

Providing Open Access to Digital Repositories for Social Sciences and Humanities Publications – Problems and Best Practices

Radoslav Pavlov, Detelin Luchev^[0000-0003-0926-5796]

Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences,
Sofia, Bulgaria
radko@cc.bas.bg, dml@math.bas.bg

Abstract. The report presents some problems and best practices for providing open access to digital repositories for scientific publications in social sciences and humanities (SSH). Editions in SSH have specifics that concerns publications in many national languages, deep archives, metadata, search, etc. As an example, the Central and Eastern European Online Library – CEEOL has been chosen, because CEEOL is a leading provider and a repository offering a highly specialized and comprehensive collection of fulltext publications in SSH. Over 75% of the editions indexed in CEEOL are open access. The study analyzes the main characteristics of CEEOL as: services, functionalities, opportunities for using its resources and the state of the Bulgarian collection.

Keywords: Open Access, Repository, Content Management System, Central and Eastern European Online Library, CEEOL, Social Sciences, Humanities, Bulgaria.

1 Introduction

Exploring the opportunities for creating and using repositories and content management systems in the social sciences and humanities (SSH) requires the study of the existing examples in this area. The “European databases and repositories for Social Sciences and Humanities research output” report shows that Bulgaria doesn’t have a national repository for SSH (Sile, L., Guns, R., Sivertsen, G. & Engels, T. 2017). Practically, however, almost all Bulgarian academic and scientific journals in SSH are stored in the repository of the Central and Eastern European Online Library (CEEOL, <https://www.ceeol.com>), Frankfurt, Germany, and have been distributed electronically by it around the world over the past two decades. The Bulgarian collection of SSH periodicals of the Bulgarian Academy of Sciences (BAS) and the main universities in Bulgaria in CEEOL is a good example of the quality, completeness and depth of the archives of these journals. The state-of-the-art, the ideas, the solutions for the SSH publications specifics, the characteristics of the Bulgarian collection and functionalities in

CEEOL could be referred as one of the best practices in the development of SSH repositories and content management systems (CMS) as well as for providing of SSH open access, because more than 75% of the content in CEEOL is open access.

The same investigation on national repositories for SSH (Sīle, L., Guns, R., Sivertsen, G. & Engels, T. 2017) presented “a rather great diversity in terms of the selection of research output to be included in a database”, some of them are collecting only open-access publications or only articles in journals, or include publications in only specific research themes; some of them are following strict procedure for inclusion, but others – not; etc. This shows that the databases use research output inclusion criteria attributable to different levels of aggregation: publication level, research output type level, academic disciplines level, institutions level. This is also a preposition for looking for a specific database and repository in SSH as a good example for a national database for Bulgarian open SSH publications.

Central and Eastern European Online Library (CEEOL) is a leading provider and a repository of academic e-journals and e-books in the humanities and social sciences from and about Central and Eastern Europe and there are good practices that can be implemented in the systems for SSH open access publications developed in Bulgaria. In the rapidly changing digital sphere CEEOL is a reliable source of adjusting expertise trusted by scholars, publishers and librarians. Currently, over 1100 publishers entrust CEEOL with their high-quality journals and e-books. CEEOL provides scholars, researchers and students with access to a wide range of academic content in a constantly growing, dynamic repository. Currently, CEEOL covers more than 2200 journals and 590.000 articles, over 2900 eBooks and 3000 Grey Literature documents.

One intrinsic specificity of SSH area is the publication in the national languages, but not in English. Compared to other repositories offering aggregated content in SSH, primarily or only published in English and mainly accessible for institutional subscribers only, CEEOL offers an immediate and direct worldwide visibility of the metadata of journal articles, eBooks and eBook chapters to every single individual, visiting the CEEOL repository, independently from institutional affiliation or library subscription.

Besides accessing www.ceeol.com, the dissemination process of content in CEEOL includes the discovery and indexing systems such as: Google Scholar, ProQuest’s Serial Solutions, Summon, Primo Central and Alma, Ebsco’s EDS and Knowledge Base, TDNet and OCLC. CEEOL works in partnership with Clarivate Analytics, the owner of Web of Science. Web of Science editors have access to the CEEOL platform, allowing for easier analysis, selection and indexing of CEEOL journals that interest the Web of Science community.

