

A semantic structuring of educational research using ontologies

Yevhenii B. Shapovalov¹, Viktor B. Shapovalov¹, Roman A. Tarasenko¹,
Stanislav A. Usenko¹ and Adrian Paschke²

¹The National Center “Junior Academy of Sciences of Ukraine”, 38-44 Degtyarivska Str., Kyiv, 04119, Ukraine

²Fraunhofer FOKUS (with support of BMBF “Qurator” 03WKDA1F), Kaiserin-Augusta-Allee 31, 10589 Berlin, Germany

Abstract. This article is devoted to the presentation of the semantic interoperability of research and scientific results through an ontological taxonomy. To achieve this, the principles of systematization and structuration of the scientific/research results in scientometrics databases have been analysed. We use the existing cognitive IT platform Polyhedron and extend it with an ontology-based information model as main contribution. As a proof-of-concept we have modelled two ontological graphs, “Development of a rational way for utilization of methane tank waste at LLC Vasylykivska poultry farm” and “Development a method for utilization of methane tank effluent”. Also, for a demonstration of the perspective of ontological systems for a systematization of research and scientific results, the “Hypothesis test system” ontological graph has created.

Keywords: cloud technologies, ontology, educational research, taxonomy, systematization

1. Introduction

Now, more than ever, science affects all aspects of human life. Latest scientific developments are often and quickly implemented in industry. However, the scientific results usually are presented in human-readable form and not in a machine-readable, so it is hard to process the knowledge using automated informational technologies.

The basic structure of a typical research paper is the sequence of Introduction, Methods, Results, and Discussion (sometimes noted as IMRAD) [30]. Each section addresses a different objective. The Introduction section motivates the research problem that was discovered or the known facts about the problem; the Method section states what authors did to discover and address the problem in a new solution, what they achieved as results in experiments is written in the Discussion section, and what they had observed is discussed in the Results section.

The most common form of science reporting is a written paper. Depending on the purpose there are a few different types of papers: Analytical Research Paper, Argumentative (Persuasive) Research Paper, Definition Paper, Compare and Contrast Paper, Cause and Effect Paper,

✉ sjb@man.gov.ua (Y. B. Shapovalov); svb@man.gov.ua (V. B. Shapovalov); tarasenko@man.gov.ua

(R. A. Tarasenko); farkry17@gmail.com (S. A. Usenko); paschke@inf.fu-berlin.de (A. Paschke)

🌐 <http://www.nas.gov.ua/UA/PersonalSite/Pages/default.aspx?PersonID=0000026333> (Y. B. Shapovalov);

<http://www.nas.gov.ua/UA/PersonalSite/Pages/default.aspx?PersonID=0000029045> (V. B. Shapovalov)

🆔 0000-0003-3732-9486 (Y. B. Shapovalov); 0000-0001-6315-649X (V. B. Shapovalov); 0000-0001-5834-5069

(R. A. Tarasenko); 0000-0002-0440-928X (S. A. Usenko); 0000-0003-3156-9040 (A. Paschke)



© Copyright for this paper by its authors, published by Academy of Cognitive and Natural Sciences (ACNS). This is an Open Access article distributed under the terms of the Creative Commons License Attribution 4.0 International (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Interpretative Experimental Research Paper, Survey Research. All the most common research papers types are shown in table 1 [31].

Table 1

The most common research papers types

Types of the Research papers	Oriented amount of words required	Specific characteristics
Analytical Research Paper	3000+	Someone poses a question and then collect relevant data from other researchers to analyse their different viewpoints.
Argumentative (Persuasive) Research Paper	3000+	The argumentative paper presents two sides of a controversial question in one paper.
Definition Paper	5000+	The definition paper describes facts or objective arguments without using any personal emotion or opinion of the author.
Compare and Contrast Paper	5000+	Compare and contrast papers are used to analyse the difference between two viewpoints, authors, subjects or stories.
Cause and Effect Paper	3000+	Cause and Effect Paper trace probable or expected results from a specific action and answer the main questions "Why?" and "What?".
Interpretative Paper	3000+*	An interpretative paper requires to use knowledge that have gained from a particular case study.
Experimental Research Paper	3000+*	This type of research paper describes a particular experiment in detail.
Survey Research Paper	5000+*	This research paper demands the conduction of a survey that includes asking questions to respondents.

* Depends on the purpose of the article and the requirements of the journal, institute, teacher

Most of the papers (but not all of them) nowadays are systemized by using scientometric databases. However, educational research reports, which use scientific methods, have not been systemized at all. Besides, scientist, unlike pupils, already know their field of research in detail and can determine by themselves their research hypothesis and they can do further analyse it by themselves. Students instead can't do this. Automated informational tools can help students in this scientific discovery and analysis tasks.

The scientific method is often used in an educational process during STEM approach by providing educational researches. This approach is only recently applied in countries such as Ukraine [47]. There are various school competitions for scientific works, such as the competition on scientific articles of the Junior academy of sciences of Ukraine and international competitions (for example, Intel ISEF). Also, the scientific method can be used during the process of creation of thesis papers (for masters' degree, bachelor's degree, etc.), pupil's research reports (for events noted before), or in simpler, but more common form of essays. In addition, students can report their results in form of scientific papers, if the level of quality of their work will be satisfactory for the scientific requirements. An overview of the types of educational research reports works are presented in table 2. The focus of this paper is on the systematization and processing

of educational research reports. The problem to be addresses is the lack of a Structuration mechanism which complicates the automated processing of the reports.

Table 2

Types of the educational research reports

Types of the educational research report	Oriented required amount of the pages	Specific characteristics	The event for which the report was prepared
Esse	In general, up to 10-15 pages	Is simple and very flexible on the content	Classes, completions of school level
Research reports	In general, up to 30-100 pages	Relatively static structure; similar to IMRAD	Competitions of Junior academy of sciences of Ukraine and Intel ISEF
Scientific paper	Declared by the source	Declared by the source	Publication in the journal
Thesis papers	In general, 40-100 pages	Relatively static structure similar to IMRAD	Defence of the qualification works

2. Literature review

To increase the convenience and efficiency of scientific data processing, structuration, and systematization of research and scientific results, the active dissemination and use of different scientometrics databases continues [44]. Specialized databases for structural science information are an integral part of the information-support system for any scientist. Scientometrics is the “quantitative study of science, communication in science, and science policy” [41] commonly referred to as the “science of science”. Scientometrics is essential to help academic disciplines understand various aspects of their research efforts, including (but not limited to) the productivity of their scholars [1, 41], the emergence of specializations [38], collaborative networks [28], patterns of scientific communications [7], and quality of research products [17]. Metric studies had developed as a subsidiary branch of Library and Information Science (LIS) over time [13]. In most cases, scientometrics models by using bibliometrics, which is a measure of the impact of publications.

