

Jun 2nd, 12:00 AM

Using Citation Data for Purchase Decisions: Analysing Citing Patterns and Journal Holdings at the Royal Institute of Technology

Cecilia Heyman Widmark
Royal Institute of Technology, cheyman@kth.se

Cecilia Heyman Widmark, "Using Citation Data for Purchase Decisions: Analysing Citing Patterns and Journal Holdings at the Royal Institute of Technology." *Proceedings of the IATUL Conferences*. Paper 1. <https://docs.lib.purdue.edu/iatul/2014/acquisition/1>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

USING CITATION DATA FOR PURCHASE DECISIONS: ANALYSING CITING PATTERNS AND JOURNAL HOLDINGS AT THE ROYAL INSTITUTE OF TECHNOLOGY

CECILIA HEYMAN WIDMARK

The Royal Institute of Technology, Sweden
cheyman@kth.se

ABSTRACT

This paper describes how citation data can be used for identifying gaps in journal holdings and in that way form a foundation for acquisitions. Citation data was matched against e-journal holdings using Web of Science™ and export files from a central knowledge base. Data for three years (2010-2012) was used, in total from 6 246 publications containing 130 090 references to 5 216 journals.

Furthermore, impact factors from Journal Citation Reports™ were added as well as information about publisher and if the journal was open access or not. The journals were also enriched with subject headings. The latter information was drawn from the database Ulrich's web™.

The output was divided according to which of the nine different schools of the institution the first author was affiliated to, each school being subject specific. (*i.e.*: *Architecture, Biotechnology, Chemistry, Computer Science, Engineering*)

Analysis of citations to journals held or not held by the library formed an excellent foundation for future demand driven purchase decisions. Also, conclusions could be drawn about citing patterns to high impact journals, how open access journals were cited and which publishers were most highly cited.

A specific analysis was performed within the life sciences as new research groups were demanding e-resources within subject areas not traditionally associated with a technological university. The data could confirm the accuracy of recent major investments as well as give support for future purchases. The data also shows how journal holdings in general match the different subject areas of the institution and in that way can provide a valuable basis for future budget discussions with the faculty.

Keywords: Citation analysis, Demand Driven Acquisitions, Collection Development

BACKGROUND

The Royal Institute of Technology (KTH) in Stockholm is the largest and oldest technical university in Sweden with education and research that spans from natural sciences to all branches of engineering. Biotechnology research at KTH is internationally well recognized. Consequently the KTH Library needs to cover a broad span of subjects when providing relevant scientific information to support the research and education within the institution.

The Library is part of one of the ten schools of KTH - The School of Education and Communication in Engineering Science. (The ECE School) The other nine schools are subject specific and main focus lies on research and education within their different areas.

The ECE School provides support for research and education for the other schools but also conducts its own educational programmes and courses mainly within pedagogy and teaching for higher education. The Library is one of three departments within ECE and is divided into three main units: Library services, Media & IT and Publishing services.

In the transition to digital collections and also because of how content is licensed new methods for collection development are necessary and in 2013 KTH Library adopted a new Media plan and policy for acquisitions. The most important result was the decision to focus on Demand

Driven Acquisitions and electronic resources. No printed material was longer to be collected and purchases of such avoided if possible.

The library receives a fair amount of purchase suggestions from researchers and students but not from all departments within KTH. In addition most publishers deliver statistics that can be useful, for instance turn-aways on non-subscribed content. Despite this we felt a need for supplementary sources - especially to cover the information needs on subjects not traditionally associated with KTH and particularly within the life sciences. Historically these subjects had been less well covered and it was necessary to expand e-journal holdings in these areas. Despite the relatively poor coverage researchers seldom approached the library with requests for new subscriptions. One explanation could be their close cooperation with colleagues from institutions which were better provided and with whom scholarly sharing could be practised. Findings in the UK Survey of Academics 2012, conducted by Ithaka S+R, Jisc, and RLUK show different options on how researchers tend to move on when access isn't immediate through subscriptions. (Figure 1)

Nevertheless it was our commission to improve journal holdings and better cover all areas of research carried out at KTH. An idea arose to analyse journals cited by researchers from the nine schools during the past few years and compare these to what journals the library subscribed to in order to identify gaps amongst highly cited (i.e. well used) e-journals.

