EconBiz-Experiences: Creating our own discovery system for business & economics

Tamara Pianos

Information Provision and Access, ZBW - German National Library of Economics Leibniz Information Centre for Economics, Kiel, Germany.
t.pianos@zbw.eu

Arne Martin Klemenz

Information Provision and Access, ZBW - German National Library of Economics Leibniz Information Centre for Economics, Kiel, Germany.
a.klemenz@zbw.eu

Abstract:

EconBiz (econbiz.de) is a subject portal for business and economic developed by the German National Library of Economics (ZBW). It advanced from a subject guide to a discovery system over the past decade. The portal provides journal articles, book chapters, working papers, monographs and conferences. A thesaurus ~~also~~ created and maintained by ZBW and its partners enhances the search experience of the portal. EconBiz is a free portal that can be used by anyone in the world. Whereas discovery systems usually focus on one university campus, EconBiz tries to provide access to publications for everyone. However, the portal functions as a discovery system for local users of ZBW. Thus improving the availability and access to documents on a local and national level as well as on an international level is an ongoing project in the development. Showing the best possible options of access is based on a complex negotiation between reference librarians, user feedback as well as metadata- and IT specialists.

This paper will also highlight the fruitful collaboration of reference librarians, metadata experts and IT specialists. The EconBiz team grew over the past few years, and in 2012 the first PhD-student in Computer Science joined the EconBiz team in order to combine new research and development with the development of the portal.

The improvement of information literacy is another aspect tightly interwoven with the portal. A Guided Walk is supposed to help students with the most pressing questions like "how to search and evaluate," "how to access the full-text" and "how to cite." This paper covers experiences related to the latest relaunch of the portal in March 2017. Some future plans will also be discussed.

Keywords: Discovery System, Subject Information, Business, Economics, Information Technology

Introduction

EconBiz is a free portal covering publications and conferences in business and economics. It is provided by the German National Library of Economics – Leibniz Information Center for Economics (ZBW). The portal covers more than 10 million publications from different databases – journal articles, books and working papers, online and print, and open access as well as licensed material. More than 4 million of those publications come from ECONIS, a database produced by ZBW. ZBW has collected and indexed publications in many different formats and languages since 1919.
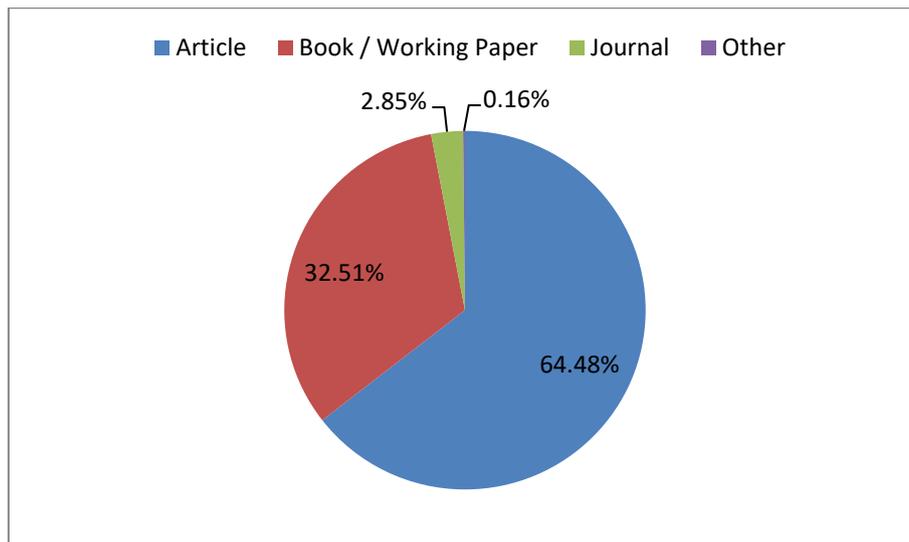
Figure 1: Publication types covered by EconBiz

