευδείς, επειδή οι αναφορές στις περισσότερες δημοσιεύσεις είναι ασυνεπείς και μεροληπτικές. Τα περισσότερα όμως και μεγαλύτερα προβλήματα της ανάλυσης αναφορών σχετίζονται με τις πηγές άντλησης δεδομένων (ISI Web of Science, Scopus, Google Scholar) και αφορούν περιορισμούς στην κάλυψη των βάσεων δεδομένων, όπως έχει ήδη αναφερθεί (ευρετηριάζεται ένας μικρός αριθμός του συνόλου των επιστημονικών περιοδικών, με χρονολογικά κενά), στην ομοιογένεια κάλυψης (διαφορετική κάλυψη μεταξύ των επιστημονικών πεδίων, μεταξύ των γλωσσών και του τύπου των δημοσιεύσεων καθώς και των χωρών προέλευσης των συγγραφέων) και σε λάθη που οφείλονται στην ανυπαρξία καθιερωμένων τύπων για τα ονόματα συγγραφέων και συλλογικών οργάνων (Hicks 1995; Meho & Yang 2007; Moed 2005; Seglen 1998; Wallin 2005). Για όλους τους παραπάνω λόγους οι περισσότεροι ερευνητές τονίζουν πως οι βιβλιομετρικοί δείκτες είναι ενδεικτικοί της απήχησης (impact) και της προβολής (visibility) των δημοσιεύσεων και σε καμία περίπτωση δεν μπορούν να αποτιμήσουν την ερευνητική ποιότητα και τάσσονται υπέρ του συνδυασμού της μεθόδου αξιολόγησης από ομότιμους με τη μέθοδο ανάλυσης αναφορών (Martin 1996; Wallin 2005). Σε κάθε περίπτωση θα πρέπει να χρησιμοποιείται συνδυαστικά ένα πλήθος βιβλιομετρικών δεικτών όπως έχει ήδη αναφερθεί, με σκοπό τη σύγκλισή τους (Martin, 1996). Οι δείκτες ανάλυσης αναφορών θα πρέπει να υφίστανται κανονικοποίηση (normalization) ως προς τον αριθμό των συγγραφέων και ως προς τον επιστημονικό κλάδο της κάθε δημοσίευσης (Wallin, 2005) και τέλος, «ασχέτως από οποιαδήποτε βάση δεδομένων αναφορών ή πηγή άντλησης δεδομένων χρησιμοποιηθεί, οι μετρήσεις θα πρέπει να πραγματοποιούνται από επαγγελματίες με θεωρητική κατανόηση και ολοκληρωμένη τεχνική κατάρτιση για τις βάσεις δεδομένων, τις τεχνικές ανάκτησης πληροφοριών, τις συντομογραφίες, τις έννοιες και την ορολογία των επιστημών που μελετούνται και σύμφωνα με τις αρχές “βέλτιστης πρακτικής” της βιβλιομετρικής» (Meho & Rogers, 2008).

Σήμερα, εργαλεία ανάλυσης αναφορών είναι εύκολα διαθέσιμα για τον υπολογισμό της μέτρησης ποικίλων επιπτώσεων για τους μελετητές, βασισμένα σε δεδομένα από τους δείκτες αναφορών. Αυτά έχουν διάφορες εφαρμογές από την ταυτοποίηση εξειδικευμένων κριτών, με σκοπό την κριτική επανεξέταση επιστημονικών ανακοινώσεων και προτάσεων επιχορήγησης έως την παροχή διαφανών δεδομένων για την υποστήριξη της αναθεώρησης της ακαδημαϊκής αξίας, της κατοχής, και των αποφάσεων προώθησης.

Ο Nicholas (1978) στο άρθρο του «Βιβλιογραφία και Βιβλιομετρία» εξήγησε τη σημασία της ανάλυσης των αναφορών και των εφαρμογών τους στη LIS (Βιβλιοθήκη και Επιστήμη των Πληροφοριών). Επεσήμανε ότι οι επιστήμονες της πληροφορίας και οι βιβλιοθηκονόμοι χρησιμοποιούν την ανάλυση αναφορών για την ποσοτική αξιολόγηση των τίτλων βασικών περιοδικών και δημοσιεύσεων καμπής

(λιγότερο χρησιμοποιημένες ή δημοσιεύσεις του περιθωρίου) σε συγκεκριμένες ειδικότητες, τις αλληλεξαρτήσεις μεταξύ συγγραφέων από διάφορα ιδρύματα και σχολές και τα συναφή δεδομένα σχετικά με την πανεπιστημιακή κοινότητα. Κάποιες πιο ρεαλιστικές εφαρμογές των, εν λόγω, πληροφοριών περιλαμβάνουν το σχεδιασμό αναδρομικών βιβλιογραφιών, την εύρεση της ηλικίας των στοιχείων που χρησιμοποιούνται στην επιστήμη (χρόνος ημίσειας ζωής), καθώς και τη σύγκριση μεταξύ της χρήσης των πρόσφατων δημοσιεύσεων σε σχέση με παλαιότερες, συγκρίνοντας την κάλυψη των δευτερευουσών υπηρεσιών που μπορούν να βοηθήσουν τους εκδότες να μετρήσουν τα επιτεύγματα τους και τον ανταγωνισμό και μπορούν να βοηθήσουν τους βιβλιοθηκονόμους στην αποτίμηση «της αποτελεσματικότητας του αποθέματος τους». Υπάρχουν επίσης ορισμένοι περιορισμοί στην αξία των δεδομένων των αναφορών. Είναι συχνά ελλιπή ή μεροληπτικά, δεδομένα που έχουν συλλεχθεί με μη αυτόματο τρόπο. Ως εκ τούτου, απαιτείται περαιτέρω έρευνα για την πλήρη κατανόηση της λογικής που διέπει τις αναφορές, η οποία θα επιτρέψει την αποτελεσματική εφαρμογή τους.

Αποκαλύπτεται λοιπόν ότι η βιβλιομετρική μέθοδος είναι πολύ χρήσιμη για την ανάλυση του αντίκτυπου της βιβλιογραφίας σε οποιοδήποτε περιοχή και ειδικότερα στην LIS είναι χρήσιμη στον καθορισμό των πολιτικών για διάφορες δραστηριότητες όπως απόκτηση, οργάνωση, αποθήκευση, εισαγωγή νέων υπηρεσιών, κατάταξη των περιοδικών, υπολογισμό του χρόνου ημίσειας ζωής της βιβλιογραφίας σε οποιοδήποτε θέμα, επιστημονικής μορφοποίησης και πολιτικών ανάπτυξης συλλογών και συναφών πολιτικών.

Είδος αποτελεσμάτων που μπορούμε να πάρουμε από μια στατιστική ανάλυση τυπικών βιβλιογραφικών αναφορών.

Έτος έκδοσης

Μια ανάλυση των ετών δημοσίευσης μπορεί για παράδειγμα να δείξει τάσεις, κατά πόσο μια μονάδα δημοσιεύει σε σύγκριση με τον υπόλοιπο κόσμο, ή με παρόμοιες μονάδες.

Τίτλος περιοδικού

Μια ανάλυση των τίτλων περιοδικών μπορεί για παράδειγμα να δώσει μια γενική εικόνα του προτύπου των δημοσιεύσεων ορισμένης μονάδας.

Ονόματα συγγραφέων

Μπορεί να είναι ενδιαφέρον να εντοπιστούν διάφοροι συγγραφείς σε μια συγκεκριμένη χώρα ή σε μια συγκεκριμένη μονάδα.

Διευθύνσεις συγγραφέων

Ανάλογη με την ανάλυση των ονομάτων των συγγραφέων, μια ανάλυση των διευθύνσεων τους μπορεί να εντοπίσει παραγωγικές χώρες, πανεπιστήμια ή άλλους οργανισμούς και να δώσει μια επισκόπηση των προτύπων συνδημοσίευσης.

Λέξεις-κλειδιά

Εάν έχουμε επαρκή αριθμό από λέξεις-κλειδιά που αποδίδονται σε μία δημοσίευση είναι δυνατό να μελετηθεί ποια θέματα εμφανίζονται συχνά στις δημοσιεύσεις του πεδίου και τις συνδέσεις μεταξύ τους, μια λεγόμενη co-word ανάλυση (ανάλυση ζευγών λέξεων). Αυτό το είδος της ανάλυσης είναι πιο εύκολο να γίνει, αν η αναφορά περιέχει όρους από ελεγχόμενο λεξιλόγιο, όπως αυτό της Medical Subject Headings (MeSH) (Ιατρικές Θεματικές Επικεφαλίδες από την National Library of Medicine, των ΗΠΑ).

Συνδυασμοί των ανωτέρω

Οι παραπάνω αναλύσεις βελτιώνονται συχνά συνδυάζοντας περισσότερες από μία αναφορές. Με το συνδυασμό ονομάτων συγγραφέων ή διευθύνσεων συγγραφέων και λέξεων-κλειδιά είναι δυνατόν να εντοπισθούν διάφοροι συγγραφείς ή οργανισμοί σε έναν συγκεκριμένο τομέα. Αυτό καθιστά επίσης δυνατή την αναγνώριση συγγραφέων που συνδέονται με κοινά θέματα, κάτι που μπορεί να εντοπίσει τα υφιστάμενα ερευνητικά δίκτυα ή τη δυνατότητα ενός νέου ερευνητικού δικτύου.

Αναλύοντας καταλόγους αναφοράς

Όταν βιβλιογραφικές εγγραφές περιλαμβάνουν καταλόγους αναφορών αναφερόμενων άρθρων, μπορούμε να επεκτείνουμε τη στατιστική ανάλυση μας σχετικά με τις συνδέσεις μεταξύ των διαφόρων δημοσιεύσεων. Σε γενικές γραμμές, οι πρόσθετες πληροφορίες που απεκτήθησαν από τους καταλόγους αναφορών μπορούν να δώσουν τουλάχιστον δύο πτυχές της βιβλιομετρικής μελέτης:

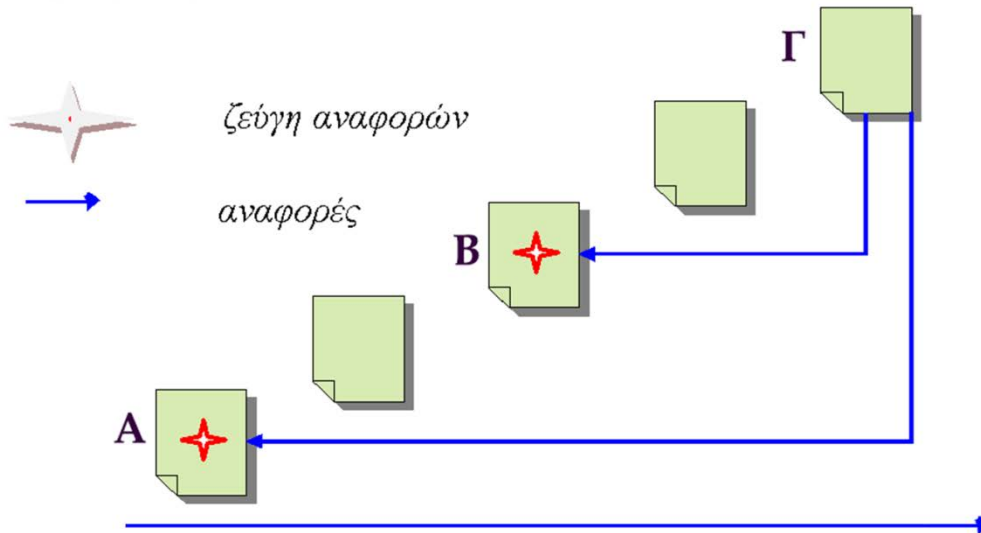
- Τη δυνατότητα να βρούμε εκδόσεις στην ίδια περιοχή, εντοπίζοντας τις δημοσιεύσεις που αναφέρουν ή αναφέρονται από τις δημοσιεύσεις που έχουν ήδη εντοπιστεί.

- Τη δυνατότητα βιβλιομετρικής αξιολόγησης της ποιότητας.

Δύο βασικά πεδία της ανάλυσης βιβλιογραφικών αναφορών.

1) Ζεύγη Αναφορών (Co-citation coupling).

Είναι μία μέθοδος που χρησιμοποιείται για να καθιερώσει μία θεματική ομοιότητα μεταξύ δύο τεκμηρίων. Εάν οι δημοσιεύσεις Α και Β αναφέρονται στη δημοσίευση Γ, θεωρούνται σχετικές μεταξύ τους, έστω κι αν δεν συνδέονται άμεσα.



Εικόνα 2 2 Τα ζεύγη αναφορών αναλύονται μέσω κοινών καταλόγων αναφορών. Οι δημοσιεύσεις με σήμανση αστέρι μπορεί να έχουν κάτι κοινό, δεδομένου ότι ένα πιο πρόσφατο άρθρο αναφέρεται σε αυτές τις δύο.

Αν οι δημοσιεύσεις Α και Β αναφέρονται μαζί σε πολλές άλλες, τότε έχουν μία ισχυρότερη σχέση, δηλαδή τόσο πιο ισχυρή είναι η παράλληλη αναφορά τους. Αυτό επεξηγείται στο Σχήμα 2.3.

Σε όσες περισσότερες δημοσιεύσεις αναφέρονται, τόσο ισχυρότερη είναι η σχέση τους.

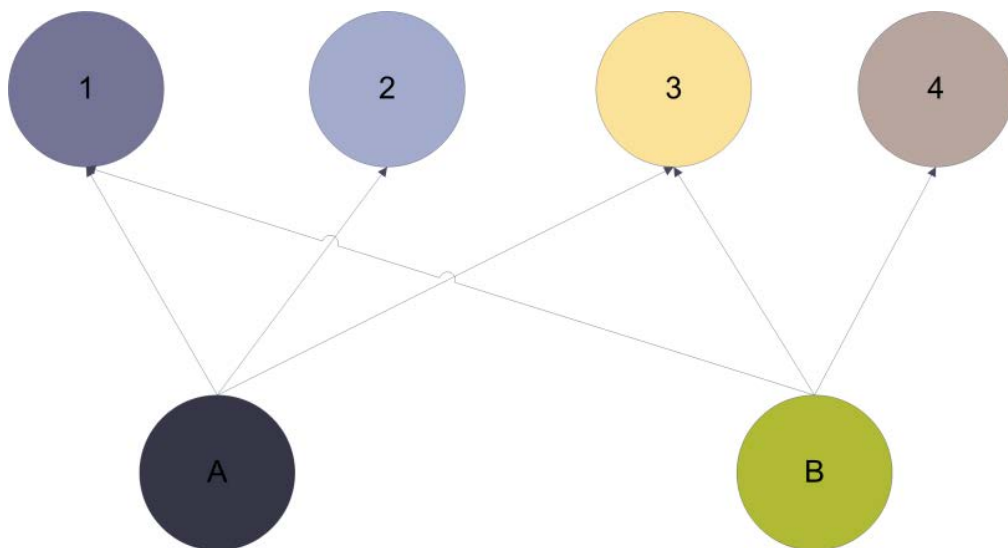
Στο σχήμα 2.1 το Α παραθέτει τα έγγραφα 1, 2 και 3. Κατά συνέπεια στα έγγραφα 1 και 2, 1 και 3, και 2 και 3 γίνεται παράλληλη ανάλυση αναφορών. Από την άλλη, το Β παραθέτει τα έγγραφα 1, 3 και 4. Συνεπώς στα έγγραφα 1 και 3, 1 και 4, και 3 και 4 υπάρχει επίσης παράλληλη ανάλυση αναφορών. Από αυτό το σχήμα μπορούμε να δούμε ότι τα έγγραφα 1 και 3 έχουν δύο δυνάμεις παράλληλης ανάλυσης αναφοράς το καθένα επειδή αναφέρονται από κοινού τόσο στο Α, όσο και στο Β. Τα έγγραφα 1 και 2, 2 και 3 έχουν μία δύναμη παράλληλης ανάλυσης αναφοράς, επειδή κάθε ζεύγος παρατίθεται μόνο από το Α, και τα έγγραφα 1 και 4, 3 και 4 έχουν επίσης μια δύναμη παράλληλης ανάλυσης αναφορών, διότι ως ζεύγος παρατίθενται μόνο από το Β. Η κυριότερη διαφορά μεταξύ βιβλιογραφικών ζευξεων και παράλληλης ανάλυσης αναφορών είναι ότι ενώ βιβλιογραφικές ζεύξεις μετρούν τη σχέση μεταξύ των πηγών, η παράλληλη ανάλυση αναφορών μετρά τη σχέση μεταξύ των εγγράφων που παρατίθενται.

Οι τεχνικές της παράλληλης ανάλυσης αναφορών έχουν εφαρμοστεί σε πολλές περιστάσεις. Ο Garfield (1974) τη χρησιμοποίησε για να αποκαλύψει την πολλαπλή ανακάλυψη των αμφιλεγόμενων οπιούχων υποδοχέων του εγκεφάλου. Ο Small (1977) χρησιμοποίησε επίσης το μοντέλο παράλληλης ανάλυσης αναφορών για να προσδιορίσει την ταυτόχρονη ανακάλυψη του πρόδρομου μορίου για το κολλαγόνο που αποκαλείται «προκολλαγόνο».

Η παράλληλη ανάλυση των αναφορών χρησιμοποιήθηκε επίσης στη μελέτη της επιδημιολογίας των επιστημονικών κλάδων από τον Small (1977) στη φυσική των σωματιδίων, από τον Garfield (1979) για τον εντοπισμό επιστημονικών παραδειγμάτων, από τους White και Griffith (1982) για την Επιστήμη, την Τεχνολογία και την Κοινωνία και τέλος, οι Earle και Vickery (1969) χρησιμοποίησαν την τεχνική αυτή για τη μελέτη της αυτο-παραγωγής και αυτο-παραπομπής των κλάδων.

Η παράλληλη ανάλυση ζευγών αναφορών έχει τους ανταγωνιστές της. Μεταξύ αυτών ο Edge (1977) που ισχυρίζεται ότι οι μελέτες της επικοινωνίας στον τομέα της επιστήμης τονίζουν τη σημασία της άτυπης επικοινωνίας κι αποδεικνύουν ότι οι επίσημες και ανεπίσημες περιοχές διαφέρουν. Το να προσπαθούμε

συνεπώς να χρησιμοποιήσουμε στοιχεία από την επίσημη περιοχή από την οποία θα προταθούν εξηγήσεις που εφαρμόζονται σ' αυτήν σημαίνει ότι «αντιστρέφεται η απαραίτητη επεξηγηματική λογική. Παρότι η συζήτησή μας πάνω στις βιβλιομετρικές μετρήσεις επικεντρώθηκε κυρίως σε επιμέρους άρθρα, περιοδικά και συγγραφείς, δύνανται να υπάρχουν διάφορα επίπεδα συσσωμάτωσης. Μονάδες ανάλυσης μπορούν να είναι βιομηχανικές οργανώσεις, ακαδημαϊκά τμήματα, εκδοτικοί οίκοι, πανεπιστήμια, πόλεις, έθνη, ήπειροι κ.λ.π. Στην πραγματικότητα, η ανάλυση αναφορών υπόκειται εύκολα σε ταξινομήσεις συχνότητας, όπου βιβλιογραφικές αναφορές είναι τα ακατέργαστα δεδομένα.

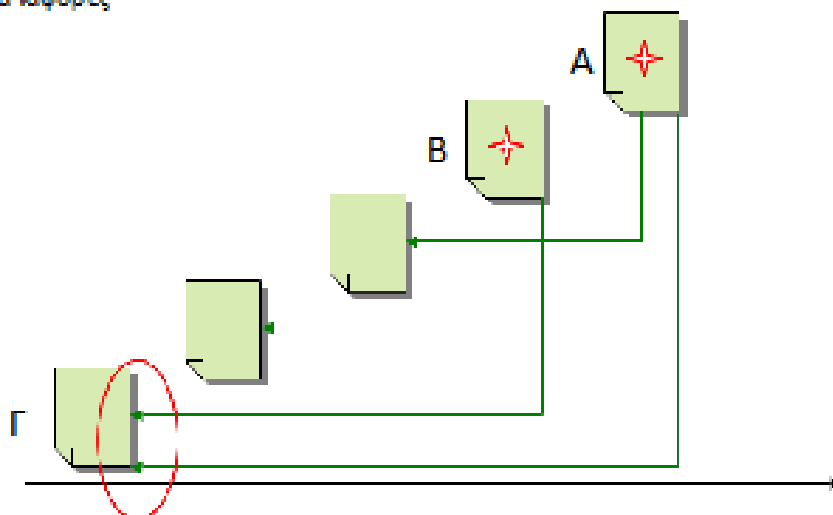


Σχήμα 2 1 Παράλληλη ανάλυση ζευγών αναφορών: οι κύκλοι με τα γράμματα αναπαριστούν τις πηγές, και οι αριθμημένοι κύκλοι αναπαριστούν τα έγγραφα που παρατίθενται.

Βιβλιογραφική Ζεύξη (Bibliographic coupling).

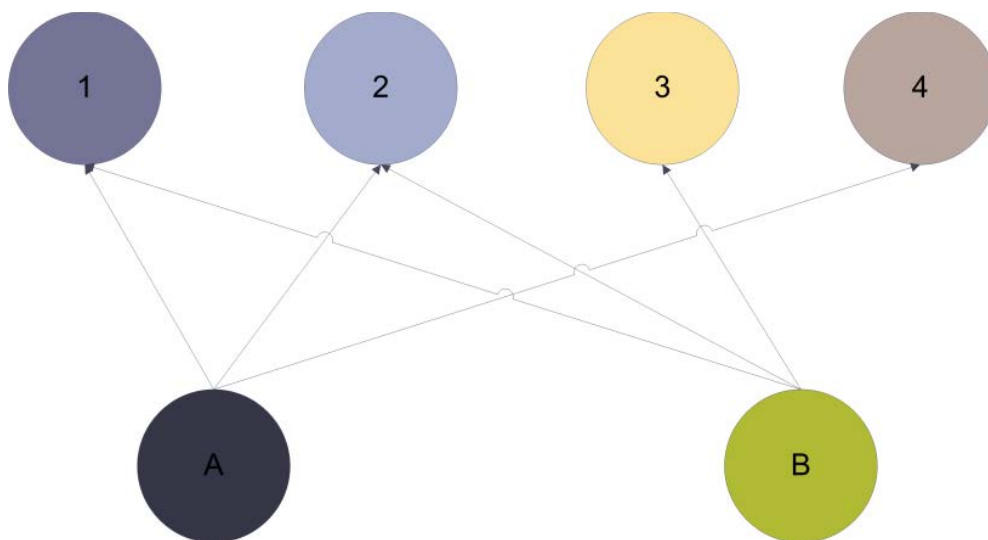
Λειτουργεί με μία παρόμοια αρχή με τα ζεύγη αναφορών, αλλά αποτελεί τον αντικατοπτρισμό των ζευγών αναφορών (σχ. 2.6). Συνδέει δύο δημοσιεύσεις που αναφέρουν τα ίδια άρθρα, έτσι ώστε εάν οι δημοσιεύσεις A και B αναφέρουν τη δημοσίευση Γ, τότε μπορούμε να πούμε ότι είναι σχετικές, έστω και αν δεν συνδέονται άμεσα μεταξύ τους. Όσα περισσότερες κοινές δημοσιεύσεις αναφέρουν τόσο ισχυρότερη είναι η σχέση τους.

✦ = βιβλιογραφική ζεύξη
→ = αναφορές



Εικόνα 2 3 Η Βιβλιογραφική ζεύξη αναλύεται μέσω κοινών αναφορών. Τα πιο πρόσφατα άρθρα με σήμανση αστέρι μπορεί να συνδέονται μεταξύ τους, δεδομένου ότι αναφέρονται στο ίδιο παλαιότερο άρθρο.

Το παρακάτω σχήμα 2.2 απεικονίζει την έννοια των βιβλιογραφικών ζεύξεων. Οι κύκλοι με τα γράμματα αναπαριστούν τα άρθρα που αποτελούν πηγή ή προσφέρουν αναφορές και οι αριθμημένοι κύκλοι αναπαριστούν τις αναφορές που παρατίθενται. Τα μόνα άρθρα που παρατίθενται από το Α κι από το Β είναι τα έγγραφα 1 και 2. Ως εκ τούτου, το Α και το Β έχουν δύο δυνάμεις ζεύξεων.

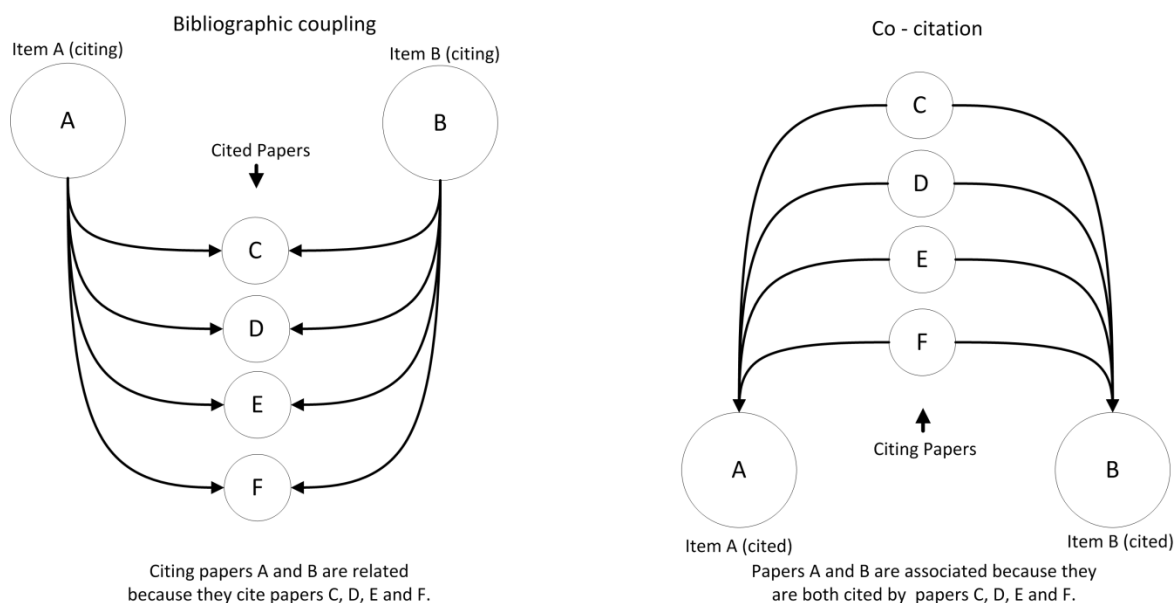


Σχήμα 2 2 Βιβλιογραφικές ζεύξεις: οι κύκλοι με τα γράμματα αναπαριστούν τις πηγές, και οι αριθμημένοι κύκλοι αναπαριστούν τα έγγραφα που παρατίθενται.

Στο πλαίσιο της μελέτης των αναφορών είχαν εξ αρχής μελετηθεί οι ονομαζόμενες βιβλιογραφικές συζεύξεις, που απαριθμούν το πλήθος των κοινών αναφορών μεταξύ δυο εργασιών (Kessler, 1965). Με τέτοιες συζεύξεις αποσκοπείτε η εκτίμηση του βαθμού συνάφειας, συγγένειας ή συσχέτισης των εργασιών. Επειδή όμως οι βιβλιογραφικές συζεύξεις τείνουν να συγκεντρώνονται γύρω από εργασίες ανασκόπησης, δεν θεωρήθηκαν ιδιαίτερα σημαντικός επιστημονομετρικός δείκτης.

Ο Kessler επίσης ανακάλυψε μια υψηλή συσχέτιση ανάμεσα σε ομάδες που σχηματίστηκαν με ζεύξη και σε ομάδες που σχηματίστηκαν με βάση την ευρετηρίαση του θέματος τους. Καθώς η βιβλιομετρική ανάλυση αυτοματοποιήθηκε περισσότερο, πολλοί προσπάθησαν να πάρουν αυτήν και τις παρακάτω τεχνικές και να σχεδιάσουν λογισμικό το οποίο θα εντοπίζει μοτίβα (patterns) και θα εγκαθιστά σχέσεις μεταξύ άρθρων. Ο Price και ο Schiminovich (1971) ξεκίνησαν να αναλύουν ηλεκτρονικές δημοσιεύσεις στα τέλη του 60 και στις αρχές του 70, καθώς μερικά άρθρα κινήθηκαν προς ψηφιακή μορφή. Αυτά ήταν τα πρώτα σημαντικά βήματα που έγιναν στο να εφαρμοστεί η βιβλιομετρική σε ηλεκτρονικές δημοσιεύσεις, με δεδομένο ότι οι πληροφορίες μπορούν να μετρηθούν στην ψηφιακή μορφή τους πολύ γρηγορότερα από ότι θα μετρηθούν στο χέρι.

Ο Martyn (1964) υποστηρίζει ότι οι βιβλιογραφικές ζεύξεις δεν είναι έγκυρος τρόπος μέτρησης μίας σχέσης, επειδή δεν γνωρίζουμε αν δύο άρθρα που παραπέμπουν σε ένα τρίτο, παραπέμπουν σε παρόμοιες πληροφορίες του. Επομένως, οι βιβλιογραφικές ζεύξεις δεν είναι παρά μια ένδειξη της «ύπαρξης πιθανότητας σχέσης μεταξύ δύο εγγράφων». Ο Tagliacozzo(1990) υποστηρίζει την ίδια ιδέα ακόμη πιο σθεναρά. «Το γεγονός ότι δύο άρθρα έχουν μια κοινή αναφορά δεν αποτελεί εγγύηση ότι και τα δύο αναφέρονται στην ίδια πληροφορία».



Σχήμα 2 3 (i) Ζεύγη αναφορών. Τα άρθρα A και B σχετίζονται γιατί αναφέρουν και τα δύο τα άρθρα C, D, E, F. (ii) Ζεύξη Αναφορών. Τα άρθρα A και B σχετίζονται γιατί και τα δύο αναφέρονται από τα άρθρα C, D, E, F.

Ανάλυση των συν-αναφορών (Co-citation Analysis)

Μεγαλύτερο ενδιαφέρον έχει η ανάλυση των συν-αναφορών, που μετρά το πλήθος των αναφορών, όχι σε μια, αλλά σε δυο εργασίες μαζί (Small&Griffith, 1974). Αυτό σημαίνει ότι με τη συχνότητα των συν-αναφορών ενός ζευγαριού εργασιών, υπάρχει η πεποίθηση ότι γίνεται μια εκτίμηση του μεγέθους της συσχέτισής τους. Το καλό στην περίπτωση αυτή είναι ότι οι συν- αναφορές ελευθερώνονται από κάποια συγκεκριμένη αρχική εργασία, αφού υπολογίζονται στο σύνολο των αναφορών που μπορούν να συλλεχθούν σε μια επιστημονική περιοχή. Επιπλέον, πιστεύεται ότι οι σχηματισμοί συγκεντρωμένων αναφορών (citation clusters) που προκύπτουν έτσι, αντανακλούν τη συγκρότηση ερευνητικών δικτύων ή ειδικοτήτων. Επομένως, ως εργαλείο πολιτικής της επιστήμης, η ανάλυση των συν-αναφορών μπορεί να χρησιμοποιηθεί για να βάλει σε μια σειρά και να αξιολογήσει τις διαφορετικές ερευνητικές δραστηριότητες ούτως ώστε να μπορούν να ληφθούν οι, κατά το δυνατόν, πιο ρεαλιστικές αποφάσεις για τη χρηματοδότησή τους.

Οι Marshakova και Small (χωριστά) βελτίωσαν τη σύζευξη πολύ περισσότερο, σημειώνοντας πως, εάν δυο αναφορές τοποθετηθούν μαζί σε μια τελική βιβλιογραφία, τότε οι δυο αναφορές θα συγγενεύουν μεταξύ τους. Όσο μεγαλύτερος είναι ο αριθμός που θα έχουν αναφερθεί μαζί, τόσο μεγαλύτερη η δύναμη συν-αναφοράς. Η μεγαλύτερη σχέση μεταξύ βιβλιομετρικής ζεύξης και συν-αναφοράς είναι πως, όταν η ζεύξη μετράει τη σχέση μεταξύ της προέλευσης των εγγράφων, η συν-αναφορά μετράει τη σχέση μεταξύ των αναφερόμενων εγγράφων. Αυτό συνεπάγεται ότι αν ένας συγγραφέας σκόπιμα διάλεξε να συσχετίσει δυο άρθρα μαζί, αυτό δεν είναι απλώς μια ένωση μεταξύ δυο άρθρων ως συζευκτικό αποτέλεσμα.

Για τον συντονισμό αυτών των τεχνικών σε λεπτότερα σημεία θα πρέπει να απαντηθούν μια σειρά ερωτημάτων σχετικά με την αναφορά:

- Είναι η αναφορά εννοιολογική ή επιχειρησιακή;
- Είναι η αναφορά οργανική ή τυπική;
- Είναι η αναφορά εξελικτικής ή πειραματικής μορφής;
- Είναι η αναφορά επιβεβαιωτική ή καταφατική;

Αυτά τα ερωτήματα αποκαλύπτουν πιστότητα και τάσεις στη μέτρηση, τέτοιες τάσεις όπως αυτή του να αρνείται (ή να προτείνει το εναλλακτικό) ενός άρθρου. Μπορούμε επίσης να ξεκινήσουμε να βλέπουμε την ωριμότητα των ιδεών, καθώς αυτές εξελίσσονται μέσα στη βιβλιογραφία.

Μοντέλο της Μεθόδου Ομαδοποίησης (block-modeling)

Μια επέκταση της ανάλυσης των αναφορών είναι το μοντέλο της μεθόδου της ομαδοποίησης, κατά το οποίο η ανάλυση των σχηματισμών των συγκεντρωμένων αναφορών ή των δικτύων των αναφορών δεν

γίνεται με βάση τις μεταξύ τους προτιμήσεις, αλλά από τις επιλογές τρίτων μερών ή προσώπων, που δεν ανήκουν στην κοινή ομάδα (Lenoir, 1979). Αντίθετα με την ανάλυση των συν- αναφορών, το μοντέλο της ομαδοποίησης κάνει μια ταξινόμηση στηριγμένη στη συνολική επίγνωση για το τι είναι σημαντικό, παρά στη μερική αλληλο-επίγνωση. Με τη μέθοδο αυτή, ο Timothy Lenoir ισχυρίζεται ότι μπορεί να γίνει μια διαχρονική ανάλυση της δυναμικής των αναφορών και να δοθεί μια πληρέστερη εικόνα της ανάπτυξης των επιστημονικών περιοχών.

Οι προηγούμενες αναλύσεις των συν-αναφορών και της ομαδοποίησης είναι δυο παραδείγματα εφαρμογής στην Επιστημομετρία των μεθόδων δικτυακής ανάλυσης, που έχουν αναπτυχθεί εκτεταμένα από την κλασική αμερικάνικη κοινωνιολογική σκέψη. Στο πλαίσιο όμως των κοινωνικών μελετών της επιστήμης και της τεχνολογίας, αναπτύχθηκε στη δεκαετία 1980 η παρακάτω δικτυακή μέθοδος.

Η θεωρία του δικτύου των δραστών (actor network theory ή ANT)

Η θεωρία αυτή, επιπλέον, οφείλει κάποια θεωρητικά σημεία στην προηγούμενη μελέτη για τα αόρατα κολέγια και τις ειδικότητες. Ειδικότερα, η θεωρία δικτύου δραστών επισημαίνει ότι οι αναλύσεις αναφορών και συν-αναφορών δεν μπορούν να ξεφύγουν από τη συμβατική τυπολογία της κλασικής κοινωνιολογίας της επιστήμης με το να επιμένουν να βλέπουν τις κοινωνικές σχέσεις μόνο με βάση τα αποκαλυπτόμενα μορφώματα αναφορών. Σε αντίθεση με αυτή τη λογική, η θεωρία δικτύου δραστών επιχειρεί να αναπτύξει μια εναλλακτική ανάλυση, που πιστεύεται ότι επικεντρώνεται περισσότερο στο θέμα του περιεχομένου της επιστήμης παρά στις τυπικές ή θεσμοποιημένες επιστημονικές σχέσεις. Αυτό γίνεται από τη θεωρία δικτύου δραστών με την υιοθέτηση της ονομαζόμενης ανάλυσης των συν-λέξεων, που είναι από κοινού συναντώμενες λέξεις-κλειδιά, συνήθως, στην περιγραφή του κειμένου της εργασίας, στην περίληψη ή στον τίτλο και τον υπότιτλο ή στον πίνακα των λέξεων- κλειδιών (Callon, Law & Rip 1986; Rip & Courtial 1984).

Προβληματισμοί και συμπεράσματα από την ανάλυση των αναφορών

Από μια άλλη οπτική γωνία, η Diana Hicks (1995) έχει προβληματισθεί για την πραγματική χρησιμότητα των μεθόδων των μελετών αναφορών για την ανάλυση της πολιτικής της επιστήμης, επειδή η υπερβολική και τυπική εμπιστοσύνη σ' αυτές μπορεί να οδηγήσει σε σφάλματα πολιτικών επιλογών. Ακόμη η Diana Hicks μαζί με τον Jonathan Potter (1991) έχουν διατυπώσει μια ιδιότυπη κριτική της ανάλυσης αναφορών, θεωρώντας την ως μηχανισμό που μπορεί να οδηγήσει στην αυτο-ρύθμιση των επιστημόνων (1991). Πέραν όμως από τους γενικότερους περιορισμούς της θεωρίας, υπάρχουν πολλά ενδιαφέροντα συμπεράσματα που μπορούν να βγουν από τη μελέτη των αναφορών. Ένα από αυτά αφορά τη χρονική κατανομή των αναφορών. Καθώς προχωρά χρονικά η επιστήμη, φαίνεται να συμπαρασύρει μαζί της και το χρονικό διάστημα που γίνονται οι περισσότερες αναφορές. Μ' αυτό θέλουμε να πούμε ότι, εξετάζοντας τις αναφορές μιας εργασίας, συνήθως οι περισσότερες από αυτές γίνονται σε εργασίες που έχουν γραφεί πρόσφατα, δηλαδή, σχετικά κοντά με την εξεταζόμενη εργασία. Αυτό σημαίνει ότι υπάρχει μια επίδραση αμεσότητας στις αναφορές, που υπαγορεύει να γίνονται αναφορές στις πρόσφατες εργασίες συχνότερα από τις παλαιότερες.

Χρόνος Ημι-ζωής Εργασίας

Ένα μέτρο εκτίμησης της επίδρασης της αμεσότητας στις αναφορές είναι ο ονομαζόμενος χρόνος ημι-ζωής μιας εργασίας, που αναφέρεται στο πλήθος των ετών, εντός των οποίων έχουν γραφεί οι μισές από τις πιο πρόσφατες αναφορές της εργασίας αυτής. Όσο πιο μικρός είναι ο χρόνος ημι-ζωής των εργασιών μιας επιστημονικής περιοχής, τόσο πιο εύκολα ξεχνιούνται ή θεωρούνται απαρχαιωμένες οι παλαιότερες συνεισφορές, κάτι που ερμηνεύεται σαν στοιχείο δυναμικής ανάπτυξης της περιοχής αυτής. Φυσικά, ο χρόνος ημι-ζωής μεταβάλλεται ανά επιστημονική περιοχή και κλάδο. Για παράδειγμα, ο Price εκτιμούσε ότι στη Φυσική και τις Βιοϊατρικές επιστήμες ο χρόνος ημι-ζωής των εργασιών είναι περίπου τρία χρόνια ή και λιγότερο, λόγω των ταχέων ρυθμών ανάπτυξης των περιοχών αυτών.

Τοπικό αποθετήριο

Ένα τοπικό αποθετήριο συνήθως δεν περιέχει καμία πληροφορία γύρω από τις αναφορές, αλλά μπορεί να περιέχει άλλες πληροφορίες που μπορούν να χρησιμοποιηθούν σε μία βιβλιομετρική μελέτη,

ανάλογα με την πολιτική για την εισαγωγή πληροφοριών δημοσίευσης. Ένα τοπικό αποθετήριο μπορεί κάλλιστα να είναι το πιο πλήρες αρχείο των δημοσιεύσεων από μια συγκεκριμένη ερευνητική μονάδα.

Τεχνικά ζητήματα με τα δεδομένα

Ανορθογραφίες

Ορθογραφικά λάθη στα δεδομένα μπορεί να συμβούν στα ονόματα των συγγραφέων, στις διευθύνσεις, στους τίτλους περιοδικών ή σε άλλες πληροφορίες, που ακολούθως θα οδηγήσουν σε λάθος αριθμούς στις αναφορές των δημοσιεύσεων. Αυτό το είδος σφάλματος είναι αρκετά συχνό και μπορεί να οφείλεται τόσο σε λάθη από τους αρχικούς συγγραφείς ή λάθη που έγιναν από τον παραγωγό της βάσης δεδομένων. Σε μεγάλες ποσότητες δεδομένων, ωστόσο, η επίδραση αυτών των λαθών είναι συχνά αμελητέα ποσότητα.

Άγνωστες διευθύνσεις

Τόσο το 2,4% των άρθρων, επιστολών, σημειώσεων και ανασκοπήσεων του Science citation index 1993-2003, όσο και το 14% στο Citation Index Κοινωνικών Επιστημών (εκτός SCI) παρουσιάζουν ελλιπή ενημέρωση σχετικά με τους συγγραφείς και τις διευθύνσεις τους (Moed, H. F. 2005, σ. 186) Το ποσοστό είναι υψηλότερο, αν περιληφθούν και άλλοι τύποι εγγράφων. Για αυτό το λόγο, οι αναφορές αυτές θα αποκλειστούν αυτόματα από πολλούς τύπους αναλύσεων.

Συνδέσεις μεταξύ των αναφορών και του αντίστοιχου άρθρου

Η αυτόματη αντιστοίχιση από τους καταλόγους αναφοράς των αντίστοιχων άρθρων πάντα αποτυγχάνει να εντοπίσει κάποιες από τις συνδέσεις. Ο Moed λέει: ' οι στατιστικές αναφορών στο επίπεδο των ατόμων, των ερευνητικών ομάδων, περιοδικών και χώρων έχουν πληγεί έντονα από προχειρότητα: στις αναφορές, στα συντακτικά χαρακτηριστικά των επιστημονικών περιοδικών, στα προβλήματα γλώσσας, στα προβλήματα συγγραφέα-ταυτοποίησης'. Σε αυτές τις περιπτώσεις ο συνολικός αριθμός των αποκλινόντων αναφορών που παρατίθενται στο Web of Science είναι περίπου 7%, αλλά μπορεί να είναι πολύ υψηλότερος σε ιδιαίτερες καταστάσεις.

Κλασμάτωση και στάθμιση

Όλοι οι συγγραφείς δεν μπορεί να θεωρηθούν ότι έχουν συμμετάσχει με την ίδια συνεισφορά στη συγγραφή μιας δημοσίευσης και ορισμένα είδη δημοσιεύσεων μπορεί να θεωρηθούν ότι έχουν μεγαλύτερη σημασία από τις άλλες σε μια συγκεκριμένη ερευνητική περιοχή. Υπάρχουν, επομένως, διαφορετικοί τρόποι για την αντιστάθμιση γι' αυτούς και παρόμοιους παράγοντες για τον υπολογισμό του αριθμού των δημοσιεύσεων ή τον αριθμό των αναφορών που αποδίδονται σε ένα συγκεκριμένο συγγραφέα ή μονάδα.

Κλασμάτωση σημαίνει ότι μόνο ένα μέρος μιας δημοσίευσης ή / και μιας αναφοράς αποδίδεται στη μονάδα ανάλυσης, σύμφωνα με κάποιο μαθηματικό υπολογισμό.

Στάθμιση σημαίνει ότι ορισμένοι τύποι των δημοσιεύσεων ή / και αναφορές θεωρούνται ότι έχουν μεγαλύτερη σημασία από τις άλλες και, ως εκ τούτου, πρέπει να τους δοθεί μεγαλύτερη "βαρύτητα". Η εφαρμογή της στάθμισης εισάγει ένα υποκειμενικό στοιχείο στην ανάλυση από τη στιγμή που ο συντελεστής με τον οποίον η δημοσίευση / αναφορά πρέπει να σταθμίζεται συναποφασίζεται από τον αναλυτή σε συνεργασία με τους εμπειρογνώμονες στο ερευνητικό πεδίο που ανήκει η δημοσίευση, η οποία πρόκειται να αξιολογηθεί.

Ο αντίκτυπος της κλασμάτωσης και στάθμισης στους δείκτες είναι σημαντικός, και θα πρέπει πάντα κάποιος να λαμβάνει υπόψη του τι θα ωφελήσει ή τι θα δυσκολέψει χρησιμοποιώντας διαφορετικούς τρόπους μέτρησης κατά την εξέταση των δεικτών. Ενδέχεται επίσης η χρήση της κλασμάτωσης και στάθμισης να αποθαρρύνει συνεργασίες, κάτι που σαν αποτέλεσμα δεν είναι επιθυμητό. Σε κάθε περίπτωση το θέμα από ερευνητική άποψη είναι άκρως ενδιαφέρον και απασχολεί έντονα τους βιβλιομέτρες.

Κλασμάτωση και στάθμιση του αριθμού των δημοσιεύσεων

Αν δεν κλασματώσεις ή δεν λάβεις υπ' όψιν τη στάθμιση μιας δημοσίευσης με οποιονδήποτε τρόπο, τότε θα αποδοθεί όπως έχουμε ήδη πρωτοειπεί ίση συνεισφορά σε κάθε ένα από τους συγγραφείς και κάθε δημοσίευση θα θεωρείται ότι είναι ίσης αξίας, όπως έχουμε ήδη αναφέρει.

Κλασμάτωση συγγραφέων ή μονάδων παραγωγής

Αν θελήσουμε να αποδώσουμε ανάλογα με τη σύνθεση της λίστας των συγγραφέων, υπάρχουν αρκετές εναλλακτικές λύσεις:

- Κλασμάτωση. (Δες 2.6.1).

- Στάθμιση στους συγγραφείς ή σε μονάδες παραγωγής.
 - Στάθμιση ανάλογα με την σειρά στην λίστα των συγγραφέων. Σε ορισμένες ειδικότητες, όπως για παράδειγμα την Ιατρική, ο συγγραφέας που έχει κάνει το μεγαλύτερο μέρος της εργασίας είναι συχνά ο πρώτος στη λίστα συγγραφέας και ο πλέον ανώτερος ερευνητής είναι στην τελευταία. Θα μπορούσαμε για παράδειγμα να δώσουμε στον πρώτο συγγραφέα 50% της έκδοσης, στον τελευταίο συγγραφέα 20% και να μοιραστούν το υπόλοιπο 30% το υπόλοιπο των συγγραφέων. Ωστόσο, αυτό διαφέρει όχι μόνο μεταξύ των κλάδων αλλά και μεταξύ των χωρών, των πανεπιστημιακών τμημάτων και των περιοδικών, με αποτέλεσμα να κάνει τη σύγκριση με άλλες μονάδες μάλλον δυσχερή.
 - Στάθμιση ανάλογα με τον αριθμό εμφάνισης ενός συγγραφέα σε διαφορετικά ιδρύματα, όταν η δημοσίευση αποδίδεται στο ίδρυμα. Γίνεται λοιπόν η υπόθεση ότι το ίδρυμα με τους περισσότερους συγγραφείς συνέβαλε περισσότερο στην εργασία για την παραγωγή της δημοσίευσης (π.χ. συγγραφέας που εμφανίζεται μέλος σε πανεπιστημιακό τμήμα και σε ερευνητικό κέντρο).
- Στάθμιση σύμφωνα με τον τύπο της δημοσίευσης.
 - Σε ορισμένες ερευνητικές περιοχές, ιδιαίτερα σε εκείνες όπου και άλλα είδη δημοσιεύσεων εκτός από τα άρθρα σε περιοδικά είναι σημαντικά, μπορούμε να αποδώσουμε περισσότερο βάρος σε ορισμένους τύπους δημοσίευσης, όπως για παράδειγμα: $P \text{ άθροισμα_στάθμισης} = \text{αριθμός των βιβλίων} * 5.0 + \text{αριθμός άρθρων σε περιοδικά} * 1.0 + \text{αριθμός κεφαλαίων βιβλίων} * 0,8$.
- Κλασμάτωση ανά τομέα.
 - Στην ανάλυση που παράγει ξεχωριστούς δείκτες για κάθε επιστημονικό πεδίο που περιλαμβάνεται, ο αριθμός των δημοσιεύσεων μπορεί να σταθμιστεί μεταξύ των διαφόρων πεδίων. Μια δημοσίευση που κατηγοριοποιείται ότι ανήκει τόσο στην ογκολογία όσο και την αιματολογία για παράδειγμα, συμβάλλει με το μισό της δημοσίευσης σε κάθε πεδίο.
- Στάθμιση ανά τομέα.
 - Αν μια δημοσίευση μπορεί να θεωρηθεί ότι ανήκει σε περισσότερα από ένα πεδία από ότι οι άλλες που έχουν ταξινομηθεί μαζί, αυτό το πεδίο μπορεί να λάβει ένα μεγαλύτερο μερίδιο της δημοσίευσης από τους άλλους τομείς. Αυτό είναι πολύ δύσκολο να εφαρμοστεί με μια κατηγοριοποίηση που γίνεται με θεματικές κατηγορίες των ISI περιοδικών, αλλά μπορεί να είναι δυνατό όταν χρησιμοποιείται για τους όρους το λεξιλόγιο MeSH για την κατάταξη του άρθρου, δεδομένου ότι σε αυτό περιλαμβάνονται διαιρέσεις των όρων σε πεδία και υποπεδία.
- Κλασμάτωση και στάθμιση του αριθμού αναφορών. Η κλασμάτωση και η στάθμιση από τους συγγραφείς (ή άλλες μονάδες παραγωγής), τους τύπους των δημοσιεύσεων και τα ερευνητικά πεδία. Οι αναφορές σε ένα άρθρο μπορεί να κλασματοποιηθούν και να σταθμιστούν με τις ίδιες μεθόδους και εκτιμήσεις, όπως ακριβώς με τον αριθμό των δημοσιεύσεων.
- Κλασμάτωση ανάλογα με το μήκος της λίστας αναφοράς.
 - Όταν ένα άρθρο αναφέρεται σε άλλο άρθρο, μπορεί να θεωρηθεί ότι η αξία της αναφοράς είναι αντιστρόφως ανάλογη με το μήκος της λίστας αναφοράς. Με άλλα λόγια, η αναφορά από ένα άρθρο με έναν μακρύ κατάλογο αναφοράς θα μπορούσε να θεωρηθεί μικρότερης αξίας από αναφορά από ένα άρθρο με μια σύντομη λίστα αναφοράς. Η πτυχή αυτή μπορεί να επιτευχθεί από κλασμάτωση της αξίας των αναφορών, διαιρώντας τον αριθμό αναφορών με τον αριθμό των αναφορών της λίστας αναφοράς. Για παράδειγμα, μια αναφορά από ένα άρθρο με 20 αναφορές θα δώσει στη δημοσίευση που αναφέρεται μια τιμή αναφοράς ίση με το $1/20 = 0,05$.
- Στάθμιση σύμφωνα με την πηγή αναφοράς.
 - Οι αναφορές σε ένα άρθρο μπορεί να προέρχονται από πολλές διαφορετικές πηγές και οι πηγές δεν μπορεί πάντα να θεωρούνται ίσης σημασίας. Μπορεί να είναι

επιθυμητό οι αναφορές από άρθρα περιοδικών με υψηλό δείκτη επιρροής να σταθμίζονται υψηλότερα από τα άρθρα των περιοδικών με χαμηλότερο δείκτη επιρροής. Ένα ιδιαίτερο είδος της στάθμισης αναφοράς ανάλογα με την πηγή γίνεται με την κατάταξη αλγόριθμου (PageRank) της Google. Εν ολίγοις, ο αλγόριθμος αυτός δίνει μεγαλύτερο βάρος σε αναφορές από τις δημοσιεύσεις που έχουν οι ίδιες υψηλό αριθμό αναφορών.

Γλώσσα

Η ανάλυση των δεδομένων του Web of Science δείχνει όπως έχουμε ήδη αναφέρει ότι τα μη-αγγλικά δημοσιεύματα αναφέρονται, κατά μέσο όρο, πολύ λιγότερο από ότι τα δημοσιεύματα στην αγγλική γλώσσα. Οι τιμές του δείκτη για τους συγγραφείς αγγλικών και μη-αγγλικών δημοσιεύσεων βελτιώνονται, όταν αποκλείονται οι μη αγγλικές εκδόσεις τους (van Leeuwenetal, 2003).

Επικρατούσες παραδόσεις κατά την απόδοση αναφοράς

Οι επικρατούσες παραδόσεις σε διαφορετικούς τομείς επηρεάζουν τον αριθμό των αναφορών που δίνονται σε διαφορετικά άρθρα. Διαφορετικοί τομείς επίσης διαφέρουν ως προς το πόσο γρήγορα ένα έγγραφο θα αναφέρεται, πόσο καιρό το ποσοστό αναφοράς θα λαμβάνει τη μεγίστη τιμή και πόσο καιρό η δημοσίευση θα συνεχίσει να αναφέρεται. Ο Eugene Garfield περιγράφει διάφορους παράγοντες που συμβάλλουν στα πεδία "δυναμικής αναφοράς". Ένας παράγοντας είναι το μήκος των καταλόγων αναφοράς, η οποία για παράδειγμα είναι δύο φορές μεγαλύτερη στη Βιοχημεία απ' ότι στα Μαθηματικά. Ένας άλλος παράγοντας είναι το μέγεθος των πεδίων. Πολλές ειδικότητες που συνδέονται με την Ιατρική και επιστήμες της ζωής έχουν υψηλό δυναμικό αναφορών. Είναι δυνατόν να πάρουμε μια γενική εικόνα των επικρατουσών παραδόσεων σε ένα συγκεκριμένο πεδίο εξετάζοντας τις λίστες αναφοράς για την επιλογή των αντικειμένων μέσα σε αυτό το πεδίο. Το μήκος των καταλόγων αναφοράς, τα θέματα των προαναφερόμενων δημοσιεύσεων και τα χρόνια της δημοσίευσής τους, σε κάποιο βαθμό, μπορούν να δώσουν μια εικόνα για την ταχύτητα αναφοράς και συμπεριφοράς του συγκεκριμένου πεδίου.

Αυτοαναφορές (self-citations)

Μελέτες έχουν δείξει ότι οι αυτοαναφορές δεν επηρεάζουν σημαντικά τα αποτελέσματα της ανάλυσης, όταν έχει μελετηθεί ένας αρκετά μεγάλος αριθμός δημοσιεύσεων. Αυτό ισχύει πιθανώς, επειδή οι περισσότεροι ερευνητές αναφέρονται σε δικές τους εργασίες με τις ίδιες συχνότητες. Σε επίπεδο ομάδας, ωστόσο, κάτι τέτοιο μπορεί πράγματι να επηρεάσει τους δείκτες.

Διαθέσιμη επιλογή για να αντισταθμιστεί αυτό, είναι η προσπάθεια να αποκλεισθούν οι αυτοαναφορές κατά τον υπολογισμό των τιμών των δεικτών, επισημαίνοντας ότι η ερμηνεία των δεικτών μπορεί να επηρεαστεί από το ποσό των αυτοαναφορών ή να υποτεθεί ότι είναι ομοιόμορφα καταναμημένα και ως εκ τούτου, να αγνοήσουμε τις επιπτώσεις τους κατά τον υπολογισμό των δεικτών. Είναι όμως πολύ δύσκολο να αφαιρεθούν οι αυτοαναφορές κατά τον υπολογισμό των δεικτών και απαιτεί δεδομένα από μια ολοκληρωμένη βιβλιογραφική βάση δεδομένων, όπως οι δείκτες αναφορών Thomson. Η διάσταση των αυτοαναφορών έχει εξαλειφθεί στις περιγραφές των δεικτών παρακάτω.

2.9. Ανάλυση Περιεχομένου (Content Analysis)

Η ανάλυση περιεχομένου ή ανάλυση κειμένου είναι μια μεθοδολογία που χρησιμοποιείται στις κοινωνικές επιστήμες για τη μελέτη του περιεχομένου της επικοινωνίας. Ο Earl Babbie (2010) ορίζει αυτή ως «τη μελέτη των καταγεγραμμένων μορφών ανθρώπινης επικοινωνίας, όπως τα βιβλία, οι ιστοσελίδες, οι πίνακες ζωγραφικής και οι νόμοι». Η ανάλυση περιεχομένου θεωρείται μια επιστημονική μεθοδολογία στις ανθρωπιστικές επιστήμες με την οποία τα κείμενα μελετούνται ως προς την πατρότητα, την αυθεντικότητα ή τη σημασία. Τα νεότερα αντικείμενα που περιλήφθηκαν είναι η φιλολογία, η ερμηνευτική και η σημειωτική. Οι βασικές ερωτήσεις της ανάλυσης περιεχομένου είναι «ποιος λέει τι, σε ποιον, γιατί, σε ποιο βαθμό και με ποιο αποτέλεσμα;». Ένας ευρύς ορισμός της ανάλυσης περιεχομένου είναι «οποιαδήποτε τεχνική για την εξαγωγή συμπερασμάτων μέσω αντικειμενικού και συστηματικού προσδιορισμού καθορισμένων χαρακτηριστικών των μηνυμάτων». Ένα άλλος ορισμός «έξι σημείων» της ανάλυσης περιεχομένου είναι: «Η ανάλυση περιεχομένου είναι μια περιληπτική, ποσοτική ανάλυση των μηνυμάτων που βασίζεται σε επιστημονική μέθοδο (που συμπεριλαμβάνει προσοχή στην αντικειμενικότητα, διυποκειμενικότητα, εκ των προτέρων σχεδιασμό, αξιοπιστία, εγκυρότητα, αναπαραγωγικότητα και έλεγχο υποθέσεων) και δεν

περιορίζεται ως προς τους τύπους των μεταβλητών που μπορεί να μετρηθούν ή το πλαίσιο στο οποίο τα μηνύματα δημιουργούνται ή παρουσιάζονται».

Η ανάλυση περιεχομένου χρησιμοποιεί αναφορές σε επιστημονικά έργα για τη δημιουργία συνδέσμων με άλλα έργα ή άλλους ερευνητές.

2.10. Βιβλιομετρικοί Χάρτες (Bibliometric Maps)

Τι σημαίνει ένας βιβλιομετρικός Χάρτης;

Ο βιβλιομετρικός χάρτης είναι μόνο μία θολή φωτογραφία των συνόρων ανάμεσα στους διαφορετικούς τομείς γνώσεων, που χρησιμεύει για γραφειοκρατική χρήση ή περιέχει συγκεκριμένες πληροφορίες, σχετικά με την πνευματική δομή εκείνων των περιοχών, ακριβώς, επειδή είχε σχεδιαστεί σύμφωνα με τις αντιλήψεις και το ενδιαφέρον του κάθε επιστημονικού κλάδου; Η πρωτοτυπία της συνεισφοράς του χάρτη δεν είναι μεγάλη. Πρέπει να τονίσουμε, όμως, το γεγονός ότι ο χάρτης δημιουργήθηκε χωρίς να ζητηθούν τα δεδομένα από μια ομάδα επιστημών, αλλά καταγράφοντας απλά διαμέσου αλγορίθμων την ανάκτηση δεδομένων, την αυθόρμητη συναίνεση ενός πολλαπλασιασμού των ανεξάρτητων στοιχείων της βιβλιογραφικής αναφοράς, τα οποία έχουν καταγραφεί σε ένα ευρετήριο αναφορών. Να σκεφτούμε παραδείγματος χάριν την υποστήριξη που θα μπορούσε να προσφέρει σε έναν ιστορικό επιστήμονα ένας βιβλιομετρικός χάρτης, που η ποιότητα της ανάγνωσης, η οποία παράγεται σε διάφορες θεματικές ενότητες, θα μπορούσε να φανεί ασυμβίβαστη ακόμα με τη διατύπωση των αρχικών ερμηνευτικών υποθέσεων.

Η πιο ενδιαφέρουσα απάντηση στις ερωτήσεις σχετικά με τους βιβλιομετρικούς χάρτες, η οποία προέρχεται από τη Βιβλιομετρία, βασίζεται στο επιστημονικό μοντέλο του Αμερικανού Φιλόσοφου Thomas Kuhn. Κατά τη γνώμη πολλών, οι βιβλιομετρικοί χάρτες θα ήταν μια εμπειρική προσέγγιση των παραδειγμάτων, που κατά Kuhn προσανατολίζουν τις δραστηριότητες της επιστημονικής έρευνας στους διαφόρους κλάδους της. Στο κουνικό μοντέλο η παραγωγή μιας νέας γνώσης δεν αποτελεί μόνο μια προοδευτική διαδικασία, αλλά μια σειρά υπομονετικών και μακρινών διαδικασιών, μιας «κανονικής» επιστήμης, οι οποίες διακόπτονται από βίαιες εννοιολογικές μεταμορφώσεις ή από «επιστημονικές» επαναστάσεις (Kuhn, 1962).

Αντίθετα ο Merton υποστηρίζει ότι οι επιστήμονες ακολουθούν έναν ηθικό κώδικα και η συμπεριφορά τους προς τις υπάρχουσες θεωρίες είναι κυρίως συντηρητική, και σε σχέση με τις νέες ιδέες αλλά και σε σχέση με τα άτομα από τα οποία προέρχονται. Αντίθετα με κάθε κριτική σκέψη, η καθημερινή έρευνα τους βασίζεται λίγο πολύ στη λύση πάζλ, δηλαδή μιας συλλογής θεωριών, μεθόδων, τεχνικών, κοινών αξιών, οι οποίες σχηματίζουν μια επιστημονική βάση για την έρευνα.

Οντας σε τέλεια αρμονία με τον ρόλο των αναφορών στην αξιολόγηση της έρευνας, οι χάρτες μας δίνουν, θα λέγαμε, περισσότερο ερωτήσεις παρά απαντήσεις, υποδεικνύοντας μεν κάτι, αλλά όχι ακριβώς αυτό που θα θέλαμε να υποδεικνύουν. Το διαφορούμενο δεν είναι αντιληπτό μόνο από την εξωτερική πείρα, από μια κρίσιμη και φιλοσοφική προοπτική, αντιθέτως διασχίζει όλη τη βιβλιομετρική παραγωγή σε όλη την έκταση της. Πράγματι, μερικές ποσοτικές αναλύσεις υποδεικνύουν τη θετική συσχέτιση μεταξύ παραδειγματικών αλλαγών στην ιστορία της επιστήμης και μετασχηματισμών στις βιβλιογραφικές συνήθειες των ερευνητών, μέχρι ακόμα και στο να αντιγράφεται η δυναμική των επιστημονικών θεωριών, από τη μορφή βιβλιομετρικών, ελκυστικών οπτικά χαρτών, σε όρους κύκλων εργασιών των παραδειγμάτων και των εννοιολογικών καταστάσεων. Άλλες έρευνες υποδεικνύουν, αντιθέτως, ότι η σύνθεση των συστάδων (cluster) των τεκμηρίων μας παραπέμπει περισσότερο από τη σύγκλιση των ιδεών, στις κοινωνικές δυναμικές μερικών ομάδων επιστημόνων και φορέων που είναι υπεύθυνοι για τις εκδόσεις αυτές. Η υπόθεση γίνεται ακόμα πιο σύνθετη, αν η ανάλυση επικεντρώνεται στις συν-εμφανίσεις των λέξεων, αντί των βιβλιογραφικών αναφορών.