To increase the quality and performance of scientometrics the ten principles of the “Leiden Manifesto of Scientometrics” have been stated [13]:

- Quantitative evaluation should support qualitative expert assessment.
- Measure performance against the research missions of the institution, group, or researcher.
- Protect excellence in locally relevant research.
- Keep data collection and analytical processes open, transparent and simple.
- Allow those evaluated to verify data and analysis.
- Account for variation by field in publication and citation practices.
- Assessment of individual research on a qualitative judgment of their portfolio.

- Avoid misplaced concreteness and false precision.
- Recognize the systemic effects of assessment and indicators.
- Scrutinize indicators regularly and update them.

Today, all existing scientometrics databases can be divided into two major groups: international and national [13, 15, 24, 34, 37, 42, 43]. The most well-known international databases are: Springer, Scopus, Web of Science, CiteseerX, Microsoft Academic, aminer, refseek, BASE (Bielefeld Academic Search Engine), WorldWideScience, JURN, Google Scholar, Google patent and others. National databases incorporate a variety of bibliographic databases, and a variety of library and university repositories. International scientometric databases are characterized by a larger scale and mandatory support for various languages, including English. Also, a characteristic feature of such databases is the availability and work with various special indices that have international recognition for example h-index [14].

As scientific publications continue to grow exponentially, also the amount of academic databases and scientometrics databases increases, which supports gaining insights into the structure and processes of science [37]. In this case, many scientific publications devoted to the principle of working scientometrics databases, and their number is growing. Thanks to them, concepts such as “metadata” of scientific articles began to be actively used in scientometrics [13, 15, 24, 34, 37, 42, 43]. Metadata is essential data about data providing information such as titles, authors, abstracts, keywords, cited references, sources, and bibliography, and other data. Metadata do not substitute the corresponding article, but it explicitly describes valuable information about the article.

By using of scientometrics systems, the contributions of researchers in the field of informatics and scientometrics were previously quantified [24]. The principal metadata indicators are: the indicators and citation indices of journals, the number of authors, the number of the publication and the degree of cooperation based on affiliation data. The disadvantage of this research is that it is devoted only to scientific articles. The authors noted that their study could not touch student’s and pupil’s research report because there is no single database where they are all located [24].

The application of the principles of the “Leiden Manifesto of Scientometrics” is stated and substantiated, which provides for transparent monitoring and support of research and encourages constructive dialogue between the scientific community and the public. In this work, the bibliometric base, which corresponds to principles of the “Leiden Manifesto of Scientometrics” has been created. The proposed bibliometric centre did not address the systematization of students and pupils’ research reports, but the authors noted the necessity of involvement of students’ and pupils’ research reports in their bibliometric centre [15].

The approach of co-word analysis has been introduced and its application in scientometrics is substantiated in [43]. The trends and patterns of scientometrics in journals has been revealed by measuring the association strength of selected keywords which represent the produced concept and idea in the field of scientometrics. Also, the authors have developed a web system for extraction of keywords from the title and abstract of the article manually. However, the web system proposed by them cannot work with research reports of students and pupils.

Another concept of analysis is iMetrics or “information metrics”. Its application in scientometrics is substantiated in [19]. iMetrics is devoted to the scientometrics of scientific journals in

the field of informatics. The authors note the possibility of applying their approach for systematization of the scientific works of students and pupils. The research related to scientometrics databases is shown in table 3.

Table 3

The research related to scientometrics databases

Subject of study	The general result of the study	Authors
Citation indices of journals, number of authors of the publication their affiliation	The contributions of researchers in the field of informatics and scientometrics	K. R. Mulla
Principles of the “Leiden Manifesto of Scientometrics”	Stated and substantiated of, “Leiden Manifesto of Scientometrics”	L. Kostenko, A. Zhabin, A. Kuznetsov, T. Lukashevich, E. Kukharchuk, T. Simonenko
Co-word analysis	The trends and patterns of scientometrics in the journals were revealed	S. Ravikumar, A. Agrahari, S. N. Singh
iMetrics (“information metrics”)	iMetrics scientometric system had provided	S. Milojevic, L. Leydesdorff

Previously, ontological graphs were used to systematize scientific articles [3, 6, 32, 36]. Systematization and structuration in such graphs is based on different approaches such as using of scientific article recommendation system [3], Scientific Articles Tagging system [6], machine learning [36], automatic summarization [32]. Also, ontologies can be to provide interoperability through semantic technologies [2]. However, none of the proposed ontological approaches for systematization and structuration is addressing the structuration of research reports of students and pupils.

None of the scientometrics database systems previously proposed [13, 15, 24, 34, 37, 42, 43] can offer a universal solution for systematization, and structured presentation of research and scientific results to pupils and students. Also, the disadvantages of all these systems are the complete lack of many parameters, that are useful for processing information about scientific works. These parameters are: the scientific novelty of the article, the practical value of the study, the hypothesis of the study, subject and object of the research. Also, existing solutions do not allow to compare research reports between each other.

This work aims to propose and justify the use of an ontological system, which permits the systematization of scientific articles with all advantages of existing scientometrics systems and without disadvantages of these systems. Which at the same time will not be deprived of the functionality of current scientometrics systems and will meet the Leiden Manifesto for Scientometrics.

We propose to use the existing cognitive IT-platform Polyhedron as technical basis for solving this problem. The core of the Polyhedron system consists of advanced and improved functions of the TODOS IT-platform described in previous works. Polyhedron is a multi-agent system which allows for transdisciplinary and acts as an interactive component in any educational and scientific research [52]. Besides, the cognitive IT-platform Polyhedron contains a function for comparison with standards which is called auditing [9, 10, 52]. Polyhedron provides: semantic web support, information systematization and ranking [11] transdisciplinary support, internal

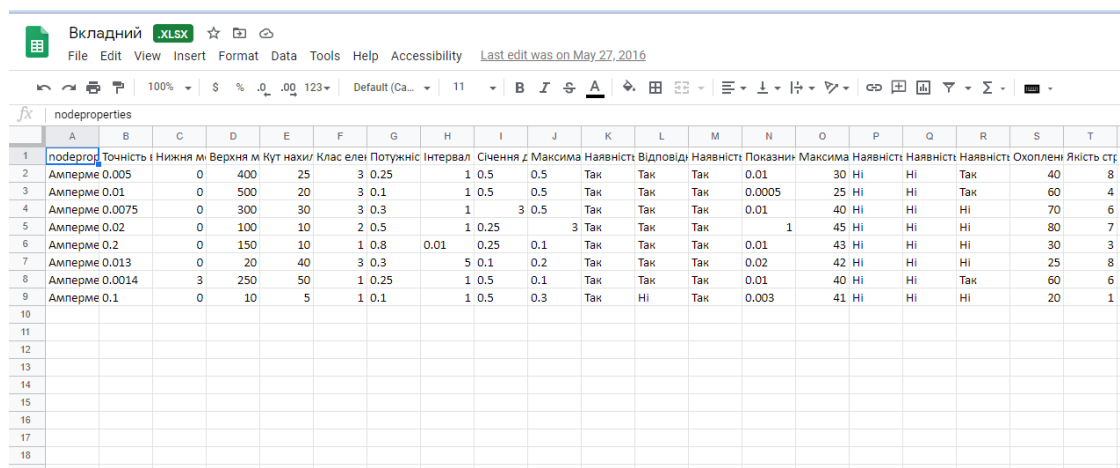
search [45] has all advantages of ontological interface tools [40], and the construction of all chains of the process of transdisciplinary integrated interaction is ensured [56]. Due to active states are hyper-ratio plural partial ordering [29, 60], the cognitive IT-platform Polyhedron is an innovative IT technology for ontological management of knowledge and information resource. The user of the Polyhedron IT system has an opportunity to use an internal search function that is more protected and reliable compared to the external one, because it provides information created by experts.