The portal is available in English and German and the main functionalities are also available in French and Spanish.
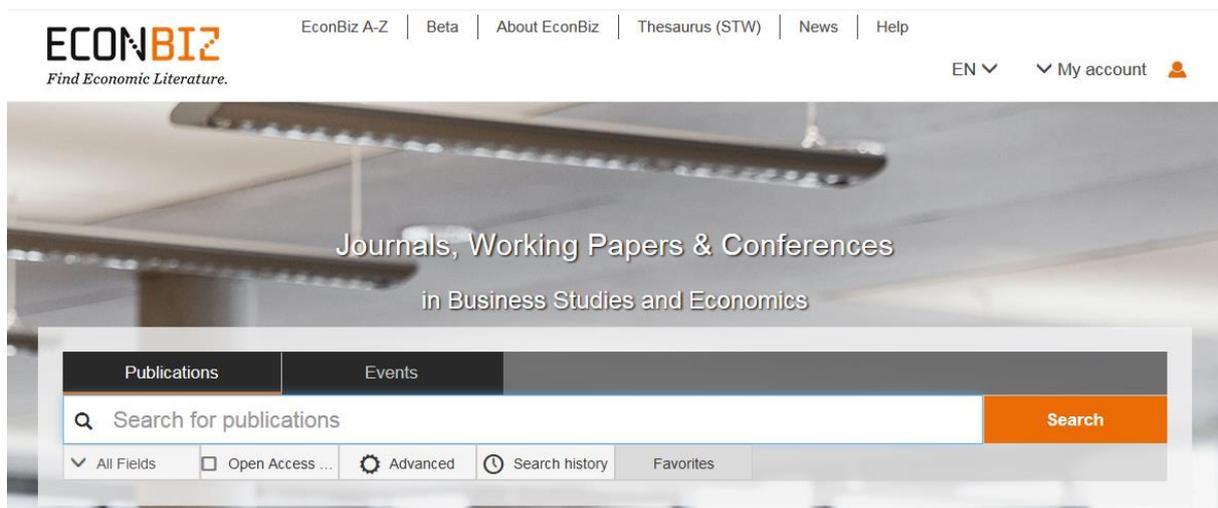
Illustration 1: Screenshot of EconBiz

EconBiz is supposed to improve the discovery of quality information related to business studies and economics. An international network of EconBiz partners provides feedback on the needs of a global user community and on usage related issues. The knowledge and developments of a number of projects (for example EC funded project EEXCESS or German

Research Foundation funded LibRank) were used to improve the EconBiz features or deepen the understanding of search and display related aspects. (Plassmeier et al. 2015)

This paper covers experiences in building the portal EconBiz and the discovery index beneath it in a collaboration of reference librarians, metadata- experts and IT and other specialists.

All-in-one device suitable for Every Purpose - EconBiz as a Discovery System

In German we have the term "eierlegende Wollmilchsau," which is literally translated as a pig that lays eggs and provides milk and wool as well.

Discovery Systems are supposed to be as easy to handle as the single search bar in Google while catering to all the special needs that researchers or librarians or vendors could think of at the same time. Ideally, researchers and students can use the system to type their query without giving much thought to the correct terminology and the system will find all relevant publications that can be easily accessed through the local library holdings or local licenses.

Like many commercial discovery systems, EconBiz offers a more or less simple search interface to access a huge discovery index. The index is not quite as big as some web-scale discovery indexes but is still large. Other than most discovery indexes it is focused on business and economics and related subjects content. Commercial web-scale discovery systems offer specific campus installations providing access to local holdings and licensed articles, whereas EconBiz has a search interface and a metadata-index but the access to local holdings functionality in the strict sense is only realized for ZBW-users. On the other hand EconBiz tries to provide access to local campus licenses wherever possible.

"A Short History of Nearly Everything" – From the late 1990s to the Present

The quote is taken from a Bill Bryson book. However, the history presented here is much shorter.

In the late 1990s two major German funding bodies, the German Research Foundation (DFG) and the Federal Ministry of Education (BMBF) started funding virtual libraries with a subject specific focus to make it easier to access all relevant quality subject information independent of their publication format. When Google Scholar was introduced in 2004, many people thought that virtual libraries would soon be superfluous. However, Google Scholar obviously does not provide the answer to every research question and is not able to provide access to all relevant publications. For example, it has been suspected that the Google algorithm favours older publications because they have more time to accumulate citations. (Martín-Martín, A et. al., 2016) In addition, all academic research in the hands of one single company with commercial interests and a non-transparent agenda behind its relevancy ranking has been criticised by some (for example Cecchino, 2010).

