

Paper. Second draft.

Document delivery as collection development.

An evaluation of the Ebrary "Patron Driven Acquisition" model at the University of Bergen Library.

By Senior Academic Librarian dr. philos Pål H. Bakka
University of Bergen Library
Library of Social Science and Music
Norway
pal.bakka@ub.uib.no

Paper, second draft.

INFORUM 2015: 21st Annual Conference on Professional Information Resources. Prague, May 26-27, 2015

Abstract

This paper analyzes and evaluates the University of Bergen Library Ebrary PDA-trial of 2012 to 2014. Its major finding is that by 2014 the volume of Ebrary e-book use was equal to 70,6 % of the comparable print circulation, with PDA books accounting for an estimated 89 % of this volume. The PDA trial must therefore be regarded as a success in terms of document delivery. In so far as collection development is concerned, no effect on the quality of the collection can be observed – except that all acquisitions actually are used by UBL institutional patrons.

History

In January 2012 Proquest invited the University of Bergen Library (UBL) to try out their “Patron driven acquisition” model for electronic books. As the library already subscribed to Proquest’s Academic Complete service, the UBL accepted the invitation by defining it as a “project”, thus limited in time and not subject to the bureaucracy of public tender rules, on Jan. 11th 2012.

As the name says, PDA is an acquisition model where the library instead of buying a static e-book collection makes a catalogue of such books available to its patrons. The library pays for the use they make of the titles in the catalogue. Payment is made by depositing money on one or more accounts, “funds”, in Ebrary, with the user fee being debited from this deposit as and when books are used. Without entering into excessive detail, the Ebrary PDA system recognizes two different modes of book usage: the Short Time Loan, a rental service for individual titles, and direct purchase, “PDA” proper, where the book is purchased the first time it is used. On the fourth STL, a book is automatically subject to a PDA-purchase. In general terms, purchasing books through STL-transactions is more expensive than purchasing through direct PDA, as the STL fee of between 10 and 25 % of the book prices is added to the price when purchasing the title. The Ebrary PDA system also functions a regular bookseller. This means that even if titles are not in the PDA-profile they can be purchased immediately if requested. This cuts document delivery time down to 30 seconds and is an excellent way of ensuring patron satisfaction with the library. It is also rather satisfying for the librarian to satisfy patron demand in this manner.

The library controls the PDA-system through the “profile” of its catalogue. It can choose freely from the titles available and delimit the catalogue by subject, publisher, age, language and price. Setting up a profile is time-consuming, but once created it will be automatically updated.

At the UBL the initial profile was a “social science” one, because the Library of Social Science and Music was chosen to be the guinea pig for the system. With the profile created and USD 15 000,- deposited, the PDA-system was activated on Friday Jan 21st . By the following Tuesday, patrons had purchased 15 titles at the cost of USD 1150,83 (+ 25 % VAT). That made the Library Director wonder whether we could afford a PDA-system. Respite was achieved by tweaking the profile to make STL the default usage mode.

In the autumn of 2012 it was decided to extend the PDA-trial to all disciplines taught at the University of Bergen, with PDA-profiles being created for the Humanities, Law, Psychology, Education and Health, Medicine and Dentistry and Science as well. Due to the overlap between Law and Social Science (the EU) and the fact that the demand for English language legal literature in Norway is limited, the “Law” profile was funded by the SSL, not the Law library budget.

In so far as any problems were encountered during the initial trial period it was caused by the legacy nature of the Norwegian library system BIBSYS, which made updating the OPAC a time-consuming and convoluted affair; a state of affairs that persisted until March 2014. As a consequence the PDA-collection was neither regularly updated nor visible in the OPAC during two thirds of the trial period. But as the numbers shown below indicates, accessibility through the OPAC really made a difference. Do not believe those that tell you that students only use google. The quintupling of the UBL PDA spending from USD 42 962.54 in 2012 to USD 254,238.37 in 2014 is proof to the contrary.

Table 1. Ebrary PDA activity at UBL, Jan. 24th 2012 to Dec 31st 2015.

Year	Month	Titles acquired				Cost (YS)/title			
		PDA	STL	PDA-4 use	Library purchase	PDA	STL	PDA-4 use	Library purchase
2012	All	207	590	91	86	85	14	95	161
2013	All	159	1213	57	165	75	16	103	126
2014	Jan	23	102	6	44	100	16	103	120
	Feb	30	192	9	18	86	16	106	102
	Mar	46	396	8	28	89	17	105	116
	Apr	75	574	24	10	92	18	115	131
	May	65	594	18	21	74	19	125	155
	Jun	41	364	23	7	87	19	119	123
	Jul	39	304	17	22	82	23	137	167
	Aug	40	344	16	10	83	36	111	133
	Sep	42	476	23	16	91	37	150	151
	Oct	96	723	45	19	91	37	116	127
	Nov	70	671	50	13	81	36	141	103
	Dec	58	442	41	6	90	31	130	182
2014		625	5182	280	214	87	29	122	132

Source: Ebrary Orders and PDA trigger reports.