Also, the proposed solution for the structuration of educational and research projects can be used together with other modern developments in the educational field, like a virtual educational experiment [5, 16, 25, 50], different tools to provide development of ICT [8, 22, 27, 55], the use of mobile Internet devices [20, 21, 23, 54], using the technology of augmented reality education [26, 46, 51, 62], online courses [57–59, 61], distance learning in vocational education and training institutions [4, 39, 49, 53], educational and scientific environments [18, 35, 45, 48].

3. Materials and methods

3.1. Ontology creation mechanism

To create ontologies in Polyhedron, Google Sheets were used to collect and structure the information (see example in figure 1). The sheets with research report data (structure file and numeric/semantic data file) have been downloaded and saved in .xls format. The files have been loaded to “editor.stemua.science”, which is part of Polyhedron. After that, the generation of the graph nodes (in .xls) with its characteristics using the data structures in the file have been carried out. The obtained graphs have been saved in .xml format and located in the database. The graphs have been filled by semantic and numeric information for ranking and filtering. Ontological edges (relations) have been formed using predicate equations, as described previously in [56].



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	nodeprop	Точність і	Нижня м	Верхня м	Кут нахил	Клас еле	Потужніс	Інтервал	Січення	Максима	Наявніс	Відповід	Наявніс	Показни	Максима	Наявніс	Наявніс	Наявніс	Охоплен	Якість ст
2	Амперме	0.005	0	400	25	3	0.25	1	0.5	0.5	Так	Так	Так	0.01	30	Hi	Hi	Так	40	8
3	Амперме	0.01	0	500	20	3	0.1	1	0.5	0.5	Так	Так	Так	0.0005	25	Hi	Hi	Так	60	4
4	Амперме	0.0075	0	300	30	3	0.3	1	3	0.5	Так	Так	Так	0.01	40	Hi	Hi	Hi	70	6
5	Амперме	0.02	0	100	10	2	0.5	1	0.25	3	Так	Так	Так	1	45	Hi	Hi	Hi	80	7
6	Амперме	0.2	0	150	10	1	0.8	0.01	0.25	0.1	Так	Так	Так	0.01	43	Hi	Hi	Hi	30	3
7	Амперме	0.013	0	20	40	3	0.3	5	0.1	0.2	Так	Так	Так	0.02	42	Hi	Hi	Hi	25	8
8	Амперме	0.0014	3	250	50	1	0.25	1	0.5	0.1	Так	Так	Так	0.01	40	Hi	Hi	Так	60	6
9	Амперме	0.1	0	10	5	1	0.1	1	0.5	0.3	Так	Hi	Так	0.003	41	Hi	Hi	Hi	20	1
10																				
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18																				

Figure 1: Google sheet with data.

3.2. Ranking tools

Taking into account that e.g. proposed reports “A” and “B” are technical, the results of the reported works can be used to provide analysis of the rationality of the implementation proposed in the concrete project. For instance, to provide it, research reports “A” and “B” were also compared with each other using ranking tool applying the following criteria: “Short-term economic perspective”, “Long-term economic prospects”. For creating a ranking the ontologies have used the module “Alternative” which is described in our previous works [11]. To provide this ranking, the nodes of a graph have been filled with semantic data grouped in semantic classes. The ranking uses grade scale from one to ten point to underline the importance coefficient.

The projects with a payback period of more than 25 years have been evaluated with 1 point, with 20–25 years of payback period with 2 points, from 15–20 years of payback period with 3 points, from 10–15 years of payback period with 4 points, 6–10 years of payback period with 5 points and with 1–5 years were evaluated as 6–10 points, respectively, by the “Economic attractiveness” criterion. A detailed evaluation for projects with 1–5 years is provided, due to its utmost interest for the investor’s “payback time”, which determines the expediency of investment.

3.3. Auditing tools

To provide an audit of hypothesis of work “A” and “B”, the “standard” graph (with which the comparison is done) and the “comparison” graph (which is compared with the “standard”) have been created. The “standard” ontology graph contains the data on hypotheses, subjects, objects of research, keywords, and other parameters, of the research reports done before. For the “standard” graph, each parameter was presented in a separate node. The content of this ontological graph “standard” is updated and supplemented constantly.

The nodes of the “comparison” graph have been represented as names of the works which need to be audited with the “standard” graph. The parameters of the work used to be audited with the “standard” graph have been located in the metadata of each separate node. The metadata type names were identical to the names of the nodes of the “standard” graph in order to enable interaction between graphs.

4. Results and discussion

The general concept of the proposed ontology-based graph model for Polyhedron research reports has a specific, logically connected structure and can be represented as an ontology. After structuration, it is possible to represent the reports’ content in simpler to understand presentation form. Besides, most results can be domain specific for each industry, and if the current standards are correctly identified, these values will be easy to compare. Also, most research in one field often use the same equipment, materials, chemicals, standard methods of analysis, literature, etc., which allow comparing these works with each other and correctly structuration them.

However, the main advantage of the proposed approach (besides structuration of the research) is the processing of results in terms of separated result parameters of the reports. This supports data analysis, further processing using ranking, and semantic data interoperability. The separation of numeric data and its location metadata class is possible due to the addresses of the same field, that is describing the process using same (or similar) parameters of the process description and result parameters description. For example, for most reports on anaerobic digestion, the process parameters are on temperature, type of substrate, reactor volume, moisture content, initial pH, parameters; the characteristics of efficiency of the process are biogas yield, methane content, average pH during the process, destruction process etc. [12].

As all research reports will be presented in a simplified form, this approach will be especially relevant for pupils and novice researchers with further potential use in the educational process or to simplify the literature review process for the new educational research.

4.1. Description of scientific works used to provide structuration

As an example, the object of the study of research report “A” is the disposal of anaerobic effluent. The subject of the research of the report is the Cultivation of *Chlorella Vulgaris* microalgae on effluent obtained after methane fermentation. The study aims to develop a method of growing *Chlorella Vulgaris* in effluent after methane fermentation. The practical significance of this scientific work is the results of this work, which will contribute to the spread of biogas technologies. Also, the proposed approach makes it possible to increase the economic benefits from the utilization of chicken manure by converting the anaerobic digestion effluent into microalgae, that have a wide range of applications. The scientific novelty of that research report is a method of utilization of anaerobic digestion effluent by using microalgae, also had obtained cultures of *Chlorella Vulgaris* that had adapted to the anaerobic digestion effluent. The working hypothesis was that the effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae *Chlorella Vulgaris*.