CEEOL’s cooperation with world-leading universities and university libraries in 25 countries makes the published Open Access journals’ content and Open Access eBooks visible and accessible potentially to more than 2,5 million students worldwide as well as to scholars, professors, researchers affiliated to those subscribing institutions. The services CEEOL provides to libraries include, besides the instant hosting for immediate access to the full-text content, also high-quality MARC records and Counter reports on usage. In 2018, CEEOL succeeded to achieve 1,3 million visits with 3,9 million page views and 1,9 million sessions.

Some of the main characteristics of CEEOL as: services, functionalities and opportunities for using its resources, are presented as ideas for the development of the repository and CMS for Bulgarian open access periodicals and books in SSH. Attention is focused on CEEOL state-of-the-art (part 2), Bulgarian collection of SSH periodicals in CEEOL (part 3), and system services and functionalities (part 4) as good practices and ideas for providing open access in digital repositories of social sciences and humanities publications.

2 CEEOL State-of-the-art

Since its inception in 1999, CEEOL has invested considerable effort to ensure that it provides a central starting point for access to the current and archive content of SSH journals from and for Central and Eastern Europe. In 2016, CEEOL developed an entirely new information system and expanded its activity in the field of non-periodical (ISBN) editions and Grey Literature (editions without ISSN or ISBN, Pavlovska, 2018) in the field of SSH and in their border areas with other scientific disciplines.

CEEOL provides institutional and individual users indexed metadata and full-text indexed issues of periodicals (as well as books and Grey literature from 2016) and, unlike ERIH PLUS (Lavik & Sivertsten, 2016), is not a formal register of journal titles in the field of SSH. CEEOL does not generate any impact factors, as it stands for the position of many other scientists and institutions (see the literature in recent research on the issue: Kulczycki, et al., 2018; Aksnes, & Sivertsen, 2019; Bayer, et al., 2019), that the specifics of SSH and their scientific publications do not suggest that they relate to both the idea and the algorithms for this type of reporting and giving rank of editions, scientists or institutions, which has been accepted in some countries to apply to natural sciences (studies of parallel cross-disciplinary comparisons about Google Scholar, Scopus and Web of Science also show this specificity (Mongeon, & Paul-Hus, 2016; Harzing, & Alakangas, 2016)).

In the end of 2018, CEEOL covered around 2,100 SSH journal titles (Fig. 1.) from 36 countries – including 1,580 open access journals – from more than 1100 international publishers with more than 590,000 indexed and archived articles, published in more than 50 languages. The CEEOL coverage is growing monthly by 5,000 newly included journal articles. Significant number of the included journals are represented with complete archival collection.

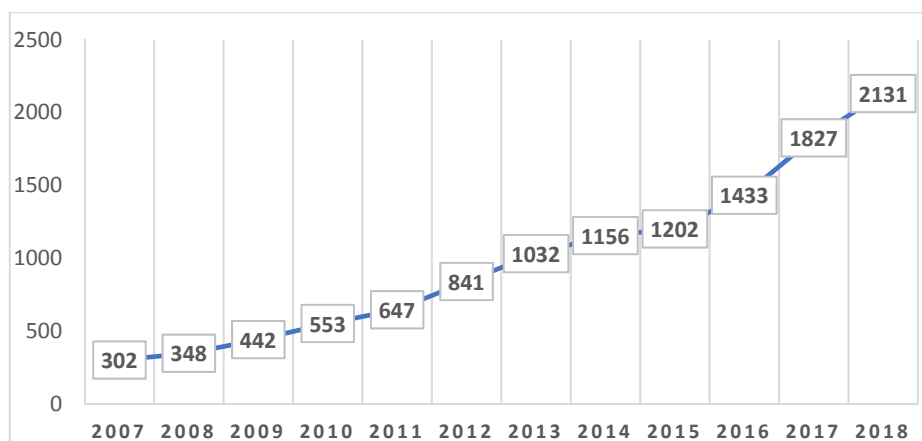


Fig. 1. CEEOL – collection development of key peer-reviewed academic and independent journals

The CEEOL eBook collection development started in 2016 and is offering an ever-growing number of eBooks – more than 2700 at the end of 2018. CEEOL Grey Literature Collection provides researchers access to more than 3,000 Grey Literature documents. CEEOL has become a valuable resource for research and exploration of the region and for training in the fields of philology, linguistics, literature, political science, history, economics, law, management, etc. CEEOL has a clear focus on SSH from CEE/SEE with more than 235,000 articles in social sciences, more than 180,000 articles in language and literature studies, more than 174,000 articles in politics and political sciences, more than 173,000 articles in history, more than 127,000 articles in economy, etc.