Notes:	Jan	Max PDA price increased to USD 200,-.
	Mar	PDA catalogue accessible in OPAC and discovery tool
	Jul	STL price raises from major publishers.

The most striking thing about the entire trial is the take-off after the catalogue became accessible through the library discovery systems in March 2014¹. STLs in particular exploded as the collection of approximately 300 000 PDA-titles became available to all patrons, not just the minority information literate enough to find the Ebrary native interface.

What to make of these numbers? How do we evaluate this conjoined acquisition and document delivery system?

Evaluating e-books: Problems of data comparability

A PDA system by its very nature combines Inter Library Lending and Collection Development into one single automated operation. That collection development should be patron driven is a truism. Libraries exist to serve their patrons immediate needs, not just their future. They emphatically do not exist just for the sake of the books alone or for the acquisition librarians. We have always bought the books the patrons asked us to buy. And some pioneering libraries have even experimented with PDA as a substitute for ILL. Just that seemed so obvious that when my library opened in 1990, such a policy was adopted for recently published titles. The decision was made more in order to ease the workload on the ILL librarian than for any other reason, was controversial from an acquisitions point of view not have been implemented at any other of the UBL lending libraries.

Spending what amounts to the price of a paperback, up to USD 45,-, for a loan of one day (or one week) may be regarded as a questionable use of a limited acquisitions budget. On the other hand, the patron's need for information is immediately satisfied without the library spending any effort on the transaction than paying the bill. There is a financial aspect as well: In 2010 the postage costs incurred by the UBL ILL-services corresponded to half of what was spent on printed books. ILL, being lending from "closed stacks", may result in the book being returned on pick-up because it did not satisfy the patron's needs.

Library lore is that 25 % of all books are never leave the shelves once put there. The UBL does not compare well in this respect. Using the 2008-2012 circulation of our 2008 acquisitions catalogued in 2008 and 2009 (97,3 % of all books ordered in 2008) as a representative sample these numbers came up:

¹ Our PDA budget estimate for 2014 was based on the 2013 level of spending. We were not prepared for the take-off in activity. But the doubling of STL fees alone cost us about USD 50 000,-. A comparable cost was incurred by the appreciation of the USD in autumn of 2014.

Table 2. UBL: Aggregate circulation 2008-2012 for 2008 acquisitions cataloged in 2008 and 2009. Titles. By UBL lending library.

Library	Titles acquired	Titles lent	All patrons (%)	Institutional patrons (%)
Humanities	3049	2095	68,7	58,8
Law	458	315	68,7	59,2
Medicine et al..	554	438	79,1	75,6
Psychology et al.	669	554	82,8	70,7
Science	902	629	69,7	60,3
Social science et al.	1617	1351	83,5	81
UBL	7249	5382	74,2	66,3

Note: Sample years are used because circulation and acquisitions data from the BIBSYS library system had to be collated manually. The year 2000 was chosen because it was the last “pre-digital” year at UBL. 2008 because it was the last year for which data were available.

There are solid reasons for the different circulation rates at the six lending libraries. They are discipline focused and thus subject to the requirements of the disciplines covered as well as the Faculties they serve. The Law Library get funding from the Law School to maintain collections that are surplus to requirements. The Humanities Library subscribes to source editions and collected works that by their very nature are not in continuous demand, but must be part of the collection of any self-respecting university library. What is notable is that institutional patron demand is significantly lower than the overall demand generated by other libraries and the general public. In the sample, the volume of interlibrary lending from the Psychology, Education and Health library was higher than that to the faculty (258 vs 239 books, 570 vs 496 loans).

At UBL books loans have been falling steadily since 1995. Table 3 shows aggregate circulation numbers for the sample period.