Against this background EconBiz started out as the "Virtuelle Fachbibliothek Wirtschaftswissenschaften /Virtual Library for Business and Economics" and went live in 2002. At the time, the portal mostly offered a number of links to different services. A few years later a metasearch covering different content collections was introduced. Not surprisingly, users – in the meantime socialized by Google - neither liked nor understood that results were sorted by databases instead of relevance (Pianos, 2010). Ultimately, in 2010 the

new EconBiz, then realized with search engine technology, went live. In 2017 a new and improved version that also included lending functionalities for ZBW users was released.

The emphasis of EconBiz developments is on:
- Search quality and services
- Providing access to publications
- Using controlled vocabulary to enhance services
- Provide more quality indicators in the future

The portal is created by ZBW using open source tools such as VuFind and Solr. Creating rather than buying the portal ensures that the content can be used in different ways and that information on content or relevancy ranking is not subject to commercial interests.

The implementation of the new EconBiz

For the 2010 relaunch we had a document with functional requirements that were to be realized by an external IT company since ZBW did not have enough IT resources at that time. From 2011 onwards more computer scientists and programmers joined ZBW, so the relaunch of the new system in March 2017 was realized with ZBW staff only.

Below the surface – Metadata aggregation and the search index VuFind and Solr

Even though the portal functionality for the 2010 relaunch of EconBiz was realized externally, elemental parts of the search index were realized by ZBW colleagues. More and more services were built around the central metadata collection and since 2012 an API grants access to most of the aggregated EconBiz content.

ZBW metadata experts mapped the fields of different bibliographic formats from the sources to a single internal format to build the metadata basis for an aggregated search index. Depending on the data source the quality and quantity of the metadata differs immensely. Besides that, the database providers sometimes change vital parts of their data or metadata formats, so that for example IDs of datasets that are assumed to be unique may no longer be unique and useful. As a result, changed or deleted datasets cannot be updated or removed from the aggregated EconBiz search index without the implementation of further alignment strategies. Therefore, these interfaces can either no longer be used or require manual effort to ensure subsequent updates which leads to a variation in metadata update frequencies for each source included in the index.

The technological basis for the EconBiz search index is Apache Solr. This search platform provides all necessary features such as faceted search, full-text search and transparent ranking algorithms to fulfill all domain and subject specific user needs.

VuFind is used as the technological basis for the search portal itself since 2013. VuFind is an open source software with a huge library user community which guarantees continuous improvements as well as frequent future releases that allow EconBiz to stay up to date with state of the art concepts in library search portals.

In the Google era, when many spelling mistakes are tolerated, it is hard to convey why metadata-quality should be an important topic. However, retrieval does not work well or at all without excellent quality metadata (Akeroyd, 2017). Also, users still long to find qualitative

and authoritative data and libraries are still the places where quality information can be found (Christensen, 2013).

Authority Data, Controlled Vocabulary and STW – a Thesaurus for Economics and Business

About one third of the content included in EconBiz uses controlled vocabularies and authority data to disambiguate for example authors and corporate bodies. The Thesaurus for Economics (STW) is a controlled vocabulary that is used for subject indexing of major parts of ZBW's collection. Based on the STW as well as further thesauri and vocabularies that are connected through mappings and alignments, the metadata in the aggregated EconBiz search index can be enriched. As a result, publications using synonyms users did not explicitly search for can also be found.

In addition to this, Authority Data from the Integrated Authority File (GND) is used for the disambiguation of authors and corporate bodies to allow for example the distinction of authors with same names. More detailed information about the authors of a publication based on the potential of linked authority data will be integrated in the portal in the near future.

What the users want – input for new developments

A number of publications and user studies as well as direct user feedback are the basis for the development and improvement of EconBiz features. In addition, an advisory board consisting of scholars and librarians provides useful insights. Colleagues from the international EconBiz partner network also make suggestions based on user feedback in their institutions. Workshops with scholars who do research in business and economics are conducted on a regular basis to learn more about their needs and changing research and publishing habits. Ideas from these workshops are taken up by ZBW researchers in order to improve the portal EconBiz or other ZBW services like the repository EconStor or the journals Wirtschaftsdienst and Intereconomics.

