

The role of educational and scientific studies taxonomies in a centralised informational web-oriented educational environment

Viktor B. Shapovalov¹, Yevhenii B. Shapovalov¹, Roman A. Tarasenko¹, Stanislav A. Usenko¹, Adrian Paschke² and Irina M. Shapovalova³

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

²Fraunhofer FOKUS (with support of BMBF “EXPAND+ER WB3”), 31 Kaiserin-Augusta-Allee, Berlin, 10589, Germany

³Secondary comprehensive school No. 69 of Kyiv, 25 Donetska Str., Kyiv, 03151, Ukraine

Abstract

Scientific and educational studies can be structured using the formalisation of the IMRAD approach, providing data interoperability. This study focuses on using study results as part of a centralised informational web-oriented educational environment. Structurisation was applied to two studies: “Development of a rational approach for utilising methane tank waste at LLC Vasytkivska chicken farm” and “Development of a strategy for utilising methane tank effluent”. Specific tools from CIT Polyhedron were used to process study data. The audit tool compares newly inputted data to existing taxonomies and highlights cases of full correspondence between elements of different works, such as study objects. The approach of integrating studies with educational ontologies, which is part of the centralised informational web-oriented educational environment, is described. The formalisation of this process is described using mathematical expressions.

Keywords

ontology, educational research, scientific studies, taxonomy, centralised educational environment

1. Introduction

Science now affects all aspects of human life more than ever before. The latest scientific developments are often quickly implemented in industry. However, scientific results are usually presented in human-readable form rather than machine-readable format, making it difficult 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 abbreviated as IMRAD) [1]. Each section addresses a different objective. The Introduction motivates the research problem or known facts about it; the Method section states what the authors did to learn about and address the issue with a new solution, what they achieved in experiments is described in the Discussion section, and their observations are discussed in the Results section.

The most common form of scientific reporting is a written paper. Depending on the purpose, there are several different paper types: Analytical Research Paper, Argumentative (Persuasive) Research Paper, Definition Paper, Compare and Contrast Paper, Cause and Effect Paper, and Interpretative Paper. The most common research paper types are shown in table 1 [2].

Nowadays, most papers (but not all) are systematised using scientometric databases. However, educational research reports, which use scientific methods, have not been systematised. Unlike pupils, scientists already know their field of research in detail and can determine their research hypothesis, and they can do further analysis by themselves. Students, instead, cannot do this. Automated informational tools can help students in these scientific discovery and analysis tasks.

AREdu 2024: 7th International Workshop on Augmented Reality in Education, May 14, 2024, Kryvyi Rih, Ukraine

✉ svb@man.gov.ua (V. B. Shapovalov); svb@man.gov.ua (Y. B. Shapovalov); tarasenko@man.gov.ua (R. A. Tarasenko);

farkry17@gmail.com (S. A. Usenko); paschke@inf.fu-berlin (A. Paschke); 2429920@gmail.com (I. M. Shapovalova)

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

0000-0002-0440-928X (S. A. Usenko); 0000-0003-3156-9040 (A. Paschke); 0000-0003-3156-9040 (I. M. Shapovalova)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Table 1

The most common research paper types.

Types of the research papers	Oriented amount of words required	Specific characteristics
Analytical Research Paper	3000+	Someone poses a question and then collects 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 traces probable or expected results from a specific action and answers the main questions “Why?” and “What?”.
Interpretative Paper	3000+*	An interpretative paper requires one to use knowledge 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.

The STEM (Science, Technology, Engineering and Mathematics) approach may be interpreted as using the scientific method in an educational process while providing academic research. This approach has only recently been applied in countries such as Ukraine. 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 (Intel ISEF). Also, the scientific method can be used during the creation of thesis papers (for masters’ degrees, bachelor’s degrees, etc.) and pupils’ research reports (for events noted before), or in simpler, but more common form of essays. In addition, students can report their results in scientific papers if the quality of their work meets scientific requirements. An overview of the types of educational research reports is presented in table 2. This paper focuses on the systematisation and processing of academic research reports. The problem to be addressed is the lack of a structuring mechanism that complicates the automated processing of such reports.