Τη δεκαετία του '80 μια ομάδα επιστημόνων της Ecole Nationale Supérieure des Mines {Εθνική Σχολή Ορυχείων} του Παρισιού εγκαίνιασε τη co-word analysis που ορίζεται ως μια στατιστική ανάλυση ζευγών λέξεων, που προέρχονται από τίτλους επιστημονικών άρθρων. Εμπνευσμένη από τη θεωρία του «ενδοδουκτικού δρώντος» ή «ANT» (Actor Network Theory) του Μπρούνο Λατούρ, η αναγνώριση και η ανάλυση συνδυασμών λέξεων έχει ως στόχο να αναδείξει τους μηχανισμούς που χρησιμοποιούνται για τη διευκόλυνση της οικοδόμησης των επιστημονικών κειμένων (Callon et al., 1986). Οι συχνές ομάδες λέξεων, που προσδιορίζονται από μία πολυπαραγοντική ανάλυση, γίνονται, πράγματι, εδώ μια υλοποίηση των προβληματικών δικτύων που εισάγονται από τους συγγραφείς για να μετακινηθεί η ισορροπία της εξουσίας στο δείκτη της επιστημονικής επικοινωνίας. Ως εκ τούτου, οι βιβλιογραφικές αναφορές, φαίνεται να

αφηγούνται πολύ διαφορετικές βιβλιομετρικές ιστορίες: οι πρώτες μιλούν για επιρροή, ποσότητα στην προοπτική των κανόνων αλλά στη συνέχεια υπόσχονται ένα πλεονεκτικό σημείο παρατήρησης πάνω στην φιλοσοφία του Kuhn, κάτω από το μεγεθυντικό φακό της co-citation analysis, οι περισσότερες μιλάνε για το σχετικισμό, τη ρητορική και τα παιχνίδια εξουσίας. Οι δύο πόλοι φαίνονται ασυμβίβαστοι, παρόλο που υπήρξε μια συναισθηματικά εκλεπτυσμένη προσπάθεια του Loet Lydesdorff (2001) να συνυπάρξουν σε ένα ενιαίο θεωρητικό πλαίσιο, υπό την έννοια της μαθηματικής θεωρίας της επικοινωνίας.

Συμπέρασμα: η Βιβλιομετρία ανιχνεύει επιπροσθέτως την εσωτερική δομή της επιστήμης (βιβλιομετρικοί χάρτες) και υποδεικνύει την αναγκαιότητα επιστημονικής αξίας, όπου και αν εμφανιστούν ίχνη (οι δείκτες που βασίζονται στις αναφορές).

Άξιο αναφοράς είναι ότι ουδέποτε δεν επετεύχθη μια καθοριστική συμφωνία ανάμεσα στους υποστηρικτές και τους επικριτές των βιβλιομετρικών δεικτών, όσον αφορά στην αξιολόγηση της ποιότητας μιας έρευνας. Υπάρχουν πάρα πολλά επιχειρήματα υπέρ της υπεροχής των ποιοτικών κρίσεων (αποφάσεων) και από την άλλη μεριά, πολλά παραδείγματα που αποδεικνύουν τα σφάλματα των ίδιων κρίσεων και την ανάγκη παραμέτρων, λιγότερο υποκειμενικών, καθώς δεν υπάρχει (δεν μπορεί να υπάρχει) ένα στοιχείο καθολικά έγκυρο που να αποδεικνύει την ανωτερότητα των βιβλιομετρικών αξιολογήσεων σχετικά με κάποιες άλλες μορφές αξιολόγησης. Οι επικριτές θα επιμένουν ότι οι μονάδες ανάλυσης της βιβλιομετρίας είναι εύθραυστες, χειραγωγούνται εύκολα και υπακούν στους νόμους που αντανakλούν τόσο τη συγκέντρωση της συμβολικής εξουσίας στην αγορά της επιστημονικής επικοινωνίας, όσο και τη συγκέντρωση οικονομικής και χρηματοπιστωτικής εξουσίας στον κόσμο των εκδόσεων (νόμος Bradford).

Βιβλιογραφία/Αναφορές

- Earle, P., & Vickery, B. (1969). Social science literature use in the UK as indicated by citations. *Journal of Documentation*, 25(2), 123–141.
- Edge, D. (1977). Why I am not a co-citationist. *Society for Social Studies of Science Newsletter*, 2, 13–19.
- Garfield, E. (1979). Controversies over opiate receptor research typify problems facing awards committees. *Current Contents*, (20), 5–18.
- Garfield, E. (1974). *ISI is Studying Structure of Science Through Co-Citation Analysis*. INST SCI INFORM INC 3501 MARKET ST, PHILADELPHIA, PA 19104.
- Glanzel, W. (2003). Bibliometrics as a research field a course on theory and application of bibliometric indicators. Retrieved from <http://nsdl.niscair.res.in/jspui/handle/123456789/968>
- Hicks, D. (1995). Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change*, 4(2), 401–424.
- Hicks, D., & Potter, J. (1991). Sociology of scientific knowledge: a reflexive citation analysis or science disciplines and disciplining science. *Social Studies of Science*, 21(3), 459–501.
- Holmes, A., & Oppenheim, C. (2001). Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management. *Information Research*, 6(2), 6–2.
- Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago press. Retrieved from https://www.google.com/books?hl=el&lr=&id=3eP5Y_OOuzwC&oi=fnd&pg=PR5&dq=The+Structure+of+Scientific+Revolutions+&ots=xU0JzboRvO&sig=K0FBI3rlRCH1wIfvTpP1n09mnfY
- Lenoir, M. (1979). Promotion of patent literature. *World Patent Information*, 1(2), 84–87.
- MacRoberts, M. H., & MacRoberts, B. R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, 61(1), 1–12.
- Martyn, J. (1964). Bibliographic coupling. *Journal of Documentation*, 20(4), 236–236.
- Sala-i-Martin, X. X. (1997). *I just ran four million regressions*. National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w6252>
- Meho, L. I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: a comparison of Scopus and Web of Science. *Journal of the American Society for Information Science and Technology*, 59(11), 1711–1726.
- Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105–2125.
- Moed, H. F. (2006). *Citation analysis in research evaluation* (Vol. 9). Springer Science & Business Media. Retrieved from <https://www.google.com/books?hl=el&lr=&id=D9SaJ6awy4gC&oi=fnd&pg=PR9&dq=Citation+analysis+in+research+evaluation.&ots=FGmUKy6NjX&sig=YJ1go9ntOthq5Zqz1lmEJLIkqGY>
- Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica Scandinavica*, 69(3), 224–229.
- Schiminovich, S. (1971). Automatic classification and retrieval of documents by means of a bibliographic pattern discovery algorithm. *Information Storage and Retrieval*, 6(6), 417–435.

- Simkin, M. V., & Roychowdhury, V. P. (2012). Theory of citing. In *Handbook of Optimization in Complex Networks* (pp. 463–505). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4614-0754-6_16
- Small, H., & Griffith, B. C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies*, 17–40.
- Small, H. G. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science*, 139–166.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269.
- Tagliacozzo, R., & others. (1967). Citations and citation indexes: a review. *Methods Inf Med*, 6, 136–142.
- van Raan, A. F. (1998). *Special Topic Issue: Science and Technology Indicators*. American Society for Information Science (ASIS).
- White, H. D., & Griffith, B. C. (1982). Authors as markers of intellectual space: Co-citation in studies of science, technology and society. *Journal of Documentation*, 38(4), 255–272.
- Wallin, J. A. (2005). Bibliometric methods: pitfalls and possibilities. *Basic & Clinical Pharmacology & Toxicology*, 97(5), 261–275.

Κεφάλαιο 3

Πρακτικός οδηγός κατασκευής και χρήσης Βιβλιομετρικών Δεικτών και Νόμων

Σύνοψη

Σε αυτό το κεφάλαιο δίδονται παραδείγματα μέτρησης και αξιολόγησης δεικτών περιοδικών και ερευνητών. Αναφέρονται τα προβλήματα που ανακύπτουν από τη χρήση των δεικτών αυτών και προτείνονται οι τρόποι αντιμετώπισής τους. Επίσης αναφέρονται διεξοδικά οι βιβλιομετρικοί δείκτες με αναλυτικά παραδείγματα για πλήρη κατανόηση εφαρμογής τους, καθώς και συνδυασμοί αυτών, ώστε να προκύψουν δείκτες πιο αντιπροσωπευτικοί για κάθε περίπτωση. Ακολούθως αναδεικνύεται ότι οι δείκτες, εκτός των άλλων, παρέχουν μετρήσεις των συνδέσεων μεταξύ των ερευνητών και των ερευνητικών περιοχών, μέσω της στατιστικής ανάλυσης των συν-δημοσιεύσεων και των αναφορών.

Εισαγωγή

Ο όρος βιβλιομετρικός δείκτης χρησιμοποιείται συχνά για τα αποτελέσματα μιας βιβλιομετρικής ανάλυσης. Ένας από τους ορισμούς του όρου δείκτη στο αγγλικό λεξικό της Οξφόρδης είναι: "Αυτό που χρησιμεύει για να δείξει ή να δώσει μια εξήγηση για κάτι" (Αγγλικό λεξικό της Οξφόρδης, 2000). Ο ορισμός αυτός εφιστά την προσοχή στο ότι τα αποτελέσματα περιγράφουν μια πραγματικότητα που είναι αρκετά πολύπλοκη για να μετρηθεί απλώς με στατιστικές ή αριθμούς. Σε ένα γλωσσάριο που παράγεται από την κοινή επιτροπή για τη διαπίστευση των Οργανισμών Υγείας και Φροντίδας στις ΗΠΑ, υπάρχει ένας άλλος ορισμός που φαίνεται πιο ενδεδειγμένος στον τρόπο με τον οποίο η λέξη δείκτης χρησιμοποιείται στη βιβλιομετρία: "1. Ένα μέτρο που χρησιμοποιείται για να καθορίσει διαχρονικά την εκτέλεση των καθηκόντων, τις διαδικασίες και τα αποτελέσματα. 2. Μία στατιστική αξία που παρέχει μια ένδειξη για την κατάσταση ή την κατεύθυνση διαχρονικά της εκτέλεσης μιας καθορισμένης διαδικασίας ή επίτευξης ενός καθορισμένου αποτελέσματος."

3.1 Δείκτες μέτρησης και αξιολόγησης περιοδικών

3.1.1 Η υπηρεσία Journal Citation Reports

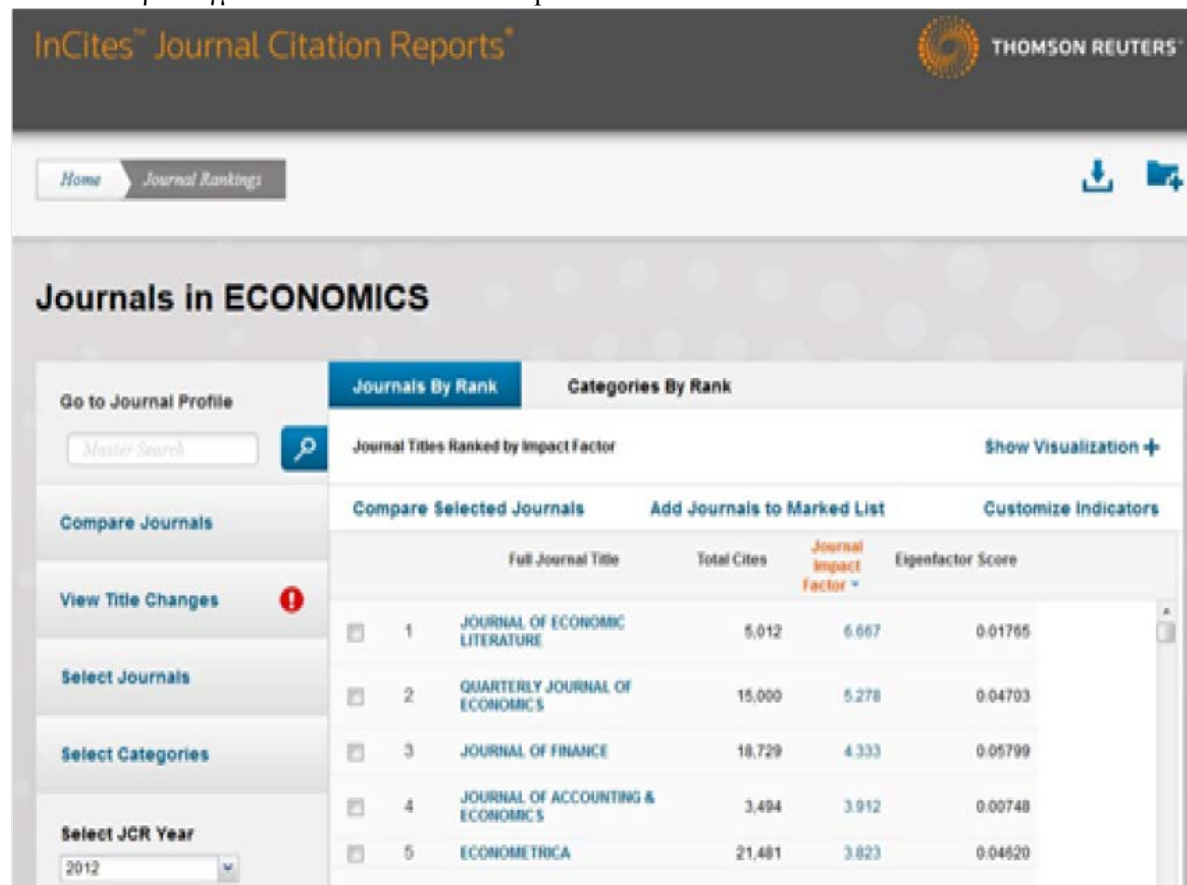
Η υπηρεσία (Thomson Scientific: Citation Products. Retrieved April 6th, 2015, from <http://scientific.thomson.com/products/categories/citation/>) που χρησιμοποιεί την πλατφόρμα Web of Science είναι ένα μοναδικό εργαλείο που επιτρέπει στον χρήστη την αξιολόγηση και τη σύγκριση περιοδικών, χρησιμοποιώντας δεδομένα από βιβλιογραφικές αναφορές περισσότερων από 12000 διεθνών επιστημονικών περιοδικών, από 3.400 και πάνω εκδότες, σε περισσότερες από 65 χώρες. Οι αναφορές των άρθρων είναι αξιόλογος δείκτης για το πόσο συχνά χρησιμοποιείται κάθε περιοδικό από τους επιστήμονες. Τα δεδομένα των αναφορών της JCR εκδίδονται ετησίως σε δύο εκδόσεις:

- JCR Science Edition, με δεδομένα από περίπου 6000 περιοδικά στις Επιστήμες Υγείας τις Φυσικές και Εφαρμοσμένες Επιστήμες, και στην Τεχνολογία.
- JCR Social Sciences Edition, με δεδομένα από περίπου 1.800 περιοδικά στις Ανθρωπιστικές Κοινωνικές, και Οικονομικές Επιστήμες.
 - ο Οι JCR δείκτες ουσιαστικά περιλαμβάνουν κάθε τομέα των Κοινωνικών Επιστημών και της Επιστήμης, της Τεχνολογίας. Με την δυνατότητα παρουσίασης ανά πεδίο:
 - των περιοδικών στα οποία υπάρχουν οι περισσότερες βιβλιογραφικές αναφορές (Most Frequently Cited Journals),
 - των περιοδικών με τον υψηλότερο αντίκτυπο (Highest Impact Journals),
 - των σπουδαιότερων περιοδικών (Hottest journals),
 - των μεγαλύτερων περιοδικών (Largest journals).

Σκοπός της εφαρμογής είναι η προβολή με ένα αντικειμενικό και συστηματικό τρόπο του προσδιορισμού της επιρροής των περιοδικών στον τομέα της επιστήμης, τον οποίο κάθε ένα υπηρετεί. Οι πληροφορίες της εφαρμογής παρέχουν το Δείκτη Επιρροής (Impact Factor) – Βλέπε. επόμενη ενότητα- και το Ευρετήριο Αμεσότητας (Immediacy Index).

Το Ευρετήριο Αμεσότητας αποτελεί το μέσο όρο των περιπτώσεων που έγινε αναφορά σε ένα δημοσίευμα το έτος κατά το οποίο δημοσιεύτηκε. Δηλαδή πόσο γρήγορα, το μέσο άρθρο ενός περιοδικού γίνεται αντικείμενο αναφοράς. Το ευρετήριο αμεσότητας έχει τη δυνατότητα εφαρμογής αφ' ενός μεν σε επίπεδο περιοδικού και αφ' ετέρου σε θεματικό επίπεδο.

Παράδειγμα του Journal Citation Report



Εικόνα 3 1 Web of Science Δείκτη ζεπιρροής (Impact Factor) περιοδικού

Πολλοί καθηγητές, όπως και μεταπτυχιακοί φοιτητές, θα επιθυμούσαν να γνωρίζουν την απήχηση που έχουν τα πνευματικά δημιουργήματά τους, κοινώς τα "άρθρα" τους. Αν και ουδείς το παραδέχεται, σίγουρα όλοι επιθυμούν να μάθουν αφ' ότου δημοσιοποιηθεί το άρθρο τους: "Θα είναι αρκετοί εκείνοι που θα το διαβάσουν; Είναι το περιοδικό το ποιο κατάλληλο για την δημοσίευσή τους;"

Ο Δείκτης επιρροής (IF-Impact Factor) αποτελεί ένα διαδεδομένο, αλλά όχι το μοναδικό, κριτήριο αξιολόγησης της αξιοπιστίας και απήχησης ενός επιστημονικού περιοδικού. Σε γενικές γραμμές υπολογίζεται από τις αναφορές, που υπάρχουν σε όλες τις εργασίες, του κάθε περιοδικού. Όσο πολυπληθέστερες είναι οι αναφορές σε κάθε εργασία του επιστημονικού περιοδικού, τόσο ποιο έγκυρο μπορεί να θεωρηθεί το περιοδικό αυτό.

Ο Δείκτης επιρροής του Web of Science, ο οποίος κατέχει το μονοπώλιο στην αξιολόγηση της ποιότητας των περιοδικών, σχεδιάστηκε από τον Eugene Garfield, γύρω στο 1960, ως ένα μέσο για να μετρηθεί η επιρροή ενός περιοδικού και δίνει μια μέση τιμή για το πόσες φορές ένα άρθρο στο περιοδικό έχει αναφερθεί. Υπολογίζεται ως ο μέσος αριθμός των αναφορών, που κατανέμονται σε ένα συγκεκριμένο έτος για τα άρθρα που δημοσιεύονται στο περιοδικό για τα δύο προηγούμενα έτη, διαιρούμενος με τον αριθμό των άρθρων που δημοσιεύονται στο περιοδικό σε αυτά ακριβώς τα δύο χρόνια. Το σύντομο χρονικό διάστημα και η χρήση μιας αναλογίας, αντί μιας συνολικής ευρετηρίασης των αναφορών, προορίζονταν για να ομαλοποιηθούν επιδράσεις που σχετίζονται με την ηλικία του άρθρου και τον χρόνο δημοσίευσής του.

3.1.1.1 Παράδειγμα εφαρμογής και Σχόλια σχετικά με τον Δείκτη

Αν πάρουμε τα δεδομένα του παρακάτω πίνακα για το περιοδικό “Journal of clinical immunoassay”, ο δείκτης IF, για το έτος 1997, υπολογίζεται ως το πηλίκο $(34+32)/(71+72)=0.462$.

Έτη	1995	1996	1997
Αριθμός αναφορών εργασιών	34	32	9
Αριθμός εργασιών	71	72	68

Πίνακας 3.1 Παράδειγμα εφαρμογής δείκτη επιρροής (IF)

Δηλαδή, τα άρθρα του περιοδικού, ένα ή δύο χρόνια πριν, είχαν κατά μέσο όρο μισή αναφορά το καθένα. Χρησιμοποιείται από το Web of Science για να εντοπιστούν τα πιο σημαντικά επιστημονικά περιοδικά σε κάθε τομέα της έρευνας και, στο πλαίσιο αυτό, χρησιμοποιείται και ως εργαλείο διαχείρισης των συλλογών των βιβλιοθηκών. Ο συντελεστής επιρροής συχνά γράφεται ως ένας αριθμός που αντιστοιχεί στον μέσο αριθμό των αναφορών σε ένα άρθρο του περιοδικού (συμπεριλαμβανομένων των διαφόρων τύπων δημοσιεύσεων ανασκόπησης και σημειώσεων). Τα περιοδικά με τους υψηλότερους δείκτες επιρροής μπορεί να φτάσουν περίπου τον αριθμό 50, ενώ περιοδικά όπως το New England Journal of Medicine, Nature και Science έχουν συντελεστές επιρροής περίπου 30. Τα περισσότερα περιοδικά έχουν συντελεστή επιρροής κάτω από τη μονάδα. Το γεγονός ότι ο συντελεστής επιρροής ISI βασίζεται στις αναφορές μόνο 1-2 ετών μπορεί να θεωρηθεί ως ένας συμβιβασμός μεταξύ της ανάγκης του να πάρουμε μια γρήγορη εκτίμηση των νέων περιοδικών και να αφήσουμε τις δημοσιεύσεις να φτάσουν στο μέγιστο αριθμό αναφορών τους (το έτος κατά το οποίο η δημοσίευση λαμβάνει το μεγαλύτερο αριθμό αναφορών τους). Τα περισσότερα άρθρα φθάνουν στο μέγιστο αριθμό αναφορών τους 3-5 χρόνια μετά τη δημοσίευσή τους, έτσι ώστε, για πολλούς τομείς έρευνας, να είναι το προτιμότερο παράθυρο αναφορών. Τα πρότυπα αναφορών διαφέρουν τόσο πολύ μεταξύ των διαφόρων ερευνητικών πεδίων, που ο παράγοντας των αναφορών δεν θα πρέπει να χρησιμοποιείται για να συγκρίνει τις επιστημονικές επιρροές των περιοδικών σε διάφορους θεματικούς τομείς. Δεδομένου ότι ο συντελεστής επιρροής είναι σχετικά εύκολο να βρεθεί και να κατανοηθεί, έχει γίνει πολύ δημοφιλής και χρησιμοποιείται συχνά για να αξιολογήσει την ποιότητα των άρθρων, μελετώντας τα περιοδικά στα οποία δημοσιεύουν οι ερευνητές, τα πανεπιστημιακά τμήματα και τα πανεπιστήμια. Τούτο είναι κάτι που αντενδείκνυται μιας και ο συντελεστής επιρροής δεν προοριζόταν για αυτό. Ωστόσο, αν θέλετε να μελετήσετε πολύ πρόσφατα δημοσιευμένα άρθρα που δεν έχουν ακόμη αναφερθεί, ο δείκτης επιρροής περιοδικού μπορεί να είναι η μόνη δυνατή περίπτωση δείκτη επιδόσεων. Ο δείκτης είναι τότε βασισμένος στην υπόθεση ότι η διαδικασία της αναφοράς είναι πιο αυστηρή σε περιοδικά υψηλού δείκτη επιρροής, που σημαίνει ότι μόνο υψηλής ποιότητας έρευνα θα γίνεται δεκτή.

Στον αντίλογο ο Sawnsen (2004), αναφέρει πως όσο μεγαλύτερος ο ανταγωνισμός για δημοσιεύσεις σε υψηλού κύρους περιοδικά, οι επιμελητές των περιοδικών αυτών τείνουν να δίνουν πολύ μεγαλύτερη έμφαση στην τεχνική αρτιότητα των υποβαλλόμενων εργασιών παρά στο περιεχόμενο των λιγότερο άρτια εργασιών που ενδεχομένως να είναι φορείς πιο σημαντικών ιδεών. Ενώ ο συνολικός αριθμός αναφορών δεν μπορεί να παράσχει μία συγκριτική εικόνα μεταξύ μεγάλων (με μεγαλύτερο αριθμό εκδόσεων ανά έτος) και μικρών περιοδικών, ο Δείκτης επιρροής τείνει να αμβλύνει το πλεονέκτημα των μεγάλων περιοδικών έναντι των μικρότερων περιοδικών που δημοσιεύονται λίγες φορές το χρόνο, καθώς και των παλαιότερων έναντι των νεότερων περιοδικών.

3.1.2. Προβλήματα εφαρμογής του Δείκτη επιρροής

Ο Δείκτης επιρροής έχει αρκετά προβλήματα, στην χρησιμοποίησή του για την αξιολόγηση ενός περιοδικού. Ας αρχίσουμε από το ίδιο το κλάσμα (Βλ. προηγούμενο παράδειγμα πίνακα 3.1). Ο παρονομαστής αναφέρεται σε «δημοσιεύσεις στο περιοδικό X το 1995-1996», δηλαδή το σύνολο των δημοσιεύσεων για το χρονικό αυτό διάστημα. Όμως όλες οι δημοσιεύσεις που φιλοξενεί ένα περιοδικό δεν λαμβάνονται υπόψη αλλά χρησιμοποιούνται για την εξαγωγή του δείκτη μόνο αυτές που κρίνει η Web of Science ότι είναι άξιες να αναφερθούν. Σε αρκετά περιοδικά όμως φιλοξενούνται και δημοσιεύσεις οι οποίες δεν είναι πρωτότυπες εργασίες ή ανασκοπήσεις, αλλά σχολιασμοί, ειδήσεις κ.ά., τα οποία ευνοούνται αφού λαμβάνονται υπόψη οι αναφορές προς αυτά μεγαλώνοντας τον αριθμητή ενώ διαιρούνται με τον αριθμό μόνο των «ερευνητικών άρθρων και των άρθρων επισκόπησης που έχουν δημοσιευθεί σε αυτό. Έτσι όμως, κάποιοι σχολιασμοί μπορεί

να αναφερθούν από άλλες δημοσιεύσεις και άρα, ο αντίκτυπος του περιοδικού να μην εκτιμηθεί δεόντως κατά αυτό τον τρόπο. Όμως το πρόβλημα μεγαλώνει όταν δεν αναφέρονται και πρωτότυπες εργασίες. Τούτο συμβαίνει στις περιπτώσεις που τα περιοδικά περιλαμβάνουν εργασίες, οι οποίες ηθελήμενα υποβαθμίζονται με σκοπό να μην συμπεριληφθούν στον παρονομαστή. Τούτο επιτυγχάνεται με πολλούς τρόπους π.χ. όταν τα άρθρα είναι εξαιρετικά μικρά σε έκταση ή δεν διαθέτουν περίληψη (abstract) ή έχουν πολύ μικρό αριθμό βιβλιογραφικών αναφορών. Έτσι η Web of Science μπορεί να θεωρήσει ότι οι δημοσιεύσεις αυτές δεν είναι κατάλληλες για αναφορά και έτσι με τον τρόπο αυτό μικραίνει ο παρονομαστής και άρα αυξάνεται το κλάσμα. Βέβαια μπορεί οι συγκεκριμένες δημοσιεύσεις που υποβαθμίστηκαν να πετύχουν πολλές αναφορές και έτσι το περιοδικό τελικά να μην αυξήσει το Δείκτη επιρροής, κάτι που είναι όμως πολύ σπάνιο.

Πρόβλημα υπάρχει επίσης και με τον αριθμητή. Αφού όσο μεγαλύτερος είναι ο αριθμητής σε σχέση με τον παρονομαστή, τόσο το κλάσμα μεγαλώνει. Επομένως οι συντάκτες (editors) των περιοδικών σε αρκετές περιπτώσεις ασκούν πίεση στους συγγραφείς να κάνουν αναφορές σε δημοσιεύσεις του περιοδικού τους, ώστε να αυξηθεί το κλάσμα. Έτσι, οι ίδιοι οι συγγραφείς των άρθρων αντιλαμβάνονται ότι, δημιουργώντας πολλές αναφορές άρθρων δημοσιευμένων στο συγκεκριμένο περιοδικό, θα είχαν περισσότερες ευκαιρίες να αυξηθεί η πιθανότητα το άρθρο τους να γίνει αποδεκτό. Υπάρχουν, βέβαια, πολλοί τρόποι για να αυξήσουμε το Δείκτη επιρροής δια μέσου του αριθμητή, αν παραλάβουμε π.χ. αρκετές ανασκοπήσεις στο περιοδικό. Αξίζει να σημειωθεί ότι τα περιοδικά με τον μεγαλύτερο Δείκτη επιρροής είναι έντυπα που περιέχουν μόνο ανασκοπήσεις, όπως το *Physiological Reviews*, με Δείκτη επιρροής που κυμαίνεται στο 37. Όπως καταλαβαίνει κανείς, είναι άδικο να συγκρίνουμε ένα περιοδικό που περιλαμβάνει ποιοτικές πρωτότυπες εργασίες με κάποιο που δημοσιεύει μόνο ανασκοπήσεις με βάση το Δείκτη επιρροής.

Υπάρχουν, όμως, και άλλα σημαντικά προβλήματα που προκύπτουν στον αριθμητή. Ο αριθμός των αναφορών δεν αντικατοπτρίζει υποχρεωτικά και την αξία μιας δημοσίευσης, αλλά περισσότερο τη δημοτικότητα της. Για παράδειγμα μια δημοσίευση στον καρκίνο ή το AIDS μπορεί να συλλέξει πολύ περισσότερες αναφορές σε σχέση με μια δημοσίευση για τη μυοπάθεια, άσχετα από την ποιότητα της έρευνας, μόνο και μόνο διότι τόσο ο καρκίνος όσο και το AIDS είναι πολύ πιο δημοφιλή θέματα από τη μυοπάθεια. Βέβαια, ο ίδιος ο όρος Δείκτης επιρροής υπονοεί ότι αξιολογεί τον αντίκτυπο και όχι την ποιότητα, αλλά αυτό το στοιχείο φαίνεται πολύ συχνά να λησμονείται από τους ίδιους τους ερευνητές. Περιοδικά με βιοϊατρικό περιεχόμενο έχουν γενικά υψηλότερο Δείκτη επιρροής από περιοδικά άλλων επιστημών, λόγω του μεγάλου ενδιαφέροντος που υπάρχει για την ιατρική επιστήμη. Επομένως, ορθά, ο δείκτης αυτός μπορεί να μετρά τον αντίκτυπο, αλλά πρέπει να θυμόμαστε ότι δεν αποτελεί κριτήριο ποιότητας. Για παράδειγμα το περιοδικό *New England Journal of Medicine* έχει Δείκτη επιρροής που κυμαίνεται γύρω στο 36-37, μεγαλύτερο από περιοδικά που παραδοσιακά δημοσιεύουν ποιοτική, πρωτότυπη, πολυθεματική, εργαστηριακή έρευνα όπως το *Science*, το *Nature* ή το *Cell*. Αυτό δεν αποδεικνύει ότι το πρώτο έχει μεγαλύτερη αξία από τα προαναφερθέντα περιοδικά, απλώς αναδεικνύει ότι η κλινική έρευνα είναι αρκετά πιο δημοφιλής από την εργαστηριακή, άσχετα από τον πραγματικό αντίκτυπο (π.χ. οι δημοσιεύσεις για τη μελέτη του ανθρώπινου γονιδιώματος έγιναν στη συντριπτική πλειοψηφία στο *Science* και το *Nature* με τεράστιο αντίκτυπο στην επιστήμη). Επίσης ένα ακόμη ενδιαφέρον στοιχείο είναι ότι στο Δείκτη Επιρροής δεν αντικατοπτρίζεται, σε πολλές περιπτώσεις, ο πραγματικός αντίκτυπος ενός περιοδικού, ιδίως αν το περιοδικό έχει πλούσιο, διδακτικό περιεχόμενο. Μπορεί, δηλαδή, να έχει αντίκτυπο σε φοιτητές, ειδικευόμενους, αλλά, λόγω μικρού αριθμού αναφορών, ο Δείκτης επιρροής να υποβαθμίζει τη βαρύτητά του στην επιστημονική κοινότητα.

Επίσης, όσον αφορά τα νέα περιοδικά για να ενταχθεί ένα νέο περιοδικό στο JCR και να αποκτήσει συντελεστή απήχησης πρέπει να περάσουν τουλάχιστον 6 χρόνια μία 3ετία υποχρεωτικής αναμονής και μετά 3ετής περίοδος εξέτασης. Έτσι πολλά άρθρα με καινοτόμες ιδέες που δημοσιεύονται σε νέα περιοδικά μένουν για μεγάλο χρονικό διάστημα εκτός των βάσεων δεδομένων αναφοράς.

Ο συντελεστής απήχησης πολλές φορές δεν ανταποκρίνεται στη σημασία που έχει ένα περιοδικό στην επιστήμη που εκπροσωπεί και ευνοεί διεπιστημονικά περιοδικά. Τα περιοδικά των οικονομικών της υγείας παραδείγματος χάριν κατατάσσονται ψηλά στην κατάταξη των περιοδικών της οικονομικής επιστήμης αφού για τον υπολογισμό του συντελεστή λαμβάνονται υπόψη οι αναφορές από όλα τα περιοδικά που περιλαμβάνονται στο WoS και όχι μόνον από τις αναφορές των περιοδικών της οικονομικής επιστήμης.

Επίσης οι αναφορές των εργασιών που δημοσιεύονται σε ένα συγκεκριμένο περιοδικό έχει αποδειχθεί ότι έχουν άμεση σχέση με τον παράγοντα του αντικτύπου του εντύπου. Ένα άλλο πρόβλημα με τη χρήση του παράγοντα του αντικτύπου για την αξιολόγηση του αντικτύπου της έρευνας είναι το γεγονός ότι υπάρχει, συχνά, ένα χρονικό διάλειμμα που μεσολαβεί πριν οι αναφορές σ' ένα άρθρο ξεκινήσουν να εμφανίζονται με μέγιστο αριθμό αναφορών ανά χρόνο, συνήθως μεταξύ 3 και 7 χρόνων μετά τη δημοσίευση.

3.1.3. Αντιμετώπιση των προβλημάτων από την εφαρμογή του Δείκτη

Για την αντιμετώπιση των προβλημάτων που ανακύπτουν από τον Δείκτη επιρροής, γίνεται μία προσπάθεια τελευταία να καθιερωθούν άλλοι δείκτες, πιο αντιπροσωπευτικοί, της αξιολόγησης και της ποιότητας ενός περιοδικού.

3.1.3.1. Υ δείκτης

Ένας δείκτης που προσπαθεί να καλύψει τα κενά του IF είναι ο «Υ δείκτης», αυτός ο δείκτης δεν χρησιμοποιεί μόνο τον αριθμό των αναφορών, όπως ο Δείκτης Επιρροής. Ο Υ δείκτης στηρίζεται, σε μεγάλο βαθμό, στον γνωστό αλγόριθμο Page Rank της Google. Ο Page Rank χρησιμοποιείται για την κατάταξη των αποτελεσμάτων αναζήτησης από τη Google. Στο Page Rank προσμετράτε ο αριθμός των ιστοσελίδων που κάνουν σύνδεση (link) σε μια ιστοσελίδα, αλλά και ο αριθμός των ιστοσελίδων που κάνουν συνδέσεις στις ιστοσελίδες αυτές κ.ό.κ. Κατά αυτόν τον τρόπο στην κατάταξη της Google δε μετράνε μόνο οι συνδέσεις που γίνονται σε μια ιστοσελίδα, αλλά και ποιά ιστοσελίδα κάνει τις συνδέσεις (δηλαδή θα ανέβει στην κατάταξη μια ιστοσελίδα αν την προτείνει για σύνδεση μια δημοφιλή ιστοσελίδα). Επομένως υπολογίζεται, στην ουσία, η δημοτικότητα της σελίδας. Οι δημιουργοί του Υ δείκτη υποστηρίζουν ότι με τον τρόπο αυτό δίνεται μια αντικειμενικότερη εικόνα της ποιότητας ενός περιοδικού, γιατί δεν υπολογίζεται μόνο ο συνολικός αριθμός των αναφορών, αλλά και ποιός κάνει την αναφορά. Δηλαδή αναφορές από έναν ερευνητή, που έχει πολλές αναφορές στο ενεργητικό του θα είναι πολύ μεγαλύτερης σημασίας από μια αναφορά από έναν επιστήμονα, ο οποίος έχει μικρό αριθμό αναφορών. Ο Υ δείκτης δεν εγκαταλείπει πλήρως τον Δείκτη Επιρροής, καθώς ο συνολικός αριθμός αναφορών θα προσμετράτε, απλά δε θα είναι ο αποκλειστικός κριτής της αξίας του περιοδικού. Να θυμίσουμε ότι μέσω της Google Scholar, ήδη, εφαρμόζεται ένα είδος άτυπης κατάταξης δημοσιεύσεων, καθώς χρησιμοποιείται το Page Rank.

Αρα, η κατάταξη του περιοδικού με βάση έναν δείκτη θα είναι κάθε φορά μια διαδικασία με ένα ποσοστό αυθαιρεσίας. Ίσως αν συνυπολογιστούν διάφοροι δείκτες, μεταξύ των οποίων και ο Δείκτης επιρροής, να έχουμε ένα πιο έγκυρο αποτέλεσμα.

3.1.4. Βαθμολογία Eigenfactor

Ένας άλλος δείκτης προς την κατεύθυνση αυτής της προσπάθειας, της αντιμετώπισης δηλαδή των προβλημάτων που προκύπτουν από τον Δείκτη επιρροής, είναι η βαθμολογία *Eigenfactor*.

Η βαθμολογία *Eigenfactor* προέκυψε από το ερευνητικό έργο που διεξήχθη το 2008, από το εργαστήριο του καθηγητή Carl Bergstrom, του Πανεπιστήμιου της Ουάσιγκτον. Η Βαθμολογία *Eigenfactor* μετρά τον αριθμό των φορών που άρθρα από το περιοδικό που εκδίδεται κατά τα τελευταία πέντε χρόνια έχουν αναφερθεί, κατά το έτος JCR. Η Βαθμολογία *Eigenfactor* ενός περιοδικού δείχνει την επιρροή του στην ερευνητική κοινότητα.

Σε αντιδιαστολή με τον Impact Factor, η βαθμολογία *Eigenfactor*:

- Μετρά τις αναφορές τόσο στις Κοινωνικές επιστήμες όσο και στις Θετικές.
- Αποκλείει τις αυτό-αναφορές. Καθώς επίσης κάθε αναφορά από ένα άρθρο ενός περιοδικού σε ένα οποιοδήποτε άλλο άρθρο του ίδιου περιοδικού δεν λαμβάνεται υπ' όψιν.
- Δίνει βάρος σε κάθε αναφορά ανάλογα με ένα στοχαστικό μέτρο του συνολικού χρόνου, που οι επιστήμονες χρησιμοποιούν για την ανάγνωση του άρθρου.

Οι βαθμολογίες *Eigenfactor* κλιμακώνονται έτσι ώστε το άθροισμα των *Eigenfactor* βαθμολογιών όλων των περιοδικών, που απαριθμούνται στο Journal Citation Reports της Thomson (JCR), να είναι 100. Το 2006, το περιοδικό Nature είχε την υψηλότερη βαθμολογία *Eigenfactor*, με βαθμολογία 1,992.

3.1.5. Βαθμολογία Επίδρασης του Άρθρου (Article Influence Score)

Η βαθμολογία επίδρασης (Article Influence Score) προέκυψε και αυτή από το ερευνητικό έργο που διεξήχθη το 2008, από το εργαστήριό του καθηγητή Carl Bergstrom (2008), του Πανεπιστήμιου της Ουάσιγκτον.

Η βαθμολογία επίδρασης συνυπολογίζει τις μετρήσεις της σχετικής επιρροής του περιοδικού με βάση το κάθε άρθρο.

Η βαθμολογία επίδοσης του άρθρου υπολογίζει τον μέσο όρο επιρροής ανά άρθρο, των δημοσιευμάτων του περιοδικού. Συνεπώς, μπορεί να συγκριθεί με το ευρέως χρησιμοποιούμενο Impact Factor της Thomson Scientific's.

Οι βαθμολογίες επίδοσης του Article Influence έχουν κανονικοποιηθεί, ώστε το μέσο άρθρο σε όλη τη βάση δεδομένων Thomson Journal Citation Reports (JCR) να έχει επιρροή άρθρου 1.00.

Το 2006, το κορυφαίο περιοδικό μέσω βαθμολογίας επίδοσης Article Influence είναι το *Annual Reviews of Immunology*, με επιρροή άρθρων 27.454. Αυτό σημαίνει ότι το άρθρο μέσου όρου σε αυτό το περιοδικό έχει 27 φορές την επιρροή του μέσου περιοδικού στο JCR.

3.1.6. Ο Αναλυτής Περιοδικού (Journal Analyzer)

Ο Αναλυτής Περιοδικού παρέχει μια εικόνα της απόδοσης περιοδικού, χρησιμοποιώντας δεδομένα αναφορών της βάσης Scopus και περιλαμβάνει δύο μετρήσεις περιοδικού, τον SJR και SNIP (Δες 4.4.5.2).

Χρησιμοποιώντας αναφορές από περίπου 19.000 τίτλους περιοδικών, από 5.000 και πλέον διεθνείς εκδότες, ο Αναλυτής Περιοδικού ξεκινά την καταμέτρηση από το έτος 1996.



Εικόνα 3 2 Παράδειγμα σύγκρισης αναφορών μεταξύ περιοδικών στην κατηγορία arts and humanities

3.1.7. Δείκτες Elsevier

Την ίδια περίοδο με τον ISI, ο εκδοτικός οίκος Elsevier δημιούργησε και έδωσε σε χρήση τα Journal Metrics τα οποία βασίζονται στα στοιχεία της υπηρεσίας Scopus. Οι μετρικές του Journal Metrics δίδουν μια εναλλακτική εικόνα της απήχησης ενός ερευνητικού περιοδικού. Οι δείκτες είναι διαθέσιμοι ελεύθερα από την ηλεκτρονική διεύθυνση www.journalmetrics.com, αλλά και διαμέσου της βάσης Scopus, η οποία αποτελεί συνδρομή του Συνδέσμου Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (Σ.Ε.Α.Β.).

Οι μετρικές του Journal Metrics είναι οι κάτωθι:

SCImago Journal Rank (SJR)

Αναπτύχθηκε από τον καθηγητή Félix de Moya, (Guerrero-Bote&Moya-Anegon, 2012), ερευνητή καθηγητή στο Ανώτατο Συμβούλιο Επιστημονικών Ερευνών. Η SJR είναι μία μετρική κύρους που βασίζεται στην ιδέα ότι «όλες οι αναφορές δεν είναι ίσες». Με την SJR, το θεματικό πεδίο, η ποιότητα και τη φήμη του

περιοδικού έχουν άμεση επίπτωση στην αξία της αναφοράς. Πρόκειται για μια παραλλαγή του μέτρου της κεντρικότητας ιδιοδιανύσματος (eigenfactor centrality), που χρησιμοποιείται στη θεωρία κοινωνικών δικτύων. Επιπλέον:

- Σταθμίζεται με την επιρροή του περιοδικού και, με αυτό τον τρόπο, κανονικοποιείται η σχέση μεταξύ των περιοδικών.
- Μηδενίζει τη χειραγωγή με την αύξηση της κατάταξης SJR δημοσιεύοντας σε επιστημονικά περιοδικά πιο έγκριτα.
- Ομαλοποιεί τις διαφορές μεταξύ θεματικών πεδίων όσον αφορά την συμπεριφορά των αναφορών.
- Περισσότερες πληροφορίες στην ηλεκτρονική διεύθυνση
- <http://www.journalmetrics.com/sjr.php>

Source Normalized Impact per Paper (SNIP)

Δημιουργήθηκε από τον καθηγητή Henk Moed, στο Κέντρο Επιστήμης και Τεχνολογίας Σπουδών (CWTS), του Πανεπιστημίου του Leiden της Ολλανδίας. Ο SNIP μετρά την απήχηση μιας αναφοράς εντός του πλαισίου αναφοράς της, σταθμίζοντας δηλαδή αναφορές με βάση το συνολικό αριθμό των αναφορών σε ένα θεματικό πεδίο. Έτσι, μια αναφορά μπορεί να έχει μεγαλύτερο αντίκτυπο σε μία θεματική ενότητα όπου οι αναφορές χρησιμοποιούνται λιγότερο και το αντίθετο. Ο δείκτης έχει σαν στόχο να γίνει ένα εργαλείο για την άμεση σύγκριση δημοσιεύσεων σε διαφορετικές γνωστικές περιοχές. Περισσότερες πληροφορίες στην ηλεκτρονική διεύθυνση <http://www.journalmetrics.com/snip.php>.

1 -50 of 17124 << First < Previous Next > Last >>										
Title	SJR	H index	Total Docs. (2008)	Total Docs. (3years)	Total Refs.	Total Cites (3years)	Citable Docs. (3years)	Cites/ Doc. (2years)	Ref./ Doc.	Country
1 Annual Review of Immunology	16,204	179	24	81	4.258	3.608	81	39.46	177.42	
2 Cell	12,802	431	535	1.570	18.726	32.388	1.023	31.86	35.00	
3 Annual Review of Biochemistry	11,602	165	31	91	4.666	3.026	91	30,25	150,52	
4 Ca-A Cancer Journal for Clinicians	11,331	69	34	118	1.848	4.772	62	76,78	54,35	
5 Nature Genetics	10,630	311	327	962	7.048	17.487	638	29,07	21,55	

Εικόνα 3 3 Παράδειγμα από το SJR κατάταξης περιοδικών 1-5 από 17,124 στο έτος 2008.

	Country	Documents	Citable documents	Citations	Self - Citations	Citations per Document	H index
1	 United States	4.307.536	4.093.725	72.315.171	33.964.623	17,29	1.023
2	 United Kingdom	1.242.415	1.149.767	17.140.454	4.282.684	14,78	619
3	 Japan	1.220.415	1.197.781	11.953.831	3.783.244	10,12	480
4	 China	1.217.169	1.210.267	3.696.504	2.038.379	4,61	237
5	 Germany	1.132.583	1.093.560	14.435.211	3.952.207	13,46	542

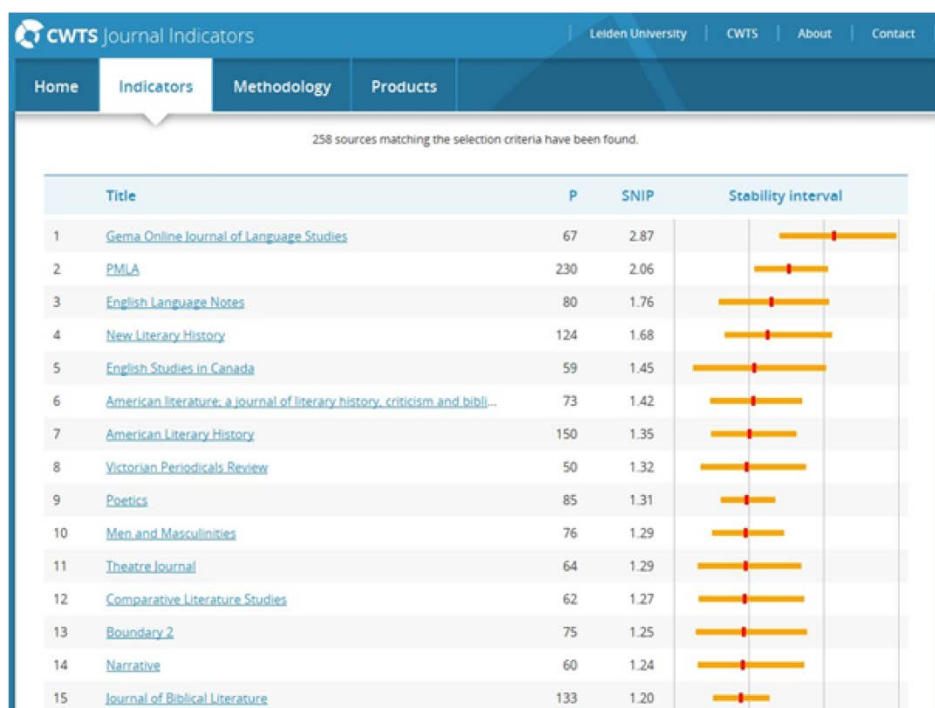
Εικόνα 3 4 Παράδειγμα από το SJR κατάταξης κρατών 1-5 από 233 χώρες από τα έτη 1996-2008.

Ο SNIP

- Λαμβάνει υπόψη τη συχνότητα αναφοράς μιας επιστημονικής περιοχής.
- Εκτιμά το πόσο γρήγορα, ένα άρθρο είναι πιθανό να έχει αντίκτυπο σε μια επιστημονικής περιοχής
- Υπολογίζει το πόσο καλά μια επιστημονική περιοχή καλύπτεται από την υποκείμενη βάση δεδομένων.
- Μετρά όλες τις πιθανές εκδοχές για ενδεχόμενη εκδοτική χειραγώγηση.

Ο SNIP αναθεωρήθηκε το 2012, ώστε να συμπεριλάβει και να εκσυγχρονίσει τους παρακάτω δείκτες.

- P: Ο αριθμός των δημοσιεύσεων από ένα περιοδικό τα τελευταία τρία (3) χρόνια.
- RIP: Η πρώτη επίδραση ανά δημοσίευση υπολογίζεται ως ο αριθμός των αναφορών, που δίδονται στην παρούσα περίοδο με δημοσιεύματα των τριών τελευταίων ετών, διαιρούμενα με το συνολικό αριθμό δημοσιεύσεων τα τελευταία τρία χρόνια. Ο RIP είναι παρόμοιος με το γνωστό Δείκτη επιρροής περιοδικού. Όπως και ο Δείκτης Επιρροής περιοδικού, Ο RIP δεν διορθώνει τις διαφορές στις πρακτικές αναφορών, μεταξύ των επιστημονικών πεδίων.
- SNIP: Η κανονικοποιημένη επιρροή της πηγής ανά δημοσίευση υπολογίζεται ως ο αριθμός των αναφορών, που δίδονται στην παρούσα περίοδο με δημοσιεύματα των τριών τελευταίων ετών, διαιρούμενα με το συνολικό αριθμό δημοσιεύσεων τα τελευταία τρία χρόνια. Η διαφορά με το RIP είναι ότι στην περίπτωση του SNIP οι αναφορές κανονικοποιούνται, ώστε να διορθωθούν οι διαφορές στις πρακτικές αναφορών μεταξύ των επιστημονικών πεδίων. Έτσι λοιπόν όσο μεγαλύτερη είναι η λίστα αναφορών του επικαλούμενου άρθρου, τόσο μικρότερη είναι η τιμή της αναφοράς που προέρχεται από την, εν λόγω, δημοσίευση.
- % Αυτοαναφορά: Είναι το ποσοστό των αυτοαναφορών μιας πηγής, η οποία υπολογίζεται ως το ποσοστό του συνόλου των αναφορών που δίνονται στον παρόντα χρόνο σε δημοσιεύματα, κατά τα τελευταία τρία χρόνια, που προέρχονται από την ίδια την πηγή.



Εικόνα 3 5 Παράδειγμα από τους CWTS Δείκτες περιοδικών του 2013, για περιοδικά καταταγμένα κατά SNIP στην Λογοτεχνία.

Year:

Field:

Journal	Nr papers (1996 - 1998)	% Reviews	SNIP	Raw impact per paper (RIP)	Relative database citation potential	Database coverage (%)	Database citation potential	% journal self citation
Review of Educational Research	56	28,57	5,135	3,48	0,7	43,4	5,5	5,6
Journal of Engineering Education	68	0	4,644	1,72	0,4	38,9	3	16,2
Review of Research in Education	19	15,79	4,565	2,47	0,5	28,9	4,4	6,4
Harvard Educational Review	54	14,81	4,126	1,72	0,4	26,3	3,4	3,2

Εικόνα 3 6 Παράδειγμα από τους CWTS Δείκτες περιοδικών του 2009 για περιοδικά καταταγμένα κατά SNIP στην εκπαίδευση.

3.1.8. Μετρήσεις Περιοδικών μέσω της υπηρεσίας Google Scholar

- **h-index**
- Η Google Scholar έχει υιοθετήσει τη μέθοδο επιρροής h-index για τις δημοσιεύσεις και τη διακύμανση h5 για πέντε πλήρη ημερολογιακά έτη.
- Ο Δείκτης h (Βλ. 3.3.1.) μιας δημοσίευσης είναι ο μεγαλύτερος αριθμός h, ώστε να υπάρχουν τουλάχιστον h άρθρα σε αυτή τη δημοσίευση που αναφέρθηκαν τουλάχιστον h φορές το κάθε ένα. Για παράδειγμα, μια έκδοση με πέντε άρθρα που αναφέρονται αντίστοιχα, 17, 9, 6, 3, και

2, έχει ως *h*-δείκτη το 3. Αφού υπάρχουν τρία άρθρα με τουλάχιστον τρεις αναφορές και όχι τέσσερα άρθρα με τουλάχιστον τέσσερις αναφορές.

- Ο *h-core* μιας δημοσίευσης είναι ένα σύνολο των πλέον αναφερόμενων *h* άρθρων από τη δημοσίευση. Για παράδειγμα, η παραπάνω δημοσίευση έχει ως *h-core* αυτόν που βασίζεται στον *h* με αριθμό τρία, εκείνα που αναφέρονται με 17, 9 και 6 αναφορές.
- Η *h-median* (διάμεση τιμή) έκδοσης είναι η διάμεση τιμή των αναφορών στον *h-core*. Για παράδειγμα, ο *h-median* της παραπάνω δημοσίευσης είναι 9. Ο *h-median* είναι ένα μέτρο της κατανομής των αναφορών στα άρθρα του *h-core*.
- Τέλος, οι *H5-index*, *h5-core* και *h5-median* έκδοσης είναι, αντίστοιχα, οι *h-index*, *h-core* και *h-median* για εκείνα μόνο τα άρθρα που δημοσιεύθηκαν στην τελευταία πενταετία.

Κορυφαία Άρθρα από τις Υψηλότερες 100 Δημοσιεύσεις

Με τη δημοσίευση του δείκτη *h5* για τις υψηλότερες 100 δημοσιεύσεις στο Google Scholar Metrics, περιλαμβάνονται επίσης τα κορυφαία άρθρα που αναφέρονται. Κατά την προβολή των υψηλότερων 100 δημοσιεύσεων, που προκύπτει αν πιέσουμε την ένδειξη 'σκορ δείκτη *h5*', θα δούμε τα πιο ιδιαίτερα αναφερόμενα άρθρα από αυτό το περιοδικό (*h5-core*). Στη συνέχεια, αν πιέσουμε την ένδειξη 'πλήθος αναφορών για οποιοδήποτε άρθρο του 'H5-core', θα δούμε ποιός έκανε αναφορά σε αυτό.

3.2. Παραδείγματα χρήσης συνδυασμού δεικτών επιρροής περιοδικών

3.2.1. Προβλήματα και Λύσεις

Έστω ότι έχουμε να συγκρίνουμε δύο διαφορετικά σύνολα άρθρων με όρους επιρροής αναφορών. Το πρώτο σύνολο αποτελείται από τα άρθρα A_1, A_2, A_3 στα περιοδικά J_1, J_2, J_3 , του τομέα της Γεωργίας στο πεδίο της Οικολογίας, της Αγρονομίας και της Κρυσταλλογραφίας αντιστοίχως. Το δεύτερο σύνολο από τα άρθρα a_1, a_2, a_3, a_4 σε περιοδικά j_1, j_2, j_3, j_4 , του τομέα της Βιολογίας. Το A_1 που ο τύπος του άρθρου είναι επισκόπηση (review) έχει $C_1=20$ αναφορές, το A_2 έχει $C_2=4$ αναφορές, το A_3 έχει $C_3=7$ αναφορές, με σύνολο αναφορών πρώτου συνόλου $C_{1sum}=31$. Ομοίως το a_1 έχει 2 αναφορές, το a_2 έχει 10 αναφορές, το a_3 έχει 3 αναφορές, το a_4 έχει 4 αναφορές, με σύνολο αναφορών δεύτερου συνόλου $C_{2sum}=19$. Οι συνολικές βαθμολογίες των δύο αναφορών δεν είναι άμεσα συγκρίσιμες, διότι εξαρτώνται από εξωτερικούς παράγοντες, μη σχετιζόμενους, όπως οι εξής: αριθμός άρθρων, ποσοστό αυτοαναφορών, ηλικία άρθρου, τύπος άρθρου, τομέας ή περιοχή έρευνας, αριθμός άρθρων. Ας δούμε πώς ένας βιβλιομέτρης θα προσπαθήσει να τα διορθώσει αυτά προκειμένου να επιτραπεί η δίκαιη σύγκριση των δύο συνόλων. Στις δύο ομάδες των άρθρων, αν για παράδειγμα τα δύο περιοδικά είναι διαφορετικού μεγέθους με περισσότερα άρθρα το πρώτο από το άλλο, το πρώτο έχει δυνητικά ένα πλεονέκτημα έναντι του άλλου (περισσότερες πιθανότητες να πάρει αναφορές). Για να διορθωθεί αυτή η αδικία θα πρέπει να υπολογισθούν οι μέσοι όροι των τιμών, δηλαδή, αριθμός αναφορών / αριθμός άρθρων.

Ο I.F. των περιοδικών, για παράδειγμα, μας επιτρέπει να συγκρίνουμε περιοδικά που δημοσιεύουν διαφορετικούς αριθμούς άρθρων και δίνεται επίσης χωρίς αυτοαναφορές στο JCR (Journal Citation Report) της Thomson Scientific. Σε ορισμένο βαθμό, η επικαλούμενη αυτοαναφορά είναι φυσιολογική, αλλά πάνω από αυτόν είναι παθολογική. Για τη διόρθωση αυτού του παράγοντα θα πρέπει να εξαλειφθούν οι αυτοαναφορές, αλλά δεν είναι τόσο εύκολο (π.χ.: στη Scopus ο δείκτης *h* είναι διαθέσιμος χωρίς αυτοαναφορές, ενώ στην πλατφόρμα WoS (Web of Science) δεν είναι. Για ένα γενικό σύνολο άρθρων μπορούμε να διορθώσουμε τόσο το μέγεθος, όσο και το ποσοστό αυτοαναφορών μέσα από τον δείκτη των αναφορών ανά άρθρο. CPP -Citations Per Paper = συνολικός αριθμός των αναφορών που έλαβαν τα άρθρα διαιρούμενα με τον αριθμό των άρθρων (p), δεν υπολογίζονται οι αυτοαναφορές,

π.χ. για το πρώτο σύνολο:

$$CPP = C_1 + C_2 + C_3 = \frac{(20 + 4 + 7)}{3} \approx 10,3$$

Περισσότερο είναι δύσκολο να διορθωθούν παράγοντες όπως: 1) Η ηλικία. Στη θεωρία, ένα παλιό άρθρο είχε περισσότερο χρόνο για να πάρει αναφορές από ένα πρόσφατο. Ορισμένα είδη άρθρων έχουν την τάση να προσελκύουν περισσότερες αναφορές, ανεξάρτητα από την ποιότητά τους (π.χ. σχόλια στη Βιοϊατρική). 2) Ο τύπος του άρθρου. Διαφορετικά ερευνητικά πεδία έχουν διαφορετικές συμπεριφορές αναφορών (π.χ. άρθρα στη Βιοχημεία συχνά περιέχουν πάνω από 50 αναφορές κατά μέσο όρο, ενώ στα Μαθηματικά 10, ως εκ τούτου, τα άρθρα Βιοχημείας έχουν περισσότερες πιθανότητες να αναφερθούν).

Για να εξαλειφθούν οι παραπάνω περιπτώσεις διορθώνουμε (ομαλοποιούμε- κανονικοποιούμε) τον πραγματικό δείκτη απήχησης των 3 άρθρων στο πρώτο σετ, διαιρεμένο με το μέσο δείκτη απήχησης όλων των άρθρων του ίδιου τύπου που δημοσιεύθηκαν την ίδια χρονιά, στα ίδια περιοδικά / ίδιων θεματικών περιοχών. Με αυτό τον τρόπο ο απόλυτος δείκτης απήχησης από τα τρία άρθρα μετατρέπεται σε ένα σχετικό δείκτη απήχησης (σε σχέση με το σύνολο παρόμοιων άρθρων που δημοσιεύθηκαν σε όλο τον κόσμο). Τέλος, συγκρίνουμε τους κανονικοποιημένους δείκτες απήχησης για τις δύο ομάδες.

Οι μέσες τιμές στο παρονομαστή, που ονομάζονται πρότυπα αναφοράς ή αναμενόμενα ποσοστά αναφορών, αντιπροσωπεύουν τον αριθμό των αναφορών που τα άρθρα (των οποίων η βαθμολογία υπολογίζεται στον αριθμητή) θα πρέπει να λάβουν, προκειμένου να ασκήσουν μια μέση επίδραση αναφοράς στον τομέα τους. Ας δούμε μερικά **παραδείγματα** των σχετικών δεικτών για το πρώτο σύνολο των 3 άρθρων. Βέβαια, σε πραγματικές καταστάσεις δεν εργαζόμαστε με ένα τόσο χαμηλό αριθμό άρθρων. Στο ακόλουθο παράδειγμα οι υπολογισμοί των προτύπων αναφοράς για περιοδικά και οι θεματικές ενότητες είναι πραγματικές. Οι υπολογισμοί έγιναν στην πλατφόρμα WoS, τον Μάρτιο του 2013, οι αυτοαναφορές δεν αποκλείονται και οι αναφορές υπολογίζονται σε τρία χρόνια μετά τη δημοσίευση. Έστω ότι $\bar{J}_1 = 0$ μέσος όρος αναφορών όλων των επισκοπήσεων (reviews) που δημοσιεύθηκαν στο περιοδικό Applied Ecology το 2003 και ισούται με 38, $\bar{S}_1 = 0$ μέσος όρος αναφορών όλων των επισκοπήσεων (reviews) με θεματική κατηγορία Οικολογία (στο WoS) το 2003 ≈ 21.9 , C_1 =αναφορές που ελήφθησαν από το πρώτο άρθρο = 20. Αντίστοιχα ο μέσος όρος αναφορών όλων των άρθρων που δημοσιεύθηκαν στο περιοδικό Agronomy το 2005 ≈ 4.7 , $\bar{J}_2 = 0$ μέσος όρος αναφορών όλων των άρθρων με θεματική κατηγορία Αγρονομία (στο WoS) το 2005 ≈ 3.9 , C_2 =αναφορές που ελήφθησαν από το δεύτερο άρθρο = 4. Και $\bar{J}_3$, ο μέσος όρος αναφορών όλων των άρθρων που δημοσιεύθηκαν στο περιοδικό Liquid Crystals το 2006 ≈ 2.9 , $\bar{S}_3 = 0$ μέσος όρος αναφορών όλων των άρθρων με θεματική κατηγορία Κρυσταλλογραφία (στο WoS) το 2005 ≈ 3.1 , C_3 =αναφορές που ελήφθησαν από το τρίτο άρθρο = 7.

Αναφορές ανά άρθρο (Citation Per Paper -CPP) / Μέσος όρος αναφορών περιοδικού (Mean Journal

$$(CitationScore - JCSm) = \frac{CPP}{JCSm} = \frac{\frac{C_1 + C_2 + C_3}{3}}{\frac{\bar{J}_1 + \bar{J}_2 + \bar{J}_3}{3}} = \frac{\frac{20 + 4 + 7}{3}}{\frac{38 + 4.7 + 2.9}{3}} = \frac{31}{45.6} \approx 0,68$$

Το αποτέλεσμα εξαρτάται από το κύρος των περιοδικών που επέλεξαν οι αρθρογράφοι (στρατηγική δημοσίευσης).

Αναφορές ανά άρθρο (Citations Per Paper-CPP)/Μέσος όρος αναφορών πεδίου (Mean Field Citation Score-FCSm)=

$$\frac{CPP}{FCSm} = \frac{\frac{C_1 + C_2 + C_3}{3}}{\frac{\bar{S}_1 + \bar{S}_2 + \bar{S}_3}{3}} = \frac{\frac{20 + 4 + 7}{3}}{\frac{21.9 + 3.9 + 3.1}{3}} \approx 1,08$$

Οι θεματικές κατηγορίες στο WoS και την Scopus εξακολουθούν να ορίζονται σε επίπεδο περιοδικού, αλλά τώρα όλα τα περιοδικά στο θέμα της κατηγορίας είναι σχετικά.

