

The Role and Challenges of the Digitalization Process in the Preservation and Development of Scientific Schools

Olga Ataeva¹[0000-0003-0367-5575], Vladimir Serebryakov²[0000-0003-1423-621X],
Natalia Tuchkova³[0000-0001-6518-5817]

¹⁻³ Dorodnicyn Computing Center FRC CSC of Russian Academy of Sciences, Vavilova str.,
40, 119333, Moscow, Russia

¹oli@ultimeta.ru, ²serebr@ultimeta.ru, ³natalia_tuchkova@mail.ru

Abstract. The problem of preserving knowledge in the digital space arose in the middle of the last century and is gradually aggravated with the increase in the flow of digital information. The scientific community is faced with new challenges related to the speed of dissemination of information and scientific ideas. Classical scientific schools, which were the basis of university education, have undergone changes in connection with distance learning and the idea of openness of science as part of the common values of society. Against the background of increasing information noise, the search for reliable data is becoming an independent activity. The period of accumulation of digitized information is being replaced by the accumulation of knowledge in an integrated form, in the form of structured data. However, the increase in the number of information resources does not improve the quality of the information provided. One of the tasks of the scientific community is to provide an expert assessment of the accumulated digital knowledge. This assessment should be based on the authority of scientific schools. The paper discusses the problem of supporting information images of traditional scientific schools and their development in the digital space. On the example of the integration of mathematical knowledge, the tasks and ways of their solution in the process of assessing and preserving digitized knowledge are formulated. A variant of specialized author indexes and thesauri of mathematical schools is proposed. The content of the digital library LibMeta and the Russian Mathematical Encyclopedia edited by I.M. Vinogradov.

Keywords¹: authors index, data integration tools, mathematical subject area thesaurus, search query.

¹ CDSSK–2020: International Conference “Common Digital Space of Scientific Knowledge”,
November 10–12, 2020, Moscow, Russia

EMAIL: oli@ultimeta.ru (Olga Ataeva); serebr@ultimeta.ru (Vladimir Serebriakov);

natalia_tuchkova@mail.ru (Natalia Tuchkova)

ORCID: 0000-0003-0367-5575 (Olga Ataeva); 0000-0003-1423-621X (Vladimir Serebriakov);

0000-0001-6518-5817 (Natalia Tuchkova)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons
License Attribution 4.0 International (CC BY 4.0)



CEUR Workshop Proceedings (CEUR-WS.org)

The period of digitalization, which determined the development of the economy, science and culture at the beginning of the 21st century, naturally affected the acquisition and perception of knowledge. Digitalization has intensified the process of “discovery”, “decolonization of knowledge” (terms recommended by UNESCO²). The idea of openness is based on the assertion of the scientific community that science is part of the overarching principles and shared values of humankind. Scientific texts began to open up to all users thanks to the open access journal. This posed a challenge for many authors accustomed to protecting scientific texts as a means of ensuring priority and authorship.

The integration of data and a scientometric approach to assessing the results in the accumulation of knowledge also generated difficulties in determining the boundaries of the belonging of works to scientific teams. The classical scientific schools, which were the basis of university education, have also undergone changes, especially in connection with distance learning and the idea of openness of knowledge. Distance education slowly entered the standard until an urgent need arose due to the COVID-19 pandemic. At the same time, it became clear that live communication cannot be ruled out in the process of transferring knowledge from person to person, and lectures have moved to the space of remote communication through well-known applications (Zoom, Microsoft Teams, etc.).

Access to paper textbooks has been reduced to a minimum, and in these conditions it is especially important to preserve the knowledge of university and academic scientific schools in an accessible modern digital form, that is, as structured data [1]. This fully applies to the preservation of the concept of authority in science, since the modern perception of knowledge is based more on clip thinking, perception of facts, and not on the consistent construction of the logic of knowledge [2].

But without defining the place of a scientific discovery in the overall picture of scientific knowledge and the scientist or team of scientists behind it, it is impossible to fully reflect the meaning and significance of this or that scientific fact. To provide such an opportunity, it is necessary to use data integration tools and tools for their expert assessment by the scientific community. In this process, it is necessary to rely on classical and newly formed scientific schools as authorities known in their field.

It is expert assessments that can serve as guidelines when searching for information during the period of training and research. However, there is a significant gap in the use of expert judgment in information resources.