More than a quarter of the articles in CEEOL are in English, indicating the orientation of the region's editions to a wider audience and the possibility of including these articles in the educational and research process in different countries. More than 6% of articles in CEEOL are in Bulgarian (Fig. 2). More than 8% of the articles are Bulgarian.

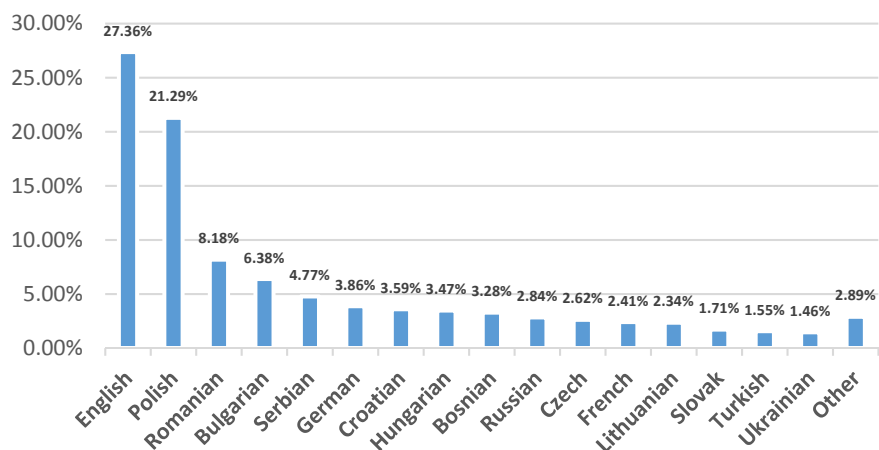


Fig. 2. Percentages of the articles in different languages in the end of 2018.

3 Specifics of the Bulgarian SSH collection of academic and scientific periodicals in CEEOL

The Bulgarian collection of academic and scientific periodicals occupies one of the significant places in CEEOL (Fig. 2). Bulgaria is on one of the top positions in the number of articles published in CEEOL by language and place of publication. Within the CEEOL database the Bulgarian editions have their content specifics. It concerns not only the high scientific value, based on the process of selection of the journals, but also the completeness of the digitized editions. Most of the Bulgarian editions have been presented since their release, which in some cases means over forty years of archives.

In the beginning of 2019 the characteristics of the Bulgarian collection are:

- 140+ peer-reviewed academic journals
- 45,000+ indexed and archived articles
- 1000+ eBooks / Grey Literature documents
- 33%+ of the Bulgarian journals – with 10 to 44 years of digital archives (see examples in Table 1.)
- 40%+ of the Bulgarian journals have complete archives.

Table 1. Examples of the depth of the digital archives of Bulgarian SSH journals in CEEOL

Title Original	Publishing House	First Year	Last Year	Archive (Years)
Българска етнология (Bulgarian Ethnology)	Institute of Ethnology and Folklore Studies with Ethnographic Museum at the Bulgarian Academy of Sciences (BAS)	1975	2018	44

Български фолклор (Bulgarian Folklore)	Institute of Ethnology and Folklore Studies with Ethnographic Museum at BAS	1975	2018	44
PALAEOBULGARICA / СТАРОБЪЛГАРИСТИКА	Cyrillo-Methodian Research Centre at BAS	1977	2018	42
Старобългарска литература (Medieval Bulgarian Literature)	Institute for Literature at BAS	1979	2018	40
Études balkaniques (Balkan Studies)	Institute of Balkan Studies and Center of Thracology "Prof. Alexander Fol" at BAS	1990	2018	29
Критика и хуманизъм (Critique and Humanism Journal)	Human and Social Studies Foundation – Sofia	1990	2018	29
Социологически проблеми (Sociological Problems)	Institute for the Study of Societies and Knowledge at BAS	1991	2018	28
Литературна Мисъл (Literary Thought)	Institute for Literature at BAS	1991	2018	28
Население (Population)	Institute for Population and Human Studies at BAS	1992	2018	27
Философски алтернативи (Philosophical Alternatives)	Institute for the Study of Societies and Knowledge at BAS	1992	2018	27
Bulgarian Historical Review	Institute for Historical Research at BAS	1993	2017	25
Архив за средновековна философия и култура (Archive for Medieval Philosophy and Culture)	Iztok-Zapad Publishing House	1994	2018	25
Divinatio	Maison des Sciences de l'Homme et de la Société (Sofia)	1997	2018	22
Икономически изследвания (Economic Studies)	Economic Research Institute at BAS	1997	2018	22
Историческо бъдеще (Historical future)	Klio Association	1997	2017	21
Психологични изследвания (Psychological Research)	Institute for Population and Human Studies at BAS	1998	2018	21
Балканистичен Форум (Balkanistic Forum)	South-West University "Neofit Rilski", Blagoevgrad	1998	2018	21