Μέσο κανονικοποιημένο σκορ αναφορών (Mean Normalized CitationScore -MNCS): ο μέσος πραγματικός δείκτης απήχησης από τα τρία άρθρα κανονικοποιείται έναντι κάθε αναμενόμενης τιμής πριν από την πρόσθεση.

$$MNCS = \frac{\frac{C1+C2+C3}{\bar{S}1+\bar{S}2+\bar{S}3}}{3} = \frac{\frac{20+4+7}{21,9+3,9+3,1}}{3} = \frac{4,19}{3} = 1,4$$

Αυτός είναι ο crown indicator.

Οι Βιβλιομετρικές κατανομές, όπως έχουμε δείξει, είναι ασύμμετρες, άρα οι μέσες τιμές είναι ελάχιστα σημαντικές. Το γεγονός οι μέσες τιμές των βαθμολογιών να είναι πανταχού παρούσες στις κατασκευές των δεικτών οφείλεται στο ότι, όταν μετράμε αναφορές ο αριθμός είναι μεγάλος κι έτσι η επίδραση της ασυμμετρίας μπορεί να αγνοηθεί. Βέβαια, μπορεί κάποιος να ισχυριστεί ότι δεν χρειάζονται καθόλου οι μέσοι όροι για να κατασκευάσουμε τους σχετικούς δείκτες, απλώς χρησιμοποιούμε εκατοστημόρια κατάταξης (percentile ranks). Για τον σκοπό αυτό:



Εικόνα 3.7 Κατανομή αναφορών για τα άρθρα της κατηγορίας επισκοπήσεις (reviews) στην θεματική περιοχή της οικολογίας το έτος 2003

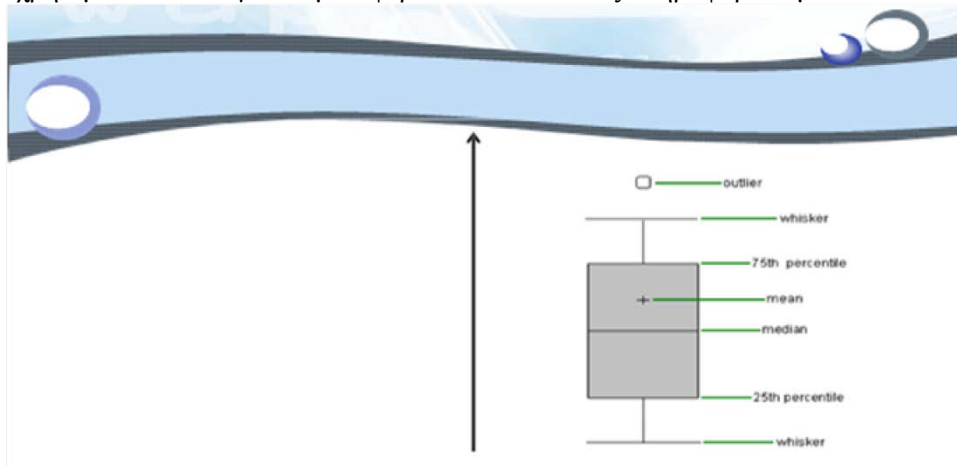
Παίρνουμε το σκορ του πρώτου άρθρου του παραδείγματος και ελέγχουμε την κατάταξη του στην κατανομή αναφορών όλων των άρθρων του ίδιου τύπου (π.χ. επισκοπήσεις) που δημοσιεύθηκαν τον ίδιο χρόνο και ανήκουν στην ίδια θεματική κατηγορία (π.χ.: ανήκει στο 10% των κορυφαίων αναφερόμενων άρθρων στον τομέα της Οικολογίας). Κάνουμε το ίδιο για όλα τα άλλα άρθρα στο πρώτο σύνολο, με σκοπό να προσδιοριστεί η (σχετική) θέση του καθενός στον δείκτη απήχησης των άρθρων στην κατανομή των παρόμοιων άρθρων, όπως προκύπτει από το ευρετήριο του δείκτη αναφορών (citation index). Κάνουμε το ίδιο για κάθε ένα από τα άλλα άρθρα στο δεύτερο σύνολο. Στο τέλος, συγκρίνουμε τα δύο σύνολα όσον αφορά τα ποσοστά των άρθρων που ανήκουν στο 5%, 10%, κ.λπ. από τα κορυφαία αναφερόμενα άρθρα στην κατηγορία τους.

Για παράδειγμα, ας υποθέσουμε ότι αυτή είναι η κατανομή των αναφορών που ελήφθησαν από όλες τις επισκοπήσεις (reviews) που δημοσιεύθηκαν το 2003, στον τομέα της Οικολογίας: Στη συνέχεια η επισκόπησή μας με 20 αναφορές στο περιοδικό της Applied Ecology είναι στο 60ο εκατοστημόριο, δηλαδή το 60% του συνόλου των επισκοπήσεων στην ίδια θεματική περιοχή έχουν λιγότερες αναφορές. Δυστυχώς, η πραγματική κατανομή αναφοράς είναι πολύ πιο πολύπλοκη (υπάρχουν συνήθως πολλά μηδενικά και πολλές επαναλαμβανόμενες τιμές αναφορών).

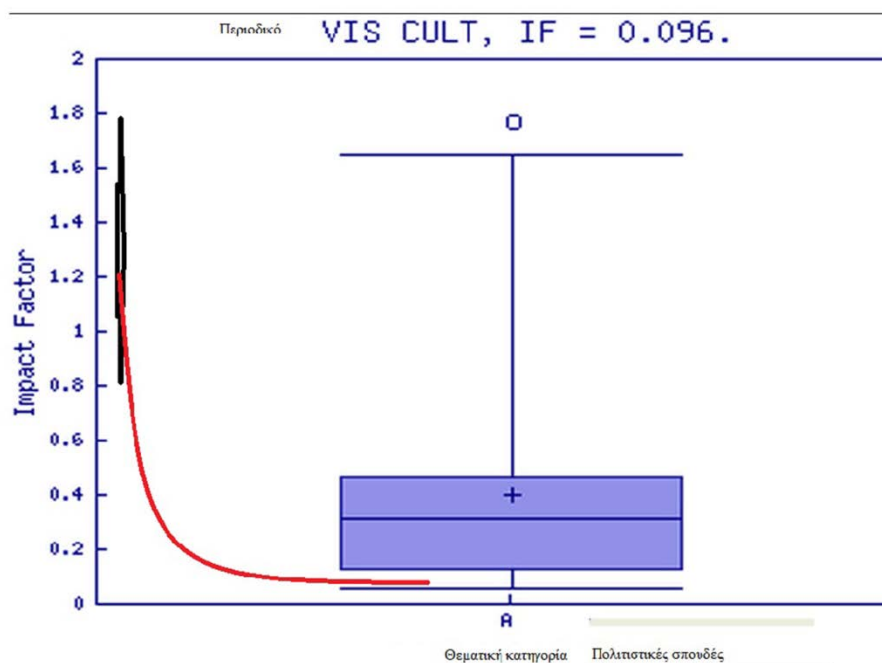
Μέχρι στιγμής έχουμε δει μόνο μία δυνατή προσέγγιση για την εξομάλυνση αναφορών. Ας δούμε και δύο άλλες νεότερες λύσεις. 1) Την κανονικοποιημένη πλευρά των αναφορών που παίρνει υπόψη της

αναφορές που λαμβάνονται από ομάδες άρθρων παρόμοιες με αυτήν του, υπό αξιολόγηση, άρθρου. Αν ορίσουμε τη θεματική περιοχή του B, όπως την περιοχή του συνόλου των τριών άρθρων που το αναφέρουν, τότε ένα κατάλληλο πρότυπο αναφοράς για τον αριθμό αναφορών του B είναι ο μέσος αριθμός των βιβλιογραφικών αναφορών σε A1, A2, A3 (δυννητική αναφορά της θεματικής περιοχή του B).

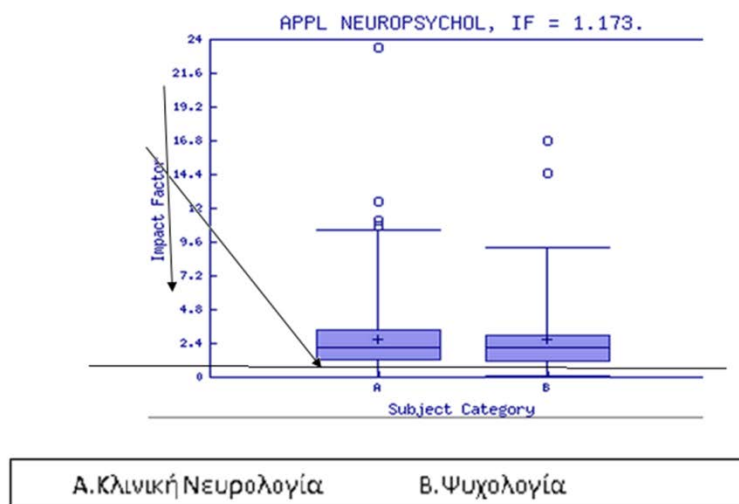
Ο IF ενός περιοδικού έχει νόημα μόνο ως σχετική ένδειξη, δηλαδή συγκρίνοντας την τιμή του, με την κατανομή της βαθμολογίας των IF's όλων των άλλων περιοδικών στην ίδια θεματική κατηγορία. Τα Boxplots του JCR χρησιμοποιούνται για να μεταφέρουν αυτό το είδος πληροφοριών με έναν οπτικό τρόπο.



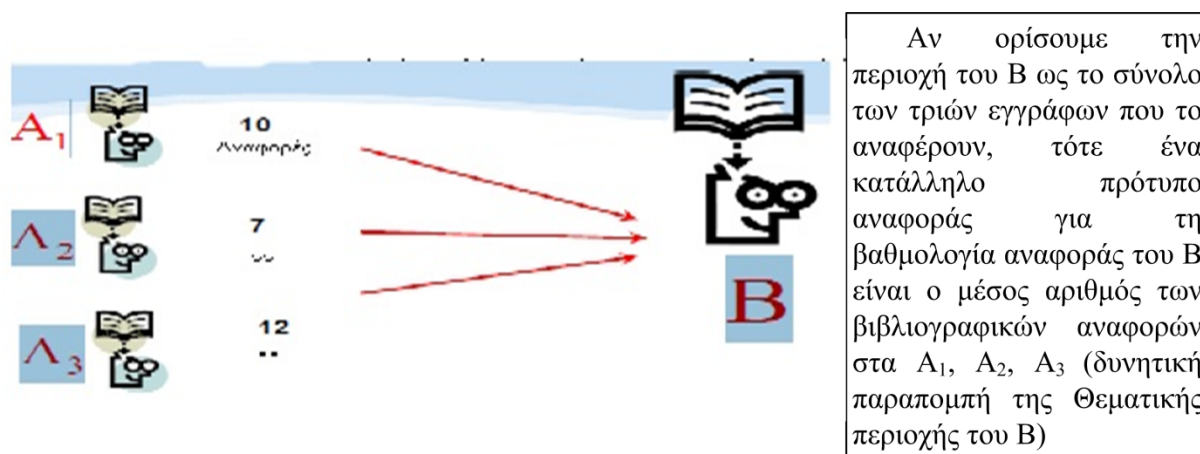
Εικόνα 3 8 Τα εκατοστημόρια και ο IF



Εικόνα 3 9 Η θέση του IF του περιοδικού *VisualCulture* του 2011 στην κατανομή των if's των περιοδικών της κατηγορίας Πολιτιστικών σπουδών

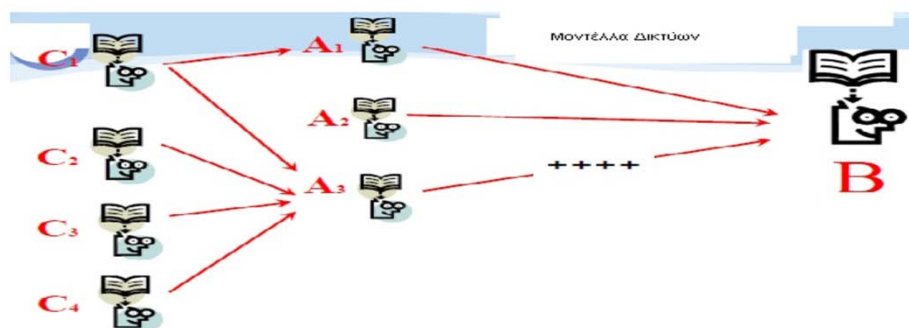


Εικόνα 3 10 Η θέση του IF του περιοδικού *AppliedNeurosciology* του 2011 είναι σχεδόν 12 φορές υψηλότερη από το 0,096 του IF του περιοδικού *VisualCulture*. Παρόλα αυτά κατατάσσεται πολύ χαμηλά μεταξύ των περιοδικών της κατηγορίας του.



Εικόνα 3 11 κανονικοποιημένη πλευρά των αναφορών

Τα μοντέλα Δικτύων



Εικόνα 3 12 Μοντέλο δικτύου αναφορών

Το άρθρο B λαμβάνει τρεις αναφορές, αλλά αυτές προέρχονται από το άρθρο A_3 που έχει βαρύτητα μεγαλύτερη από τα υπόλοιπα άρθρα, επειδή το A_3 είναι ένας εξέχων κόμβος στο δίκτυο (έχει τέσσερις αναφορές, σε αντίθεση με το A_1 και το A_2). Αυτό το μοντέλο είναι παρόμοιο με τη μοντελοποίηση σταθμισμένης αναφοράς Page Rank της Google. Οι δείκτες περιοδικών που ομαλοποιούν τις διαφορές στις αναφορές εξαιτίας των ιδιαιτεροτήτων των κλάδων των επιστημών βασίζονται στα δύο μοντέλα: α) Στην

κανονικοποιημένη πλευρά των αναφορών ή κανονικοποιημένη πηγή επιρροής ανά άρθρο (Source Normalized Impact per Paper -SNIP) βασισμένη σε δεδομένα της βάσης Scopus). β) Στο μοντέλο δικτύου αναφορών βασισμένο στο SCImago Journal Rank (SJR) σε δεδομένα της βάσης Scopus και στο EigenFactor βασισμένο σε δεδομένα της πλατφόρμας WoS.

Ο SNIP του περιοδικού J είναι ο λόγος

$$[(RIP - \text{Row ImPact}) \text{ Μέση επίδραση ανά άρθρο του περιοδικού J}] / [(CPOT) \text{ δυναμικό αναφοράς της θεματικής περιοχής του J}]$$

Πιο πρακτικά, ο λόγος με αριθμητή τον μέσο αριθμό των αναφορών που ελήφθησαν, σε ένα δεδομένο έτος, από όλα τα άρθρα που δημοσιεύονται στο περιοδικό J κατά τα τρία προηγούμενα έτη (όμοιος με τον IF) και παρονομαστής τον μέσο αριθμό αναφορών που εμφανίζονται στις βιβλιογραφίες, οι οποίες αναφέρουν το περιοδικό J. Υπολογίζονται μόνο αναφορές σε άρθρα που ευρετηριάζονται στη βάση δεδομένων Scopus (ο αριθμός τους ονομάζεται Citation Potential-CPOT), των οποίων το έτος δημοσίευσης είναι εντός του εύρους των τριών ετών, που χρησιμοποιείται για τον υπολογισμό του RIP. Στον τελικό υπολογισμό του SNIP το CPOT κάθε μεμονωμένου περιοδικού ομαλοποιείται πλήρως διαιρούμενο με την μέση τιμή του CPOT, όπως αυτό έχει υπολογιστεί σε όλα τα περιοδικά της βάσης δεδομένων. Η ομαλοποιημένη τιμή ονομάζεται σχετική βάση δεδομένων CPOT.

Υπολογισμός του Citation Potential-CPOT του 2011 για το περιοδικό J με IF 2.0

α/α. Περιοδικού που αναφέρει	Σύνολο αναφορών	Βάρη	I.F's των περιοδικών
1	100	100/200=0,5	1,000
2	60	60/200=0,3	2,500
3	30	30/200=0,15	0,800
4	10	10/200=0,05	1,400
Σύνολο	200	1	

Πίνακας 3.2 Αναφορές περιοδικών της ίδιας θεματικής περιοχής του J που αναφέρουν το 2011 τα άρθρα που έχουν δημοσιευτεί στο περιοδικό J το 2010, 2009.

$$CPOT = 0.5 * 1.0 + 0.3 * 2.5 + 0.15 * 0.8 + 0.05 * 1.4 = 1.44$$

Τότε 2.0/1.44=1.39.

Δηλαδή ο JIF 2.0 είναι 39% μεγαλύτερος από το CPOT της θεματικής περιοχής του περιοδικού. Οι επιστήμες ζωής έχουν υψηλό IF και υψηλό CPOT. Οι τέχνες έχουν χαμηλό IF και χαμηλό CPOT.

Ο SJR κάνει χρήση ενός σύνθετου τύπου με βάση ένα πιθανοθεωρητικό μοντέλο που δίνει βάρος σε αναφορές σύμφωνα με τη θέση του άρθρου στο παγκόσμιο δίκτυο βιβλιογραφικών συνδέσεων (κεντρικότητα ιδιοδιανύσματος). Αν τα δύο περιοδικά έχουν την ίδια βαθμολογία αναφορών, ο τύπος εκχωρεί μεγαλύτερη αξία σε εκείνο του οποίου οι αναφορές προέρχονται από περιοδικά με υψηλότερες τιμές SJR.

3.3. Οι Βιβλιομετρικοί δείκτες: Ένας πρακτικός οδηγός κατασκευής και χρήσης για ερευνητές

Πολλοί ερευνητές της βιβλιομετρίας τονίζουν τη σημασία του να μην θεωρούμε τα αποτελέσματα της κάθε βιβλιομετρικής ανάλυσης ως «αλήθειες». Ο όρος βιβλιομετρικοί δείκτες συχνά χρησιμοποιείται για να σημειωθεί το γεγονός ότι τα αποτελέσματα περιγράφουν μια αρκετά πολύπλοκη πραγματικότητα, που πρέπει να μετρηθεί μόνο με στατιστικές ή αριθμούς. Οι Βιβλιομετρικές μέθοδοι περιέχουν τόσο πολλές απλοποιήσεις, που παρέχουν μόνο μια πολύ περιορισμένη εικόνα της έρευνας, που προσπαθούν να περιγράψουν. Είναι σημαντικό να δούμε τους βιβλιομετρικούς δείκτες ως ένα από τα πολλά εργαλεία που πρέπει να χρησιμοποιούνται από τους αρμόδιους, με ειδικές γνώσεις, βαθμολογητές, σχετικούς με τις ερευνητικές περιοχές που περιλαμβάνονται στην ανάλυση. Αυτό παραδείγματος χάριν είναι εμφανές, όταν οι δημοσιεύσεις που περιλαμβάνονται σε μια αξιολόγηση περιέχουν πολύ ασυνήθιστα ή νέα ερευνητικά δεδομένα. Αυτές οι εργασίες δεν θα έχουν ακόμη γίνει αντικείμενο αναφοράς, πράγμα που σημαίνει ότι οποιαδήποτε αξιολόγηση βασισμένη αποκλειστικά στους βιβλιομετρικούς δείκτες δεν θα ανακαλύψει την ενδεχόμενη δυναμική των εμπλεκόμενων ερευνητικών ομάδων.

Κανείς βιβλιομετρικός δείκτης δεν έχει την δύναμη να χρησιμοποιείται μόνος αποκομμένος από τους άλλους. Αρκετοί δείκτες πρέπει πάντα να συνδυαστούν για να επιτευχθεί μια πιο ολοκληρωμένη εικόνα της επιστημονικής παραγωγής μιας μονάδας (Visser, 2003; Nederhof & vanRaaij, 2003). Ο δείκτης Crown πρέπει, για παράδειγμα, πάντα να συνοδεύεται από τον λεγόμενο top δείκτη, που δείχνει αν η μέση τιμή των αναφορών σε δημοσιεύσεις της μονάδας οφείλεται σε μερικά πολύ ιδιαίτερα αναφερόμενα άρθρα ή η πλειοψηφία των δημοσιευμάτων αναφέρονται λίγο άνω του μέσου όρου και από έναν ποσοτικό δείκτη για να δείξει πόσες δημοσιεύσεις περιλαμβάνονται στην ανάλυση.

3.3.1. h-index

J.E. Hirsch δημιούργησε τον δείκτη h, το 2005, στο άρθρο του με τίτλο "Ένας δείκτης για την ποσοτικοποίηση της επιστημονικής ερευνητικής παραγωγής ενός ατόμου." Σε αυτό το θεμελιώδες άρθρο, ο Hirsch προσπάθησε να απαντήσει στο ερώτημα: "Πώς μπορεί κάποιος να ποσοτικοποιήσει τις σωρευτικές επιπτώσεις και τη σημασία των αποτελεσμάτων της επιστημονικής έρευνας ενός ατόμου;" Ο δείκτης h είναι ο αριθμός των δημοσιεύσεων (h), που αποδίδονται στην υπό ανάλυση μονάδα, κατά τη διάρκεια του χρονικού διαστήματος που αναλύθηκαν, που έχει τουλάχιστον h αναφορές. Υπολογίζεται από τον οργανισμό Web of Science, βρίσκοντάς τον στην περιοχή απολογισμός αναφορών στην αρχή μιας αναζήτησης. Στη Scopus, ο χρήστης θα βρει τον h-index με τη διεξαγωγή μιας αναζήτησης και επιλέγοντας ως υπερσύνδεσμο το όνομα ενός συγγραφέα στην ενότητα έρευνα. Ο δείκτης έχει το ίδιο μειονέκτημα με τους άλλους βασικούς δείκτες, αφού εξομοιώνει συγκρίνοντας τις ερευνητικές εργασίες διαφόρων τύπων, που δημοσιεύθηκαν σε διαφορετικές χρονικές περιόδους, σε εντελώς διαφορετικά θέματα. Ο δείκτης μεροληπτεί υπέρ των παλαιών ερευνητών με πολλά άρθρα, δεδομένου ότι αυτά είχαν στο περισσότερο χρόνο για να αναφερθούν. Ωστόσο, αυτή η μεροληψία μετριάζεται κάπως από το γεγονός ότι κάθε νέο άρθρο ενός συγγραφέα με υψηλό δείκτη h πρέπει να έχει υψηλό επίπεδο αναφοράς για να μεγαλώσει ο, ήδη, υψηλός h-index. Μια άλλη επίκριση για τον h-index είναι ότι θέτει σε μειονεκτική θέση τους επιστήμονες με σύντομη καριέρα, δεδομένου ότι ο δείκτης h δεν μπορεί να είναι μεγαλύτερος από τον αριθμό των δημοσιευμένων άρθρων, αφού δεν έχει σημασία πόσο σημαντικά και πολύ αναφερόμενα είναι τα άρθρα τους. Ο δείκτης h-index προορίζεται για να διακρίνει πραγματικά αξιόλογους επιστήμονες σε αντίθεση με αυτούς που απλώς δημοσιεύουν πάρα πολλά άρθρα.

- Δεν επηρεάζεται από μεμονωμένα άρθρα που έχουν πάρει πολλές αναφορές.
- Λειτουργεί ορθά μόνο για όταν συγκρίνει επιστημονικές εργασίες που βρίσκονται στο ίδιο επιστημονικό πεδίο. Γιατί οι αναφορές διαφέρουν σε μεγάλο βαθμό μεταξύ των διαφόρων επιστημονικών τομέων.
- Για να συγκρίνουμε του h-δείκτες εξομαλύνοντας τις τιμές τους, τους διαιρούμε με ένα δεύτερο παράγοντα όπως είναι τα χρόνια που έχουν περάσει από την απόκτηση του διδακτορικού.

3.3.2. G -index

Επειδή ο h-Index μένει ανεπηρέαστος από άρθρα με υψηλούς ή χαμηλούς δείκτες βιβλιογραφικών αναφορών, αν χρησιμοποιηθεί οποιοσδήποτε ποσοτικός δείκτης για να μετρήσει την αξία ενός ερευνητή θα πρέπει να υπολογίσει την επιρροή ή αντίκτυπο ή απήχηση) των σημαντικότερων άρθρων του.

Φανταστείτε ότι τρεις συγγραφείς έχουν τις παρακάτω ‘επιδόσεις’ που οδηγούν στον ίδιο h-Index τον αριθμό 3. Δεν φαίνεται πολύ λογικό και δίκαιο αυτό.

	Εργασία 1η	Εργασία 2η	Εργασία 3 ^η	Εργασία 4η
Συγγραφέας Α	3 αναφορές	6 αναφορές	100 αναφορές	2 αναφορές
Συγγραφέας Β	3 αναφορές	6 αναφορές	100 αναφορές	
Συγγραφέας Γ	400 αναφορές	150 αναφορές	3 αναφορές	3 αναφορές

Πίνακας 3.3

Για αυτό το λόγο προτάθηκε από τον Leo Egghe (Egghe, 2006) μία βελτιωμένη έκδοση του h –index, ο G-Index, που λαμβάνει υπόψη του την απόδοση των σημαντικότερων άρθρων κάθε συγγραφέα, διαφοροποιώντας τους σε μια πιο ‘δίκαιη’ κατάταξη.

Πρόσβαση στην Web of Knowledge στον υπερσύνδεσμο <http://www.wokinfo.com> και στη Scopus στον <http://www.info.sciverse.com/scopus>

Υπολογισμός των δεικτών h και G

Αναφορές Άρθρων	Σειρά κατάταξης	Συσσωρευτικές αναφορές	Σειρά κατάταξης στο τετράγωνο
47	1	47	1
42	2	89	4
37	3	126	9
36	4	162	16
21	5	183	25
18	6	201	36
17	7	218	49
16	8	234	64
16	9	250	81
16	10	266	100
15	11	281	121
13	12	294	144
13	13	307	169
13	14	320	196
13	15	333	225
12	16	345	256
12	17	357	289
12	18	369	324
12	19	381	361
11	20	392	400

Πίνακας 3.4 Υπολογισμός των δεικτών h και g ενός ερευνητή

Κατατάσσουμε τα άρθρα σε φθίνουσα σειρά ξεκινώντας με αυτό που έχει τις περισσότερες αναφορές 47 τον αριθμό. Στην δεύτερη γραμμή στην στήλη συσσωρευτικές αναφορές θα έχουμε τον αριθμό 89 που προκύπτει από το άθροισμα των 47 αναφορών του 1^{ου} άρθρου και των 42 του 2^{ου} και στην 4^η στήλη τον αριθμό 4 που προκύπτει από τον αριθμό της κατάταξης του άρθρου 2^ο από άποψη αναφορών υψωμένο στο τετράγωνο. Ο h-index ισούται με 13, γιατί 13 άρθρα έχουν τουλάχιστον 13 αναφορές. Ο g-index είναι 19, αφού είναι η τελευταία διαβάθμιση για την οποία ισχύει: αριθμός συσσωρευτικών αναφορών μεγαλύτερος από σειρά κατάταξης στο τετράγωνο, δηλαδή 381>361 που ισχύει για τη σειρά κατάταξης 19.

3.4. Χρησιμοποίηση Δεδομένων για Παραγωγή Δεικτών

Το ερώτημα είναι: Ποιους επιπλέον βιβλιομετρικούς δείκτες μπορεί να προσφέρει η ανάλυση; Ή πώς να χρησιμοποιούμε τα δεδομένα για την παραγωγή των διαφόρων δεικτών; Η απάντηση σε αυτά τα ερωτήματα δίδεται από την κατασκευή των παρακάτω δεικτών (Σαχίνη κ. ά., 2014):

- **Ο Αριθμός των δημοσιεύσεων (Number of publications)**

Είναι ο πρώτος από τους βασικότερους βιβλιομετρικούς δείκτες που δείχνει τον όγκο παραγωγής των ερευνητικών εργασιών που αποδίδονται σε ένα επιστήμονα ή σε, μια ερευνητική ομάδα, ή έναν φορέα, ή ένα επιστημονικό πεδίο ή μια χώρα, κατά τη διάρκεια μιας καθορισμένης χρονικής περιόδου.

- **Ο Αριθμός των αναφορών (Number of citations)**

Ο επόμενος από τους δύο (2) πιο βασικούς δείκτες που δείχνει την αναγνωρισιμότητα ή την επιρροή των ερευνητικών δημοσιεύσεων που έλαβαν αναφορές, κατά τη διάρκεια μιας καθορισμένης χρονικής περιόδου.

- **Ο αριθμός δημοσιεύσεων και αναφορών ανά ερευνητή**

Αυτός ο αριθμός είναι ένα σχετικό μέτρο δημοσιεύσεων και αναφορών ανά ερευνητή. Δείχνει το αποτέλεσμα της επιστημονικής παραγωγής σε σχέση με τους πόρους που επενδύονται αντισταθμίζοντας, με το μέγεθος του.

- **Μερίδιο (%) δημοσιεύσεων (Share of publications)**

Το επί τοις εκατό μερίδιο των δημοσιεύσεων υπολογίζεται ως το ποσοστό του αριθμού των ερευνητικών εργασιών της χώρας επί του συνολικού αριθμού εργασιών άλλων χωρών που είναι μέλη διεθνών οργανισμών, όπως ο Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης ή μιας επιστημονικής περιοχής επί του συνολικού αριθμού των ερευνητικών δημοσιεύσεων μιας χώρας ή μιας κατηγορίας οργανισμών επί του συνολικού αριθμού δημοσιεύσεων των οργανισμών στο σύνολο των επιστημονικών πεδίων ή ενός οργανισμού επί του συνολικού αριθμού των ερευνητικών δημοσιεύσεων της κατηγορίας στην οποία αυτός ανήκει.

- **Ποσοστό (%) δημοσιεύσεων που λαμβάνουν αναφορές (% cited papers)**

Υπολογίζεται ως το ποσοστό επί τοις εκατό των άρθρων που έχουν λάβει τουλάχιστον μία αναφορά. Ο υπολογισμός λαμβάνει χώρα σε χρονικά διαστήματα επικαλυπτόμενων χρονικών ετών στις εξής κατηγορίες: στο σύνολο της χώρας, ανά κατηγορία οργανισμού.

- **Μερίδιο (%) αναφορών (% Share of citations)**

Υπολογίζεται ως το ποσοστό (%) του αριθμού των αναφορών (χρονικά διαστήματα) που έλαβαν οι επιστημονικές δημοσιεύσεις: της χώρας επί του συνολικού αριθμού των αναφορών που έλαβαν οι δημοσιεύσεις των χωρών μελών ενός οργανισμού ή ενός φορέα επί του συνολικού αριθμού των αναφορών που έλαβαν οι δημοσιεύσεις της κατηγορίας στην οποία ανήκει.

- **Δείκτης απήχησης (Citation impact)**

Ο δείκτης απήχησης είναι ο μέσος όρος αναφορών ανά δημοσίευση και υπολογίζεται ως ο λόγος του αριθμού των αναφορών που καταγράφονται σε ορισμένη χρονική περίοδο προς το συνολικό αριθμό των δημοσιεύσεων της ίδιας χρονικής περιόδου.

- **Σχετικός δείκτης απήχησης (Relative citation impact)**

Ο σχετικός δείκτης απήχησης συγκρίνει την απήχηση των δημοσιεύσεων μιας οντότητας (π.χ. Ιόνιο Πανεπιστήμιο) σε σχέση με την απήχηση των δημοσιεύσεων ενός καθορισμένου συνόλου αναφοράς (π.χ. σύνολο Ελληνικών Πανεπιστημίων) και προκύπτει από τη διαίρεση των αντίστοιχων δεικτών απήχησης. Όταν η τιμή του σχετικού δείκτη απήχησης είναι μεγαλύτερη από 1, οι δημοσιεύσεις της οντότητας που εξετάζεται έχουν μεγαλύτερη απήχηση από το μέσο όρο του συνόλου αναφοράς. Ο δείκτης αυτός δεν λαμβάνει υπόψη τις διαφοροποιήσεις στις πρακτικές αναφορών στις διαφορετικές επιστημονικές περιοχές.

- **Σχετικός δείκτης απήχησης - κανονικοποιημένος βάσει επιστημονικού πεδίου (Field normalized citation score)**

Ο δείκτης απήχησης μετά από κανονικοποίηση, βάσει του αριθμού διαφορετικών επιστημονικών θεματικών περιοχών, συγκρίνει την απήχηση μιας δημοσίευσης σε σχέση με την απήχηση που έχουν οι δημοσιεύσεις σε παγκόσμιο επίπεδο στην ίδια επιστημονική περιοχή. Όταν η τιμή του σχετικού δείκτη απήχησης είναι μεγαλύτερη από 1, οι δημοσιεύσεις της οντότητας που εξετάζεται έχουν μεγαλύτερη απήχηση

από τον παγκόσμιο μέσο όρο. Υπολογίζονται οι σχετικοί δείκτες απήχησης μετά από κανονικοποίηση για: το σύνολο δημοσιεύσεων μιας χώρας ή τις δημοσιεύσεις μιας κατηγορίας φορέων ανά επιστημονικό πεδίο ή το σύνολο των δημοσιεύσεων ενός φορέα ή τις δημοσιεύσεις ενός φορέα ανά επιστημονικό πεδίο.

3.5. Εξειδικευμένα παραδείγματα:

- **Ο δείκτης του στέμματος (crown indicator)**

Ο δείκτης του στέμματος μετρά τον επιστημονικό αντίκτυπο μιας ομάδας συγγραφέων. Συγκρίνει το μέσο αριθμό των αναφορών των δημοσιεύσεων της ομάδας με τον μέσο αριθμό των αναφορών των δημοσιεύσεων των διεθνών εκδόσεων του ίδιου έτους, στην ίδια θεματική περιοχή, ίδιων τύπων εγγράφου. (Karolinska Institutet. <http://ki.se/start>). Συνήθως γράφεται ως δεκαδικός αριθμός, που δείχνει τη σχέση με τον παγκόσμιο μέσο όρο, δηλαδή το 0,9 δείχνει ότι οι δημοσιεύσεις που αναφέρονται είναι 10% λιγότερες από τον παγκόσμιο μέσο όρο και το 1.2 δείχνει ότι οι αναφορές είναι 20% περισσότερες.

- **To Top 5%**

Το Top 5% παρουσιάζει το ποσοστό των δημοσιεύσεων που αποδίδεται σε μια ομάδα συγγραφέων που ανήκουν στο 5% των πιο αναφερομένων δημοσιεύσεων στον κόσμο από το ίδιο έτος, στην ίδια θεματική περιοχή, ίδιων τύπων εγγράφου. Είναι ακριβώς όπως ο δείκτης 'crown', γράφεται ως δεκαδικός αριθμός και δείχνει τη σχέση με τον παγκόσμιο μέσο όρο. Μια τιμή πάνω από 1 δείχνει ότι η υπό ανάλυση μονάδα έχει τις περισσότερες δημοσιεύσεις της μεταξύ των κορυφαίων 5% από τον παγκόσμιο μέσο όρο, ενώ μια τιμή κάτω από 1 ότι έχει τις λιγότερες. Το Top 5% χρησιμοποιείται συχνά ως συμπλήρωμα του δείκτη 'crown'. Δείχνει εάν μια υψηλή τιμή του 'crown indicator' επιτυγχάνεται διαμέσου πολύ ιδιαίτερα αναφερόμενων άρθρων ή ένας μεγαλύτερος αριθμός άρθρων αναφέρεται άνω του μέσου όρου. Μπορεί επίσης να προσδιορίζει ιδιαίτερα υψηλά αναφερόμενα άρθρα από μια ομάδα με χαμηλή τιμή 'crown indicator', του οποίου οι κορυφαίες δημοσιεύσεις διαφορετικά θα ήταν απαρατήρητες. Στο συγκεκριμένο ερευνητικό πεδίο απαιτούνται γνώσεις για να αποφασισθεί ποιο από τα δύο πρότυπα δημοσίευσης είναι η καλύτερη ένδειξη για έρευνα υψηλής ποιότητας.

- **Αριθμός δημοσιεύσεων με υψηλή απήχηση (P Top X%)**

Πρόκειται για τον αριθμό των επιστημονικών δημοσιεύσεων που σε παγκόσμιο επίπεδο και ανά έτος κατατάσσονται υψηλά στην ποσοστιαία κατάταξη των δημοσιεύσεων στο αντίστοιχο επιστημονικό πεδίο. Η κατάταξη πραγματοποιείται βάσει του αριθμού αναφορών. Ο δείκτης υπολογίζεται για χρονικά διαστήματα πενταετίας, αφορά τον αριθμό των δημοσιεύσεων που κατατάσσονται παγκοσμίως στο 1%, 5%, 10%, 25% και 50% των δημοσιεύσεων με την υψηλότερη απήχηση και αναφέρεται: στο σύνολο της χώρας ή ανά φορέα.

- **Ποσοστό (%) δημοσιεύσεων με υψηλή απήχηση (Top X%)**

Υπολογίζεται ως το ποσοστό (%) του αριθμού των δημοσιεύσεων μιας οντότητας που ανά έτος χαρακτηρίζονται ως δημοσιεύσεις με υψηλή απήχηση επί του συνολικού αριθμού των δημοσιεύσεων της οντότητας το ίδιο έτος. Όταν η ποσοστιαία κατανομή των δημοσιεύσεων με υψηλή απήχηση της εξεταζόμενης οντότητας προσεγγίζει ή υπερβαίνει την αντίστοιχη παγκόσμια κατανομή 1%, 5%, 10%, 25% και 50%, η οντότητα θεωρείται ότι προσεγγίζει ή υπερβαίνει αντίστοιχα τον παγκόσμιο μέσο όρο. Ο δείκτης υπολογίζεται για χ χρονικά διαστήματα και αναφέρεται: στο σύνολο της χώρας ή ανά φορέα. http://metrics.ekt.gr/sites/metrics/files/bibliometric_analysis.pdf

- **Μέσος Δείκτης Επιρροής**

Οι δημοσιεύσεις σε επιστημονικά περιοδικά υψηλού δείκτη επιρροής συχνά θεωρείται ότι είναι υψηλής ποιότητας. Δεν είναι ασυνήθιστο για τους ερευνητές να κληθούν να παράσχουν πληροφορίες σχετικά με "το μέσο δείκτη επιρροής" τους, δηλαδή τη μέση τιμή των δεικτών επιρροής των περιοδικών που έχουν δημοσιεύσει, όταν υποβάλλουν αίτηση για χρηματοδότηση έρευνας ή για μια νέα θέση. Μερικές φορές, μια ερευνητική μονάδα ή ένα πανεπιστήμιο εμφανίζει επίσης πόσες εκδόσεις τους έχουν σε περιοδικά με πολύ υψηλούς δείκτες επιρροής ISI, δηλαδή η 20η ή η 40η πιο υψηλή θέση στην κατάταξη ISI, ως ένδειξη της ποιότητας της έρευνας που παράγεται από τους συντάκτες της συγκεκριμένης μονάδας. Ωστόσο, ο δείκτης επιρροής ενός περιοδικού δεν μπορεί να προβλέψει τον αριθμό των αναφορών που κάθε μεμονωμένη δημοσίευση θα λάβει. Συχνά περίπου 20% των δημοσιεύσεων σε ένα περιοδικό λαμβάνουν το 80% των αναφορών και πολλά άρθρα αναφέρονται 0-1 φορές ακόμα και σε περιοδικά υψηλού δείκτη όπως έχουμε ήδη πρωτοπει.

- **Δημοσιεύσεις χωρίς αναφορές**

Το ποσοστό των δημοσιεύσεων που παραμένουν χωρίς αναφορές μετά από ένα ορισμένο χρονικό διάστημα μπορεί να θεωρηθεί το αντίθετο του κορυφαίου δείκτη 5%. Εάν ο δείκτης 'crown' των δημοσιεύσεων μιας ομάδας έχει μεγάλη τιμή, πληροφορίες σχετικές με ένα μεγάλο αριθμό δημοσιεύσεων χωρίς αναφορές συνεπάγεται ότι έγιναν μεγαλύτερες προσπάθειες σε μερικές «εργασίες σημαία».

3.6. Σφάλματα Στρεβλώσης και Αδυναμίες Μέτρησης Αναφορών

Υπάρχουν δύο σημαντικοί τομείς σφάλματος που μπορεί να στρεβλώσουν την καταμέτρηση αναφορών. Ένας είναι η σκόπιμη χειραγώγηση του συστήματος από τους επιστήμονες, που προσπαθούν να εξασφαλίσουν τον υψηλότερο, δυνατό, αριθμό αναφορών των άρθρων τους και ο δεύτερος τομέας ανακρίβειας είναι συνυφασμένος με το πώς τα άρθρα αυτά παρατίθενται, ευρετηριάζονται και αναζητούνται.

Η θεμελιώδης παραδοχή του κάθε βιβλιομετρικού συστήματος είναι ότι η αναφορά αποτελεί μια μονάδα θετικής αποδοχής, αλλά δεν είναι σαφές ότι η συσχέτιση μεταξύ της ποιότητας του ενός εγγράφου και του αριθμού των αναφορών σε αυτό είναι αρκετά ισχυρή, ώστε να αποτελέσει τη βάση για μια αυστηρή αξιολόγηση των ατόμων ή φορέων.

Είναι γνωστό ότι οι συγγραφείς θα θελήσουν να δημοσιεύουν τις εργασίες τους σε περιοδικά υψηλού δείκτη επιρροής, αφού παρόμοιες στρατηγικές χρησιμοποιούνται από τους ερευνητές, ώστε να διασφαλίζεται με αυτόν τον τρόπο ότι θα υπάρξουν αναφορές για τις εργασίες τους. Πρώτη αδυναμία που αναφέρεται είναι η εσφαλμένη αναφορά. Σύμφωνα με τους Simkin και Roychowdhury (2012), μόνο το 20% των εργασιών στην πραγματικότητα διαβάζονται από τους συγγραφείς που τις αναφέρουν. Έχει μελετηθεί λεπτομερώς και έχει βρεθεί να είναι ευρέως διαδεδομένη η χρήση ακατάλληλων αναφορών προς στήριξη των ισχυρισμών σε έγκριτα άρθρα. Είναι δύσκολο να προσδιορισθεί τι ωθεί σε τόσο υψηλά ποσοστά εσφαλμένης αναφοράς. Οι συνέπειες του φαινομένου αυτού στη Βιβλιομετρία είναι το γεγονός ότι κάποια άρθρα υπολογίζονται όταν δεν θα έπρεπε, ενώ αντίστοιχα, η έρευνα που θα έπρεπε να αναφερθεί, δεν λαμβάνεται υπόψη δεόντως.

Δεύτερη αδυναμία είναι οι αναφορές επιστημονικών τεκμηρίων τα οποία περιέχουν ανακρίβειες ή λάθος συμπεράσματα. Οι περισσότεροι συγγραφείς, κάποια στιγμή, αναφέρονται σε έρευνες που έρχονται σε αντίθεση με τις δικές τους. Στην πλειονότητα των περιπτώσεων μια τέτοια εργασία είναι θετική για την επιστήμη, απλώς με διαφορετικά αποτελέσματα, αλλά θα υπάρξουν επίσης περιπτώσεις όπου η άλλη μελέτη είναι απλά λάθος. Φυσικά, τα επιχειρήματα που παρατίθενται για τη στήριξη ενός ισχυρισμού θα μπορούσαν επίσης να είναι λανθασμένα. Σε αυτήν την περίπτωση ο συγγραφέας έχει αποτύχει να δει το λάθος, αν και υπάρχουν σχεδόν σίγουρα περιπτώσεις όπου ένας 'αδίστακτος' συγγραφέας αναφέρεται σκόπιμα σε ένα υποστηρικτικό δεδομένο, ακόμη και αν αυτό αμφισβητεί τα συμπεράσματα. Η αυξανόμενη χρήση της μετά-ανάλυσης μπορεί άθελά της να προωθήσει λάθη, διότι οι μετά-αναλύσεις κατά κανόνα δεν κάνουν διάκριση μεταξύ «καλών» και «κακών» εργασιών. Επιπλέον ορισμένες μετά-αναλύσεις βασίζονται σε δευτερογενή δεδομένα και, ως εκ τούτου, μπορεί να παραποιήσουν ακούσια τα αποτελέσματα της έρευνας.

Η τρίτη αδυναμία αφορά την αναζήτηση και την πρόσβαση. Δεν είναι όλα τα περιοδικά εξίσου προσιτά και είναι λογικό να υποτεθεί ότι οι εργασίες που είναι δύσκολο να βρεθούν, αναφέρονται λιγότερο. Αυτό συνήθως επηρεάζεται και από τις αποφάσεις είτε των βιβλιοθηκών, οι οποίες προκύπτουν, όσον αφορά το πλήθος των τεκμηρίων, από το μέγεθος των προϋπολογισμών τους ή από το εάν ο συγγραφέας έχει την οικονομική δυνατότητα να αντέξει την ανοικτή πρόσβαση. Επίσης στις βιβλιοθήκες μπορεί να δοθεί προτεραιότητα σε έναν πυρήνα περιοδικών υψηλής επιρροής κι έτσι, έμμεσα, ο δείκτης επιρροής ενός περιοδικού, ο οποίος θεωρείται γενικά φτωχός, επηρεάζει την πιθανότητα να χρησιμοποιηθεί ως αναφορά στην εργασία ενός ατόμου. Ακόμα ένα έγγραφο που εμφανίζεται αμέσως στις μηχανές αναζήτησης θα μπορούσε να έχει απλώς έναν αποτελεσματικό τίτλο ή λέξεις κλειδιά. Ένα πιο κατάλληλο άρθρο μπορεί να παραλειφθεί διότι περιλαμβάνει όρους που αναζητούνται σπανιότερα. Έτσι λοιπόν, ακόμη και αν ένα περιοδικό είναι διαθέσιμο, οι επιστημονικές εργασίες που περιέχει μπορεί να αγνοηθούν.

Η τέταρτη αδυναμία έχει να κάνει με τη συλλογή και τη μέτρηση. Υπάρχει κάποια αβεβαιότητα ως προς το πώς ακριβώς οι αριθμοί των αναφορών προσδιορίζονται σε φορείς όπως Thomson ISI Web of Science για τον καθορισμό του δείκτη επιρροής των περιοδικών. Ο αριθμός των αναφορών που λαμβάνει ένα τεκμήριο επίσης εξαρτάται από το ηλεκτρονικό σύστημα που χρησιμοποιείται. Για παράδειγμα οι Meho & Yang, (2007) διαπίστωσαν σημαντικές διαφορές μεταξύ των Google Scholar, Web of Science και Scopus.

Η πέμπτη και τελευταία αδυναμία, αλλά όχι λιγότερο υπολογίσιμη, έχει να κάνει με τα λάθη στις λίστες αναφοράς. Υπάρχει μια σειρά από συγγραφείς που έχουν δημοσιεύσει εργασίες δίνοντας λεπτομέρειες για τα τυπογραφικά και άλλα τέτοια λάθη – τα οποία συχνά αποκαλούνται «λάθη αναφορών» - που βρέθηκαν

στους καταλόγους αναφοράς των περιοδικών. Λαμβάνοντας τα δεδομένα από τις 5 ξεχωριστές μελέτες βιοϊατρικών περιοδικών, ο μέσος αριθμός των λανθασμένων αναφερόμενων άρθρων είναι ένα εντυπωσιακό 34,28 %. Τέτοια λάθη μπορεί να κάνουν την ανάκτηση του επιθυμητού άρθρου δύσκολη και μερικές φορές αδύνατη, μειώνοντας έτσι την πιθανότητα του εσφαλμένου εγγράφου που αναφέρθηκε να χρησιμοποιηθεί ως αναφορά. Μια αναφορά που δεν εμφανίζεται σε μια βάση δεδομένων, επειδή το όνομα του συγγραφέα έχει ορθογραφικά λάθη, για παράδειγμα, παραβλέπεται. Θα ήταν σχετικά εύκολο να διορθωθούν αυτά τα είδη των σφαλμάτων, αλλά η χρονοβόρα φύση της διαδικασίας σημαίνει ότι αυτό, πιθανώς, δεν θα συμβεί μέχρι να υπάρξει μεγαλύτερο κίνητρο για τους εκδότες, ώστε να το αλλάξουν ή να αναπτυχθεί λογισμικό που θα ελέγχει αυτόματα αναφορές για λάθη. Πάντως, σύμφωνα με τη Διεθνή Επιτροπή των Συντακτών Ιατρικού Τύπου, η ευθύνη ανήκει στους συγγραφείς, οι οποίοι πρέπει να βεβαιωθούν ότι οι αναφορές τους είναι σωστές.

Τέλος, τα σφάλματα και οι ασυνέπειες που παρατέθηκαν μπορούν να οδηγήσουν σε λανθασμένη εκτίμηση και ανακρίβειες, που μπορούν είτε να αυξήσουν ή να μειώσουν τον πραγματικό αριθμό των αναφορών που λαμβάνει μια εργασία-άρθρο. Μετρώντας αναφορές ή αξιοποιώντας τη χρήση των δεικτών που βασίζονται σε μετρήσεις (π.χ. h-index, g-index), επιφανειακά φαίνεται να είναι αντικειμενική και αποτελεσματική μέθοδος που ποσοτικοποιεί τη «ποιότητα» της εργασίας-άρθρου, όπως και του συγγραφέα του. Δυστυχώς, η μέτρηση των αναφορών μπορεί να αντανakλά επίσης μια ολόκληρη σειρά από λάθη και δεδομένα που στρεβλώνουν το μετρικό σε σημείο, που δεν πρέπει να ληφθεί υπόψη εκ πρώτης όψεως. Κλείνοντας διαπιστώνουμε αυτό που αναφέραμε και στην εισαγωγή, ότι δηλαδή οι δείκτες επιρροής και οι αναφορές δεν συνιστούν ένα σύστημα που βασίζεται εξ' ολοκλήρου σε αυτά που χρειάζεται για να αξιολογήσει κάποιος τη συμβολή ενός επιστήμονα. Παρ' όλα αυτά, αναγνωρίζεται ότι υπάρχουν επίσης ελλείψεις σε επιτροπές αξιολόγησης και κατανοείται, ως εκ τούτου, ότι η ζήτηση για τα ποσοτικά μέτρα είναι βέβαιο πως θα παραμείνει υψηλή. Δεδομένης της πολυπλοκότητας του έργου αξιολόγησης, μια πολύπλευρη διαδικασία που ενσωματώνει τη Βιβλιομετρία, την αξιολόγηση από ομότιμους, καθώς και άλλους δείκτες είναι ένας συμβιβασμός που θα μπορούσε να είναι αρεστός σε όλα τα εμπλεκόμενα μέρη. Σύμφωνα με τον Garfield, από τα 38 εκατομμύρια τεκμήρια που αναφέρθηκαν από το 1900 μέχρι το 2005, μόνο το 0,5% αναφέρθηκαν περισσότερες από 200 φορές, ενώ τα μισά δεν αναφέρθηκαν καθόλου. Με αυτό το μάλλον συγκλονιστικό στατιστικό δεδομένο ο Garfield μας αναγκάζει να αναγνωρίσουμε ότι μετρήσεις μας εστιάζουν πρώτον, στο 50% των τεκμηρίων τα οποία αναφέρονται και δεύτερον, στο μικρό ποσοστό εκείνων που αναφέρονται ιδιαίτερα.

Σύγκριση εργασιών σε διαφορετικά επιστημονικά πεδία τα οποία εντάσσονται σε διαφορετικά υποδείγματα δημοσιεύσεων και αναφορών

Για τη συγκριτική αξιολόγηση της απήχησης των δημοσιεύσεων χρησιμοποιείται ο σχετικός δείκτης απήχησης. Σε κάθε άλλη περίπτωση υπολογισμού δεικτών απήχησης (π.χ. ανά κατηγορία φορέων, ανά φορέα, ανά επιστημονικό πεδίο) χρησιμοποιείται ο σχετικός δείκτης απήχησης μετά από κανονικοποίηση βάσει επιστημονικού πεδίου (Δές 3.4). Ανάλογα με τους επιστημονικούς τομείς, υπάρχουν σημαντικές διαφοροποιήσεις στην πρακτική δημοσίευσης, στο χρόνο απαξίωσης των ερευνητικών αποτελεσμάτων και στα μοτίβα αναφοράς δημοσιεύσεων (citation patterns). Σε αυτήν την περίπτωση, οι αναφορές κάθε δημοσίευσης διαιρούνται με τον μέσο όρο των αναφορών που λαμβάνουν παγκοσμίως οι δημοσιεύσεις της εκάστοτε εξειδικευμένης θεματικής περιοχής και του αντίστοιχου έτους δημοσίευσης. Με τον τρόπο αυτό υπολογίζονται οι σχετικοί δείκτες απήχησης για κάθε κατηγορία φορέων, φορέα ή επιστημονικό πεδίο. Για παράδειγμα, ο συνολικός δείκτης απήχησης ενός φορέα προκύπτει από τον μέσο όρο των τιμών για όλες τις δημοσιεύσεις του φορέα (συνολικός κανονικοποιημένος δείκτης απήχησης).

Αντίστοιχα, ο σχετικός δείκτης απήχησης σε ένα επιστημονικό πεδίο προκύπτει από τον μέσο όρο των τιμών για τις δημοσιεύσεις που εντάσσονται στο επιστημονικό αυτό πεδίο κ.ο.κ. Με τον τρόπο αυτό λαμβάνονται υπόψη οι διαφοροποιήσεις στον αριθμό αναφορών τόσο στα διάφορα επιστημονικά πεδία όσο και σε σχέση με τον χρόνο δημοσίευσης. Όταν η τιμή του σχετικού δείκτη απήχησης είναι μεγαλύτερη από 1, οι δημοσιεύσεις της οντότητας που εξετάζεται έχουν μεγαλύτερη απήχηση από τον παγκόσμιο μέσο όρο.

Εργαλεία μέτρησης και αξιολόγησης συγγραφέα

1) SciVal

Το SciVal είναι μια συνεισφορά βασισμένη σε ένα εργαλείο ερευνητικής προσπάθειας εκτίμησης της επίδρασης που χρησιμοποιεί δεδομένα από το Scopus. Μας επιτρέπει να αξιολογήσουμε μεμονωμένους ερευνητές, ομάδες από ερευνητές και ιδρύματα βασισμένα σε μια ποικιλία από βιβλιομετρικά δεδομένα. Αυτό

συγκεντρώνει πληροφορίες για 4600 ερευνητικά ιδρύματα και 220 χώρες. Παρέχεται μέσω μιας συνδρομής που, εκτός των άλλων, προσφέρει:

- επιστημονική παραγωγή,
- καταμέτρηση αναφοράς,
- αναφορές ανά δημοσιεύσεις,
- προαναφερόμενες δημοσιεύσεις,
- σταθμισμένο πεδίο αναφοράς επιδράσεων,
- αποτελέσματα σε ανώτατα εκατοστημόρια,
- h-index,
- καταμέτρηση κατηγορίας περιοδικού,
- αριθμοί αναφερομένων χωρών,
- συνεργασία,
- επίπτωση συνεργασίας,
- ακαδημαϊκή συνεργασία.

2) **Publish or Perish**

Η Anne-Will Harzing, Καθηγήτρια Διοικητικής Επιστήμης στο Πανεπιστήμιο της Μελβούρνης, δημιούργησε το Publish or Perish (PoP) (Δημοσίευση ή Φθορά) για να βοηθήσει τους ενδιαφερόμενους ερευνητές που αναζητούν βιβλιομετρικά δεδομένα. Το PoP είναι ένα ελεύθερο πρόγραμμα, διαθέσιμο για download, που μαζεύει τα δεδομένα από το Google Scholar με βάση το όνομα του συγγραφέα. Οι χρήστες μπορούν να αφαιρέσουν χειροκίνητα τα αρχεία για να βελτιώσουν τα δεδομένα, όπως η δυνατότητα που προσφέρεται τώρα από το Google Scholar Citations.

Χρήση του Publish or Perish (PoP)

Η PoP μπορεί να υπολογίσει πολλά μετρικά, συμπεριλαμβανομένων εναλλακτικών λύσεων για το h-index. Ωστόσο, επειδή λίγοι άνθρωποι είναι εξοικειωμένοι με υπολογισμούς του h-index, εναπόκειται στους χρήστες να εξηγήσουν αυτά τα μετρικά σε μεγαλύτερα ακροατήρια.

Οδηγίες

Για την πραγματοποίηση της εργασίας θα χρειαστείτε την βοήθεια της εφαρμογής Google Scholar (<http://scholar.google.com>). Για να μπορέσετε να εξάγετε τους βιβλιομετρικούς δείκτες που θα σας ζητηθούν παρακάτω θα χρειαστεί να κατεβάσετε την εφαρμογή "Publish or Perish", της Anne Harzing. Η διεύθυνση από την οποία θα μπορέσετε να κατεβάσετε την εφαρμογή είναι η ακόλουθη:

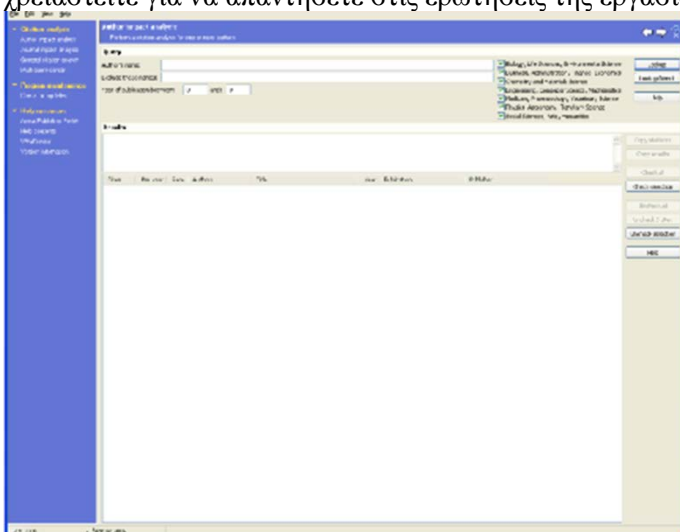
<http://www.harzing.com/download/PoPSetup.exe>

Το παρακάτω παράθυρο απεικονίζει τον διάλογο εγκατάστασης του προγράμματος στον υπολογιστή σας.



Εικόνα 3.13 Διάλογος εγκατάστασης του προγράμματος στον υπολογιστή

Με το που ολοκληρωθεί η εγκατάσταση του προγράμματος θα είσατε σε θέση να χρησιμοποιήσετε τη συγκεκριμένη εφαρμογή για την πραγματοποίηση της εργασίας. Η παρακάτω εικόνα υποδεικνύει τη κύρια διεπαφή της εφαρμογής την οποία και θα χρησιμοποιήσετε για την εξαγωγή των δεδομένων, τα οποία θα χρειαστείτε για να απαντήσετε στις ερωτήσεις της εργασίας.



Εικόνα 3.14 Οθόνη κύριας διεπαφής της εφαρμογής για την εξαγωγή δεδομένων

Τα πεδία της ομάδας Query που βλέπετε στην παραπάνω εικόνα έχουν ως εξής:

Author's Name: Το όνομα του συγγραφέα, όπως αυτό απεικονίζεται στην ερευνητική εργασία που θα σας δοθεί να αναλύσετε.

Exclude these names: Αν και εφόσον επιθυμείτε να εξαιρέσετε κάποια από τα ονόματα στις εργασίες που θα εμφανιστούν από την αναζήτηση σας, συμπληρώστε τα σε αυτό το πεδίο.

Year of Publication between: Δώστε το εύρος της χρονικής διάρκειας στην οποία θα γίνει η αναζήτηση των ερευνητικών εργασιών που θέλετε να πραγματοποιήσετε. Δίπλα από τα πεδία αυτής της ομάδας θα δείτε μια σειρά από πεδία check box που αφορούν τις ερευνητικές περιοχές. Η προεπιλεγμένη κατάσταση είναι όλα επιλεγμένα. Ωστόσο στην εκφώνηση της άσκησης θα σας δίνεται η ακριβής θεματική περιοχή. Για αυτό το λόγο θα πρέπει να έχετε επιλεγμένα μόνο το πεδίο ή πεδία στα οποία ανταποκρίνεται η θεματική περιοχή που αναζητάτε. Τέλος, για να πραγματοποιήσετε την αναζήτηση πατήστε το κουμπί **Lookup**, στην πάνω αριστερή γωνία της διεπαφής του προγράμματος.

Εφαρμογή

Η ερευνητική εργασία "Authoritative Sources in a hyperlinked environment" του Καθηγητή John Kleinberg του Πανεπιστημίου Cornet της Νέας Υόρκης θεωρείται ένα κλασικό κομμάτι ερευνητικής εργασίας στο πεδίο των μηχανών αναζήτησης και της αναζήτησης πληροφορίας στο διαδίκτυο.

Η πλήρης βιβλιογραφική εγγραφή της εργασίας με βάση το πρότυπο APA δίνεται ακολούθως:

Kleinberg, John (1999). *Authoritative Sources in a Hyperlinked Environment*. Journal of the ACM. Vol 46, Issue 5. pp 604-632.

Με τη χρήση του προγράμματος Publish or Perish και για το πεδίο: Engineering, Computer Science and Mathematics να απαντήσετε στα παρακάτω ερωτήματα. Για την απάντησή σας να συμπεριλάβετε την εκφώνηση του ερωτήματος.

Ερωτήματα

- 1) Να βρείτε πόσες εργασίες από τον ίδιο συγγραφέα έχουν δημοσιευτεί πριν από την ημερομηνία έκδοσης της συγκεκριμένης εργασίας. [Υποβοήθηση: Χρησιμοποιήστε το πεδίο εύρους της ημερομηνίας]
- 2) Χρησιμοποιώντας τα παραπάνω δεδομένα, να αναπαραστήσετε σε γράφημα τύπου barchart τον αριθμό ερευνητικών δημοσιεύσεων ανά έτος του ίδιου συγγραφέα, μέχρι και το έτος δημοσίευσης της συγκεκριμένης εργασίας (1999). Ο οριζόντιος άξονας θα είναι αυξητικά ταξινομημένος ανά έτος (π.χ 1996, 1997, 1998 κ.τ.λ.).
- 3) Επαναλάβετε τη διαδικασία για όλες τις δημοσιεύσεις του συγκεκριμένου συγγραφέα έως το έτος 2008 και αναπαραστήστε τις σε καινούριο γράφημα με εύρος από το 1999 μέχρι το 2008.
- 4) Για τις εργασίες από το 1999 μέχρι το 2008, να βρείτε τον αριθμό αναφορών της κάθε μίας (χρησιμοποιήστε τη τιμή του πεδίου Cites δίπλα στον τίτλο της κάθε εργασίας).
- 5) Να αναπαραστήσετε σε γράφημα μορφής XYscatter plot τον αριθμό των εργασιών (άξονας X) και τον αριθμό των αναφορών (άξονας Y). Π.χ. το ζεύγος χ,ψ μπορεί να είναι (1, 19), (2, 45), (3, 56) κ.τ.λ.
- 6) Να βρείτε τον δείκτη σημαντικότητας κατά Hirsch (H) του ερευνητή και να τον παραστήσετε με ευθεία γραμμή στο γράφημα.
- 7) Ο δείκτης m υπολογίζει την ερευνητική παραγωγικότητα του ερευνητή ως εξής: $m = h/N$, όπου h ο δείκτης Hirsch και N η χρονική περίοδος που αντιστοιχεί στην διαφορά της τελευταίας δημοσίευσης από την αρχική δημοσίευση στο γράφημα της μορφής XYscatter plot. Να υπολογίσετε τον δείκτη m για τον συγκεκριμένο ερευνητή.