Table 3. UBL. Total circulation. Ebrary-compatible libraries. 2009-2012

	Own collection					Interlibrary loans				
	2009	2010	2011	2012	Total	2009	2010	2011	2012	Total
Humanities	45858	41671	39658	40029	215829	10187	9076	8550	8529	46879
Law	7232	6829	7247	7435	36673	679	837	664	576	3495
Medicine et al.	6230	6703	6645	5958	32075	551	528	507	380	2716
Psychology et al.	14639	14591	13245	12206	69317	1399	1278	1207	937	6314
Science	9119	8796	8496	8141	44278	1153	770	753	530	4499
Social Science et al.	24228	23375	22483	21554	111973	3915	2574	2712	2081	16915
Total	107306	101965	97774	95323	510145	17884	15063	14393	13033	80818

Note: Excluding the “non Ebrary compatible” Law and Medicine Libraries as well as all fiction.

Narrowing on comparable entities, English language monographs acquired in 2000 and 2008 lent to faculty and students at the UBL, the 2611 acquisitions of 2000 were lent a total of 13497 times, annualized at a rate for 1,23 % over the period. The corresponding numbers for the 4152 2008-acquisitions are 8747 and 3,68 %. If the 264 768 interlibrary borrowings made by UBL is included in the document delivery total – as they should be – the ratio drops to 1,01 and 1,92 % respectively. If the actual acquisition decisions were based on the same criteria as in the Ebrary PDA – purchase on the fourth usage episode – the total acquisition for 2008 would have been 942 – 22,7 % of the actual number. These titles however, account for 64,,6 % of the circulation of the titles purchased in 2008.

The table below shows the UBL circulation frequencies confirming library lore, with 27,1 % never lent and 22,1 % being lent 4 or more times. The sample is “Ebrary-comparable” books: English language scientific monographs published by “international” (US, British and Dutch) publishers purchased by the Humanities, Science, Social science or Psychology library at the; thus excluding fiction and reference work as well as the acquisitions of the Law and Medical libraries. The former because it purchases predominantly non-English language literature, the latter because the average price of medical literature was higher than the price limit of the UBL PDA trials.

Tab. 4. Lending by institutional user categories 2008-2012. 2008 acquisitions. Documents

Library	Aquisitions	Circulation		User category		
		Never	loans > 3	Undergraduates	Graduate students	Faculty
Humanities	1696	558	260	494	559	689
Psychology	565	164	136	114	279	205
Science	643	228	120	136	273	243
Social sciences	1274	244	409	573	700	932
All	4178	1194	925	1317	1811	2069

Based on circulation data the UBL collection development policy is a three-quarters success. Total success is neither obtainable nor desirable. A balanced academic collection has to include literature that may gather dust on the shelves for thirty years or more. It is, however, no goal in itself to swamp the stacks with catalogue heroes. In so far as the number of unfulfilled ILL-requests is less than 1 %, UBL document delivery may be regarded as more of a success. But the ILL numbers reflect past sins of omission in collection development; sins that may very well be present as well.

If patrons are given access to a substantial catalogue, as is the case in and PDA-collection, collection development occurs as and when patrons use books. And as they get the books they need, the Library discharges its document delivery responsibility at the same time.

In evaluating the PDA-model two questions must be answered:

- Quantity: Do the users use the system? How does Ebrary e-book document delivery compare to traditional methods?
- Quality: Do the books available meet the academic standards of library collection development?

How much is much? Comparing circulation and use.

As the UBL only subscribed to Ebrary Academic Complete from 2009, the circulation data from Ebrary do not cover the exact same period as that of the 2008 acquisition sample. I intend to work around this using UBL accession data in addition to the circulation data, basing my tentative conclusions on the assumption that the 2008 sample is representative of the general pattern of book circulation at UBL. That is, one years' worth of acquisitions will produce 3,68 % of the circulation for any given four year period.

Year	Undergraduates	Graduate students	Faculty	All
2000	23495	22118	16896	62509
2001	23900	21294	16705	61899
2002	25057	23140	17428	65625
2003	23166	26458	17986	67610
2004	21552	29358	19925	70835
2005	21046	34466	18644	74156
2006	18836	32659	18604	70099
2007	17065	30042	18849	65956
2008	19487	27593	20530	67610
2009	18927	25931	20845	65703
2010	18077	24497	19486	62060
2011	16724	23552	18477	58753
2012	15042	21335	16512	52889
All	262374	342443	240887	845704

Aggregate loans to institutional users at the four libraries in the sample decreased by 18,1 % from 2000 to 2012. The number of student and staff was stable throughout the period. Faculty and graduate borrowing has fluctuated

over the period, but remained stable until 2009/2010. Aggregate lending dropped by 24,2, with graduate borrowing declining rather less and staff borrowing more than the average. As UoB electronic resources, including Ebrary books, are available to students and faculty only, a decline in the lending of physical books to institutional users is to be expected. The decline in borrowing after 2009/2010 is steeper than prior to that date.

