Infotheca: Journal for Digital Humanities - 2000-2026 -

UDC 050:02(497.11)INFOTEKA
DOI 10.18485/infotheca.2026.26.1.4

ABSTRACT: *Infotheca*, Journal for Digital Humanities, is a multidisciplinary journal which publishes original papers that are subject to review by the members of an international Editorial Board. It started publishing in 2000 as *Infotheca: Journal for Informatics and Librarianship*, and until today was transformed in many ways that are going to be presented in this paper. From issue 1/2 in 2007, the subtitle has been changed into *Journal for Librarianship and Informatics*, and from issue 1 in 2014, into Journal for Digital Humanities. *Infotheca* is published biannually in Open Access, it is bilingual (papers are published in Serbian and English), and covers topics such as librarianship, informatics and digital humanities. Thus, it became an excellent language resource for various types of research.

KEYWORDS: scientific journals, scientific publishing, *Infotheca*, University Library „Svetozar Marković“, library and information science, digital humanities, computer linguistics, COBISS, Biblisha, Wikidata

PAPER SUBMITTED: 20 February 2026
PAPER ACCEPTED: 06 April 2026

Jelena Andonovski
ORCID 0009-0007-9764-5364
andonovski@unilib.rs

University Library "Svetozar Marković" Belgrade, Serbia

1 Introduction

Infotheca is a multidisciplinary journal in digital humanities which publishes original papers that are subject to review by the members of an international Editorial Board. It also publishes reviews of events and publications in this scientific field. The journal is in Open Access, which means that all content is available free of charge to the user or institutions. It is the copyright holder of published articles, under the conditions defined in Creative Commons

license Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)[1], and users are allowed to read, download, copy, distribute, print, search, or add links to the full texts of the articles in this journal without asking for permission. Content of the journal is published bilingually (in Serbian and English), biannually, and the publisher is the University Library "Svetozar Marković," in cooperation with the Faculty of Philology of the University of Belgrade[2] and the Serbian Academic Library Association (SALA)[3]. In its 26 years of existence, 25 volumes, 39 issues and nearly 600 papers have been published.

From its inception until today, *Infotheca* has been changed in different ways that have affected its profile. The changes concerned not only the structure and quality of the journal and published papers, but also the scientific community that publishes its research results there, which over time has become significantly broader than at the beginning. The last ten years were especially challenging for the editorial team because the journal was changed step by step, following technological innovations in scientific publishing. In addition, all significant changes in the scientific journal categorization process in the Republic of Serbia, in accordance with European and global scientific publishing policies, have largely shaped the quality of *Infotheca*. In accordance with that, the editorial team makes great efforts to fulfill the criteria from the Rulebook on Categorization and Ranking of Scientific Journals[4]. Until 2024, the Journal was categorized as M53, and according to the List of Categorized Scientific Journals of Domestic Publishers for 2025[5], the categorization is now M52. This is a great success, considering that the journal ranks high in the category of domestic scientific journals for electronics, telecommunications and information technology.

For 26 years, *Infotheca* has undergone various changes (more in Section 2), and today it is completely online and in Open Access, with online editing and publishing system (more in Section 3), it is part of different databases such as the COBISS catalogue (more in Section 4), Bibliša (more in Subsection 5.2), Wikidata (more in Subsection 5.3), Wikipedia. Also, it is an important digital language resource in various research areas, such as computational linguistics (more in Section 5).

1. Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)
2. University of Belgrade, Faculty of Philology
3. Serbian Academic Library Association
4. Rulebook on Categorization and Ranking of Scientific Journals
5. List of Categorized Scientific Journals of Domestic Publishers for 2025. Accessed 9. 2. 2026.

2 *Infotheca* – Historical Overview

From 1995 to 2000, the Serbian Academic Library Association (SALA) had been published *Information of the Serbian Academic Library Association* (ISALA), an information bulletin for librarians, documentalists, information specialists and other professionals in libraries and INDOC (information and documentation) centers and other units within the universities of Belgrade, Niš and Novi Sad. The bulletin was launched with the aim of providing important information to library and information experts in higher education and research institutions. It was published quarterly, in A4 format, in Serbian, and the editor-in-chief was Milena Matić, the head of the Supervision Department at the University Library "Svetozar Marković".

Interest in publishing papers in ISALA was great, not only in the field of library and information science, but also in the whole scientific research community of university libraries in Serbia. However, the regular sections of the paper did not have the capacity to publish such an increased volume of information, as it had not been designed for that purpose. Also, the need arose to provide more space for scientific and professional articles. Thus, at the Conference "Theoretical Aspects of Scientific Information in the 21st Century and the Real Possibilities of Our Libraries" held within the Sixth Assembly of the SALA, ISALA Editorial Board, together with the SALA Executive Board decided that ISALA should be transformed into *Infotheca: Journal for Informatics and Librarianship*. The goal of the new journal was to publish: original papers about the application of information technology in librarianship, scientific information in general, the impact of informatics on the modern development of academic and special libraries, as well as new knowledge in information and library science. The first issue was published in September 2000, and the journal has been published regularly ever since. Until the first issue for the year 2014, the publisher was only the SALA; from that issue onward, the University Library "Svetozar Marković" and the Faculty of Philology of the University of Belgrade, also joined as publishers. The first change of the subtitle, into *Journal for Librarianship and Informatics*, occured in issue 1/2 (2007). At the same time, as the journal started to be published in English (in a printed volume in reverse), the title in English on the back cover was determined, as well as the subtitle in issue 1 (2012).

Over time, *Infotheca* began to publish papers from the broader field of digital humanities, so, from the first issue for the year 2014, the subtitle was changed into *Journal for Digital Humanities. Infotheca* became the first and

only scientific journal in the Republic of Serbia with this title, and the only one that exclusively publishes papers in this scientific field.

2.1 Editors-in-chief

From the inception in 2000 until issue 1/2 in 2007, the editor-in-chief was Dušan Surla, professor at the Faculty of Sciences at the University of Novi Sad. For the next 16 years (until the first issue in 2023), the editor-in-chief was Cvetana Krstev, professor of Computer Science in the Department of Library and Information Science at the Faculty of Philology of the University of Belgrade, and also a member of the Language Resources and Technologies Society (JeRTeh)[6]. From the second issue in 2023, editors-in-chief have been Ranka Stanković, professor at the of Applied Mathematics and Informatics at the Faculty of Mining and Geology of the University of Belgrade, and Dr. Aleksandra Trtovac, library advisor in the Cataloging Department at the University Library "Svetozar Marković".

As digital humanities is a „transdisciplinary scientific area... a heterogeneous field of research between IT, cultural studies and humanities in general" (Pajxepт 2014, 22), i.e. a multidisciplinary scientific field which encompasses the use of digital resources in the humanities, it is changing the way scientific work is conducted. As noted in (*The Digital Humanities Manifesto 2.0* 2026; Вранеш 2014) "universities are called upon to shape natively digital models of scholarly discourse for the newly emergent public spheres of the present era". In this context, the articles published in *Infotheca* are also multidisciplinary, which was the main reason for appointing two editors-in-chief from different, but related scientific research fields.

The implementation of the open source software for editing and publishing journals in Open Access, Open Journal System (OJS)[7] (more in Section 3), and the transition of the entire journal editorial policy to this software from issue 1/2 (2016), created the need for an editor of the online edition. From the beginning, it has been Dr. Jelena Andonovski, library advisor in the Cataloging Department at the University Library "Svetozar Marković", with occasional assistance from her colleague Ivana Gavrilović, a librarian advisor in the Department for Digitization and Cultural Presentation at the Library.