The object of the study of the research report “B” is the disposal of anaerobic digestion effluent. The subject of the research is the processing of anaerobic digestion effluent into humates by the autocatalytic catalysis method. The study aims to establish regularities of processing of the solid fraction, which had obtained during the process of methane fermentation of chicken manure by autocatalytic catalysis method. The practical significance of this scientific work is that the study indicates the possibility of acquiring salts of humic and fulvic acids by the autocatalytic catalysis method. This approach makes it possible to increase the economic benefits from the disposal of chicken manure by converting the anaerobic digestion effluent into a more valuable product with a wide range of applications. Its scientific novelty is that potassium humate had firstly obtained from anaerobic digestion effluent and for the first time the efficiency of receiving humates from the solid fraction of anaerobic digestion had investigated and the main regularities of the process determined. The working hypothesis was that the solid fraction of methane fermentation of chicken manure can be recycled by the autocatalytic catalysis method.

For both research report “A” and “B”, as a substrate for anaerobic digestion have used the chicken manure from the same poultry farm. In this case, chicken manure and its effluent, which has obtained by anaerobic digestion, were analysed by the same methods and indicators. Such indicators were: “ash and dry content”, “Determination of volatile fatty acids content” (in

terms of acetic acid), “Determination of ammonium nitrogen content with Nessler’s reagent”. The equipment which has used to determine these indicators was also the same. Therefore, has considered how these works can be structured and integrated by using of the cognitive IT-platform Polyhedron. All examples of the usage ontological nodes the obtained graphs for further potential information processing are presented in table 4.

4.2. Structuration of the scientific works using ontologies

For the presentation of possibilities and systematization of the research report we have applied a ontological taxonomy for students’ works “A” and “B”. The general view of the obtained graphs is shown in figure 2 [56].

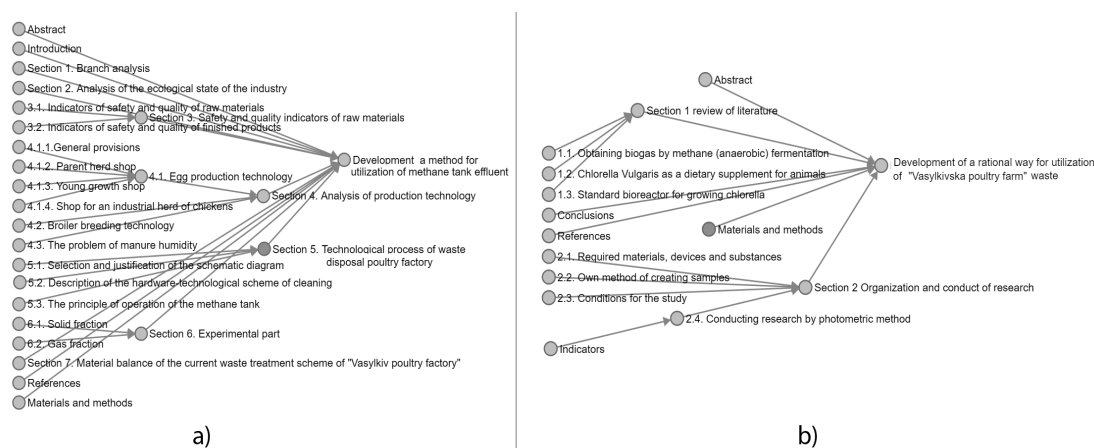


Figure 2: The general view of the (a) research report “A” (b) research report “B” ontological graph.

A separate node called “Abstract” has been created, which contains all the necessary metadata of the work such as “Object of the study”, “Subject of study”, “The aim of the study”, “Practical value”, “Scientific novelty”, “Keywords” and “Hypothesis of scientific works” in form of the attributes. All metadata have been used to provide filtering and ranking.

The “Materials and methods” node, which contains all the materials was used to perform the experiments. Every approach has been divided into the separate attribute of the node. This allows concentrating the reader’s attention, and it helps to process the data with each other. In further researchers, this mechanism will be described in detail. The general view of both works’ “Material and Methods” node is shown in figure 3 [56].

For each ontological node that duplicate sections of the research report, and that contain specific indicators after analysing, additional separate leaf nodes with these results have created. In this leaf node, all the issues are held in the form of semantic and numeric data. These results are automatically available for filtering, auditing and ranking. An example of this leaf node is shown in figure 4.

Table 4

Examples of the usage of the educational research element in ontology

Element of the educational research	Example	The role of the node in the resulting graph	Using of the data
Title	Node: "Development a method for utilization of anaerobic digestion effluent"	Parent node	Used only for structuration
Object	Node: Abstract Class: Object (object is only one per report) Value: Anaerobic digestion; Value: Microalgae's growth Value: Disposal of the waste	Located in Abstract node; each object presented as attribute	Used for the audit; to provide literature review; to link reports for each other with same data; to identify novelty and plagiarism
Subject	Node: Abstract Class: Subject Value: The processing of anaerobic digestion effluent into humates by the autocatalysis method	Located in Abstract node; each object presented as attribute	Same as previous
Hypothesis	Node: Abstract Class: Hypothesis Value: Effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae <i>Chlorella Vulgaris</i>	Located in Abstract node; each object presented as attribute	Same as previous
Keywords	Node: Abstract Class: Keywords Value1: Biogas; Value2: Anaerobic digestion Value3: Microalgae	Located in Abstract node; each object presented as attribute	Same as previous
Sections, Abstract, Introduction	Node: Introduction; Class1: Text; Value1: text itself; Class2: Biogas production in literature, ml/g of VS; Value2: 368; Class3: methane content, %; Value3: 59	Each section presented in separated nodes; all text is presented in separate class of metadata, based on type of data	Used for representing of the main text of the educational reports; structuration and navigation
Materials and methods	Node: Materials and methods Class1: Method1; Value1: Desorption1; Class2: Method2; Value2: Desorption2	Located single node; each method is separated class of metadata	Used to provide links between the reports used same method by indexing and search
Concrete results and parameters of the research	Node: Results Class1: pH; Value1: 7.3; Class2: Decomposition, %; Value2: 87	Located a in separate node; each parameter is separated class of metadata	Used for the creation of the single ranking tool to systemize results from same field
Economic data	Node: Economic data Class: Payback period, years; Value: 5.3	Located the separate node; payback period presented in metadata	Used to provide comparison of the approaches to assess investment attractiveness
References	Node: Li et al. 2018, Chen 2003, Sergienko et al. 2016	Each report (paper) located in separate node	Used to link reports used same reference with each other

