EconBiz — Meeting User Needs with New Technology

Tamara Pianos

ZBW, German National Library of Economics, Leibniz Information Centre for Economics, Düsternbrooker Weg 120, 24105 Kiel, Germany, t.pianos@zbw.eu

Abstract

Virtual libraries try to combine traditional library services with new document types and services.¹ The first generation of virtual libraries mostly tried to offer services based on a library-centric view of information retrieval systems.² New virtual libraries try to concentrate on user’s needs, but this is often easier said than done. Restrictions like copyright laws, technical limitations and the like often make it difficult to meet user requirements. A number of studies documented these needs: easy-to-use, comprehensive yet focussed search, and easy access to print and online documents, subject specific, yet not too restricted to specific areas.

The new EconBiz-portal, relaunched in August 2010, has a disciplinary focus on business and economics and related subjects. It includes about 6 million records from different databases. Based on search-engine technology Lucene/Solr, combined with a metadata framework developed by the ZBW, it allows fast, convenient and complex searches. The integration of the Standard-Thesaurus-for Economics supports researchers by suggesting key words and related terms. Information on the availability of the documents is also included. Documents can either be accessed online or ways are shown to material that is available in print only. Journals Online & Print, a service developed by the German Electronic Journals Library (EZB) and the German Union Catalogue of Serials (ZDB) is included to provide easy access to all forms of journals.³ In addition, services like an event calendar, a tutorial on how to find information and an online-reference desk help to cater to the user’s complex needs.

The new EconBiz-portal was developed by the ZBW in close cooperation with the USB Cologne. Major parts of the search engine framework were developed by a company specialized in information technology.
This paper elaborates on the extraction of users’ requirements from different studies, the deduction of functional requirements, and, finally, the implementation of the portal with all its ups and downs.

**Key Words:** user requirements; user needs; subject portal; virtual library

## The User — the Unknown/Well-known/Erratic Being

What does the user want? And who is the user? Or rather who are our users? Virtual libraries address a number of different users: from digital native students who think they know how to use the internet, but often show a remarkable lack of information competency, to middle-aged researchers who started research when card catalogues and librarians where your best way to find what you wanted. One expects the latter group to be familiar with terms like ‘key word’ or ‘thesaurus’ etc., and to know about Boolean searches. According to a ZBW study, there are also quite a number of ‘older’ researchers who display a lack of sophisticated search skills, and therefore the groups cannot easily be divided into those who know how to do an expert search and those who do not. In addition, according to a new JISC study the Google generation is a myth, too. Impatience, or rather the now-or-never-attitude towards getting desired documents immediately, is common among all age groups now. Many (younger) researchers are members of community networks like Researchgate. If libraries do not supply them with what they need, they know colleagues around the world who will.

A number of studies tried to find out what researchers want and how they search. The studies range from local small-scale questionnaires of individual libraries or specific subject fields to national, international and interdisciplinary studies of the behaviour and preferences of students and researchers (see Figure 1). Examples at an international level are the OCLC study (OCLC, 2005) (interdisciplinary) and NEEO (NEEO, 2008) (subject-specific: economics). There have also been a couple of studies at a national interdisciplinary level. In Germany, these were ADL (BMBF, 2002), SteFi (Klatt, 2001), SSG (te Boekhorst, 2003a and 2003b) and a large-scale project at the Humboldt University (HU) (Havemann, 2006); in the UK it was RSLG (RSLG, 2003). The ZBW conducted a subject-specific study for ZBW services. This study will not be published completely, but the major findings will be presented here.
There seem to be a number of obvious trends in user habits and expectations (get it easy, fast, for free). However, there are also a number of differences in search preferences. Sometimes the age of the researcher seems to make a difference, sometimes the subject field. Older scholars in social sciences prefer bibliographical data and printed journals or librarians’ recommendations while older researchers in natural sciences prefer electronic journals, and (surprisingly) younger researchers in natural sciences prefer to browse shelves with printed books (Havemann and Kaufmann, 2006, p. 81).

**Getting to Know the Users**

*What they want and like*

- easy search tools
- fast searches, fast answers
• relevance ranking (no matter what the algorithm behind it is)
• comprehensive databases/collections that find everything that is important or relevant to them at the moment
• quality information
• books and journal articles are seen as the most important sources of quality information
• direct access to the desired documents
• lists of recommended articles on certain subjects
• context-sensitive support
• drill-downs/facetted search options
• folders to collect items
• serendipity.