6. Language Resources and Technologies Society
7. Open Journal System (OJS)

2.2 Publication Frequency, Language and Article Categorization

Although it can be concluded that *Infotheca* journal was published continuously for two and a half decades, the dynamics were diverse. In the first decade, it was published once a year as a double issue, except in 2000 and 2003, when two independent issues were published. In the next year, 2004, a double issue and a special issue were published, and in 2005, issue 3. From 2010 to 2014, it was regularly published twice a year. A milestone was 2015, when an open source and free software for the management of peer-reviewed academic journals, Open Journal System (OJS), was implemented (more in Section 3). As the editorial team needed some time to adapt to this change, the journal was not published this year. Since a double issue in 2016, the journal has continued to come out regularly and is entirely edited using this software. Until 2025, *Infotheca* has been published regularly, but not always twice a year. In 2020, a double issue (1/2) was published, in 2022 and 2024, one issue each, while in 2025, a double issue was also published (1/2).

Regarding the journal language, papers are published entirely bilingually, in Serbian and English. At the very beginning, papers were mostly published only in Serbian, with a title and abstract in English, as mandatory data. Until 2004, some papers categorized as "review paper" or "professional paper" were published in Serbian and English, or the authors themselves requested that their papers be published in both languages. Papers by foreign authors were always published bilingually, as in the case of the paper *Academic research in Serbia and available database resources* by Sam Brooks in issue 2, in 2003. A similar case was for the translation of foreign documents, such as the *IFLA Internet Manifesto*, published in the same issue, and with papers on projects such as TEMPUS. The first fully bilingual issue was a special issue published in 2004[8], dedicated to the results of the first completed TEMPUS project in Serbia. Since then, some papers by Serbian authors have begun to be published bilingually, and this has become the established practice since issue 1, in 2011. Until today, the exception was the second issue in 2019, and the second issue in 2021, which were special and published only in English.

It is important to point out that, from its inception until today, every article has included a UDC number as mandatory metadata, while the papers have been grouped into categories in the table of contents (Communications, Publications, Professional meetings, Articles, News etc.). Within the papers themselves, the categorization was not explicitly mentioned, unless it was a "review paper" or "professional paper", and even then, not consistently. The

8. Vol. 5, special issue (2004)

relevant category was indicated with the appearance of the first paper in each issue. Since 2009, when the *Act on Scientific Journal Editing*[9] entered into force, the criteria for journal editing have become significantly stricter, and the articles' categorization has become mandatory metadata. So, from the first issue in 2010, it has been indicated at the beginning of each paper in *Infotheca*.

2.3 Format and Design

Initially, *Infotheca* was published in A4 format. With issue 1/2 in 2007, the format was changed to B5, and from issue 1/2 (2016) to A5, which is still in use. Until 2007, text of the papers published bilingually was printed in double columns, with Serbian and English versions in parallel. Since 2007, when the English version was established, papers have been printed in double columns, but the English version is printed in reverse, with a separate cover page. With the format change in 2016, the English version has continued to be published in reverse, but, due to the reduced format and impracticality, the text is no longer printed in double columns.

In addition to the format change, the design was also changed with issue 2 (2013). From the beginning of the journal, publishing and page layout design were done using the software *Indesign*. Since Prof. Cvetana Krstev became editor-in-chief, this job was done by Vanja Radulović from the University of Belgrade Information Center (IC), and at that time, a master's student at the Department of Library and Information Science at the Faculty of Philology, University of Belgrade. After her, for a very short period (until 2018), this job was done by Nataša Matović, a graphic editor at the University Library "Svetozar Marković".

Since the first issue for the year 2018, LaTeX has been used for publishing and page layout design. Because of that, the journal design was changed and prepared according to the LLNCS.CLS document format. The class LLNCS is an extension of the standard LaTeX *article* class[10], which Springer[11] has adapted for the purpose of preparing papers for submission to the conference

9. Ministry of Science and Technological Development has adopted the Act on Scientific Journal Editing (110-00-17/2009-01, from 9 July 2009) with the aim of helping editorial teams get better quality and improve domestic scientific periodicals, thereby enabling greater inclusion of journals in the scientific information system, both domestically and internationally. Accessed 1 April 2026,

10. LaTeX class *article*

11. Springer

proceedings series *Lecture Notes in Computer Science (LNCS)*.[12] Style files as templates and example articles are preloaded into Overleaf[13], allowing authors to create LNCS-formatted documents prior to paper submission. Overleaf is an open-source online collaborative LaTeX editor that enables online editing and page layout design of the paper before submission. It is necessary that all authors and journal collaborators (editors, reviewers, copyeditors, proofreaders, editorial board members, and technical editors) have an account created on the platform. All changes to one paper, including reviewers' remarks and comments, author corrections, and editorial revisions, can be made in a single document from which the final version is produced and published.

Based on this approach, Prof. Cvetana Krstev assisted Branislava Šandrih, at that time Assistant Professor of Computer Science in the Department of Library and Information Science at the Faculty of Philology, University of Belgrade, in adapting the Springer LLNCS.CLS document format for *Infotheca*, using Springer guidelines[14]. The adapted document class, with templates and article examples, was uploaded to Overleaf. So, since issue 1 (2018), authors familiar with LaTeX who wish to prepare their papers in Overleaf before submission can use these templates. While most of the authors are ready for this step, because many journals use it, many reviewers and copyeditors are still hesitant; therefore, it will be necessary for the *Infotheca* editorial team to pay more attention to their training in the future.

In addition to changes in the design of *Infotheca*, the cover design introduced in issue 1/2 (2007) and is still in use today, was also created by Vanja Radulović.

3 Online Edition

In the early years, *Infotheca* had only a printed edition. But in 2004, during the first TEMPUS project in Serbia, the website of the University Library "Svetozar Marković" was created. As the headquarters of the SALA, *Infotheca's* only publisher at the time, was located in the Library, this moment was used to create the journal's website as well. Since then, it has been published online as well. Also, it was assigned the ISSN number not only for the printed (1450-9687), but also for the online edition (2217-9461).

12. Lecture Notes Computer Science (LNCS)
13. Overleaf
14. Instructions for Using Springer's llncs Class for Computer Science Proceedings Papers

At the end of 2011, a new website of the University Library was created, including a new *Infotheca* website, hosted on the new domain *infoteka.bg.ac.rs*. In 2014/2015, another website was created on the domain *infoteka.unilib.bg.ac.rs*. In addition to the website, in 2015 the open source software for editing and publishing journals in Open Access, Open Journal System (OJS), was implemented. OJS is a free software for the management of peer-reviewed academic journals, created by the Public Knowledge Project[15], and released under the GNU General Public License[16]. It enables the whole issue to be prepared and published online in one system. In other words, the entire researcher-to-reader workflow for submission, peer review, copyediting and production is done in the OJS, where the final issue is published. The system operates though sending email notifications in every step, so editors and collaborators must be registered users, with their own accounts. Another advantage of the OJS is that published articles and issues are fully indexed in global discovery services like Google Scholar, Crossref, DOAJ, and many others, and references can be downloaded in various citation formats: ABNT, APA, BibTeX, CBE, EndNote - EndNote format (Macintosh & Windows), MLA, ProCite - RIS format (Macintosh & Windows), RefWorks, Reference Manager - RIS format (Windows only), Turabian, etc. The electronic issue 1/2 (2016) of *Infotheca* was the first fully prepared and published issue using OJS, but it was also posted on the existing website. The editorial team, as well as collaborators (authors, reviewers, copyeditors, proofreaders, technical editors), needed some time to get used to this editorial system; therefore, the issues up to 2019 were published simultaneously on the OJS and the journal's website. Since 2020, the online edition has been available only on the OJS. Over time, the entire *Infotheca* archive was transferred into the OJS, and the old website was officially shut down in 2024.