Годишник на Асоциация за антропология, етнология и фолклористика „Онгъл“ (Yearbook of “Ongal” Association for Anthropology, Ethnology and Folklore Studies)	“Ongal” Association for Anthropology, Ethnology and Folklore Studies	2000	2018	19
Българско музиковедие (Bulgarian Musicology)	Institute of Art Studies at BAS	2000	2018	19
Проблеми на изкуството (Art Studies Quarterly)	Institute of Art Studies at BAS	2001	2018	18
Икономическа Мисъл (Economic Thought)	Economic Research Institute at BAS	2001	2018	18
Исторически преглед (Historical Review)	Institute for Historical Research at BAS	2001	2018	18
Езиков свят - Orbis Linguarum (Linguistic World)	South-West University “Neofit Rilski”, Blagoevgrad	2001	2018	18
Правна Мисъл (Legal Studies)	Institute for the State and the Law at BAS	2002	2018	17
Електронно списание „Диалог“ (Dialogue - E-Journal)	D. A. Tsenov Academy of Economics	2002	2018	17
Scripta & e-Scripta	Institute for Literature at BAS	2003	2018	16
Научни трудове на Института за държавата и правото (Research Works of the Institute for Legal Studies)	Institute for the State and the Law at BAS	2004	2018	15
Поредица „Софийски диалози“ (Sofia Dialogues Series)	Maison des Sciences de l’Homme et de la Société (Sofia)	2005	2018	14
Икономика и управление (Economics and Management)	South-West University “Neofit Rilski”, Blagoevgrad	2005	2018	14
CAS Sofia Working Paper Series	Centre for Advanced Study Sofia (CAS)	2007	2018	12

CEEOL seeks to give its users the opportunity to access the full streams of digitized scientific journals, and some of the journals, including some Bulgarian ones, do not exist in their entirety in any other e-library with their printed or digital copies, as they do in CEEOL.

The Bulgarian collection of journals in CEEOL has been steadily developing in since 2015 due to the ability of more and more university journals to meet the requirements

for scientific publishing (Fig. 3.) and to be added to the academic journals of institutes of the Bulgarian Academy of Sciences in CEEOL.

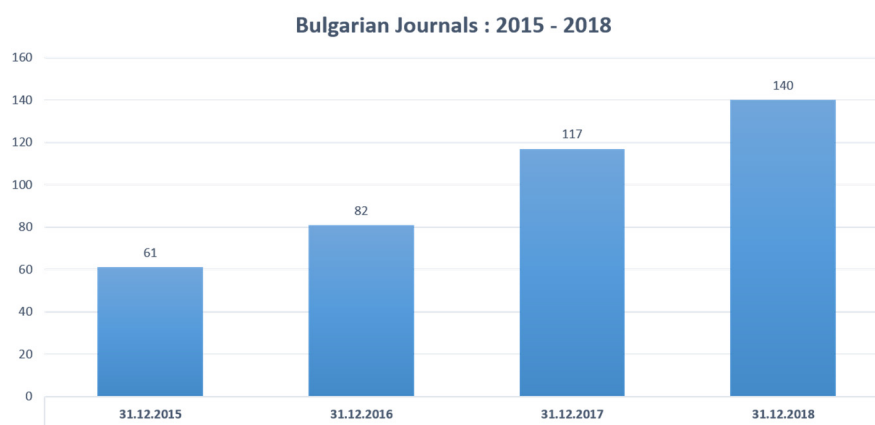


Fig. 3. Bulgarian collection of journals in CEEOL: 2015 – 2018

CEEOL database includes paid and open access (OA) journals. Specific for the Bulgarian collection is the balance between paid and OA journals (Fig. 4.) In the end of 2018 for the first time OA journals were more than the paid ones, showing a tendency for the periodicals' orientation to Open Access.