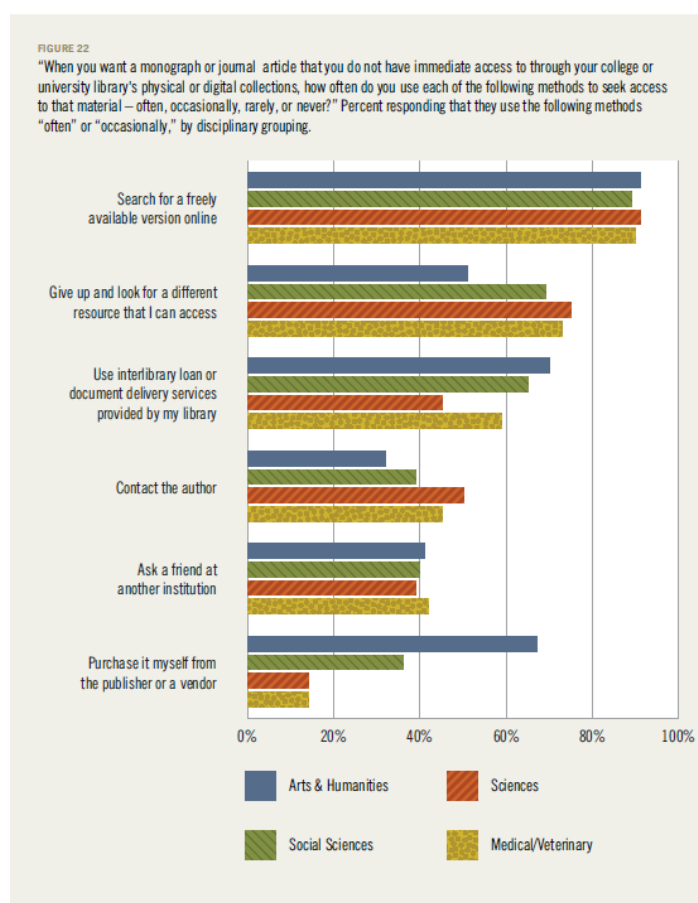


Figure 1: Findings on UK researchers' behaviour when access to material is not immediate.¹

¹ <http://www.sr.ithaka.org/research-publications/ithaka-sr-jisc-rluk-uk-survey-academics-2012>

METHOD

Records were withdrawn from Web of Science™ for all publications authored by at least one KTH affiliated researcher during three years, 2010-2012. The data was analysed with BibExcel™² that extracts the references included in each record. The data was then formatted and imported to MS Excel™. In total there were 6 246 publications containing 130 090 references to 5 216 journals.³ Only publications with less than 30 authors were part of the analysis.

Thereafter e-journal holdings for the library were exported from the central knowledge base (CKB) of the link resolver. After the data had been cleaned up it could be matched with the citation data. Each journal held or not held by the library was listed as well as the number of citations (i.e. references) each journal had received. The result was then easy to sort and filter regarding number of citations and if the library had access or not.

Furthermore, impact factors from Journal Citation Reports™ were added as well as information about publisher and if the journal was open access or not. The journals were also enriched with subject headings. The latter information was drawn from Directory of Open Access Journals (DOAJ) and the database Ulrich's web™.⁴

The output was also divided according to which of the nine different schools of the institution the first author was affiliated to, each school being subject specific. For the overall analysis the output was considered as a whole but dividing the data could for example confirm that the library had done less well covering the need for the School of Biotechnology.

RESULTS

We could see that the majority of gaps existed within medicine and life sciences. One explanation is the emergence of Science for Life Laboratory (SciLifeLab) which is a huge investment from the Swedish Government and a joint research collaboration between four major academic institutions – Karolinska Institutet, Stockholm University, Uppsala University and The Royal Institute of Technology. The publications with authors affiliated to SciLifeLab and KTH were an important share of the total for KTH since these researchers since 2010 have been publishing a growing number of journal articles.