This paper considers the problem of identifying scientific schools based on the content of a digital library and proposes a technology for establishing semantic links between scientific publications, as a sign of integration or separation of scientific schools. The aim of a study is to propose a way to designate the belonging of research to a particular school in order to preserve authorship and priorities in science. Naturally, this data analysis cannot be carried out without artificial intelligence methods and machine learning algorithms. Experience with the content of the LibMeta library [3,

² https://en.unesco.org/sites/default/files/open_science_brochure_en.pdf

4], including the digitized version of the encyclopedia in Russian [5], allows using vector and other learning algorithms to solve this problem.

1 The value of the experience of scientific schools in the digital space of knowledge

1.1 On the development of domestic mathematical schools

It is known that since the time of Leonhard Euler [6], the relations between domestic and foreign scientists, especially in the field of mathematics, have undergone different periods. The history of the development of science shows that in the 18 and 19 centuries Russian and European scientists continued to communicate and compete in scientific discoveries. There were periods when the language of science was German, then French.

At the same time, Russia developed its own mathematical schools, which can be read about in a wonderful short survey by L.D. Faddeev and I.A. Lavrov [7], including the period of the 20th century. Continuing to cite this work [7], we recall that the Moscow Mathematical Society (<http://mms.mathnet.ru/>) was created in 1864. This happened earlier than London, 1865 (<https://www.lms.ac.uk/>) or the USA in 1888 (<https://www.ams.org/>).

1.2 The impact of digitalization on the integration, disunity and structuring of knowledge

The digitalization period is associated with the creation of integrated bibliographic resources and databases. However, with the advent of digital publications and scientific bibliographic databases, English acquired the status of a scientific language. One of the oldest mathematical journals in the world, founded by the Moscow Mathematical Society in 1866, “Sbornik: Mathematics” (<https://iopscience.iop.org/journal/1064-5616>), has been published since 1995 in collaboration with the London Mathematical Society in English.

In particular, a huge number of electronic journals have appeared, most of which are not included in international citation bases, which leads to their isolation and lowers the ratings of these publications. It is significant that even well-known academic Russian publications from the series “Reports of the Academy of Sciences” periodically dropped out of citation in WoS and Scopus, and now individual recommended English-language versions of articles published by Springer Nature are taken into account (<https://www.springer.com/journal/11472>). The issues of integration and disunity of scientific schools became more obvious with the advent of digital resources and the development of methods for analyzing these data [8]. The monograph [8] describes examples of faith in science as a whole by scientists of the 19th century and split at the end of the 20th century. The philosophical substantiation of this phenomenon is given. In general, these periods are associated both with the development

of various scientific schools and standards, and with the idea of “main” and “non-main” sciences.

The monograph [9] analyzes the approach to the development of “STEM-sciences” (STEM – stands for science, technology, engineering and mathematics and refers to any subjects that fall under these four disciplines) and “others”. The author draws attention to the processes of isolation taking place in science and society, to the need to consider all scientific areas as equally important and calls on research teams to “effective cooperation across the borders that usually divide people and scientists”. All these features, naturally, were reflected in the development of domestic mathematical sciences and scientific schools. That is, against the background of absorption, characteristic of the integration of digital resources, individual research teams and schools strive to preserve original copyright research, priorities and scientific publications.

When searching for reliable data, the use of personal research experience, relevant accumulated bibliographic collections and research results that are actively disseminated on the Internet comes to the fore. Expert assessments can serve as criteria when searching for information. The creation and support of digital images of scientific schools, as landmarks of reliable knowledge, is an important task of modern information technologies.

One of the features of scientific schools is the prevailing “common language”. On this terminological basis, it is possible to form thesauri, characteristic of scientific schools. Author’s collections, bibliographic indexes reflect the affiliation of subject domain (SbD) and a particular scientific school. The author of a unique project for the creation of dictionaries and thesauri in mathematics, Professor of the University of Toronto Kenneth O. May, owns the thesis: “You cannot do scientific work, and systematically work in any area without dictionaries in this area” [10]. This was said when it was not possible to analyze digital author's collections, but it was already clear that the terminological description of the SbD is an important stage in the structuring of knowledge. SbD terminological description, creation of dictionaries and thesauri are available only thanks to the expert knowledge of professionals [10].

Since the search and analysis of data in Internet resources is becoming an independent type of activity [11], the opinion of recognized authorities of scientific schools is of particular importance in the flow of digital information. Scientific schools are landmarks in the digital space, their value is the examination of SbD information images.

2 Integration of mathematical knowledge in a digital environment and examples of implementation

2.1 Features of the integration of scientific subject areas

The purpose of integrating mathematical knowledge in information systems is to perform the main function of the utility of information systems, namely, to provide the

most complete response to the user's information request and to meet the user's information needs [12]. This is fully in line with the UNESCO recommendations.