Table 2

Types of 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
Essay	In general, up to 10-15 pages	Simple and very flexible in content	Classes, school-level competitions
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 a journal
Thesis papers	In general, 40-100 pages	Relatively static structure similar to IMRAD	Defence of qualification works

2. Literature review

The active dissemination and use of different scientometric databases continue to increase the convenience and efficiency of scientific data processing, structuring, and systematisation of research and scientific results. Specialised 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” [3], 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 [3, 4], the emergence of specialisations [5], collaborative networks [6], patterns of scientific communications [7], and quality of research products [8]. Metric studies have been developed as a subsidiary branch of Library and Information Science (LIS) [9]. Often, scientometrics applies bibliometrics, which measures the impact of publications.

To increase the quality and performance of scientometrics, the ten principles of the “Leiden Manifesto for Scientometrics” [9] have been stated:

1. Quantitative evaluation should support qualitative expert assessment.
2. Measure performance against the research missions of the institution, group, or researcher.
3. Protect excellence in locally relevant research.
4. Keep data collection and analytical processes open, transparent and simple.
5. Allow those evaluated to verify data and analysis.
6. Account for variation by field in publication and citation practices.
7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
8. Avoid misplaced concreteness and false precision.
9. Recognise the systemic effects of assessment and indicators.
10. Scrutinise indicators regularly and update them.

Today, all existing scientometric databases can be divided into two major groups: international and national [9, 10, 11, 12, 13, 14, 15]. 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 Patents, and others. National databases incorporate a variety of bibliographic databases and a variety of library and university repositories. International scientometric databases are characterised 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 multiple unique indices that have international recognition, for example, the h-index [16].

As scientific publications continue to grow exponentially, the number of academic databases and scientometric databases increases, which supports gaining insights into the structure and processes of science [14]. In this case, many scientific publications are devoted to the working principles of scientometric databases, and their number is growing. Thanks to them, concepts such as “metadata” of scientific articles began to be actively used in scientometrics [9, 10, 11, 12, 13, 14, 15]. Metadata is essential data about data providing information such as titles, authors, abstracts, keywords, cited references, sources, bibliography, and other data. Metadata does not substitute the corresponding article, but it explicitly describes valuable information about the report.

By using scientometric systems, researchers’ contributions in informatics and scientometrics were previously quantified [11]. The principal metadata indicators are:

- The indicators and citation indices of journals.
- The number of authors.
- The number of publications.
- 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 cover students’ and pupils’ research reports because there is no single database where they are all located.

The application of the principles of the “Leiden Manifesto for Scientometrics” is stated and substantiated, providing transparent monitoring and support of research and encouraging 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 center did not address the systematisation of students’ and pupils’ research reports. Still, the authors noted the necessity of involvement of students’ and pupils’ research reports in their bibliometric center.

The approach of co-word analysis has been introduced, and its application in scientometrics is substantiated in [12]. The trends and patterns of scientometrics in journals have 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 [17]. iMetrics is devoted to the scientometrics of scientific journals in the field of informatics. The authors note the possibility of applying their approach to the systematisation of the scientific works of students and pupils. The research related to scientometric databases is shown in table 3.

Table 3

Research related to scientometric databases.

Subject of study	The general result of the authors’ study
Citation indices of journals, number of authors of the publication their affiliation	The contributions of researchers in the field of informatics and scientometrics [11]
Principles of the “Leiden Manifesto for Scientometrics”	Stated and substantiated “Leiden Manifesto for Scientometrics” [10]
Co-word analysis	The trends and patterns of scientometrics in the journals were revealed [12]
iMetrics (“information metrics”)	iMetrics scientometric system has been provided [17]

Previously, ontological graphs were used to systematise scientific articles [18, 19, 20, 21]. Systematisation and structuring in such graphs are based on different approaches, such as using a scientific article recommendation system [18], Scientific Articles Tagging system [19], machine learning [20], and automatic summarisation [21]. Also, ontologies can be used to provide interoperability through semantic technologies [22]. However, none of the proposed ontological approaches for systematisation and structuring addresses the structuring of research reports of students and pupils.