According to the Rules of Procedure and the Manner of Evaluation and Quantitative Presentation of Scientific Research Results of Researchers, DOI (Digital Object Identifier)[17] became mandatory metadata for articles in scientific journal online editions, which meant that articles in Infotheca were

15. Public Knowledge Project

16. GNU General Public License

17. A digital object identifier (DOI) is a string of alphanumeric characters used to uniquely identify electronic documents. DOI is a permanent, and thus more reliable method of referring to an electronic document than the URL. It was introduced to solve the problems of URI addresses of electronic documents, such as the existence of multiple different addresses for the same e-document, address unreliability, and the

required to have DOI identifiers. The linking of article metadata, DOI identifiers, and web locations (URL), that is DOI numbers registration, is done via the service CrossRef[18]. Since the University Library's publishing activity is not extensive enough to justify a direct agreement with the CrossRef agency, the Faculty of Philology of the University of Belgrade, as one of the Infotheca's publishers and an institution which already had a contract with the agency, assumed responsibility for DOI registration for Infotheca. The journal thus became part of the repository doiFil of the/ Faculty of Philology, University of Belgrade.[19] Since then, when one issue is ready for publication, *Infotheca's* online editor prepares metadata required for DOI registration and sends it to the DOI editor at the Faculty of Philology (Miloš Utvić, Assistant Professor at the Department for Library and Information Science) for the review and registration. For the registration, editor prepares the following metadata: title, author(s) name(s), ORCID[20], journal title, ISSN, eISSN, year, volume, issue, pages, language, article number in the issue, DOI number, DOI URL and the full-text URL. After registration, the DOI structure is: *10.18485/infotheca.year.volume.issue.article_number_in_the_issue*. Since issue 1/2 (2016), the DOI identifiers have been assigned to all scientific and professional papers in English and thanks to that, the papers, their authors, and *Infotheca* itself, have become "visible" to the professionals and scientific community, which significantly contributes to its international recognition (Milinković 2016, 76).

4 *Infotheca* in COBISS

Infotheca, like every publication published in the Republic of Serbia, has its own bibliographic record in the COBISS catalogue. A bibliographic record is a metadata set about library resources that contains data elements necessary for identifying and retrieving a resource, as well as additional supporting information, presented in a standardized bibliographic format. Regarding journals and serials in general, in addition to basic metadata (ISSN number, frequency, title, editor-in-chief, inception, publisher, important notes on the serial, link to the online edition record if available, etc.) and the signature

like. In 2012, the International Organization for Standards (ISO) adopted the ISO 26324:2012 standard, which defines the DOI identifier system.

18. CrossRef

19. doiFil / Faculty of Philology, University of Belgrade

20. From the second issue in 2023, ORCID number is mandatory metadata with the author(s) name, email and affiliation.

mark, the number of published issues is also significant. COBISS also enables the creation of analytical records for all published articles in one journal, which are linked to the journal's record.

Regarding *Infotheca*, there are two COBISS records. The first one is for the printed edition (ID 166890503)[21] with the following metadata: source type (continuous source), record type (printed source), bibliographic level (serial publication), ISSN number (1450-9687), language, country of publication, frequency, title and intellectual responsibility (journal title, editor-in-chief), publisher, important notes about the journal, link to the preceeding publication (regarding *Infotheca*, link is connected to the COBISS record for ISALA – ID 103379463[22] via title and ISSN number), link to the other edition in other medium (link is connected to the COBISS record for the online edition via title and ISSN number), parallel title (English title in *Infotheca*), UDC number and the journal's URL. In addition to this metadata, *Infotheca's* COBISS record also has a signature mark *M Č 1046*. As it is a serial publication, signature marks are also assigned to individual issues, including metadata such as year, volume and issue. After this work is done, all issues have signature marks in the form *M Č 1046/year* (for example, *M Č 1046/2010*).

The corresponding record exists for the online edition (ID 192501004)[23], with the following metadata: source type (continuous source), record type (online source), bibliographic level (serial publication), ISSN number (2217-9461), language, country of publication, frequency, metadata about the online source, title and intellectual responsibility (journal title, editor-in-chief), publisher, online source characteristics, important notes about the online source, other edition in the other medium (link is connected to the COBISS record for the printed edition via title and ISSN number), parallel title (English titile in *Infotheca*), UDC number and URL.

In addition to the record for the *Infotheca* itself, COBISS also contains analytic records for a certain number of papers published in it. At the time of writing the paper, the shared cataloging system included analytic records for 381 papers published in *Infotheca*, of which 238 are in the local COBISS catalogue (University Library catalogue). Papers are described using the following metadata: record type (textual, printed or online source)[24],

21. Permanent link https://plus-legacy.cobiss.net/cobiss/sr/sr/bib/166890503
22. Permanent link https://plus.cobiss.net/cobiss/sr/sr/data/cobib/103379463
23. Permanent link https://plus-legacy.cobiss.net/cobiss/sr/sr/bib/192501004
24. This metadata depends on whether the analytical record is created for an article published in print or electronic edition of a journal.

bibliographic level (analytical level (component part)), material typology (original scientific paper, review paper, professional paper, review), journal ISSN number (using this metadata, analytic record is connected to the journal's bibliographic record, in this case, with one of *Infotheca's* records)[25], language, country of publication, title and intellectual responsibility (article title, author(s) name(s)), physical description (pages, illustration data if applicable, volume, issue and year), notes, parallel title, subject headings, UDC, author(s) metadata (linked to the records in the CONOR database), and full-text URL.

Processed articles from Infotheca can be retrieved from COBISS in two ways. One is by using advanced search and creating a search query „HI=166890503 OR SN=1450-9687“ or „HI=192501004 OR SN=2217-9461“. The other way is to find record for the journal itself via title or COBISS ID number, accessing the list of all linked analytical records - "component parts". In addition to searching and metadata display with access to full texts, it is also possible to export metadata from COBISS in various formats, one of which is XML. However, detailed information retrieval, such as full-text search or metadata grouping by some criteria, is not possible, which is why Infotheca has been integrated into other databases, as will be shown below.

Processed articles from *Infotheca* can be retrieved from COBISS in two ways. One is by using advanced search and creating a search query „HI=166890503 OR SN=1450-9687“ or „HI=192501004 OR SN=2217-9461“. The other way is to find record for the journal itself via title or COBISS ID number, accessing the list of all linked analytical records - "component parts". In addition to searching and metadata display with access to full texts, it is also possible to export metadata from COBISS in various formats, one of which is XML. However, detailed information retrieval, such as full-text search or metadata grouping by some criteria, is not possible, which is why *Infotheca* has been integrated into other databases, as will be shown below.

5 *Infotheca* as an Important Language Resource in Different Research Areas

For years, *Infotheca* has published significant results in the field of library and information science, but for more than a decade, it has been publishing

25. In this case, the analytical record is linked to one of two catalog records for the journal Infotheca, depending on whether the issue was published in print or electronic format.

significant results in the digital humanities, making it a very important resource for researchers in these fields. Furthermore, it is published bilingually and in Open Access, and as the editor-in-chief for more than a decade, Prof. Cvetana Krstev recognized that its content could become an important resource for applying different language tools for Serbian processing developed by the JeRTeh Society and could be used in research in the field of computational linguistics. Therefore, the bilingual (English-Serbian) parallel corpus *INFOteka* was created (more in Section 5.1), which became part of the multilingual digital library Bibliša (more in Section 5.2). From that moment on, different tools for natural language processing have been applied, especially those in the field of information and terminology retrieval and extraction from domain corpora.

As the first corpus that became part of the Bibliša digital library, *INFOteka* was used for testing the system performance and search capabilities. From the beginning, Bibliša has offered metadata search and full-text search, enabling the composition of queries using both simple and multiword keywords in Serbian and English, with the possibility of expanding search queries both morphologically and semantically, using different lexical and terminological resources. Along with this, the corpus was also used for testing the display options for search results and the entire interface environment, after which the possibilities of improving the entire system were analyzed in order to make it more than just a simple search tool (Stanković et al. 2012; Stanković, Obradović, and Trtovac 2012).