Information technology does not stop moving along the path of providing search resources for scientific and educational purposes. Some online libraries (for example, <https://cyberleninka.ru/>) and resources (for example, <https://www.researchgate.net/>, <https://www.semanticscholar.org/>) have already begun to offer a service for the selection of relevant works on topics similar to the request. Search engines accumulate samples, compare queries, and return search results based on previous queries. This is justified, since only the specialist himself, the researcher, can adequately formulate his information request, and it is necessary to provide him with the opportunity to see the results of queries related to the topic. For mathematical works, this is especially important, since they use the formula language when searching, as the simplest way to express the specifics of SbD, and you can get results for similar symbolic expressions. Let us note the distinctive features of mathematical publications, such as the presence of formulas and the specifics of presentation (the presence of statements, theorems, proofs, etc.). During preliminary analytical and synthetic processing of mathematical texts, in addition to creating dictionaries, thesauri and sets of keywords for indexing, it is necessary to compile lists of formulas and relationships between them, contexts where they occur. The connections between statements and formulas are dictated by the logic of the mathematical SbD itself.

Research towards the creation of integrated resources of mathematical knowledge consists of several sections, such as the development and creation of:

- methods and technologies for representing mathematical knowledge in digital libraries as structured data;
- technologies for using symbolic information in search queries and for transferring data;
- technologies for representing formulas in computer algebra systems;
- evidence systems;
- platforms to support users in the study and use of knowledge of mathematical SbD;
- technologies of searching for links between mathematical publications to identify the mutual influence of research of various research teams.

All of the above can be summarized as the creation of systems for supporting actual mathematical knowledge in the scientific digital space and managing this data based on SbD semantic links. For the integration of information images of sections of mathematics, it is necessary to take into account the listed aspects. Namely, to use representative dictionaries with the interpretation of terms, formulas and connections between them as the basis of the SbD information retrieval thesaurus and the inclusion of author indexes in them to reflect the role of individual authors and scientific schools in SbD ontologies [12, 13]. In Fig. 1 shows an example (for Russian language works) of a nominal terminological dictionary-index LibMeta, where you can trace the links: Author \rightarrow Terms; Term \rightarrow Authors.

View object

В П Радченко

Type [Person](#)

URI <http://libmeta.ru/mathnet/person/47>

Related objects

Author (publication)	Получение аналитических решений нелинейных задач теплопроводности на основе введения дополнительных граничных условий
Author (publication)	Памяти Олега Александровича Репина
Author (publication)	Моделирование ползучести и релаксации остаточных напряжений в поверхностно упрочненных элементах статически неопределимых стержневых систем

[View full list](#)

Linked keywords

[нелинейные задачи](#) [аналитические методы](#) [фронт температурного возмущения](#) [дополнительные граничные условия](#) [интеграл теплового баланса](#) [трехэлементная статически неопределимая система](#) [поверхностное упрочнение](#) [остаточные напряжения](#) [ползучесть](#) [релаксация напряжений](#) [структурные модели](#) [реологические модели](#) [вязкоупругость](#) [ползучесть](#) [дробное исчисление](#)

[View full list](#)

Fig. 1. Nominal terminological dictionary-index

Figure 1 shows how, based on the index, you can get lists of keywords for further analysis of the relationship between authors and subject areas. The author's keywords are formed on the basis of the relationship “is the author”, fixed in the ontology of the subject area, according to the rule “if the person is the author of the publication, then the keywords of the publication are included in the set of the author's keywords.”

Within the framework of this work, the problem of analyzing the content associated with both an individual author and the entire content of the library as a whole is being solved. The semantic library produces: preliminary processing of the author's texts, to bring them to a form suitable for semantic analysis; highlighting the author's area of interest based on the available data on publications and connections; formation of the thesaurus of the addressee of the author based on the links of the SbD thesaurus; analysis of the network of co-authors and bibliographic lists. Based on the results of the implementation of these stages, it is possible to form lists of scientists from SbD and evaluate their contribution to the development of SbD within the framework of the digital library content.

Remark. The concept of scientific schools in the usual representation in the real world is broader, and we do not pretend to replace recognized authorities in scientific fields, based on the accumulated content of the library. The paper considers only the available limited digital content, on the basis of which the corresponding conclusions are drawn. Obviously, the presence of recognized authorities with a sufficiently large amount of data should manifest itself in the analysis of the digital content of any SbD and will correlate with the situation in the real scientific world.