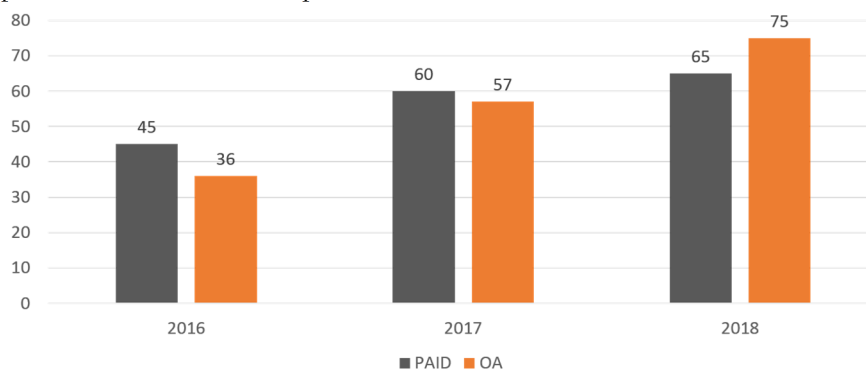


Fig. 4. Bulgarian paid and open access journals in CEEOL

4 Services and Functionalities

4.1 CEEOL Services and Functionalities for Publishers

Since 2016, the new CEEOL system provides a dedicated module for each publishing house, where it has the opportunity to independently develop and manage its collection

of editions following the guidelines and requirements for CEEOL's content management. A new level is the capabilities and benefits of publishers available in the previous system, such as, for example:

- inclusion of all publications in full-text electronic format in the original language concerned;
- support for indexed and referenced bibliographic information database available through the Internet (via CEEOL, as well as other detection and indexing systems: Google Scholar, ProQuest's Serial Solutions, Summon, Primo Central and Alma, Ebsco EDS and Knowledge Base) and offering the opportunity to purchase e-publications electronically;
- the possibility of digitizing (converting in electronic form) valuable publications that a publishing house has only in paper form, if they'll be included in CEEOL;
- the ability to distribute the editions globally free of charge;
- enabling use of full-text indexed archives CEEOL in applying for inclusion in Web of Science;
- complete and detailed information on the downloaded articles, books and chapters of books (by title and number) and on the cash balance of the sales of the paid publications;
- tutorial, initial training, technical assistance and consultancy for work with the specialized software of the CEEOL content management system for registration of publications, etc.

After submitting an application for inclusion in CEEOL (in the For Publishers section - Apply to Include Your Journal), each new periodical passes a check whether it meets the scientific publishing requirements that are comparable to those in the Web of Science and is approved or recommended for necessary adjustments.

The Content Management module (Fig. 5) enables publishers to manage the entire publishing process through their own content-administrators, who are delegated the respective rights to work with the database and with the eBooks (see Fig. 6), the Journals-Issues and the Grey Literature content registration and management.