An important example of research in which *INFOteka* is used is the doctoral dissertation of Dr. Aleksandra Trtovac (Тртовац 2016, 2017). In this research, the corpus was used in the field of information retrieval and extraction from the library and information science domain. Since *INFOteka* is a domain corpus from this field, and also part of the Bibliša digital library[26], the dissertation analyzed in detailed the information retrieval possibilities within this library and tested its search capabilities. It also investigated lexical resources used to enhance and refine users' queries (Stanković et al. 2015), especially the morphological e-dictionary of the Serbian language (Krstev 2008; Krstev and Vitas 2009), which contains an electronic terminological dictionary from library and information science into which the Dictionary of Librarianship (Kovačević, Injac, and Begenišić 2004) was integrated. The evaluation of the obtained results helped supplement the electronic terminological dictionary from library and information science,

26. At the time of the research, the corpus included parallelized issues from 2007 to 2014.

as well as other lexical resources, with new terms and lexical units found in the *INFOteka* corpus during the research. Also, it helped in the development of a new resource, Biblimir, which should improve Bibliša's search capabilities (Stanković, Obradović, and Trtovac 2012).

The next example of research in which the journal *Infotheca* was used is the study of (Stanković and Davidovć 2021). Within this research, Wikidata was created not only for the journal itself, but also for each article published in it which, at the time of the research, was part of the *INFOteka* corpus in Bibliša. Wikidata was created based on existing metadata about the journal, scholarly articles, and their authors in the corpus. After its creation, links were added to the journal articles on the Bibliša website, and the co-authorship graph was aslo presented as an interactive page. Metadata from Bibliša were exported into CSV format, and then the tools OpenRefine and QuickStatements were used. At the end, a HTML page integrating Wikidata Query Service with Bibliša was prepared. Also, the queries retrieving tables of the latest published articles, frequency of keywords in articles, pictures of authors, author profile table, co-authorship graphs, distribution of authors by sex, etc., were written. More about *Infotheca's* Wikidata is given in Section 5.3.

The latest research example is presented in (Stanković 2026) in which textometry and a comparative analysis of the parallel corpus *INFOteka* were presented, with the aim of examining the lexical, structural, and thematic characteristics of the corpus itself. In this research, individual metadata for Serbian and English subcorpora exported from Bibliša were used, as well as full texts of articles in both languages. By applying different methods (frequency analysis, analysis of collocations, time progression, specification, and topic modeling), the subcorpora were analyzed individually and comparatively; some characteristics and differencies were presented, and the obtained results confirmed that parallel bilingual corpora are a valuable and important resource for linguistics and interdisciplinary research.

5.1 Parallel Corpus *INFOteka*

A parallel corpus is a multilingual collection consisting of one or more texts in one language (the source language) and their translations into one or more different languages (the target language). Parallel corpora are created by pairing semantically equivalent structural elements of texts in the source and target languages, resulting in a set of Translation Units (TUs), each consisting of two or more Translation Unit Variants (TUVs). The Natural

Language Processing Group, University of Belgrade, now the JeRTeh Society, has, since its inception, been developing parallel corpora containing Serbian, either as a source or a target language. In this context, it has developed a parallelization procedure consisting of several steps, which were also applied in the creation of the INFOteka parallel corpus:

1. text preparation and segmentation into alignment units (segments),
2. alignment of units,
3. visualization of parallelized texts, control and correction of produced pairs,
4. generation of parallelized texts in TMX (Translation Memory eXchange)[27] format,
5. splitting the TMX file into individual XML files,
6. verticalization of individual texts and corpus creation.

The first step, text preparation and segmentation into alignment units (segments), involves structuring the texts by marking sections, paragraphs and segments (in practice, most often sentences). Segments are used as translation unit variants, but the alignment of corresponding text parts is performed at all three levels to ensure greater precision, which is necessary for the parallelization program. Papers from *Infotheca* were marked using the following tags: <body> for the entire text, for sections, <head> for titles and subtitles, <p> for paragraphs, and <seg> for marking segments (sentences). After annotation, two XML files were generated in accordance with the Text Encoding Initiative (TEI)[28]: one for the source language (in this case, English) and one for the target language (in this case, Serbian). Both XML files were validated using the appropriate Document Type Definition (DTD)[29].

The files generated in the first step are then parallelized using different tools for the automatic generation of aligned segments. The members of the Belgrade Natural Language Processing Group opted for software packages developed at the "Laboratoire Lorrain de Recherche en Informatique et ses Applications" (LORIA)[30] in France: Xalign, for automatic alignment of segments, and Concordancier, for the visualization and manual correction

27. TMX is an ISO standard (ISO 24616:2018) that defines the storage of so-called translation memories and their exchange between different software translation tools, as well as between different companies that maintain translation memories.
28. Text Encoding Initiative - TEI
29. Document Type Definition - DTD
30. Laboratoire Lorrain de Recherche en Informatique et ses Applications – LORIA

of alignment errors. These tools are integrated into the ACIDE environment for generating aligned parallel texts (Obradović, Stanković, and Utvić 2008). ACIDE consists of two modules: Alignment and TMX. The Alignment module integrates Xalign and Concordancier. The default input consists of well-formed and validated XML files in the source and target languages, and the output is a new XML file with aligned translation unit variants (segments). Parallelization aims for 1-1 alignment, i. e. establishing a direct correspondence between equivalent segments in the source and target languages. ACIDE was used for the parallelization of articles in Infotheca, and after all necessary corrections, 1-1 alignment was achieved.

The fourth step, the generation of a parallelized text in TMX format, is carried out in the second module of ACIDE. This module produces a file in TMX format which consists of the header and the document body. The header contains metadata describing the parallelized text, while the document body consists of a translation units set containing one or more semantically equivalent translation unit variants (Figure 1).

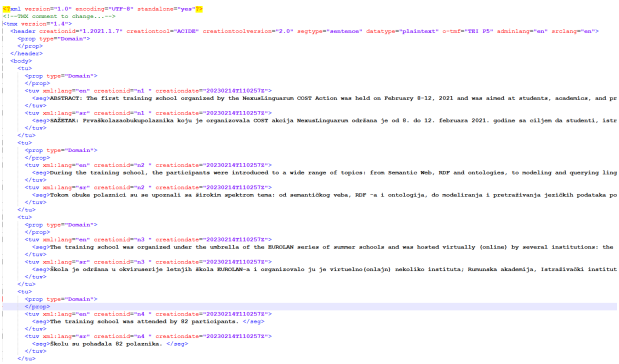

Figure 1. Example of the parallelized text in TMX format

The fifth and sixth steps involve splitting the TMX file into individual texts in XML format and verticalizing individual XML files. Spliting means that the generated TMX file is decomposed into separate XML files for each language (source and target), containing information on translation units. Produced individual XML files are used as input for the paralell corpus. A separate corpus is created for each language, and connections are made that enable parallel search. Before corpus creation, verticalization is performed,

meaning that each token (word, number, punctuation mark, XML tag, XML comment, etc.) is placed on a separate line. These two steps are important when IMS OCWB[31] is used for corpus creation, as was the case with corpora such as SrpKor (Utvić 2013), SrpFranKor (Vitas and Krstev 2004, 2006), SrpEngKor (Krstev and Vitas 2011), developed by the Belgrade Natural Language Processing Group. However, for Infotheca and other corpora that are part of the Bibliša digital library, the generated TMX file is sufficient. However, both steps were done in the case of *Infotheca*.

At the time of writting, the *INFOteka* parallel corpus included all parallelized issues from 2007 to 2025. According to data from Bibliša, the corpus contains 228 articles and 30,062 aligned segments (sentences). Although the total number of articles in the corpus is 228, only 207 are fully parallelized, as issue 2 (2019) and issue 2 (2021) were published only in English as special issues and could not be parallelized. However, metadata for these articles were entered in Bibliša in both languages, thereby enriching the INFOteka corpus in English, which is important for information retrieval in this scientific field.