Εύρεση του h-Index δύο ερευνητών από δύο διαφορετικές βάσεις

Υπολογίστε τον h-Index χρησιμοποιώντας τις πηγές δεδομένων a,b:

- Web of Science
- Scopus

Για τους ερευνητές

Jorge E. Hirsch, Professor of Physics, University of California, San Diego

και

Peter Vinkler, Senior Scientist and Scientific Secretary at the Chemical Research Center of the Hungarian Academy of Sciences, Budapest, Hungary

Επιπροσθέτως:

Υπολογίστε, στη συνέχεια, τον h-index του Peter Vinkler ως χημικού και ως βιβλιομέτρη.

Εύρεση κατάλληλου περιοδικού για δημοσίευση

Ένας επιστήμονας (βιοχημικός) επιθυμεί να δημοσιεύσει τα τελευταία αποτελέσματα του στο πεδίο "macromolecular assemblies". Ποια περιοδικά θα του προτείνατε για δημοσίευση, με βάση ποια κριτήρια (όπως π.χ. το πεδίο εφαρμογής και την προβολή);

Μετά την πρώτη σύσταση, αυτός / αυτή τονίζει ότι τα ευρήματα είναι πολύ καινοτόμα και ότι αυτός / αυτή ενδιαφέρεται για μια πολύ γρήγορη διάδοση. Τι άλλο θα μπορούσε να συστήσετε; Συζητήστε διαφορετικές στρατηγικές δημοσίευσης.

Στόχοι μάθησης

- Συνδυασμός του Web of Science και του JCR για τις στρατηγικές δημοσίευσης.
- Εξοικείωση με το αντικείμενο της κατάταξης κατά ISI και αντιμετώπιση των δυσκολιών των πολλαπλών πεδίων στόχευσης των περιοδικών.
- Κατανόηση των επιπτώσεων για τις ταξινομήσεις περιοδικών (διαφορετικές θέσεις κατάταξης του ίδιου περιοδικού, σε διάφορες κατηγορίες).

Στόχοι μάθησης της εφαρμογής

- Να κατανοηθεί ότι η κάλυψη των αναφορών διαφέρει μεταξύ των βάσεων δεδομένων.
- Να κατανοηθεί ότι οι διαφορές αυτές επηρεάζουν την έκβαση των υπολογισμών του h-index και ότι πρέπει να δηλωθεί ρητά η βάση δεδομένων που χρησιμοποιείται.
- Να ευαισθητοποιηθεί για τα προβλήματα και τις ασυνέπειες που δημιουργούνται όταν γίνεται σύγκριση των h-δείκτων, που υπολογίζονται με τη χρήση διαφορετικών βάσεων δεδομένων.
- Να κατανοηθεί ότι η τιμή του δείκτη επηρεάζεται από το χρονικό διάστημα ή τα χρόνια της εκδοτικής δραστηριότητας του ερευνητή.
- Να ευαισθητοποιηθεί στο γεγονός ότι ένας νεαρός ερευνητής, που μόλις ξεκινά τη σταδιοδρομία του, βρίσκεται σε μειονεκτική θέση σε σύγκριση με ένα ερευνητή με περισσότερη εμπειρία.
- Να ευαισθητοποιηθεί ότι ο δείκτης πρέπει να χρησιμοποιηθεί μόνο για συγκρίσεις σε ερευνητές στην ίδιο στάδιο της σταδιοδρομίας τους.
- Να κατανοηθεί η γνώση των διαφορών στα διαφορετικά πεδία των επιστημών.
- Να κατανοηθεί η γνώση των διαφορετικών συνηθειών δημοσίευσης.

Scholarometer

- Το Scholarometer αποτελεί πρόσθετο για τον Firefox ή τον Chrome και παρέχει μια εναλλακτική για το PublishorPerish, αν χρησιμοποιείται το GoogleScholar ως η βασική πηγή δεδομένων για τη βιβλιομετρική ανάλυση.
- Αναπτύχθηκε από τους Diep Thi Hoang και Fil Menczer της Σχολής Πληροφορικής και Υπολογιστών της Ινδιάννα.
- Μπορείτε να βρείτε περισσότερες πληροφορίες σχετικά με αυτό στο <http://scholarometer.indiana.edu>
- Είναι αρκετά νέο.
- Λειτουργεί ως πλευρική μπάρα στον Firefox και παρέχει παρόμοιες διευκολύνσεις με το PublishorPerish, αλλά σε πολύ περιορισμένο σύνολο βιβλιομετρικών αναλύσεων.
- Σημαντικά χαρακτηριστικά είναι η ικανότητα να εξάγει διάφορες μορφές ονομάτων, να απομακρύνει ανεπιθύμητες μορφές του ονόματος, να αναδεικνύει διπλοεγγραφές άρθρων και να απομακρύνει μεμονωμένα στοιχεία από το σύνολο των αποτελεσμάτων τα οποία δεν αντιστοιχούν στο σωστό συγγραφέα – αυτό βελτιώνει τις περιορισμένες ικανότητες του PublishorPerish που το κάνει ιδιαίτερα δύσκολο για τη χρήση συνωνυμίων.
- Άλλο ένα χαρακτηριστικό αυτού του προϊόντος είναι η χρήση του Web 2.0. όταν αναζητούμε ένα συγγραφέα. Οι ερευνητές μπορούν να εντοπίσουν ετικέτες (tags) για την αναζήτηση τομέων ενδιαφέροντος και αυτές μπορούν να εμφανιστούν στο twitter και να χρησιμοποιούνται σε κάποια μετρικά εργαλεία. Αυτό το στοιχείο του προϊόντος είναι ακόμα σε πειραματικό στάδιο.

Web of Science

Το Web of Science εισήγαγε ένα χαρακτηριστικό που ονομάζεται Researcher ID. Αυτό επιτρέπει σε έναν συγγραφέα να δημιουργήσει ένα (δωρεάν) μοναδικό προσωπικό αναγνωριστικό, το οποίο μπορεί να συνδεθεί με τις δημοσιεύσεις του στο Web of Science, ανεξάρτητα από τις διακυμάνσεις στη χρήση του

ονοματός του και ανεξάρτητα από το πότε ένα συγκεκριμένο έγγραφο δημοσιεύθηκε. Ο Researcher ID βοηθά να προσδιοριστεί μοναδικά ένας μεμονωμένος ερευνητής, κάτι που μπορεί να είναι δύσκολο για τους ανθρώπους με ένα κοινό επώνυμο. Ο Researcher ID εξασφαλίζει, επίσης, τη σωστή απόδοση των εγγράφων και την ακρίβεια των βιβλιομετρικών αναλύσεων.

Για να εγγραφείτε και δημιουργήσετε ένα δικό σας ResearcherID.:

- Είσοδος στο Web of Science μέσω του καταλόγου ιστοσελίδας της Βιβλιοθήκης.
- Από την πρώτη σελίδα πηγαίνετε στο τμήμα που ονομάζεται “ Προσαρμόσετε την εμπειρία σας” και, στη συνέχεια, επιλέξτε “Εγγραφή”.
- Όταν έχετε δημιουργήσει ένα λογαριασμό και συνδεθείτε, επιλέξτε Researcher ID από τη γραμμή εργαλείων στο πάνω μέρος της οθόνης.
- Συμπληρώστε τις λεπτομέρειες, όπως ζητήθηκε για να ολοκληρώσετε την εγγραφή σας.

3.7 Βιβλιομετρικοί Νόμοι: εφαρμογή σε εμπειρικά σύνολα δεδομένων

3.7.1. Ο Νόμος του Zipf

Ο πιο ισχυρός, ευρέως φάσματος, νόμος της Βιβλιομετρίας είναι ο νόμος του Zipf. Ο νόμος του Zipf βρίσκει εφαρμογή σε πληθώρα επιστημονικών κλάδων που ασχολούνται με τη φυσική γλώσσα όπως αυτών της γλωσσολογίας και ειδικότερα της ποσοτικής γλωσσολογίας και της υπολογιστικής ψυχολογολογίας. Ο νόμος βρίσκει εφαρμογή εκτός από τη φυσική γλώσσα και σε πληθώρα άλλων επιστημών όπως στη μουσική, στα υπολογιστικά συστήματα, στο διαδίκτυο, στα φυσικά και βιολογικά συστήματα. Ο Zipf ισχυρίστηκε πως ο νόμος του είναι βασισμένος στη βασική προγνωστική ανθρώπινη συμπεριφορά: προσπαθεί να ελαχιστοποιήσει την προσπάθεια. Ως εκ τούτου, το έργο του Zipf εφαρμόζεται σχεδόν σε κάθε τομέα όπου εμπλέκεται η ανθρώπινη παραγωγή. Ο νόμος του Zipf περιγράφει την σχέση μεταξύ της συχνότητας εμφάνισης λέξεων σε ένα σώμα κειμένου (corpus) και της κατάταξής τους. Η μαθηματική έκφραση του νόμου είναι:

$$r \cdot f = c$$

όπου το r αντιστοιχεί στην κατάταξη μιας λέξης, το f στην συχνότητα εμφάνισης της λέξης και το c αντιστοιχεί σε μια σταθερά, η οποία εξαρτάται από το σώμα κειμένου.

Τοποθετώντας, λοιπόν, τις λέξεις ενός σώματος κειμένου σε φθίνουσα σειρά, ξεκινώντας από αυτήν με την μεγαλύτερη συχνότητα, τότε η δεύτερη συχνότερη λέξη θα εμφανίζεται περίπου τις μισές φορές από ότι η πρώτη και η τρίτη συχνότερη λέξη περίπου 1/3 φορές από ότι η πρώτη κ.ό.κ. Έτσι πολλαπλασιάζοντας την κατάταξη r με τη συχνότητα f της κάθε λέξης, η σταθερά c θα πρέπει να παραμένει περίπου ίδια για κάθε λέξη. Προκύπτει, λοιπόν, μια αντιστρόφως ανάλογη σχέση μεταξύ κατάταξης και συχνότητας των λέξεων ενός κειμένου.

Στην εργασία του Wyllys (1981) ο νόμος του Zipf, εκτός από την παραπάνω αλγεβρική του έκφραση, περιγράφεται επίσης ως ισοδύναμος με τη γραφική αναπαράσταση:

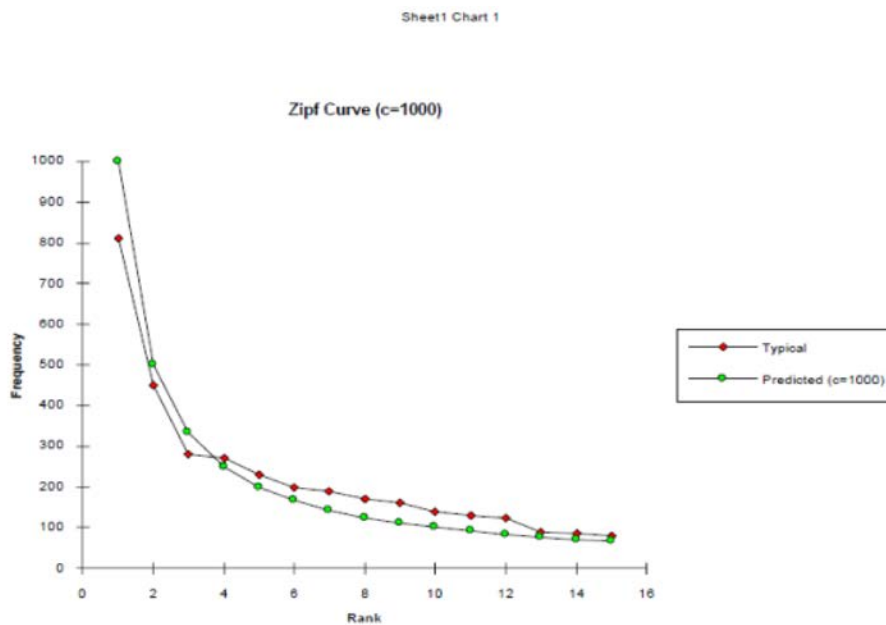
$$\log r + \log f = \log c$$

όπου στη σχεδίαση των ζευγών σημείων που προκύπτουν, ο λογάριθμος της κατάταξης r τοποθετείται στον οριζόντιο άξονα και ο λογάριθμος της συχνότητας f στον κάθετο άξονα και έτσι τα σημεία σχηματίζουν μια ελαφρά καμπύλη γραμμή, γνωστή και ως καμπύλη Zipf (Zipf's curves).

Ο Wyllys αναφέρει επίσης ότι ο υπολογισμός με τον νόμο του Zipf έχει πιο έγκυρα αποτελέσματα κυρίως όσον αφορά την κατάταξη λέξεων με μεσαία τάξη εμφάνισης, παρά για τις λέξεις με πολύ υψηλή ή χαμηλή συχνότητα εμφάνισης. Ακόμη αναφέρει ότι η εργασία του Zipf δείχνει πως το μέγεθος του δείγματος θα πρέπει να αποτελείται από τουλάχιστον 5000 λέξεις, ώστε το $r \cdot f$ να είναι σταθερό, ακόμη και για τις μεσαίες κατατάξεις.

Όροι	Κατάταξη όρων	Συχνότητα εμφάνισης όρων	Υπολογισμός $C = r \cdot f$	Αναμενόμενη συχνότητα $C=1000$
The	1	810	810	1000
Of	2	450	900	500
A	3	280	840	333
Information	4	270	1080	250
Is	5	230	1150	200
To	6	200	1200	167
And	7	190	1330	143
That	8	170	1360	125
As	9	160	1440	111
In	10	140	1400	100
We	11	130	1430	91
Be	12	125	1500	83
Or	13	90	1170	77
May	14	85	1190	71
by	15	80	1200	67

Πίνακας 3.5 Παράδειγμα κατανομής λέξεων σώματος κειμένου σύμφωνα με τον νόμο του Zipf



Εικόνα 3.15 Η καμπύλη του Zipf για τις πραγματικές και τις προβλεπόμενες τιμές του παραδείγματος

3.7.2. Ο Νόμος του Samuel Clement Bradford 1934

Ο νόμος αναφέρει πώς τα άρθρα, σε ένα συγκεκριμένο θεματικό πεδίο, κατανέμονται στα περιοδικά. Ο στόχος του είναι να αναπτύξει μια μέθοδο για την εξακρίβωση των πιο παραγωγικών περιοδικών σε ένα θεματικό πεδίο και να διαχειριστεί αυτό που ο ίδιος απεκάλεσε “χάος των εγγράφων” (documentary chaos). Χρησιμοποιεί ως μία γενική οδηγία προς τους βιβλιοθηκονόμους για τον προσδιορισμό των βασικών τίτλων περιοδικών σε μία συγκεκριμένη θεματική περιοχή. Δηλώνει ότι τα περιοδικά σε ένα συγκεκριμένο θεματικό πεδίο-όπως έχουμε ήδη αναφέρει- μπορούν να διαιρεθούν σε τρεις ομάδες, η κάθε μία εκ των οποίων περιέχει

τον ίδιο αριθμό άρθρων. Η πρώτη βασική ομάδα αποτελείται από έναν σχετικά μικρό σε αριθμό πυρήνα τίτλων περιοδικών για το πεδίο αυτό, που συγκεντρώνει περίπου το ένα τρίτο του συνόλου των άρθρων, η δεύτερη ομάδα περιέχει τον ίδιο αριθμό άρθρων (1/3) με την πρώτη, αλλά ένα μεγαλύτερο αριθμό τίτλων περιοδικών, και η τρίτη ομάδα περιέχει τον ίδιο αριθμό άρθρων (1/3) με τη δεύτερη και πρώτη και ένα ακόμη μεγαλύτερο αριθμό τίτλων περιοδικών. Η μαθηματική σχέση του αριθμού των περιοδικών στον πυρήνα της πρώτης ομάδας είναι μία σταθερά k και στη δεύτερη ζώνη η σχέση είναι k^1 . Ο Bradford εξέφρασε πως ο αριθμός των περιοδικών στις τρεις παραπάνω ομάδες αυξάνει με την αναλογία $k^0 : k^1 : k^2, \dots$ (π.χ. αν έχουμε μια ομάδα 390 άρθρων που αναφέρονται πάνω σε ένα θέμα, δημοσιευμένων σε 39 περιοδικά και τα 130 βρίσκονται στα τρία πρώτα περιοδικά -από άποψη αριθμού δημοσιεύσεων σε αυτήν την κατηγορία- τότε η αναλογία σε συνεχόμενες ομάδες των 130 άρθρων, θα ήταν η παρακάτω:

Τα πρώτα 3 περιοδικά από το σύνολο των 39 αυξάνουν με k^0 τα επόμενα 3 με k^1 τα επόμενα 3 με k^2 κ.ο.κ. Άρα $3 \cdot k^0 + 3 \cdot k^1 + 3 \cdot k^2 = 39$ ή $3(k^0 + k^1 + k^2) = 39$ ή $k^0 + k^1 + k^2 = 13$ ή $k + k^2 + 1 = 13$ ή $k^2 + k - 12 = 0$ ή $k = 3$.

Άρα ο αριθμός των περιοδικών ανά ομάδα θα είναι: 3, 9, 27

αριθμός περιοδικών	Αρ. δημοσί/ων Ανά περ/κό	Συνολικός αρ δημοσιεύσεων
3 {	1	60
	2	70
9 {	1	30
	2	50
	2	18
	4	32
27 {	10	60
	7	35
	5	20
	5	15

Εικόνα 3.16 Ο αριθμός των περιοδικών που κατανέμεται σε τρεις ομάδες σε κάθε θεματικό πεδίο

3.7.3. Νόμος του Lotka

Περιγράφει τη συχνότητα δημοσίευσης των συγγραφέων ενός συγκεκριμένου τομέα. Δηλώνει ότι ο αριθμός των συγγραφέων που έχει παράγει x άρθρα είναι περίπου $1/x^2$ αυτών που δημιουργούν μόνο ένα και η αναλογία των συνεισφερόντων που έχουν μία συμβολή είναι περίπου 60,8%. Αυτό σημαίνει ότι από όλους τους συγγραφείς σε ένα συγκεκριμένο τομέα, 60,8% θα έχουν μόνο μία δημοσίευση και 15,2% θα έχουν δύο δημοσιεύσεις ($1/2^2$ επί 60,8), 6,8% των συγγραφέων θα έχουν τρεις δημοσιεύσεις ($1/3^2$ επί 60,8), και ούτω καθ' εξής. Σύμφωνα με τον νόμο της επιστημονικής παραγωγικότητας του Lotka, μόνο 6 στους 1000 συγγραφείς σε ένα τομέα θα παράγουν περισσότερα από δέκα άρθρα (Palmquist, 2005).

Ο νόμος του Lotka δίδεται από την σχέση $f(x) = k/x^a$

Όπου a, k σταθερές με τιμές >0 :

Το x παίρνει τιμές: $x=1, 2, \dots$

Το $f(x)$ είναι ο αριθμός (ή το ποσοστό) των συγγραφέων με x δημοσιεύσεις

Οπότε παίρνοντας όπου $a=2$, $k=608$ (τιμές που έβαλε ο Lotka) για συλλογές έστω 1000 συγγραφέων, μπορούμε να προβλέψουμε πόσοι συγγραφείς $f(x)$ έχουν γράψει x δημοσιεύσεις (πόσοι μία δημοσίευση, πόσοι δύο κ.ό.κ)

Αναμενόμενος αριθμός

Αριθμός συγγραφέων

Ποσοστό συγγραφέων

δημοσιεύσεων

1	$608 \cdot 1000 / 1^2 = 608$	60.8%
2	$608 \cdot 1000 / 2^2 = 152$	15.2%
3	$608 \cdot 1000 / 3^2 = 68$	6.8%
4	$608 \cdot 1000 / 4^2 = 38$	3.8%
5	$608 \cdot 1000 / 5^2 = 24$	2.4%
6	$608 \cdot 1000 / 6^2 = 17$	1.7%

Πίνακας 3.6 Συχνότητα κατανομής της επιστημονικής παραγωγικότητας (δημοσίευσης των συγγραφέων σε ένα συγκεκριμένο τομέα).

Ωστόσο ο Λότκα δεν λαμβάνει υπόψη του τον αντίκτυπο, μόνο τους αριθμούς παραγωγής. Επιπλέον το 1974 ο Voos διαπίστωσε ότι στην επιστήμη της Πληροφορίας, η αναλογία επί του παρόντος ήταν $1:n^{3.5}$ (Voos, 1974). Έτσι μπορούμε να πούμε ότι ο νόμος του Λότκα ίσως να μην είναι σταθερός στην τιμή του εκθέτη της δύναμης, αλλά στον τύπο του αντίστροφου τετράγωνου. Η πρόκληση, τότε, για εμάς θα είναι να βρούμε τον σωστό εκθέτη σε διαφορετικά μέσα και πεδία.

Βιβλιογραφία

- Altmann, G. (1980). Prolegomena to Menzerath's law. *Glottometrica*, 2, 1–10.
- Bensman, S. J. (2007). GARFIELD AND THE IMPACT FACTOR: THE CREATION, UTILIZATION, AND VALIDATION OF A CITATION MEASURE Part 2, The Probabilistic, Statistical, and Sociological Bases of the Measure By. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.160.6965>
- Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The Eigenfactor™ metrics. *The Journal of Neuroscience*, 28(45), 11433–11434.
- Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.
- Bradford, S. C. (1985). Sources of information on specific subjects. *Journal of Information Science*, (10), 173–180.
- Garfield, E., & Merton, R. K. (1979). *Citation indexing: Its theory and application in science, technology, and humanities* (Vol. 8). Wiley New York. Retrieved from <http://www.garfield.library.upenn.edu/cifwd.html>
- Glänzel, W. (2003). Bibliometrics as a research field. Retrieved May, 25, 2006.
- Guerrero-Bote, V. P., & Moya-Anegón, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics*, 6(4), 674–688.
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.
- Karolinska Institutet. <http://ki.se/start>
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of Washington Academy Sciences*. Retrieved from <http://doi.apa.org/?uid=1926-10141-003>
- Lundberg, J. (2006). *Bibliometrics as a research assessment tool: impact beyond the impact factor*. Institutionen för lärande, informatik, management och etik, LIME/Department of Learning, Informatics, Management and Ethics (Lime). Retrieved from <https://publications.ki.se/xmlui/handle/10616/39489>
- Lundberg, J. (2007). Lifting the crown—citation z-score. *Journal of Informetrics*, 1(2), 145–154.
- Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105–2125.
- Meho, L. I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: a comparison of Scopus and Web of Science. *Journal of the American Society for Information Science and Technology*, 59(11), 1711–1726.
- Moed, H. F. (2006). *Citation analysis in research evaluation* (Vol. 9). Springer Science & Business Media. Retrieved from <https://www.google.com/books?hl=el&lr=&id=D9SaJ6awy4gC&oi=fnd&pg=PR9&dq=Citation+analysis+in+research+evaluation&ots=FGmUKy4OmY&sig=KNfS4Drm--Wku-EOTSW3aAvPPSE>
- Moed, H., De Bruin, R., & Van Leeuwen, T. H. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics*, 33(3), 381–422.
- Nederhof, A. J., & Van Raan, A. F. (1993). A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy*, 22(4), 353–368.
- Nordstrom, L. (1990). 'Bradford's law' and the relationship between ecology and biogeography. *Scientometrics*, 18(3-4), 193–203.

- Otlet, P. (1920). *L'Organisation internationale de la bibliographie et de la documentation*. Inst. Internat. de Bibliogr.
- PubMed Overview (2006, June 30, 2006). Retrieved Sept 20th, 2006, from <http://www.ncbi.nlm.nih.gov/entrez/query/static/overview.html>
- Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica Scandinavica*, 69(3), 224–229.
- Sentinel Event Glossary of Terms. (2006). Retrieved Sept 27th, 2006, from http://www.jointcommission.org/SentinelEvents/se_glossary.htm
- Simkin, M. V., & Roychowdhury, V. P. (2012). Theory of citing. In *Handbook of Optimization in Complex Networks* (pp. 463–505). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4614-0754-6_16
- Thomson Scientific: Citation Products. Retrieved April 6th, 2006, from <http://scientific.thomson.com/products/categories/citation/>
- Van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., & Van Raan, A. F. (2003). The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57(2), 257–280.
- Van Raan, A. F. (1998). *Special Topic Issue: Science and Technology Indicators*. American Society for Information Science (ASIS).
- Visser, J. A. (2003). AMH signaling: from receptor to target gene. *Molecular and Cellular Endocrinology*, 211(1), 65–73.
- What's the Difference Between MEDLINE and PubMed? Fact Sheet. (2006, 21 December 2004). Retrieved Sept 20th, 2006, from http://www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html
- Wyllis, R. E. (1981). Empirical and theoretical bases of Zipf's law. *Library Trends*, 30(1), 53–64.
- Zipf, G. K. (1949). Human behavior and the principle of least effort. Retrieved from <http://psycnet.apa.org/psycinfo/1950-00412-000>
- Σαχίνη, Ε., Μάλλιου Ν., Χούσος Ν. (2014). *Βιβλιομετρική Ανάλυση: Μεθοδολογική Προσέγγιση* ΕΚΤ, Εθνικό Κέντρο Τεκμηρίωσης 2^η Έκδοση.

Κεφάλαιο 4

Περίληψη

Η χρήση σε τεχνικές βιβλιομετρίας, κανόνες και φόρμουλες σε ένα "καθαρό" περιβάλλον – πληροφοριών οργανωμένων σε πληροφοριακά συστήματα και πώς οι ποσοτικές μέθοδοι που είχαν σχεδιαστεί για τη βιβλιομετρική ανάλυση των βιβλιογραφικών αναφορών των άρθρων, που δημοσιεύονται σε επιστημονικά περιοδικά, μπορούν να εφαρμοστούν και στο Web, με τη χρήση μηχανών αναζήτησης για την παροχή δεδομένων.

ΕΙΣΑΓΩΓΙΚΑ

Το διαδίκτυο αποτελεί στις μέρες μας μια πολύ πλούσια πηγή πληροφόρησης. Απαρτίζεται, στην ουσία, από διασυνδεδεμένες μεταξύ τους ηλεκτρονικές σελίδες, οι οποίες προστίθενται διαρκώς σ' αυτό, με έναν τρόπο αρκετά ακαθόριστο και άναρχο. Έτσι η ποσότητα της πληροφορίας που περιέχεται σ' αυτό αυξάνεται συνεχώς και με πολύ γρήγορους ρυθμούς.

Πολλοί είναι εκείνοι που συμμετέχουν στη διαμόρφωση του διαδικτύου, καθώς ο καθένας μπορεί, πλέον, να φτιάξει μια ιστοσελίδα εύκολα και χωρίς να ακολουθήσει κάποιο συγκεκριμένο πρότυπο. Έτσι, άτομα με διαφορετικό μορφωτικό επίπεδο, κουλτούρα, γλώσσα και πολλές φορές διαφορετικό σκοπό, καλό ή δόλιο, δημιουργούν συνεχώς καινούριο υλικό, εμπλουτίζοντας και αλλάζοντας την εικόνα του. Επίσης, συνεχώς, αυξανόμενος είναι και ο αριθμός των καινούριων χρηστών του διαδικτύου, οι οποίοι αναζητούν τη σωστή πληροφορία για το θέμα που τους ενδιαφέρει.

Για να εκτιμήσουμε κάπως το μέγεθος του διαδικτύου, αρκεί να λάβουμε υπόψη το ότι, ήδη, από το 1997 αριθμούσε περί τις 150.000.000 ιστοσελίδες. Επίσης, το 2000, ο αριθμός των ιστοσελίδων έφτασε περίπου τα 300.000.000, ενώ προστίθεντο κάθε μέρα περί το 1.000.000 ιστοσελίδες. Το 2002, ακολούθως, ο αριθμός έφτασε τα 8,9 δις μόνο στο δημόσια προσβάσιμο διαδίκτυο. Το 2008 η Google ανέφερε ότι η νέα καταμέτρηση έδειχνε πως υπήρχαν 1.000.000.000.000 ιστοσελίδες! Σήμερα ο αριθμός καταχωρήσεων ιστοσελίδων έχει φτάσει τα 4,5 εκατομμύρια κάθε μήνα. Είναι, λοιπόν, κατανοητή η τεράστια ποσότητα πληροφορίας που βρίσκεται στον ιστό σήμερα. Αυτό, ίσως, μπορεί να θεωρηθεί και το κυριότερο προτέρημά του, διότι ο καθένας μπορεί πλέον να βρει κάτι σχετικό με αυτό που ψάχνει (αν και εκτιμάται ότι το 99% της πληροφορίας που βρίσκεται στον ιστό, δεν ενδιαφέρει το 99% περίπου των ανθρώπων!).

Από την άλλη μεριά όμως, λόγω αυτής της ραγδαίας ανάπτυξης, το διαδίκτυο απέκτησε γρήγορα μια φαινομενικά άναρχη δομή. Έτσι, σε μια προσπάθεια αναζήτησης κάποιας συγκεκριμένης πληροφορίας, ο χρήστης δυσκολευόταν στο να ανακαλύψει τις ιστοσελίδες που ταίριαζαν στο αντικείμενο που μελετούσε. Αυτό το τελευταίο αποτελεί και το κύριο πρόβλημα σε γενικές επερωτήσεις (broad-topic queries), γνωστό και ως «πρόβλημα της αφθονίας», (abundance problem), σύμφωνα με το οποίο το πλήθος των σελίδων που μπορούν να επιστραφούν σε μια επερώτηση, ως σχετικές, είναι μακράν πολύ μεγάλο, για να το διαχειριστεί κάποιος.

Ένα άλλο βασικό πρόβλημα στην αναζήτηση στο διαδίκτυο είναι και αυτό που αφορά αρκετά συγκεκριμένες και στοχευμένες επερωτήσεις. Εδώ υπάρχει αυτό που αποκαλείται ως το «πρόβλημα της σπανιότητας» (scarcity problem), δηλαδή, σ' αυτήν την περίπτωση, υπάρχουν λίγες ιστοσελίδες που περιέχουν την επιθυμητή πληροφορία και είναι συχνά δύσκολο να καθοριστεί η ταυτότητά τους.

Αυτές κυρίως οι δυσκολίες γέννησαν την ανάγκη να δημιουργηθούν κατάλληλα εργαλεία, τα οποία θα μπορούσαν να ερευνούν συνεχώς το διαδίκτυο και σε μια ενδεχόμενη ερώτηση κάποιου χρήστη για ένα συγκεκριμένο θέμα, θα ήταν σε θέση να του παρέχουν τις σωστές ιστοσελίδες ή, εν γένει, τη σωστή πληροφόρηση. Αυτά τα εργαλεία είναι οι μηχανές αναζήτησης (search engines).

Για να κατανοήσουμε τον τρόπο με τον οποίο δουλεύουν οι μηχανές αναζήτησης, θα πρέπει πρώτα να προσεγγίσουμε επιγραμματικά δύο ζητήματα, την ίδια τη δομή του διαδικτύου και τον τρόπο με τον οποίο κάνουν τις αναζητήσεις τους οι χρήστες.

Κατά πρώτον, στο διαδίκτυο προστίθενται συνεχώς και με γρήγορους ρυθμούς καινούριες ιστοσελίδες, αλλάζοντας έτσι, αρκετά συχνά, τη γενική εικόνα του. Ακόμα, οι ήδη υπάρχουσες ιστοσελίδες ενδέχεται κι αυτές να ανανεώνουν τη δομή τους. Επίσης η πληροφορία που παρουσιάζεται μέσω των ιστοσελίδων είναι αρκετά συγκεχυμένη, καθώς για παράδειγμα, άλλες σελίδες εντυφούν σχολαστικά σε ένα

θέμα, ενώ άλλες για το ίδιο θέμα, απλώς, κάνουν μια νύξη. Ακόμα, σε πολλές περιπτώσεις, η πληροφόρηση που παρέχεται μπορεί να υπηρετεί συγκεκριμένους σκοπούς διαφήμισης, οπότε να μην είναι εντελώς έγκυρη.

Από την άλλη μεριά, η συμπεριφορά του τυχαίου χρήστη όσον αφορά στις επερωτήσεις (queries) προς τις μηχανές αναζήτησης, δεν είναι ίδια με αυτή προς μια κλασική βάση δεδομένων. Εκεί ο χρήστης πρέπει να δομήσει μια πιο σύνθετη επερώτηση, για να έχει αποδοτικά αποτελέσματα. Στις μηχανές αναζήτησης, όμως, λόγω και του τεράστιου πλήθους των χρηστών και της μικρής ή μεγαλύτερης άγνοιας που έχουν πολλοί γύρω από τη χρήση των υπολογιστών, τα ερωτήματα είναι κατά κύριο λόγο μικρά, αρκετά ασαφή και γενικόλογα, αποτελούμενα συνήθως από τρεις τέσσερις λέξεις και χωρίς, τις περισσότερες φορές, να σχηματίζουν λογικές προτάσεις.

Η φύση, όμως, της ίδιας της αναζήτησης μπορεί να εγείρει διάφορα προβλήματα. Ένα παράδειγμα είναι η περίπτωση που κάποιος χρήστης θέλει να ψάξει κάτι πολύ συγκεκριμένο, όπως το αν κάποιος φυλλομετρητής (browser) υποστηρίζει ένα εξεζητημένο πρόγραμμα. Το πρόβλημα που υφίσταται εδώ είναι ότι δεν υπάρχουν πολλές ιστοσελίδες που περιέχουν πληροφορία σχετική με τη συγκεκριμένη ερώτηση και έτσι, είναι δύσκολος ο προσδιορισμός της ταυτότητάς τους ή της τοποθεσίας τους στον ιστό.

Μία δεύτερη περίπτωση είναι όταν ένας χρήστης αναζητά κάτι γενικό, όπως πληροφορίες για μια γλώσσα προγραμματισμού. Εδώ, δεν υπάρχει το πρόβλημα των λίγων ιστοσελίδων, αλλά αντίθετα, ο αριθμός αυτών που μπορούν να θεωρηθούν ως σχετικές είναι αρκετά μεγάλος, για να μπορέσει να τις διαχειριστεί ο χρήστης. Έτσι, θα πρέπει να βρεθούν τρόποι να επιλέγονται, ανάμεσα από μια τεράστια συλλογή από σχετικές ιστοσελίδες, εκείνες που θεωρούνται ως οι πιο αυθεντικές και έγκυρες.

Από νωρίς, λοιπόν, οι σχεδιαστές των μηχανών αναζήτησης είχαν να αντιμετωπίσουν και να ξεπεράσουν αυτά τα προβλήματα, προκειμένου να παρέχουν στους χρήστες εργαλεία που θα τους βοηθούσαν να χρησιμοποιούν αποτελεσματικά το διαδίκτυο. Κατά την πάροδο των ετών παρουσιάστηκαν τα εργαλεία και οι αλγόριθμοι, πολλά από τα οποία οδήγησαν σε επιτυχημένα εμπορικά πακέτα.

Από το 1996 ξεκίνησαν οι προσπάθειες εξερεύνησης της φύσης και των περιεχομένων του Web, τη δομή των συνδέσεων και των μηχανών αναζήτησης, με την εφαρμογή νέων μεθοδολογιών στα περιεχόμενα των κειμένων. Το πεδίο αυτό άρχισε να αναπτύσσεται, όταν διαπιστώθηκε ότι μπορούμε να χρησιμοποιήσουμε τεχνικές βιβλιομετρίας, κανόνες και φόρμουλες σε ένα "καθαρό" περιβάλλον πληροφοριών οργανωμένων σε πληροφοριακά συστήματα και πως οι ποσοτικές μέθοδοι που είχαν σχεδιαστεί για τη βιβλιομετρική ανάλυση των βιβλιογραφικών αναφορών των άρθρων, που δημοσιεύονται σε επιστημονικά περιοδικά, μπορούσαν να εφαρμοστούν και στο Web, με τη χρήση εμπορικών μηχανών αναζήτησης, για την παροχή δεδομένων. Αυτό ξεκίνησε αφ' ότου η διεπαφή αναζήτησης της Alta Vista είχε τη δυνατότητα σύνθετων αναζητήσεων με τη χρήση Boolean τελεστών και τη δυνατότητα μελέτης των URLs και των ιδιοτήτων των συνδέσεων. (Thelwall, et. al, 2005). Η συλλογή δεδομένων από το Web είναι πλέον πολύ διαδεδομένη και δυναμική. Για τη συλλογή δεδομένων που αφορούν τεκμήρια του Web, δικτυακούς τόπους ή τη δομή μερών του Web χρησιμοποιούνται υπάρχοντα εργαλεία ανάκτησης πληροφοριών, που είναι κυρίως οι μηχανές αναζήτησης, οι οποίες βέβαια απέχουν πολύ από το να είναι τέλειες, όπως ήδη προαναφέρθηκε.

• Μοντέλα σύνθετων αναζητήσεων του ιστού

Ένα πρώτο μοντέλο, το πιο απλό ίσως απ' όλα, είναι το Boolean μοντέλο. Σύμφωνα μ' αυτό, η μηχανή αναζήτησης διαχειρίζεται τις λέξεις-κλειδιά που θα θέσει ο χρήστης με βάση δύο τιμές, την τιμή true (ή 1), αν η λέξη-ή η σειρά των λέξεων- είναι μέσα σε ιστοσελίδα της συλλογής (έτσι όπως ακριβώς την πληκτρολόγησε ο χρήστης) και την τιμή false (ή 0), αν δεν είναι. Έτσι ανακαλεί τις ιστοσελίδες για τις οποίες ισχύει η τιμή true.

Μια επέκταση αυτού του μοντέλου, για να ξεπεραστεί αυτή η απολυτότητα του «είναι ή δεν είναι», στηρίζεται στη θεωρία των ασαφών συνόλων (fuzzy set theory), σύμφωνα με την οποία επιτρέπεται μερική συμμετοχή ενός στοιχείου στο σύνολο, με βάση κάποια συνάρτηση.

Ένα δεύτερο μοντέλο είναι το μοντέλο διανυσματικού χώρου (vector space model). Σύμφωνα με αυτό, κάθε έγγραφο μιας συλλογής μπορεί να παρασταθεί από ένα n-διάστατο διάνυσμα ενός διανυσματικού χώρου, ο οποίος είναι βαθμονομημένος από n το πλήθος κανονικοποιημένα διανύσματα-όρους (πιο συγκεκριμένα, κάθε έγγραφο μπορεί να γραφεί σαν γραμμικός συνδυασμός των διανυσμάτων που αντιπροσωπεύουν τους όρους). Έτσι, αν αναπαραστήσουμε και την επερώτηση του χρήστη σαν διάνυσμα στον ίδιο διανυσματικό χώρο, τότε μπορούμε να χρησιμοποιήσουμε εργαλεία και έννοιες των διανυσματικών χώρων, όπως η ευκλείδεια απόσταση, το εσωτερικό γινόμενο ή το συνημίτονο της γωνίας δύο διανυσμάτων,

για να μετρήσουμε την «απόσταση» του κάθε διανύσματος-εγγράφου από το διάνυσμα-επερώτηση. Έτσι, όσο πιο «κοντά» είναι ένα διάνυσμα, τόσο πιο σχετικό είναι με την επερώτηση.

Ένα τρίτο μοντέλο είναι το πιθανοθεωρητικό μοντέλο (probabilistic model). Σύμφωνα με αυτό, για να εξεταστεί αν ένα αντικείμενο της συλλογής είναι συναφές με την επερώτηση του χρήστη, υπολογίζονται δύο δεσμευμένες πιθανότητες, της σχετικότητας και της μη-σχετικότητας του εγγράφου με την επερώτηση ($Pr(rel)$ και $Pr(nonrel)$). Δηλαδή, για έναν όρο t , υπολογίζονται αρχικά οι πιθανότητες :

P (ο όρος t υπάρχει στο έγγραφο/το έγγραφο είναι σχετικό)

και

P_0 (ο όρος t υπάρχει στο έγγραφο/το έγγραφο είναι μη-σχετικό)

και ακολούθως με στατιστικά εργαλεία, υπολογίζεται ένα μέτρο ομοιότητας (similarity) του εγγράφου με την επερώτηση q .

Επιπρόσθετα, το μοντέλο κάνει χρήση δύο ακόμα παραμέτρων, a_1 και a_2 , για να αναπαραστήσει την απώλεια σε ακρίβεια από την ανάκτηση ενός μη σχετικού ντοκουμέντου και τη μη ανάκτηση ενός σχετικού ντοκουμέντου αντίστοιχα.

Παρ' όλη όμως την επιτυχή λειτουργία τους, αυτοί οι αλγόριθμοι είχαν και ένα βασικό μειονέκτημα, που αφορούσε ακριβώς στον ίδιο το σχεδιασμό τους. Το να ψάχνει μια μηχανή αναζήτησης αποκλειστικά στο κείμενο μιας ιστοσελίδας, για να αποφανθεί αν αυτή είναι σχετική με μια επερώτηση του χρήστη, ενέχει και κάποια προβλήματα ή συγχύσεις.

Μια περίπτωση είναι λέξεις που ενώ συλλαβίζονται το ίδιο, εν τούτοις έχουν εντελώς διαφορετικό νόημα. Αν για παράδειγμα στην επερώτηση μας συμπεριλάβουμε τη λέξη «σκληρός», μιλώντας για δίσκο ηλεκτρονικού υπολογιστή, η μηχανή αναζήτησης ενδέχεται να μας επιστρέψει στα αποτελέσματά της και σελίδες άσχετες με πληροφορική, οι οποίες όμως έχουν τη λέξη «σκληρός» (εκτός κι αν σκεφτούμε από την αρχή να βάλουμε όλη την έκφραση «σκληρός δίσκος» μέσα σε ομοιωματικά, οπότε η περίπτωση λάθους ελαχιστοποιείται). Ή ακόμα, ψάχνοντας για ιστορικά πρόσωπα και πληκτρολογώντας το όνομά τους στη μηχανή, εκείνη μπορεί να μας επιστρέψει κάλλιστα σελίδες με ευρετήρια οδών, μόνο και μόνο επειδή το όνομα αυτού του προσώπου θα έχει δοθεί σε δρόμο μιας πόλης.

Μια άλλη περίπτωση είναι οι απώλειες που έχουμε από συνώνυμα λέξεων. Αν, για παράδειγμα, ψάχνουμε σελίδες με τη λέξη «ανάγωγος» (όχι με τη μαθηματική έννοια, αλλά με τη συνήθη έννοια της μη κόσμιας συμπεριφοράς) ενδέχεται να χάσουμε σελίδες που περιέχουν τη λέξη «αγενής», παρ' όλο που και αυτή η λέξη έχει την ίδια σημασία και συνεπώς, θα μπορούσε να καλύψει την πληροφοριακή μας ανάγκη. Ένα άλλο, συναφές παράδειγμα, που ίσως απαντάται συχνά, είναι και αυτό με τα συνώνυμα «δημοφιλής-διάσημος».

Μια τρίτη περίπτωση, κατά την οποία μπορούμε να έχουμε σημαντική απώλεια πληροφορίας αφορά πολλές ιστοσελίδες που αναφέρονται σε κάποιο συγκεκριμένο θέμα, ενδέχεται, ωστόσο, να μην έχουν μέσα στο κείμενό τους την ονομασία του θέματος αυτού. Για παράδειγμα, οι κεντρικές σελίδες πολλών βιομηχανιών αυτοκινήτων δεν έχουν στο κείμενό τους τη λέξη «αυτοκινητοβιομηχανία», παρότι εντάσσονται εξολοκλήρου σ' αυτόν τον τομέα. Όπως επίσης το ίδιο μπορεί να συμβεί αν για λέξεις κλειδιά βάλουμε τους όρους «μηχανές αναζήτησης». Δεν είναι σίγουρο ότι οι κεντρικές σελίδες των διασημότερων μηχανών αναζήτησης θα εμφανιστούν στις πρώτες θέσεις.

Αυτές οι δυσκολίες, που δημιουργούνταν από τη φιλοσοφία κατασκευής των αλγορίθμων, έδωσαν το έναυσμα για τη μελέτη του διαδικτύου από μια διαφορετική σκοπιά. Έτσι λοιπόν άρχισε να μοντελοποιείται το διαδίκτυο με τη βοήθεια της θεωρίας γραφημάτων.

Σύμφωνα με τη θεωρία γραφημάτων, το διαδίκτυο παρίσταται σαν ένα κατευθυνόμενο γράφημα, στο οποίο οι κόμβοι αποτελούν τις ιστοσελίδες, ενώ οι ακμές αποτελούν τις διασυνδέσεις (links) μεταξύ τους. Έτσι, αν μία ιστοσελίδα συνδέεται με ένα link με μίαν άλλη, τότε στο γράφημα θα υπάρχει ένα βέλος (ακμή) από τον ένα κόμβο, που αναπαριστά τη μια ιστοσελίδα, προς τον άλλο κόμβο, που παριστά τη δεύτερη ιστοσελίδα.

Το βασικό πλεονέκτημα που έχει αυτή η θεώρηση είναι ότι μπορούμε να εξάγουμε συμπεράσματα για τη γενική σημαντικότητα μιας ιστοσελίδας, ανεξαρτήτως των λέξεων-κλειδιών που πληκτρολογούμε κάθε φορά σε μια μηχανή αναζήτησης. Άλλωστε, μια επιτυχής αναζήτηση είναι αυτή που φέρνει στις πρώτες θέσεις, όχι μόνο τις πιο σχετικές ιστοσελίδες με αυτό που ψάχνουμε, αλλά τις σχετικές ιστοσελίδες με τη μεγαλύτερη ποιότητα και σημαντικότητα.

Αυτός ο προσδιορισμός της γενικής σημαντικότητας μιας ιστοσελίδας μπορεί να γίνει βάσει μιας απλής, αλλά ουσιαστικής διαπίστωσης, η οποία αξιοποιήθηκε αρχικά στην εργασία του Kleinberg (1999) και μετέπειτα των Page και Brin (1999) (οι οποίοι είναι και οι «πατέρες» του google) και την οποία αρκετοί

παραλληλίζουν με τα citations που γίνονται στις ακαδημαϊκές εργασίες (αν και διατυπώνεται ότι υπάρχουν διαφορές). Σύμφωνα με αυτή, ο δημιουργός μιας ιστοσελίδας, όταν φτιάχνει ένα link προς μια δεύτερη ιστοσελίδα (και δε μιλάμε για τα links που υπάρχουν καθαρά για λόγους πλοήγησης μέσα στην ιστοσελίδα, όπως τα “back”, “home” κ.τ.λ.), στην ουσία εκδηλώνει την προτίμησή του προς την ιστοσελίδα που δείχνει. Έτσι, λοιπόν, αυτή αποκτά μια κάποια σημαντικότητα. Επομένως, όσο περισσότερα links δείχνουν προς μια ιστοσελίδα, τόσο πιο αυξημένη θεωρείται η σημαντικότητα αυτής.

Εδώ, βέβαια, θα πρέπει να διατυπωθούν και κάποιες παγίδες που έχει αυτή η θεώρηση. Κατά πρώτον, όπως ελέγχθη και στην προηγούμενη παράγραφο, ένα link ενδέχεται να μη δείχνει προς μια άλλη ιστοσελίδα, αλλά να υπάρχει καθαρά για λόγους διευκόλυνσης της πλοήγησης μέσα στην ίδια ιστοσελίδα. Έτσι, η αξία που έχει αυτός ο σύνδεσμος είναι μακράν μικρότερη της αξίας ενός συνδέσμου προς διαφορετική ιστοσελίδα.

Κατά δεύτερον, δεν πρέπει να συγχέεται η έννοια της σημαντικότητας που έχει μια ιστοσελίδα, ως προς κάποιο θέμα, με την έννοια της δημοφιλίας μιας ιστοσελίδας. Ένα link ενδέχεται να δείχνει μια ιστοσελίδα, όχι επειδή θεωρείται σημαντική για κάποιο θέμα, αλλά επειδή είναι δημοφιλής γενικά.

Μια άλλη παγίδα είναι και το γεγονός ότι ένα link μπορεί να δείχνει κάπου καθαρά για εμπορικούς ή διαφημιστικούς σκοπούς, χωρίς να σημαίνει ότι αυτό που δείχνει είναι σημαντικό. Η ακόμα, μπορεί ο δημιουργός μιας ιστοσελίδας να καταδεικνύει κάποια άλλη, μόνο ως παράδειγμα προς αποφυγή και όχι επειδή τη θεωρεί σημαντική. Το link, όμως, προς αυτή τη σελίδα υπάρχει.

Αντιμετωπίζοντας αυτές τις δυσκολίες, ο Kleinberg οδηγήθηκε στο να διαχωρίσει τις ιστοσελίδες σε δύο κατηγορίες, τα authorities και τα hubs. Η πρώτη κατηγορία αφορά ιστοσελίδες, οι οποίες κοντολογίς δέχονται πολλά links από άλλες (inlinks), οπότε θεωρούνται ως σημαντικές, έχουν δηλαδή κύρος. Η δεύτερη κατηγορία αναφέρεται στις ιστοσελίδες εκείνες που «μοιράζονται» την προτίμησή τους σε πολλές άλλες σελίδες, έχουν δηλαδή πολλά links προς άλλες σελίδες (outlinks). Στην εργασία αυτή, ο Kleinberg έδειξε ότι υπάρχει μια σχετική ισορροπία μεταξύ αυτών των δύο τύπων ιστοσελίδων και παρουσίασε έναν αλγόριθμο που βρίσκει αυτούς τους δύο τύπους, αξιοποιώντας, όπως ελέγχθη παραπάνω, τη δομή των διασυνδέσεων.

Ο αλγόριθμος αυτός, όμως, βασίζεται σε ένα αρχικό σύνολο σελίδων, οι οποίες αποκτώνται από μια βασισμένη στο κείμενο (text-based) μηχανή αναζήτησης, για κάθε ερώτημα του χρήστη. Μπορούμε να πούμε, λοιπόν, ότι η έννοια «σημαντικότητα» παρουσιάζεται συναρτήσει κάθε φορά του ερωτήματος του χρήστη.

Ένα βήμα παραπέρα προχώρησαν τη θεώρηση αυτή οι Page και Brin, παρουσιάζοντας τον αλγόριθμο pagerank, τον πυρήνα της διάσημης μηχανής αναζήτησης google, η οποία τα τελευταία χρόνια βρίσκεται στην πρώτη θέση στις προτιμήσεις των χρηστών, με πολύ μεγάλη διαφορά από τις άλλες μηχανές αναζήτησης. Χαρακτηριστικό παράδειγμα, τον Αύγουστο του 2009 και μόνο στην Αμερική, το ποσοστό προτίμησης στις αναζητήσεις έφτασε το 64,6%, με τη δεύτερη εταιρεία στη σειρά να αποσπά μόλις το 16%.

Σύμφωνα με τους Page και Brin, μια σελίδα δεν χρειάζεται να έχει μόνο πολλά links που δείχνουν προς αυτή, για να είναι σημαντική. Αρκεί να δεικτοδοτείται και από μία μόνο σελίδα, η οποία είναι ήδη σημαντική, όπως για παράδειγμα μια κυβερνητική ιστοσελίδα ή οι κεντρικές ιστοσελίδες πολύ μεγάλων και γνωστών εταιρειών. Έτσι, λοιπόν, μία σελίδα που ενδέχεται να λαμβάνει links από εκατό (τυχαίο νούμερο) μικρές σελίδες, θα αποκτήσει, βεβαίως, μια κάποια σημαντικότητα. Μια άλλη, όμως, που δεικτοδοτείται από το επίσημο site του ελληνικού κοινοβουλίου, για παράδειγμα, και μόνον από αυτό, μπορεί να λάβει μεγαλύτερο βαθμό σημαντικότητας.

Η διαδικασία που ακολουθεί τελικά το google για να εμφανίσει στον χρήστη τις σελίδες που τον ενδιαφέρουν (άσχετα από ποια μαθηματική σκοπιά έχει θεωρηθεί ο βασικός αλγόριθμος) αποτελείται από δύο βήματα, όπως αναφέρεται και σε σχετική σελίδα της εταιρείας [<http://www.google.com/intl/el/corporate/tech.html>]. Πρώτα εξετάζει ολόκληρη τη δομή των συνδέσεων του παγκόσμιου ιστού, για να καθορίσει ποιες σελίδες είναι πιο σημαντικές (η δουλειά του pagerank) και μετά διενεργεί μια ανάλυση αντιστοίχισης υπερκειμένου, για να καθορίσει ποιες σελίδες είναι συναφείς με την εκάστοτε αναζήτηση. Έτσι, σύμφωνα πάντα με την εταιρεία, συνδυάζεται η συνολική σημαντικότητα με τη συνάφεια για το συγκεκριμένο ερώτημα του χρήστη.

Μια μεταγενέστερη προσπάθεια για να δοθούν στον χρήστη ποιοτικότερα αποτελέσματα στις αναζητήσεις του, είναι και ο TrustRank, ένας αλγόριθμος που δεν προτείνει μια εντελώς καινούρια θεώρηση του διαδικτύου, αλλά απλώς σε ένα αρχικό σύνολο σελίδων (το οποίο μπορεί να εξαχθεί από μια άλλη μηχανή αναζήτησης) υπολογίζει, με κάποια εργαλεία, την πιθανότητα για κάθε σελίδα να είναι είτε αξιόπιστη, είτε κακόβουλη (spam). Ξεκινώντας από αυτό το αρχικό σύνολο και βασιζόμενος στην πρωταρχική ιδέα ότι οι καλές σελίδες σπάνια θα δείχνουν σε κακές, επεκτείνει το σύνολο σε ένα μεγαλύτερο, το οποίο περιέχει σελίδες που με μεγάλη πιθανότητα είναι αξιόπιστες.

Ο αλγόριθμος αυτός γεννήθηκε, σύμφωνα πάντα με την αντίστοιχη ακαδημαϊκή εργασία, από το γεγονός ότι ο pagerank δεν περιέκλειε καμία πληροφορία που να αφορούσε την ποιότητα της κάθε σελίδας, αλλά ούτε και «τιμωρούσε» κακές σελίδες.

4.1. Ανοικτή Πρόσβαση (Open Access)

Τα πανεπιστήμια δεν μπορούν να δώσουν στους υποτρόφους πρόσβαση σε κάθε ακαδημαϊκό περιοδικό στον κόσμο. Οι συνδρομές είναι αρκετά ακριβές, ιδιαίτερα σε περιόδους λιτότητας, πράγμα που σημαίνει ότι οι ερευνητικές βιβλιοθήκες πρέπει να κάνουν μια επιλογή. Είναι δίκαιο να υποθέσουμε ότι το επιστημονικό σύστημα, στο σύνολό του, θα ήταν πολύ πιο αποτελεσματικό, αν οι ερευνητές είχαν πρόσβαση σε κάθε έρευνα. Οι υπέρμαχοι της ελεύθερης πρόσβασης έχουν αναπτύξει μια σειρά από συστήματα για την παροχή τους. Η ιδέα είναι να δοθεί στους μελετητές δωρεάν πρόσβαση σε ένα συγκεκριμένο σύνολο δημοσιεύσεων και, ενδεχομένως, να επεκταθεί η πρόσβαση σε κάθε ερευνητικό έργο.

Η ιδέα αυτή έχει ελκύσει την προσοχή πολλών μελετητών, όπως φαίνεται από τον αριθμό των σχετικών δημοσιεύσεων στα επιστημονικά περιοδικά. Αξίζει, επίσης, να ληφθεί υπόψη από τους ιθύνοντες, γιατί θα μπορούσε να έχει σημαντική επίδραση στο σύστημα της επιστήμης, στην παρούσα κατάστασή του. Η Ευρωπαϊκή Επιτροπή ξεκίνησε πρόσφατα μια μελέτη σχετικά με την ανοικτή πρόσβαση και τη βελτιστοποίηση του συστήματος δημοσίευσης της έρευνας, στο πλαίσιο των προσπαθειών της για τη δημιουργία ενός Ευρωπαϊκού Χώρου Έρευνας και την ενίσχυση του διεθνούς προφίλ της ευρωπαϊκής έρευνας (CORDIS 2004).

• Ποικιλία των συστημάτων

Η ανοικτή πρόσβαση μπορεί να πάρει διάφορες μορφές. Σύμφωνα με μια πρόσφατη πρωτοβουλία της Public Library of Science (PLOS), οι ερευνητές πρέπει να χρησιμοποιούν τη δική τους έρευνα, επενδύοντας στη δημοσίευση της στο ακαδημαϊκό περιοδικό *Public Library of Science (PLOS) Biology*, το οποίο διατίθεται δωρεάν. Ένα παρόμοιο μοντέλο είχε ήδη ξεκινήσει από το Περιοδικό κλινικών Ερευνών (Journal of Clinical Investigation -JCI), το οποίο είναι διαθέσιμο στο διαδίκτυο δωρεάν, από το 1996 (Varki, 2003). Οι Harnad και άλλοι (2004a) αναφέρουν τις δωρεάν εκδόσεις, που δραστηριοποιούνται υπό αυτήν την αρχή, ως τα "χρυσά" περιοδικά. Μια άλλη προσέγγιση των περιοδικών είναι να συνεχίσουν να χρεώνουν για τις συνδρομές, αλλά να επιτρέπουν την ανάρτηση του άρθρου στο Διαδίκτυο. Οι Harnad και άλλοι (2004a) αναφέρονται σε αυτά ως "πράσινα" περιοδικά. Αυτή η προσέγγιση εξαρτάται σε κάποιο βαθμό από την ειλικρίνεια και την καλή θέληση των ερευνητικών βιβλιοθηκών, η οποία εναπόκειται στη συνέχιση της συνδρομής σε ακαδημαϊκά περιοδικά, ακόμη και αν τα άρθρα τους έχουν αναρτηθεί.

Μια μέθοδος που φαίνεται πιο ελπιδοφόρα είναι η αυτό-αρχειοθέτηση ή το κατέβασμα των άρθρων σε αρχεία των ιδρυμάτων, τα οποία μπορούν να δεσμεύονται από το πρωτόκολλο Open Archive Initiative (OAI) (2004a). Ειδικές μηχανές έρευνας μπορούν να σαρώσουν τα αρχεία, σχηματίζοντας ένα ενιαίο αρχείο. Επιπλέον τα αρχεία μπορούν να παρέχουν αναφορές ή λήψεις κι έτσι να προστίθενται άλλες λειτουργίες, που είναι ιδιαίτερα χρήσιμες για τη βιβλιομετρία. Σύμφωνα με πληροφορίες, 100 πανεπιστήμια έχουν, τη στιγμή αυτή, τα δικά τους αρχεία (2004b).

• Τρέχουσα κατάσταση και επιπτώσεις

Η σχετικά πρόσφατη έννοια της ανοικτής πρόσβασης έχει αποκτήσει γρήγορα αντίκτυπο στην επικαιρότητα. Σύμφωνα με τους Harnad κ.α. (2004a), το 80% των ακαδημαϊκών περιοδικών είναι "πράσινο". Αυτό επιτρέπει στους συγγραφείς να αρχειοθετούν τα άρθρα τους που υπόκεινται σε περιορισμούς, οι οποίοι ποικίλουν σε αριθμό και έκταση. Ένα άλλο 5% των περιοδικών είναι "χρυσό": η πρόσβαση σε αυτά είναι πλήρης και δωρεάν. Το (PLOS Biology είναι ένα παράδειγμα). Το Science-Metrix χρησιμοποίησε τον κατάλογο των περιοδικών ανοικτής πρόσβασης, (Directory of Open Access Journals-DOAJ) για να εξακριβωθεί ο αριθμός των ανοικτής πρόσβασης ("χρυσός") περιοδικών στις βάσεις δεδομένων και βρέθηκαν 529 στις Social Sciences and Humanities (SSH) βάσεις, σε σύγκριση με τις 436 βάσεις στο Natural Science and Engineering (NSE).

Τι αντίκτυπο έχει η ανοικτή πρόσβαση στη χρήση της έρευνας; Υπάρχει μια απόκλιση των ευρημάτων και των απόψεων για το θέμα αυτό. Μια σειρά από μελέτες έχουν συγκρίνει τον αριθμό των αναφορών που κατανέμονται σε ανοικτής πρόσβασης άρθρα με τον αριθμό που έλαβαν άρθρα μη ανοικτής πρόσβασης. Ο Lawrence (2001) κάλυψε 119.924 άρθρα του συνεδρίου στην επιστήμη των υπολογιστών που

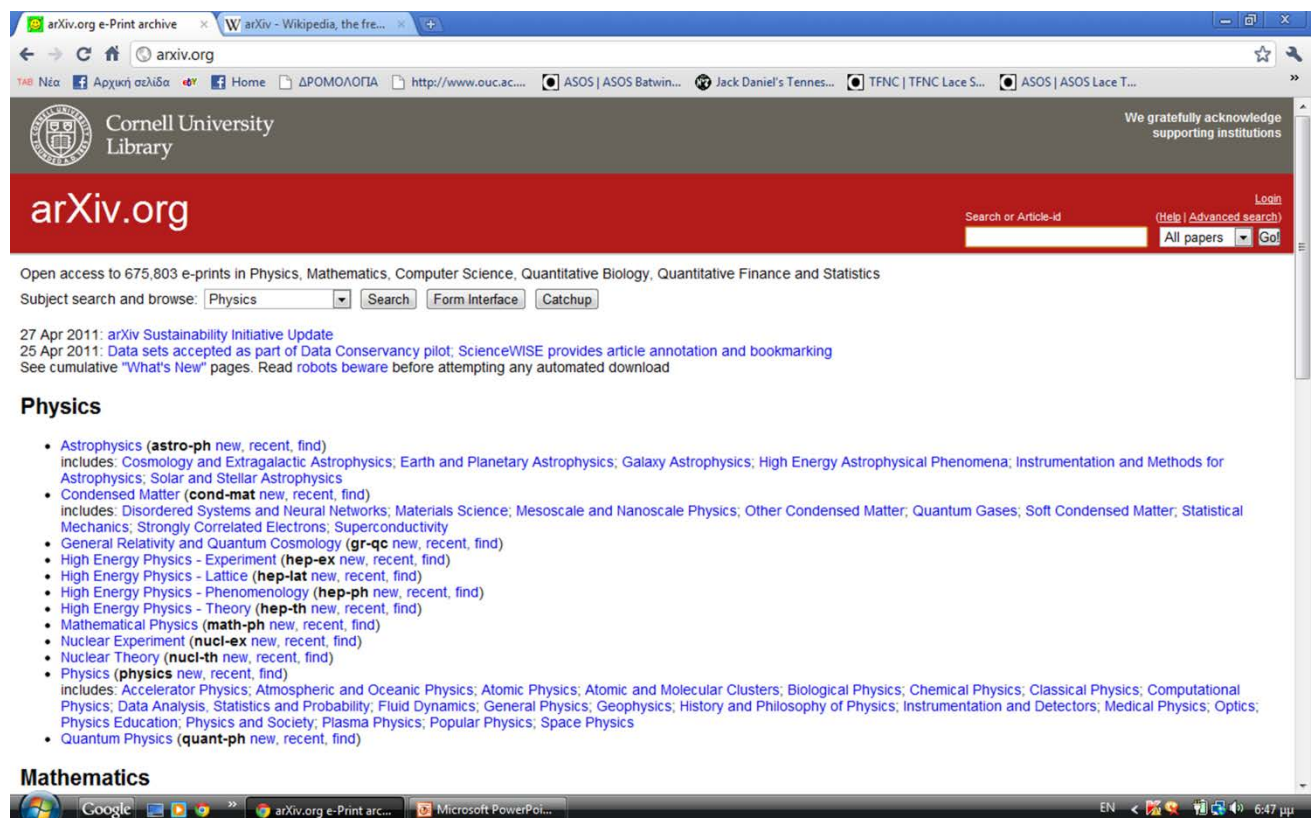
παρήχθησαν από το 1989 έως το 2000. Ο μέσος αριθμός των αναφορών που έχει αποδοθεί σε άρθρα τα οποία δεν είναι άμεσα διαθέσιμα (on line) ήταν 2,74, σε σύγκριση με το 7,03 για τα άμεσα διαθέσιμα άρθρα στο διαδίκτυο. Οι Harnad et al. (2004a) αναφέρουν ότι οι Kurtz (2004) και Odlyzko (2002) έκαναν παρόμοιες εκτιμήσεις για την αστρονομία και τα μαθηματικά. Μια κοινοπραξία που αποτελείται από το παρατηρητήριο επιστήμης και τεχνικής 'Observatoire des sciences et des technique' (OST) του Πανεπιστημίου του Southampton και το Πανεπιστήμιο του Oldenburg ανέλαβε την εκπόνηση μελέτης σχετικής με το θέμα και τα προκαταρκτικά ευρήματα δείχνουν ότι τα άρθρα της φυσικής που διατίθενται δωρεάν στο ηλεκτρονικό αρχείο ArXiv παρατίθενται 2,5 έως 5 φορές συχνότερα (Brody et al. 2004; Harnad et al., 2004a). Η Thomson ISI εξέτασε το ζήτημα σε πρώτο επίπεδο και έχει παράγει τα δικά της στατιστικά στοιχεία (Testa and McVeigh, 2004). Οι Testa και McVeigh έλεγξαν τους παράγοντες των επιπτώσεων των ακαδημαϊκών περιοδικών στις φυσικές επιστήμες και στη μηχανή (Neuron Specific Enolase – NSE) και διαπίστωσαν ότι ο συντελεστής για τα 148 περιοδικά ανοικτής πρόσβασης που καλύφθηκαν ήταν γενικά συγκρίσιμος με αυτόν των συνδρομητικών περιοδικών. Οι συγγραφείς καταλήγουν στο συμπέρασμα ότι τα περιοδικά ανοικτής πρόσβασης απολαμβάνουν μόνο ένα μικρό πλεονέκτημα έναντι των ομολόγων τους. Έτσι τα ευρήματά τους αντιστρατεύονται αυτά που περιγράφηκαν νωρίτερα. Η Thomson ISI επισημαίνει, ωστόσο, ότι τα αποτελέσματά της έρευνας τους είναι προκαταρκτικά. Είναι ενδιαφέρον να σημειωθεί ότι ο οργανισμός δίνει μια καλή βαθμολόγηση των επιδόσεων στο PLoS Biology για τα τέσσερα πρώτα του τεύχη.