In total we lacked 456 journals that had received 5 citations or more during the time period analysed. When the results had been more closely examined by staff dedicated to e-resource acquisitions it was obvious that the costs for subscribing to that large number of journals would be very high. We therefore made a selection and set the limit to journals which had received 30 citations or more. The approximate cost for these subscriptions was estimated to EUR 41 700. Of these 49 journals 12 had ceased or transferred to Open Access. (*Figure 2*)

As we had data on publishers included in the output we could quite quickly detect a number of journals that probably were part of publisher package deals and identify such packages, some of which were available through the Swedish national consortia.

Since the main goal of the study was to identify gaps and possible purchases we focused on this work, but the supplementary data that was added also illustrated other aspects - for instance usage of high impact factor journals and Open Access publications.

² A toolbox developed by Olle Persson, Umeå universitet. <http://www8.umu.se/inforsk/Bibexcel/>

³ Only journals with an identifier (ISSN-number) are included in the sum above. The sum of references and publications are not deduplicated so publications can be co-authored between the nine schools and therefore included multiple times.

⁴ The methodology for the analysis has been developed by Peter Sjögarde and Nils Jansson at KTH Library whom the author would like to express gratitude to. Contact information:
Peter Sjögarde, Bibliometric analyst: sjogarde@kth.se
Nils Jansson, Librarian: nilsjan@kth.se

A	B	C	D	E	F	G
Citations	Abbreviated title	Title	ISSN	KTH Sci	Price	Comment
245	MDL CELL PROTEOMICS	Molecular & Cellular Proteomics	1535-9476	BIO	6650	price 2013
138	CANCER RES	Cancer Research	0008-5472	BIO	see row 35	
132	J NEUROSCI	Journal Of Neuroscience	0270-6474	CSC	3172	price 2014
125	IRONMAK STEELMAK	Ironmaking & Steelmaking	0301-9233	ITM	10070	price 2013
94	APPL OPTICS	Applied Optics	1559-128X	ICT	9061	price 2013
90	J NEUROPHYSIOL	Journal Of Neurophysiology	0022-3077	CSC	10981	price 2013
78	AVIAT SPACE ENVIR M	Aviation Space And Environmental Medicine	0095-6562	STH	1944	price 2014
77	GENOME RES	Genome Research	1088-9051	BIO	11970	price 2013
75	METALL TRANS	Metallurgical Transactions	0026-086X	ITM	see row 36	
74	METALL TRANS B	Metallurgical Transactions B - Process Metallurgy	0360-2141	ITM	see row 36	
69	METALL TRANS A	Metallurgical Transactions A - Physical Metallurgy And Materials Science	0360-2133	ITM	see row 36	
62	AM J PHYSIOL-HEART	American Journal Of Physiology-Heart And Circulatory Physiology	0363-6135	SCI	7810	price 2013
62	CIRCULATION	Circulation	0009-7322	STH	18129	approx. price 2014
60	CLIN CANCER RES	Clinical Cancer Research	1078-0432	BIO	see row 35	
59	J CLIN ONCOL	Journal Of Clinical Oncology	0732-183X	BIO	8833	price 2013
58	J IRON STEEL I	Iron And Steel Institute. Journal	0021-1567	ITM	9061	price 2013
57	AM J PHYSIOL	American Journal Of Physiology (Consolidated)	0002-9513	SCI	30318	price 2013
56	J IMMUNOL	Journal Of Immunology	0022-1767	BIO	9754	price 2014
54	CIRC RES	Circulation Research	0009-7330	SCI	21935	approx. price 2014
50	T METALL SOC AIME	Metallurgical Society Of Aime. Transactions	0543-5722	ITM	see row 36	
41	Z METALLKD	Zeitschrift Fuer Metallkunde	0044-3093	ITM	13523	price 2013
39	J ASS ASPHALT PAVINC	Association Of Asphalt Paving Technologists. Journal	0270-2932	ABE	1364	price 2013
38	NEW ENGL J MED	New England Journal Of Medicine	0028-4793	BIO	33477	price 2013
38	SOIL SCI SOC AM J	Soil Science Society Of America Journal	0361-5995	ABE	4747	price 2013