5.2 *INFOteka* in Bibliša

Bibliša is a web tool developed by the Belgrade Human Language Technology Group aimed at enhancing search capabilities in digital libraries of e-journals (Stanković et al. 2012). Over time, parallel corpora have been incorporated into Bibliša, such as the Serbian-German Literary Parallel Corpus – SrpNemKor (Андоновски 2019, 2021),[32] and the Serbian-Italian Literary Parallel Corpus ItSrKor[33] (Moderc et al. 2023), among others. The main feature of Bibliša is that it enables users to compose queries using both simple and multiword keywords in more than one language, which can be expanded semantically and morphologically using different supporting monolingual and multilingual lexical and terminological resources (Stanković et al. 2012; Stanković, Obradović, and Trtovac 2012; Stanković et al. 2015; Stanković et al. 2016).

Each digital object in Bibliša is described using appropriate metadata. Its structure was originally prepared to describe a journal, its individual issues, and each article individually (Stanković et al. 2016, 1711), but over time

31. IMS OCWB is free, open source software developed for working with text corpora.
32. SrpNemKor
33. ItSrKor

the structure has been changed, so that other types of digital objects can be described, such as those objects in parallel literary corpora (Андоновски 2019, 2021). When loading journal articles, it is first necessary to describe the issue with the following metadata: identification number (ID), volume, issue, month, and year of publication. In this way, a collection of a certain issue of the journal is created, into which articles are then uploaded and described using the following metadata, in both Serbian and English: article identification number (ID), UDC number, author(s) name(s) (if there is more than one author they are listed in the order they appear in the paper), contact information (email), affiliation, Wikidata reference, article title, category, pages, abstract, keywords, link to the full text, and, where applicable, information about the translator. After entering the metadata, the TMX file is uploaded; it must contain the article ID so that it can be successfully uploaded. After this step, users are allowed to view the metadata (Figure 2), the TMX format of the article (unregistered users can only see the first nine aligned segments) (Figure 3), a link to the article's full text (link to the PDF in OJS), a link to the Bag of Words[34], and a link to Wikidata (more in the next Section). In addition to individual article metadata, users can also see the whole list of articles that are part of the *INFOteka* corpus.[35]

5.3 *Infotheca* and Wikidata

Besides Bibliša, Infotheca has its own Wikipedia page[36], as well as Wikidata entries both for the journal itself, and for all articles that are part of the corpus. In this way, the visibility of the journal on the web, as well as of the Serbian language, is increased.

Wikidata is a multilingual open knowledge base developed by the Wikimedia Foundation with the aim of enabling the storage and retrieval of structured multilingual data from Wikipedia in an interoperable, machine-readable format. As Wikidata is based on the principles and guidelines of the Semantic Web, the database has records in the form of statements in RDF [37] triples: subject (item) – predicate (relation) – object (value) (Schreiber and Raimond 2014). The subject (item) can be any topic (person, object, place,

34. Bag of Words represents a set of words in one text with the calculated frequency of their occurrence in that text, regardless of their grammatical forms and word order in the sentence.

35. Corpus INFOteka

36. Инфотека

37. Resource Description Framework

Figure 2. Metadata view in Bibliša for an article published in *Infotheca*

concept). It has a unique, persistent identifier, which is assigned automatically when the record is created and cannot be changed later. Also, each QID is unique, which means that it can be assigned only once. A data item identifier is a positive integer with an upper-case Q as a prefix (QID) and may, in addition to being linked to a title and a description, have a number of aliases and statements (claims, expressions) representing its properties and values (Stanković and Davidovć 2021, 90). Predicate (relation) is a property of a subject and, like the subject, has a persistent identifier: a positive integer with an upper-case P as a prefix (PID). The object provides additional information about the subject and can be another item, a string, a URL, etc.

In the Wikidata base, *Infotheca* is an item with the QID "Q25460443"[38] (Figure 4) and has certain numbers of statements. Table 1 presents a comparative overview of the statements for the item Q25460443 in natural language (English), their annotation with ID identifiers in the Wikidata statement format.

As mentioned, all papers that are part of *INFOteka* are individual items in Wikidata with statements such as "published in (P1433) Infotheca: Jour-

38. Infotheca: Journal for Digital Humanities (Q25460443)

Figure 3. TMX view for an article published in *Infotheca*

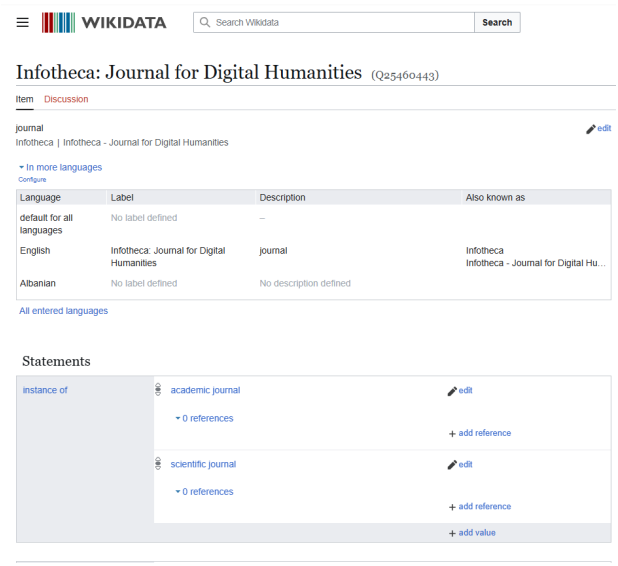

Figure 4. Item Q25460443 (Infotheca) in Wikidata

nal for Digital Humanities (Q25460443)" (Figure 5). Table 2 presents a comparative overview of selected statements for the item "The Language of Food and Its Dictionary" (Q122583889)[39] in natural language (English), their annotation with ID identifiers, and their Wikidata statement format.

It is possible to search metadata in Wikidata using the Wikidata Query Service environment[40] by composing SPARQL queries (Simple Protocol and RDF Query Language)[41]. There are numerous query examples that can be used by users who are not familiar with the SPARQL language, and query results can be visualized in various ways (Андоновски 2019, 144–149), (Andonovski 2023, 43–51), (Ikonić Nešić, Stanković, and Rujević 2023, 63–65). In 2025, the full Wikidata graph was split into two distinct graphs; the main graph and the scholarly graph. The scholarly graph includes all triples for items that contain (Lubiana, Rasberry, and Mietchen 2025):

39. The Language of Food and Its Dictionary
40. Wikidata Query Service
41. SPARQLis a protocol and query language for RDF databases

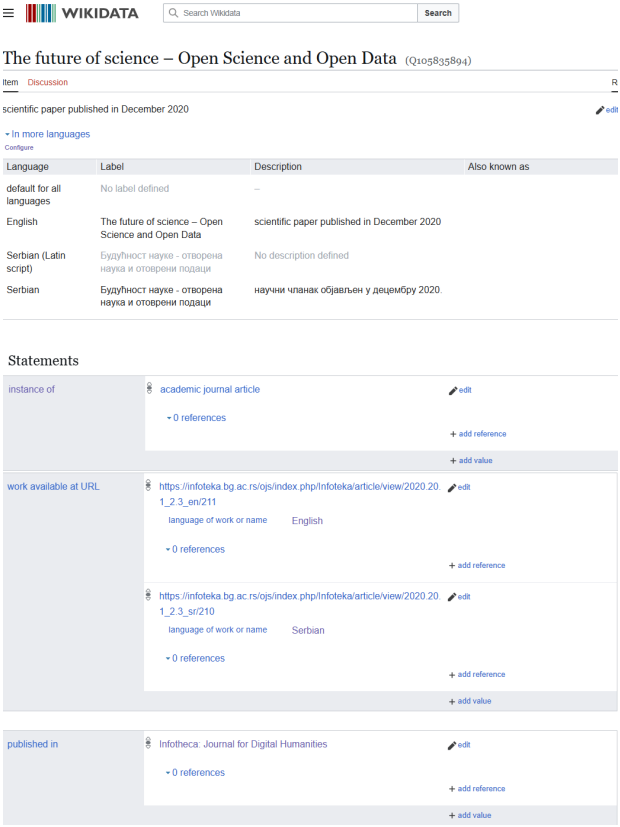

Figure 5. Property "published in (P1433)" with the value Q25460443 (Infotheca: Journal for Digital Humanities)