Είναι σημαντικό να παρακολουθείται η εξέλιξη της προσέγγισης της ανοικτής πρόσβασης. Προς το παρόν, η χρήση της δεν είναι ευρέως διαδεδομένη και οι επιπτώσεις της δεν είναι σαφώς καθορισμένες, αλλά θα μπορούσε να έχει σημαντική επίδραση στην έρευνα των επιστημονικών εκδόσεων. Οι τρόποι με τους οποίους θα μπορούσε να επηρεάσει την εφαρμογή της βιβλιομετρίας παρουσιάζονται παρακάτω (Βλ. 4.2.).

• Βάσεις Δεδομένων Ανοικτής Πρόσβασης

arXiv.org

Το arXiv.org είναι μια μηχανή αναζήτησης επιστημονικών εγγράφων σε ηλεκτρονική μορφή (e-print), που αφορούν θέματα μαθηματικών, φυσικής, επιστήμης ηλεκτρονικών υπολογιστών, κβαντικής βιολογίας, αστρονομίας και στατιστικής. Δημιουργήθηκε το 1991 και λειτουργούσε στο Εθνικό εργαστήριο Los Alamos, όμως αργότερα μεταφέρθηκε στο πανεπιστήμιο Cornell.



Εικόνα 4.1 Ενδεικτική ιστοσελίδα της μηχανής

CiteSeer^x beta

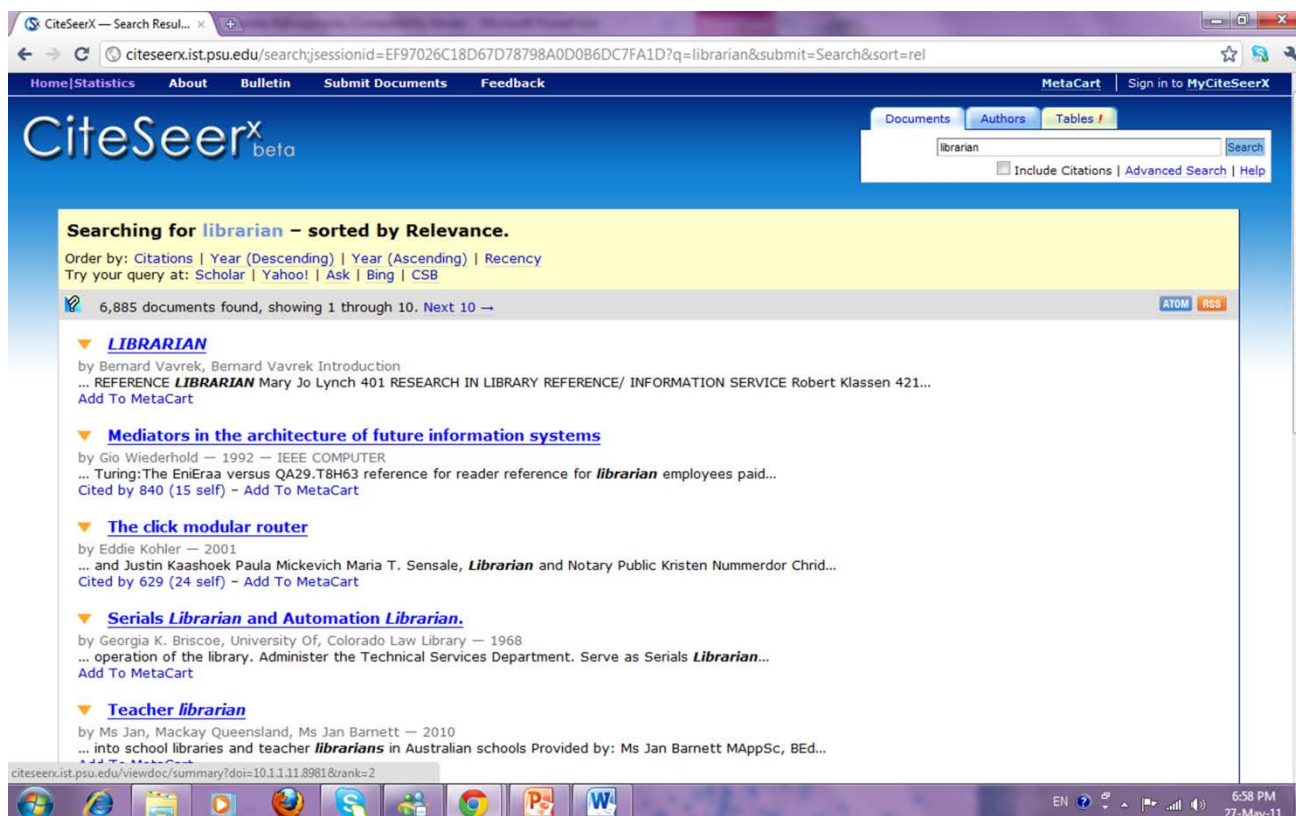
Εικόνα 4 2 Το λογότυπο της μηχανής

Το CiteSeer είναι μια μηχανή αναζήτησης, ψηφιακή βιβλιοθήκη και αποθετήριο επιστημονικών και ακαδημαϊκών αρχείων που επικεντρώνεται στους ηλεκτρονικούς υπολογιστές και στις πηγές πληροφόρησης. Δημιουργήθηκε από τον Δρ. Isaac Council και τον Δρ. C. Lee Giles στα πανεπιστήμια the College of Information Sciences and Technology και Pennsylvania State University. Το CiteSeer βαθμολογήθηκε ως το νούμερο ένα αποθετήριο στον κόσμο το 2010 και εκεί υπάρχουν, μέχρι στιγμής, πάνω από 1,5 εκατομμύρια έγγραφα, με σχεδόν 1,5 εκατομμύρια συγγραφείς και πάνω από 30 εκατομμύρια παραπομπές.



Εικόνα 4 3 Κεντρική ιστοσελίδα μηχανής

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN



Εικόνα 4 4 Ενδεικτική ιστοσελίδα της μηχανής

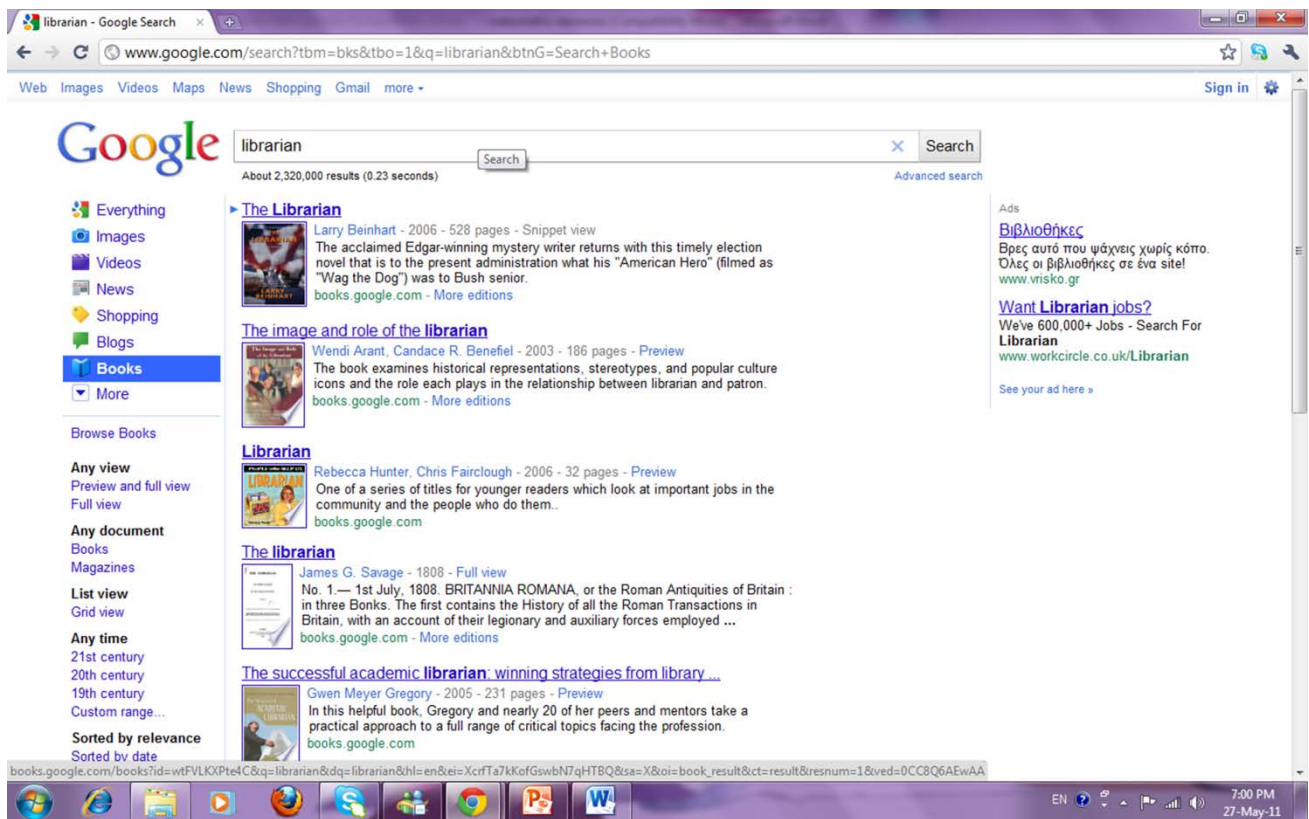


Εικόνα 4 5 Το λογότυπο της μηχανής

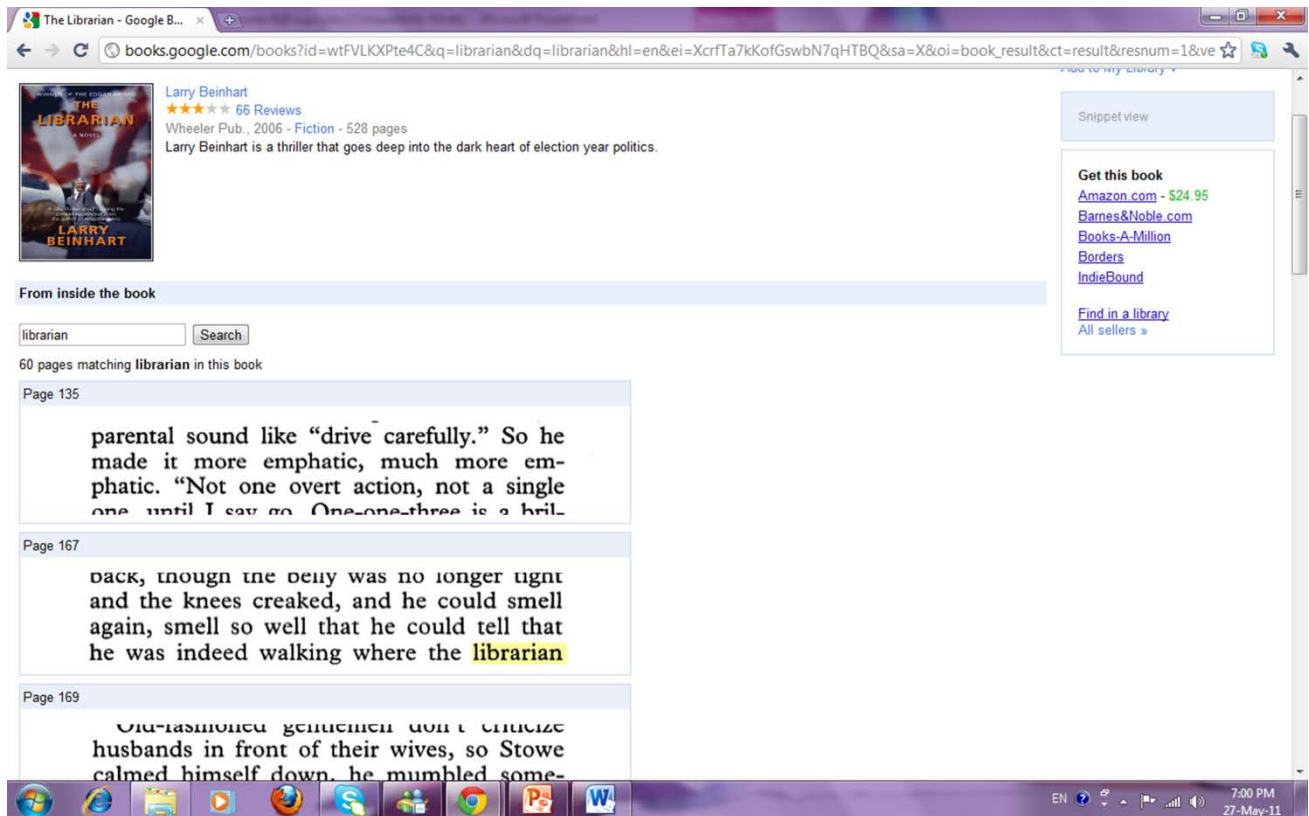
Το Google books είναι μια μηχανή αναζήτησης βιβλίων τα οποία έχουν ψηφιοποιηθεί και αποθηκευθεί στη βάση δεδομένων της Google. Η υπηρεσία αυτή ήταν γνωστή ως Google print - όπως παρουσιάστηκε στο Frankfurt Book Fair τον Οκτώβριο του 2004 - και αργότερα ως Google Book Search - όπως ανακοινώθηκε το Δεκέμβριο του 2004.

Οι εγγεγραμμένοι χρήστες μπορούν να δουν τα βιβλία των οποίων η Google έχει τα δικαιώματα και να έχουν πρόσβαση στις σελίδες τους στην οθόνη του ηλεκτρονικού τους υπολογιστή online ή κατεβάζοντάς τα (ebook). Παρόλ' αυτά η Google έχει κάποια όρια στις σελίδες που μπορεί ο κάθε χρήστης να κατεβάσει και να δει για λόγους ασφαλείας. Στις περιπτώσεις που η Google δεν έχει τα δικαιώματα ή δεν είναι γνωστό το ποιος είναι ο συγγραφέας του βιβλίου, εμφανίζονται μόνο 2-3 γραμμές σε κάθε σελίδα.

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN



Εικόνα 4 6 Ενδεικτική ιστοσελίδα της μηχανής



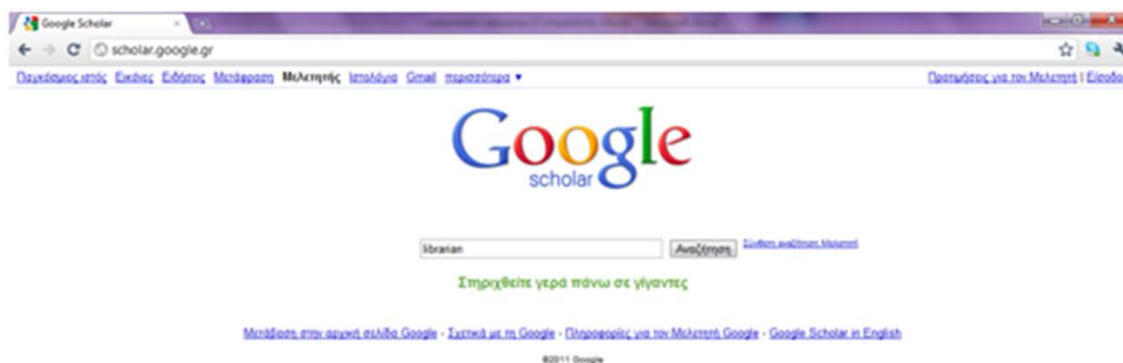
Εικόνα 4 7 Ενδεικτική ιστοσελίδα της μηχανής



Εικόνα 4 8 Το λογότυπο της μηχανής

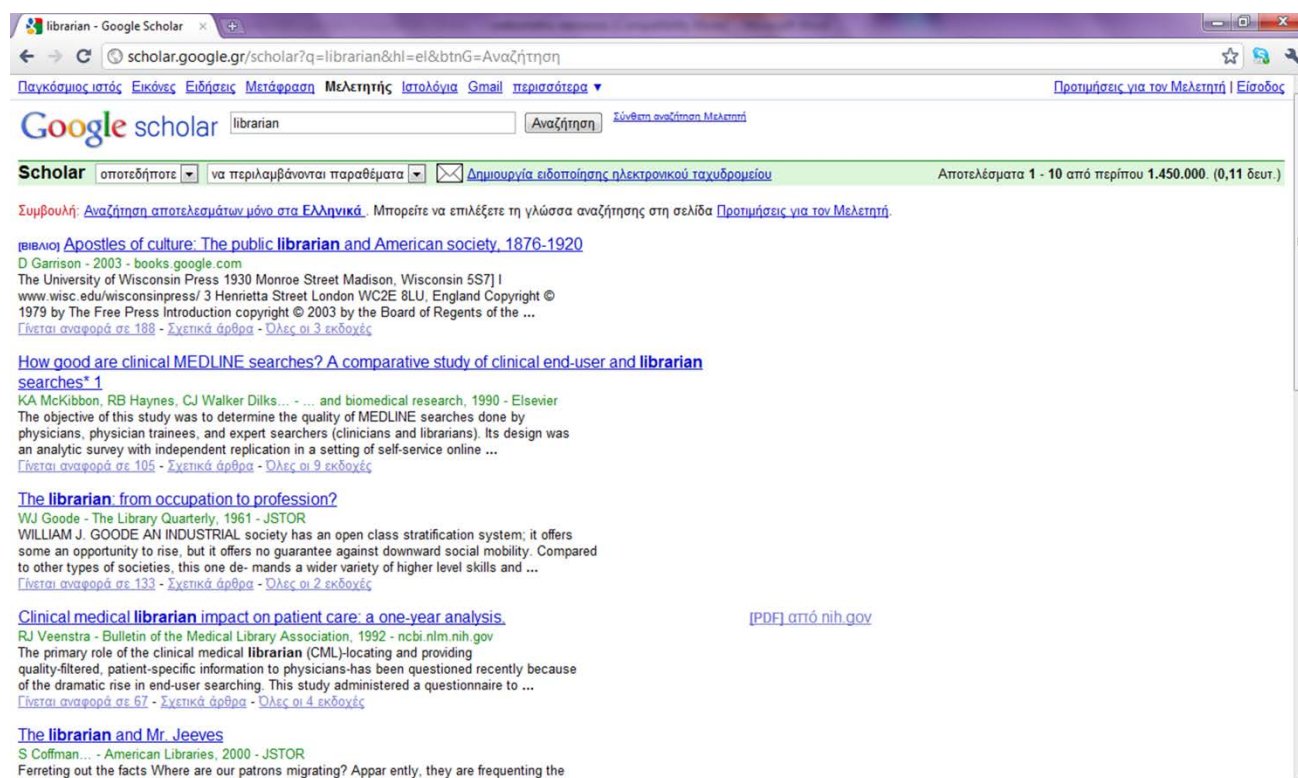
Υπηρεσία της γνωστής εταιρίας Google, για την παροχή δωρεάν μηχανής αναζήτησής στο ευρύ κοινό. Αυτή η μηχανή ειδικεύεται στο ψάξιμο επιστημονικής βιβλιογραφίας, όπως πτυχιακές, papers, τεχνικές αναφορές (κ.ά.) από διάφορους κλάδους.

Τα αποτελέσματα που επιστρέφονται ανασύρονται από διάφορες πηγές όπως είναι ακαδημαϊκοί εκδότες, επαγγελματικές κοινότητες, πανεπιστήμια καθώς και αποθήκες υλικού που δεν έχουν εκδοθεί ακόμα. Φυσικά η αναζήτηση συμπεριλαμβάνει και επιστημονικά άρθρα που είναι διαθέσιμα στο Web.



Εικόνα 4 9 Κεντρική σελίδα μηχανής

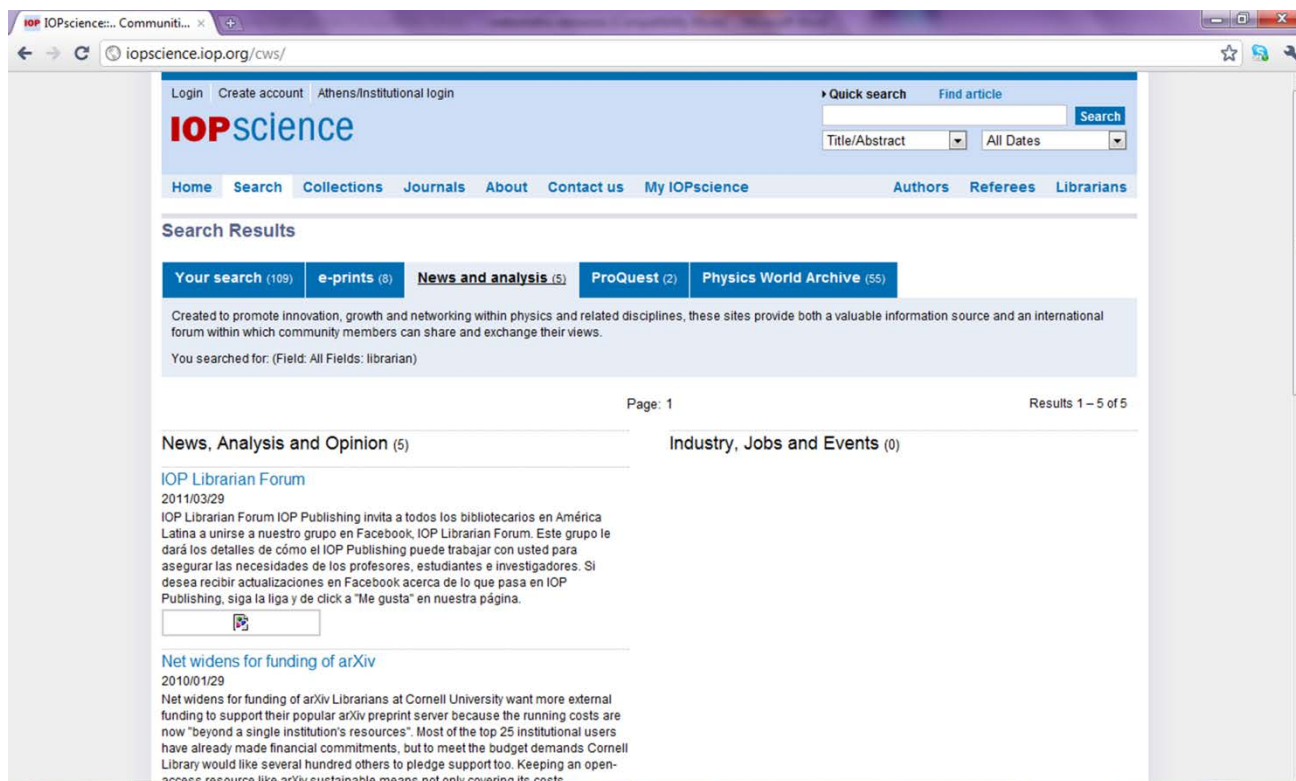
Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN



Εικόνα 4 10 Ενδεικτική ιστοσελίδα της μηχανής



Εικόνα 4 11 Το λογότυπο της μηχανής



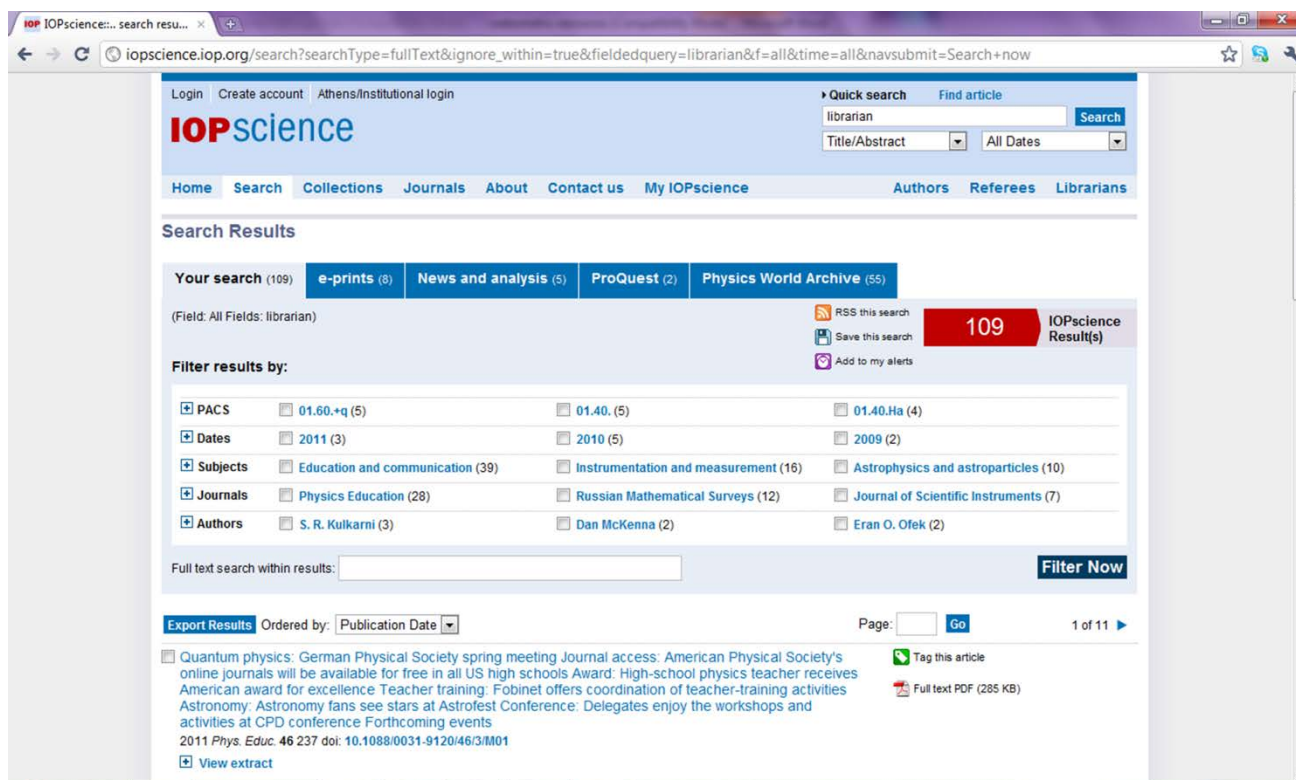
Εικόνα 4 12 Ενδεικτική ιστοσελίδα της μηχανής

Η υπηρεσία του Institute of Physics (IoP Electronic Journals) του εκδότη του Institute of Physics εξασφαλίζει την ελεύθερη on-line πρόσβαση στο περιεχόμενό της, στις περιλήψεις και στα πλήρη κείμενα (fulltext) άρθρων επιστημονικών περιοδικών. Τα περιοδικά καλύπτουν το πεδίο κυρίως, του χώρου της Φυσικής. Το δε περιβάλλον της αλληλεπίδρασης είναι στα Αγγλικά.

Διαθέτει:

- Μηχανισμό αναζήτησης με ταυτόχρονη ομαδοποίηση των αποτελεσμάτων.
- Αναλυτικές πληροφορίες για κάθε τίτλο περιοδικού.
- Δυναμικό και εύρηστο οδηγό βοήθειας.
- Εξατομικευμένες υπηρεσίες (απαιτείται εγγραφή).
- Υπηρεσίες ειδοποίησης με τεχνολογία Rich Site Summary (RSS).
- Παρέχει επίσης τη δυνατότητα:
- Απλής και σύνθετης αναζήτησης με χρήση Boolean λογικής και πεδίων αναζήτησης.
- Βελτίωση αναζήτησης.
- Προσαρμογής εμφάνισης αποτελεσμάτων.
- Αποθήκευσης αποτελεσμάτων αναζήτησης και αποστολής μέσω email.
- Εξαγωγή των βιβλιογραφικών στοιχείων των αποτελεσμάτων αναζήτησης σε μορφή συμβατή με προγράμματα διαχείρισης βιβλιογραφικών αναφορών (EndNote, ProCite, Reference Manager).
- Δημιουργίας ιστορικού αναζητήσεων (εξατομικευμένη υπηρεσία).

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN

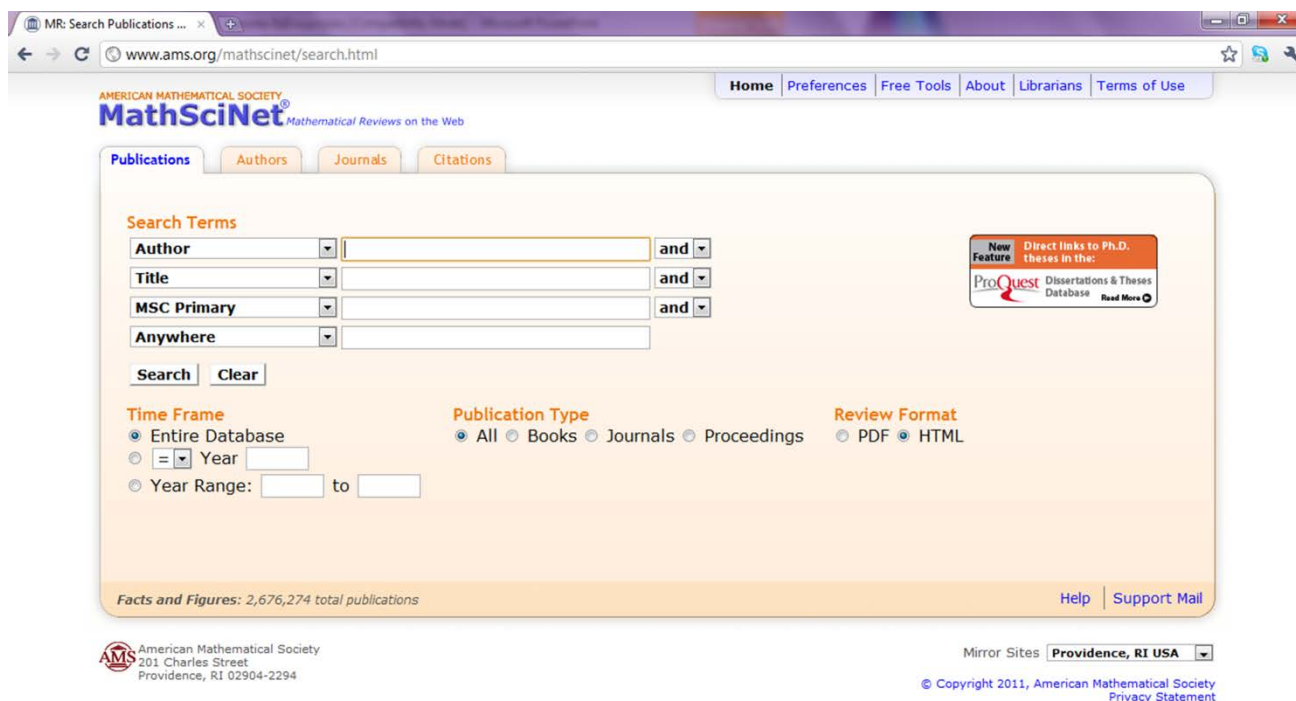


Εικόνα 4 13 Κεντρική σελίδα μηχανής



Εικόνα 4 14 Το λογότυπο της μηχανής

Το Mathematical Sciences Network (MathSciNet) αποτελεί τη μεγαλύτερη βιβλιογραφική βάση δεδομένων που καλύπτει την παγκόσμια μαθηματική βιβλιογραφία από το 1940 μέχρι σήμερα. Περιλαμβάνει περιλήψεις και κριτικές για άρθρα 383.458 συγγραφέων και ευρετηριάζει 1799 τρέχοντα περιοδικά. Εκτός από την περίληψη και την κριτική για το κάθε άρθρο, προσφέρεται σύνδεσμος στο πλήρες κείμενο σε ηλεκτρονική μορφή (όπου αυτό είναι διαθέσιμο και με την προϋπόθεση ότι η Βιβλιοθήκη έχει τα απαραίτητα δικαιώματα πρόσβασης). Μέχρι σήμερα, υπάρχουν σύνδεσμοι στο πλήρες κείμενο 213.098 πρωτότυπων άρθρων.



Εικόνα 4 15 Κεντρική σελίδα μηχανής



Εικόνα 4 16 Κεντρική σελίδα μηχανής

Η βάση δεδομένων της Εθνικής Υπηρεσίας Αεροναυπηγικής και Διαστήματος (NASA'S Astrophysics Data System) είναι μια βάση δεδομένων χρηματοδοτούμενη από τη NASA, η οποία περιλαμβάνει περιλήψεις άρθρων από τέσσερις κατηγορίες:

1. Αστρονομία και αστροφυσική
2. Εγκατάσταση και χρήση αστρονομικών οργάνων
3. Φυσική και γεωφυσική
4. Προδημοσιεύσεις του Los Alamos.

Εκτός των περιλήψεων, έχει αρχίσει να προσφέρεται, επίσης, δωρεάν πρόσβαση στο πλήρες κείμενο άρθρων πολλών περιοδικών.



Εικόνα 4 17 Κεντρική σελίδα μηχανής



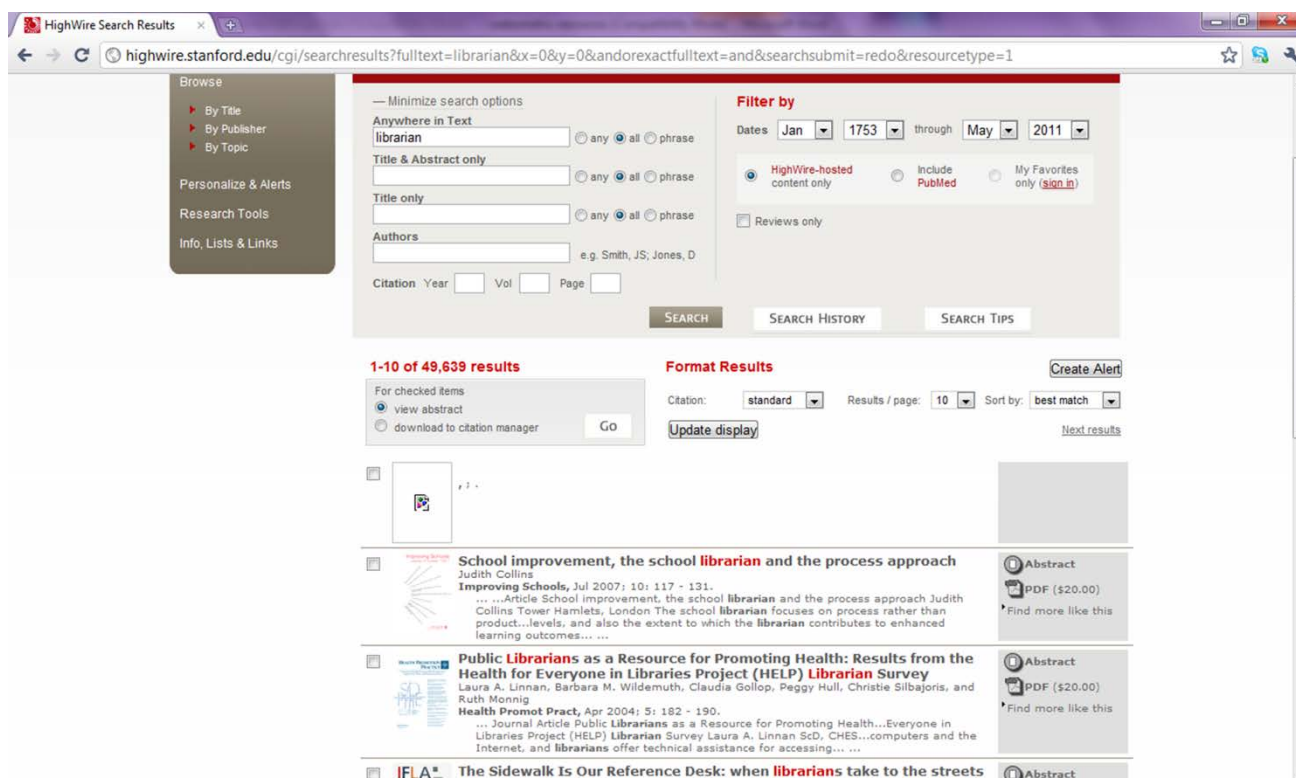
Εικόνα 4 18 Το λογότυπο της μηχανής

Το **HighWire Press** αποτελεί τμήμα του Stanford University Libraries και φιλοξενεί έναν μεγάλο αριθμό άρθρων ανοιχτής πρόσβασης (πάνω από 800,000 άρθρα).

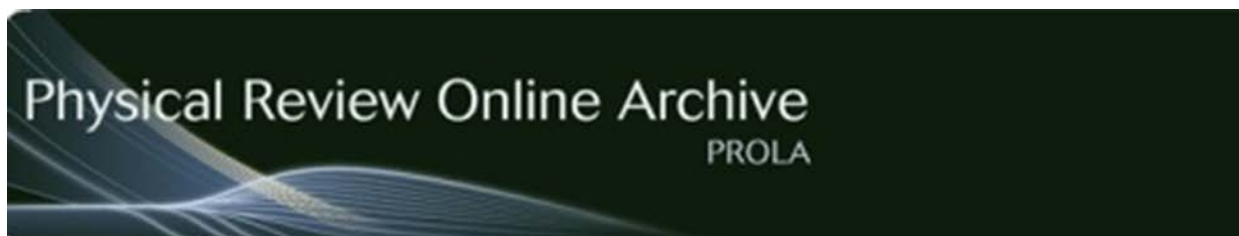


Εικόνα 4 19 Κεντρική σελίδα μηχανής

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN

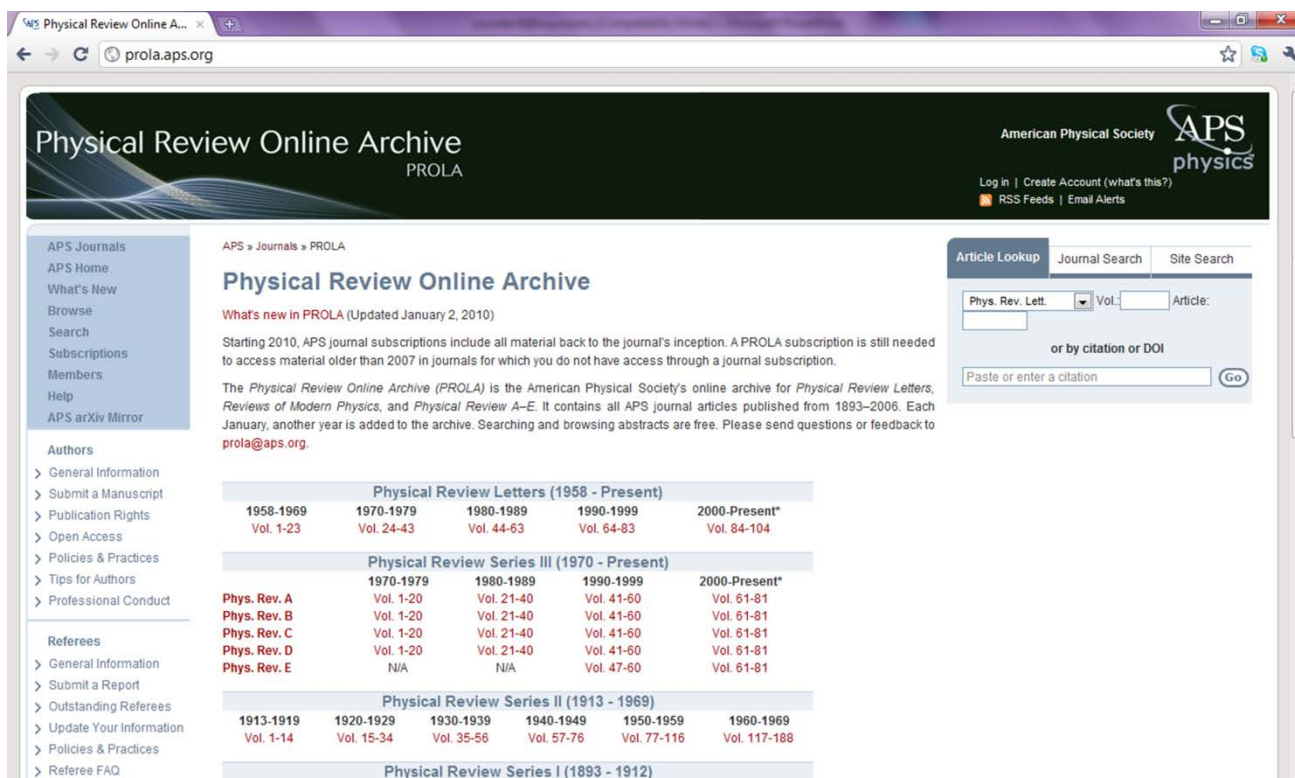


Εικόνα 4 20 Κεντρική σελίδα μηχανής



Εικόνα 4 21 Το λογότυπο της μηχανής

PROLA είναι το **Physical Review Online Archive** της **American Physical Society**. Είναι η απτή έκφραση για τη διασφάλιση της άμεσης και μακροπρόθεσμης πρόσβασης στο σύνολο του περιεχομένου του περιοδικού, που έχει δημοσιευθεί. Υπάρχει συμφωνία με τη Βιβλιοθήκη του Κογκρέσου για τη θέσπιση ενός αποθηκευτικού χώρου όλου του υλικού PROLA.

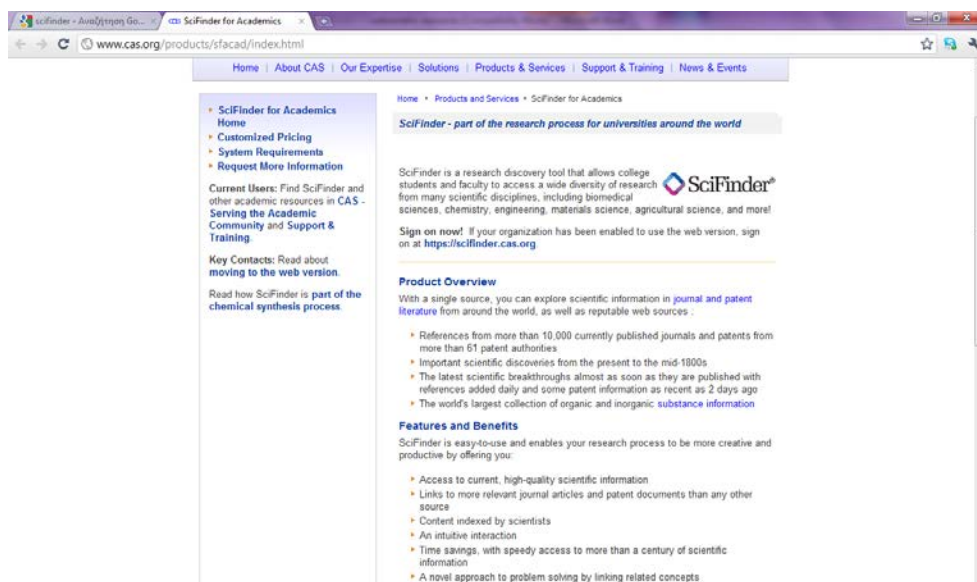


Εικόνα 4 22 Κεντρική σελίδα μηχανής



Εικόνα 4 23 Το λογότυπο της μηχανής

Το **SciFinder** είναι ένα εργαλείο έρευνας που επιτρέπει σε φοιτητές και καθηγητές να έχουν πρόσβαση σε μεγάλη ποικιλία ερευνών από πολλούς επιστημονικούς κλάδους, συμπεριλαμβανομένων των βιοϊατρικών επιστημών, της χημείας, της μηχανικής, της επιστήμης των υλικών, της γεωργικής επιστήμης και πολλών άλλων.

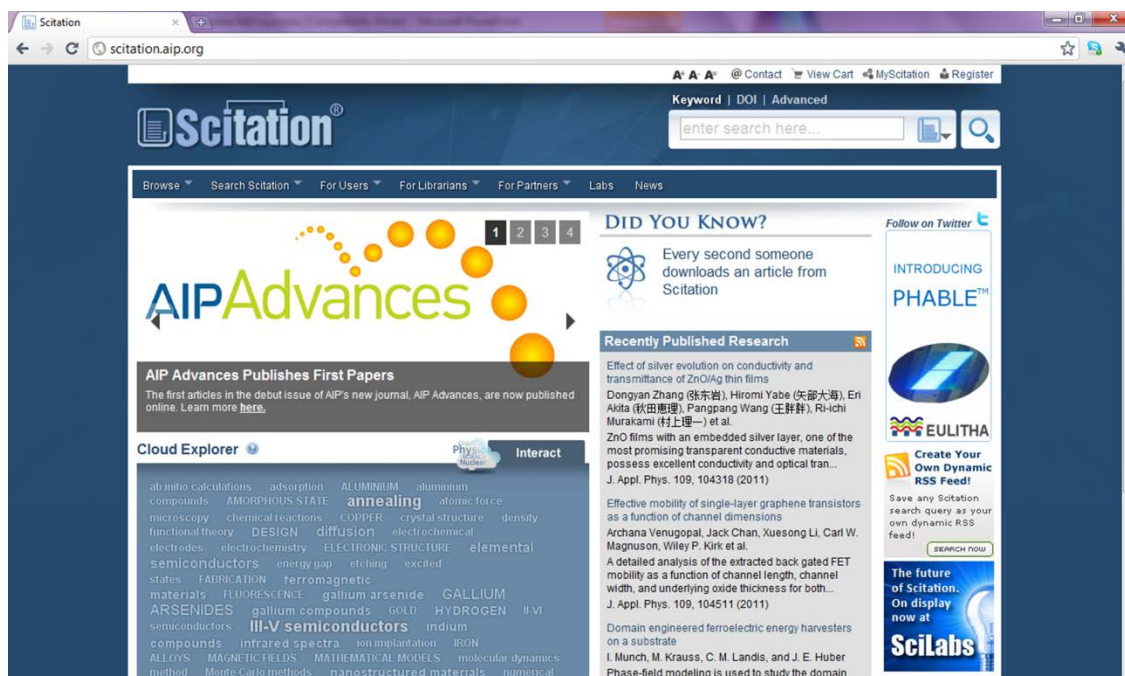


Εικόνα 4 24 Κεντρική σελίδα μηχανής



Εικόνα 4 25 Το λογότυπο της μηχανής

Η βιβλιογραφική βάση δεδομένων **SPIN** (Searchable Physics Information Notices) είναι μια υπηρεσία ευρετηρίασης και περίληψης που παράγεται από το Αμερικανικό ίδρυμα φυσικής (aip). Η εστίαση της SPIN περιγράφεται ως η σημαντικότερη περιοχή φυσικής έρευνας.



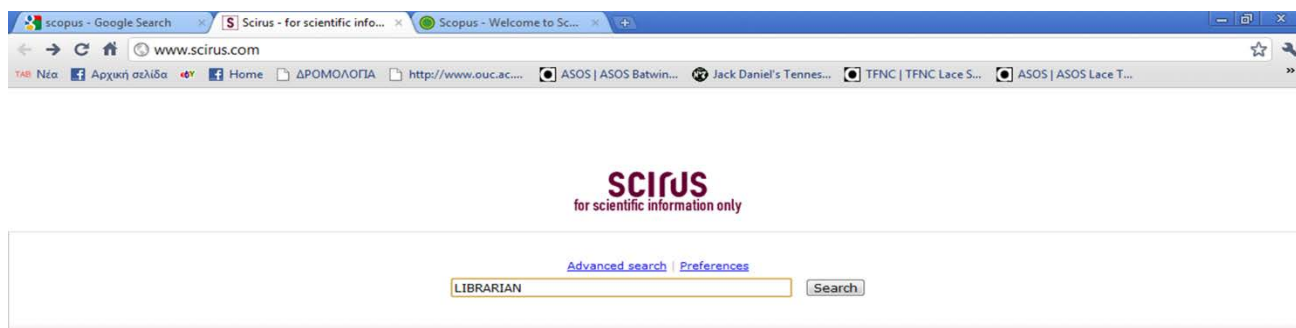
Εικόνα 4 26 Κεντρική σελίδα μηχανής

SCIRUS
for scientific information only

Εικόνα 4 27 Το λογότυπο της μηχανής

Είναι το πιο ολοκληρωμένο επιστημονικό, ερευνητικό εργαλείο στο διαδίκτυο. Περιέχει πάνω από 40 εκατομμύρια επιστημονικά άρθρα.

Επιτρέπει στους ερευνητές να αναζητούν, εκτός από το περιεχόμενο του περιοδικού, τις ιστοσελίδες των επιστημόνων, τα διπλώματα ευρεσιτεχνίας του ιδρυματικού αποθετηρίου και της ιστοσελίδας πληροφοριών.



SCIRUS is the most comprehensive scientific research tool on the web. With over 410 million scientific items indexed at last count, it allows researchers to search for not only journal content but also scientists' homepages, courseware, pre-print server material, patents and institutional repository and website information.

[SciTopics - expert generated knowledge sharing service for the scientific community](#)

[Latest Scientific News - from New Scientist](#)

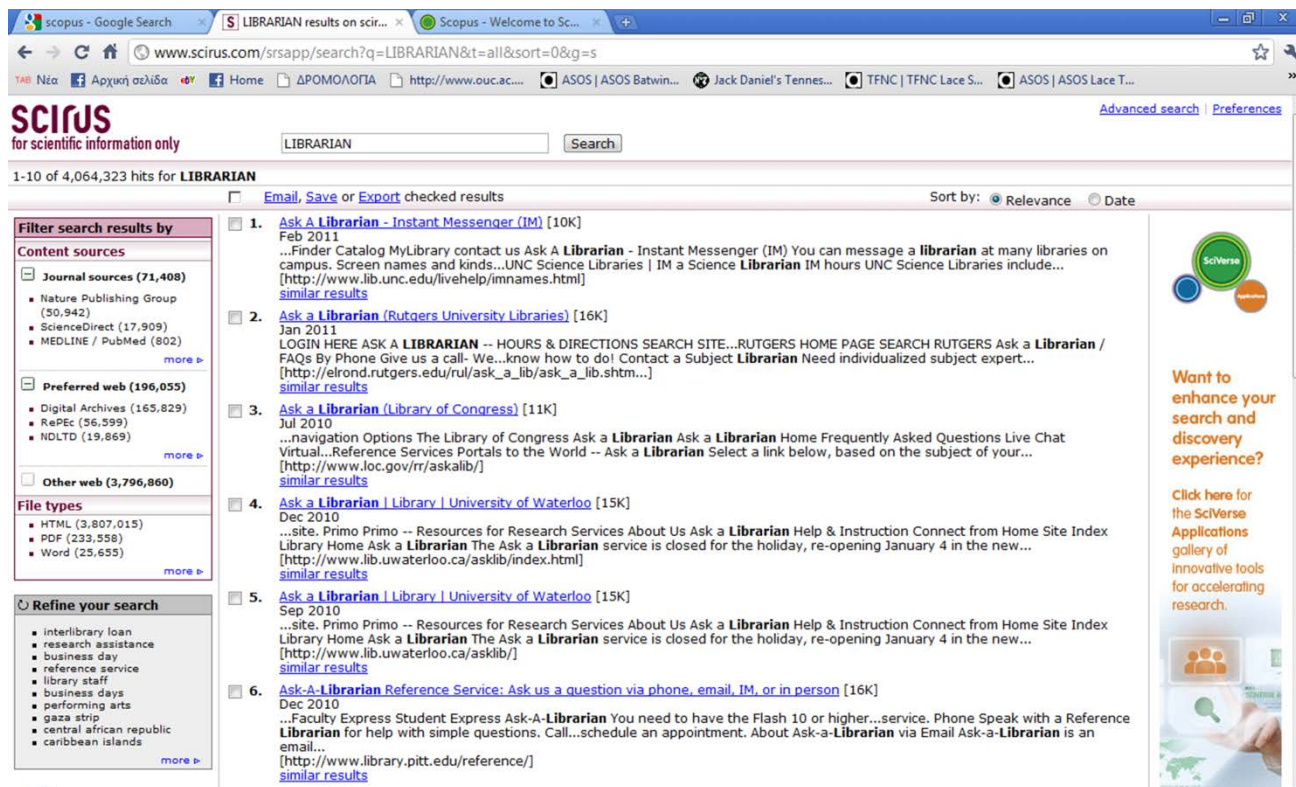
[Downloads](#) | [Submit website](#) | [Scirus newsletter](#) | [Help](#) | [Library partners](#) | [Contact us](#)

[About us](#) | [Advisory board](#) | [Privacy policy](#) | [Terms & Conditions](#) | [Newsroom](#)

Powered by [FAST](#) © Elsevier 2011

Εικόνα 4 28 Κεντρική σελίδα μηχανής

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN



Εικόνα 4 29 Ενδεικτική σελίδα μηχανής



Εικόνα 4 30 Το λογότυπο της μηχανής

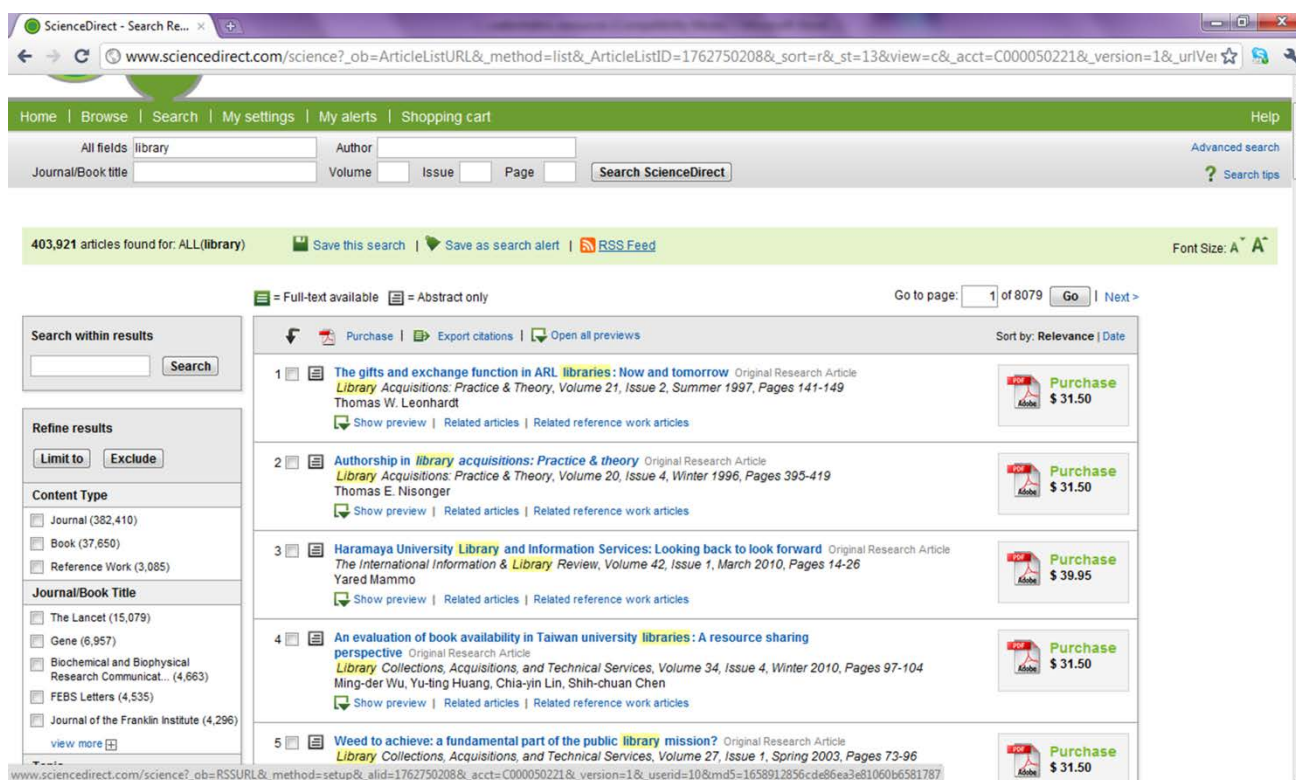
Το **ScienceDirect** είναι μια από τις μεγαλύτερες, απευθείας σύνδεσης, συλλογές δημοσιευμένης επιστημονικής έρευνας στον κόσμο. Κατασκευάστηκε από τον εκδότη του *Elsevier* και περιέχει σχεδόν 10 εκατομμύρια άρθρα, πάνω από 2.500 περιοδικά και πάνω από 6.000 e-books, τις εργασίες αναφοράς, τις σειρές βιβλίων και τα εγχειρίδια που εκδίδονται από τον *Elsevier*.

Τα άρθρα ομαδοποιούνται στα πιο κάτω κύρια τμήματα:

- Φυσικές επιστήμες
- Εφαρμοσμένη μηχανική
- Βιολογικές επιστήμες
- Επιστήμες υγείας
- Κοινωνικές επιστήμες

Για τα περισσότερα άρθρα σχετικά με τον ιστοχώρο, οι περιλήψεις είναι ελεύθερα διαθέσιμες, ενώ η πρόσβαση στο πλήρες κείμενο του άρθρου (σε PDF και HTML για τις νεώτερες δημοσιεύσεις) απαιτεί συνδρομή.

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN

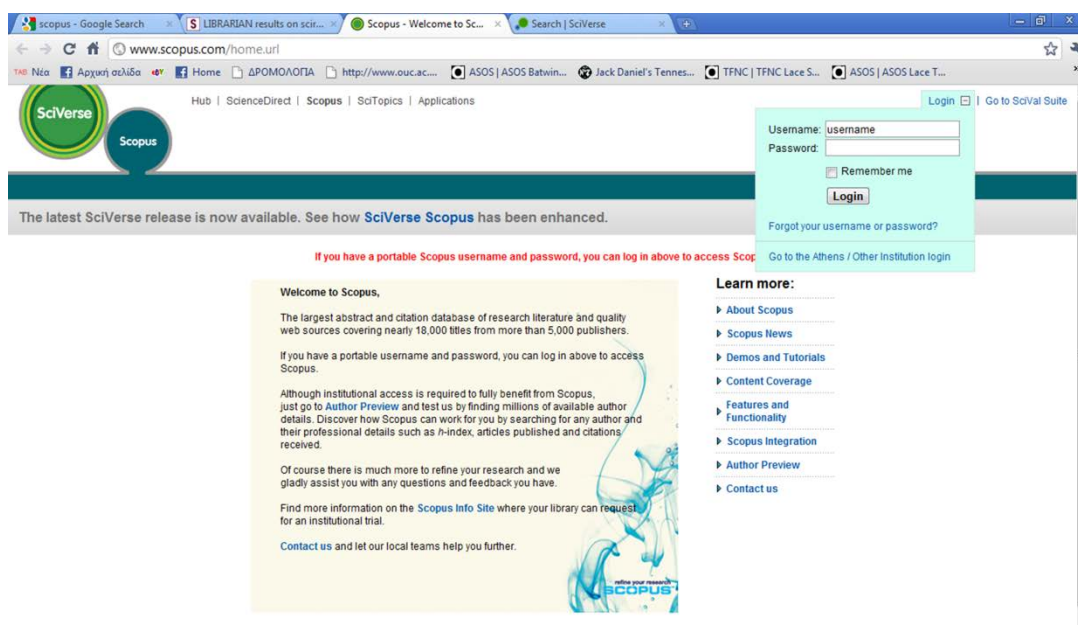


Εικόνα 4 31 Ενδεικτική σελίδα μηχανής



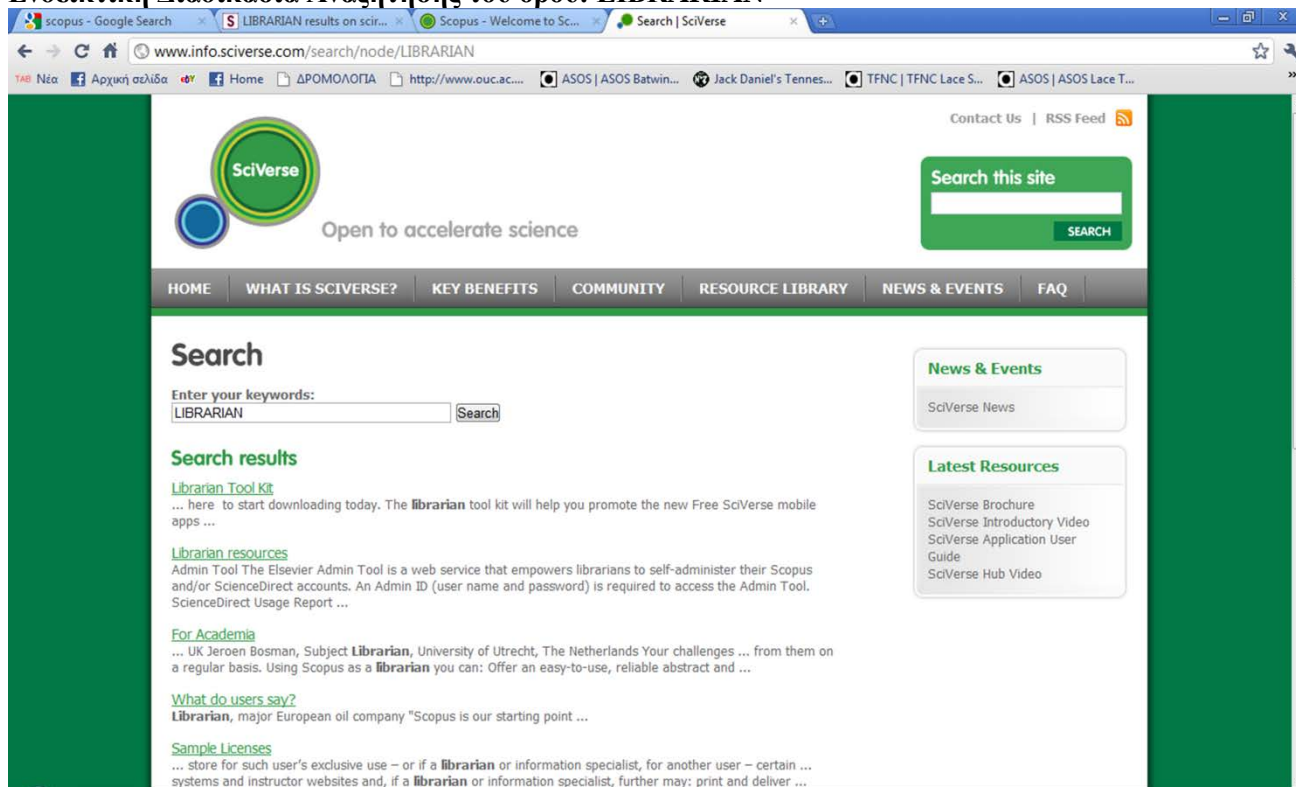
Εικόνα 4 32 Το λογότυπο της μηχανής

Η Scopus όπως έχει ήδη αναφερθεί είναι η μεγαλύτερη βάση δεδομένων (database) για περιλήψεις (abstracts) και αναφορές (citations) ερευνητικών δημοσιεύσεων και ποιοτικών πηγών του διαδικτύου. Η Scopus παρέχει ταχεία, εύκολη και ουσιαστική αναζήτηση βιβλιογραφίας, με καθημερινή ανανέωση.