Using Research and Researchers to Improve EconBiz Services

ZBW has three professors in Computer & Information Sciences. They and their PostDocs, PhD- or MA and BA- students frequently use EconBiz content to do experiments. The EconBiz product management also suggests projects that may improve individual features. The portal has a beta section were tools can be shown in an early development stage.

ZBW also frequently participates in national and international projects funded for example by the European commission. Since the portal is our own, we can easily use at least a major part of the content for these projects without many restrictions. ZBW has a division devoted to interacting with scholars in economics and business studies.

OPAC Functionalities

Some of the challenges are that EconBiz serves local, national and international audiences which have to be kept in mind especially when it comes to access to licensed or printed publications.

EconBiz has OPAC functionalities for local ZBW users and can be considered as the local discovery system, but is more like a subject database for all other users. This means, for example also that the functionality after login differs according to the user profile.

The major developments of the OPAC integration circled around the logins and lending information including status information on books that were e.g. ready to be picked up or on route from Hamburg to Kiel etc.

The Patrons Account Information API interface (PAIA) was used to implement the service. Users with a simple EconBiz account can only use their account to save favorites or searches whereas users with a ZBW-EconBiz account also find information on loans and reservations or can order volumes directly from the stacks.

Functionalities provided after login: EconBiz-Login

Functionalities provided after Login: ZBW-EconBiz-Login

Illustration 2: Different options for different user groups

Availability Options as a Crucial Service Component

On an abstract level most users want to find, identify, select and obtain publications relevant to their research as described in the FRBR user tasks. As service institutions "[l]ibraries have a strategic interest in the tools and technologies that facilitate the discovery of and access to the resources for the communities that they serve" (Breeding, NISO, p.1). Thus, giving users easy access to relevant publications is a key feature or at least goal of any discovery system. As EconBiz aims to offer the best possible access to a publication independently of the format of the publication, a variety of access options needs to be covered. Some users might have access to a printed version of a journal article through their library while other users have the right to directly access the electronic version through their institution. If these options are unavailable document delivery or interlibrary loan may still be a way to obtain access.
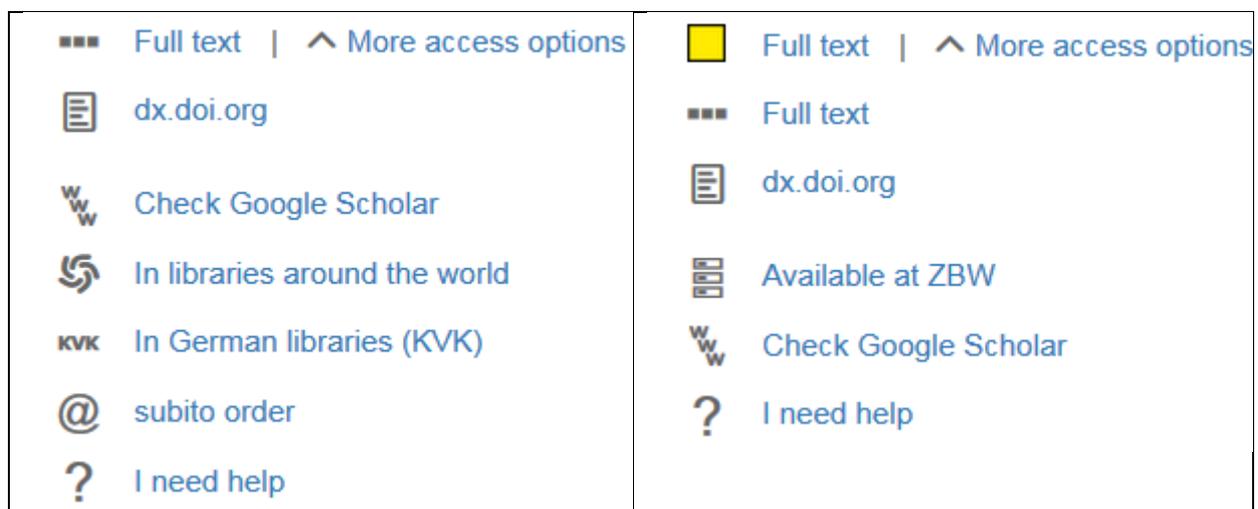
Illustration 3: Display of different availability options

Depending on the format and the location of the user, different options for gaining access to publications are available. Access not only depends on a license and the correct information on who is eligible to access a publication based on IP addresses or other options, but also on the correct metadata since link resolving does not work if ~~e.g.~~ page or volume numbers are not displayed correctly. In order to create the list of availability options several colleagues from different backgrounds have worked together over a longer period of time optimizing and changing the logic behind it.