Infotheca: Journal for Digital Humanities instance of scientific journal, academic journal, text corpus. Its inception was in 2000. Publishers are: University of Belgrade, Faculty of Philology, University Library "Svetozar Marković" and the Serbian Academic Library Association. ISSN number is 1450-9687 subject named as Infoteka, distribution format printed matter and 2217-9461 subject named as Infoteka (Online), distribution format online publication.	
Infotheca: Journal for Digital Humanities (Q25460443) instance of (P31) scientific journal (Q5633421), academic journal (Q737498), text corpus (Q461183). inception (P571) „2000". publisher (P123): University of Belgrade Faculty of Philology (Q3542768), University Library "Svetozar Marković" (Q683389) and Serbian Academic Library Association (Q98821991). ISSN (P236) „1450-9687" subject named as (P1810) Infoteka, distribution format (P437) printed matter (Q1261026). ISSN (P236) „2217-9461" subject named as (P1810) Infoteka (Online), distribution format (P437) online publication (Q1714118).	Q25460443 P21 Q5633421; Q737498; Q461183; P123 Q3542768; Q683389; Q98821991; P236 „1450-9687" P1810 „Infoteka"; P437 Q1261026; P236 „2217-9461"; P1810 „Infoteka (Online)"; P437 Q1714118.

Table 1. Comparative representation of selected statements in natural language, using identifiers, and in the Wikidata statement format for item Q25460443.

1. statements with the property "instance of (P31)"[42] that matches any one of the 49 scholarly QIDs: for example, "scholarly article" (Q13442814) and 24 subclasses of it, "thesis" (Q1266946) and 15 subclasses of it, as well as "erratum" (Q1348305), "dissertation" (Q1385450), "comment" (Q58897583), "research report" (Q59387148), "field study report" (Q1402850), "conference poster" (Q54670950), "scientific note" (Q114613919) and, perhaps surprisingly, "Bachelor of Literature" (Q112585758), and
2. 2. statements with the property "publication type of scholarly work (P13046)".[43]

Accordingly, there are now two query services. For searching data in the main graph, the previously mentioned environment is used, while for searching data in the scholarly graph a separate query service is available at https://query-scholarly.wikidata.org/ (*Wikidata:SPARQL query service/WDQS graph split*). This query service is used for searching the data regarding *Infotheca*, because it has the statement "instance of (P31): academic

42. Wikidata: instance of (P31)
43. Wikidata: publication type of scholarly work (P13046)

The Language of Food and Its Dictionary instance of academic journal article. Author Duško Vitas. Publication date July 2023, published in Infotheca: Journal for Digital Humanities, language of work or name English and Serbian. DOI 10.18485/INFOTHECA.2023.23.1.1.	
The Language of Food and Its Dictionary (Q122583889) instance of (P31) academic journal article (Q18918145). Author (P50) Duško Vitas (Q12751371). Publication date (P577) „July 2023“. published in (P1433) Infotheca: Journal for Digital Humanities (Q25460443). language of work or name (P407) English (Q1860) and Serbian (Q9299). DOI (P356) 10.18485/INFOTHECA.2023.23.1.1.	Q122583889 P31 Q18918145; P50 Q12751371; P577 „јул 2023“; P1433 Q25460443; P407 Q1860; Q9299; P356 10.18485/INFOTHECA. 2023.23.1.1.

Table 2. Comparative representation of selected statements in natural language, using identifiers, and in the Wikidata statement format for item Q122583889.

journal (Q737498); scientific journal (Q5633421); text corpus (Q461183)", while articles have the statement "instance of (P31): academic journal article (Q18918145)".

We will illustrate the use of SPARQL queries with three examples. The first lists all papers from *Infotheca* published in the period 2018-2022; the results are presented as a timeline[44] and are sorted by the date of publication (Figure 6). The second one lists all papers published in *Infotheca* that have more than three authors; the results are presented as a tree structure and sorted by the number of authors decreasing[45] (Figure 7). The third query lists all authors who have published more than three papers in *Infotheca*. This query was formulated in two ways: the first one lists authors and the number of their published papers for the period 2007-2025, and the query results are presented as a table (Figure 8);[46] the second one lists all authors and the papers they have published, and the query results are presented as an interactive graph (Figure 9).[47]

Example 1:

```
SELECT DISTINCT ?item ?itemLabel ?volume ?issue ?publication_date
WHERE { # paper published (P1433) in Infotheca (Q25460443)
```

44. Visualization of the first SPARQL query in the form of a timeline, https://w.wiki/KcGc

45. Visualization of the second SPARQL query in the form of a tree, https://w.wiki/KcNQ

46. Visualization of the third SPARQL query in the form of a table, https://w.wiki/M5sN

47. Visualization of the third SPARQL query in the form of an interactive graph, https://w.wiki/KcU\protect\TU\textdollar

```
   ?item wdt:P1433 wd:Q25460443;
      # volume of publication (P478)
      wdt:P478 ?volume;
      # issue of publication (P433)
      wdt:P433 ?issue;
      #publication_date (P577)
      wdt:P577 ?publication_date.
FILTER (?publication_date >= "2018-06-01T00:00:00Z"^^ xsd:dateTime
&& ?publication_date <= "2022-06-01T00:00:00Z"^^ xsd:dateTime)
SERVICE wikibase:label
{bd:serviceParam wikibase:language "[AUTO_LANGUAGE],mul,en".} }
ORDER BY DESC (?publication_date)
```

Example 2:

```
SELECT DISTINCT ?item ?itemLabel(COUNT (?author) AS ?authorCount)
WHERE { #paper published (P1433) in Infotheca (Q25460443)
   ?item wdt:P1433 wd:Q25460443;
      # author (P50)
      wdt:P50 ?author.
SERVICE wikibase:label
{bd:serviceParam wikibase:language "[AUTO_LANGUAGE],mul,en".} }
GROUP BY ?item ?itemLabel
HAVING (COUNT(?author) > 3)
ORDER BY DESC (?authorCount)
```

Example 3a:

```
SELECT ?authorLabel(COUNT(DISTINCT ?item) AS ?publicationCount)
WHERE { # paper published (P1433) in Infotheca (Q25460443)
   ?item wdt:P1433 wd:Q25460443;
      # author (P50)
      ?item wdt:P50 ?author.
# Get the author label
SERVICE wikibase:label
{bd:serviceParam wikibase:language "[AUTO_LANGUAGE],mul,en".} }
GROUP BY ?author ?authorLabel
# Filter for more than 3 papers
HAVING (COUNT(DISTINCT ?item) > 3)
ORDER BY DESC (?publicationCount)
```

Example 3b:

```
SELECT ?author ?authorLabel ?item ?itemLabel
WHERE { #Find publications in Infotheca (Q25460443)
   ?item wdt:P1433 wd:Q25460443;
      # author (P50)
      wdt:P50 ?author.
# Filter for authors with more than 3 publications in this journal
{ SELECT ?author (COUNT(?item) AS ?publicationCount)
WHERE { # Find publications in Infotheca (Q25460443)
   ?item wdt:P1433 wd:Q25460443;
      # author (P50)
      wdt:P50 ?author.
}
GROUP BY ?author
```

```
HAVING (COUNT(?item) > 3)
}
      # Labels
SERVICE wikibase:label
{bd:serviceParam wikibase:language "[AUTO_LANGUAGE],mul,en".} }
ORDER BY DESC (?publicationCount)
```