ILL at the four libraries declined as shown by the table below:

	2009	2010	2011	2012	2013	2014	All
Humanities	10187	9603	8549	8529	7427	6841	51136
Psychology	1399	1372	1207	937	647	434	5996
Science	1153	850	754	630	587	388	4362
Social Science	3915	2767	2712	2081	2002	1397	14874
All	16654	14592	13222	12177	10663	9060	76368

ILL book delivery declined even faster with an overall decline by 36,4 % from 2009 to 2012, followed by a further decline of 34,0 % between 2012 to 2014.

It is hard not to attribute the decline in book lending at the UBL to an increasing availability of e-books to institutional patrons. The number of books accessible through Academic Complete tripled to 120 000 from 2008 to 2014. The Ebrary PDA profile added in the region of 250 000 titles to the UBL e-book catalogue. By the time the entire PDA collection became accessible to the patrons in March 2014, the UBL Ebrary collection comprised in the region 400 000 titles.

Measuring the effect of this on document delivery as unambiguously as in “loans” is another matter entirely, however. Consider the two frequency distributions below:

Fig. 1. UBL: 2008 acquisitions. Ebrary comparable titles. Circulation 2008-2012. Loan frequencies.

(N=4152. m=2,17. Max=38)

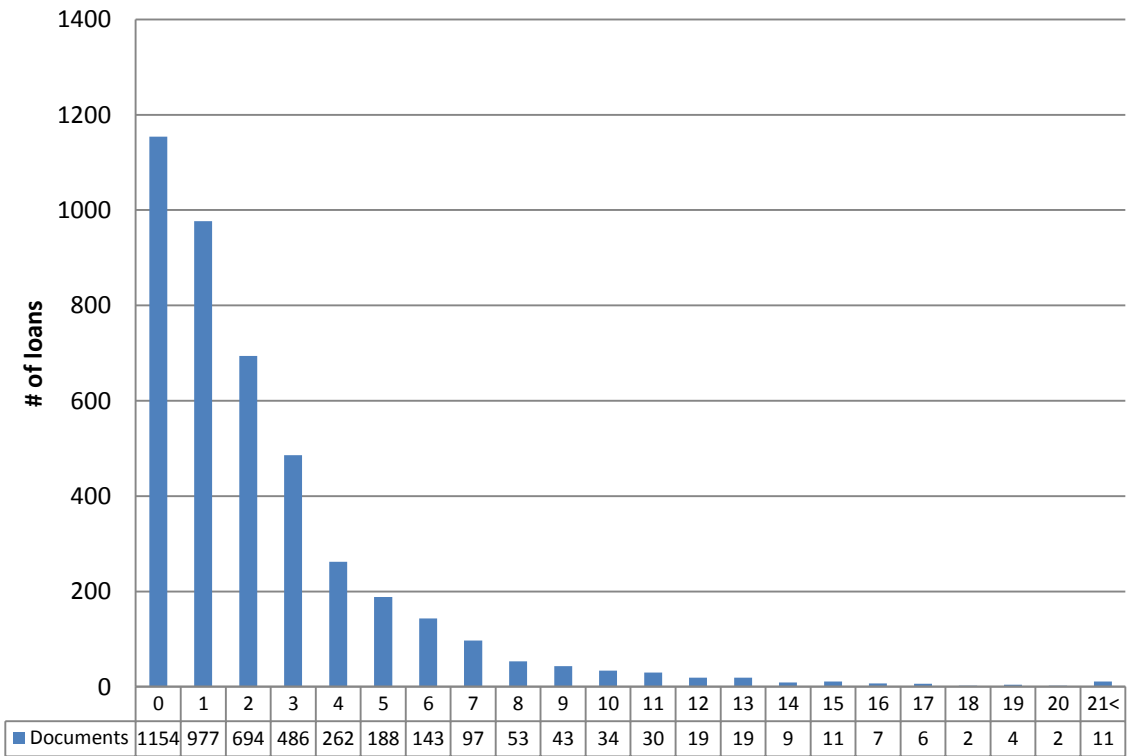
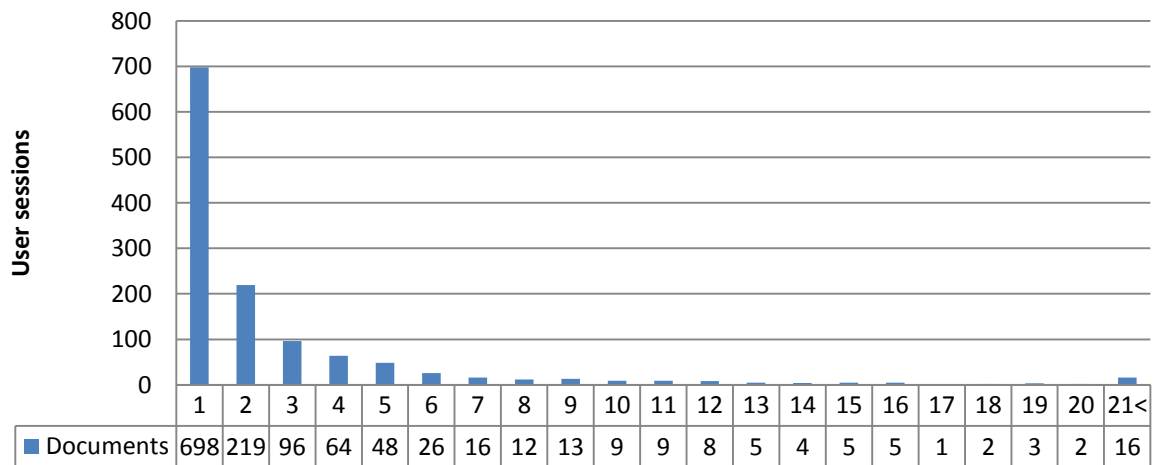