What they (usually) don’t like

• result lists sorted by database (which usually happens in a meta-search/parallel search in different databases)
• complex, complicated gateways/portals
• having to check different portals or databases to find what they want
• long manuals
• give research data away
• share too much information in anonymous big groups
• logins.

How they usually search/work

• search Google or Google Scholar
• type one or two words and look at the first set of results
• search some of the major databases available at their universities
• search library catalogues/websites
• put things in folders (annotate, file away).

What they use for research

A recent study among researchers and students of business and economics shows (again) that Google is the number one entry point for research. When another study showed that many researchers start with Google, some German library colleagues asserted that the study was wrong and that this could not be true or even if it was true, serious scholars would not admit to
using Google for research. However, several studies in different fields seem to prove over and over again that many or even most scholars start their research with Google (Table 1).

Table 1: Entry points for searches.

<table>
<thead>
<tr>
<th>Entry Points</th>
<th>Researchers</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search engines</td>
<td>50%</td>
<td>54%</td>
</tr>
<tr>
<td>Google</td>
<td>31%</td>
<td>39%</td>
</tr>
<tr>
<td>Google Scholar</td>
<td>24%</td>
<td>18%</td>
</tr>
<tr>
<td>Yahoo</td>
<td>–</td>
<td>1%</td>
</tr>
<tr>
<td>(Bibliographical) databases</td>
<td>43%</td>
<td>35%</td>
</tr>
<tr>
<td>EBSCO</td>
<td>18%</td>
<td>10%</td>
</tr>
<tr>
<td>JSTOR</td>
<td>9%</td>
<td>3%</td>
</tr>
<tr>
<td>ScienceDirect</td>
<td>9%</td>
<td>5%</td>
</tr>
<tr>
<td>SSRN</td>
<td>8%</td>
<td>–</td>
</tr>
<tr>
<td>RePEc</td>
<td>3%</td>
<td>–</td>
</tr>
<tr>
<td>Library OPAC</td>
<td>21%</td>
<td>46%</td>
</tr>
<tr>
<td>Statistics</td>
<td>16%</td>
<td>20%</td>
</tr>
<tr>
<td>Specific portals/databases</td>
<td>8%</td>
<td>1%</td>
</tr>
<tr>
<td>(e.g. subject portals, OECD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publishers/publisher’s catalogues</td>
<td>8%</td>
<td>3%</td>
</tr>
<tr>
<td>(e.g. beck-online, Elsevier)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online journals</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>Papers (e.g. Handelsblatt, Spiegel)</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>3%</td>
<td>18%</td>
</tr>
<tr>
<td>Online journals</td>
<td>6%</td>
<td>1%</td>
</tr>
</tbody>
</table>

How often they use certain databases

Based on the frequency of use Google is also the number one entry point (Figure 2).

Users’ general attitudes

Students and researchers want and expect everything to be available online. Even though looking for information is an important part of their work and
even though they tend to use search engines a lot, many of them often do not think about how search engines work. They may not always be happy with the results they get, but they do not know about any alternative ways of searching or how to use search operators etc. (Figure 3).

Fig. 3: Attitudes towards research and search engines.

<table>
<thead>
<tr>
<th>Attitudes</th>
<th>Researchers</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Everything should be available online.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for information is an important part of my work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I do not think about how search engines work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If rewarding, I am ready to register.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I only use 2 to 3 databases/search engines.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Rights Management may be acceptable.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As good as it gets

Google and Google Scholar seem to fulfil many user needs better than other search portals or libraries etc., although they are not considered to be perfect either. Young students tend to be satisfied or even very satisfied with what search engines offer them (JISC, 2008, c. 7) (Figure 4). However, user needs for filter options, expert search options and subject-specific search options are not fulfilled by Google, even though Google by now offers a few facets. This is traditional library turf; libraries should use their advantages creatively and intelligently.

Fig. 4: Fulfilling user needs: for libraries it is ‘the library as physical place/building’ | library online services’.

<table>
<thead>
<tr>
<th>Fulfilling User Needs</th>
<th>Google</th>
<th>Google scholar</th>
<th>Libraries</th>
<th>Databases</th>
</tr>
</thead>
<tbody>
<tr>
<td>anywhere/anytime</td>
<td>++</td>
<td>++</td>
<td>-/+</td>
<td>++</td>
</tr>
<tr>
<td>Results</td>
<td>+</td>
<td>+</td>
<td>+/+</td>
<td>+</td>
</tr>
<tr>
<td>All-inclusive</td>
<td>+</td>
<td>+</td>
<td>++/+</td>
<td>++</td>
</tr>
<tr>
<td>Easy access</td>
<td>++</td>
<td>++</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>Easy search</td>
<td>++</td>
<td>++</td>
<td>+/+</td>
<td>-</td>
</tr>
<tr>
<td>Filter options</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>Intelligent search</td>
<td>++</td>
<td>++</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Expert options</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Subject specific</td>
<td>-</td>
<td>+</td>
<td>+/+</td>
<td>+</td>
</tr>
</tbody>
</table>