None of the scientometric database systems previously proposed [9, 10, 11, 12, 13, 14, 15] can offer a universal solution for systematisation 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 valuable parameters 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 for comparing the metadata about the research reports between each other.

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

As Proof of Concept (PoC), we propose to use the existing cognitive IT platform Polyhedron as the 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. The Polyhedron is a multi-agent system that allows for transdisciplinarity and acts as an interactive component in educational and scientific research [23]. Besides, the cognitive IT platform Polyhedron contains a function for comparison with standards which is called auditing [23, 24, 25]. Polyhedron provides: semantic web support, information systematisation and ranking [26], transdisciplinary support, and

internal search [27], has all advantages of ontological interface tools [28], and the construction of all chains of the process of transdisciplinary integrated interaction is ensured [29]. Due to active states for hyper-ratio plural partial ordering [30, 31], the cognitive IT platform Polyhedron is an innovative IT technology for ontological management of knowledge and information resources, regardless of the standards of their creation. The user of the Polyhedron IT system has an opportunity to use an internal search function that has more views than the external one because it provides information created by experts.

Also, the proposed solution for the structuring of educational and research projects can be used together with other modern developments in the academic field, like a virtual educational experiment [32], different tools to provide development of ICT [33], the use of mobile Internet devices [34], using the technology of augmented reality education [35], online courses [36], distance learning in vocational education and training institutions [34], educational and scientific environments [27].

As was investigated before, the main elements of educational studies are represented by IMRAD nodes and their specific subnodes related to a particular study [37]. They may be described by a set of formulas. According to the theory of using IMRAD, each examination consists of an Introduction, Methods, Results, and Discussion (in terms of informational systems, the discussion is charged to processing – P):

$$\{I, M, R, P\} \in S \quad (1)$$

where I – node of ontology that integrates data related to introduction; M – subject of study: node of ontology that integrates data related to methods; R – node of ontology that integrates data related to results; P – results of study's results processing.

Each scientific study contains specific data structured by IMRAD, and it may be represented as a set of tuples (corteges) that describe elements of specific studies. The equations 2 and 3 are used to describe representing two different studies structured by IMRAD:

$$S_I = \langle I_I, M_I, R_I, P_I \rangle \quad (2)$$

$$S_{II} = \langle I_{II}, M_{II}, R_{II}, P_{II} \rangle \quad (3)$$

Two different studies integrated into a single ontology will be described as the sum of IMRAD elements. Such representation is shown in equation 4:

$$\langle S_I, S_{II} \rangle = \langle I_I, M_I, R_I, P_I, I_{II}, M_{II}, R_{II}, P_{II} \rangle \quad (4)$$

In such case, some specific elements of studies are overlapping and other are not. For example, the Method section of two different studies represented in form of an ontology using IMRAD will be as follows:

$$M_I = M_a, M_b, M_c, M_d \quad (5)$$

$$M_{II} = M_b, M_d, M_f \quad (6)$$

In such representation M_b and M_d belong to both studies, and it is possible to use them as linking nodes to connect the two studies:

$$M_b \in M_I, M_{II} \quad (7)$$

$$M_d \in M_I, M_{II} \quad (8)$$

It is worth noting that such representation of studies leads to the ontologisation of the studies' data. The most specific terms may be used to connect to different types of ontology-based knowledge, for example, educational programmes. However, such an approach was not conducted before. This study also aims to provide interoperability between different ontology-based knowledge systems using terms used in conducted studies and other knowledge systems (on the example of educational systems).

3. Materials and methods

3.1. Ontology creation mechanism

To create ontologies in Polyhedron, Google Sheets were used to collect and structure the information by an expert who took the data manually (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”, part of Polyhedron. After that, the generation of the graph nodes (in .xls) with their characteristics using the data structures in the file has been carried out. The obtained graphs have been saved in .xml format and located in the database. The graphs have been filled with semantic and numeric information for ranking and filtering. Ontological edges (relations) have been formed using predicate equations, as described previously in [29].