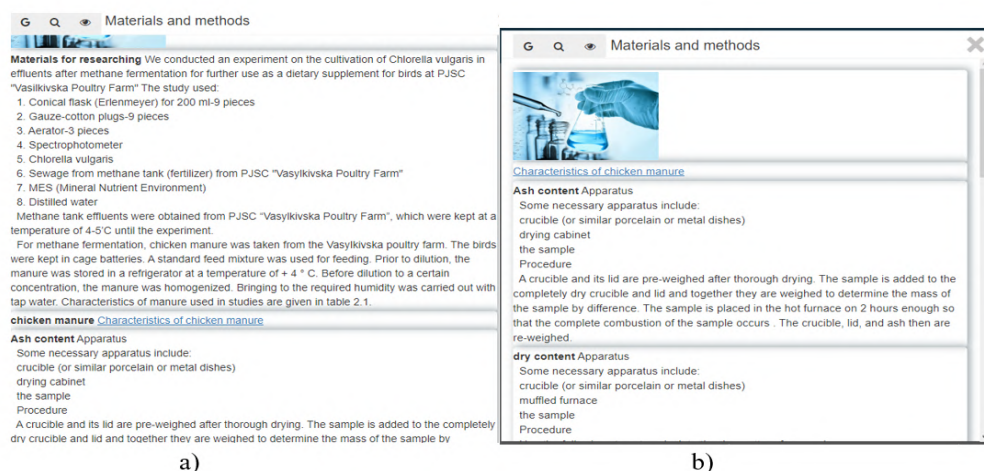


Figure 3: The general view of a) research report “A” b) research report “B” “Materials and methods” node.

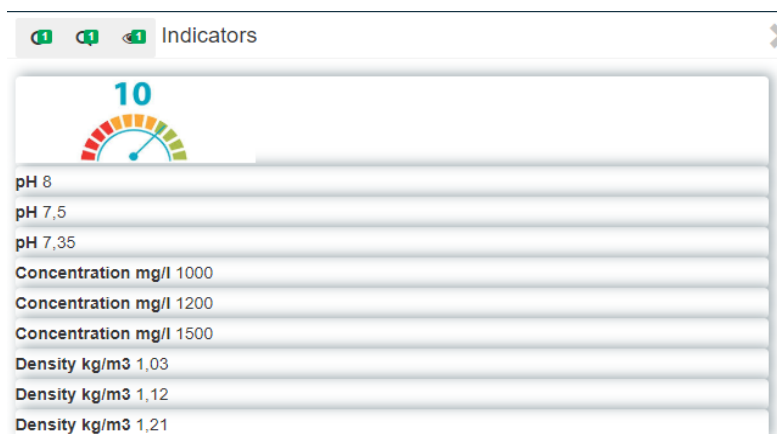


Figure 4: An example of leaf node with indicators after analysing.

5. Information processing of the research report using Polyhedron tools

5.1. Using an audit tool to test a hypothesis

The audit tool [9, 10, 52] can be used to compare the hypotheses, subjects, objects of research, keywords, and other parameters of the research reports. To demonstrate the capabilities of the audit tool, the focus is on auditing only hypotheses. A model version of the “standard” ontology has been created, which contains metadata from the “Abstract” node of the research reports “A” ontological graph. This ontology had a simple structure without branches with the parent node being named “Abstract”. The child nodes duplicate metadata from the “Abstract” node of the

research reports “A”.

The “comparison” ontology has been created with the child nodes which contain the following hypothesis: the effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae *Spirulina Platensis* (hypothesis 1), the effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae *Chlorella Vulgaris* (hypothesis 2), the effluent obtained after anaerobic digestion cannot use it as a nutrient medium for microalgae *Chlorella Vulgaris* (hypothesis 3). The hypothesis 2 node also contain some metadata. This ontology also had a simple structure without branches with the parent node is the “Hypothesis test system”. The general view of the obtained ontology of the comparison and the ontology of the standard in taxonomic form is shown in figure 5.

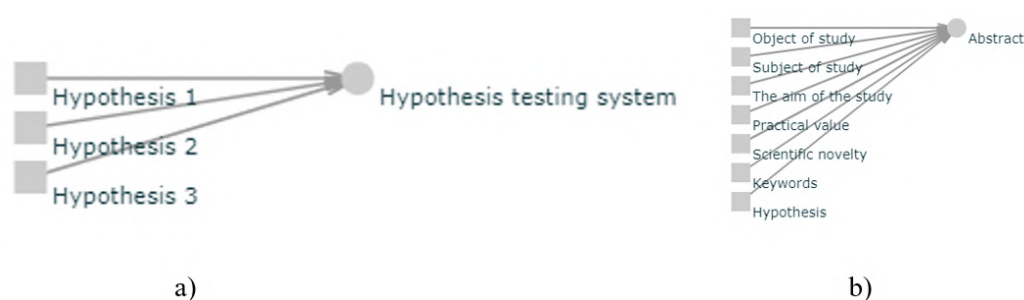


Figure 5: General view of in the taxonomic form the ontology of the “comparing” (a) and (b) the ontology of the “standard”.

Using the function of the audit the system has checked the hypothesis to be true or false. Those indicators which do not correspond to the standard have been colored by red. Thus, this solution will allow not only to test the hypothesis of these scientific works, but also to check other metadata that have already been set by using information from the “Abstract” node (see figure 6).

5.2. Analysing of the research reports result on the practice value

Research report “A” and research report “B” have been compared with each other by the following criteria “Short-term economic perspective”, “Long-term economic prospects”. According to section 2 of the research report “A”, the payback period of project “A” is five years, which corresponds to 6 points according to the criterion “Economic attractiveness”. This parameter is better for the project described in report “B” with a payback period of four years and three months which corresponds to 5 points on “Economic attractiveness”. The system provides raking of the results. In case, if there will be a large amount of the data, the instrument, will be useful to quickly and effectively evaluate the projects on “Economic attractiveness”. Besides, in further research, the other criteria will be justified and used to provide data management on the educational research, which will make the tool more functional. The general view of the ranking result is presented in figure 7.

Hypothesis testing system

(Аудит)

Враховуються властивості

#	Показники	Одиниця виміру	ДСТУ	Зразки		
				supposition 1	supposition 2	supposition 3
Abstract						
1	Object of study	Object of study	Chlorella vulgaris		Chlorella vulgaris	
2	Subject of study	Subject of study	Cultivation of Chlorella vulgaris microalgae on effluents obtained after methane fermentation.		Cultivation of Chlorella vulgaris microalgae on effluents obtained after methane fermentation.	
3	The aim of the study	The aim of the study	Developing a method of growing Chlorella Vulgaris in effluents after methane fermentation.		Developing a method of growing Chlorella Vulgaris in effluents after methane fermentation.	
4	Practical value	Practical value	The results of this work will contribute to the spread of biogas technologies. This approach makes it possible to increase the economic benefits from the utilization of bird droppings by converting the anaerobic digestion effluents into microalgae that have a wide range of applications.		The results of this work will contribute to the spread of biogas technologies. This approach makes it possible to increase the economic benefits from the utilization of bird droppings by converting the anaerobic digestion effluents into microalgae that have a wide range of applications.	
5	Scientific novelty	Scientific novelty	A method of utilization of methane tank effluent using microalgae is proposed. Cultures of Chlorella Vulgaris were adapted to the methane tank effluent.		A method of utilization of methane tank effluent using microalgae is proposed. Cultures of Chlorella Vulgaris were adapted to the methane tank effluent.	
6	Keywords	Keywords	microalgae		Chlorella Vulgaris	
7	Hypothesis	Hypothesis	The effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae Chlorella Vulgaris.	The effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae Spirulina Platensis.	The effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae Chlorella Vulgaris.	The effluent obtained after anaerobic digestion can not used it as a nutrient medium for microalgae Chlorella Vulgaris.