2.2 Examples of representing mathematical knowledge in a digital environment

The work [14] presented the SbD thesaurus “ordinary differential equations” and continues to develop its extension to the area of “partial differential equations” as part of the general mathematical resource on “equations of mathematical physics and related areas” [15–18]. The test implementation is carried out on the basis of the LibMeta digital library [3, 4]. To accumulate personal scientific knowledge within the framework of mathematical SbD, the concept of the addressee's thesaurus is used [19]. It should be noted that the theoretical developments of domestic scientists in the field of processing and presentation of digital SbD information traditionally outstrip their practical implementation and some foreign developments. However, the lack of integrated data from domestic publications often does not allow the implementation of these developments in full.

Some foreign and domestic resources related to the topic under study can be noted: Digital Library of Mathematical Functions (dlmf.nist.gov); EMIS (www.emis.de); MacTutor History of Mathematics archive (turnbull.mcs.st-and.ac.uk); MathDiss database (mathdiss.mathguide.de); SciGuide (<http://www.prometeus.nsc.ru/sciguide/>); NIST Digital Library of Mathematical Functions (<https://dlmf.nist.gov/>); the European Mathematical Information Service (<https://www.emis.de/>); ELibM – Electronic Library of Mathematics; Zentralblatt MATH (<https://zbmath.org>); All-Russian portal Math-Net.Ru (http://www.mathnet.ru/index.phtml?&option_lang=eng); EqWorld (<http://eqworld.ipmnet.ru/indexr.htm>); OntoMathPro [20, 21].

Zentralblatt MATH demonstrates on its bibliographic database the advantages of using symbolic information, formulas. ResearchGate (<https://www.researchgate.net/>) – integrator of scientific bibliography based on author profiles.

<https://zbmath.org>, – is the closest to the claimed approach, and contains, in particular, a search by formulas in the annotations to publications. The filling technique proposed in this approach, as in ResearchGate, assumes the use of open source sources.

Among all the others, the Wolfram Research portal (<http://www.wolfram.com>), rich in computational methods and equipped with its own mathematical programming language, stands out.

The main domestic (Russian) resource is the Encyclopedia of Mathematics [5] https://www.encyclopediaofmath.org/index.php/Main_Page – a digitized version of the encyclopedia edited by I.M. Vinogradov, which, unfortunately, exists separately from other information resources, both in Russian and in English.

It should be noted that now this encyclopedia is available from the digital library LibMeta [3]. Is this made possible by the data structure of this library as an integrated system, where the connection of mathematical concepts with sources is realized. One such source is the Encyclopedia of Mathematics.

Many well-known groups in the scientific community are engaged in the creation of information systems, one way or another related to the representation of mathematical knowledge, research and publications in the digital space (for example, [22–24]).

This is determined by the demand for mathematical knowledge, competencies and research results in many areas of science, education and business.

However, there is no overarching integrated mathematical resource even in perspective. Moreover, the strongest competition generated by the economic component associated with the ratings of scientists leads to the disunity of scientific schools.

The work [25] presents the history of the creation of information resources of the department of the State Public Scientific Technical Library of the SB RAS. It examines the path of development of databases on eminent scientists and scientific schools headed by them and their representation in digital form.

2.3 Notation and formulas as a reflection of research approaches

It is known that different scientists in different parts of the world conducted similar mathematical research, created their own schools and recorded the results in different ways. The loudest story about the dispute between Newton and Leibniz on the contribution of each to the development of mathematical analysis [26].

The work [14] provides an overview of sources devoted to the main research trends in the use of indexing of publications based on taxonomy and the inclusion of formulas in a search query. Since the representation of formulas is still used as the main language of mathematics, the analysis of a text without taking into account the representation of formulas is not enough. And the presentation of formulas in the context of a scientific school is the main characteristic of its language. Therefore, experts separate the problems of representing formulas for storage, visualization from the problems of finding publications with formulas, and the presentation of formulas plays a decisive role.

Methods of coding and analysis of symbolic expressions, indexing of mathematical articles with keywords and formulas are used to save data. Contextual search systems are being developed taking into account taxonomy. These developments constitute a direction that combines the use of semantic links and recursive search of indexed documents with formulas. This approach involves the creation of a preliminary data structure and markup of documents for fast search and further recursive refinement of the query after prefetching.

The digital semantic library LibMeta develops on these principles. It implements the thesaurus for ordinary differential equations, which is built into the SbD ontology and is associated with publications and authors (Fig. 2).