Fig. 2. Ebrary usage 2009-2012. Titles published 2008. User session frequencies.

(N=1262. m= 2,95. Max=99)



Compared to figure 1 above, the most obvious difference is that e-book 0-usage is not measured. In evaluating PDA non-usage is not an issue. It is in evaluating print circulation. If one user session e-books are treated as 0 loans, the frequency distribution look very much alike, except that e-books have a longer tail. In fact the proportion of titles with more than 20 user sessions in the sample is over five times the number of books with a similar level of circulation (1,4 vs 0,26 %). The books in question are in variably titles with chapters on course syllabi. An e-book is a far more flexible document than a printed book when it comes to accessing single chapters, for reading, printing or downloading purposes.

Then: How should e-book usage be measured? E-books are universally accessed through a catalogue, they are on “closed stacks” and must be “ordered”. Once ordered they can be browsed as if they were on an “open stack”. The user session combines borrowing and browsing. In my experience, working with the collection in a distant magazine from 2006 until 2008, when the Social Science Library was closed for rebuilding, 25 % of the books ordered by patrons were not collected or returned immediately on inspection.

The total number of user sessions for the 1262 titles in the 2008 Ebrary sample was 3728. The total number of loans for the 4152 in the 2008 print sample was 8747. The means are 2,95 and 2,17 respectively. A loan is a loan. But how do we separate e-book shelf browsing from a “proper” loan?

One simple way is to assume that mean loan circulation of a book is that of a printed book. In that case the 3728 Ebrary user sessions produced 2739 “loans”.

If the book lending mean is deflated to include books that have actually circulated, 2998 out of the 4152, giving a mean of 3,25 for print book circulation. But on that basis the number of ebrary loans becomes higher than actual number of user sessions, 4101.

Table 7. Ebrary 2009-2012. Titles published 2008. All uses.

User sessions	Page Views	Page/ Copy/Print	m/ PCP	Downloads	Documents
3728	60391	10354	18,8	195	1257

The bookkeeping problem is of course the 1-user sessions. But on average the 698 one-user session in the sample produced 11,88 pages viewed, copied or printed, as well as 31 downloads². The average result of a user session was 18,8

² To illustrate usage patterns: The two one use, zero pageview sessions in the sample produced 5 downloads. There appears to be no obvious intuitive connection between viewing, printing and downloading in user behaviour

pages viewed, copied or printed for a grand total of 70745 pages. If ten pages is considered a legitimate threshold for using, as opposed to browsing, a book, the number of books above the threshold is 678, giving a “document deflator” of 53,89 %. The number of user sessions is 3025, or 81,14 % of the total, the “use deflator”.

Defining a “loan” as a user session producing 10 pages or more read, copied or printed, gives a total of 3 025 “loans” for the sample. With the 236 downloads added, this produces a grand total of 3261 Ebray e-book “loans” from the sample. The same sample for UBL printed books produced 9747 loans. Ebrary circulation thus corresponded to 33,46 % of print lending in the period.

In terms of documents used, the deflation of Ebray usage reduced the number of documents circulated from 1262 to 680, 22,93 % of the 2964 UBL 2008 acquisitions borrowed by its own users. This means that in terms of documents circulated, Ebray usage was 1,5 times higher than print circulation in the sample.

First conclusion: A success in terms of document delivery.