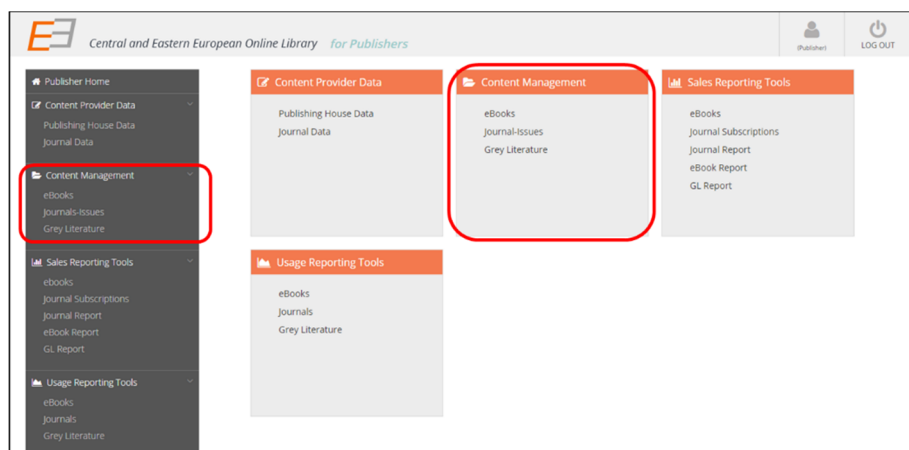


Fig. 5. Content Management Module

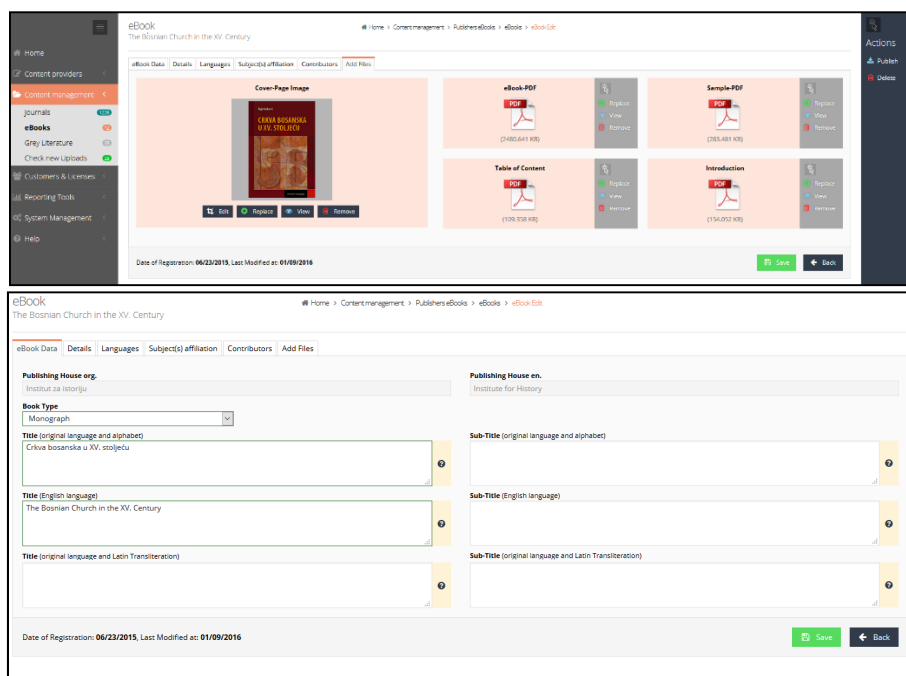


Fig. 6. Example of content management in CEEOL (eBooks)

The interface of the CEEOL Content Management System is convenient, easy and enjoyable to work with. The publisher access level provides a tutorial as well as systematized guidelines and requirements for registration of CEEOL publications in the relevant language.

Regarding the registration of the articles of periodical editions, metadata include: Article Type; Title (original language and writing system); Title (English language; Number of pages; Page Range; Keywords; Summary/Abstract, etc. (Fig. 7).

In separate sections, articles are also registered according to Language, Subject(s) affiliation and Contributors, where the metadata are structured according to the specifics of SSH publications (for example, there are 14 main research areas and more than 300 special research areas in the section Subject(s) affiliation), and the article file is attached (as a fulltext in searchable .pdf format). All metadata are in English, except the original names of the publisher and the journal, as well as the original title of the article (if not in English).

4.2 CEEOL Search Services and Functionalities

Each user has free, full and immediate access to the metadata for the relevant article (Fig. 8), and when downloading a publication as the first page, CEEOL provides the bibliographic information required for quoting.

CEEOL enables several ways to reach the required information.

Article
Journal: Strategies for Policy in Science and Education, Year:2019, Issue:2

Home > Content management > Journals > Issues > Issue/Articles > Article Edit

Article Languages Subject(s) affiliation Contributors Add Files

Article Type
Scholarly Article

Title (original language and alphabet)
Съвременни технонаучни инфраструктури - същност, перспективи и политики

Title (English language)
Modern Technoscientific Infrastructures - Nature, Perspectives and Policies

Sub-Title (original language and alphabet)

Sub-Title (English language)

Number of pages 14 **Page Range** 127-140 **Price in €** 2.5 €

Keywords
technoscience; knowledge and technology transfer; innovation infrastructures; cluster approach; industry 4.0; science and innovation policy

Summary / Abstract
Some basic ideas in the contemporary concept of technoscience are pointed out. Technoscience is considered as a new phenomenon based on the convergence of science and technology and their gradual hybridization, leading to the creation of an inseparable symbiotic link between them. It involves the emergence of a common target orientation, a strong convergence of scientific research and technological innovation, existence of similar objects of study and treatment, and the achievement of similar results and products. Another important feature of technoscience is the growing role of the social factor for its development. Some modern technological and organizational innovation infrastructures such as clusters, techno-parks, "smart cities", "living labs", etc., are presented, considering