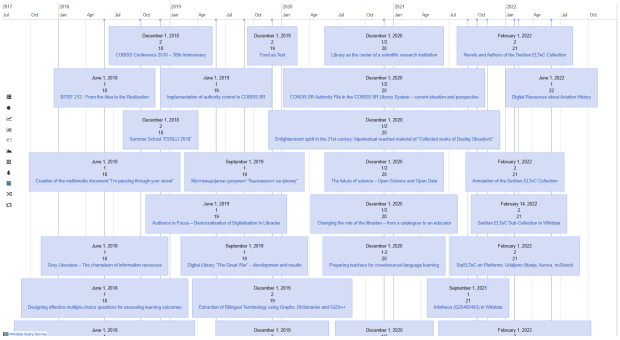

Figure 6. Visualization of the SPARQL query in the form of a time scale

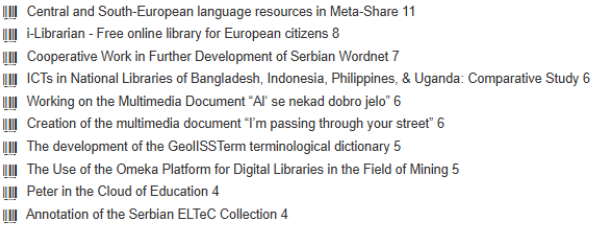

Figure 7. Visualization of the SPARQL query in the form of a tree

6 Conclusion

The paper presents all changes of the journal *Infotheca*, from its inception to the present day, and its transformation from a traditional resource into

author	publicationCount
wd:Q99281502	17
wd:Q97199587	12
wd:Q99281834	11
wd:Q12751371	8
wd:Q99281544	6
wd:Q99281474	5
wd:Q99368929	5
wd:Q99281736	4
wd:Q99463989	4

Figure 8. Visualization of the SPARQL query in the form of a table

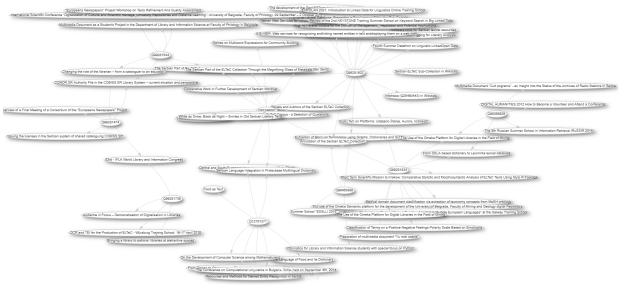

Figure 9. Visualization of the SPARQL query in the form of an interactive graph

a fully online Open Access publication. The implementation of the Open Journal System (OJS) has enabled the editorial team and all collaborators (authors, reviewers, copyeditors, proofreaders, and technical editors) to participate in one place in the preparation of *Infotheca* issues. In addition, by using the LaTeX package with templates for page layout design prepared according to the document class used by Springer, authors have the possibility to submit their texts already prepared in the publishing format. For this purpose, Overleaf is used as a free collaborative LaTeX editor, allowing submitted papers to be simultaneously reviewed and technically edited by *Infotheca's* technical editors with granted access. With the implementation of OJS and Overleaf, all activities related to issue preparation can be carried out without the exchange of a large number of emails, which is a great technical improvement of the journal.

To facilitate faster, easier, and simpler information retrieval within *Infotheca*, the journal's archive has been integrated into different databases. The fact that the papers are published bilingually and in Open Access, as well as the fact that Prof. Cvetana Krstev has been the editor-in-chief for more than a decade, were the main characteristics that influenced the creation of the bilingual parallel corpus *INFOteka*, which gave the journal a different form. The fact that it is also a domain-specific corpus has had an additional influence on it becoming a digital language resource very important in research across different scientific fields, including information and terminology retrieval and extraction from domain corpora, and computational linguistics. Furthermore, by integrating the corpus into Bibliša, a wider circle of researchers has been enabled to search the corpus content using metadata search or full-text search composing queries in more than one language that can be expanded, both semantically and morphologically, using different supporting monolingual and multilingual lexical and terminological resources. Users can also view the search results in various ways.

On the other hand, the creation of Wikidata entries for all papers in the parallel corpus, which can also be accessed via Bibliša, has enabled a different way of information retrieval regarding *Infotheca*, as well as alternative presentation of the obtained results. Unlike Bibliša, which also supports full-text search, Wikidata focuses on metadata, i.e., structural metadata. Therefore, information retrieval regarding *Infotheca* in this database is of exceptional importance, as it enables researchers to formulate complex queries and get results that cannot be obtained from Bibliša or COBISS. In addition, as Wikidata is part of the Linked Open Data (LOD) cloud, the journal, and also the Serbian language, are integrated into this huge ecosystem on the

web. Its presence in Wikipedia further increases both the journal's visibility, and the visibility of the Serbian language.

Infotheca will continue to publish scientific results and professional expertise in the field of digital humanities in the Republic of Serbia and abroad, as well as reviews of significant events and publications in this field, striving to achieve the highest possible quality. In the upcoming period, the development of the journal will procede in several directions. Planned activities include bibliometric and citation analysis, the completion of the parallel corpus by adding issues from 2000 to 2006 so as to complete the archive, and, in addition to regular analytical processing, retrospective processing, which will include assigning DOI to works published prior to 2016.

In 2025, the 25th volume was published, marking a quarter century of the journal's existence, as well as an important period in the work of the University Library "Svetozar Marković". The Library has played a crucial role in the continuity of *Infotheca* since its inception. In addition to being one of the publishers, it also serves as the headquarters of the SALA. Particularly important is the engagement of the librarians and IT specialists employed at the Library, who, through their work and editorial activities in *Infotheca*, contribute to its sustainability and quality. The work of all librarians, including the work of Serbian and English proofreaders employed at the Library, is in the scope of their everyday professional duties, while external collaborators involved in journal editing, work on a voluntary basis. This demonstrates the great enthusiasm and commitment of all collaborators to maintain such a journal within a small yet significant scientific community.

References

Andonovski, Jelena. 2023. "SrpNemKor and Linked Open Data." *Infotheca – Journal for Digital Humanities* 23 (1): 31–58. https://doi.org/https://doi.org/10.18485/infotheca.2023.23.1.2.

Ikonić Nešić, Milica, Ranka Stanković, and Biljana Rujević. 2023. "Serbian ELTeC Sub-Collection in Wikidata." Preuzeto 18. 2. 2026, https://doi.org/10.18485/infotheca.2021.21.2.4, *Infotheca –- Journal for Digital Humanities* 23 (2): 60–87.

Kovačević, Ljiljana, Vesna Injac, and Dobrila Begenišić. 2004. *Bibliotekarski terminološki rečnik : englesko-srpski, srpsko-engleski.* Beograd : Narodna biblioteka Srbije.

Krstev, Cvetana. 2008. *Processing of Serbian – Automata, Texts and Electronic dictionaries.* Belgrade: Faculty of Philology, University of Belgrade.

Krstev, Cvetana, and Duško Vitas. 2009. "An Effective Method for Developing a Comprehensive Morphological e-Dictionary of Compounds." In *Proceedings of Lexis and Grammar Conference, Bergen,* 204–212. Preuzeto 19. 2. 2026, https://poincare.matf.bg.ac.rs/~cvetana/biblio/Krstev-Vitas-LGC09.pdf.