UBL Ebray usage for the entire period looks like this:

Tab. 8: Ebray 2009-2014. Aggregate annual “circulation”.							
Year	Pages viewed	Pages copied	Pages printed	Unique titles	User sessions	Chapter/Range downloads	Full title downloads
2009	196601	2276	37144	6735	11660		
2010	187902	3173	48406	6479	11707		
2011	241317	3462	32896	6801	14298	249	92
2012	366250	3993	35389	11700	23004	3215	389
2013	382059	2988	30507	12262	25037	5232	509
2014	713187	6284	151423	25614	48330	7573	1101
	2087316	22176	335765	69591	134036	16269	2091

Using the use deflator from the sample, the 134036 user sessions can be regarded as having produced 108757 “loans”. The actual valid user sessions “loan count” of ten pages or more viewed, copied or printed is 114112. The underestimate is still within two standard deviations (4,2 %). Actual use amounts to 132472 over five years. Total number of books used was 18889.

Aggregate circulation at the four “Ebrary comparable” libraries at the UBL in the same period was 492129. In 2014, however, gross circulation at the Humanities, Psychology, Social Science and Science libraries amounted to 80731 documents. Total Ebrary use in the same year translated into 47412; 38738 loan equivalents and 8674 downloads.

The specific PDA numbers for that year are equally interesting.

Tab. 9. UBL. 2014. Ebrary. PDA as a proportion of total use

	Titles	User sessions	Total Pages V/C/P use	Downloads
Ebrary	16697	48330	870894	8674
PDA	4927	33413	775317	7360
PDA %	29,5	69,13	89,03	84,85

While PDA books only accounted for three out of ten Ebrary books used, they accounted for between 70 and 90 % of the use. In particular the “less advanced” use – reading – is something patrons do when the open PDA-books³. Of particular importance is that the PDA collections contain more recent publications than Academic Complete. Of 4895 recent, 2012 and 2013, used from Ebrary in 2014, 2288 were PDA books. These numbers compare favourably with the 9006 “Ebrary-comparable” books bought by the UBL in the two years.

In general it seems that the e-book is a format that really simplifies document delivery. And a PDA model is the only rational way of organizing e-book acquisitions in any library that actually believes that it exists for the sake of the patrons, not for the sake of the acquisition librarians.

Second conclusion: The wisdom of crowds (ie publishers and patrons).

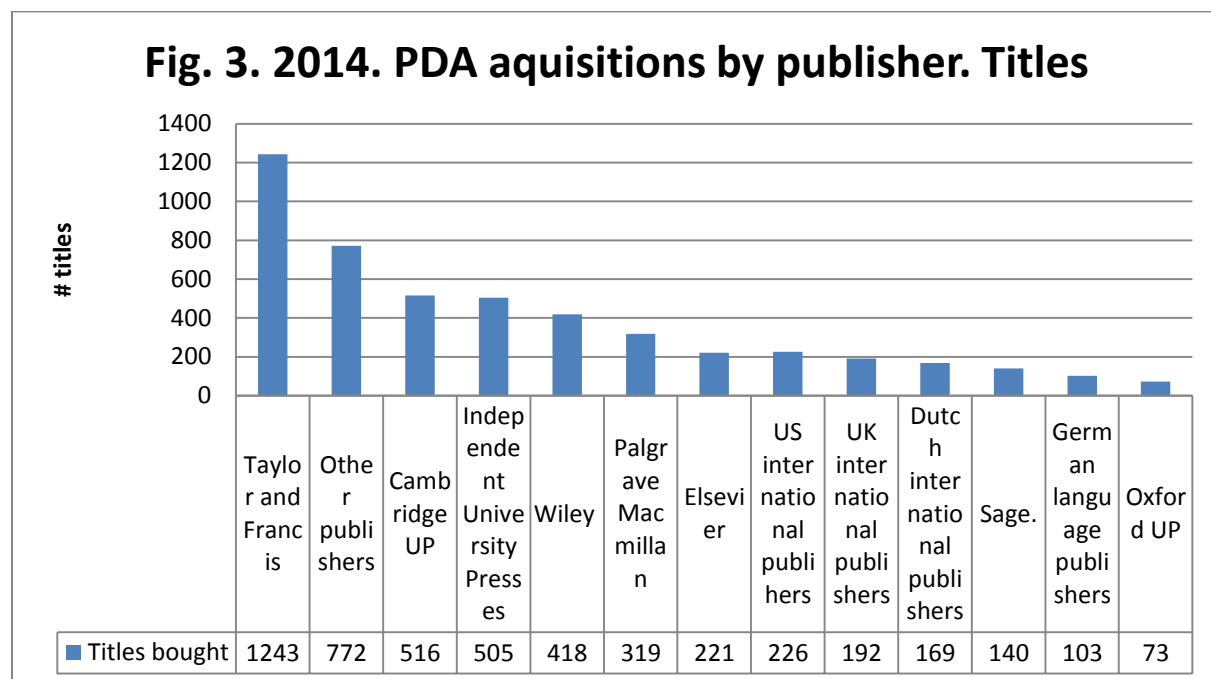
In traditional acquisition models a specific decision is made for each monograph (or collection) that is acquired. In a PDA model acquisition decisions take the form of explicit parameters of, the PDA-profiles. As all acquisition decisions eventually boil down to whether or not somebody else than the acquisition librarian would like to read the book, the decision to buy will always be based on an implicit “profile” of the patrons, their preferences, and the universe of available books. In this profiling of the knowledge universe, publishers play a major role, because they create cosmos out of the knowledge chaos. They not only organize by specializing in certain subjects. They guarantee the quality of the books they sell.