Εικόνα 4 33 Κεντρική σελίδα μηχανής

Ενδεικτική Διαδικασία Αναζήτησης του όρου: LIBRARIAN



Εικόνα 4 34 Κεντρική σελίδα μηχανής

4.2. Βιβλιομετρικοί μέθοδοι στον Ιστό.

Οι δημοσιεύσεις θεωρούνται οι στοιχειώδεις μονάδες των επιστημονικών πληροφοριών και η κύρια πηγή των δεικτών. Η ποικιλία των νέων μοντέλων της επικοινωνίας της σχετικής με το ηλεκτρονικό δίκτυο και η ετερογένειά του Διαδικτύου, ωστόσο, θολώνει τα παραδοσιακά σύνορα μεταξύ της τυπικής και της άτυπης επικοινωνίας. Το διαδίκτυο είναι ένας χώρος πληροφοριών αρκετά διαφορετικός από τον κοινό επιστημονικό και τις επαγγελματικές βάσεις δεδομένων κι οι ομοιότητες μεταξύ ηλεκτρονικών και παραδοσιακών μέσων είναι συχνά επιφανειακή.

Η πιο σημαντική διαφορά μεταξύ των έντυπων μέσων ενημέρωσης και του web είναι ότι ο χρόνος παίζει διαφορετικό ρόλο στο web. Μια πρόσθετη, θεμελιώδης διαφορά μεταξύ της εκτύπωσης και της ανάλυσης που στηρίζεται στο web είναι η δυνατότητα μιας σχεδόν συνεχούς αλλαγής του περιεχομένου στο web, με αποτέλεσμα μια εντελώς διαφορετική αντίληψη της γήρανσης των πληροφοριών.

Παρά το γεγονός ότι η αναλογία μεταξύ των βιβλιομετρικών και web μετρικών φαινομένων είναι περιορισμένη, πολλές βιβλιομετρικές μετρήσεις και μοντέλα μπορούν επίσης να εφαρμοστούν σε web μετρικές, όπως:

- Ο αριθμός των επισκεπτών των ιστοσελίδων μπορεί να περιγραφεί από σημειακές στοιχειώδεις διαδικασίες ή νέες που δημιουργούνται από σπάνιες σωρευτικές διαδικασίες στο web.

- Η συχνότητα των λήψεων μπορεί να περιγραφεί από μετρικά μοντέλα που χρησιμοποιούνται στο πλαίσιο των αιτημάτων αντίγραφου.

Όσον αφορά στις βιβλιομετρικές αναλύσεις, υπάρχουν αναλογίες αλλά και αποκλίσεις από τα αντίστοιχα μοντέλα αναφορών.

Η «Webographic σύζευξη» θα μπορούσε να είναι ένα πολλά υποσχόμενο εργαλείο για την απεικόνιση των δομών του web και τον εντοπισμό συμπλεγμάτων στο διαδίκτυο.

Μελέτες φαινομένων από τον «μικρό κόσμο» είναι πιο σχετικές στην webομετρία από ό,τι στη βιβλιομετρία, όπου ελάχιστα έχουν γίνει σε αυτόν τον τομέα.

Στο πλαίσιο της ανάλυσης και χαρτογράφησης των δομών του δικτύου στο web, πολλές τυπικές στατιστικές τεχνικές που χρησιμοποιούνται στην βιβλιομετρία μπορούν να εφαρμοστούν.

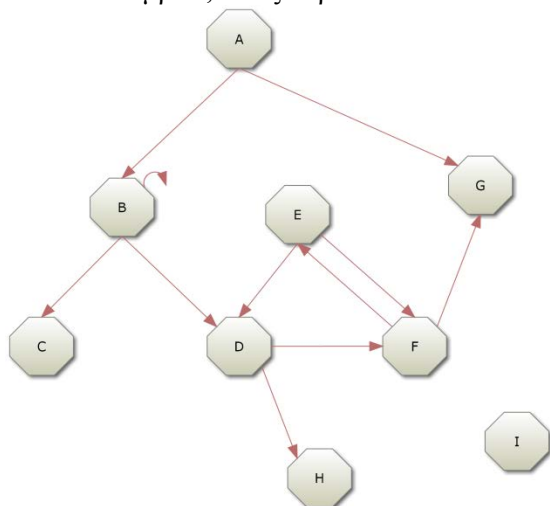
Από όλα αυτά μπορούμε να συμπεράνουμε ότι οι δομές στο διαδίκτυο είναι πιο περίπλοκες από ό, τι οι αντίστοιχες βιβλιογραφικές τους. Ωστόσο, η κύρια διαφορά μεταξύ των έντυπων μέσων ενημέρωσης και του web μπορούν να συνοψιστούν ως ακολούθως. Οι περισσότερες βιβλιομετρικές διαδικασίες είναι σωρευτικές δεδομένων των δημοσιεύσεων, οι αναφορές είναι μη αναστρέψιμες και οι βιβλιογραφικές συνδέσεις δεν μπορούν να αφαιρεθούν από τη στιγμή που έχουν συμβεί. Αντίθετα, ο ιστός, από την άποψη τόσο του περιεχόμενου όσο και των συνδέσεων, βρίσκεται σε μόνιμη αλλαγή.

Η Βιβλιομετρία είναι άμεσα σχετιζόμενη με τον Παγκόσμιο Ιστό, το μεγαλύτερο δίκτυο των εγγράφων στην ιστορία. Η επισκόπηση παραδειγμάτων σε αυτή την ενότητα περιλαμβάνει την ανάλυση αρχείων καταγραφής συμβάντων χρήσης διακομιστή Web όλων των ειδών, την αρχιτεκτονική των πόρων του συστήματος, καθώς και τη δημιουργία περιεχομένου για να μεγιστοποιηθεί η συμβατότητα με βιβλιομετρικές μεθόδους. Με την προσαρμογή πόρων Παγκοσμίου Ιστού για τη δύναμη της βιβλιομετρίας μπορούμε να δημιουργήσουμε πιο ακριβή δεδομένα και να καθιερώσουμε μεγαλύτερη αυστηρότητα για την οργάνωση της αρχιτεκτονικής του συστήματος.

4.2.1 Ορολογία

4.2.1.1 Βασική ορολογία συνδέσμων

Οι αρχικές εξερευνητικές φάσεις ενός νέου πεδίου, όπως η webometrics, αναπόφευκτα οδηγούν σε ποικιλία των όρων που χρησιμοποιούνται. Για παράδειγμα, ένας σύνδεσμος που έχει προέλθει από έναν κόμβο του web (εδώ ο δικτυακός όρος "κόμβος" καλύπτει μια ενότητα ανάλυσης όπως μια σελίδα του web, έναν φάκελο, έναν ιστότοπο ή μια ολόκληρη περιοχή υψηλού επιπέδου μιας χώρας) ονομάστηκε, εισερχόμενος σύνδεσμος, *inboundlink*, *inwardlink*, *backlink*, and *citation*. Ο τελευταίος όρος δημιουργήθηκε (McKiernan, 1996; Rousseau, 1997) με σαφή παραπομπή σε ανάλυση βιβλιογραφικών αναφορών. Ένα παράδειγμα μιας περισσότερο προβληματικής ορολογίας είναι τα δύο αντίθετα νοήματα του εξωτερικού συνδέσμου - είτε ως σύνδεσμος που δείχνει εκτός του ιστοτόπου είτε ως σύνδεσμος που δείχνει σε έναν ιστότοπο. Προτείνουμε τη συνεπή βασική webometric ορολογία των Björneborn και Ingwersen για τις σχέσεις των συνδέσμων μεταξύ των web κόμβων, όπως παρατίθενται εν συντομία στο σχήμα. 4.1



Σχήμα 4.1 Τα γράμματα μπορούν να αναπαριστούν διαφορετικά web επίπεδα κόμβων, για παράδειγμα, ιστοσελίδες, φακέλους web, ιστοτόπους, ή υψηλού επιπέδου περιοχές χωρών ή γενικών τομέων. (Björneborn & Ingwersen, 2004)

Η προτεινόμενη ορολογία έχει ρίζες στη θεωρία γραφημάτων, στην ανάλυση κοινωνικών δικτύων και στη βιβλιομετρία.

- B έχει έναν εξωτερικό σύνδεσμο από το A (outlink).
- B έχει έναν εσωτερικό σύνδεσμο στο C (inlink).
- B συνδέεται με τον εαυτό του.

- Ε και F είναι αμοιβαίως συνδεδεμένα.
- Α έχει ένα εγκάρσιο εξωτερικό σύνδεσμο προς G: λειτουργεί ως συντόμευση.
- Η συνδέεται με το Α μέσω ευθείας σύνδεσης μονοπατιού.
- Ι δεν έχει ούτε εσωτερικό ούτε εξωτερικό σύνδεσμο, είναι απομονωμένο.
- Β και Ε συν-συνδέονται (co-linked) από το D, Β και G έχουν κοινούς εξωτερικούς συνδέσμους (ζεύγη αναφορών).
- C και D συν-συνδέονται στο Β, Β και Ε έχουν κοινούς εσωτερικούς συνδέσμους.
- Το Α συνδέεται μεταβατικά με το C και το D μέσω του Β (βιβλιογραφική ζεύξη).

Οι όροι *outlink* και *inlink* χρησιμοποιούνται πολύ συχνά σε μελέτες γύρω από την επιστήμη των υπολογιστών (Pirolli et al., 1996; Chen et al., 1998; Broder et al., 2000). Ο όρος *outlink* υποδηλώνει πως ένας κατευθυνόμενος σύνδεσμος φαίνεται από τον κόμβο στόχο, ανάλογα με τη χρήση του όρου παραπομπή στη βιβλιομετρία. Μια αναλογία υπάρχει ανάμεσα στους όρους *inlink* και *citation*, με τον κόμβο πηγή να αποτελεί το σημείο που βλέπει ο θεατής. Ένας σύνδεσμος που ξεπερνά τα σύνορα ενός ιστοτόπου, όπως ο σύνδεσμος Β στο σχ. 4.1. παρακάτω, ονομάζεται *siteoutlink* ή *siteinlink*, ανάλογα την οπτική του θεατή.

Στο Web, ο όρος *selflinks* χρησιμοποιείται για ένα μεγαλύτερο εύρος σκοπών σε σχέση με τις *self-citations* στην επιστημονική λογοτεχνία. Τα *selflinks* σελίδων δείχνουν από μια ενότητα σε μια άλλη, μέσα στην ίδια σελίδα. Τα *selflinks* ιστοτόπων (γνωστά και ως *internallinks*) αποτελούν τυπικούς δείκτες πλοήγησης από τη μια σελίδα σε μια άλλη, μέσα στον ίδιο ιστότοπο. Οι περισσότεροι σύνδεσμοι στο Web συνδέουν ιστοσελίδες που περιέχουν συγγενή θέματα (Davison, 2000). Παρόλα αυτά, κάποιοι σύνδεσμοι μπορούν να ξεφύγουν από το τυπικό σχέδιο σύνδεσης στη «γειτονιά» web κόμβων και να συνδέσουν μη συγγενικές περιοχές. Αυτοί έχουν ορισθεί ως *trans versallinks* (Björneborn, 2001, 2003) και λειτουργούν ως *cross-topic* συντομεύσεις, που μπορεί να επηρεάσουν τα λεγόμενα φαινόμενα του μικρόκοσμου του web (βλ. την ενότητα σχετικά με τους μικρούς κόσμους).

Οι δύο συνδεδεμένοι web κόμβοι C και D στην εικ. 2.1 με *co-inlinks* από τον ίδιο κόμβο – πηγή Β είναι ανάλογοι με τη βιβλιομετρική έννοια *co-citation* (Small, 1973). Αντίστοιχα, οι δύο συνδεδεμένοι κόμβοι Β και Ε έχουν *co-outlinks* στον ίδιο κόμβο – στόχο και αναλογούν στο *bibliographic coupling* (Kessler, 1963). Το *Co-links* προτείνεται ως γενικός όρος που θα καλύπτει και τις 2 έννοιες.

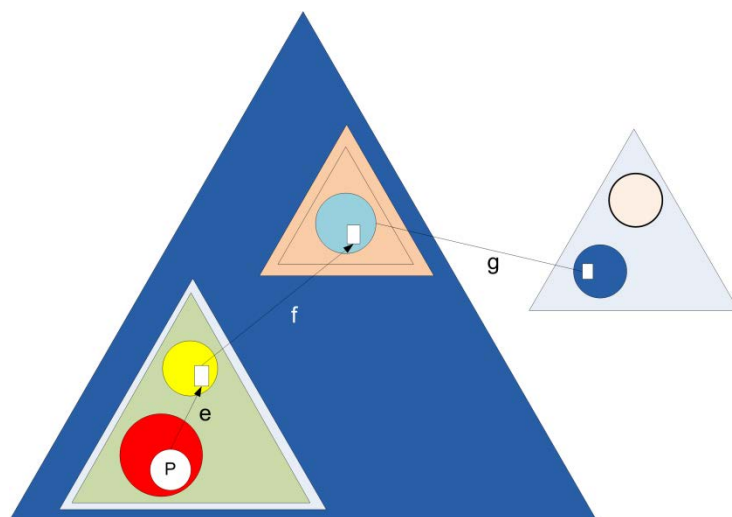
Η ανάλυση *co-citation* (συν-αναφορών) βασίζεται στην υπόθεση ότι εάν δύο αναφορές παρατίθενται μαζί, σε μια τελευταία βιβλιογραφία, τότε οι δύο αυτές αναφορές σχετίζονται μεταξύ τους. Η κυριότερη βελτίωση μεταξύ της βιβλιομετρικής ζεύξης και των συν-αναφορών (*co-citation*) είναι ότι, ενώ η σύζευξη μετρά τη σχέση μεταξύ των εγγράφων που αποτελούν τις πηγές για άλλα έγγραφα, η μέθοδος *co-citation* μετρά τις σχέσεις μεταξύ των αναφορών.

4.2.1.2 Βασική ορολογία web Κόμβων και διαγράμματα

Στις μελέτες *webometric* μπορεί να αποδειχτεί χρήσιμη η οπτικοποίηση των διαφορετικών μονάδων ανάλυσης, για παράδειγμα στο Εναλλακτικό Μοντέλο Κειμένου (Alternative Document Model) που περιγράφεται παρακάτω.

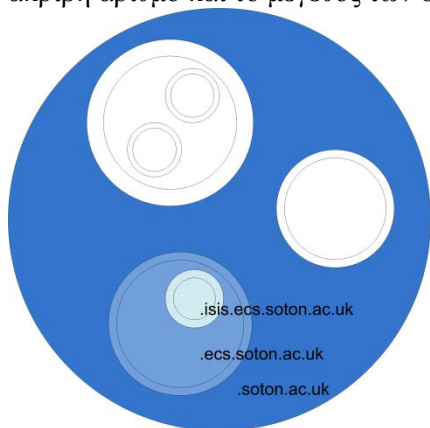
Το σχήμα 4.2 δείχνει ένα διάγραμμα που απεικονίζει κάποια βασικά δομικά στοιχεία στο πλαίσιο ενός σταθερού web κόμβου (Björneborn & Ingwersen, 2008). Στο διάγραμμα υπάρχουν 4 βασικά επίπεδα web κόμβων, που αποτυπώνονται με απλά γεωμετρικά στοιχεία: *τετράπλευρα* (ιστοσελίδες), *διαγώνιες γραμμές* (web φακέλοι), *κύκλοι* (ιστότοποι) και *τρίγωνα* (χώρα ή υψηλού επιπέδου γενικές περιοχές, *Top-Level Domains* (TLD's)- το κομμάτι μετά την τελεία π.χ. .gr .com). Υπό-επίπεδα μέσα στα τέσσερα επίπεδα κόμβων σημειώνονται με επιπλέον σύνορα στην αντίστοιχη γεωμετρική φιγούρα. Για παράδειγμα, ένα τρίγωνο με διπλή γραμμή συνοριακή υποδηλώνει μια γενική περιοχή 2^{ου} επιπέδου (SLD), γνωστή και ως υπό-TLD, που ανατίθεται από πολλές χώρες σε τμήματα της, εκπαιδευτικά, εμπορικά, κυβερνητικά και άλλα, για παράδειγμα .ac.uk, .co.uk, .ac.jp, .edu.au.

Το υπεραπλουστευμένο διάγραμμα web κόμβων στην εικόνα 4.2 δείχνει μια σελίδα (page) *P* που βρίσκεται στον φάκελο (directory) ενός υπό-site σε μια υπό-TLD. Αυτή η σελίδα έχει ένα *outlink* ιστοτόπου, *e*, σε μια σελίδα ενός ιστοτόπου στην ίδια υπό-TLD. Η *outlinked* σελίδα μετατρέπεται σε *outlinking f*, σε μια σελίδα ενός ιστοτόπου σε κάποιο άλλο υπό-TLD για την ίδια χώρα. Το μονοπάτι συνδέσμων *e-f-g* καταλήγει σε μια σελίδα ενός ιστοτόπου σε κάποια άλλη TLD.



Σχήμα 4 2 Απλοποιημένο διάγραμμα δικτυακού κόμβου που απεικονίζει βασικά επίπεδα δικτυακού κόμβου..(Björneborn & Ingwersen, 2004).

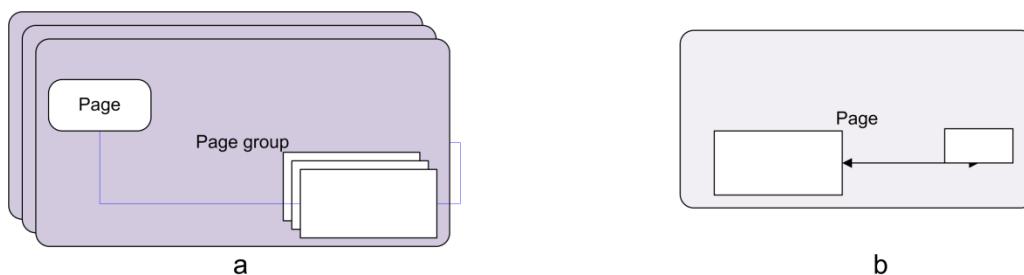
Κοιτώντας καλύτερα σε ένα και μόνο ιστότοπο, αυτό θα μπορούσε να αποτελέσει πολλές υπό-ενότητες στο σχήμα των υπό-ιστοτόπων, υπό-υπόιστοτόπων, κ.λ.π., όπως υποδεικνύεται από ιεραρχικά παράγωγα ονόματα περιοχών. Για παράδειγμα, όπως φαίνεται στο σχ. 4.3, ο υπό-υπό-ιστότοπος `isis.ecs.soton.ac.uk` βρίσκεται ανάμεσα στον υπό ιστότοπο `ecs.soton.ac.uk`, ένας από τους υπο ιστοτόπους του University of Southampton, UK (`soton.ac.uk`). Υπό-ιστότοποι και υπό-υπό ιστότοποι διαφαίνονται ως κύκλοι με διπλές και τριπλές συννοριακές γραμμές αντίστοιχα. Κατώτερα υποεπίπεδα λογικά θα διαφαίνονταν με αύξηση στον αριθμό των οριακών γραμμών. Για λόγους απλότητας, το διάγραμμα δεν αντικατοπτρίζει τον ακριβή αριθμό και το μέγεθος των στοιχείων.



Σχήμα 4 3 Απλοποιημένο διάγραμμα δικτυακού κόμβου μιας ιστοσελίδας που περιέχει υπο-σελίδες και υπο-υπο-σελίδες.(Björneborn & Ingwersen, 2004).

Ενώ κάποιοι ιστότοποι υποδιαίρονται σε παράγωγα ονόματα περιοχών, όπως φαίνεται παραπάνω, άλλοι ιστότοποι ανιχνεύουν τον ίδιο τύπο υποενοτήτων σε καταλόγους (φακέλους). Προφανώς, τέτοιου είδους, δυσανάλογος καταμερισμός και πρακτικές ονομασίας καθιστούν πιο πολύπλοκη τη σύγκριση στις μελέτες webometrics. Στην εικ. 4.4. κατάλογοι, υπό-κατάλογοι, κ.λ.π., σημειώνονται από μια ή περισσότερες διαγώνιες γραμμές που μοιάζουν με καθέτους των URL και αντιστοιχούν στον αριθμό επιπέδων των καταλόγων κάτω από το αρχικό επίπεδο του URL.

Οι ιστοσελίδες μπορούν επίσης να συνίστανται από υπό-στοιχεία όπως ενότητες κειμένου, πλαίσια κ.λ.π. Πρόσθετες ομάδες δείχνουν αυτά τα υπό-στοιχεία των σελίδων ως τους στόχους της σελίδας *selflinkh* και της σελίδας *outlinkI* από τις δύο αδελφές ιστοσελίδες στον ίδιο κατάλογο, όπως φαίνεται στην εικόνα 4.4a. Πιο πολύπλοκες και πολυάριθμες συνδέσεις μέσα σε έναν ιστότοπο ή υπό-ιστότοπο κ.λ.π. μπορούν να διαφανούν από συνδυασμούς στοιχείων στην εικ. 4.4a&b, που δείχνει συνδέσμους μεταξύ σελίδων που βρίσκονται είτε σε διαφορετικά επίπεδα καταλόγων (εικ. 4.4a), είτε σε συγγενικούς καταλόγους στο ίδιο επίπεδο (εικ. 4.4b) στην ιεραρχία αρχείων του ιστοτόπου.



Σχήμα 4.4 Απλοποιημένα διαγράμματα δικτυακού κόμβου μιας ιστοσελίδας και μιας υπο-ιστοσελίδας με τις συνδέσεις μεταξύ των διαφόρων επιπέδων καταλόγου συμπεριλαμβανομένων των επιμέρους στοιχείων της σελίδας. (Björneborn & Ingwersen, 2004).

Φυσικά, κάθε διαγραμματική αναπαράσταση μεγάλης κλίμακας δομών υπερκειμένου κινδυνεύει να είναι τόσο σύνθετη, ώστε να μην έχει πρακτική χρήση – πόσο μάλλον για να μεταφραστεί με οποιονδήποτε ποσοτικό τρόπο. Παρόλα αυτά, τα προτεινόμενα διαγράμματα web κόμβων, με τα απλά και διαισθητικά τους γεωμετρικά σχήματα, πρόκειται να χρησιμοποιηθούν έτσι ώστε να δώσουν έμφαση και να απεικονίσουν ποσοτικές διαφορές ανάμεσα στα επίπεδα των διερευνούμενων web κόμβων, στη webometric μελέτη.

4.3 Εναλλακτικά μοντέλα εγγράφων

Όταν μετράμε αντικείμενα στο web, πρέπει να ληφθεί μια απόφαση για την πιο κατάλληλη μονάδα μέτρησης. Συχνά αυτή θα είναι είτε η σελίδα web ή το site του web. Πόσες ιστοσελίδες επιχειρήσεων βρίσκονται στις ΗΠΑ; Πόσες σελίδες υπάρχουν στην ιστοσελίδα του Stanford University; Και τα δύο, όμως, έχουν τα προβλήματα τους και είναι πιθανές κάποιες εναλλακτικές.

Υπάρχουν τεχνικά προβλήματα στην αναγνώριση των web pages, κάποια σχετίζονται με τον ορισμό του ίδιου του web, ορισμένα έχουν να κάνουν με HTML και άλλα με την ευρύτερη ερμηνεία του τι θα έπρεπε να είναι έγγραφο. Ένα HTML πρόβλημα είναι πως μια μόνο σελίδα μπορεί να αποτελείται από πολλά αρχεία, εάν χρησιμοποιείται το σύνολο πλαισίων HTML tag. Η μέτρηση των αρχείων, ασχέτως με το αν βρίσκονται σε σύνολο πλαισίων, αποτελεί πιθανή λύση, αφού τα ατομικά πλαίσια σε ένα σύνολο πλαισίων μπορούν να αλλάξουν ανεξάρτητα (Thelwall, 2002a).

Ένα εννοιολογικό πρόβλημα είναι πως ο συγγραφέας στο web μπορεί να επιλέξει σε πόσες σελίδες θα μοιράσει το έργο του. Ένα online βιβλίο, για παράδειγμα, θα μπορούσε να παρουσιαστεί σαν μια τεράστια ιστοσελίδα ή σε χιλιάδες ενότητες αποτελούμενες από ιστοσελίδες. Μια ιδανική λύση θα ήταν ίσως η μέτρηση των σελίδων ή συλλογών σελίδων βάσει το είδος, ώστε και στις δύο (2) περιπτώσεις αυτό που θα μετριούνταν θα ήταν ένα στιγμιότυπο του είδους του online βιβλίου. Αλλά η ταυτοποίηση του είδους στο web είναι ένα πολύ πολύπλοκο, άλυτο πρόβλημα (Crowston & Williams, 2000; Rehm, 2002) και έχουν αναπτυχθεί αυτοδίδακτες τεχνικές ως υποκατάστατο, οι Alternative Document Models (ADMs) (Thelwall, 2002b). Υπάρχουν 4 βασικά ADMs σε χρήση: η σελίδα, ο φάκελος, το πεδίο και το site ADM.

Το ADM σελίδας είναι η προκαθορισμένη ενότητα, που βασίζεται σε ατομικά αρχεία. Με τον φάκελο ADM, όλες οι σελίδες στον ίδιο (εικονικά) φάκελο αναγνωρίζονται μέσα από ένα μονοπάτι αρχείου URL και μετριούνται σαν μια ενότητα. Στο παράδειγμα του Online βιβλίου, εάν οι ενότητες σελίδων web αποθηκεύονταν σε έναν φάκελο, τότε η χιλιοστή σελίδα του online βιβλίου θα μετριούνταν σαν μια μόνο ενότητα, σε έναν web φάκελο. Προφανώς αυτή είναι μια αυτοδίδακτη μέθοδος, αφού ένα άτομο μπορεί επίσης να αποθηκεύσει ολόκληρη βιβλιοθήκη σε έναν και μόνο φάκελο. Το ADM πεδίο (παρόμοιο με την υπό-ιστοσελίδα στην εικ. 4.3) συγκεντρώνει όλες τις σελίδες με το ίδιο όνομα πεδίου στο URL τους σε μια κοινή οντότητα. Η ιστοσελίδα ADM (αρχικά ADM πανεπιστημίου) επιτρέπει πολλά ονόματα πεδίων στο ίδιο έγγραφο, ξεκαθαρίζοντας μόνο την κατάληξη του ονόματος του πεδίου. Η ιστοσελίδα ADM του University of Southampton στην εικόνα 4.3 θα περιλάμβανε όλες τις σελίδες από τις υπό-σελίδες, υπό-υπό- σελίδες, κ.λπ. με URLs που θα μπορούσαν να περιέχουν ονόματα πεδίων, που καταλήγουν σε .soton.ac.uk.

Άλλη μια ενότητα ανάλυσης έχει προταθεί και χρησιμοποιηθεί εκτεταμένα από τον Aguillo (1998), το “κάθισμα” (seat), όπως αναφέρεται και στην περιγραφή στο άρθρο του. Αυτό είναι παρόμοιο με το site ADM και είναι σύμφωνο με τις γενικές παραδοχές της φράσης «web site», αλλά περιλαμβάνει στοιχεία εκφραστικού ύφους. Το seat αποτελεί συλλογή ιστοσελίδων που έχουν δημιουργηθεί από ένα πρόσωπο ή οργανισμό για συγκεκριμένο σκοπό. Τυπικά, τα seats θα είναι μεγαλύτερα από μια μονή σελίδα και μπορεί να αποτελούν συλλογή σελίδων που σχετίζονται με ένα άτομο, όπως και αν φιλοξενούνται στο web. Τα seats

καμιά φορά θα είναι τα ίδια με τα directory, domain και site ADMs, αλλά είναι πιο ευέλικτα. Ως αποτέλεσμα, η έννοια είναι ταυτόχρονα πιο ισχυρή και δυσκολότερο να εφαρμοστεί αυτόματα στην πράξη.

Όταν μετράμε έγγραφα ή συνδέσμους, τα ADMS δίνουν την επιλογή να μετρήσουμε σε διαφορετικά επίπεδα συσσωμάτωσης. Για παράδειγμα, ο αριθμός των domains σε ένα site πανεπιστημίου θα μπορούσε να μετρηθεί, όπως και ο αριθμός των directories που περιέχουν ένα σύνδεσμο σε ένα δοσμένο WEB SITE. Αυτή η πρόσθετη ευελιξία μπορεί να χρησιμοποιηθεί για να βρεθούν ανωμαλίες στα δεδομένα και τα εννοιολογικά προβλήματα με τη μέτρηση των σελίδων, όπως αναφέρεται παρακάτω στις υπό-ενότητες για τις πανεπιστημιακές και τμηματικές μελέτες. Η επιλογή ADM για συγκεκριμένες μελέτες εξαρτάται από μοτίβα χρήσης του web για την περιοχή που καλύπτεται. Δύο(2) τεχνικές έχουν προταθεί για να βρεθεί η καλύτερη. Η πρώτη είναι να συγκρίνονται συσχετίσεις ανάμεσα σε στατιστικές περιλήψεις που μετριοούνται με διαφορετικά ADMs και από εξωτερική πηγή δεδομένων (Thelwall & Harries, 2003a; Thelwall & Wilkinson, 2003a). Η δεύτερη αποτελεί μια θεωρητική προσέγγιση που δεν απαιτεί αυτές, χρησιμοποιώντας αντ' αυτών υποθέσεις σχετικά με τη διεργασία που αποτελεί την αιτία δημιουργίας του εγγράφου και εισαγωγικών κειμένων για την αξιολόγηση της έκτασης, που οι υποθέσεις συναντούν τα δεδομένα (Thelwall, 2004a; Thelwall, 2004b).

4.4 Έρευνα και συλλογή δεδομένων του Παγκόσμιου Ιστού

Από τότε που δημιουργήθηκε ο Παγκόσμιος Ιστός, έχουν υπάρξει πολυάριθμες μελέτες για τον προσδιορισμό των δημογραφικών στοιχείων τόσο των χρηστών όσο και της χρήσης του (συλλογή πληροφοριών που αφορούν τη χρήση της τεχνολογίας του, τη χρήση των λιστών συζήτησης-discussion lists- ή την κατανόηση της χρήσης του). Οι περισσότερες από αυτές τις προσπάθειες είναι παρόμοιες στις βιβλιογραφικές τεχνικές και αποκαλύπτουν συγκρίσιμα αποτελέσματα. Με την εξέταση αυτών των ερευνών μπορούμε να αναπτύξουμε μεθόδους για να βελτιώσουμε ή να αυξήσουμε τη γενική βιβλιομετρία για να μάθουμε ποιοι είναι οι χρήστες και τι κάνουν. Η έρευνα μπορεί να γίνει με τη χρήση ποσοτικών μεθόδων, όπως με ερωτηματολόγια που στέλνονται είτε με τον παραδοσιακό τρόπο είτε μέσω ηλεκτρονικού ταχυδρομείου, είτε και με συνδυασμό των δύο μεθόδων. Ένας άλλος τρόπος έρευνας είναι η δημιουργία ειδικού χώρου στο site, όπου οι επισκέπτες του μπορούν να συμπληρώσουν τα στοιχεία που τους ζητούνται. Άλλες μέθοδοι που έχουν χρησιμοποιηθεί είναι οι ανοικτές συνεντεύξεις ή η ανάλυση του περιεχομένου συγκεκριμένων δικτυακών τόπων.

Οι έρευνες Παγκόσμιου Ιστού της Georgia Tech, σε σχέση με τα γραφικά, την οπτικοποίηση και την ευχρηστία, είναι ο ακρογωνιαίος λίθος της δημογραφίας στον Παγκόσμιο Ιστό. Μέσω της έξυπνης χρήσης του προγραμματισμού, της συλλογής δεδομένων, καθώς και της εκτεταμένης στατιστικής ανάλυσης, αναπτύχθηκε ένα προφίλ χρηστών του παγκόσμιου ιστού, καθώς και των προτιμήσεών τους. Χρησιμοποιώντας φόρμες HTML, οι ποιοτικές πληροφορίες μπορούν να συλλεχθούν επιπλέον για τα μεγάλα ποσά των ποσοτικών δεδομένων. Τεχνικές όπως η υπερδειγματοληψία και τυχαίες δειγματοληψίες (μεταξύ των χρηστών) εφαρμόζονται επίσης για την αφαίρεση, όσο το δυνατόν, περισσότερων ασαφειών. Έρευνες όπως αυτές θα μπορούσαν να επεκταθούν, ώστε στην πραγματικότητα να συγκεντρωθούν πιο συγκεκριμένες μετρήσεις παγκόσμιου Ιστού, παρατηρώντας πώς οι χρήστες πραγματικά αντιδρούν και χρησιμοποιούν συγκεκριμένα έγγραφα παγκόσμιου Ιστού, όχι ρωτώντας τους. Καθώς αυτές οι τεχνικές είναι εκλεπτυσμένες, μπορούν να αυτοματοποιηθούν σε μια προσπάθεια προς προσαρμόσιμες διεπαφές Παγκόσμιου Ιστού.

4.4.1 Προσωπικοί Web Crawlers

Οι μηχανές αναζήτησης (search engines π.χ. Google), χρησιμοποιούν προγράμματα τα οποία ονομάζονται crawlers, για να διασχίσουν όλες τις ιστοσελίδες του WWW. Αυτές εκτελούν αναδρομικά την εξής διαδικασία: a) Κατέβασε την σελίδα X, b) Βρες όλους του επόμενους συνδέσμους (links) από την X, c) Για κάθε σύνδεσμο εκτέλεσε τα βήματα a-c

Αν μια μελέτη ασχολείται με ένα μόνο μικρό web site, τότε ο ερευνητής μπορεί να είναι σε θέση να βρει όλες τις σελίδες μονάχα ξεφυλλίζοντας το, αν έχει λόγο να πιστεύει πως όλες οι σελίδες του site θα είναι προσβάσιμες. Σε μεγαλύτερα site, μια εναλλακτική είναι η χρήση μιας εμπορικής μηχανής αναζήτησης (περισσότερα σχετικά με αυτό αργότερα) ή ενός web crawler. Ο crawler ή αράχνη ή ρομπότ είναι ένα κομμάτι λογισμικού που μπορεί αυτόματα και επαναληπτικά να κατεβάζει σελίδες, να εξάγει και να κατεβάζει τους συνδέσμους των (Heydon & Najork, 1999; Arusu et al., 2001; Thelwall, 2001a). Για παράδειγμα, θα

μπορούσε να δοθεί το URL της αρχικής σελίδας ενός site και έπειτα να επιστρέψει όλες τις σελίδες του site που μπορούν να βρεθούν, ακολουθώντας τους συνδέσμους της αρχικής σελίδας. Παρόλα αυτά, η κάλυψη του crawler θα είναι λιγότερη από 100%. Οι crawlers βρίσκουν νέες σελίδες, ακολουθώντας τους συνδέσμους και είναι πιθανό να χάσουν αρκετές απομονωμένες σελίδες και συλλογές σελίδων. Οι σελίδες που είναι συνδεδεμένες με άλλες μπορεί επίσης να χαθούν από τον crawler, αν τα links είναι σε μορφή που ο crawler δεν καταλαβαίνει, για παράδειγμα JavaScript, αν οι σελίδες είναι προστατευμένες με κωδικό, αν προσωρινά ο server δεν λειτουργεί ή αν ο ιδιοκτήτης ζητήσει η σελίδα να μην παρατηρείται από crawler, δηλαδή μέσω του Robot Exclusion Protocol.

Σε ένα μεγάλο site πανεπιστημίου, οι σελίδες που δημιουργούνται από ατομικά μέλη του προσωπικού ή φοιτητές θα παραβλέπονταν, εκτός αν αναβαθμιζόνταν στο site από συσκευές, όπως λίστες του προσωπικού / των φοιτητών που παράγονται από τον server online ή συνδέσμους που προστίθενται χειρωνακτικά σε επίσημες σελίδες, ίσως από web administrators. Οι crawlers που τρέχουν από εμπορικές μηχανές αναζήτησης μπορούν προσωρινά να αντισταθμίσουν αυτό το πρόβλημα, προσθέτοντας στο ψάξιμο τους τα URLs όλων των σελίδων που προηγουμένως μαζεύτηκαν από το site, τα URL που αποκτήθηκαν από Links από άλλα site και κάθε URLs που εισήχθησαν στη μηχανή αναζήτησης από χρήστες (Thelwall, 2001b). Όλοι οι crawlers μπορούν επίσης να χρησιμοποιήσουν τεχνικές όπως ένα μέγιστο βάθος ψαξίματος, μέγιστο μονοπάτι βάθους ή μέγιστο αριθμό σελίδων που επιστρέφουν ανά site για να αποφύγουν τις άπειρες επαναλήψεις. Οι εμπορικές μηχανές αναζήτησης μπορεί, επίσης, να έχουν πολιτικές ψαξίματος που επηρεάζονται από οικονομικούς παράγοντες, με το πιο ακραίο παράδειγμα να είναι τα σχήματα που πλήρωσε για καταχώρηση (pay-for-inclusion). Οι διαφορετικές πολιτικές ψαξίματος μπορούν να δώσουν πολύ διαφορετικά αποτελέσματα για το ίδιο site, δίνοντας έμφαση στην σημαντικότητα της πλήρους αναφοράς των παραμέτρων ψαξίματος.

4.4.2 Εμπορικές μηχανές αναζήτησης

Οι εμπορικές μηχανές αναζήτησης έχουν χρησιμοποιηθεί εκτεταμένα για webometric έρευνα και είναι ιδιαίτερα χρήσιμες για μετρήσεις σελίδων σε μεγάλες περιοχές του web. Μερικές, συμπεριλαμβάνοντας την AltaVista, επιτρέπουν αναπτυγμένα ερωτήματα με την Boolean λογική, ώστε να συνδυάσουν ερωτήματα πληροφοριών σχετικά με σελίδες, συνδέσμους και περιοχών (Almind & Ingwersen, 1997) που μπορούν να χρησιμοποιηθούν για να επιστρέψουν μετρήσεις σελίδων και σελίδες συνδέσμων που ικανοποιούν τα δοσμένα κριτήρια. Αυτή είναι η υπηρεσία που γέννησε την webometric. Οι σελίδες βοήθειας αναλυτικής αναζήτησης των μηχανών αναζήτησης θα έπρεπε να λαμβάνονται υπόψη για λεπτομέρειες του πλήρους εύρους των ερωτημάτων, που είναι διαθέσιμες για κάθε μηχανή.

- Οι εμπορικές μηχανές έχουν και αρκετά μειονεκτήματα. Δεν καλύπτουν όλο το web ή καλύτερα όλο το δημόσια ευρετηριασμένο web. Οι αλγόριθμοι ψαξίματος και αναφοράς τους αποτελούν εμπορικά μυστικά, κάτι που είναι πρόβλημα για τη χρήση τους στην επιστημονική έρευνα. Τα αποτελέσματα τους στο παρελθόν βρέθηκαν αναξιόπιστα καθώς διακυμαίνονται ανάμεσα σε μικρά χρονικά διαστήματα (Ingwersen, 1998; Rousseau, 1999; Snyder & Rosenbaum, 1999; Bar-Ilan, 1999; Mettrop & Nieuwenhuysen, 2001). Για τον μετριασμό αυτού, η AltaVista βρέθηκε ως η πιο αξιόπιστη (Thelwall, 2001c; Vaughan & Thelwall, 2003a) και ακόμη πως παρείχε καλή κάλυψη των ακαδημαϊκών web site (Thelwall, 2001b).
- Παρόλα αυτά, οι ερευνητές θα έπρεπε πάντοτε να κάνουν βήματα για την αξιολόγηση της ακρίβειας και της κάλυψης των αποτελεσμάτων, γιατί οι αλγόριθμοι μπορούν να υπόκεινται σε αλλαγές χωρίς προειδοποίηση. Η ακρίβεια μπορεί να ελεγχθεί τεστάροντας ένα δείγμα σελίδων που επιστρέφονται από την AltaVista, για να διασφαλιστεί πως περιέχουν τα χαρακτηριστικά που διευκρινίστηκαν στο ερώτημα. Η κάλυψη είναι πιο δύσκολο να ελεγχθεί, αλλά για κάθε δοσμένη μελέτη μπορεί να είναι πιθανό να χρησιμοποιηθεί ο χώρος φιλοξενίας της AltaVista: εντολή για τον έλεγχο του αριθμού των σελίδων που ευρετηριάζονται από την AltaVista, σε ένα σύνολο από site κλειδιά. Ο Rousseau (1999) πρότεινε επίσης στατιστικές τεχνικές για να αντισταθμίσει τις διακυμάνσεις. Οι Björneborn & Ingwersen (2001) πρότειναν να χρησιμοποιήσουν δειγματισμό και φιλτράρισμα για να καθαρίσουν τα δεδομένα από τις μηχανές αναζήτησης, λαμβάνοντας υπόψη ότι είναι πιθανό αυτό να είναι δύσκολο να εφαρμοστεί.

Μια εναλλακτική στρατηγική είναι η χρήση πολλαπλών μηχανών αναζήτησης, αλλά κάτι τέτοιο θα αύξανε τη θεωρητική κάλυψη, φέρνοντας το βάρος των σελίδων που ευρετηριάζονται πιο ψηλά (αφού θα

μετριοῦνταν πάνω από μια φορά) και η διεργασία θα γινόταν πιο χρονοβόρα για τον έλεγχο τον μεθόδων λειτουργίας, ακρίβειας και κάλυψης των μηχανών που χρησιμοποιούνταν. Εναλλακτικά, διαφορετικές μηχανές αναζήτησης μπορεί να χρησιμοποιούνται παράλληλα, συγκρίνοντας τα μεταξύ τους αποτελέσματα, ώστε να βρεθεί αν οι εσωτερικές διαφορές αποτελούν σημαντικό παράγοντα όσον αφορά το αποτέλεσμα της έρευνας. Αυτή είναι μια προσέγγιση που αξίζει να ληφθεί σοβαρά υπόψη από κάθε έρευνα που χρησιμοποιεί τις εμπορικές μηχανές. Φυσικά, η ομοιότητα στα αποτελέσματα δεν αποκλείει την πιθανότητα και οι 2 μηχανές να έχουν παρόμοια ελαττώματα, πιθανώς για τους τεχνικούς λόγους που αναφέρθηκαν.

Ένα μελλοντικό ζήτημα για την εγκυρότητα των δεδομένων από τις μηχανές αναζήτησης μπορεί να αποτελούν τα εμπορικά μοντέλα που αλλάζουν, ιδίως αν περιλαμβάνουν 'pay-for-inclusion and pay-for-placement' πληρωμή για καταχώρηση και πληρωμή για τοποθέτηση. Δεν υπάρχει κάποια λογική αιτία γιατί αυτό δεν θα έπρεπε να επεκταθεί σε πανεπιστήμια και ερευνητικές ομάδες που πληρώνουν για βελτιωμένες θέσεις σε μεγαλύτερες μηχανές. Προφανώς κάτι τέτοιο θα προκαλούσε προβλήματα για όσους χρησιμοποιούν δεδομένα από μηχανές αναζήτησης, για να αξιολογήσουν την άμεση επιρροή (online impact). Ακόμη η συνεχόμενη παροχή προηγμένων διεπαφών (interface) αναζήτησης δεν θα έπρεπε να θεωρείται δεδομένη, αφού μπορεί να σταματήσουν, αν κριθεί πως δεν αποφέρουν κέρδος. Αυτές οι σκέψεις αποτελούν κίνητρο για τους ερευνητές, ώστε να αναπτύξουν τεχνικές συλλογής δεδομένων που δεν βασίζονται στην AltaVista ή οποιαδήποτε άλλη μηχανή.

Οι εμπορικές μηχανές αναζήτησης δείχνουν εθνικούς προϋπολογισμούς σε κάλυψη ιστοσελίδων. Μια μελέτη τεσσάρων χωρών έδειξε ότι τα site των ΗΠΑ ήταν πολύ καλύτερα καλυμμένα από τρεις μεγάλες μηχανές, σε σχέση με την Κίνα, την Ταϊβάν και τη Σιγκαπούρη (Vaughan & Thelwall, 2004).

Τα δεδομένα διαστήματα χρόνων μπορούν να αποκτηθούν από κάποιες μηχανές και από διαδικτυακή βιβλιοθήκη (Internet Archive). Για παράδειγμα, η AltaVista επιτρέπει αναζητήσεις για σελίδες σύμφωνα με την ημερομηνία τελευταίας αλλαγής τους ή δημιουργίας τους (Leydesdorff & Curran, 2000). Αυτά αποτελούν μονάχα μερικά δεδομένα, γιατί η αναζήτηση θα αγνοήσει την πραγματική ημερομηνία δημιουργίας, επιστρέφοντας μόνο την ημερομηνία τελευταίας αλλαγής. Η μηχανή Internet Archive Wayback παρέχει αξιόπιστα δεδομένα διαστημάτων χρόνων, γιατί καταγράφει όλες τις ημερομηνίες στις οποίες μια σελίδα βρέθηκε, κρατώντας αντίγραφα κάθε ευρετηριασμένης εκδοχής, έτσι ώστε οι αλλαγές να μπορούν να εντοπιστούν στο πέρασμα του χρόνου (Vaughan&Thelwall, 2003).

Το αρχείο μπορεί να χρησιμοποιηθεί για τη γνωστή ημερομηνία της δημιουργίας για κάθε σελίδα, αλλά όχι για την ακριβή ημερομηνία δημιουργίας της, γιατί η πρώτη που καταγράφεται είναι όταν πρωτοβρέθηκε από τον crawler του Archive και όχι όταν ακριβώς δημιουργήθηκε. Ο Björneborn (2003) απέκτησε ενδεικτικές ηλικίες διαφορετικών τοπολογικών συστατικών των ακαδημαϊκών web site της Μ. Βρετανίας βασισμένες στις πρώτες ημερομηνίες ευρετηρίασης από το αρχείο. Ένα μειονέκτημα της Wayback Machine είναι πως δεν επέτρεπε σφαιρικά Boolean ερωτήματα πέρα από απλές ταυτίσεις με μοτίβα μπαλαντέρ 'wildcard matches', αντίθετα με την AltaVista. Μια εναλλακτική είναι η χρήση crawler και η συλλογή δεδομένων για μια προκαθορισμένη λίστα web σελίδων ή site (2002), αλλά αυτό δεν είναι δυνατό για αναδρομικές μελέτες.

4.4.3 Web Log Files

Εκτός του να μελετάει κανείς το ίδιο το web, είναι επίσης χρήσιμο να συλλέγει δεδομένα για το πώς οι άνθρωποι αλληλεπιδρούν με αυτό, δηλαδή για το πώς χρησιμοποιείται. Αυτή η συλλογή δεδομένων μπορεί να περιλάβει μελέτες ανθρώπων που χρησιμοποιούν το web. Άλλη πηγή πληροφοριών είναι τα αρχεία web server log. Αυτά καταγράφουν τα ερωτήματα που στέλνονται στους web servers από τους browser των χρηστών (και των web crawlers) και μπορούν να εξορύξουν χρήσιμες πληροφορίες για το πώς χρησιμοποιούν το site. Οι τυπικές εφαρμογές της log ανάλυσης του server αποσκοπούν στο να βρεθεί ποιες σελίδες επισκέπτονται πιο συχνά ή για να αναγνωριστούν μοτίβα όσον αφορά το σερφάρισμα, με σκοπό να βελτιωθεί η πλοήγηση στο site. Το βασικό μειονέκτημα, από την άποψη της webometric, είναι πως τα log files τυπικά καλύπτουν ένα site ή όλα τα site που ανήκουν στον ίδιο server και συχνά έχουν περιορισμένη πρόσβαση και γενικά δεν αποτελούν καλή πηγή δεδομένων για χρήση μελέτης των web site άλλων ατόμων.

Τα log files των μηχανών αναζήτησης μπορούν να δώσουν πληροφορίες για το τι ενδιαφέρει έναν χρήστη και πού πήγε για να το βρει. Γενικά αποτελούν μυστικό, αλλά χρήσιμες μελέτες έχουν διεξαχθεί από ερευνητές, στους οποίους δόθηκε πρόσβαση σε δείγματα αρχείων log files (Silverstein *et al.*, 1999; Spink *et al.*, 2001; Spink *et al.*, 2002).

Υπάρχουν πολλές έμφυτες δυσκολίες στην ανάλυση των αρχείων web log. Τα κύρια προβλήματα είναι στην αναγνώριση ενός χρήστη και στον καθορισμό μιας συνεδρίας αναζήτησης. Στα παραδοσιακά συστήματα ανάκτησης, οι χρήστες πρέπει να κάνουν log in και log out, ώστε να αναγνωρίζεται πλήρως η κάθε συνεδρία αναζήτησης. Το web στερείται ταυτότητας, κάτι που καθιστά την ταυτοποίηση εξαιρετικά δύσκολη. Τα κύρια μέτρα που χρησιμοποιήθηκαν για το καθορισμό ενός μοναδικού χρήστη web είναι τα cookies (μικρές ποσότητες πληροφορίας που αποθηκεύονται από το μηχάνημα του χρήστη) και IP διευθύνσεις-(Internet Protocol address), είναι ένας μοναδικός αριθμός που χρησιμοποιείται από συσκευές για τη μεταξύ τους αναγνώριση-. Οι Silverstein et al. (1999) χρησιμοποίησαν την προηγούμενη τεχνική για τη μελέτη των χρηστών AltaVista. Για ερωτήματα στα οποία ο χρήστης απενεργοποίησε τα cookies, χρησιμοποίησαν το ζευγάρι «domain IP/web browser used» ως υποκατάστατο του cookie. Αναγνώρισαν πως η μέθοδος cookie είναι ατελής, γιατί διαφορετικοί άνθρωποι μπορεί να χρησιμοποιήσουν τον ίδιο browser π.χ. ένα δημόσιο computer σε μια βιβλιοθήκη. Η παρακολούθηση της συμπεριφοράς αναζήτησης ενός χρήστη στο χρόνο είναι ακόμη πιο δύσκολη, γιατί το άτομο μπορεί να χρησιμοποιεί διαφορετικά computer π.χ. ένα στη δουλειά κι ένα στο σπίτι. Η αναγνώριση χρήστη από τη διεύθυνση IP είναι ακόμη λιγότερο ακριβής, γιατί πολλοί χρήστες μοιράζονται μια μόνο IP σε μεγάλους *internet-service-provider* (ISP's) και ορισμένοι οργανισμοί αναθέτουν κυμαινόμενες (floating) IP διευθύνσεις σε ηλεκτρονικό υπολογιστή. Η υποσχόμενη πλευρά είναι πως 96% των ερωτημάτων στη μελέτη (Silverstein et al., 1999) έχει πληροφορία cookie. Στην εργασία τους ανέλυσαν τα δεδομένα με δύο (2) τρόπους, έναν για όλο το σύνολο δεδομένων και έναν για αυτά με πληροφορία cookie. Τα δύο (2) αποτελέσματα έμοιαζαν πολύ.

Μια κοινή μέθοδος για τον καθορισμό συνεδρίας αναζήτησης είναι το «time out». Οι Silverstein κ.α. (1999) όρισαν τη συνεδρία ως μια «σειρά ερωτημάτων από έναν μόνο χρήστη που πραγματοποιήθηκαν σε μικρό εύρος χρόνου». Η συνεδρία τερματίζεται μετά από 5 λεπτά μη δραστηριοποίησης του χρήστη. Η μέθοδος time out (ορισμός χρονικού ορίου στην λειτουργία) χρησιμοποιείται αρκετά από προμηθευτές εμπορικών βάσεων δεδομένων web, για να συλλέξουν στατιστικές χρηστών, αλλά το μήκος χρόνου ποικίλλει.

Υπάρχουν και άλλες προκλήσεις στην ανάλυση των log files. Τα αρχεία αυτά μπορεί να μην καταγράφουν την πλήρη διεργασία αναζήτησης όταν το caching (κρυφή μνήμη) ή ο proxy server (διακομιστής μεσολάβησης) χρησιμοποιείται. Ακόμη είναι δύσκολος ο διαχωρισμός ερωτημάτων αναζήτησης από ανθρώπους ή από ρομπότ. Οι Silverstein et al. (1999) ανακάλυψαν ότι ερωτήματα από ρομπότ στα logs της AltaVista έδωσαν σειρά περιέργων αποτελεσμάτων, που θα έπρεπε να αφαιρεθούν πριν την ανάλυση. Ακόμη είναι πολύ δύσκολο να καθοριστούν τα δημογραφικά στοιχεία του χρήστη, όπως η γεωγραφική τοποθεσία.

4.4.4 Χρήση δειγματοληπτικών μεθόδων

Η έρευνα στην ολότητα του Web είναι αδύνατη, λόγω του τεραστίου όγκου του, της πολύπλοκης δομής του, αλλά και των πάσης φύσεως τεχνικών δυσκολιών,. Για αυτό τον λόγο, όταν οι επιστήμονες δεν καταφέρνουν να αρκεστούν στη χρησιμοποίηση ενός μικρού αριθμού πηγών, καταφεύγουν στη χρήση δειγματοληπτικών μεθόδων (2001).

Ορισμένες έρευνες που διεξήχθησαν με αυτό τον τρόπο βασίστηκαν στη δημιουργία υποτιθέμενων 'τυχαίων URLs' ή και δειγμάτων από λέξεις σταθμισμένων λεξιλογίων του παγκοσμίου ιστού, ώστε να μετρηθεί η απόκριση που προσφέρουν οι διάφορες μηχανές αναζήτησης. Η εφαρμογή της συνεχούς ανάκτησης όλων των τεκμηρίων που υπάρχουν στο παγκόσμιο ιστό και σχετίζονται με κάποιο ειδικό θέμα ή με κάποια συγκεκριμένη χώρα, δίδει την δυνατότητα στους ερευνητές για τη δημιουργία τυχαίων δειγμάτων από το σύνολο αναφοράς (2002).

4.4.5 Ανάκτηση σελίδων δειγματοληπτικά

Η ανάγκη συλλογής στοιχείων από το παγκόσμιο ιστό με μεθόδους δειγματοληψίας μας οδήγησε στη κατασκευή ειδικών αλγορίθμων που δίδουν την δυνατότητα ανάκτησης ιστοσελίδων, βασιζόμενες στην απαιτούμενη τυχαία επιλογή, σ' αυτές τις περιπτώσεις . Η κύρια προσέγγιση είναι η επιλογή σελίδων να πραγματοποιείται μέσα από ένα, όσο το δυνατόν, μεγαλύτερο αρχικό σύνολο. Ένας πολύ γνωστός τρόπος είναι η φυλλομέτρηση του ερευνητή σε διάφορες ιστοσελίδες, είτε τυχαία είτε ακολουθώντας τις συνδέσεις των ιστοσελίδων που επισκέπτεται. Σε οποιαδήποτε περίπτωση όμως, ο ερευνητής θα πρέπει να επισκεφτεί πολύ περισσότερες ιστοσελίδες από εκείνες που θα επιλέξει στο τέλος (2003).

4.4.6 Ανάκτηση δικτυακών τόπων δειγματοληπτικά

Η μελέτη πολλών δικτυακών τόπων που επιλέγονται μέσα από ένα προκαθορισμένο και συγκεκριμένο μεγάλο σύνολο απαιτεί την εφαρμογή συγκεκριμένης μεθόδου, που εξασφαλίζει την τυχαιότητα της επιλογής. Αν τα URL's είναι συγκεκριμένα, τότε μπορεί να χρησιμοποιηθεί οποιαδήποτε μέθοδος που επιλέγει τυχαίους αριθμούς μέσα από ένα σύνολο, αφού παραδείγματος χάριν έχουν ήδη αριθμηθεί ή μπορεί να γίνει η επιλογή τους με βάση τις IP διευθύνσεις. Υπάρχουν, όμως, και άλλες περιπτώσεις που οι παραπάνω μέθοδοι δεν δύνανται να εφαρμοστούν. Σε κάθε περίπτωση, πάντως, είναι αναγκαία η χρησιμοποίηση της καταλληλότερης μεθόδου που μας θωρακίζει με την απαραίτητη αξιοπιστία, αλλά και προϋποθέτει, όσο γίνεται λιγότερη καταφυγή σε πηγές εκτός του παγκοσμίου ιστού.

4.4.7 Μελέτες σε μεγάλα τμήματα του παγκοσμίου ιστού

Οι μελέτες που προϋποθέτουν την έρευνα σε πολύ μεγάλα τμήματα του παγκοσμίου ιστού μπορούν να κάνουν χρήση των εξής εναλλακτικών λύσεων (2002):

- Χρήση προγράμματος ανίχνευσης του παγκοσμίου ιστού 'crawler' για την ανάκτηση των ιδίων των ιστοσελίδων.
- Πρόσβαση στις αναζητήσιμες όλων των ιστοσελίδων που βρίσκουμε στο αρχείο του διαδικτύου μιας μη κερδοσκοπικής ψηφιακής βιβλιοθήκης 'Internet Archive' (www.archive.org).
- Αναζητήσεις μέσω χρήσης "crawler" στο σύνολο μιας συγκεκριμένης μηχανής αναζήτησης.
- Αναζητήσεις με χρήση σύνθετων ερωτημάτων σε μηχανές αναζήτησης, που ορίζουν επακριβώς τις προς μέτρηση ιστοσελίδες, η περίπτωση αυτή εφαρμόζεται εκεί που δεν είναι αναγκαία η διερεύνηση του πλήρους περιεχόμενου των ιστοσελίδων.

4.5 Ανάλυση Βιβλιομετρική στο παγκόσμιο ιστό

Αφού συλλέξουμε τα δεδομένα, ασχέτως της εφαρμοζόμενης μεθόδου καταγραφής τους, η επόμενη φάση είναι αυτή της ανάλυσής τους. Εδώ χρησιμοποιούμε κυρίως απλές διαδικασίες και συγκεκριμένη στατιστική και μαθηματική διερεύνηση. Σε πολλές άλλες περιπτώσεις εφαρμόζεται η χρήση ειδικών βιβλιομετρικών νόμων, μεθόδων ή μοντέλων που απαιτούνται στην ανάλυση των δεδομένων για το παγκόσμιο ιστό.

4.5.1 Ανάλυση βιβλιογραφικών αναφορών (citation analysis)

Μια από τις κυριότερες μελέτες σε αυτό το πεδίο είχε να κάνει με την αξιολόγηση και ανάλυση των ηλεκτρονικών περιοδικών, δια της καταμέτρησης του αριθμού των βιβλιογραφικών αναφορών τους. Άλλες όμως έρευνες που ακολούθησαν ασχολήθηκαν με τις αναφορές που απέσπασαν οι ηλεκτρονικές πηγές (μηνύματα ηλεκτρονικού ταχυδρομείου, κάθε είδους δημοσιεύσεις και ηλεκτρονικές πηγές διαθέσιμες στο Internet, μηνύματα σε λίστες συζήτησης). Επιπροσθέτως από την αρχή ακόμη οι ερευνητές διέγνωσαν πως σε κάθε ηλεκτρονική πηγή οι συνδέσεις που υφίστανται δικαιολογημένα θα μπορούσαν να είναι το ανάλογο των βιβλιογραφικών αναφορών στα συμβατικά τεκμήρια.

Μια από τις πολλές έρευνες σε αυτό το πεδίο σε συγκεκριμένα ηλεκτρονικά περιοδικά (2003), με τη χρήση της μηχανής αναζήτησης AltaVista, μέτρησε τις συνδέσεις που γίνονταν από το υπόλοιπο του παγκοσμίου ιστού. Τα συμπεράσματα κατέληξαν πως δεν υπήρχαν αρκετές ομοιότητες μεταξύ των συντελεστών απήχησης (impact factors), όπως αυτοί ορίζονται από το Institute for Scientific Information (ISI) και όπως αυτοί προκύπτουν από τη μελέτη των υπερσυνδέσεων στο παγκόσμιο ιστό. Δηλαδή οι υπερσυνδέσεις των ηλεκτρονικών περιοδικών λειτουργούν με διαφορετικό τρόπο απ' ότι οι βιβλιογραφικές αναφορές στο συμβατικό περιβάλλον των τεκμηρίων. Έρευνες όμως που ακολούθησαν, (Vaughan & Hysen, 2002; Thelwall, et Al., 2003) μας έδειξαν πως τελικά υπάρχουν πολλές ομοιότητες και πως αυτό οφείλεται στην ολοκλήρωση που απέκτησε, στο μεσοδιάστημα ο παγκόσμιος ιστός.

Πολλές μηχανές αναζήτησης, όπως οι Google, AltaVista, και Hotbot δίδουν τους Ενιαίους Εντοπιστές Πόρων 'URL's' που αναφέρουν στις βάσεις δεδομένων τους. Αυτό μας παρέχει τη δυνατότητα ελέγχου της παροχής σελίδων που έχουν σύνδεση προς μια συγκεκριμένη σελίδα "backlink", άμεσα σε έναν δικτυακό τόπο και όχι δια μέσου της μηχανής αναζήτησης. Ασφαλώς τεράστια σημασία έχουν οι έρευνες που δεν επικεντρώνονται στις υπερσυνδέσεις ως ανάλογες των παραδοσιακών βιβλιογραφικών αναφορών, αλλά

αναφέρονται στη έρευνα τους διαμέσου του συστήματος, που δημιουργεί αυτόματα ένα ευρετήριο αναφορών της βιβλιογραφίας σε ηλεκτρονική μορφή ‘Autonomous Citation Indexing’ (ACI). Αυτό το ευρετήριο των αναφορών μας δίνει την δυνατότητα στην αναζήτηση συγχρόνως και στα άρθρα και στις αναφορές τους.

4.5.2 Ανάλυση ζευγών λέξεων και αναφορών (co-word and co-citation analysis)

Σε αυτήν την περίπτωση ελέγχεται η περίπτωση εφαρμογής στην πράξη των τεχνικών των ζευγών αναφορών ή λέξεων στο περιβάλλον του παγκόσμιου ιστού, στην περίπτωση που οι συνδέσεις θεωρούνται ως αναφορές, για την αξιολόγηση και τον εντοπισμό παρόμοιων ή συναφών ιστοσελίδων. Αυτό πραγματοποιείται με τη χρήση εμπορικών μηχανών αναζήτησης και των δεδομένων που ανακτώνται από την σχετική επεξεργασία. Επίσης πραγματοποιείται μελέτη των πολυδιάστατων κλιμάκων (MDS) και των πινάκων συσχέτισης (correlation matrix) μέσω της δημιουργίας ενός χάρτη αντιστοίχισης ζευγών αναφορών (co-citation map). Ένας εναλλακτικός τρόπος μελέτης τους είναι η επεξεργασία ειδικών αλγόριθμων.