What users think about specific search engines and databases

What is most liked about Google is that it is easy to use and that it tolerates typing mistakes. These are obviously very important fields of competition. On the other hand, Google does not fare too well when it comes to relevant results or high-quality results. No other database mentioned delivered
relevant results only or high-quality results only. These results seem to suggest that we should go for easy search rather than quality results since Google seems to be far more successful with its easy search than other databases that focus on quality results (Figure 5).

**Fig. 5: Characteristics of search engines and databases.**

![Characteristics of Search Engines and Databases](image)

**More Difficulties for Information Providers**

There are a number of circumstances that make it difficult for libraries to communicate with their users and to display or offer everything they have got:

- Words mean different things in different contexts or are not intelligible to the average user at all, but if you want to offer more complex services it is difficult to work your way around words like: full-text, database, resources, browse, electronic journal etc.
- Licence situations are difficult to explain: why can you access something at the university and not at home?
- Copyright laws are getting more and more difficult: it is not easy to convey to users that libraries sometimes have to send photocopies
instead of online documents and that this is so because they have to abide by laws, not because they are technically incapable of sending electronic material.

- Moral dilemmas of librarians: they tend to take copyright laws, privacy protection etc. seriously, which makes it more difficult to 'just' send everything they have in their collection to anyone. They are often not too happy about cooperating with Google but often enough they cannot do without either.

- Including all the possible databases that might be relevant for the subject means that you also tend to include the same document twice or more often.

**Getting Started — Planning the New EconBiz**

We know what the users want and we know about difficulties and problems: now what? Of course, the ultimate aim would be to go for easy search as well as comprehensive content and high quality and to offer any support available. In other words we tried to make it simple, accessible, affordable (Chad, 2009, c. 64). To this end, we used the tools, devices or strategies listed in Table 2.

*Table 2: What users want and how to deal with it.*

<table>
<thead>
<tr>
<th>What they want</th>
<th>And how to deal with it</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Fast and easy search, result sets sorted by relevance</td>
<td>• Use search engine technology</td>
</tr>
<tr>
<td>• Easy to use</td>
<td>• Google-like entry page</td>
</tr>
<tr>
<td>• Finding exactly what they want</td>
<td>• Offer complex search options as well as an easy entry point</td>
</tr>
<tr>
<td></td>
<td>• Spend lots and lots of time homogenizing heterogeneous metadata</td>
</tr>
<tr>
<td></td>
<td>• Offer filter options</td>
</tr>
<tr>
<td>• Tolerate typing mistakes</td>
<td>• Use search engine technology (problems discussed below)</td>
</tr>
<tr>
<td>• Be able to access documents right away</td>
<td>• Integrate as much open access material as possible</td>
</tr>
<tr>
<td></td>
<td>• Show different availability options, e.g., by including new national services like Journals Online &amp; Print (JOP)</td>
</tr>
</tbody>
</table>
Based on the knowledge of several user studies, a couple of colleagues got together and wrote down the functional requirements for the new EconBiz. It was clear that we needed search engine technology and a Content Management System for different portal functions. We decided to use Lucene/Solr and Typo 3 because they are both open source and they have substantial developer communities. Furthermore, there are over forty different subject portals in Germany that all contribute to vascoda and a number of these portals also uses Lucene and/or Typo 3; thus we could count on community support.\textsuperscript{7}

Content

Users want a comprehensive search at one go; they do not like having to check several different portals and databases. That is one reason why they like Google so much: it seems to cover everything, and it actually covers a great percentage of relevant content. Fortunately, the databases of the two EconBiz partners already cover a great deal of relevant information. However, we tried (and keep trying) to include even more relevant content. We include databases that have a focus on business and economics and related social sciences. So far, we have included the following databases and their metadata:

- ECONIS (4,000,000)
- USB Catalogue (900,000)
- RePec (900,000)
- Online Contents Business and Economics (2,700,000)
- Emerald Databases (60,000)
- SSRN (soon to be included) (270,000)
- Selected Internet Resources (29,000)

The plan is to include more and more databases relevant for business and economics as we go along.