The core of the PDA-profile is the choice of publisher, a choice that allows the librarian to live out all her prejudices. As the collections available for profiling have been vetted beforehand by the bookseller, creating a knowledge based profile through choosing the most suitable publishers presents no problem. In so

³ 713 187 pages viewed corresponds to 71318 articles of 3500 words, corresponding to 7,5 % of UBL aggregate periodicals use. When counting we assume that the act of downloading and/or copying actually is an act of reading.

far problems arise it is because some publishers are reluctant to embark on PDA-arrangements. As the figure below reveals, availability appears to be the most important criterion for patrons, making “being there” the most important sales trick for any publisher.

Taylor and Francis tops the bestseller list of the UBL patrons and Oxford University Press is placed last because OUP prefer to sell its e-books through other channels than PDA arrangements. Not because T&F books have a consistently higher quality. In so far as the major publishers are concerned their PDA rank-order is different from their print rank-order. Where the PDA differs from the print acquisition is in the number of titles bought from “minor” publisher, not unexpectedly American. Forty percent of the acquisitions are made from other publishers than the big six⁴.

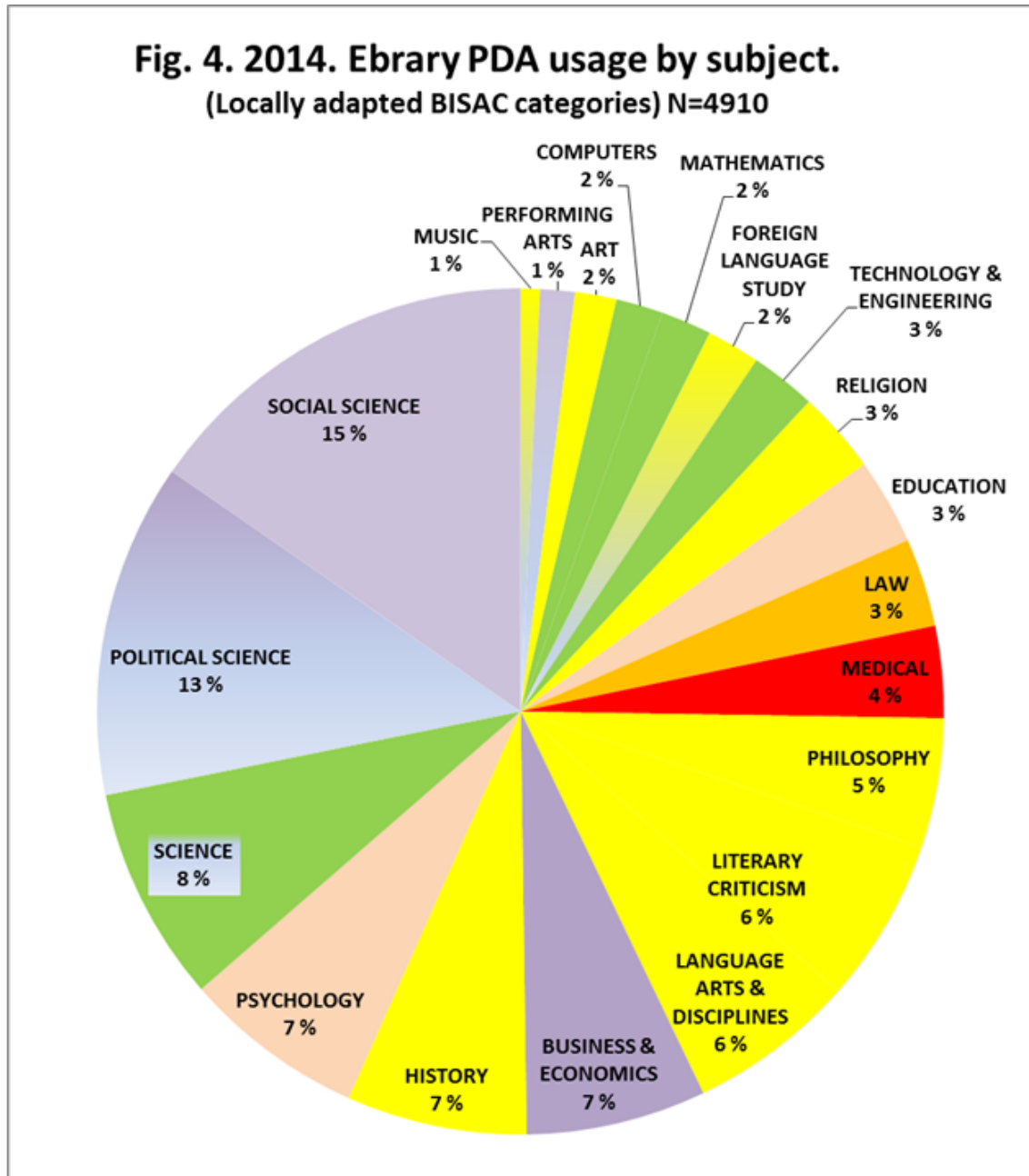


The impression is that this is a higher proportion than in the print acquisitions. It means that patrons are able to access a broader spectrum of literature, a larger part of the book universe, than is the case. Depending on the supply of e-books, a library adopting a PDA model will also be in a position to deliver documents in languages only the patrons can read. I would personally favour a development in this direction. But for the time being Ebrary offers Mexican and French Canadian literature, as well as a wide selection of German.

⁴ UBL subscribes to the SpringerLink e-book collection. The publisher is not included in the PDA-profile.

As noted above. The use patrons make of the PDA collections is the proof of their quality. Use validates the collection development decisions made when creating the profiles.

Another avenue of approach is analyzing use by subject to establish whether or not the PDA collection “behaves” as the print collection does. This is what the breakdown of the 2014 PDA usage into BISAC subjects informs us of:



If the the BISAC classes are combined (permutated, actually) into the six neat pigeonholes of the UBL acquisition budget, the overall distribution actually conforms to pattern of circulation at the UBL:

Table 10. UBL. Book circulation, by budgetary unit. 2014