Communication - Different Stakeholder in the development of the new EconBiz

In the late 1990s EconBiz started as a project with only a few people with designated tasks. As the portal grew and covered more and more aspects, so grew the number of people involved, as well as the variety of professions involved. Diverse topics like user needs, information literacy, metadata quality, Search Engine Optimization (SEO), reference, lending, tools and technology like Solr VuFind and relevancy ranking, metadata integration, controlled vocabulary, use of authority files, system stability and more need to be dealt with. Ever since the OPAC functionalities were to be added to the portal the reference/information services librarians also had a say in the implementation of features surrounding the former OPAC. The local, national and international views need to be reconciled and functions have to be implemented in a way that no user group is confused by one word or another.

Communication is an important part of implementing new services. A lot of mistakes can be made along the way and negotiating the best way to discuss and communicate the changes is an ongoing process. The complex world of discovery brings together experts and all-rounders and one word can have multiple meanings in different contexts. With a number of different stakeholders involved during the implementation phases it was always important to figure out who should be involved in which decision at which stage. Experts feel frustrated if decisions are made without consulting them first. Sometimes experts are involved in so many decisions though that they do not want to be the bottleneck either. In a complex system with many interconnections, it is not always clear either which modules/topics/experts may be affected by a decision, so it might come down to asking the whole team: "Does anyone see a problem here?" Meeting arrangements were adapted a number of times along the way in order to avoid boring people on the one hand while making sure that everyone has the information they need for their work packages on the other. The crucial question is often "Who needs how much and what kind of information for an informed decision?" Colleagues from diverse professional backgrounds have different approaches and attitudes, e.g. in their love for detail, adherence to (cataloguing) rules or tolerance for mistakes. The product management might say "80% is good enough for me" and may eventually revoke this because one particular feature does not work due to some minor missing detail. One lesson learned in this context is to negotiate as early as possible where fuzziness of certain aspects can be tolerated and where this might lead to major problems later on.

In 2012 the first PhD student with a background in computer science joined the EconBiz team. One aspect of his work is to discuss and translate options around new developments, IT infrastructure and research as well as the product management and the rest of the team. In 2014 another PhD student also with a computer science background joined the team as well. One of her tasks is to deliberate further developments regarding metric indicators and altmetrics between researchers and product management. Communication and translating

detailed questions of for example server configurations into implications for product features or aligning research questions and product features are part of their job descriptions.

In their presentation at the German Bibliothekartag (the major German librarian conference) two German colleagues suggested that introducing a discovery system also implies a lot of changes for many parts of the library from cataloguing to reference (Blenkle/Bodem 2014). Thus, creating a discovery system in a library often not only involves communication on technical aspects but also on why and how some aspects are changed. As the development of EconBiz from a website with references to discovery system was spread over several years and more and more people were involved in the actual planning through the years, the changes did not occur very suddenly. The developments and reasons behind the changes were communicated in an internal wiki, e-mails and internal training sessions. In these sessions, a number of questions were asked like, "Why is this feature missing/different now?" Following up on all of those questions is a lot of work (and not all questions are fully answered yet) but all colleagues who eventually have to work with the system will probably feel better if their concerns are not ignored. If a feature was important in a particular context and is no longer available there should be at least an answer why it is difficult or even impossible to maintain this particular feature - or why this feature may not be useful in a new context.

Plans for the Future

Apart from generally improving features and access options as an ongoing process there are also plans to support researchers' decisions on quality of publications.

Feedback from users, as well as the EconBiz advisory board, shows that quality indicators are an important topic, especially since the number of publications and journals increased immensely. Plans to implement some kind of "journal landscape" are currently discussed. The journal landscape is supposed to provide useful information on one journal in relation to related journals in the same subject area. Questions like "How long has this journal been around?," "Who uses and cites this journal?," "How popular is this journal?" and many more are supposed to be answered in this context.
One goal is to have a visualization of the information on the different journals that helps researchers to make decisions on the quality of journals. This feature will be developed in a close cooperation of ZBW researchers and the EconBiz team.