Developing the new EconBiz

When the functional requirements were more or less fixed, it was obvious that we did not yet have enough IT developers in the ZBW to develop the complete portal on our own within a reasonable time span. So we started
looking for a company that was able to implement some of the major functions for us. We found a German company called iSearch that is specialized in search engines and information retrieval, so we asked them to develop some of the major components for us. Shortly afterwards, in the summer of 2009, we were able to secure the help of a new IT colleague who worked in another library for several years and knows all about the problems of trying to feed heterogeneous sources into one portal. He is supported by two metadata specialists who do the mapping of the different formats and keep pointing out metadata problems, e.g., why browsings or drill-downs would not work as planned because the functions are not supported by the metadata.

Together, these three people are responsible for feeding the metadata framework. The incoming data is converted to an intermediate format which is roughly based on the metadata standards Dublin Core, OpenURL and MODS. In a second step, XML files that are suitable for feeding the Lucene/Solr index are generated from this internal representation.

The metadata framework is important for a couple of vital functions. Besides getting the data into the index and displaying it, functions like availability and exports (for reference management) is/will be realized through the framework.

Information on the availability of documents will be provided in a number of ways. Wherever possible, direct links to open access material are to be included. Journal articles are to be found through Journals Online & Print (JOP). JOP is a service provided by the German Electronic Journals Library (EZB) and the German Union Catalogue of Serials (ZDB). If you click on the URL provided by the service you can find out if a journal is available in print or as an electronic journal in your local library (via IP check). In addition, links to subito and the vascoda availability service are provided.

Easy entry

We tried to make the entry as easy as possible — although with every new aspect we talked about, we were tempted to add another button on the home page making it less and less easy again.
Spelling mistakes

Something users really like about Google and co. is that they tolerate spelling mistakes, so the new EconBiz was supposed to have some ‘Did you mean …?’ function as well. We soon realized that this was easier said than done. First we told our developing company to offer alternatives if there was a zero-result list. This worked reasonably well, but we found out at the same time that our databases also contained a number of fairly widely made spelling mistakes. If you typed, e.g. ‘managment’ instead of ‘management’ you would still find more than fifty sets of metadata containing this word. We therefore needed a function that would display those fifty sets and still ask ‘Did you mean: management?’ to show users that a few hundred thousand more sets could be found by using the right word. The developers adapted the software and it worked perfectly well for the manag(e)ment phenomena, but if you were looking for the buzz-word ‘viral marketing’, which is a correct term, you would now get suggestions like ‘Did you mean “virallinen market- ing”?’ because both words individually were present in the index more often that viral marketing but combined they did not make more sense than ‘viral marketing’. Something was changed again, and the service seems to be more helpful now. This is just one example of how difficult it can really be to make life easier for our users.

Relevance ranking

Librarians tend to say that it is better (and more honest) to sort result lists by date of publication rather than by relevance, because metadata usually does
not provide enough information for a good and convincing relevance ranking. But users usually say they want results ranked by relevance. What to do? First of all, we provide the possibility to choose between relevance ranking and date. Our ranking will mostly be based on numbers taken from a publication on the Heidelberg OPAC (Langenstein, 2009) with a few minor changes. Journal articles from peer-reviewed journals will also be prioritized hopefully in the near future. We will publish our criteria for relevance ranking on our website, so that at least our priorities will be traceable for everyone who is interested in how it works. In the future, we would like to offer options for users so that they can define their own sets of priorities for relevance ranking.

Usability tests

When a first (almost presentable) version of the portal was ready, we invited a few users to test the portal. They gave a number of useful comments as to which terms were misleading and which things they expected to find or could not find. They also made it clear that some things we thought were user-friendly were not user-friendly at all or at least open to misinterpretations:

- We thought the looking-glass symbol was useful for users, so we put it on the search button. But they expected to find additional functionality besides the search behind this symbol. Since the image on the search button confused users, we plan to get rid of it (Figure 7):

Fig. 7: Search buttons.

- The button for an expert search was in the upper navigation, but the users wanted it closer to the search button, so we plan to move that too.
- The users wanted an individual button for the result list, so we plan to insert one.
- In our project group we talked a long time about buttons to delete your search versus buttons to send your search and we thought we had found a good solution, but the users did not agree, at least not when the buttons were too close to the other fields in the expert search, so we had to move them a bit further down (Figure 8).
There are a number of smaller comments we got out of our first round of usability tests.

A couple of things obviously worked as we intended them to work after all. A happy user commented on the home page: ‘It is good that there are not too many links on this page that distract me.’ (Figure 9)
Other things we offer our users

From most studies we know that retrieving articles and books is what users mostly want from information portals. Apart from that, there is a need for more information literacy and customized help in different situations. A browsing option for internet resources is also offered by EconBiz.