The library contains a module that allows you to search by formulas, taking into account their semantics. To support searching, the *Formula* class has been modeled, which allows you to store the original formula string from the source where it was obtained. The string can be in Content MathML, Presentation MathML, LaTeX format. If necessary, the number of types for representing formulas in different notations can be easily expanded. The *Formula class* has relationships with the information objects that make up the library content and with the thesaurus concepts. Thus, it is possible to build a network of formula links with the various objects that make up the library content. Each formula can be supplemented with keywords.

The open-source editor of mathematical formulas VisualMathEditor by David Grima [27, 28] was used directly to implement the search interface. This editor allows the user to type a formula without knowing the intricacies of working with LaTeX or MathML.

Figure 2 shows a part of the hierarchical and horizontal links of the thesaurus between the terms on the basis of which the rules are formed linking different resources of the system with the terms of the subject area from the thesaurus, for example, “if a person is the author of the publication, then the terms of the thesaurus found in his publications are included in the set of the author's terms”.

View thesaurus links

Thesaurus of ordinary differential equations

- Unperturbed motion**
association > [DE0136](#) [DE0142](#) [DE0137](#) [DE0174](#)
synonymy > [RDE0139](#) [RDE0140](#) [RDE0143](#)
- Unperturbed solution**
association > [DE0136](#) [DE0142](#) [RDE0110](#) [DE0137](#) [DE0174](#)
synonymy > [RDE0140](#) [RDE0143](#)
- Phase trajectory**
association > [DE0136](#) [RDE0104](#)
- Autonomous ODE System**
association > [DE0136](#) [DE0103](#) [DE0128](#)
synonymy > [DE0142](#) [RDE0110](#) [RDE0135](#) [RDE0139](#) [RDE0140](#) [RDE0143](#) [RDE0104](#) [DE0174](#)

Fig. 2. Presenting a thesaurus in a digital library

Search by formulas involves searching by the structure of a formula and using its semantic links.

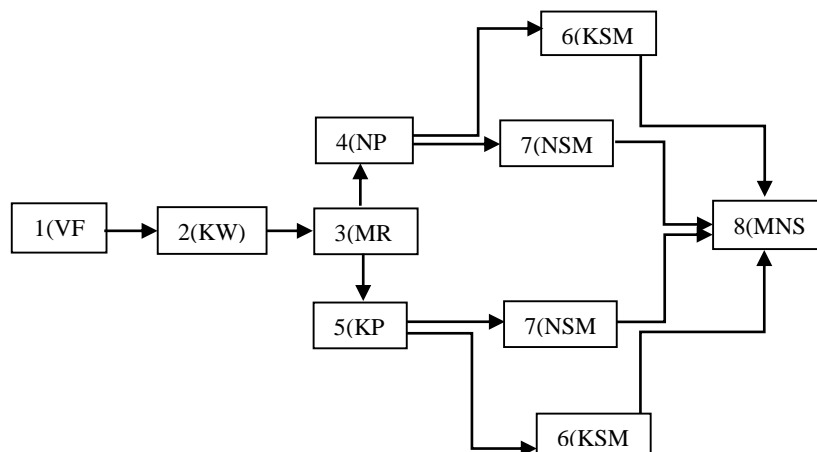


Fig. 3. Link diagram for searching by formulas in general

The diagram (Fig. 3) schematically shows the path that a formula takes in the process of establishing semantic links, depending on whether it is already in the LibMeta dictionary or has been encountered for the first time in new publications.

Explanations for the connection diagram (Fig. 3): verbal formulation of the problem (VF); presentation of the problem in the form of a set of keywords (KW) (descriptor, synonyms, associative terms); mathematical record of the problem (MR); new problem (NP); known problem (IZ); known solution method (KSM); new solution method (NSM); mathematical notation of the solution (MNS).

Information about the authors is implemented in the digital library LibMeta through the concept of "Resource". Associated with this concept are publications, keywords, and thesaurus concepts.

2.4 Composition of the author's index for mathematical SbD

The mechanism for the formation of author indexes to one degree or another is implemented in all modern bibliographic databases. Nevertheless, this is not enough to adequately reflect the importance of personality in science and use it in the educational process as a significant element in SbD.

Research shows that a person in science should be reflected in the context of links with specific achievements in SbD [28]. Of course, one cannot declare all authors to be the founders of scientific schools, but one can indicate connections with the concepts of the SbD thesaurus. This allows one to take into account the logic of research development and belonging to a certain scientific school. Such a possibility is realized only in the presence of an integrated system in which, in addition to bibliographic data, SbD ontologies are present, which distinguishes the LibMeta digital library [3, 16, 17, 31].