Figure 6: General view of the audit results in the “Hypothesis test system” ontology.

6. Discussion

The proposed database follows the “Leiden Manifesto of Scientometrics”. In the obtained ontological database quantitative evaluation can be supported by qualitative expert assessment. Additionally, this ontological database can unite the research missions of the institution, group, or researcher and protect excellence in internally relevant research. The ontological form of research reports can keep data collection and analytical processes open, transparent, and simple. Because all metadata is contained in a separate node that can be expanded and supplemented. Thus the obtained ontological database can also account for variations, e.g. in publication and citation practices and it can provide a base assessment of individual researchers in a qualitative judgment of their portfolio. Because all ontological graphs are validated by experts, in this way it is possible to avoid misplaced concreteness, including false precision and recognize the systemic effects of all assessment and indicators. In addition, in the obtained ontological database indicators can be scrutinized regularly and updated. Furthermore, the proposed ontology-based research reports can be integrated in a single environment – ontology repositories, as it was proposed before [33].

The process starts from the paper creation, for this stage we can use various text editors,


Postion	Name	Short ecomonic perspective, points
1	Development a method for utilization of methane tank effluent 	8
2	Development of a rational way for utilization of meta-tank waste at PJSC Vasytkivska poultry farm 	6

Figure 7: General view of the ranking result.

for example, word or google doc. Then expert or author of the paper will formulate metadata, which is necessary for the ontology. For this purpose, the author will use Microsoft Excel or Google Sheets. Then, an editor needs to add information in the graph, in our occasion it is the IT Platform Polyhedron. And last, but not least it is possible to use the “Alternative” system, which includes Audit, Filtering and Ranking instruments. All proposed instruments are illustrated in the workflow diagram in figure 8.

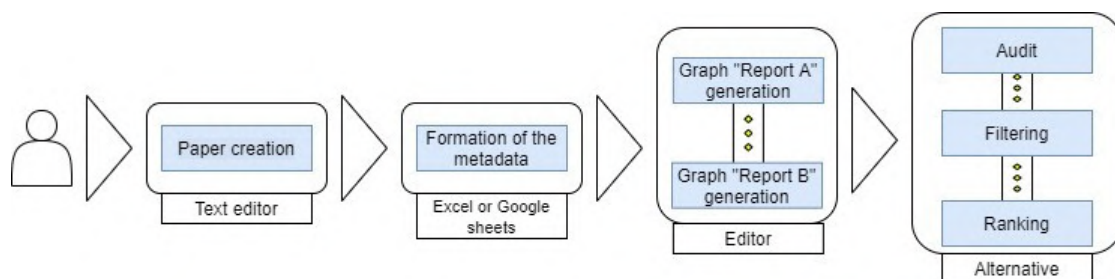


Figure 8: Workflow diagram of the creation of structured ontologies on scientific reports and their processing.

7. Conclusions

An ontological approach for the systematization of scientific works has been proposed, which also ensures their interoperability. A method of research reports structuration using digital taxonomies (ontologies) has been developed. It supports using the native structure of the reports to define hierarchical relations of the nodes. Concrete parameters were added as metadata

(semantic, numeric, pictures and links) of the nodes to provide processing using Polyhidron tools. Ranging and filtering were used for semantic and numeric metadata processing. Obtained results provide interoperability between different research reports (including educational). The obtained ontological approach follows the “Leiden Manifesto of Scientometrics”.

Further research will be devoted to provide even better interoperability between research works by providing generation of one single taxonomy that provides hierarchization by same methods, literature and results of the reports and its processing using both, methods proposed in the research and newly developed ones.

References

- [1] Abramo, G., D’Angelo, C.A. and Costa, F.D., 2011. University-industry research collaboration: A model to assess university capability. *Higher education*, 62(2), pp.163–181. Available from: <https://doi.org/10.1007/s10734-010-9372-0>.
- [2] Alnemr, R., Paschke, A. and Meinel, C., 2010. Enabling reputation interoperability through semantic technologies. *Acm international conference proceeding series*. Available from: <https://doi.org/10.1145/1839707.1839723>.
- [3] Amami, M., Faiz, R., Stella, F. and Pasi, G., 2017. A graph based approach to scientific paper recommendation. *Proceedings - 2017 IEEE/WIC/ACM international conference on web intelligence, WI 2017*. pp.777–782. Available from: <https://doi.org/10.1145/3106426.3106479>.
- [4] Bobyliev, D.Y. and Vihrova, E.V., 2021. Problems and prospects of distance learning in teaching fundamental subjects to future mathematics teachers. *Journal of physics: Conference series*, 1840(1), p.012002. Available from: <https://doi.org/10.1088/1742-6596/1840/1/012002>.
- [5] Bondarenko, O., Pakhomova, O. and Lewoniewski, W., 2020. The didactic potential of virtual information educational environment as a tool of geography students training. *Ceur workshop proceedings*, 2547, pp.13–23.
- [6] Boughareb, D., Khobizi, A., Boughareb, R., Farah, N. and Seridi, H., 2020. A Graph-Based Tag Recommendation for Just Abstracted Scientific Articles Tagging. *International journal of cooperative information systems*, 29(03), p.2050004. Available from: <https://doi.org/10.1142/S0218843020500045>.
- [7] Braun, T., Glänzel, W. and Schubert, A., 2001. Publication and cooperation patterns of the authors of neuroscience journals. *Scientometrics*, 51(3), pp.499–510. Available from: <https://doi.org/10.1023/A:1019643002560>.
- [8] Fedorenko, E., Velychko, V., Stopkin, A., Chorna, A. and Soloviev, V., 2019. Informatization of education as a pledge of the existence and development of a modern higher education. *Ceur workshop proceedings*, 2433, pp.20–32.
- [9] Globa, L., Kovalskyi, M. and Stryzhak, O.Y., 2019. Increasing Web Services Discovery Relevancy in the Multi-ontological Environment. *Advances in intelligent systems and computing*, 342, pp.335–345. Available from: <https://doi.org/10.1007/978-3-319-15147-2>.
- [10] Globa, L., Sulima, S., Skulysh, M., Dovgyi, S. and Stryzhak, O., 2020. Architecture and Operation Algorithms of Mobile Core Network with Virtualization. *Mobile computing*. IntechOpen, pp.1–22. Available from: <https://doi.org/10.5772/intechopen.89608>.