37	MED SCI SPORT EXER	Medicine And Science In Sports And Exercise	0195-9131	STH	9037	price 2013
37	TAPPI J	Tappi Journal	0734-1415	CIIC	2517	
34	AM MINERAL	American Mineralogist	0003-004X	ITM	6991	price 2014
34	BLOOD	Blood	0006-4971	BIO	10709	price 2013
34	CHEMUSUSCHEM	Chemsuschem	1864-5631	CHE	26482	price 2013
34	PROTEIN ENG	Protein Engineering	0269-2139	BIO	11629	price 2013
34	STUD SURF SCI CATAL	Studies In Surface Science And Catalysis	0167-2991	CHE		ScienceDirect
34	Z PHYS	Zeitschrift Fuer Physik Nytt namn: European Physical Journal A - online only	0044-3328	SCI	31230	price 2013
30	RADIOLOGY	Radiology	0033-8419	SCI	13969	price 2014
30	COMBO - INCL CANCER RESEARCH AND CLINICAL CANCER RESEARCH (SINGLE-SITE)				18416	price 2013
30	METALLURGICAL AND MATERIALS TRANSACTIONS A & B - online only				36322	price 2013
					380 100 SEK (41 761 EUR)	
Abbreviated title	Title	ISSN	KTH Sci	Price	Comment	
40	AAPG BULL	Aapg Bulletin	0149-1423	ITM		Open access
37	ARCH EISENHUTTENW	Archiv Fuer Das Eisenhuettenwesen	0003-8962	ITM		Ceased
31	J ENVIRON QUAL	Journal Of Environmental Quality	0047-2425	ABE		Ceased
47	J PULP PAP SCI	Journal Of Pulp And Paper Science	0826-6220	CHE		New title name, already held
53	J PULP PAP SCI	Journal Of Pulp And Paper Science	0826-6220	SCI		Ceased
35	PROG ELECTROMAGN	Progress In Electromagnetics Research-Pier	1559-8985	ICT		Open Access
34	PROG ELECTROMAGN	Progress In Electromagnetics Research-Pier	1559-8985	EES		Open Access
31	PROG THEOR PHYS	Progress Of Theoretical Physics	0033-068X	SCI		Open Access
64	SCAND J METALL	Scandinavian Journal Of Metallurgy: Processes And Materials Engineering	0371-0459	ITM		Ceased
37	STEEL RES	Steel Research	0177-4832	ITM		Ceased
45	TAPPI	Tappi	0039-8241	CHE		Ceased
69	TETSU TO HAGANE	Tetsu To Hagane-Journal Of The Iron And Steel Institute Of Japan	0021-1575	ITM		Open Access
41	T IRON STEEL JPN	Transactions Of The Iron And Steel Institute Of Japan	1881-1183	ITM		Ceased
417	ISIJ INT	Isij International	0915-1559	ITM		Open Access
31	ARSENIC ENV	Arsenic In The Environment. Proceedings	2154-6568	ABE		Already held

Figure 2: Most cited journal not held by KTH including estimated price.

CONCLUSIONS

Analysis of citations to journals held or not held by the library formed an excellent basis for demand driven purchase decisions. Such models are developing and evolving for e-books but less so for e-journals. The aim was to identify gaps amongst e-journals, mainly within areas we knew to be poorly covered and from which the library tended to have less contact with researchers. Requests from users as well as turn-away statistics are both valuable for demand driven acquisitions for e-journals. Citation data is to be seen as a complement to the other two but is as thus a useful and fairly easy method to use. Once the methodology had been developed it was not very time-consuming.

Apart from this the outcome can also show if subscriptions and purchases are relevant and well-motivated by identifying highly cited journals that are held by the library. Furthermore the results can assist to detect journals that the library perhaps should cancel. In all cases the data can prove valuable in discussions regarding media budget and media holdings with faculty.

In the time of print journals it was a tedious and slow process to decide upon renewals, cancellations and new subscriptions. With the data and tools available today libraries can approach these matters more efficiently.

As said, the models for Demand Driven Acquisitions are mainly offered for e-book purchases. Perhaps the next step for publishers is to offer similar solutions on journal article level? Until then, we will have to rely on alternative methods to keep the journal holdings optimal, up-to-date and highly relevant.