Thus, in the author's index, it is necessary to include, in addition to the list of publications, a list of keywords with thesaurus links. These minimal connections with SbD will give an idea of the nature of the author's research and, among other things, will reflect the development of SbD and the scientific school over time. For mathematical SbD, such an index should be provided with an index of formulas too, which will facilitate recognition of authorship among similar publications. Figure 4 shows how, based on LibMeta content (for Russian language works), to establish links between an author and SbD.

View object

Метод решения краевой упругопластической задачи о растяжении полосы с концентраторами напряжений с учетом локальных областей пластического разрушения материала

Type [Publication](#)

URI <http://libmeta.ru/mathnet/publication/vsgtu1366>

Attributes

[Annotation](#) - Предложен метод решения связанной краевой задачи деформирования твёрдого тела из пластически разрушающегося материала с учётом его повреждённости. Исходными данными для расчёта являются поля деформаций и напряжений, полученные по результатам моделирования поведения условно неповреждённой конструкции при действии фиктивных сил.

[Author](#) - [С В Горбунов](#)

[Author](#) - [В П Радченко](#)

[MSC](#) - [74R20 - Anelastic fracture and damage](#) [linked objects](#) [linked concepts](#)

[MSC](#) - [74C05 - Smallstrain, rateindependent theories \(including rigid-plastic and elastoplastic materials\)](#) [linked objects](#) [linked concepts](#)

[MSC](#) - [74G70 - Stress concentrations, singularities](#) [linked objects](#) [linked concepts](#)

[Category](#) - [Механика деформируемого твёрдого тела](#) [linked objects](#) [linked concepts](#)

Related tags [краевая задача](#) [концентратор напряжений](#) [закритическая стадия деформирования](#) [плоское напряжённое состояние условно повреждённая среда](#) [пластичность](#) [экспериментальная проверка](#)

Thesaurus terms found in the text

[Полоса](#) (Математическая энциклопедия) - Original title
[Полос метод](#) (Математическая энциклопедия) - Original title
[Область](#) (Математическая энциклопедия) - Original title
[Решение](#) (Математическая энциклопедия) - Original title
[Краевая задача](#) (Математическая энциклопедия) - Original title
[Теория](#) (Математическая энциклопедия) - Annotation

Fig. 4. Links of the author's publications to subject areas

2.5 Link analysis

Quantitative and qualitative assessments of data and knowledge of a digital library are associated, first of all, with goal-setting and the choice of criteria. For the preservation and development of scientific schools, the issue of the reliability of digital data is one of the key ones. Along with quantitative metrics set by citation data, modern research is aimed at a qualitative analysis of the accumulated bibliographic arrays.

The features of the LibMeta digital library as a data integrator make it possible to represent the topology of links, take into account their number and conduct a qualitative analysis of the accumulated data. The data is presented as a multidimensional matrix and analyzed in the direction of link vectors [29]. This approach differs from taking into account only “linear” relationships in assessments, since it allows you to analyze data by a set of features.

It is also necessary to define what is meant by the quality of data and knowledge in a digital library. One of the definitions of data quality is that we are talking about “bad” data if data are missing, incorrect or inappropriate in any context [30]. Since digital knowledge is structured data [1], what has been said applies to the quality of

knowledge, with the only difference that additional connections (associative, semantic) and expert assessments can be used to assess the quality of knowledge.

The objectives of a qualitative analysis of bibliographic arrays in a digital library is to obtain a number of estimates: "contribution" of authors to publications and research in this SbD and SbD authors; attribution of authorship in controversial cases; identification of the rating of authors and publications; identification of leading scientists and basic research; identification of new directions in research in SbD and science in general.

Appointment of the received marks: verification of databases and knowledge; identification of new and key studies; identification of researchers, authorities in SbD, on whose knowledge you can rely in the learning process.

For counting and evaluations, basic data such as thesauri, keywords, classifiers, lists of publications and authors are used. All indicators are taken into account with the distribution by years and are contained in a related form in the SbD ontology of the digital library, the analysis of the links of which is carried out in the work [31].

3 Conclusion

The idea of combining classical knowledge and new advances in science is based on the use of the addressee's thesaurus and an automated author's index as part of the SbD ontology. The difference from the modern resources presented in the review is that along with the automated index, functionality is being developed for the analysis of publications and other scientific materials, for further research and assessment of their quality.

On the one hand, this helps to reduce the number of duplicates of digital data, works and results, and on the other hand, it does not require multiple repeated searches on the Internet. The proposed methods for assessing data quality demonstrate how these links are used to determine the affiliation of SbD studies and authors of publications and scientific schools.