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	підергод	Точність	Нижня м	Верхня м	Кут нахил	Клас еле	Потужність	Інтервал	Січення з	Максима	Нааявність	Відповідь	Нааявність	Показник	Максима	Нааявність	Нааявність	Нааявність	Охоплені	Якість ст
2	Амперме 0.005	0	400	25	3	0.25	1	0.5	0.5	Так	Так	Так	0.01	30	Ні	Ні	Так	40	8	
3	Амперме 0.01	0	500	20	3	0.1	1	0.5	0.5	Так	Так	Так	0.0005	25	Ні	Ні	Так	60	4	
4	Амперме 0.0075	0	300	30	3	0.3	1	3	0.5	Так	Так	Так	0.01	40	Ні	Ні	Ні	70	6	
5	Амперме 0.02	0	100	10	2	0.5	1	0.25	3	Так	Так	Так	1	45	Ні	Ні	Ні	80	7	
6	Амперме 0.2	0	150	10	1	0.8	0.01	0.25	0.1	Так	Так	Так	0.01	43	Ні	Ні	Ні	30	3	
7	Амперме 0.013	0	20	40	3	0.3	5	0.1	0.2	Так	Так	Так	0.02	42	Ні	Ні	Ні	25	8	
8	Амперме 0.0014	3	250	50	1	0.25	1	0.5	0.1	Так	Так	Так	0.01	40	Ні	Ні	Так	60	6	
9	Амперме 0.1	0	10	5	1	0.1	1	0.5	0.3	Так	Ні	Так	0.003	41	Ні	Ні	Ні	20	1	
10																				

Figure 1: Google sheet with data.

3.2. Ranking tools

Considering that, e.g., proposed reports “A” and “B” are technical, the results of the reported works can be used to analyse the rationality of the implementation proposed in the concrete project. For instance, to offer it, research reports “A” and “B” were also compared with each other using a 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”, described in [26]. The nodes of a graph have been filled with semantic data to provide this ranking.

The ranking uses a grade scale from one to ten points to underline the relevance 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 investment expediency.

3.3. Auditing tools

To provide an audit of the 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 constantly updated and supplemented.

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 terms of the nodes of the “standard” graph to enable interaction between graphs.

3.4. Using centralised informational web-oriented educational environment concept and ensuring interdisciplinarity

The developed ontologies were saved in the same environment where elements of the centralised informational web-oriented educational environment were saved. Its features were used to provide interoperability with educational programs, methods, equipment, ontology-based didactical materials, and other ontology tools. As all such ontologies had the same graphs’ nodes names, we provided the integration between elements of the centralised informational web-oriented educational environment and proposed structuring of academic studies. We used the same nodes and provided links with each graph that it contains. For example, the term temperature regime that is used in educational programmes is connected with all academic programmes in physics (part of topic energy, thermal energy), chemistry (amount of topic energy of reaction), and scientific study graph that was conducted by a young researcher on biogas production research called temperature regime. Also, we can link this term with a method called ensuring of requested temperature by a thermostat and with equipment dry air thermostat. So, for this, we are using the term temperature regime to provide an interdisciplinarity approach that is related to different fields of science and to varying types of data (educational plans, equipment that is used, specific methods and specific personal studies).

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 structuring, it is possible to describe the reports’ content in a 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 uses the same equipment, materials, chemicals, standard methods of analysis, literature, etc., which allow comparing these works with each other and correctly structuring them.

However, the main advantage of the proposed approach (besides structuring the research) is processing 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, describing the process using the same (or similar) parameters of the process description and result parameters description. For example, for most reports on anaerobic digestion, the process parameters are temperature, type of substrate, reactor volume, moisture content, initial pH; the characteristics of efficiency of the process are biogas yield, methane content, average pH during the process, destruction process, etc. [38].