- [11] Gorborukov, V., Stryzhak, O.Y., Franchuk, O. and Shapovalov, V.B., 2018. Ontological representation of the problem of ranking alternatives. *Mathematical modeling in economics*, 4, pp.49–69. Available from: <https://doi.org/10.1017/CBO9781107415324.004>.
- [12] Ivanov, V., Shapovalov, Y.B., Stabnikov, V., Salyuk, A.I., Stabnikova, O., Rajput Haq, M. ul, Barakatullahb and Ahmed, Z., 2020. Iron-containing clay and hematite iron ore in slurry-phase anaerobic digestion of chicken manure. *Aims materials science*, 6(5), pp.821–832.
- [13] Khasseh, A.A., Soheili, F., Moghaddam, H.S. and Chelak, A.M., 2017. Intellectual structure of knowledge in iMetrics: A co-word analysis. *Information processing & management*, 53(3), pp.705–720. Available from: <https://doi.org/10.1016/j.ipm.2017.02.001>.
- [14] Kinouchi, O., Soares, L.D. and Cardoso, G.C., 2018. A simple centrality index for scientific social recognition. *Physica a: Statistical mechanics and its applications*, 491, pp.632–640. Available from: <https://doi.org/10.1016/j.physa.2017.08.072>.
- [15] Kostenko, L., Zhabin, A., Kuznetsov, A., Lukashevich, T., Kukharchuk, E. and Simonenko, T., 2015. Scientometrics: A Tool for Monitoring and Support of Research. *Science and science of science*, (3), pp.88–94.
- [16] Lavrentieva, O., Arkhypov, I., Kuchma, O. and Uchitel, A., 2020. Use of simulators together with virtual and augmented reality in the system of welders' vocational training: Past, present, and future. *Ceur workshop proceedings*, 2547, pp.201–216.
- [17] Lawani, S.M., 1986. Some bibliometric correlates of quality in scientific research. *Scientometrics*, 9(1-2), pp.13–25. Available from: <https://doi.org/10.1007/BF02016604>.
- [18] Merzlykin, P., Popel, M. and Shokaliuk, S., 2017. Services of SageMathCloud environment and their didactic potential in learning of informatics and mathematical disciplines. *Ceur workshop proceedings*, 2168, pp.13–19.
- [19] Milojević, S. and Leydesdorff, L., 2013. Information metrics (iMetrics): a research specialty with a socio-cognitive identity? *Scientometrics*, 95(1), pp.141–157. Available from: <https://doi.org/10.1007/s11192-012-0861-z>.
- [20] Modlo, Y., Semerikov, S., Bondarevskiy, S., Tolmachev, S., Markova, O. and Nechypurenko, P., 2020. Methods of using mobile Internet devices in the formation of the general scientific component of bachelor in electromechanics competency in modeling of technical objects. *Ceur workshop proceedings*, 2547, pp.217–240.
- [21] Modlo, Y., Semerikov, S., Shajda, R., Tolmachev, S., Markova, O., Nechypurenko, P. and Selivanova, T., 2020. Methods of using mobile internet devices in the formation of the general professional component of bachelor in electromechanics competency in modeling of technical objects. *Ceur workshop proceedings*, 2643, pp.500–534.
- [22] Modlo, Y., Semerikov, S. and Shmeltzer, E., 2018. Modernization of professional training of electromechanics bachelors: ICT-based Competence Approach. *Ceur workshop proceedings*, 2257, pp.148–172.
- [23] Modlo, Y.O., Semerikov, S.O., Nechypurenko, P.P., Bondarevskiy, S.L., Bondarevskaya, O.M. and Tolmachev, S.T., 2019. The use of mobile Internet devices in the formation of ICT component of bachelors in electromechanics competency in modeling of technical objects. *Ceur workshop proceedings*, 2433, pp.413–428.
- [24] Mulla, K.R., 2012. Identifying and mapping the information science and scientometrics analysis studies in India (2005-2009): A bibliometric study. *Library philosophy and practice*, pp.1–18.

- [25] Nechypurenko, P., Selivanova, T. and Chernova, M., 2019. Using the cloud-oriented virtual chemical laboratory VLab in teaching the solution of experimental problems in chemistry of 9th grade students. *Ceur workshop proceedings*, 2393, pp.968–983.
- [26] Nechypurenko, P., Starova, T., Selivanova, T., Tomilina, A. and Uchitel, A., 2018. Use of augmented reality in chemistry education. *Ceur workshop proceedings*, 2257, pp.15–23.
- [27] Nechypurenko, P., Stoliarenko, V., Starova, T., Selivanova, T., Markova, O., Modlo, Y. and Shmeltser, E., 2020. Development and implementation of educational resources in chemistry with elements of augmented reality. *Ceur workshop proceedings*, 2547, pp.156–167.
- [28] Newman, M., 2011. The structure of scientific collaboration networks. *The structure and dynamics of networks*. Princeton: Princeton University Press, vol. 9781400841, pp.221–226. Available from: <https://doi.org/10.1515/9781400841356.221>.
- [29] Nicolescu, B. and Ertas, A., 2013. *Transdisciplinary, Theory Practice*.
- [30] Oriokot, L., Buwembo, W., Munabi, I. and Kijjambu, S., 2011. The introduction, methods, results and discussion (IMRAD) structure: A Survey of its use in different authoring partnerships in a students' journal. *Bmc research notes*, 4(1), p.250. Available from: <https://doi.org/10.1186/1756-0500-4-250>.
- [31] Paperpale, 2020. Types of research papers. Available from: <https://paperpile.com/g/types-of-research-papers/> [Accessed 2020-09-17].
- [32] Parveen, D., 2018. *A Graph-based Approach for the Summarization of Scientific Articles*. Ph.D. thesis.
- [33] Paschke, A. and Schäfermeier, R., 2018. OntoMaven - Maven-based ontology development and management of distributed ontology repositories. *Advances in intelligent systems and computing*, 626, pp.251–273. 1309.7341, Available from: https://doi.org/10.1007/978-3-319-64161-4_12.
- [34] Pavlovskiy, I., 2017. Using Concepts of Scientific Activity for Semantic Integration of Publications. *Procedia computer science*, 103(October 2016), pp.370–377. Available from: <https://doi.org/10.1016/j.procs.2017.01.123>.
- [35] Pererva, V., Lavrentieva, O., Lakomova, O., Zavalniuk, O. and Tolmachev, S., 2020. The technique of the use of Virtual Learning Environment in the process of organizing the future teachers' terminological work by specialty. *Ceur workshop proceedings*, 2643, pp.321–346.
- [36] Perraudin, N., 2017. *Graph-based structures in data science : fundamental limits and applications to machine learning*. Ph.D. thesis. Available from: <https://doi.org/10.5075/epfl-thesis-7644>.
- [37] Perron, B.E., Victor, B.G., Hodge, D.R., Salas-Wright, C.P., Vaughn, M.G. and Taylor, R.J., 2017. Laying the Foundations for Scientometric Research: A Data Science Approach. *Research on social work practice*, 27(7), pp.802–812. Available from: <https://doi.org/10.1177/1049731515624966>.
- [38] Pianta, M. and Archibugi, D., 1991. Specialization and size of scientific activities: A bibliometric analysis of advanced countries. *Scientometrics*, 22(3), pp.341–358. Available from: <https://doi.org/10.1007/BF02019767>.
- [39] Polhun, K., Kramarenko, T., Maloivan, M. and Tomilina, A., 2021. Shift from blended learning to distance one during the lockdown period using Moodle: test control of students' academic achievement and analysis of its results. *Journal of physics: Conference series*, 1840(1), p.012053. Available from: <https://doi.org/10.1088/1742-6596/1840/1/012053>.