4.5.3 Ανάλυση περιεχομένου (content analysis)

Η ανάλυση περιεχομένου των πηγών του παγκοσμίου ιστού και του διαδικτύου χρησιμοποιούνται για την δυνατότητα εμπέδωσης του περιεχομένου τους. Στις πολλές και διαφορετικές έρευνες που έχουν γίνει, έχουν αναλυθεί τα περιεχόμενα των ιστοσελίδων, μηνυμάτων λιστών συζήτησης, και δικτυακών τόπων. Στις έρευνες αυτές εξετάζεται η ωφελιμότητα του περιεχομένου των ανωτέρω πεδίων, οι μέθοδοι επικοινωνίας και διάδρασης μεταξύ των συμμετεχόντων, η σχετικότητα των αναφορών, τα χαρακτηριστικά των περιεχομένων, η δομή τους και άλλα σχετικά.

Η ανάλυση περιεχομένου και η απλή μέτρηση των ιστοσελίδων προσομοιάζει αρκετά με την ανάλυση των παραδοσιακών τεκμηρίων, αφού οι συνδέσεις προς και από μια ιστοσελίδα μπορούν να σχετισθούν και να καταμετρηθούν όπως οι συμβατικές αναφορές. Όμως ο παγκόσμιος ιστός, λόγω της φύσης του και της δυναμικής του, συχνά παρέχει σελίδες που έχουν δημιουργήσει ταυτόχρονες συνδέσεις μεταξύ τους, κάτι που δε θα μπορούσε να συμβεί στα έντυπα εκ των πραγμάτων. Η κάλυψη από τις μηχανές αναζήτησης της ολότητας του παγκόσμιου ιστού μπορεί να διερευνηθεί, ακριβώς όπως η κάλυψη όλων των επιστημονικών περιοχών από τις βάσεις δεδομένων των βιβλιογραφικών αναφορών. Μια άλλη σημαντική αντίθεση της εφαρμογής βιβλιομετρικών μεθόδων στο παγκόσμιο ιστό είναι πως, ενώ στα παραδοσιακά επιστημονικά δημοσιεύματα υπάρχει μια εκ των προτέρων αξιολόγησή τους, αυτό στο παγκόσμιο ιστό εμφανίζεται ελάχιστα, με συνέπεια να μειώνεται αυτόματα και η ποιότητα του περιεχομένου του.

Ένα πολύ βασικό χαρακτηριστικό του παγκόσμιου ιστού που το διαφοροποιεί από τα παραδοσιακά περιβάλλοντα δημοσίευσης είναι ο χρόνος. Ο χρόνος στο παγκόσμιο ιστό παρέχει μια ξεχωριστή διάσταση στη δημοσίευση, αφού αφ’ ενός σε πολύ λίγο χρόνο είναι δυνατόν κανείς να δημοσιεύσει την έρευνα του, αλλά και, πολύ συχνά, εργασία που έχει καταχωρηθεί, παύει να υπάρχει.

4.5.4 Αξιολογώντας με τη χρησιμοποίηση είτε υπαρχόντων ή νέων μεθόδων

Σε αυτό το στάδιο αξιολογείται η αποτελεσματικότητα της σχετικότητας και της ακρίβειας της ανάκτησης πληροφοριών από τις μηχανές αναζήτησης στο παγκόσμιο ιστό, όπως αυτές αξιολογούνται με την παραδοσιακή τεχνική από τους χρήστες, που έκαναν την αναζήτηση των δημοσιεύσεων. Κρίνονται, δηλαδή, οι μηχανές αναζήτησης αυτές καθαυτές. Σε αυτήν την προσπάθεια προστίθενται, επιπλέον, και νέοι τρόποι καταμέτρησής τους, όπως η μέθοδος των Page Narin “PageRank”, που αν και βασίζεται στην κλασική μέθοδο της ανάλυσης των αναφορών, αντί να αθροίζει μόνο τον αριθμό των συνδέσεων προς ένα συγκεκριμένο τεκμήριο, συνυπολογίζει και την ποιότητα του τεκμηρίου το οποίο προσφέρει αυτή τη σύνδεση. Επιπλέον προτείνεται η αξιολόγηση της ποιότητας των μηχανών αναζήτησης, που ενδέχεται να βασίζεται στις συνδέσεις που γίνονται προς αυτήν είτε στον αριθμό των επισκεπτών της και τη θεματική ενότητα, από την οποία αυτοί προέρχονται.

4.5.6 Υπολογισμός και επεξήγηση των δεικτών

Στα έντυπα περιοδικά για να υπολογιστεί η απήχηση που έχει ένα συγκεκριμένο περιοδικό, άρα και η φήμη και ενδεχομένως η ποιότητά του, υπολογίζονται οι συντελεστές απήχησης ‘impact factors’. Ορίζονται σχετικοί συγκεκριμένοι δείκτες για το Web. Ο Ingwerson, το 1998, προσδιόρισε το δείκτη απήχησης παγκοσμίου ιστού ‘Web Impact Factor’ (WIF) ως ακολούθως σύμφωνα με τους Bar και λοιποί (2001):

- Ανάλογα με τις ιστοσελίδες που έχουν σύνδεσμο σ' αυτόν τον δικτυακό τόπο ή στη συγκεκριμένη χώρα
- Ανάλογα με τον αριθμό των ιστοσελίδων του συγκεκριμένου δικτυακού τόπου ή της χώρας

Ακριβώς όπως και ο κλασικός συντελεστής απήχησης των περιοδικών του WoS, ο δείκτης απήχησης παγκοσμίου ιστού ενός συγκεκριμένου δικτυακού τόπου ή μιας χώρας δείχνει την ένταση της δυνατότητας εύρεσης και χρήσης του στο παγκόσμιο ιστό. Ο συντελεστής απήχησης ενός δικτυακού τόπου είναι το κλάσμα με αριθμητή τον αριθμό των συνδέσεων που παρέχονται προς ένα δικτυακό τόπο, και παρανομαστή τον αριθμό των ιστοσελίδων αυτού του δικτυακού τόπου. Μάλιστα ο Rousseau (1997) χρησιμοποιεί τον όρο που παραπέμπει στην αναφορά το "sitiation", και ο οποίος ασχολείται με την ιστομετρική ανάλυση των αναφορών και αναλύει τα επαναλαμβανόμενα χαρακτηριστικά κατανομής των δικτυακών τόπων και τις εισερχόμενες συνδέσεις. Όσον αφορά στη χρήση των WIF's, έχουν γίνει οι παρακάτω παρατηρήσεις σύμφωνα με τους Bjorneborn & Ingwersen (2001):

- Όπως γίνεται και με τις αναφορές των εντύπων τεκμηρίων, οι συνδέσεις των δικτυακών τόπων μπορεί να πραγματοποιούνται για διάφορους λόγους. Έτσι λοιπόν, οι συνδέσεις είναι μία ατελής προσέγγιση της γενικής απήχησης ενός δικτυακού χώρου σε ολόκληρο το παγκόσμιο ιστό.
- Οι συμβατικοί συντελεστές απήχησης βασίζονται στην έννοια του τεκμηρίου, το οποίο είναι μία αυτοτελής οντότητα. Δεν συμβαίνει το ίδιο όμως με μια ιστοσελίδα, η οποία, δεν καθορίζεται τόσο εύκολα και διαφέρει ανάλογα με τα πεδία που περιέχουν διάφορες πληροφορίες.
- Ο Ενιαίος Εντοπιστής Πόρων (URL) δεν δείχνει απαραίτητα τη θέση μιας ιστοσελίδας.
- Οι βάσεις δεδομένων των μηχανών αναζήτησης δεν είναι πάντα άμεσα ενημερωμένες.

4.6 Μοντέλα βιβλιομετρικής ανάλυσης στο Παγκόσμιο Ιστό

Ο Rousseau (1997), απέδειξε πως ο βιβλιομετρικοί νόμοι είναι εφαρμόσιμοι και στο Παγκόσμιο Ιστό. Προσάρμοσε δε τον νόμο του Lotka και στον αριθμό των αναφορών σε ένα δικτυακό τόπο και στον αριθμό των δεδομένων των ανακτώμενων σελίδων σε κάθε δικτυακό τόπο. Απέδειξε επίσης πως ο νόμος του Bradford εφαρμόζεται και στον προσδιορισμό βασικών ομάδων πληροφόρησης π.χ. blogs. Ο Hubert και άλλοι έδειξαν πως η συμπεριφορά των χρηστών του Web ακολουθεί τον νόμο της διασποράς του Zipf.

4.6.1 Τοπολογικά μοντέλα και εξόρυξη στο Παγκόσμιο Ιστό (web mining)

Στη θεωρία των γράφων, ένα γράφημα είναι μια μαθηματική περιγραφή ενός δικτύου- Ειδικότερα ο όρος δίκτυο μπορεί να αναφέρεται σε οποιοδήποτε αλληλοσυνδεδεμένη ομάδα ή σύστημα-που αποτελείται από ένα σύνολο V κορυφών (vertices), ή σημείων, ή κόμβων, και ένα σύνολο E ακμών (edges), ή τόξων, ή γραμμών. Μια ακμή (σχέση), είναι ένα ζεύγος (u, v) από κορυφές. Παραδείγματος χάριν οι καθηγητές του Ιονίου Πανεπιστημίου που αντιστοιχούνται από τους συνδεδεμένους ή μη κόμβους (nodes), που συνδέονται ή όχι μεταξύ τους με γραμμές. σε σχέσεις (relations) π.χ. συνεργασίας κάθε καθηγητή. Οι κόμβοι αυτοί μπορεί να είναι άνθρωποι (όπως το παραπάνω παράδειγμα κοινωνικού δικτύου), τεκμήρια (συνδεδεμένα σε ένα δίκτυο αναφορών) ή έννοιες (συνδεδεμένες σε έναν θησαυρό). Σύμφωνα με τη θεωρία γραφημάτων, το διαδίκτυο παρίσταται σαν ένα κατευθυνόμενο γράφημα, στο οποίο οι κόμβοι αποτελούν τις ιστοσελίδες, ενώ οι ακμές αποτελούν τις διασυνδέσεις (links) μεταξύ τους. Έτσι, αν μία ιστοσελίδα συνδέεται με ένα link με μίαν άλλη, τότε στο γράφημα θα έχω ένα βέλος (ακμή) από τον ένα κόμβο, που αναπαριστά τη μια ιστοσελίδα, προς τον άλλο κόμβο που παριστά τη δεύτερη ιστοσελίδα. (2001).

4.6.2 Μοντέλα σύνδεσης

Ένα πολύ ενδιαφέρον στοιχείο στην έρευνα για την κατανόηση και την αξιολόγηση του παγκόσμιου ιστού αποτελούν τα δεδομένα και ο μηχανισμός που ενεργοποιείται στην φορά και την δυναμική που δημιουργούν οι συνδέσεις. Η έρευνα, για να συμπεράνει συγκεκριμένα ποσοτικά χαρακτηριστικά, καταγράφει: α) τον αριθμό των συνδέσεων προς και από ένα δικτυακό τόπο, β) τις ιστοσελίδες που διαθέτει, γ) τις επισκέψεις που

δέχεται ή δ) τις σελίδες που δέχονται επισκέψεις σε ένα δικτυακό τόπο. Συμπερασματικά έχει διαπιστωθεί πως μία ιδιότητα που ισχύει είναι αυτό που αποκαλείται «συνεχής ανάπτυξη και η προνομιακή σύνδεση», δηλαδή ο πλούσιος δικτυακός τόπος έχει την τάση να γίνεται πλουσιότερος με άλλα λόγια η γνωστή μας από το 1^ο κεφάλαιο αρχή του σωρευτικού πλεονεκτήματος ή διαφορετικά η κατά Ματθαίον επίδρασης (2003).

4.6.3 Δίκτυα “Small-Worlds”

Πολλά συστήματα, είτε φυσικά (π.χ. δίκτυα νευρώνων σε έναν οργανισμό) είτε κατασκευασμένα από ανθρώπους (π.χ. δίκτυο ηλεκτροδότησης μιας πόλης ή ακόμα και τα διάφορα κοινωνικά δίκτυα), τα οποία μπορούν να μοντελοποιηθούν μέσω της θεωρίας γράφων, έχει παρατηρηθεί ότι διέπονται από το «small-world» φαινόμενο, δηλαδή είναι γραφήματα που παρουσιάζουν αρκετά ισχυρές ομαδοποιήσεις κόμβων (οι οποίοι π.χ. έχουν όμοιες ιδιότητες) των οποίων οι αποστάσεις μεταξύ δύο τυχαίων κόμβων δεν είναι μεγάλες.

Τι είναι, όμως, αυτό το φαινόμενο και πώς οι επιστήμονες άρχισαν να αντιλαμβάνονται ότι διέπει πολλά συστήματα που εμφανίζονται στην καθημερινή ζωή;

Στα τέλη της δεκαετίας του '60, έγινε ένα πείραμα στο πανεπιστήμιο του Χάρβαρντ, από τον Stanley Milgram. Πήρε κάποια γράμματα που προορίζονταν για ένα γνωστό του χρηματιστή στη Βοστώνη και τα διένειμε σε ένα τυχαίο πλήθος ανθρώπων σε μια άλλη πόλη. Οι οδηγίες ήταν ότι τα γράμματα έπρεπε να σταλούν στον τελικό παραλήπτη τους, τον χρηματιστή, περνώντας από το ένα άτομο στο άλλο, ενώ επιπλέον έπρεπε ο κάθε αποστολέας να γνωρίζει τον (ενδιάμεσο) παραλήπτη, στον οποίο θα έστελνε το γράμμα, χρησιμοποιώντας το μικρό του όνομα. Από τη στιγμή λοιπόν που ο τυχαίος, αρχικός αποστολέας ήταν σχεδόν βέβαιο ότι δεν ήξερε προσωπικά τον χρηματιστή, η καλύτερη επιλογή που είχε ήταν να στείλει το γράμμα σε κάποιον που γνώριζε και που νόμιζε ότι εκείνος θα ήταν «πιο κοντά» στο να ξέρει τον χρηματιστή. Αυτή η διαδικασία, προφανώς, θα επαναλαμβανόταν, έως ότου φτάσουν τα γράμματα στον τελικό τους προορισμό.

Ένας λογικός αριθμός γραμμάτων έφτασε στον χρηματιστή κι έτσι ο Milgram μπόρεσε να διαπιστώσει ότι χρειάστηκε ένας μέσος όρος περίπου έξι βημάτων από τον αρχικό αποστολέα μέχρι τον τελικό παραλήπτη. Αυτό έκανε τον Milgram να γενικεύσει (ίσως λίγο άκομψα) το αποτέλεσμα του πειράματος στο γεγονός ότι ανάμεσα σε δύο τυχαία άτομα, οπουδήποτε στον κόσμο, παρεμβάλλονται κατά μέσο όρο έξι «γνωριμίες» που τα χωρίζουν. Ομολογουμένως πολύ μικρός αριθμός, αν λάβουμε υπόψη το συνολικό πληθυσμό της γης. Από την άλλη, βέβαια, αν υποθεθεί ότι ο καθένας έχει κατά μέσο όρο γύρω στις εκατό γνωριμίες (μια ματιά στον κατάλογο των επαφών του κινητού τηλεφώνου είναι αρκετή για να πεισθεί κανείς γι' αυτό) και ότι ο συνολικός πληθυσμός της γης είναι περίπου στα 7 δις (<http://www.census.gov/ipc/www/popclockworld.html>) είναι κατανοητό ότι η παραπάνω γενίκευση είναι σε μεγάλο βαθμό βάσιμη, αφού αποδεικνύεται μαθηματικά.

Αυτή η διαπίστωση ονομάστηκε «six degrees of separation» και μεταγενέστερα παρατηρήθηκε και σε άλλες περιπτώσεις κοινωνικών δικτύων (social networks). Έτσι, στον χώρο των ηθοποιών του Χόλυγουντ, η ορολογία αυτή παραφράστηκε ως «six degrees of Kevin Bacon». Στην περίπτωση αυτή, ξεκινώντας από το γνωστό ηθοποιό, ψάχνουμε να βρούμε πόσο «απέχει» ένας τυχαίος ηθοποιός από τον Bacon. Με τον όρο «απέχει», εννοείται η σύνδεση του τυχαίου ηθοποιού μέσω άλλων, οι οποίοι έχουν συμπρωταγωνιστήσει σε ταινίες μαζί του. Έτσι, ο ηθοποιός Α έχει βαθμό ένα, αν έχει παίξει σε μια ταινία μαζί με τον Bacon, ενώ ένας άλλος ηθοποιός Β έχει βαθμό δύο, αν έχει παίξει σε ταινία με έναν ηθοποιό, ο οποίος με τη σειρά του έχει συμπρωταγωνιστήσει με τον Bacon κ.ο.κ.

Μια παραλλαγή της ίδιας φιλοσοφίας είναι και ο βαθμός Erdős. Ο Erdős ήταν ένας πολυγραφότατος μαθηματικός, ο οποίος είχε δημοσιεύσει αρκετά συχνά μαζί με άλλους επιστήμονες κι από άλλους τομείς, πολλές ακαδημαϊκές εργασίες. Έτσι, ένας επιστήμονας έχει αριθμό Erdős ένα, αν έχει δημοσιεύσει κάποιο άρθρο κατ' ευθείαν μαζί με τον Erdős, ενώ έχει βαθμό δύο, αν έχει δημοσιεύσει άρθρο μαζί με έναν επιστήμονα, ο οποίος με τη σειρά του έχει δημοσιεύσει μαζί με τον Erdős ένα άρθρο κ.ο.κ.

Και στις δύο αυτές δημοφιλείς περιπτώσεις παρατηρείται το φαινόμενο που προέκυψε και στο πείραμα του Milgram. Οποιοσδήποτε τυχαίος ηθοποιός (ή επιστήμονας στη δεύτερη περίπτωση) απέχει, κατά μέσο όρο, έξι ενδιάμεσους «σταθμούς» από το αρχικό άτομο.

Μια τέτοια γενίκευση, βέβαια, δεν μπορεί να θεωρηθεί σε απόλυτο βαθμό αξιόπιστη, διότι ενδέχεται σε κάποιες άλλες περιπτώσεις κοινωνικών ή άλλων δικτύων να παρεκκλίνουν λίγο τα αποτελέσματα ή το μονοπάτι ανάμεσα από δύο τυχαίους κόμβους (με ό,τι αντιπροσωπεύει ο όρος κόμβος) να είναι κάπως μεγαλύτερο. Εν τούτοις, και οι τρεις περιπτώσεις μας δίνουν να καταλάβουμε ότι δύο τυχαία επιλεγμένα

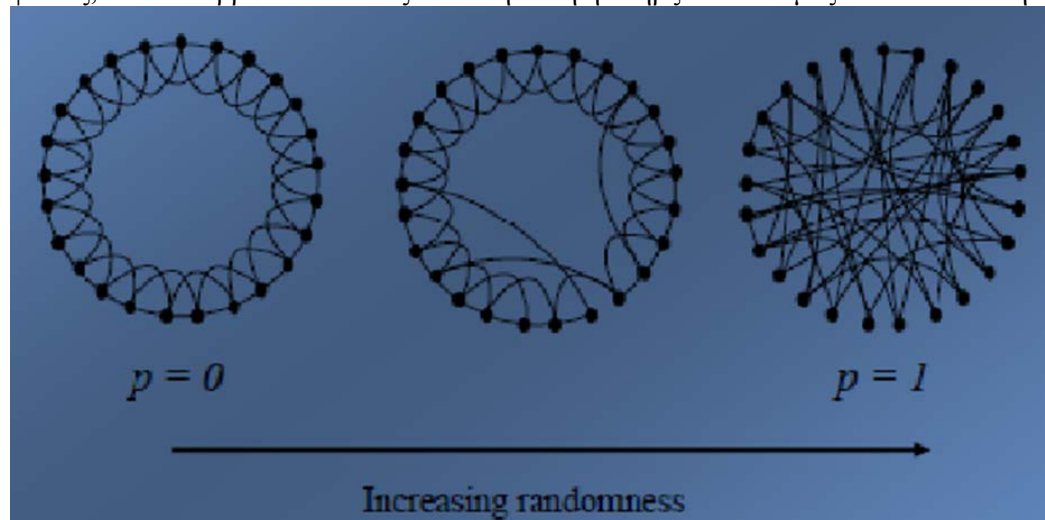
άτομα μπορούν να συνδεθούν με μια σχετικά μικρή αλυσίδα ενδιάμεσων «σταθμών», παρά τον τεράστιο αριθμό του συνολικού πληθυσμού της γης. Αυτό το φαινόμενο, λοιπόν, είναι γνωστό ως «small-world effect».

Με σκοπό την πιο αυστηρή εξήγηση του φαινομένου αυτού, οι επιστήμονες άρχισαν να το μοντελοποιούν βάσει της θεωρίας γράφων. Σκοπός είναι να κατασκευαστούν γραφήματα (ή να ευρεθούν ήδη υπάρχοντα), των οποίων οι κόμβοι να είναι σε πρώτη φάση αρκετά ισχυρά ομαδοποιημένοι, ενώ από την άλλη το μονοπάτι ανάμεσα σε δύο τυχαίους κόμβους να είναι σχετικά μικρό. Έτσι, τα ως άνω παραδείγματα που αναφέρθηκαν μπορούν να μοντελοποιηθούν με γραφήματα και ως εκ τούτου να παραχθούν: όχι μόνο το δίκτυο συνεργασίας ηθοποιών (six degrees of Kevin bacon), αλλά και το δίκτυο συνεργασίας επιστημόνων (αριθμός Erdos) και άλλα, όπως το ηλεκτρικό δίκτυο μιας πόλης, ένα δίκτυο κινητής τηλεφωνίας ή ένα σιδηροδρομικό δίκτυο, το δίκτυο νευρώνων του οργανισμού, όπως επίσης και αυτό το διαδίκτυο.

Μία πρώτη προσέγγιση είναι η υλοποίηση ενός τυχαίου γράφου. Εδώ οι συνδέσεις είναι εντελώς τυχαίες ανάμεσα στους κόμβους. Αυτό απ' ενός σημαίνει ότι το μονοπάτι ανάμεσα σε ένα τυχαίο ζευγάρι κόμβων είναι μικρό, καθώς είναι πολύ πιθανό να υπάρχει σύνδεση ανάμεσα σε δύο απομακρυσμένους κόμβους, απ' ετέρου όμως, έχουμε απουσία ομαδοποίησης κόμβων (με τον όρο ομαδοποίηση εννοείται το γεγονός ότι αν δύο κόμβοι είναι γειτονικοί με έναν άλλον, τότε είναι και μεταξύ τους ή όπως χαρακτηριστικά λέγεται, δύο φίλοι ενός ατόμου μπορεί να είναι και μεταξύ τους φίλοι κι αυτό λόγω πάλι της τυχειότητας των συνδέσεων, οι οποίες δεν έχουν και κάποιο βάρος ή σημασία για τους συνδεόμενους κόμβους). Αυτή η υλοποίηση, όπως είναι φανερό, δεν αντιπροσωπεύει πλήρως τα small-world γραφήματα, καθώς λείπει η ομαδοποίηση των κόμβων.

Από την άλλη, υπάρχει και η υλοποίηση ενός πλήρως κανονικοποιημένου γραφήματος. Εδώ υπάρχει ισχυρή ομαδοποίηση κόμβων, αλλά τα μήκη των μονοπατιών μεταξύ δύο τυχαίων κόμβων είναι συνήθως αρκετά μεγάλα. Φανερό είναι πάλι ότι ούτε και αυτή η περίπτωση είναι πλήρως αντιπροσωπευτική. Τα small-world γραφήματα μπορούμε να πούμε ότι βρίσκονται κάπου ενδιάμεσα στις δύο «ακραίες» περιπτώσεις.

Οι Watts και Strogatz (1998) παρουσίασαν ένα μοντέλο για τα small-world γραφήματα, το οποίο ίσως είναι πιο κοντά στην κοινή διαίσθηση που υπάρχει γι' αυτά. Πρόκειται για ένα κανονικοποιημένο γράφημα, το οποίο όμως έχει και ένα βαθμό τυχειότητας. Το γράφημα αυτό μπορεί να δημιουργηθεί ως εξής: ξεκινώντας από ένα πλήρως κανονικοποιημένο γράφημα, αλλάζουμε τυχαία μερικές συνδέσεις κάποιων κόμβων, με σκοπό κυρίως την ένωση κάποιων απομακρυσμένων κόμβων (που θα παίζουν το ρόλο της «γέφυρας» μεταξύ μιας ομάδας κόμβων με μια άλλη). Έτσι, παραμένει, κατά μέσο όρο, η ισχυρή ομαδοποίηση που υπάρχει στα πλήρως κανονικοποιημένα γραφήματα, αλλά με αυτές τις καινούριες συνδέσεις επιτυγχάνεται και η συντόμευση των μονοπατιών, φαινόμενο που παρατηρείται στα τυχαία γραφήματα. Ο συνδυασμός αυτών των δύο μας δίνει, εν τέλει, ένα small-world γράφημα. Η έννοια αυτή της «γέφυρας» μπορεί εύκολα να ερμηνευθεί από τη σκοπιά των κοινωνικών δικτύων. Πολλοί από μας, εκτός από τους κοντινούς φίλους (και με τον όρο «φίλος» μπορεί να εννοηθεί ο κάτοικος του γειτονικού σπιτιού, ο συνάδελφος στη δουλειά ή το άτομο που μας έχει συστήσει ένας φίλος μας) έχουμε και κάποιους άλλους φίλους, οι οποίοι βρίσκονται ίσως σε άλλη πόλη ή υπήρξαν κοντά μας σε κάποια άλλη περίοδο της ζωής μας.



Εικόνα 4 35 Πηγή: Small world, DJWatts, SHStrogatz – Nature, 1998

[http://cssp.us/pdf/D_Watts_Six_Degrees_CSSP.pdf] Η εικόνα παρουσιάζει παραστατικά τις δύο ακραίες περιπτώσεις γραφημάτων και ενδιάμεσα ένα small-world γράφημα.

Μια άλλη προσέγγιση, παρόμοια με αυτή των Watts και Strogatz, των Newman και Watts (1999) ξεκινά πάλι από ένα κανονικοποιημένο γράφημα, αλλά αντί να αλλάξει κάποιες συνδέσεις, προσθέτει επιπλέον συνδέσεις σε τυχαία επιλεγμένους κόμβους, για να πετύχει τη συντόμευση των μονοπατιών.

Η όλη φιλολογία γύρω από αυτό το φαινόμενο θα ήταν μάλλον ανούσια, αν δεν γινόταν λόγος για τα οφέλη που ενδέχεται να έχει μια τέτοια προσέγγιση των, πάσης φύσεως, δικτύων. Και πρώτα πρώτα στην επικοινωνία μεταξύ των ατόμων. Η οποιαδήποτε ανταλλαγή πληροφορίας, όπως ειδήσεις, αστεία, φήμες κ.ά. λαμβάνει χώρα ανάμεσα σε μεμονωμένες οντότητες που παίζουν το ρόλο των ενδιαμέσων σταθμών. Έτσι, μια πληροφορία μπορεί να διαδοθεί ταχύτερα σε ένα δίκτυο, όπου το μέγιστο μονοπάτι ανάμεσα σε δύο τυχαίους κόμβους του (σε κοινωνικά δίκτυα είναι τα άτομα) είναι έξι ενδιάμεσοι σταθμοί, παρά σε ένα άλλο, όπου αυτός ο αριθμός είναι πολύ μεγαλύτερος.

Μία άλλη, ίσως πιο κρίσιμη περίπτωση, είναι η εξάπλωση των ασθενειών, οι οποίες διαδίδονται κι αυτές από την επαφή ατόμου με άτομο (κόμβοι δικτύου). Είναι, λοιπόν, σημαντικό να γνωρίζει ο όποιος ερευνητής, αν το δίκτυο των ατόμων διέπεται από το small-world φαινόμενο, για να μπορεί να βγάλει συμπεράσματα για την ταχύτητα εξάπλωσης.

Επίσης τα αποτελέσματα γύρω από το small-world φαινόμενο μπορούν να φανούν χρήσιμα και στα δίκτυα κινητής τηλεφωνίας, όπου αντί να δημιουργούνται καινούριοι σταθμοί αναμετάδοσης, αρκεί να αναδιατάσσονται κάποιες συνδέσεις κόμβων του δικτύου, με σκοπό τη δημιουργία ενός small-world δικτύου, το οποίο θα έχει καλύτερη συμπεριφορά σε σχέση με το ήδη υπάρχον. Μια παραλλαγή αυτού, βέβαια, είναι και το ίδιο το διαδίκτυο, όπου μια τέτοια αναδιάταξη μπορεί να είναι καταλυτική στη γρηγορότερη αποστολή e-mails ή στη γρηγορότερη εύρεση μιας ιστοσελίδας ή μιας πληροφορίας γενικότερα.

Όσον αφορά στο διαδίκτυο, μια ακαδημαϊκή εργασία έχει δείξει ότι το small-world φαινόμενο μπορεί να παίξει σημαντικό ρόλο στη δημιουργία μιας πιο αποδοτικής μηχανής αναζήτησης.

Σύνηθες είναι το φαινόμενο ιστοσελίδες με συναφές περιεχόμενο, εκτός από τις συνδέσεις που έχουν μεταξύ τους (δεδομένου του περιεχομένου τους), να έχουν και συνδέσεις με κάποιες άλλες σελίδες διαφορετικού περιεχομένου. Αυτό είναι ίσως καταλυτικό για να μελετηθεί το διαδίκτυο από τη σκοπιά του small-world φαινομένου, καθώς τα δύο βασικά χαρακτηριστικά του (ομαδοποίηση και σύντομα μονοπάτια) υπάρχουν και σ' αυτό το γράφημα.

Μια παρατήρηση γύρω από τη φύση των συνδέσεων των ιστοσελίδων είναι ότι οι υπερσύνδεσμοι αποτελούν και προτίμηση προς τη σελίδα που δείχνουν (κάτι που έχει αναφερθεί και σε προηγούμενο κεφάλαιο). Έτσι, μια καλή σελίδα είναι πολύ πιθανόν να δείχνει σε μια άλλη καλή σελίδα, συναφούς μάλιστα περιεχομένου. Με αυτόν τον τρόπο δημιουργούνται ομάδες ποιοτικών ιστοσελίδων σχετικές με κάποιο θέμα. Μέσα σε μια τέτοια ομάδα υπάρχουν και κάποιες σελίδες που αποτελούν καλά «κεντροειδή» (centers), με την έννοια ότι η απόστασή τους από τις άλλες σελίδες είναι κατά μέσο όρο ελάχιστη (απ' ότι η απόσταση δύο άλλων, τυχαίων ιστοσελίδων της ομάδας).

Τέτοιες ομάδες μπορούν να θεωρηθούν και τα SCCs (Strongly Connected Components) του γράφου του διαδικτύου. Τα SCCs είναι κομμάτια του γράφου που αποτελούνται από κόμβους, για τα οποία ισχύει ότι ανάμεσα σε δύο τυχαίους κόμβους υπάρχει μονοπάτι από τον έναν στον άλλο, αλλά και από τον δεύτερο προς τον πρώτο. Ένα SCC συνήθως περιέχει σελίδες που ανήκουν στον ίδιο ιστότοπο, αλλά υπάρχουν και SCCs που συνδέουν διαφορετικούς ιστοτόπους. Μια μηχανή αναζήτησης θα πρέπει να δίνει περισσότερη βαρύτητα σε τέτοια SCCs, καθώς συνδέσεις μεταξύ διαφορετικών ιστοτόπων δίνουν καλύτερη πληροφορία για τη σημαντικότητα τους, παρά συνδέσεις ανάμεσα σε σελίδες του ίδιου ιστοτόπου.

Έτσι, αντί η μηχανή να παρουσιάζει μια λίστα από σελίδες που βρίσκονται στον ίδιο ιστότοπο, θα μπορούσε να παρουσιάζει μόνο το κεντροειδές από κάθε SCC. Άρα θα μπορούσαν να ελαχιστοποιηθούν τα αποτελέσματα που παρουσιάζονται στον χρήστη (οπότε ο χρήστης να μην κουράζεται ψάχνοντας πολλά αποτελέσματα σε μια μεγάλη λίστα), χωρίς όμως να μειώνεται η κάλυψη του θέματος που τον αφορά. Ένα κεντροειδές, βέβαια, ενδέχεται να μην είναι η σελίδα που αρμόζει τέλεια στην αναζήτηση του χρήστη. Η ιδιότητα, όμως, που αυτό έχει (η ελάχιστη κατά μέσο όρο απόσταση από όλες τις άλλες σελίδες της ίδιας ομάδας) δίνει τη δυνατότητα στο χρήστη να φτάσει στην επιθυμητή σελίδα ίσως και με ένα μόνο βήμα.

Επίσης άτυπες παρατηρήσεις δείχνουν ότι σελίδες που ανήκουν σε μικρά SCCs είτε τείνουν να εστιάζουν σε περιορισμένης έκτασης θέματα, είτε δεν έχουν πολλά inlinks ή outlinks (ειδικά για τα inlinks, αυτό υποδηλώνει ότι ίσως δεν είναι και τόσο σημαντικές). Έτσι, αν η μηχανή λάβει υπόψη της και το μέγεθος του κάθε SCC που ανήκει το κεντροειδές, μπορεί να τα ταξινομεί σύμφωνα με αυτό το μέγεθος, παρουσιάζοντας στον χρήστη πρώτα τα πιο-σύμφωνα με αυτή τη θεώρηση-σημαντικά.

4.6.4 Ανάλυση ακολουθούμενης διαδρομής (path analysis)

Με τη μελέτη των ακολουθούμενων διαδρομών από τον ερευνητή για την μελέτη των διασυνδεδεμένων ιστοσελίδων, βιβλιογραφικών αναφορών ή δικτυακών τόπων σε κάποια βάση δεδομένων, μπορεί να διερευνηθεί η τακτική με την οποία κάποιος μπορεί να ξεκινήσει από ένα θεματικό πεδίο και να οδηγηθεί σε ένα άλλο, π.χ. να ξεκινήσει από τη Βοτανολογία και να καταλήξει στην Φαρμακολογία για την αντιμετώπιση της χοληστερίνης. Η προσέγγιση αυτή αναδεικνύει δεδομένα που επιτρέπουν το σχεδιασμό της συμπεριφοράς των ερευνητών και τις διασυνδέσεις μεταξύ συγγενών αλλά και διαφορετικών θεματικών περιοχών, δεδομένα που βοηθούν στην διασαφήνιση των επιστημονικών πεδίων και στην ανακάλυψη νέων περιοχών επιστημονικής έρευνας (2001).

4.6.5 Τοπολογική μοντελοποίηση και εξόρυξη του Web

Οι Φυσικοί και οι επιστήμονες της πληροφορίας προσπάθησαν να οικοδομήσουν μοντέλα του ιστού μέσω των συνδέσεων, για να εξαχθούν αφηρημένα μοτίβα. Όλες αυτές οι προσεγγίσεις μπορούν να ονομαστούν τοπολογικές, επειδή αντιμετωπίζουν τον ιστό σαν ένα μαθηματικό γράφημα, ενώ παραβλέπουν χωρικές σχέσεις και τα περιεχόμενα της σελίδας. Ο ιστός είναι το μεγαλύτερο δίκτυο για την όποια τοπολογική πληροφορία είναι διαθέσιμη και το Web, με τον τρόπο αυτό, έχει καταστεί ένα πεδίο δοκιμών για πολλές προσπάθειες μοντελοποίησης (2002).

Το θεωρητικό γράφημα έχει χρησιμοποιηθεί στη βιβλιομετρία από τη δεκαετία του 1960, για την ανάλυση δικτύων αναφορών και άλλων δικτύων πληροφοριών (Garner, 1967; Nance, Korfhage & Bhan, 1972; Hummon & Doreian, 1989; Durner, Ellis & Willett, 1996).

Δυο αλγόριθμοι κλειδιά έχουν χρησιμεύσει για να υπογραμμιστεί το δυναμικό υπερ-σύνδεσης της δομής: η ανάλυση του Kleinberg (1999) που δημιούργησε τον αλγόριθμο HITS (Hypertext-Induced topic search) και ο αλγόριθμος PageRank της Google (1998). Ένας λιγότερο γνωστός αλγόριθμος, ο Constrained Spreading Activation (CSA) (1987) είχε επίσης αντίκτυπο στην επιστήμη των υπολογιστών, όταν εφαρμόζεται σε συνδέσμους. Σκοπός της παραμέτρου HITS είναι να αναγνωρίσει τις πιο χρήσιμες ιστοσελίδες για κάθε συγκεκριμένο θέμα. Ο αλγόριθμος προσδιορίζει σχετικές σελίδες αντιστοιχίζοντας κείμενο, αλλά βρίσκει την περισσότερη σχετική, μέσω μιας επαναληπτικής σύνδεσης με βάση τον αλγόριθμο αυτό. Σε γενικές γραμμές, επιχειρεί να βρει «αρχές», με πολλούς εσωτερικούς συνδέσμους από το θέμα της και πολλούς «διανομείς», σελίδες με πολλούς εξωτερικούς συνδέσμους για το θέμα των σελίδων. Καλές «αρχές» και «διανομείς» είναι και τα δύο επιθυμητά.

Στον PageRank έχουμε αναφερθεί στα 'Μοντέλα σύνθετων αναζητήσεων του ιστού'.

4.6.6 Μέτρηση του Web 2.0

Το Web 2.0 είναι ένας όρος που επινοήθηκε από τον εκδότη Tim O'Reilly, κυρίως, για να αναφερθεί στους ιστότοπους που κατευθύνονται σε ένα περιεχόμενο που ενδιαφέρει τους καταναλωτές, όπως τα blogs, η Wikipedia και οι ιστότοποι κοινωνικής δικτύωσης. Η αύξηση του όγκου του διαδικτυακού περιεχομένου που δημιουργήθηκε από απλούς χρήστες έχει προκαλέσει μία βιομηχανία πληροφοριών για την αγορά, αλλά και ανάγκη για εξονυχιστική μέτρηση της έρευνας. Η ιδέα πίσω απ' αυτά είναι η εξόρυξη δεδομένων: αφού τόσο πολλοί άνθρωποι έχουν καταγράψει άτυπες σκέψεις online σε διάφορες μορφές (formats), όπως τα blogs, τα chatrooms, οι πίνακες ανακοινώσεων και οι ιστότοποι κοινωνικής δικτύωσης, θα έπρεπε να είναι δυνατή η εξαγωγή πληροφοριών, όπως οι αντιδράσεις των καταναλωτών στα προϊόντα ή στα παγκόσμια γεγονότα. Προκειμένου να αντιμετωπιστούν ζητήματα όπως αυτά, έχει αναπτυχθεί νέο λογισμικό από μεγάλες εταιρείες όπως το Web Fountain της IBM και το Pulse της Microsoft. Επιπλέον έχουν δημιουργηθεί ή προσαρμοστεί εξειδικευμένες εταιρείες διαδικτυακής πληροφόρησης, όπως η Nielsen Buzz Metrics και η Market Sentinel.

Ένα καλό παράδειγμα μιας ερευνητικής πρωτοβουλίας για την αξιοποίηση των μέσων μαζικής ενημέρωσης που δημιουργούνται από τους καταναλωτές είναι μία προσπάθεια πρόβλεψης πωλήσεων βιβλίων, με βάση τον όγκο των συζητήσεων στο blog τους. Ωστόσο, οι προβλέψεις είχαν μόνο περιορισμένη επιτυχία, ίσως γιατί συχνά οι άνθρωποι μιλάνε στο blog για τα βιβλία μετά την ανάγνωσή τους, όταν θα είναι πολύ αργά για να προβλεφθεί μία αγορά. Ο Gruhl κ.ά. (1999) ανέλυσαν τον όγκο της συζήτησης για την επιλογή των θεμάτων σε χώρο ιστολογίου (blogspace), βρίσκοντας αρκετές διαφορετικές μορφές. Για παράδειγμα, ορισμένα θέματα συζητήθηκαν για ένα μικρό χρονικό διάστημα μόνο, ενώ άλλα συζητιούνταν συνεχώς, με ή χωρίς περιστασιακές εκρήξεις επιπλέον συζήτησης. Μία μελέτη προσανατολισμένη στις

κοινωνικές επιστήμες επεδίωξε να φτιάξει αναδρομικά χρονοδιαγράμματα με τα σημαντικότερα γεγονότα από τα blogs και τις συζητήσεις των ειδήσεων, βρίσκοντας αυτό να είναι δυνατό σε περιορισμένο βαθμό.

Εκτός από το στυλ έρευνας με εξόρυξη δεδομένων, έχουν υπάρξει πολλές μελέτες των ιστοσελίδων Web 2.0, με σκοπό την περιγραφή του περιεχομένου τους και την εξήγηση της συμπεριφορά των χρηστών σε αυτά. Μία μεγάλης κλίμακας μελέτη των πρώτων χρόνων του Facebook παρέχει την πιο ολοκληρωμένη επισκόπηση των δραστηριοτήτων του χρήστη. Τα δεδομένα προήλθαν από το Φεβρουάριο του 2004 έως το Μάρτιο του 2006, όταν το Facebook ήταν ένας ιστότοπος κοινωνικής δικτύωσης αποκλειστικά για τους σπουδαστές κολεγίων των ΗΠΑ. Οι χρήστες φαίνονταν να ταιριάζουν το Facebooking τους με τον κανονικό τρόπο χρήσης του υπολογιστή τους ενώ μελετούσαν, αντί να κατανέμουν / αφιερώνουν διαφορετικές χρονικές στιγμές. Από την άποψη της γεωγραφίας της φιλίας, τα μέλη κυρίως χρησιμοποιούσαν το Facebook για να επικοινωνούν με άλλους φοιτητές στο ίδιο κολέγιο, παρά με άλλους σχολικούς φίλους σε μακρινά πανεπιστήμια. Αυτό υποδηλώνει ότι η κοινωνική δικτύωση είναι μία επέκταση της επικοινωνίας χωρίς σύνδεση (offline communication), παρά της προώθησης ριζικά νέων γεωγραφιών επικοινωνίας (geographies of communication), αν και το τελευταίο έχει ενεργοποιηθεί από την τεχνολογία του Facebook. Αυτό το συμπέρασμα υποστηρίζεται από την ποιοτική έρευνα σ' έναν άλλο δημοφιλή ιστότοπο, τον MySpace.

Μία βιβλιομετρική μελέτη του MySpace ερεύνησε έμμεσα τα επίπεδα δράσης, αλλά εστίασε στα προφίλ των μελών. Μεταξύ άλλων διαπιστώσεων, η μελέτη έδειξε ότι περίπου το ένα τρίτο των εγγεγραμμένων μελών είχαν πρόσβαση στην ιστοσελίδα κάθε εβδομάδα και η μέση αναφερθείσα ηλικία ήταν 21 ετών. Παρόλο που άλλη έρευνα έδειξε ότι οι στενοί φίλοι του MySpace είχαν την τάση να αντανακλούν offline φιλίες / φιλίες χωρίς σύνδεση (offline friendships), τόσο οι χρήστες άνδρες όσο και οι γυναίκες προτιμούσαν να έχουν, επί το πλείστον, φίλους - γυναίκες. Άλλη μελέτη εξέτασε τη γεωγραφία της φιλίας, διαπιστώνοντας ότι η πλειοψηφία των φίλων έτειναν να ζουν μέσα σε απόσταση εκατό μιλίων, αν και μία μειοψηφία ζούσε στην ίδια πόλη ή κομμόπολη.

Τέλος, έχουν δημοσιευθεί πολλές στατιστικές σχετικά με το Web 2.0 από εταιρείες έρευνας αγοράς. Παρά την αβέβαιη προέλευση αυτών των δεδομένων, τα αποτελέσματα μερικές φορές φαίνονται λογικά και επίσης, λόγω του κόστους απόκτησης των δεδομένων, φαίνεται απίθανο να αναπαραχθούν από ακαδημαϊκούς ερευνητές. Ένα παράδειγμα είναι η ανακοίνωση από τη HitWise ότι το MySpace είχε εκτοπίσει το Google ως ο ιστότοπος με τη μεγαλύτερη επισκεψιμότητα από τους χρήστες των ΗΠΑ, το Δεκέμβριο του 2007. Τα δεδομένα της έρευνας αναφέρθηκε ότι προέρχονταν από δύο εκατομμύρια χρήστες του διαδικτύου των ΗΠΑ, μέσω μιας συμφωνίας μεταξύ της HitWise και των Παρόχων Υπηρεσιών Διαδικτύου (Internet Service Providers) των χρηστών. Η δημοσιοποίηση των αποτελεσμάτων των αναλύσεων της επισκόπησης παρέχει χρήσιμη δημοσιότητα στη HitWise και πολύτιμες πληροφορίες στους διαδικτυακούς ερευνητές.

4.6.7 Χρησιμοποιώντας εξυπηρετητές Παγκόσμιου Ιστού

Σε γενικές γραμμές, οι εξυπηρετητές του παγκόσμιου Ιστού μπορούν να στείλουν οποιαδήποτε φόρμα μέσω των που έχει οριστεί ως MIME τύπος (Multi purpose Internet Mail Extensions). Πρότυπες μορφές περιλαμβάνουν κείμενα, γραφικά, CGI (Common Gateway Interface) προγράμματα και βοηθητικές εφαρμογές Java. Τα δυο τελευταία είναι προγράμματα που εκτελούνται είτε στον πελάτη, είτε στον εξυπηρετητή, ή και στους δύο για να επιτελέσουν κάποια επιπλέον λειτουργία.

Ένα από τα βασικά προβλήματα της χρήσης της Βιβλιομετρίας στους εξυπηρετητές Παγκόσμιου Ιστού είναι ότι κανένα στάδιο πληροφορίας δεν καταγράφεται μεταξύ των χρήσεων στον εξυπηρετητή. Τα τρέχοντα πρωτόκολλα εξυπηρετητή αντιμετωπίζουν κάθε πρόσβαση χωριστά, από προεπιλογή κανένα μονοπάτι χρηστών που ακολούθησαν αυτοί για να φτάσουν στη συγκεκριμένη πηγή δεν καταγράφεται μέσω των πόρων. Αυτό θα είναι παρόμοιο με την καταμέτρηση αναφορών, αλλά χωρίς να είναι γνωστός ο αριθμός των εγγράφων πηγών, από όπου αυτές οι αναφορές προέρχονται. Κάποιοι χρήστες ενδέχεται να απομακρύνονται, όταν οι συνδέσεις είναι περιττές ή εξυπηρετούνται εφραπτόμενες πληροφορίες (διπλά εικονίδια πλοήγησης στο επάνω και στο κάτω μέρος των σελίδων) και αυτό δεν θεωρείται εξίσου σημαντικό με άλλα δεδομένα στα αρχεία καταγραφής.

Ένα άλλο σημαντικό πρόβλημα είναι πως οι ευρέσεις σε έναν εξυπηρετητή(server) δεν αντανακλούν πραγματική χρήση, αφού πολλοί χρήστες ίσως να μην χρησιμοποιούν μια πηγή, όταν η σύνδεσή της απαιτεί μεγάλο χρόνο ή διακόπτεται και οδηγούνται στη χρήση άλλων ανακτηθέντων δεδομένων. Για αυτούς τους ίδιους λόγους μετρητές του Παγκόσμιου Ιστού είναι πιθανόν πολύ ανακριβείς. Μια τελευταία απειλή για την εγκυρότητα μετρικής πληροφορίας είναι η έλευση των ρομπότ, που μπορούν να έχουν πρόσβαση σε μια σελίδα και σε κάθε αριθμό πιθανών πόρων αυτής (ακόμα και πολλαπλές φορές). Τα προβλήματα αυτά

δείχνουν ότι πολλά από τα στοιχεία που συγκεντρώθηκαν πρόσφατα στο Παγκόσμιο Ιστό είναι εξαιρετικά ακριβή. Ωστόσο με τους συνδυασμούς των εξυπηρετητών αρχείων καταγραφής και του κατάλληλου σχεδιασμού των πόρων του συστήματος, μπορούν να συγκεντρωθούν πιο σχετικά δεδομένα (2005).

4.7 Σχεδιάζοντας Βιβλιομετρία

Το πρώτο βήμα είναι η πλήρης χρήση όλα των πιθανών αρχείων καταγραφής του εξυπηρετητή. Σε ένα τυπικό διακομιστή ιστού με λειτουργικό σύστημα UNIX, υπάρχουν τέσσερα πιθανά είδη αρχείων καταγραφής:

1. **Server-based**– το αρχείο καταγραφής που δημιουργήθηκε από τον ίδιο τον εξυπηρετητή.
2. **Proxy-based**– μέσω ενός τοίχου προστασίας (επιπλέον επίπεδο προστασίας εξυπηρετητή), ή κάποιοι εξυπηρετητή με ελεγχόμενο σύστημα πρόσβασης.
3. **Client-based**– μέσω κώδικα εφαρμογής πελάτη που μπορεί να καταγράψει και να διαβιβάσει πληροφορίες στον εξυπηρετητή, όπως τα cookies.
4. **Network-based**– τα προγράμματα που ελέγχουν το επίπεδο ασφαλείας δικτύων και την πρόσβαση στον εξυπηρετητή (1996).

Οι περισσότερες επιλογές αρχείου καταγραφής είναι ελαχιστοποιημένες για να συλλέξουν και τα πλήρη δεδομένα, υποστηρίζοντας πως οι ευρέσεις απόδοσης είναι αμελητέες σε σχέση με τον πλούτο των ακριβών πληροφοριών που συλλέγονται. Είναι επίσης πιθανό να μετατοπιστούν οι επιλογές καταγραφής για ορισμένες εκδηλώσεις του Παγκόσμιου Ιστού ή να αντιμετωπίσουν συγκεκριμένα, γνωστά προβλήματα απόδοσης. Ακόμα η περιοδική συλλογή αρχείου καταγραφής μπορεί να αποδειχτεί χρήσιμη, όταν αναλύεται μια ιστοσελίδα.

Μία άλλη (αν και ακριβή) επιλογή είναι να χρησιμοποιήσουμε ένα άλλο σύστημα δικτύωσης του εξυπηρετητή, για να συγκεντρώσουμε πληροφορίες αρχείου. Εναλλακτικά, κάποιοι έξυπνοι προγραμματιστές λογισμικού εξυπηρετητή βασικά αποκαλύπτουν το περιεχόμενο σε άλλον εξυπηρετητή και επιτρέπουν στον αρχικό εξυπηρετητή παγκόσμιου Ιστού να συγκεντρώσει μητρικές πληροφορίες καταγραφής.

Άλλες εσωτερικές πληροφορίες εξυπηρετητή μπορούν να συλλεχτούν, όπως ένα αρχείο καταγραφής ηλεκτρονικού ταχυδρομείου (από και προς τον διαχειριστή του συστήματος, το μέγεθος των δεδομένων και οι χρόνοι και όχι κατ' ανάγκην το περιεχόμενο), τα αρχεία καταγραφής χρήσης του προγράμματος (ακόμη και σε επίπεδο συντήρησης), καθώς και η χρήση του δίσκου (άμεσες πληροφορίες συστήματος σχετικά με ποικίλες διαμορφώσεις κρυφής μνήμης).

4.7.1 Βέλτιστη Διαμόρφωση Εξυπηρετητή Παγκόσμιου Ιστού

Υπάρχει μία σειρά από τυποποιημένες μεθόδους για να χρησιμοποιείς, διαμορφώνοντας έναν εξυπηρετητή Παγκόσμιου Ιστού, ώστε να συγκεντρώσει ολοκληρωμένες μετρήσεις:

1. Ανακατεύθυνε τα προγράμματα για να βρεις το ερώτημα που υποβάλλει το HTTPREFER.(είναι ένα πεδίο κεφαλίδα HTTP που προσδιορίζει τη διεύθυνση της ιστοσελίδας ,δηλαδή το URI)
2. Χρησιμοποίησε μια βάση δεδομένων για να αποθηκεύσεις και να δημοσιεύσεις δυναμικά το περιεχόμενο του παγκόσμιου Ιστού φτιάχνοντας συνεπή οργάνωση ιστού, που δίνει ένα επιπλέον αρχείο καταγραφής για την καταγραφή των δεδομένων χρήσης.
3. Δημιούργησε στάδιο πληροφοριών προγραμματίζοντας μέσω Network Service Access Point Identifier (NSAPI) ή NSAPI είναι τεχνολογίας ActiveX, για να παρακολουθήσεις τη διαδρομή ενός χρήστη σε πραγματικό χρόνο, μέσω της ιστοσελίδας.
4. Χρησιμοποίησε ένα FTP (FILE TRANSFER PROTOCOL) εξυπηρετητή για τη μεταφορά αρχείων, την απελευθέρωση πόρων εξυπηρετητή παγκόσμιου ιστού, δίνοντας πιο συγκεκριμένα αρχεία μεταφοράς αρχείων.

4.7.2 Διαχείριση αρχείων καταγραφής

Η ειρωνεία του Παγκόσμιου Ιστού είναι ότι υπάρχουν σχεδόν απεριόριστες δυνατότητες για τη συλλογή, χρήση και την καταγραφή σχεσιακών πληροφοριών, ενώ υπάρχουν λίγες εναλλακτικές μέθοδοι για την κατεργασία και τη σύγκριση δεδομένων. Είναι το κενό που η Βιβλιομετρία μπορεί να καλύψει για την ερμηνεία των δεδομένων του Παγκόσμιου Ιστού.

Προσθετικά αρχεία καταγραφής εξυπηρετητών μπορεί να μην αντικατοπτρίζουν πλήρως τις συνεδρίες των χρηστών στον Παγκόσμιο Ιστό ή να μην εκπροσωπούν την ακριβή προσαρμογή μεταξύ εγγράφων αλλά να την προσεγγίζουν. Και πάλι, λόγω της ψηφιακής φύσης τους, οι πόροι αυτοί είναι ανοιχτοί για την έρευνα στην επεξεργασία όλων των ειδών. Ωστόσο, πρέπει να ληφθεί μέριμνα για τη διατήρηση αρχείων και τα αποτελέσματά τους. Μερικές γενικές μέθοδοι που εξασφαλίζουν την αξιοπιστία του αρχείου καταγραφής περιλαμβάνουν:

- Τακτική δημιουργία αντιγράφων ασφαλείας, με βάση το μοτίβο του εξυπηρετητή.
- Αποθήκευση των αποτελεσμάτων ανάλυσης του αρχείου καταγραφής και των ίδιων των αρχείων καταγραφής.
- Εκκίνηση νέων αρχείων καταγραφής σε μια έγκαιρη και διαχειρίσιμη βάση.
- Δημοσίευση αποτελεσμάτων και καταγραφή πληροφοριών για σύγκριση.

Η μορφή του αρχείου καταγραφής είναι αρκετά πυκνή. Αυτό πρέπει να λαμβάνεται υπόψη κατά το σχεδιασμό της επεξεργασία καταγραφής και κατά την καταγραφή μεθόδων. Ωστόσο, το κόστος αποθήκευσης μειώνεται με ταχείς ρυθμούς, ιδίως σε σύγκριση με την αξία της ποιότητας των πληροφοριών σχετικά με τις επιδόσεις και τη χρήση συστημάτων καταγραφής διακομιστή.

Εργαλεία ανάλυσης καταγραφών είναι διαθέσιμα όπως: WWWStat (πρόγραμμα γραμμένο σε Perl), GetStats (πίνακας επιστροφής στατιστικών του server), και διάφορα Perl σενάρια (γλώσσα διαθέσιμη για σχεδόν όλα τα λειτουργικά συστήματα) που επιτρέπουν τον έλεγχο μιας ή περισσότερων εφαρμογών, για να βοηθήσουν τη διαδικασία και να προβάλουν τα στατιστικά του διακομιστή. Πρόσθετα εργαλεία εμπορικής ανάλυσης διακομιστή προσφέρουν μια σειρά από πρόσθετες επιλογές, συμπεριλαμβανομένης της ανάλυσης με μετα-δεδομένα, συμπίεση και αυτόματη περιοδική ανάλυση.

Μια τυπική ανάλυση καταγραφής σωρευτικού δείγματος αποδεικνύεται ότι παρουσιάζει μια σειρά από επικαλυπτόμενα αποτελέσματα στην καταμέτρηση της απόδοσης και την πρόσβαση.

Δυστυχώς, σχεδόν κάθε είδος προγράμματος εξυπηρετητή και κάθε μεμονωμένη γεννήτρια καταγραφής του διακομιστή δημιουργούν ελαφρώς διαφορετικές μορφές. Μια προσπάθεια να αλλάξει αυτό βρίσκεται σε εξέλιξη από την επιτροπή του παγκόσμιου ιστού, ώστε να συμφωνήσει πάνω σε μια εκτεταμένη μορφή αρχείου καταγραφής. Η νέα μορφή θα πρέπει να χαρακτηρίζεται από:

- ταχύτητα στην καταγραφή,
- ακρίβεια για το τεκμήριο και τη χρήση συνδέσμου (δίνοντας, επιτέλους, περισσότερες πληροφορίες σχετικά με τη βιβλιομετρία),
- παροχή τυποποιημένων γραμμών βάσης για σύγκριση, ακόμη και με γραφικά,
- αυτόματη εγγραφή της παραπομπής HTTP,
- καταγραφή της συχνότητας ενός γεγονότος,
- καταγραφή του πώς οι ISP (Παροχείς Υπηρεσιών Διαδικτύου) μπορούν να εγκαταστήσουν κρυφή μνήμη (cache) προσωρινά και να διαφοροποιούν πολλούς χρήστες.

4.7.3 Ο J. S. Downie (Καθηγητής στο Πανεπιστήμιο του Illinois) και η χρήση του Παγκόσμιου Ιστού

Πολλές από αυτές τις ιδέες χρησιμοποιούνται σε μια πρώιμη προσπάθεια του Downie (1996) να εφαρμοστεί η Βιβλιομετρία στον παγκόσμιο ιστό. Αυτές οι προσπάθειες αναλύουν, κυρίως, τις παρακάτω κατηγορίες :

1. **Ανάλυση από χρήστες** για την περαιτέρω διερεύνηση και ανακάλυψη δημογραφικών στοιχείων και απόδοση στοιχείων βασισμένων στο:
 - ποιος (οργάνωση),
 - πού (Περιοχή),
 - τι (Περιηγητής Πελάτης).

2. Αιτήσεις αναλύσεων σε σχέση με το περιεχόμενο

Ανάλυση, με βάση τα Byte, για τη μέτρηση της ακατέργαστης απόδοσης σε σχέση με συγκεκριμένα χρονοδιαγράμματα.

Η δύναμη αυτών των τεχνικών είναι ότι μπορούν να συμπτυχθούν, για να αναπτύξουν ένα αναλυτικό σενάριο από την περιήγηση(εις) ενός χρήστη στις ιστοσελίδες και τα στοιχεία τους, προβλήματα και πράξεις. Κατά τη χρήση βιβλιομετρικών τεχνικών ανάλυσης, ο Downie ανακάλυψε ότι οι αιτήσεις προσαρμόζονταν σε μια Zipfian κατανομή από έναν πίνακα συχνότητας κατάταξης.

Αλλά αποτελέσματα επιβεβαιώνουν ότι η χαμηλής ποιότητας παραμετροποίηση του εξυπηρετητή στον παγκόσμιο ιστό και η έλλειψη δικαιωμάτων πρόσβασης ή η χρήση πλήρων αρχείων καταγραφής μπορεί να φράξουν περαιτέρω τα αποτελέσματα. Επίσης είναι αξιοσημείωτο ότι ο Downie διερευνά ζητήματα ηθικής στις παρατηρήσεις, κάτι που πολλοί διαχειριστές ιστοχώρων και επαγγελματίες συστημάτων πληροφόρησης δεν έχουν, έως τώρα, αναλογιστεί.

4.7.4 Βέλτιστη παραμετροποίηση περιεχομένου παγκόσμιου ιστού

Με βάση τις τεχνικές παραμετροποίησης εξυπηρετητών και την αληθινή δουλειά και εμπειρία του Downie, μπορούν να χρησιμοποιηθούν μερικές από τις μεθόδους για την οργάνωση του παγκόσμιου ιστού με τον βέλτιστο τρόπο, όπως :

- Κατοχή πληροφοριών επικοινωνίας από ερωτήσεις, σχόλια και δικαιώματα παρατήρησης.
- Κατοχή δηλώσεων σκοπών σχετικά με την ιστοσελίδα και πληροφορίες εύκολα διαθέσιμες για να βοηθήσουν τους χρήστες να βρουν ό,τι χρειάζονται στην ιστοσελίδα.
- Χρήση μοναδικών αρχείων και ονομάτων φακέλων.
- Χρήση σαφούς και συνεπούς αρχιτεκτονικής πόρων.
- Συνέπεια στον πολλαπλασιασμό των υπερσυνδέσεων.
- Χρήση τύπων Mime (Multipurpose Internet Mail Extensions).
- Έλεγχος ή απαγόρευση πρόσβασης ρομπότ αναζήτησης και αραχνών αναζήτησης.

4.7.5 Συλλογή Βιβλιομετρικών δεδομένων

Η τεχνολογία της πληροφορικής είναι επίσης πιθανό να συλλέξει πληροφορίες σε σχέση με μια ιστοσελίδα, με το να βρίσκει πληροφορίες εξωτερικά από την ιστοσελίδα σε ποικίλες πηγές του Παγκόσμιου Ιστού. Αν μια ιστοσελίδα αναπτύχθηκε αξιοποιώντας τις παραπάνω μεθόδους, μπορούν να συλλεχτούν περισσότερα δεδομένα από εξωτερικές ιστοσελίδες. Ένα καθαρό παράδειγμα είναι η χρήση μηχανής αναζήτησης για την εύρεση αναφορών στην ιστοσελίδα γενικότερα, είτε με συγκεκριμένες διευθύνσεις ιστοσελίδων (URL). Αυτή η διαδικασία γίνεται απλούστερη, αφού τα ονόματα των σελίδων και τα ονόματα των καταλόγων είναι σχετικά μοναδικά. Ωστόσο, προτείνεται για αποτελεσματικότερη αναζήτηση, η χρήση διαφορετικών μηχανών αναζήτησης για μια ποικιλία από δεδομένα σε διαφορετικό χρόνο και γεωγραφικά τμήματα. Κάθε μηχανή αναζήτησης έχει συγκεκριμένο χρόνο ανανέωσης των ευρετηρίων αναζήτησης για να φιλτράρει πλεονάζουσες και διπλότυπες, περιττές πληροφορίες, χρησιμοποιώντας αυτό σαν πλεονέκτημα στη συλλογή δεδομένων.

Προοδευτικά, αναζητώντας και περιπλανώμενοι στις υπερσυνδέσεις σε οποιαδήποτε σελίδα που είναι στις αναφορές, μπορούμε να αποκαλύψουμε λιγότερες προφανείς συνδέσεις και αναφορές. Μια ανάστροφη πορεία ανάλυσης των διευθύνσεων ιστοσελίδων(URL) μπορεί να οδηγήσει στη δική της Βιβλιομετρική Zipfian κατανομή τύπων και περιεχομένου, αποκαλύπτοντας χαρακτηριστικά ονομάτων, φακέλων και αρχείων.

Για ανάλυση υψηλής ακρίβειας, υπηρεσίες όπως της DejaNews μπορούν να δείξουν μηνύματα και πρόσβαση σε στατιστικά δεδομένων συγκεκριμένης χρήσης του Παγκόσμιου Ιστού. Τέλος, η αξιοποίηση ενός προσαρμοσμένου ρομπότ ή αράχνης αναζήτησης για την αναζήτηση εσωτερικών και εξωτερικών ιστοσελίδων, για τις κοινοτυπίες και τις αναφορές, μπορεί επίσης να αποδώσει νέα δεδομένα.