Own (author's, personal or collective) collection includes vocabularies of terms that are semantically related to the SbD vocabulary, thus keeping personal data and SbD up to date. This approach leads to the reduction of gaps and duplication in SbD ontologies due to the publication activity of the addressee and taking into account new terms and connections in the process of accumulating the results of author's research, which is important for the development of knowledge in the digital representation of scientific schools.

Creation of a personal environment for scientific research on the basis of individual bibliographic collections and the results collected by the user and the author in the process of research makes it possible to increase the efficiency of using digital information. This will also reduce the time and number of calls to information resources, which is important, including in the educational process.

The work is presented in the framework of the implementation of the theme of the state assignment "Mathematical methods of data analysis and forecasting" FRC CSC

of RAS and partially supported by grants #20-07-00324 and #18-29-10085mk of the Russian Foundation of Basic Research.

References

1. Gavrilova, T.A., Horoshevskij, V.F.: Bazy znaniy intellektual'nyh sistem. SPb.: Piter, 384 s. (2000).
2. Volkodav, T., Semenovskikh, T.: Dichotomy of the 'clip thinking' phenomenon, Proceedings of ICEPS 2017 (International Conference on Education, Psychology, and Social Sciences) 2 August 2017 Bangkok, Thailand. Chulalongkorn Business School, Chulalongkorn University. 4. P. 345–353 (2017) ISSN: 2518-2498.
https://www.researchgate.net/publication/320225378_Dichotomy_of_the_%27Clip_Thinking%27_Phenomenon, last accessed 2021/09/29.
3. Ataeva, O.M., Serebryakov, V.A.: Ontologiya cifrovoj semanticheskoy biblioteki LibMeta. Informatika i eyo primeneniya 12(1), S. 12–20 (2018).
<https://doi.org/10.14357/19922264180101>
4. Ataeva, O.M., Sererbryakov, V.A., Tuchkova, N.P.: Using Applied Ontology to Saturate Semantic Relations. Lobachevskii Journal of Mathematics 42(8). P. 1776–1785 (2021).
<https://doi.org/10.1134/S1995080221080059>
5. URL: https://dic.academic.ru/contents.nsf/enc_mathematics,
URL: https://encyclopediaofmath.org/wiki/Main_Page, last accessed 2021/09/29
6. Krylov, A.N.: Leonard Euler 1707–1783. Sbornik statej i materialov k 150-letiyu so dnya smerti. Izdatel'stvo Akademii nauk SSSR. S. 1–28 (1935).
7. Faddeev, L.D., Lavrov, I.A.: Rossijskie matematicheskie shkoly. Vestnik Rossijskoj Akademii Nauk 69 (5). S. 391–397 (1999).
8. Galison, P., Stump, D.J. (Eds.): The Disunity of Science: Boundaries, Contexts, and Power. Stanford, CA: Stanford University Press (1996).
9. Posselt, J.R.: Equity in Science Representation, Culture, and the Dynamics of Change in Graduate Education, Stanford, CA (2020).
10. May, K.O.: Historiography: A Perspective for Computer Scientists. Invited address to International Research Conference on the History of Computing, June 10, 1976, Los Alamos (1976).
11. Lande, D.V.: Poisk znaniy v Internet. Professional'naya rabota. Moscow: Izdatel'skij dom "Vil'yams", 272 s. (2005).
12. Shrejder, Yu.A.: Tezaurusy v informatike i teoreticheskoy semantike. Nauchno-Tekhnicheskaya Informaciya. Ser. 2. (3). S. 21–24 (1971).
13. Ion, P.D.F.: The Global Digital Mathematics Library and the International Mathematical Knowledge Trust. P.D.F. Ion, S.M. Watt. In: Geuvers H., England M., Hasan O., Rabe F., Teschke O. (Eds). Intelligent Computer Mathematics. CICM 2017. Lecture Notes in Computer Science 10383. Springer, Cham (2017).
https://doi.org/10.1007/978-3-319-62075-6_5
14. Moiseev, E.I., Muromskij, A.A., Tuchkova, N.P.: Tezaurus informacionno-poiskovyj po predmetnoj oblasti: obyknovennye differencial'nye uravneniya. Moscow: MAKS Press. 116 s. (2005).
15. Moiseev, E.I., Muromskij, A.A., Tuchkova, N.P.: Internet i matematicheskie znaniya: predstavlenie uravnenij matematicheskoy fiziki v informacionno-poiskovoj srede. Moscow: MAKS Press, 80 s. (2008).