Digital Object Identifier (DOI)

Fig. 7. Metadata for article description at registration in CEEOL

Стратегии на образователната и научната политика

Strategies for Policy in Science and Education

Publishing House: Национално издателство за образование и наука „Аз-буки“

Subject(s): Social Sciences

Frequency: 6 issues

Print ISSN: 1310-0270


Online-ISSN: 1314-8575

Status: Active

Issues/Articles Journal Information

2016 | 2017 | 2018 | 2019

Issue No. 1/27 Issue No. 2/27

 **Year:** 2019
Volume: 27
Number: 2

[Table of Content](#)

Articles list

Съвременни технонаучни инфраструктури - същност, перспективи и политики

Съвременни технонаучни инфраструктури - същност, перспективи и политики
(Modern Technoscientific Infrastructures - Nature, Perspectives and Policies)

Author(s): Ludmila Ivancheva
Language: Bulgarian
Subject(s): Social Sciences, Education, Library and Information Science, Communication studies, Sociology, Theory of Communication, Vocational Education, History of Education, Educational Psychology, Social Informatics, Globalization, Sociology of Education
Issue: 2/27/2019
Page Range: 127-140
No. of Pages: 14
Keywords: technoscience; knowledge and technology transfer; innovation infrastructures; cluster approach; industry 4.0; science and innovation policy
Summary/Abstract: Some basic ideas in the contemporary concept of technoscience are pointed out. Technoscience is considered as a new phenomenon based on the convergence of science and technology and their gradual hybridization, leading to the creation of an inseparable symbiotic link between them. It involves the emergence of a common target orientation, a strong convergence of scientific research and technological innovation, existence of similar objects of study and treatment, and the achievement of similar results and products. Another important feature of technoscience is the growing role of the social factor for its development. Some modern technological and organizational innovation infrastructures such as clusters, techno-parks, "smart cities", "living labs", etc., are presented, considering them as an essential prerequisite for the sustainable development of technoscience. Key policy initiatives and measures to support them are presented at both European and national level.

Price: 2.50 €

[Add to Cart](#)

[Preview](#)

Академичното самоуправление в България - традиции и предизвикателства

Action Research in Higher Education: Fostering Transition toward a Sustainable Economy: Trainees in Entrepreneurship at Three Bulgarian

Fig. 8. An example of the metadata for the description of the article in CEEOL

The simple search can be performed across the entire database (Content search), including the metadata and full text indexed in all publications, based on the language string (characters, words or phrases), or can be limited to search the publisher metadata (Publisher), name of the book (Book Title), name of the journal (Journal Title), keywords (Key Words), ISSN, ISBN, author name (Author Name) etc. The ability to search for sequences (Title series) is very convenient for reaching collections with a particular thematic focus.

Complex search provides multiple possibilities for concretizations and narrowing of the results (Fig. 9).

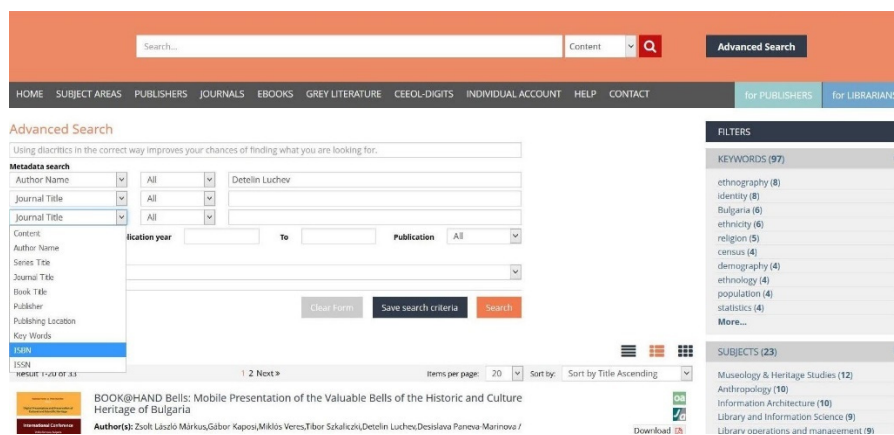


Fig.9. Search Services in CEEOL

In addition to searching, access to the required results can also be achieved on the basis of special pages browsing, where all Subjects Areas, Publishers, Journals, eBooks, Gray Literature or series of books and papers are structured. The pages themselves have the additional functionality to support browsing by means of restrictions by country, first letter, etc.