4.8 Altmetrics

Ο όρος «αλτμετρική» επινοήθηκε το 2010 ως ένας ομαδικός, εναλλακτικός όρος των μέτρων, που είναι εξαρτημένος από αναφορές και από το επίπεδο μέτρων του εντύπου (2012.)ορίζοντας, «τη δημιουργία και τη μελέτη των νέων μετρήσεων με βάση τον Κοινωνικό Παγκόσμιο Ιστό (Social Web) για την ανάλυση και την ενημέρωση των δεδομένων». Ο όρος παρουσιάστηκε από εξέχουσες προσωπικότητες στον κόσμο της βιβλιομετρίας, όπως οι Jason Priem και Heather Piwowar (2013).

Οι altmetrics, όπως υποδηλώνει το όνομα, είναι εναλλακτικές μετρήσεις του παραδοσιακού παράγοντα επιρροής και των άλλων δεικτών. Οι Altmetrics μπορούν να διαδοθούν από ένα ευρύ φάσμα δικτύων κοινωνικών μέσων μαζικής ενημέρωσης, όπως τα blogs, το Twitter και το Mendeley. Όσον αφορά στις αποδείξεις που δείχνουν επιστημονικό αντίκτυπο, χρησιμοποιώντας altmetrics, εκτός από τις παραδοσιακές μετρικές, μπορούμε να προσφέρουμε μια συνολική και πιο άμεση εικόνα του επιστημονικού έργου καθενός.

4.8.1 Πηγές Altmetrics

Altmetrics.org. Αυτό το δωρεάν Web site είναι ένα κεντρικό κομβικό σημείο για πληροφορίες σχετικά με την αυξανόμενη κίνηση των altmetrics. Το altmetrics.org διατηρεί δεσμούς με τα νέα online εργαλεία για τον υπολογισμό των επιπτώσεων. Πρόσβαση: <http://altmetrics.org>.

•**Impact Story** (IS-Ιστορία των Επιπτώσεων). Παλαιότερα γνωστό ως Συνολική Επίπτωση. Η Ιστορία Επιπτώσεων (IS) είναι ένα εργαλείο ανοιχτού κώδικα για την υποστήριξη των βασιζόμενων σε URL δημοσιεύσεων, μέσω του συνυπολογισμού των on line altmetrics. Οι χρήστες δημιουργούν συλλογές των υλικών μέσω online αναγνωριστικών, όπως το Google Scholar Profiles, τα ψηφιακά αναγνωριστικά αντικειμένου ‘Digital Object Identifier’ (DOI’s) και τα PubMed IDs. Η Ιστορία των Επιπτώσεων (IS) χρησιμοποιεί πάνω από μια ντουζίνα APIs για να ψάξει μετρικά σε αυτά τα συλλεγόμενα στοιχεία, με πηγές που κυμαίνονται από τα δημοφιλή μέσα κοινωνικής δικτύωσης σε επιστημονικά εργαλεία, όπως τα Mendeley και PLoS. Τα στοιχεία αυτά, στη συνέχεια, αποδίδονται σε κατηγορίες επιπτώσεων, όπως σε γενικά / υψηλά «αναφερόμενα», «συνιστώμενα» ή «που συζητήθηκαν». Αυτή η πηγή είναι ιδιαίτερα χρήσιμη για τους ερευνητές που δημοσιεύουν στους μη παραδοσιακούς χώρους. Πρόσβαση: <http://impactstory.it>.

•**PLoS Article Level Metrics.** Το Public Library of Science (PLoS) έχει αναδειχθεί ως το κορυφαίο αποθετήριο περιοδικών ανοιχτής πρόσβασης, εν μέρει λόγω των υψηλών παραδοσιακών παραγόντων επιπτώσεών του. Ωστόσο, το PLoS προσφέρει μια εναλλακτική στους παραδοσιακούς αντίκτυπους με τη μορφή του Article Level Metrics, το οποίο παρακολουθεί την επιρροή των μεμονωμένων PLoS άρθρων, από τη στιγμή που θα το κατεβάσει κάποιος και θα το αναφέρει στα social media και τα blogs. Το PLoS παρακολουθεί επίσης τις εσωτερικές μετρήσεις του άρθρου, συμπεριλαμβανομένων των σχολίων, των σημειώσεων και των κατατάξεων. Ενώ είναι ένας πολύτιμος πόρος για τις επιπτώσεις, μόνο τα PLoS άρθρα επωφελούνται από τα μετρικά του. Παρόλα αυτά, ο πόρος αυτός αντιπροσωπεύει ένα σημαντικό νέο δρόμο για τα μετρικά, τα οποία στο μέλλον οι εκδότες είναι πιθανό να αναπαράγουν. Διαθέσιμο δωρεάν online. Πρόσβαση: <http://article-level-metrics.PLoS.org/>.

•**Reader Meter.** Το Reader Meter είναι ένα δωρεάν εργαλείο που «προμηθεύεται» τις επιπτώσεις από την επεξεργασία των δεδομένων του αναγνωστικού κοινού από το Mendeley. Δημιουργήθηκε από τον Dario Taraborelli του Ιδρύματος Wikimedia. Η ιστοσελίδα λειτουργεί με τη σύνταξη εκθέσεων με βάση τα ονόματα των συγγραφέων, τα οποία στη συνέχεια υποβάλλονται σε επεξεργασία μέσω του Mendeley API. Κάθε έκθεση δημιουργεί πληροφορίες, όπως το «HR-Index», το «GR-Index», το «Total Bookmarks» και τα «Top Publications αναγνωσιμότητας» ενός συγγραφέα. Το Reader Meter επικρίθηκε από ορισμένους στην κοινότητα των altmetrics, για την κατάρτιση των δεδομένων αποκλειστικά από το Mendeley. Ωστόσο, υπάρχουν σχέδια για την ενσωμάτωση δεδομένων από πολλαπλές τοποθεσίες διαχείρισης αναφορών, όπως το CiteULike. Πρόσβαση: <http://readermeter.org/>.

4.8.2 Επιστημονικά δίκτυα Altmetrics

Academia.edu. Το Academia.edu είναι μια δωρεάν online πλατφόρμα κοινοποίησης άρθρων, που ενθαρρύνει τους ακαδημαϊκούς ώστε να αυξήσουν την προβολή τους και την παρακολούθηση της έρευνας μέσα από αυτήν και μεταξύ των ακαδημαϊκών του δικτύου της. Με σχεδόν 2 εκατομμύρια προφίλ ακαδημαϊκών και 1.500.000 ανεβασμένα άρθρα, το academia.edu έχει γίνει ένας δημοφιλής παίκτης στον κόσμο των online αποθετηρίων. Οι μετρικές των αποτελεσμάτων για αυτό το site είναι παρόμοιες με εκείνες που προσφέρονται από πολλά blogs και περιλαμβάνουν προβολές προφίλ, τις θεάσεις του εγγράφου και

επισκεψιμότητα με βάση κάθε χώρα. Ακολουθώντας μια άλλη αυξανόμενη τάση για τα επιστημονικά δίκτυα, η ιστοσελίδα προσφέρει επίσης δυνατότητες προσανατολισμένες προς την κοινωνική αλληλεπίδραση, όπως το εργαλείο "κάνε μια ερώτηση". Πρόσβαση: <http://www.academia.edu/>.

•**Mendeley**. Το Mendeley είναι ένα σχετικά πρόσφατο ξεκίνημα από την ίδια εταιρεία που δημιούργησε το Last.fm. Οι ερευνητές με προφίλ μπορεί να χαρτογραφήσουν τις απόψεις και τις λήψεις της έρευνάς τους μέσω της πύλης, να συμμετέχουν σε ομάδες και να δουν δημοφιλή άρθρα σχετικά με τους τομείς τους. Το Mendeley έχει γίνει ιδιαίτερα ελκυστικό στις θετικές επιστήμες, από τις οποίες το μεγαλύτερο μέρος των χρηστών του το χαιρετίζει. Ωστόσο, με την ενσωμάτωση των δεδομένων του Mendeley σε περισσότερα εργαλεία altmetrics, κατά πάσα πιθανότητα, θα γίνει δημοφιλές και σε άλλους κλάδους επίσης. Το Mendeley είναι δωρεάν με αναβαθμίσεις που κοστίζουν σε αποθήκευση και είναι διαθέσιμο τόσο online, όσο και για download. Πρόσβαση: <http://www.mendeley.com>.

•**Social Science Research Network (SSRN)**. Το SSRN είναι ένα online αποθετήριο άρθρων, το οποίο πρόσφατα ήταν στο νούμερο ένα στην Ιστοσελίδα της Κατάταξης Παγκόσμιων Αποθετηρίων για το 2012. Περιλαμβάνει τρία βασικά χαρακτηριστικά: μια βάση δεδομένων με περισσότερες από 400.000 περιλήψεις, μια μεγάλη ηλεκτρονική συλλογή άρθρων και 20 εξειδικευμένα θεματικά δίκτυα μέσω των οποίων οι καταχωρημένοι χρήστες μπορούν να προωθήσουν το έργο τους και να συνδεθούν με δωρεάν περιλήψεις και άρθρα. Αν και έχει επαινεθεί για την ικανότητά του να διευκολύνει την προώθηση της έρευνας, το SSRN έχει επίσης επικριθεί για την αυστηρότητα της πολιτικής του, που ορισμένοι θεωρούν ασφυκτική, σε σύγκριση με τα αναδυόμενα επιστημονικά δίκτυα. Τα ειδικά τοποθετημένα μετρικά του για τα «κορυφαία άρθρα», «κορυφαίους συγγραφείς» και τα «κορυφαία ιδρύματα» παραμένουν το κλειδί για τις σχολές κοινωνικών επιστημών. Πρόσβαση: <http://www.ssrn.com>.

•**VIVO**. Το VIVO είναι μια ελεύθερη, κατάλληλη για κατέβασμα δεδομένων εφαρμογή σημασιολογικού ιστού που αποσκοπεί στη διευκόλυνση της έρευνας. Αρχικά αναπτύχθηκε στο Cornell και καλεί τα ιδρύματα να ανεβάζουν τα δεδομένα που σχετίζονται με το προφίλ της σχολής, για να χαράξουν χρήσιμους δεσμούς μεταξύ των ερευνητών. Το VIVO δεν υποστηρίζει άμεσα τα μετρικά με επίκεντρο το χρήστη, αλλά έχει τη δυνατότητα να είναι ένα ισχυρό εργαλείο για τη συλλογή ερευνητικών μετρικών επιστημονικού επιπέδου. Μέχρι σήμερα, μόνο λίγα μεγάλα ιδρύματα έχουν εφαρμόσει το VIVO, καθώς απαιτεί σημαντικές γνώσεις προγραμματισμού, αλλά και μεγάλη δέσμευση. Πρόσβαση: <http://vivoweb.org>.

4.8.3 Τα Blogs και τα μέσα ενημέρωσης Altmetrics

•**Citation Culture**. Αυτό το blog είναι η δημιουργία του Paul Wouters, διευθυντή του Centre for Science and Technology Studies του Πανεπιστημίου του Leiden (LU). Γραμμένο από τον Wouter και έναν συνάδελφο καθηγητή του στο LU, το blog αυτό είναι αφιερωμένο στη συζήτηση της ακαδημαϊκής επίδρασης, από την ανάλυση των αναφορών στην ευρύτερη αξιολόγηση της έρευνας σε όλα τα πανεπιστήμια. Πρόσφατες δημοσιεύσεις έθιξαν θέματα, όπως η ανθρωπιστική βιβλιομετρία και τα επιστημονικά altmetrics. Μολονότι οι πληροφορίες σχετικά με την ιστοσελίδα είναι εξαιρετικές και λεπτομερείς, οι δημοσιεύσεις γίνονται με ρυθμό ένα με δύο κάθε μήνα. Πρόσβαση: <http://citationculture.wordpress.com/>.

•**Ιστοσελίδα του Jason Priem**. Ο Jason Priem είναι ένας υποψήφιος διδάκτορας στο University of North Carolina-Chapel Hill's School of Information and Library Science και ο συνιδρυτής της Ιστορίας των Επιπτώσεων. Ο Priem έχει αναδειχθεί ως ένας από τους πιο ένθερμους υποστηρικτές των altmetrics. Τα ενδιαφέροντά του αφορούν μια ποικιλία θεμάτων altmetrics, συμπεριλαμβανομένων του μέλλοντος της επιστημονικής επικοινωνίας, του κινήματος των ανοιχτών δεδομένων, καθώς και των δικαιωμάτων δημιουργού. Πρόσβαση: <http://jasonpriem.org/>.

•**Scholarly Kitchen**. Ιδρύθηκε από την Society for Scholarly Publishing. Το Scholarly Kitchen διαθέτει διαχειριστή blog που παρουσιάζει ιδέες για τρέχοντα θέματα της επιστημονικής δημοσίευσης και επικοινωνίας. Αν και δεν είναι αυστηρά επικεντρωμένο στην βιβλιομετρία, πολλοί από τους «σεφ» της ιστοσελίδας διαθέτουν εμπειρία στον διεπιστημονικό τομέα μεταξύ των επιδράσεων και των εκδόσεων. Η ιστοσελίδα προσφέρει, επίσης, χρήσιμα φίλτρα κατηγοριών, όπως το «Μετρικά & Αναλυτικά», το οποίο περιλαμβάνει πάνω από 280 δημοσιεύσεις και συνεχίζει ακάθεκτο. Πρόσβαση: <http://scholarlykitchen.sspnet.org/>.

4.8.4 Υποστήριξη της έρευνας Altmetrics

•**Elsevier Bibliometrics Research Program (EBRP).** Το EBRP σχεδιάστηκε από τον Elsevier ως ένας τρόπος για να αποκτήσουν οι ερευνητές της βιβλιομετρίας δωρεάν πρόσβαση σε μεγάλες ποσότητες δεδομένων. Τα διαθέσιμα στοιχεία περιλαμβάνουν τη δημοσίευση μεταδεδομένων από το Scopus, δεδομένα χρήσης και δεδομένα πλήρους κειμένου από το ScienceDirect. Οι ερευνητές αιτούνται τα δεδομένα και οι επιτυγχόντες λαμβάνουν ένα σύνολο δεδομένων που έχουν σχεδιαστεί ειδικά για το έργο τους από τον Elsevier. Παραδείγματα επιτυχημένων έργων στην ιστοσελίδα είναι ιδιαίτερα χρήσιμα σε όσους ενδιαφέρονται για τρέχοντα θέματα altmetrics, όπως η σχέση μεταξύ κατεβασμένων άρθρων και αναφορών. Πρόσβαση: <http://ebrp.elsevier.com/index.asp>.

•**OII Tool kit for the Impact of Digitised Scholarly Resources** (εργαλειοθήκη ΟΙΙ για τις επιπτώσεις των Ψηφιοποιημένων Επιστημονικών Πόρων). Αυτή η εργαλειοθήκη αναπτύχθηκε από το Oxford Internet Institute για να βοηθήσει τους συγγραφείς, τους εκδότες και τους βιβλιοθηκονόμους να μάθουν περισσότερα για τη μέτρηση των επιδράσεων των ψηφιακών δεδομένων. Η ιστοσελίδα χωρίζεται σε τρεις ενότητες: μελέτες περιπτώσεων, ποσοτικές μέθοδοι και ποιοτικές μέθοδοι. Οι δύο τελευταίες ενότητες προσδιορίζουν και συζητούν τις μεθοδολογικές υποκατηγορίες, όπως τη Βιβλιομετρία / Επιστημομετρία και την ανάλυση περιεχομένου. Οι εισφορές στην εργαλειοθήκη ενθαρρύνονται με τη μορφή άρθρων και σχολίων, τα οποία μπορούν να υποβάλλονται μετά τη δημιουργία ενός δωρεάν λογαριασμού χρήστη. Πρόσβαση: <http://microsites.oii.ox.ac.uk/tidsr/welcome>.

Βιβλιογραφία

- Aguillo, I. F. (1998). STM information on the Web and the development of new Internet R&D databases and indicators. In *ONLINE INFORMATION-INTERNATIONAL MEETING-* (pp. 239–244). LEARNED INFORMATION LTD.
- Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the World Wide Web: Methodological approaches to 'webometrics'. *Journal of Documentation*, 53(4), 404–426.
- Arasu, A., Cho, J., Garcia-Molina, H., Paepcke, A., & Raghavan, S. (2001). Searching the web. *ACM Transactions on Internet Technology (TOIT)*, 1(1), 2–43.
- Bar, Judit, Ilan Blyma and Bluma C. Peritz (2002), Informetric theories and methods for exploring the internet: an analytical survey of recent research literature, *Library Trends*, 50, 3 (371-392).
- Björneborn, L., & Ingwersen, P. (2001). Perspective of webometrics. *Scientometrics*, 50(1), 65–82.
- Björneborn, L. (2001). Small-world linkage and co-linkage. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia* (pp. 133–137). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=504252>
- Björneborn, L. (2004). *Small-world link structures across an academic web space: a library and information science approach*. Det Informationsvidenskabelige AkademiDanish School of Library and Information Science, Institut østInstitut øst. Retrieved from http://pure.iva.dk/ws/files/31034741/lennart_bjorneborn_phd.pdf
- Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.
- Boudourides, M. A., Sigrist, B., & Alevizos, P. D. (1999). Webometrics and the self-organization of the European Information Society. *Draft Report, Task*, 2, 17–19.
- Casserly, M. F., & Bird, J. E. (2003). Web citation availability: analysis and implications for scholarship. *College & Research Libraries*, 64(4), 300–317.
- Brody, T., Carr, L., Gingras, Y., Hajjem, C., Harnad, S., & Swan, A. (2007). Incentivizing the open access research web: publication-archiving, data-archiving and scientometrics. *CTWatch Quarterly*, 3(3). Retrieved from <http://eprints.ecs.soton.ac.uk/14418>
- Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., ... Wiener, J. (2000). Graph structure in the web. *Computer Networks*, 33(1), 309–320.
- Brody, T., Harnad, S., & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060–1072.
- Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5303–5310.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377.
- CitNetExplorerTutorial. Διαθέσιμοσε: www.citnetexplorer.nl/gettingstarted/. Ανακτήθηκετην 5η Μαΐου 2015
- Davison, B. D. (2000). Topical locality in the web. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 272–279). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=345597>
- Downie, S. J. (2013). Infometrics and the World Wide Web: A Case Study and Discussion. In *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI*. Retrieved from <http://www.cais-acsi.ca/ojs/index.php/cais/article/view/19>
- Egghe, L. (2000). New informetric aspects of the Internet: some reflections-many problems. *Journal of Information Science*, 26(5), 329–335.

- Garner, B. J. (1967). Models of urban geography and settlement location. *Models in Geography*, 303–360.
- Gruhl, D., Bailey, J., Spohrer, J., & Maglio, P. P. (2007). *Steps toward a science of service systems. Computer*, (1), 71–77.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236–243.
- Ingwersen, P. (2000). The international visibility and citation impact of Scandinavian research articles in selected social science fields: the decay of a myth. *Scientometrics*, 49(1), 39–61.
- Ingwersen, P., Larsen, B., & Noyons, E. (2001). Mapping national research profiles in social science disciplines. *Journal of Documentation*, 57(6), 715–740.
- Ingwersen, P., & Wormell, I. (1999). Publication behaviour and international impact: Scandinavian clinical and social medicine, 1988–96. *Scientometrics*, 46(3), 487–499.
- Ingwersen, P., & Christensen, F. H. (1997). Data Set Isolation for Bibliometric Online Analyses of Research Publications: Fundamental Methodological Issues. *JASIS*, 48(3), 205–217.
- Ingwersen, P. (2006). Webometrics: ten years of expansion. Retrieved from <http://eprints.rclis.org/archive/00007554>
- Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10(6). Retrieved from <http://eprints.ecs.soton.ac.uk/10207>
- Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., ... Hilf, E. R. (2008). The access/impact problem and the green and gold roads to open access: An update. *Serials Review*, 34(1), 36–40.
- Harnad, S., Brody, T., & Vallières, F. (n.d.). The access/impact problem and the green and gold roads to open access. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/00987913.2004.10764930>
- Heydon, A., & Najork, M. (1999). Mercator: A scalable, extensible web crawler. *World Wide Web*, 2(4), 219–229.
- Hubert, B., Gérard, J.-C., Bisikalo, D. V., Shematovich, V. I., & Solomon, S. C. (2001). The role of proton precipitation in the excitation of auroral FUV emissions. *Journal of Geophysical Research: Space Physics (1978–2012)*, 106(A10), 21475–21494.
- Hummon, N. P., & Dereian, P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks*, 11(1), 39–63.
- Klavans, R., & Boyack, K. W. (2006). Quantitative evaluation of large maps of science. *Scientometrics*, 68(3), 475–499.
- Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5), 604–632.
- Koehler, W., & others. (2004). A longitudinal study of Web pages continued: a consideration of document persistence. *Information Research*, 9(2), 9–2.
- Kochen, M. (1989). *The small world*. Ablex Norwood, NJ. Retrieved from <http://www.socsci.uci.edu/~bgrofman/R35%20Grofman%20and%20Feld.%201988.Toward%20a%20sociometric%20theory%20of%20rep....pdf>
- Kurtz, M. J. (2004). *Restrictive access policies cut readership of electronic research journal articles by a factor of two, National Policies on Open Access (OA) Provision for University Research Output: an International meeting*. Southampton University, Southampton UK. 19 February 2004.
- Larson, R. R. (1996). Bibliometrics of the World Wide Web: An exploratory analysis of the intellectual structure of cyberspace. In *Proceedings of the Annual Meeting-American Society for Information Science* (Vol. 33, pp. 71–78). Retrieved from <http://free-journal.umm.ac.id/files/file/asis96.ps>
- Lawrence, S. (2001). Online or invisible. *Nature*, 411(6837), 521.

- Leydesdorff, L., & Curran, M. (2000). Mapping university-industry-government relations on the Internet: the construction of indicators for a knowledge-based economy. *Cybermetrics*, 4(1), 1–17.
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362.
- Li, X., Thelwall, M., Musgrove, P., & Wilkinson, D. (2003). The relationship between the WIFs or inlinks of Computer Science Departments in UK and their RAE ratings or research productivities in 2001. *Scientometrics*, 57(2), 239–255.
- Li, X. (2003). A review of the development and application of the Web impact factor. *Online Information Review*, 27(6), 407–417.
- McKiernan, P. (1996). *Historical evolution of strategic management* (Vol. 2). Dartmouth Publishing.
- Mettrop, W., & Nieuwenhuysen, P. (2001). Internet search engines-fluctuations in document accessibility. *Journal of Documentation*, 57(5), 623–651.
- Milgram, S. (1967). The small world problem. *Psychology Today*, 2(1), 60–67.
- Nance, R. E., Korfhage, R. R., & Bhat, U. N. (1972). Information networks: definitions and message transfer models. *Journal of the American Society for Information Science*, 23(4), 237–247.
- Newman, M. E., & Watts, D. J. (1999). Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4), 341–346.
- Odlyzko, A. (2002). The rapid evolution of scholarly communication. *Learned Publishing*, 15(1), 7–19.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: bringing order to the Web. Retrieved from <http://ilpubs.stanford.edu:8090/422>
- Palmquist, Ruth A. Bibliometrics, University of Texas, <http://www.gslis.utexas.edu/~palmauis/courses/biblio.html> [πρόσβαση 24-7-2015].
- PERITZ, J. B.-I. B. (2002). Informetric theories and methods for exploring the Internet: An analytical survey of recent research literature. *ILLINOIS*, 50(3), 371–392.
- Pirolli, P., Pitkow, J., & Rao, R. (1996). Silk from a sow's ear: extracting usable structures from the Web. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 118–125). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=238450>
- Piwowar, H. (2013). Altmetrics: Value all research products. *Nature*, 493(7431), 159–159.
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv Preprint arXiv:1203.4745*. Retrieved from <http://arxiv.org/abs/1203.4745>
- Rehm, G. (2002). Towards automatic Web genre identification: a corpus-based approach in the domain of academia by example of the Academic's Personal Homepage. In *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on* (pp. 1143–1152). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=994036
- Rousseau, J.-J., & Gourevitch, V. (1997). *Rousseau: 'The Social Contract' and Other Later Political Writings*. Cambridge University Press. Retrieved from https://www.google.com/books?hl=el&lr=&id=kcvseZCgQKMC&oi=fnd&pg=PR7&dq=%27The+Social+Contract%27and+Other+Later+Political+Writings&ots=CcVmpfl95w&sig=hawAx1ppaz_L88iuaCVFAFJ73PU
- Rousseau, R. (1997). *Sitations: an Exploratory Study*. *Cybermetrics*, [Online], 1 (1).
- Rousseau, R. (2002). Journal evaluation: Technical and practical issues. *Library Trends*, 50(3), 418–439.
- Skupin, A., Biberstine, J. R., & Börner, K. (2013). Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PloS One*, 8(3), e58779.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269.

- Snyder, H., & Rosenbaum, H. (1999). Can search engines be used as tools for web-link analysis? A critical view. *Journal of Documentation*, 55(4), 375–384.
- Spink, A., Wolfram, D., Jansen, M. B., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3), 226–234.
- Spink, A., Jansen, B. J., Wolfram, D., & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. *Computer*, 35(3), 107–109.
- Testa, J., & McVeigh, M. E. (2004). The impact of open access journals: A citation study from Thomson ISI. Çevrimiçi Elektronik Adres: [Http://www. Thomsonscientific. Com/media/presentrep/acropdf/impact-Oa-Journals. Pdf](http://www.Thomsonscientific.Com/media/presentrep/acropdf/impact-Oa-Journals.Pdf) [15.10. 2005'te Erişildi].
- Thelwall, M. (2001). Extracting macroscopic information from web links. *Journal of the American Society for Information Science and Technology*, 52(13), 1157–1168.
- Thelwall, M. (2002). Conceptualizing documentation on the Web: An evaluation of different heuristic-based models for counting links between university Web sites. *Journal of the American Society for Information Science and Technology*, 53(12), 995–1005.
- Thelwall, M. (2003). What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, 8(3), 8–3.
- Thelwall, M., & Harries, G. (2003). The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university. *Journal of the American Society for Information Science and Technology*, 54(7), 594–602.
- Thelwall, M., & Harries, G. (2004). Do the web sites of higher rated scholars have significantly more online impact? *Journal of the American Society for Information Science and Technology*, 55(2), 149–159.
- Thelwall, M., & Wilkinson, D. (2004). Finding similar academic Web sites with links, bibliometric couplings and colinks. *Information Processing & Management*, 40(3), 515–526.
- Thelwall, M., Vaughan, L., & Björneborn, L. (2005). Webometrics. *Annual Review of Information Science and Technology*, 39(1), 81–135.
- Turnbull, Don. Bibliometrics and the World-Wide Web, University of Toronto, <http://www.ischool.utexas.edu/~donturn/research/bibweb.html> [πρόσβαση 24-7-2007].
- Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring Scholarly Impact* (pp. 285–320). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-10377-8_13
- Van Raan, A. F. (2001). Bibliometrics and Internet: Some observations and expectations. *Scientometrics*, 50(1), 59–63.
- Vaughan, L., & Hysen, K. (2002). Relationship between links to journal Web sites and impact factors. In *Aslib Proceedings* (Vol. 54, pp. 356–361). MCB UP Ltd. Retrieved from <http://www.emeraldinsight.com/doi/abs/10.1108/00012530210452555>
- Vaughan, L., & Thelwall, M. (2004). Search engine coverage bias: evidence and possible causes. *Information Processing & Management*, 40(4), 693–707.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440–442.
- Williams, M. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society*, 16(3), 201–215.
- Wilson, C. S. (1999). Informetrics. *Annual Review of Information Science and Technology* (ARIST), 34, 107–247.
- Wormell, I. (1998). Informetrics: an emerging subdiscipline in information science. *Asian Libraries*, 7(10), 257–268.

Varki, S., & Wong, S. (2003). Consumer involvement in relationship marketing of services. *Journal of Service Research*, 6(1), 83–91.

Σίτας, Α. Κώνστα, Ο. Ντούλια, Π.(2007). *Βιβλιομετρική-Πληροφοριομετρική-Ιστομετρική*. Δημοσιευμένο στο: Μιχάλης Τζεκάκης: ο οραματιστής. Ένωση Ελλήνων Βιβλιοθηκονόμων και Επιστημόνων της Πληροφόρησης, Xerox,44-63.

Διαδικτυογραφία

<http://arxiv.org/>

<http://citeseerx.ist.psu.edu/>

<http://books.google.com/>

<http://scholar.google.com/>

<http://iopscience.iop.org/page/archive>

<http://www.ams.org/mathscinet/>

<http://adswwww.harvard.edu/>

<http://highwire.stanford.edu/>

<http://prola.aps.org/>

<http://www.cas.org/products/sfacad/index.html>

http://scitation.aip.org/search_scitation?key=SPIN

<http://www.scirus.com/>

<http://www.sciencedirect.com/>

<http://www.scopus.com/home.url>

<http://www.google.com/>

EYPETHPIO

A

Atlas of Science, 26
Author's Name, 111

B

block-modeling, 83

C

chatrooms, 33, 152
Citations, 6, 10, 24, 37, 67, 89, 90, 99, 100, 110
cited papers, 106
co-linked, 138
cookies, 144, 153
correlation matrix, 146

F

Field normalized citation score, 107

H

Highest Impact Journals, 91
Hottest journals, 91
hybrid, 72
h-γράφημα, 77
h-δείκτης, 77

L

Largest journals, 91
Library and Information Science, 9, 28, 56, 159, 166

M

Most Frequently Cited Journals, 91

P

publication and citation pattern, 77
Publish or Perish, 9, 110, 111

S

six degrees of separation, 148

A

αλληλεπίδραση, 47, 48, 157
Ανάλυση Περιεχομένου, 6, 87
ανάλυση συν-αναφορών, 72
Αναλυτής Περιοδικού, 7, 95
αναφερόμενο άρθρο, 51
Αναφορές εντός κειμένου, 52
ανοικτή πρόσβαση, 29, 109, 124, 125
αξιολόγηση από ομότιμους, 13, 35, 37, 40, 47, 109
απευθείας σύνδεσης, 135
Αποδοτικότητα, 72

αριθμός των αναφορών, 34, 39, 47, 60, 64, 72, 92, 94, 95, 97, 99, 109, 125
αριθμός των δημοσιεύσεων, 41, 60, 85, 97, 104
Αυτοαναφορές, 86

B

Βαθμολογία Επίδρασης του Άρθρου, 7, 95
βαθμός Erdős, 149
βιβλιογραφική έρευνα, 49
Βιβλιογραφική Ζεύξη, 81
Βιβλιομετρία, 12, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 36, 37, 40, 45, 47, 48, 59, 72, 73, 76, 78, 87, 88, 108, 109, 154, 155, 158
βιβλιομετρικοί δείκτες, 13, 34, 35, 40, 60, 61, 73, 76, 78, 91, 104

Δ

δείκτης του στέμματος, 107
Διαδικασία Αξιολόγησης της Έρευνας, 28, 32
δίπλωμα ευρεσιτεχνίας, 53
δυναμικό αναφοράς, 103

E

εκθέσεις αναφορών, 76
εξόρυξη στο Παγκόσμιο Ιστό, 147
εξωτερικό σύνδεσμο, 138
επικοινωνίας χωρίς σύνδεση, 152
Επιστήμη των Πληροφοριών, 78
Επιστημομετρία, 26, 29, 30, 31, 83, 158
επιστημονικά πεδία, 12, 13, 17, 36, 60, 61
ερευνητικό έργο, 20, 49, 59, 95, 124
εσωτερικό σύνδεσμο, 138
Ευρετήριο Αμεσότητας, 92

Z

Ζεύγη Αναφορών, 79

Θ

θεωρία του δικτύου των δραστών, 83
θεωρία των γράφων, 148

I

Ιστομετρία, 31, 32, 33
Ιχνηλάτης αναφορών, 76

K

Κλασμάτωση και στάθμιση, 84, 85, 86
κοινωνικά δίκτυα, 148, 150
Κυβερνομετρία, 33

Λ

λογοκλοπή, 49, 53

Μ

μηχανές αναζήτησης, 109, 120, 121, 122, 123, 142, 143, 145, 146, 147

Π

Παγκόσμιος Ιστός, 27, 141

Παραγωγικότητα, 72

Περιγραφική Βιβλιομετρία, 34

περιοδικά, 12, 19, 20, 21, 26, 30, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 45, 46, 50, 53, 60, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 74, 75, 77, 80, 85, 91, 93, 94, 96, 98, 99, 100, 103, 104, 108, 112, 115, 120, 121, 124, 125, 129, 130, 135, 146, 147

Πληροφοριομετρία, 25, 29, 30, 31

ποιοτική ανάλυση, 17

ποσοτική ανάλυση, 12, 13, 17, 24, 59, 87

πρόβλημα της αφθονίας, 120

πρότυπα, 15, 37, 53, 68, 76, 77, 93, 100, 107

Πρότυπα δημοσίευσης, 76

Σ

συγκριτική αξιολόγηση, 25, 45

συν-αναφορά, 24, 26, 35, 83

συνδέεται με τον εαυτό του, 138

σύνδεση, 16, 26, 37, 50, 76, 94, 146, 147, 148, 149, 152

Σχεσιακή Βιβλιομετρία, 35

Σχετικός δείκτης απήχησης, 107

Τ

Τοπικό αποθετήριο, 84

Φ

φαινόμενο κατά Ματθαίον, 23, 24

Φαινόμενο Ματίλντα, 24

φυλλομετρητής, 121

Χ

Χάρτης Αναφορών, 76

Βιβλιογραφία/Αναφορές

- Aguillo, I. F. (1998). STM information on the Web and the development of new Internet R&D databases and indicators. In *ONLINE INFORMATION-INTERNATIONAL MEETING-* (pp. 239–244). LEARNED INFORMATION LTD.
- Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the World Wide Web: Methodological approaches to 'webometrics'. *Journal of Documentation*, 53(4), 404–426.
- Altmann, G. (1980). Prolegomena to Menzerath's law. *Glottometrica*, 2, 1–10.
- Andersen, H. (2000). Influence and reputation in the social sciences-how much do researchers agree? *Journal of Documentation*, 56(6), 674–692.
- Arasu, A., Cho, J., Garcia-Molina, H., Paepcke, A., & Raghavan, S. (2001). Searching the web. *ACM Transactions on Internet Technology (TOIT)*, 1(1), 2–43.
- Archambault, É., & Gagné, É. V. (2004). The use of bibliometrics in the social sciences and humanities. Montreal: Science-Metrix.
- Bar, Judit, Ilan Blyma and Bluma C. Peritz (2002), Informetric theories and methods for exploring the internet: an analytical survey of recent research literature, *Library Trends*, 50, 3 (371-392).
- Bensman, S. J. (2007). GARFIELD AND THE IMPACT FACTOR: THE CREATION, UTILIZATION, AND VALIDATION OF A CITATION MEASURE Part 2, The Probabilistic, Statistical, and Sociological Bases of the Measure By. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.160.6965>
- Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The Eigenfactor™ metrics. *The Journal of Neuroscience*, 28(45), 11433–11434.
- Björneborn, L. (2001). Small-world linkage and co-linkage. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia* (pp. 133–137). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=504252>
- Björneborn, L. (2004). *Small-world link structures across an academic web space: a library and information science approach*. Det Informationsvidenskabelige AkademiDanish School of Library and Information Science, Institut østInstitut øst. Retrieved from http://pure.iva.dk/ws/files/31034741/lennart_bjorneborn_phd.pdf
- Björneborn, L., & Ingwersen, P. (2001). Perspective of webometrics. *Scientometrics*, 50(1), 65–82.
- Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.
- Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.
- Bokos, M. P. G., Papavlasopoulos, N. K. S., & Avlonitis, M. (2007). Specific selection of FFT amplitudes from audio sports and news broadcasting for classification purposes. Retrieved from <http://www.emis.ams.org/journals/JGAA/accepted/2007/Poulos+2007.11.1.pdf>
- Bonitz, M. (1982). *Scientometrie, Bibliometrie, Infor-metrie*. Zentralblatt für Bibliothekswesen, 96(1), 19-24.
- Boudourides, M. A., Sigrist, B., & Alevizos, P. D. (1999). Webometrics and the self-organization of the European Information Society. *Draft Report, Task*, 2, 17–19.
- Bradford, S. C. (1985). Sources of information on specific subjects. *Journal of Information Science*, (10), 173–180.

- Broadus, R. N. (1987). Early approaches to bibliometrics. *Journal of the American Society for Information Science (1986-1998)*, 38(2), 127.
- Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., ... Wiener, J. (2000). Graph structure in the web. *Computer Networks*, 33(1), 309–320.
- Brody, T., Carr, L., Gingras, Y., Hajjem, C., Harnad, S., & Swan, A. (2007). Incentivizing the open access research web: publication-archiving, data-archiving and scientometrics. *CTWatch Quarterly*, 3(3). Retrieved from <http://eprints.ecs.soton.ac.uk/14418>
- Brody, T., Harnad, S., & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060–1072.
- Brookes, B. C. (1990). Biblio-, sciento-, infor-metrics?? what are we talking about? Retrieved from <https://uhdspace.uhasselt.be/dspace/handle/1942/857>
- Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis.
- Cassery, M. F., & Bird, J. E. (2003). Web citation availability: analysis and implications for scholarship. *College & Research Libraries*, 64(4), 300–317.
- Cattell, J. M. (1906). A statistical study of american men of science. III. The distribution of american men of science. *Science*, 732–742.
- Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5303–5310.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377.
- CitNetExplorerTutorial. Διαθέσιμοσε: www.citnetexplorer.nl/gettingstarted/. Ανακτήθηκετην 5ηΜαΐου 2015
- Clemens, E. S., Powell, W. W., McIlwaine, K., & Okamoto, D. (1995). Careers in print: Books, journals, and scholarly reputations. *American Journal of Sociology*, 433–494.
- Cole, S. (2004). Merton's Contribution to the Sociology of Science. *Social Studies of Science*, 34(6), 829–844.
- Comité national d'évaluation de la recherche (CNER). (2002). *Évaluation de la recherche publique dans les établissements publics français*. Paris.
- Crane, D. (1967). The gatekeepers of science: Some factors affecting the selection of articles for scientific journals. *The American Sociologist*, 195–201.
- Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. *Journal of Documentation*, 53(3), 263–273.
- Davison, B. D. (2000). Topical locality in the web. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 272–279). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=345597>
- Downie, S. J. (2013). Infometrics and the World Wide Web: A Case Study and Discussion. In *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI*. Retrieved from <http://www.cais-acsi.ca/ojs/index.php/cais/article/view/19>
- Earle, P., & Vickery, B. (1969). Social science literature use in the UK as indicated by citations. *Journal of Documentation*, 25(2), 123–141.
- Edge, D. (1977). Why I am not a co-citationist. *Society for Social Studies of Science Newsletter*, 2, 13–19.
- Egghe, L. (2000). New informetric aspects of the Internet: some reflections-many problems. *Journal of Information Science*, 26(5), 329–335.
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.

- Egghe, L., & Rousseau, R. (1990). Introduction to informetrics: Quantitative methods in library, documentation and information science. Retrieved from <http://eprints.rclis.org/handle/10760/6011>
- Elkana, Y., & others. (1978). *Toward a metric of science*. Wiley. Retrieved from <http://agris.fao.org/agris-search/search.do?recordID=US201300538654>
- Fernberger, S. W. (1917). On the number of articles of psychological interest published in the different languages. *The American Journal of Psychology*, 141–150.
- Garfield, E. (1974). *ISI is Studying Structure of Science Through Co-Citation Analysis*. INST SCI INFORM INC 3501 MARKET ST, PHILADELPHIA, PA 19104.
- Garfield, E. (1979). Controversies over opiate receptor research typify problems facing awards committees. *Current Contents*, (20), 5–18.
- Garfield, E. (1986). THE 250 MOST-CITED AUTHORS IN THE ARTS-AND-HUMANITIES CITATION INDEX, 1976-1983. *Current Contents*, (48), 3–10.
- Garfield, E., & Merton, R. K. (1979). *Citation indexing: Its theory and application in science, technology, and humanities* (Vol. 8). Wiley New York. Retrieved from <http://www.garfield.library.upenn.edu/cifwd.html>
- Garner, B. J. (1967). Models of urban geography and settlement location. *Models in Geography*, 303–360.
- Gingras, Y. (1984). La valeur d'une langue dans un champ scientifique. *Recherches Sociographiques*, 25(2), 286–296.
- Glänzel, W. (1996). A bibliometric approach to social sciences. National research performances in 6 selected social science areas, 1990–1992. *Scientometrics*, 35(3), 291–307.
- Glanzel, W. (2003). Bibliometrics as a research field a course on theory and application of bibliometric indicators. Retrieved from <http://nsdl.niscair.res.in/jspui/handle/123456789/968>
- Glänzel, W. (2003). Bibliometrics as a research field. Retrieved May, 25, 2006.
- Glänzel, W., & Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing & Management*, 35(1), 31–44.
- Grayson, L., & Gomersall, A. (2003). A difficult business: finding the evidence for social science reviews. *ESRC UK Centre for Evidence Based Policy and Practice. Working Paper*, 19. Retrieved from <http://www.kcl.ac.uk/sspp/departments/politiceconomy/research/cep/pubs/papers/assets/wp19.pdf>
- Gross, P. L., & Gross, E. M. (2007). College libraries and chemical education Science. 1927; 66 (1713): 385-389. *Smith DR. Historical Development of the Journal Impact Factor and Its Relevance for Occupational Health. Ind Health*, 45(6), 730–42.
- Gruhl, D., Bailey, J., Spohrer, J., & Maglio, P. P. (2007). Steps toward a science of service systems. *Computer*, (1), 71–77.
- Guerrero-Bote, V. P., & Moya-Anegón, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics*, 6(4), 674–688.
- Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10(6). Retrieved from <http://eprints.ecs.soton.ac.uk/10207>
- Harnad, S., Brody, T., & Vallières, F. (n.d.). The access/impact problem and the green and gold roads to open access. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/00987913.2004.10764930>
- Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., ... Hilf, E. R. (2008). The access/impact problem and the green and gold roads to open access: An update. *Serials Review*, 34(1), 36–40.
- Heydon, A., & Najork, M. (1999). Mercator: A scalable, extensible web crawler. *World Wide Web*, 2(4), 219–229.

- Hicks, D. (1995). Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change*, 4(2), 401–424.
- Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44(2), 193–215.
- Hicks, D. (2005). The four literatures of social science. In *Handbook of quantitative science and technology research* (pp. 473–496). Springer. Retrieved from http://link.springer.com/chapter/10.1007/1-4020-2755-9_22
- Hicks, D., & Potter, J. (1991). Sociology of scientific knowledge: a reflexive citation analysis of science disciplines and disciplining science. *Social Studies of Science*, 21(3), 459–501.
- Hodgson, G. M., & Rothman, H. (1999). The editors and authors of economics journals: A case of institutional oligopoly? *The Economic Journal*, 109(453), 165–186.
- Holmes, A., & Oppenheim, C. (2001). Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management. *Information Research*, 6(2), 6–2.
- Hubert, B., Gérard, J.-C., Bisikalo, D. V., Shematovich, V. I., & Solomon, S. C. (2001). The role of proton precipitation in the excitation of auroral FUV emissions. *Journal of Geophysical Research: Space Physics* (1978–2012), 106(A10), 21475–21494.
- Hummon, N. P., & Dereian, P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks*, 11(1), 39–63.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236–243.
- Ingwersen, P. (2000). The international visibility and citation impact of Scandinavian research articles in selected social science fields: the decay of a myth. *Scientometrics*, 49(1), 39–61.
- Ingwersen, P. (2006). Webometrics: ten years of expansion. Retrieved from <http://eprints.rclis.org/archive/00007554>
- Ingwersen, P. E. R. (1997). *The central international visibility of Danish and Scandinavian research 1988-1996*. Centre for Informetric Studies, Royal School of Librarianship.
- Ingwersen, P., & Christensen, F. H. (1997). Data Set Isolation for Bibliometric Online Analyses of Research Publications: Fundamental Methodological Issues. *JASIS*, 48(3), 205–217.
- Ingwersen, P., & Wormell, I. (1999). Publication behaviour and international impact: Scandinavian clinical and social medicine, 1988–96. *Scientometrics*, 46(3), 487–499.
- Ingwersen, P., Larsen, B., & Noyons, E. (2001). Mapping national research profiles in social science disciplines. *Journal of Documentation*, 57(6), 715–740.
- Karolinska Institutet. <http://ki.se/start>
- Klavans, R., & Boyack, K. W. (2006). Quantitative evaluation of large maps of science. *Scientometrics*, 68(3), 475–499.
- Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5), 604–632.
- Kochen, M. (1989). *The small world*. Ablex Norwood, NJ. Retrieved from <http://www.socsci.uci.edu/~bgrofman/R35%20Grofman%20and%20Feld.%201988.Toward%20a%20sociometric%20theory%20of%20rep....pdf>
- Koehler, W., & others. (2004). A longitudinal study of Web pages continued: a consideration of document persistence. *Information Research*, 9(2), 9–2.
- Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago press. Retrieved from https://www.google.com/books?hl=el&lr=&id=3eP5Y_OOuzwC&oi=fnd&pg=PR5&dq=The+Structure+of+Scientific+Revolutions+&ots=xU0JzboRvO&sig=K0FBI3rlRCH1wIfvTpP1n09mnfY

- Kurtz, M. J. (2004). *Restrictive access policies cut readership of electronic research journal articles by a factor of two, National Policies on Open Access (OA) Provision for University Research Output: an International meeting*. Southampton University, Southampton UK. 19 February 2004.
- Larivière V. and Godin B. (2001). *Profil de la production scientifique du Service canadien des forêts*. Report prepared for Canadian Forest Service.
- Larson, R. R. (1996). Bibliometrics of the World Wide Web: An exploratory analysis of the intellectual structure of cyberspace. In *Proceedings of the Annual Meeting-American Society for Information Science* (Vol. 33, pp. 71–78). Retrieved from <http://free-journal.umm.ac.id/files/file/asis96.ps>
- Lawrence, S. (2001). Online or invisible. *Nature*, 411(6837), 521.
- Lenoir, M. (1979). Promotion of patent literature. *World Patent Information*, 1(2), 84–87.
- Lewison, G. (2001). Evaluation of books as research outputs in history of medicine. *Research Evaluation*, 10(2), 89–95.
- Leydesdorff, L. (2003). Can networks of journal-journal citations be used as indicators of change in the social sciences? *Journal of Documentation*, 59(1), 84–104.
- Leydesdorff, L., & Curran, M. (2000). Mapping university-industry-government relations on the Internet: the construction of indicators for a knowledge-based economy. *Cybermetrics*, 4(1), 1–17.
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362.
- Li, X. (2003). A review of the development and application of the Web impact factor. *Online Information Review*, 27(6), 407–417.
- Li, X., Thelwall, M., Musgrove, P., & Wilkinson, D. (2003). The relationship between the WIFs or inlinks of Computer Science Departments in UK and their RAE ratings or research productivities in 2001. *Scientometrics*, 57(2), 239–255.
- Line, M. B. (1981). The structure of social science literature as shown by a large-scale citation analysis. *Social Science Information Studies*, 1(2), 67–87.
- López Yepes, J. (1995). La documentación como disciplina: teoría e historia. actual. y ampli. *Panplona: EUNSA*.
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of Washington Academy Sciences*. Retrieved from <http://doi.apa.org/?uid=1926-10141-003>
- Lundberg, J. (2006). *Bibliometrics as a research assessment tool: impact beyond the impact factor*. Institutionen för lärande, informatik, management och etik, LIME/Department of Learning, Informatics, Management and Ethics (Lime). Retrieved from <https://publications.ki.se/xmlui/handle/10616/39489>
- Lundberg, J. (2007). Lifting the crown—citation z-score. *Journal of Informetrics*, 1(2), 145–154.
- MacRoberts, M. H., & MacRoberts, B. R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, 61(1), 1–12.
- Martyn, J. (1964). Bibliographic coupling. *Journal of Documentation*, 20(4), 236–236.
- McKiernan, P. (1996). *Historical evolution of strategic management* (Vol. 2). Dartmouth Publishing.
- Meho, L. I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: a comparison of Scopus and Web of Science. *Journal of the American Society for Information Science and Technology*, 59(11), 1711–1726.
- Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105–2125.

- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press. Retrieved from <https://www.google.com/books?hl=el&lr=&id=zPvcHuUMEMwC&oi=fnd&pg=PR9&dq=The+Sociology+of+Science:+Theoretical+and+Empirical+Investigations&ots=x5QIQmbavM&sig=hfWufcJoat-4WNmuCGe6Dh2hDz0>
- Mettrop, W., & Nieuwenhuysen, P. (2001). Internet search engines-fluctuations in document accessibility. *Journal of Documentation*, 57(5), 623–651.
- Milgram, S. (1967). The small world problem. *Psychology Today*, 2(1), 60–67.
- Moed, H. F. (2006). *Citation analysis in research evaluation* (Vol. 9). Springer Science & Business Media. Retrieved from <https://www.google.com/books?hl=el&lr=&id=D9SaJ6awy4gC&oi=fnd&pg=PR9&dq=Citation+analysis+in+research+evaluation.&ots=FGmUKy6NjX&sig=YJ1go9ntOthq5Zqz1lmEJLIKqGY>
- Moed, H. F., Luwel, M., & Nederhof, A. J. (2002). Towards research performance in the humanities. *Library Trends*, 50(3), 498–520.
- Moed, H., De Bruin, R., & Van Leeuwen, T. H. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics*, 33(3), 381–422.
- Nacke, O. (1979). Informetrie: Ein neuer Name für eine neue Disziplin. *Nachrichten Für Dokumentation*, 30(6), 219–226.
- Nalimov, V. V. (1970). Influence of mathematic statistics and cybernetics on the methodology of scientific investigations. *Zavodskaya Laboratoriya*. 36 (10). *English Translation in Industrial Laboratory*, 36(10), 1549–1558.
- Nalimov, V. V., & Mulczenko, Z. M. (1969). *Naoukometria [Scientometrics. The study of science as an information process]*. Moscow: Nauka
- NALIMOV, V. V., I. V. KORDON, A. YA. KORNEEVA (1971), *Geograficheskoe Raspredelenie Nauchnoi Informatsii. [Geographic Distribution of Scientific Information.]* Informatsionnye Materialy. Moscow: an SSSR Nauchnyi Sovet po Kompleksnoi Probleme Kibernetiki. [Informational Papers. Moscow: Soviet Academy of Science, Scientific Council on Cybernetics. 2 : 3–37.
- Nance, R. E., Korfhage, R. R., & Bhat, U. N. (1972). Information networks: definitions and message transfer models. *Journal of the American Society for Information Science*, 23(4), 237–247.
- Nederhof, A. J., & Noyons, E. C. (1992). International comparison of departments' research performance in the humanities. *Journal of the American Society for Information Science*, 43(3), 249–256.
- Nederhof, A. J., & Zwaan, R. A. (1991). Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. *Journal of the American Society for Information Science*, 42(5), 332–340.
- Nederhof, A. J., & Van Raan, A. F. (1993). A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy*, 22(4), 353–368.
- Nederhof, A. J., Zwaan, R. A., De Bruin, R. E., & Dekker, P. J. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioral sciences: A comparative study. *Scientometrics*, 15(5-6), 423–435.
- Nederhof, A., Luwel, M., & Moed, H. (2001). Assessing the quality of scholarly journals in linguistics: An alternative to citation-based journal impact factors. *Scientometrics*, 51(1), 241–265.
- Newman, M. E., & Watts, D. J. (1999). Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4), 341–346.
- Nordstrom, L. (1990). 'Bradford's law' and the relationship between ecology and biogeography. *Scientometrics*, 18(3-4), 193–

- Kostagiolas, P., Papadaki, E., Kanlis, G., & Papavlasopoulos, S. (2013). Responding to crises with alliances: evidence from an academic library survey in Greece. *Advances in Librarianship*, 36, 247–279.
- Nordstrom, L. (1990). 'Bradford's law' and the relationship between ecology and biogeography. *Scientometrics*, 18(3-4), 193–203.
- Odlyzko, A. (2002). The rapid evolution of scholarly communication. *Learned Publishing*, 15(1), 7–19.
- Otlet, P. (1920). L'Organisation internationale de la bibliographie et de la documentation. Inst. Internat. de Bibliogr.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: bringing order to the Web. Retrieved from <http://ilpubs.stanford.edu:8090/422>
- Palmquist, Ruth A. Bibliometrics, University of Texas, <http://www.gslis.utexas.edu/~palmauis/courses/biblio.html> [πρόσβαση 24-7-2015].
- Papavlasopoulos, S., & Poulos, M. (2012). Neural network design and evaluation for classifying library indicators using personal opinion of expert. *Library Management*, 33(4/5), 261–271.
- Pareto, V., & Page, A. N. (1971). Translation of Manuale di economia politica ('Manual of political economy'). *AM Kelley*.
- PERITZ, J. B.-I. B. (2002). Informetric theories and methods for exploring the Internet: An analytical survey of recent research literature. *ILLINOIS*, 50(3), 371–392.
- Pirolli, P., Pitkow, J., & Rao, R. (1996). Silk from a sow's ear: extracting usable structures from the Web. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 118–125). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=238450>
- Piwowar, H. (2013). Altmetrics: Value all research products. *Nature*, 493(7431), 159–159.
- Price, D.J. de Solla (1976). *A general theory of bibliometric and other cumulative advantage processes*. Journal of the American Society for Information Science 27 (5), 292–306. doi:10.1002/asi.4630270505
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv Preprint arXiv:1203.4745*. Retrieved from <http://arxiv.org/abs/1203.4745>
- Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, (25), 348–349.
- PubMed Overview (2006, June 30, 2006). Retrieved Sept 20th, 2006, from <http://www.ncbi.nlm.nih.gov/entrez/query/static/overview.html>
- Rehm, G. (2002). Towards automatic Web genre identification: a corpus-based approach in the domain of academia by example of the Academic's Personal Homepage. In *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on* (pp. 1143–1152). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=994036
- Restivo, S. P. (1994). *Science, society, and values: Toward a sociology of objectivity*. Lehigh University Press. Retrieved from https://www.google.com/books?hl=el&lr=&id=h0z5zdPAR0gC&oi=fnd&pg=PR9&dq=Toward+a+Sociology+of+Objectivity&ots=AGBJKUR--t&sig=ek_adOiOf0XtgGKw-uPS65F850c
- Rousseau, J.-J., & Gourevitch, V. (1997). *Rousseau: 'The Social Contract' and Other Later Political Writings*. Cambridge University Press. Retrieved from https://www.google.com/books?hl=el&lr=&id=kcvseZCgQKMC&oi=fnd&pg=PR7&dq=%27The+Social+Contract%27and+Other+Later+Political+Writings&ots=CcVmpfI95w&sig=hawAx1ppaz_L88iuaCVFAFJ73PU
- Rousseau, R. (1997). *Sitations: an Exploratory Study*. *Cybermetrics*, [Online], 1 (1).
- Rousseau, R. (1997). *Sitations: an Exploratory Study*. *Cybermetrics*, [Online], 1 (1).
- Rousseau, R. (2002). Journal evaluation: Technical and practical issues. *Library Trends*, 50(3), 418–439.

- Royle, P., & Over, R. (1994). The use of bibliometric indicators to measure the research productivity of Australian academics. *Australian Academic & Research Libraries*, 25(2), 77–88.
- Rubio, A. V. (1992). Scientific production of Spanish universities in the fields of social sciences and language. *Scientometrics*, 24(1), 3–19.
- Sala-i-Martin, X. X. (1997). *I just ran four million regressions*. National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w6252>
- Schiminovich, S. (1971). Automatic classification and retrieval of documents by means of a bibliographic pattern discovery algorithm. *Information Storage and Retrieval*, 6(6), 417–435.
- Schoepflin, U. (1992). Problems of representativity in the social sciences citation index. *Representations of Science and Technology. Leiden: DSWO*, 177–188.
- Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica Scandinavica*, 69(3), 224–229.
- Sentinel Event Glossary of Terms. (2006). Retrieved Sept 27th, 2006, from http://www.jointcommission.org/SentinelEvents/se_glossary.htm
- Shapiro, F. R. (1992). Origins of bibliometrics, citation indexing, and citation analysis: the neglected legal literature. *Journal of the American Society for Information Science*, 43(5), 337–339.
- Simkin, M. V., & Roychowdhury, V. P. (2012). Theory of citing. In *Handbook of Optimization in Complex Networks* (pp. 463–505). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4614-0754-6_16
- Skupin, A., Biberstine, J. R., & Börner, K. (2013). Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PloS One*, 8(3), e58779.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269.
- Small, H. G. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science*, 139–166.
- Small, H., & Griffith, B. C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies*, 17–40.
- Snyder, H., & Rosenbaum, H. (1999). Can search engines be used as tools for web-link analysis? A critical view. *Journal of Documentation*, 55(4), 375–384.
- Spink, A., Jansen, B. J., Wolfram, D., & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. *Computer*, 35(3), 107–109.
- Spink, A., Wolfram, D., Jansen, M. B., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3), 226–234.
- Tagliacozzo, R., & others. (1967). Citations and citation indexes: a review. *Methods Inf Med*, 6, 136–142.
- Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing & Management*, 28(1), 1–3.
- Testa, J., & McVeigh, M. E. (2004). The impact of open access journals: A citation study from Thomson ISI. Çevrimiçi Elektronik Adres: [Http://www. Thomsonscientific. Com/media/presentrep/acropdf/impact-Oa-Journals. Pdf](http://www.Thomsonscientific.Com/media/presentrep/acropdf/impact-Oa-Journals.Pdf) [15.10. 2005'te Erişildi].
- Thelwall, M. (2001). Extracting macroscopic information from web links. *Journal of the American Society for Information Science and Technology*, 52(13), 1157–1168.
- Thelwall, M. (2002). Conceptualizing documentation on the Web: An evaluation of different heuristic-based models for counting links between university Web sites. *Journal of the American Society for Information Science and Technology*, 53(12), 995–1005.

- Thelwall, M. (2003). What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, 8(3), 8–3.
- Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1–116.
- Thelwall, M., & Harries, G. (2003). The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university. *Journal of the American Society for Information Science and Technology*, 54(7), 594–602.
- Thelwall, M., & Harries, G. (2004). Do the web sites of higher rated scholars have significantly more online impact? *Journal of the American Society for Information Science and Technology*, 55(2), 149–159.
- Thelwall, M., & Wilkinson, D. (2004). Finding similar academic Web sites with links, bibliometric couplings and colinks. *Information Processing & Management*, 40(3), 515–526.
- Thelwall, M., Vaughan, L., & Björneborn, L. (2005). Webometrics. *Annual Review of Information Science and Technology*, 39(1), 81–135.
- Thomson Scientific: Citation Products. Retrieved April 6th, 2006, from <http://scientific.thomson.com/products/categories/citation/>
- Turnbull, Don. Bibliometrics and the World-Wide Web, University of Toronto, <http://www.ischool.utexas.edu/~donturn/research/bibweb.html> [πρόσβαση 24-7-2007].
- van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring Scholarly Impact* (pp. 285–320). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-10377-8_13
- Van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., & Van Raan, A. F. (2003). The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57(2), 257–280.
- van Raan, A. F. (1998). *Special Topic Issue: Science and Technology Indicators*. American Society for Information Science (ASIS).
- van Raan, A. F. (2001). Bibliometrics and Internet: Some observations and expectations. *Scientometrics*, 50(1), 59–63.
- Varki, S., & Wong, S. (2003). Consumer involvement in relationship marketing of services. *Journal of Service Research*, 6(1), 83–91.
- Vaughan, L., & Hysen, K. (2002). Relationship between links to journal Web sites and impact factors. In *Aslib Proceedings* (Vol. 54, pp. 356–361). MCB UP Ltd. Retrieved from <http://www.emeraldinsight.com/doi/abs/10.1108/00012530210452555>
- Vaughan, L., & Thelwall, M. (2004). Search engine coverage bias: evidence and possible causes. *Information Processing & Management*, 40(4), 693–707.
- Visser, J. A. (2003). AMH signaling: from receptor to target gene. *Molecular and Cellular Endocrinology*, 211(1), 65–73.
- Wallin, J. A. (2005). Bibliometric methods: pitfalls and possibilities. *Basic & Clinical Pharmacology & Toxicology*, 97(5), 261–275.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440–442.
- Webster, B. M. (1998). Polish sociology citation index as an example of usage of national citation indexes in scientometric analysis of social sciences. *Journal of Information Science*, 24(1), 19–32.
- Weinberg, B. H. (1997). The earliest Hebrew citation indexes. *Journal of the American Society for Information Science*, 48(4), 318–330.

- What's the Difference Between MEDLINE and PubMed? Fact Sheet. (2006, 21 December 2004). Retrieved Sept 20th, 2006, from http://www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html
- White, H. D., & Griffith, B. C. (1982). Authors as markers of intellectual space: Co-citation in studies of science, technology and society. *Journal of Documentation*, 38(4), 255–272.
- White, H. D., & McCain, K. W. (1989). Bibliometrics. *Annual Review of Information Science and Technology*, 24, 119–186.
- Williams, M. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society*, 16(3), 201–215.
- Wilson, C. S. (1999). Informetrics. *Annual Review of Information Science and Technology (ARIST)*, 34, 107–247.
- Winclawska, B. M. (1996). Polish sociology citation index (principles for creation and the first results). *Scientometrics*, 35(3), 387–391.
- Wormell, I. (1998). Informetrics: an emerging subdiscipline in information science. *Asian Libraries*, 7(10), 257–268.
- Wyllys, R. E. (1981). Empirical and theoretical bases of Zipf's law. *Library Trends*, 30(1), 53–64.
- Yitzhaki, M. (1998). The 'language preference' in sociology: Measures of 'language self-citation', 'relative own-language preference indicator', and 'mutual use of languages'. *Scientometrics*, 41(1-2), 243–254.
- Zipf, G. K. (1949). Human behavior and the principle of least effort. Retrieved from <http://psycnet.apa.org/psycinfo/1950-00412-000>
- Μπουντουρίδης, Μωυσής Α. (1999). <http://hyperion.math.upatras.gr/courses/sts/>, Πανεπιστημίου Πατρών
- Σαχίνη, Ε., Μάλλιου Ν., Χούσος Ν. (2014). *Βιβλιομετρική Ανάλυση: Μεθοδολογική Προσέγγιση* ΕΚΤ, Εθνικό Κέντρο Τεκμηρίωσης 2^η Έκδοση.
- Σίτας, Α. Κώνστα, Ο. Ντούλια, Π. (2007). Βιβλιομετρική-Πληροφοριομετρική-Ιστομετρική. Δημοσιευμένο στο: Μιχάλης Τζεκάκης: ο οραματιστής. Ένωση Ελλήνων Βιβλιοθηκονόμων και Επιστημόνων της Πληροφόρησης, Xerox, 44-63.

Δικτυογραφία

- <http://arxiv.org/>
- <http://citeseerx.ist.psu.edu/>
- <http://books.google.com/>
- <http://scholar.google.com/>
- <http://iopscience.iop.org/page/archive>
- <http://www.ams.org/mathscinet/>
- <http://adswww.harvard.edu/>
- <http://highwire.stanford.edu/>
- <http://prola.aps.org/>
- <http://www.cas.org/products/sfacad/index.html>
- http://scitation.aip.org/search_scitation?key=SPIN

<http://www.scirus.com/>

<http://www.sciencedirect.com/>

<http://www.scopus.com/home.url>

<http://www.google.com/>

<http://www.harzing.com/pop.htm>. Retrieved Sept 2th, 2015,

Karolinska Institutet. <http://ki.se/start>