16. Moiseev, E.I., Muromskij, A.A., Tuchkova, N.P.: O tezaurose predmetnoj oblasti smeshannye uravneniya matematicheskoy fiziki. CEUR Workshop Proceedings 2260. P. 395–405 (2018). <https://doi.org/10.20948/abrau-2018-43>
17. Tuchkova, N.P., Ataeva, O.M.: Podhody k izvlecheniyu znanij v nauchnyh predmetnyh oblastyah. Informacionnye i matematicheskie tekhnologii v nauke i upravlenii. 2 (18). S. 5–18 (2020). <https://doi.org/10.38028/ESI.2020.18.2.001>
18. Ataeva, O., Serebryakov V., Tuchkova, N.: Query Expansion Method Application for Searching in Mathematical Subject Domains. Proceedings of the 21st Conference on Scientific Services & Internet (SSI–2019) Novorossiysk–Abrau, Russia, September 23–28, 2019. 2543. P. 38–48 (2020). URL: <http://ceur-ws.org/Vol-2543/rpaper04.pdf>, urn:nbn:de:0074-2543-4, last accessed 2021/09/29.
19. Zhdanova, G.S., Kolobrodova, E.S., Polushkina, V.A., Chernyj, A.N.: Slovar' terminov po informatike na russkom i anglijskom yazykah. Moscow: Nauka. 360 s. (1971).
20. Elizarov, A., Kirillovich, A., Lipachev, E., Nevzorova, O.: Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management. In: Kalinichenko, L., Kuznetsov, S.O., Manolopoulos, Y. (eds.) DAMDID/RCDL 2016. CCIS. 706. P. 33–46. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57135-5_3
21. Elizarov, A., Kirillovich, A., Lipachev, E., Nevzorova, O.: Mathematical Knowledge Management: Ontological Models and Digital Technology. In: CEUR Workshop Proceedings 1752. P. 44–50 (2016). URL: <http://ceur-ws.org/Vol-1752/paper08.pdf>, last accessed 2021/10/01.
22. Kohlhase, M.: Mathematical knowledge management: Transcending the one-brain barrier with theory graphs. European Mathematical Society (EMS) Newsletter, June 2014, P. 22–27 (2014). URL: <https://kwarc.info/people/mkohlhase/papers/ems13.pdf>, last accessed 2021/09/29.
23. Miller, B.R.: Technical aspects of the digital library of mathematical functions. B.R. Miller, A. Youssef.: Ann. Math. Artif. Intell. 38 (1). P. 121–136 (2003). URL: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=150850, last accessed 2021/09/26
24. Kalenov, N.E., Serebryakov, V.A.: Ob ontologii Edinogo cifrovogo prostranstva nauchnyh znanij. Informacionnye Resursy Rossii 5. S. 10–12 (2020).
25. Yudina, I.G., Bazyleva, E.A.: Kompleksnye informacionnye resursy akademicheskoy biblioteki: modernizaciya i razvitie. Bibliosfera 4. S. 56–63 (2018). <https://doi.org/10.20913/1815-3186-2018-4-56-63/>
26. Meli, D.B.: Equivalence and Priority: Newton versus Leibniz: Including Leibniz's Unpublished Manuscripts on the Principia. Clarendon Press (1993).
27. Grimm S., Abecker, A., Völker, J., Studer, R.: Ontologies and the Semantic Web. In: Domingue J., Fensel D., Hendler J.A. (eds) Handbook of Semantic Web Technologies. Springer, Berlin, Heidelberg (2011).
28. <http://visualmatheditor.equatheque.net/>, last accessed 2021/09/29.
29. Barsegyan, A.A., Kupriyanov, M.S., Stepanenko, V.V., Holod, I.I.: Metody i modeli analiza dannyh: OLAP i Data Mining. SPb.: BHV-Petersburg, 331 s. (2004).
30. Singh, R., Singh, K.A.: Descriptive Classification of Causes of Data Quality Problems in Data Warehousing. IJCSI International Journal of Computer Science Issues 7(3) (2) (2010).
31. Ataeva, O., Serebryakov, V., Tuchkova, N.: Author's Identification within the Subject Area in the Semantic Library. Proceedings of the 22nd Conference on Scientific Services & Internet (SSI–2020) Novorossiysk–Abrau (online), Russia, September 21–25, 2020. 2784. P. 12–22 (2020). URL: <http://ceur-ws.org/Vol-2784/rpaper02.pdf>, urn:nbn:de:0074-2784-8, last accessed 2021/09/29.