A Word on Information Literacy

The ultimate end to have a high quality system that automatically understands and enhances user queries and delivers all relevant results to a simple/single query has not yet been reached. Some might ask if this is the right goal after all. In the meantime, EconBiz integrates ways to convey some basic concepts of information literacy in a playful manner that deal with the most important aspects of a student's research life. A Guided Walk was implemented to take students along their paths of "searching and evaluating," "access to full-texts" and how to "cite correctly."

The Guided Walk is also used to explain why some things like e.g. access to licensed full-text is sometimes difficult and that this is not necessarily the libraries' fault.

Conclusion

User feedback for EconBiz from local users as well as international partners is mostly quite positive. There are a number of features that still need improvement and it is still quite a long way to achieve some kind of perfection. However, under the particular circumstances with a big institution backing the system and researchers in computer sciences doing projects to improve individual aspects of the portal or using data from the portal as a testbed, the decision to create rather than buy was the right one. The use of open source tools provides the freedom to adapt services and distribute content in many possible ways.

With many people inside and outside of ZBW involved in one way or another, the importance of communication will increase rather than decrease in the future, so the communication channels will be continually improved.

Acknowledgments

URLs to services mentioned in the text:

More information on EconBiz: https://www.econbiz.de/eb/en/about/abouten/
More information on the EconBiz partner network:
https://www.econbiz.de/eb/en/about/econbiz-partner-network/econbiz-partner-network/
EconBiz Beta: https://www.econbiz.de/eb/en/beta/about-beta/
EconStor: https://www.econstor.eu/
FRBR User Tasks: https://archive.ifla.org/VII/s13/frbr/frbr3.htm
Guided Walk – Information Literacy: https://www.econbiz.de/eb/en/gw/#c3296
Intereconomics: http://www.intereconomics.eu/
STW: http://zbw.eu/stw/version/latest/about.en.html
VuFind: https://vufind.org/vufind/
Wirtschaftsdienst: http://www.wirtschaftsdienst.eu/

References

Blenkle, M. and Bodem, C.: Mission Possible? Erfahrungen & Empfehlungen zur Einführung von Discovery-Systemen, Bibliothekartag 2014, (in German) https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/1498

Blenkle, M.; Ellis, R.; Haake, E.: Only the first results count: user-feedback-modified relevance ranking in E-LIB Bremen, Insights, 28(2), pp. 75–80 (2015) - http://doi.org/10.1629/uksg.235

Breeding, M.: The Future of Library Resource Discovery: A white paper commissioned by the NISO Discovery to Delivery (D2D) Topic Committee NISO Whitepaper: http://www.niso.org/apps/group_public/download.php/14487/future_library_resource_discovery.pdf

Cecchino, N. J.: Google Scholar : Reviewed. - In: Journal of the Medical Library Association 98 (2010) 4, S. 320 - 321. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1324783

Christensen, A.: Next generation catalogues: what do users think? In: Chambers, Sally (ed.): Catalogue 2.0: The Future of the Library Catalogue. London: Facet, 2013. http://www.facetpublishing.co.uk/downloads/file/Next-generation%20catalogues%20-%20what%20do%20users%20think.pdf

Martín-Martín, A., Orduna-Malea, E., Ayllón, J. M., & Delgado López-Cózar, E.. Back to the past: on the shoulders of an academic search engine giant. Scientometrics. http://doi.org/10.1007/s11192-016-1917-2

Pianos, T., EconBiz — Meeting User Needs with New Technology. LIBER Quarterly. 20(1), pp.4–24. (2010). DOI: http://doi.org/10.18352/lq.7972

Plassmeier, K., Borst, T., Behnert, C., & Lewandowski, D.: Evaluating Popularity Data for Relevance Ranking in Library Information Systems. In: Proceedings of the 78th ASIS&T Annual Meeting (Vol. 51) (2015). :

http://www.librank.info/wp-content/uploads/2016/03/Poster_final.5.2.2.pdf