The results of each search or browse can be further selected on the basis of filters (separate and complex), such as: Keywords, Subject, Authors, Languages, Content Type, Access type, etc. The sequence of categories in some of the filters (Keywords, Subjects, and Authors) is displayed based on the frequency of the category in the results, while others provide with all possible options for selecting the results based on the metadata. These filters are particularly useful in the new functionality of the system in terms of browsing by thematic areas, journals, etc. (see Fig 9, on the right). For example, if an user selects an area in the Subject Areas browse page (or Subjects areas to the fourth level of specification) that is of interest to him, he can filter it by Content Type through one or more of the options: Journal, eBook, Grey Literature, Book Chapter, Journal Article, and/or apply other filters in terms of language (e.g. Bulgarian), access (Open Access), an author (Authors) and so on.

The diverse and fast ways to search and access the required information in metadata and in the full texts of journals and books characterize CEEOL as an electronic environment that is adequate to the skills of every Internet user and reader.

5 Conclusion

In the present paper, Central and Eastern European Online Library (CEEOL) was selected as an example of good practice for providing open access to digital repositories for scientific publications in the humanities and social sciences. It is the primary database providing full-text access to academic journals and books in electronic form in the humanities and social sciences from and for Central and Eastern Europe, with more than 75% of the editions indexed in CEEOL providing open access. Access to CEEOL by many search and indexing systems such as Google (which links to CEEOL are always among the top 10), Google Scholar, ProQuest's Serial Solutions, Summon, Primo Central and Alma, Ebsco's EDS and Knowledge Base, etc., as well as a partnership with Clarivate Analytics, provides additional opportunities for publishers and authors. The services, functionalities and capabilities offered by CEEOL as an information system are a prerequisite for the growing interest in academic, university and other scientific publishing houses, as well as for subscriptions by university and public libraries all over the world. A trend is the growing interest of institutional subscribers from Central and Eastern Europe in CEEOL. The data on the long-standing use of CEEOL database in three Bulgarian universities and the trial access in more than ten other Bulgarian universities, characterized CEEOL as an opportunity and necessity for the training and research process in the field of SSH.

The CEEOL content management system can be considered as a good practice for providing open access to digital repositories for social sciences and humanities publications in Bulgaria and many of the concepts of its CMS realization can be included and used in the developed systems for open access to the publications in SSH area.

Acknowledgements.

This work was partially supported by the Bulgarian Ministry of Education and Science under National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security", approved by DCM No 577, 17 August 2018.

References

- Aksnes, D.W. & Sivertsen, G. (2019). A criteria-based assessment of the coverage of Scopus and Web of Science. *Journal of Data and Information Science*, 4:1, 1–21, DOI: 10.2478/jdis-2019-0001
- Bayer, F., et al. (2019) Investigating SSH Research and Publication Practices in Disciplinary and Institutional Contexts. A Survey-Based Comparative Approach in Two

- Universities. *Frontiers in Research Metrics and Analytics*, 4:1. DOI: 10.3389/frma.2019.00001
- Harzing, A.W. & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison, *Scientometrics*, 106 (2), 787–804. DOI: 10.1007/s11192-015-1798-9
- Kulczycki, E., et al. (2018) Publication patterns in the social sciences and humanities: evidence from eight European countries. *Scientometrics*, 116 (1), 463–486
- Lavik, G.A.V. & Sivertsen, G. (2017). ERIH PLUS – Making the SSH visible, searchable and available. *Procedia Computer Science*, 106, 61-65.
- Mongeon, P. & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106 (1): 213–228. DOI: 10.1007/s11192-015-1765-5
- Pavlovska, E. (2018). Grey Literature” in Electronic Archives. *Digital Presentation and Preservation of Cultural and Scientific Heritage*. Vol. 8, Sofia, Bulgaria: Institute of Mathematics and Informatics – BAS, 105-108.
- Sîle, L., Guns, R., Sivertsen, G. & Engels, T. (2017). *European databases and repositories for Social Sciences and Humanities research output. Report*. DOI: 10.6084/m9.figshare.5172322.v2.

Received: June 05, 2019
Reviewed: July 10, 2019
Finally Accepted: July 23, 2019