Krstev, Cvetana, and Duško Vitas. 2011. "An Aligned English-Serbian Corpus." In *ELLSIIR Proceedings (English Language and Literature Studies: Image, Identity, Reality), Volume I, Belgrade, 4-6 December 2009,* edited by Nenad Tomović and Jelena Vujić, 495–508. Preuzeto 12. 2. 2026, https://poincare.matf.bg.ac.rs/čvetana/biblio/AlignedCorpus-full-final.pdf. Belgrade: Faculty of Philology, University of Belgrade.

Lubiana, Tiago, Lane Rasberry, and Daniel Mietchen. 2025. "The Wikidata Query Service Split and its Impact on the Scholarly Graph." In *SEMANTiCS Conference 2025, September 03–05, 2025, Vienna, Austria.* Preuzeto 30. 3. 2026, https://ceur-ws.org/Vol-4064/PD-paper3.pdf.

Milinković, Milena. 2016. "The Bibliometric and Citation Analyses of the SPATIUM Journal." Преузето 01. 4. 2026, https://infoteka.bg.ac.rs/ojs/index.php/Infoteka/article/view/2016.16.1_2.4_en, *Infotheca – Journal for Digital Humanities* 16 (1–2): 72–98.

Moderc, Saša, Ranka Stanković, Aleksandra Tomašević, and Mihailo Škorić. 2023. "An Italian-Serbian sentence aligned parallel literary corpus." Preuzeto 20. 3. 2026, =http://elib.mi.sanu.ac.rs/files/journals/ncd/43/ncdn43p78-91.pdf, *Review of the NCD,* no. 43, 78–91.

Obradović, Ivan, Ranka Stanković, and Miloš Utvić. 2008. "Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen." Chap. Integrisano okruženje za pripremu paralelizovanog korpusa, 563–578. Preuzeto 13. 2. 2026, https://rgf.bg.ac.rs/LicnePrezentacije/ivan_obradovic/Radovi/GRAZ_2007a.pdf. Münster: LitVerlag.

Schreiber, Guus, and Yves Raimond, eds. 2014. *RDF 1.1 Primer. W3C Note.* Preuzeto 18. 2. 2026, https://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/.

Stanković, Ranka. 2026. "Textometric and comparative analysis of the bilingual corpus of the journal *Infotheca*." *Infotheca – Journal for Digital Humanities* 26 (1). https://doi.org/https://doi.org/10.18485/infotheca.2026.26.1.2.

Stanković, Ranka, and Lazar Davidovć. 2021. "Infotheca (Q25460443) in Wikidata." *Infotheca – Journal for Digital Humanities* 21 (1): 87–98. https://doi.org/https://doi.org/10.18485/infotheca.2021.21.1.5.

Stanković, Ranka, Cvetana Krstev, Biljana Lazić, and Dalibor Vorkapić. 2015. "A bilingual digital library for academic and entrepreneurial knowledge management." In *Proceeding of 10th International Forum on Knowledge Asset Dynamics — IFKAD 2015: Culture, Innovation and Entrepreneurship: connecting the knowledge dots, Bari, Italy, 10-12 June 2015,* edited by JC Spender, Giovanni Schiuma, and Vito Albino, 1764–1777. Preuzeto 9. 2. 2026, https://poincare.matf.bg.ac.rs/~cvetana/biblio/IFKAD_RS_CK_BL_DV-fin.pdf.

Stanković, Ranka, Cvetana Krstev, Ivan Obradović, Aleksandra Trtovac, and Miloš Utvić. 2012. "A Tool for Enhanced Search of Multilingual Digital Libraries of E-journals." In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12),* edited by Nicoletta et al. Calzolari, 1710–1717. Preuzeto 17. 2. 2026, =http://www.lrec-conf.org/proceedings/lrec2012/pdf/375_Paper.pdf. Istanbul: ELRA.

Stanković, Ranka, Cvetana Krstev, Duško Vitas, Nikola Vulović, and Olivera Kitanović. 2016. "Keyword-Based Search on Bilingual Digital Libraries." In *Semantic Keyword-Based Search on Structured Data Sources - Second COST Action IC1302 International KEYSTONE Conference, IKC 2016, Cluj-Napoca, Romania, September 8-9, 2016. Revised 285 Selected Papers,* edited by A. Calì, D. Gorgan, and M. Ugarte, 112–123. Preuzeto 9. 2. 2026, https://poincare.matf.bg.ac.rs/~cvetana/biblio/IFKAD_RS_CK_BL_DV-fin.pdf. Springer.

Stanković, Ranka, Ivan Obradović, and Aleksandra Trtovac. 2012. "An Approach to Development of Bilingual Lexical Resources." In *Fifth Balkan Conference in Informatics BCI 2012, Novi Sad, Serbia, September 16-20, 2012,* edited by Zoran Budimac, Mirjana Ivanović, and Miloš Radovanović, 101–104. Preuzeto 17. 2. 2026, https://dr.rgf.bg.ac.rs/files/original/16b86b4ac88bd56a73332f45cac634abb00e6379.pdf. Novi Sad: Faculty of Sciences, Department of Mathematics / Informatics.

The Digital Humanities Manifesto 2.0. 2026. Preuzeto 10. 2. 2026, https://othone.org/Manifesto_V2.pdf.

Utvić, Miloš. 2013. "Izgradnja referentnog korpusa savremenog srpskog jezika: doktorska disertacija." Preuzeto 14. 2. 2026, http://phaidrabg.bg.ac.rs/o:10061. PhD diss., Univerzitet u Beogradu, Filološki fakultet.

Vitas, Duško, and Cvetana Krstev. 2004. "A lexical approach to text alignment using Intex." In *Intex pour la linguistique et le traitement automatique des langues,* edited by Claude Muller, Jean Royaute, and Max Silberztein, 249–263. Preuzeto 12. 2. 2026, https://poincare.matf.bg.ac.rs/čvetana/biblio/krstev-vitas.pdf. Besancon: Presses Universitaires de Franche Comte.

Vitas, Duško, and Cvetana Krstev. 2006. "Literature and Aligned Texts." In *Readings in Multilinguality,* edited by Milena Slavcheva, Galia Angelova, and Kiril Simov, 148–155. Preuzeto 12. 2. 2026, https://poincare.matf.bg.ac.rs/čvetana/biblio/CvDv-Paskaleva.pdf. Sofia: Institute for Parallel Processing, Bulgarian Academy of Sciences.

Wikidata:SPARQL query service/WDQS graph split. Preuzeto 20. 2. 2026, https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/WDQS_graph_split.

Андоновски, Јелена. 2019. "Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса." Преузето 12. 2. 2026, https://phaidrabg.bg.ac.rs/view/o:22874. PhD diss., Докторска дисертација. Универзитет у Београду, Филолошки факултет.

Андоновски, Јелена. 2021. *Паралелни српско-немачки корпус књижевних текстова: израда, проналажење информација и семантички веб.* Београд: Универзитетска библиотека „Светозар Марковић".

Вранеш, Александра. 2014. "Дигитална хуманистика и савремене библиотеке." Преузето 10. 2. 2026, https://infoteka.bg.ac.rs/ojs/index.php/Infoteka/article/view/2014.15.1.1_sr, *Инфотека – часопис за дигиталну хуманистику* 15 (1): 4–15.

Рајхерт, Рамон. 2014. "Дигитална хуманистика. Увод у историју, теорију и развој културе знања засноване на подацима." Преузето 10. 2. 2026, https://infoteka.bg.ac.rs/ojs/index.php/Infoteka/article/view/335, *Инфотека – часопис за дигиталну хуманистику* 15 (2): 22–34.

Тртовац, Александра. 2016. "Дескриптори метаподатака и дескриптори садржаја у проналажењу информација у дигиталним библиотекама." Преузето 10. 2. 2026, https://phaidrabg.bg.ac.rs/o:12605. PhD diss., Универзитет у Београду, Филолошки факултет.

Тртовац, Александра. 2017. *Проналажење информација у дигиталним библиотекама.* Београд: Универзитетска библиотека „Светозар Марковић".