	PDA	Print
Humanities	32,02	40,54
Law	3,38	8,11
Medicine et al.	3,5	7,26
Psychology et al.	10,24	13,33
Science	14,48	8,17
Social science et al.	36,37	22,58

As noted, medicine and law are underrepresented in the UBL Ebrary PDA. And the distinction between media studies, politics and anthropology of religion on the one hand and digital humanities, contemporary history and history of religion is extremely blurred, making the difference between the humanities and the social science library in their relative shares of PDA and print circulation is mainly a result of the BISAC classification – even if the social sciences are more “Ebrary-compatible” than the humanities. There is no Norwegian language literature in Ebrary. More interesting is the observation that e-books are 50 % more popular in the natural sciences than printed books, if the circulation data of the Science library is believable. That the e-book activity associated with the Psychology library is less than their print circulation is a mystery, but may be a result of their very high level of print circulation.

But overall pattern of Ebrary PDA activity does not differ significantly or inexplicably from that of the print circulation, leading to the conclusion that the wisdom of library patrons apply to their use of e-books as well as their use of printed books.

By way of a conclusion.

This is a very preliminary work. The main argument, that e-book usage already has outstripped the use of printed books, is definitely jury-rigged. It lacks solid foundations in hard data, even if the “deflator” developed to make use and circulation data comparable is less inaccurate than one would fear.

But the main argument, that PDA is the only tenable acquisition model for e-books, is not affected by the problems of exact comparisons between the two formats. A PDA-model allows a library to discharge its responsibilities towards its patrons in a far more efficient way than any traditional circulation model. PDA is to the delivery of monographs what the digitized periodical was to articles. Fair and efficient document delivery mandates the same “just in case, just in time” approach to monographs as to periodicals.

What is certain is that the quality of the patron-acquired monographs is identical to that of the library-acquired. PDA as automated collection development does not pose a threat to the quality of the collection if PDA-profiles are created with care by librarians that know their subjects.

Document delivery and collection development are equally core routine library activities. To the extent they can be automated, they ought to be automated; to the benefit of the librarians no less than the library patrons.

Which is why PDA is a very good idea.