- [40] Popova, M. and Stryzhak, O.Y., 2013. Ontological interface as a means of presenting information resources in the GIS environment. *Scientific notes of the taurida national university. v. i. vernadsky.*, 65(26), pp.127–135.
- [41] Ramesh Babu, A. and Singh, Y.P., 1998. Determinants of research productivity. *Scientometrics*, 43(3), pp.309–329. Available from: <https://doi.org/10.1007/BF02457402>.
- [42] Ramirez, M.C. and Devesa, R.A.R., 2019. A scientometric look at mathematics education from Scopus database. *Mathematics enthusiast*, 16(1-3), pp.37–46.
- [43] Ravikumar, S., Agrahari, A. and Singh, S.N., 2015. Mapping the intellectual structure of scientometrics: a co-word analysis of the journal *Scientometrics* (2005–2010). *Scientometrics*, 102(1), pp.929–955. Available from: <https://doi.org/10.1007/s11192-014-1402-8>.
- [44] Semerikov, S., Pototskyi, V., Slovak, K., Hryshchenko, S. and Kiv, A., 2018. Automation of the export data from Open Journal Systems to the Russian Science Citation Index. *Ceur workshop proceedings*, 2257, pp.215–226.
- [45] Shapovalov, V., Shapovalov, Y., Bilyk, Z., Atamas, A., Tarasenko, R. and Tron, V., 2019. Centralized information web-oriented educational environment of Ukraine. *Ceur workshop proceedings*, 2433, pp.246–255. Available from: <http://ceur-ws.org/Vol-2433/paper15.pdf>.
- [46] Shapovalov, V., Shapovalov, Y., Bilyk, Z., Megalinska, A. and Muzyka, I., 2020. The Google Lens analyzing quality: An analysis of the possibility to use in the educational process. *Ceur workshop proceedings*, 2547, pp.117–129. Available from: <http://www.ceur-ws.org/Vol-2547/paper0>.
- [47] Shapovalov, Y., Shapovalov, V., Andruszkiewicz, F. and Volkova, N., 2020. Analyzing of main trends of STEM education in ukraine using stemua.science statistics. *Ceur workshop proceedings*, 2643, pp.448–461. Available from: <http://ceur-ws.org/Vol-2643/paper26.pdf>.
- [48] Shapovalov, Y., Shapovalov, V. and Zaselskiy, V., 2019. TODOS as digital science-support environment to provide STEM-education. *Ceur workshop proceedings*, 2433, pp.232–245. Available from: <http://ceur-ws.org/Vol-2433/paper14.pdf>.
- [49] Shokaliuk, S., Bohunenko, Y., Lovianova, I. and Shyshkina, M., 2020. Technologies of distance learning for programming basics on the principles of integrated development of key competences. *Ceur workshop proceedings*, 2643, pp.548–562.
- [50] Slipukhina, I., Kuzmenkov, S., Kurilenko, N., Mienailov, S. and Sundenko, H., 2019. Virtual educational physics experiment as a means of formation of the scientific worldview of the pupils. *Ceur workshop proceedings*, 2387, pp.318–333.
- [51] Striuk, A., Rassovytska, M. and Shokaliuk, S., 2018. Using Blippar augmented reality browser in the practical training of mechanical engineers. *Ceur workshop proceedings*, 2104, pp.412–419.
- [52] Stryzhak, O.Y., Gorborkov, V., Franchuk, O. and Popova, M., 2014. Ontology of the choice problem and its application in the analysis of limnological systems. *Ecological safety and nature management*, pp.172–183.
- [53] Syvyi, M., Mazbayev, O., Varakuta, O., Panteleeva, N. and Bondarenko, O., 2020. Distance learning as innovation technology of school geographical education. *Ceur workshop proceedings*, 2731, pp.369–382.
- [54] Tkachuk, V., Yechkalo, Y., Semerikov, S., Kislova, M. and Hladyr, Y., 2021. Using Mobile ICT for Online Learning During COVID-19 Lockdown. In: A. Bollin, V. Ermolayev, H.C. Mayr, M. Nikitchenko, A. Spivakovsky, M. Tkachuk, V. Yakovyna and G. Zholtkevych,

- eds. *Information and communication technologies in education, research, and industrial applications*. Cham: Springer International Publishing, pp.46–67.
- [55] Tkachuk, V., Yechkalo, Y., Semerikov, S., Kislova, M. and Khotskina, V., 2020. Exploring student uses of mobile technologies in university classrooms: Audience response systems and development of multimedia. *Ceur workshop proceedings*, 2732, pp.1217–1232.
 - [56] Velichko, V., Popova, M., Prihodnyuk, V. and Stryzhak, O.Y., 2017. TODOS is an IT platform for the formation of transdisciplinary information environments. *Weapons systems and military equipment*, 1(49), pp.10–19.
 - [57] Vlasenko, K., Chumak, O., Lovianova, I., Kovalenko, D. and Volkova, N., 2020. Methodical requirements for training materials of on-line courses on the platform "Higher school mathematics teacher". *E3s web of conferences*, 166. Available from: <https://doi.org/10.1051/e3sconf/202016610011>.
 - [58] Vlasenko, K., Kovalenko, D., Chumak, O., Lovianova, I. and Volkov, S., 2020. Minimalism in designing user interface of the online platform "Higher school mathematics teacher". *Ceur workshop proceedings*, 2732, pp.1028–1043.
 - [59] Vlasenko, K., Volkov, S., Sitak, I., Lovianova, I. and Bobyliev, D., 2020. Usability analysis of on-line educational courses on the platform "Higher school mathematics teacher". *E3s web of conferences*, 166, p.10012. Available from: <https://doi.org/10.1051/e3sconf/202016610012>.
 - [60] Volckmann, R., 2007. Transdisciplinarity: Basarab Nicolescu Talks with Russ Volckmann. *Lancet neurology*, 6(9), p.76. Available from: [https://doi.org/10.1016/S1474-4422\(07\)70211-9](https://doi.org/10.1016/S1474-4422(07)70211-9).
 - [61] Yahupov, V.V., Kyva, V.Y. and Zaselskiy, V.I., 2020. The methodology of development of information and communication competence in teachers of the military education system applying the distance form of learning. *Ceur workshop proceedings*, 2643, pp.71–81.
 - [62] Zelinska, S., Azaryan, A. and Azaryan, V., 2018. Investigation of opportunities of the practical application of the augmented reality technologies in the information and educative environment for mining engineers training in the higher education establishment. *Ceur workshop proceedings*, 2257, pp.204–214.