Information literacy is covered by LOTSE, a cooperative online tutorial by a number of German libraries which offers specific help on how to search and access information, how to write your first paper etc. for specific subjects. If you do not know what a bibliography is LOTSE will explain what it is and will point out the most important bibliographies in your subject field. If you would like to get in touch with colleagues, LOTSE will point out ways to do this, like mailing lists or weblogs etc. and will name the most important sources in your subject field. EconDesk is a helpdesk provided by the ZBW. Users can ask questions on the phone, via email or chat (Krüger, 2008). A small widget (see Figure 10) is included next to the result set after a search. It can be used to start a chat directly. The chat is available Monday to Friday from 10 am to 5 pm. When the chat is closed, users can still use the email function to get support. EconDesk also provides users with information from databases and statistical material upon request.

Fig. 10: EconDesk chat widget.
Browsing Internet resources

EconBiz contains about 29,000 handpicked internet resources with a focus on business and economics. These resources can not only be searched, users are also able to browse through them (Figure 11).

Fig. 11: Browsing internet resources (English interface available soon).

Internet resources can be browsed in the following categories: all, events, portals, databases, institutions, researchers’ websites. Within these categories users can choose between, e.g., business, economics, industries, document type or country. A calendar of events is included to make it easier for users to find important conferences etc. in their particular subject fields.
Lessons Learned

It is quite easy to summarize the lessons learned: it always takes longer than you think. You have to know what and how to specify what. While I am writing this paper, I am still waiting for a couple of crucial features to be implemented. When we specified the requirements for the portal, we discussed and over-specified some features. Sometimes these were features that we had to cut out completely because some major external preconditions were missing. For example, we wanted to give users the opportunity to choose their library from a pull-down selection of options to be able to learn about the licence situations in their libraries. Later on, we realized that it must be completely confusing for users to choose a library and find themselves in different licence situations depending on whether they access the system from a library computer or from home. On the other hand, we hid major work packages in innocent looking phrases like ‘services a and b need to be included in the portal’.

Conclusion and Perspective

Trying to create a portal according to users’ needs was an interesting experience including lots of heated discussions on the actual needs of the users. After all, we are all users and tend to think that our individual preferences are universal. Despite the number of studies on user needs, there is always room for interpretation, always room for one more or one less button or functionality. We tried to make the best of it, but we also found out that a portal with several different databases and services is a very complex system. If you want to get rid of one functionality because it does not seem useful, you might unwillingly kill a completely different functionality. If you want to offer users a number of possibilities, you are likely to create more and more buttons and choices and the ‘make it easy’ paradigm is lost once again.

Once the portal is finished we will have to start (usability) testing all over again to keep improving the service. We are pretty sure we want to include more Web 2.0 functionalities in the portal. Optimizing the portal will never be truly finished, and, hopefully, we meet the user’s needs with most of our actions and changes.
References


Websites Referred to in the Text

EconBiz, http://www.econbiz.de/
EconDesk, http://www.zbw.eu/e_services/e_econdesk.htm
EZB, http://rzblx1.uni-regensburg.de/ezeit/
JOP, Journals Online & Print, http://www.zeitschriftendatenbank.de/services/journals-online-print.html
LOTSE, http://lotse.uni-muenster.de/
USB Cologne, http://www.ub.uni-koeln.de/
Vascoda, http://www.vascoda.de/
ZBW, http://www.zbw.eu/
ZDB, http://www.zeitschriftendatenbank.de/

Notes

1 For more information on virtual libraries in Germany see Depping (2007).
2 See Pasternack (2006) for the librarians’ approach to offering information through subject libraries.
3 At the present date (June 15th 2010), the realization of two features (thesaurus, availability) is still pending. They ought to be realized by the end of June 2010.

4 When looking at search behaviour with web search engines, we find that the average query length is 1.7 words for German language queries. For the English language, queries are a bit longer, but this has to do with the specifics of the German language, where there is heavy use of compound words. For example, ‘search engine user behaviour’ in English is four words, while the German Suchmaschinenutzerverhalten is just one word. Therefore, approximately 50% of German queries consist of just one word, while with English queries, the percentage is a bit lower (Lewandowski 2008, p. 262).


6 There have been a couple of tests with words users understand, do not understand or tend to misinterpret. For English terms see: http://www.jkup.net/terms-studies.html.

7 A list of subject portals in Germany can be found at www.vascoda.de. Vascoda started out as an interdisciplinary portal giving access to the accumulated data of the subject portals. In the near future it will be changed into a community platform for the different subject portals.