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

4.1. Description of scientific works used to provide structuring

For example, the object of the study of research report “A” is the disposal of anaerobic effluent. The subject of the report’s research 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 the effluent after methane fermentation. The practical significance of this scientific work is the results, which will contribute to the spread of biogas technologies. Also, the proposed approach makes it possible to increase the economic benefits of utilising chicken manure by converting the anaerobic digestion effluent into microalgae with a wide range of applications. The scientific novelty of that research report is a method of utilisation 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 could be used as a nutrient medium for microalgae *Chlorella Vulgaris*.

The object of the research report “B” study 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 the solid fraction obtained during the methane fermentation of chicken manure by the 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 of chicken manure disposal 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. For the first time, the efficiency of receiving humates from the solid fraction of anaerobic digestion was investigated, and the main regularities of the process were determined. The working hypothesis was that the solid fraction of methane fermentation of chicken manure can be recycled by the autocatalytic catalysis method.

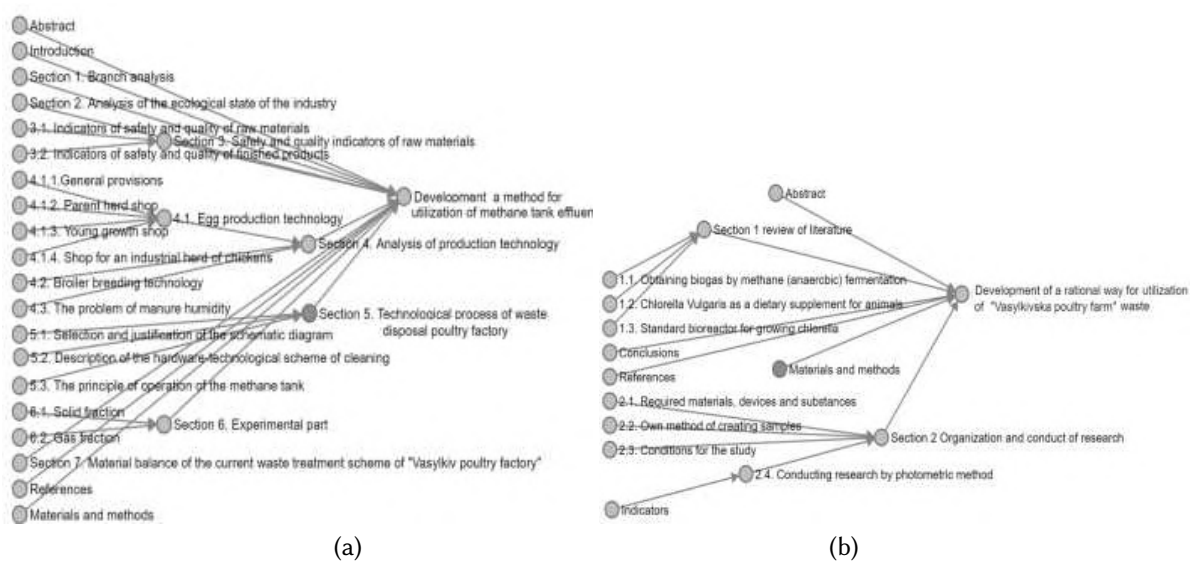


Figure 2: The general view of the (a) research report “A” (b) research report “B” ontological graph.

For both research reports, “A” and “B”, chicken manure from the same poultry farm has been used as a substrate for anaerobic digestion. In this case, chicken manure and its effluent, which has been 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 used to determine these indicators was also the same. Therefore, how these works can be structured and integrated using the cognitive IT platform Polyhedron has been considered. All examples of ontological nodes in the obtained graphs for further potential information processing are presented in Table 4.

4.2. Structuring of the scientific works using ontologies

To present possibilities and systematisation of the research report, we have applied an ontological taxonomy for students’ works “A” and “B”. The general view of the obtained graphs is shown in figure 3 [29].

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 utilisation 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 systemise 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

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 the form of the attributes. All metadata has 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. For further researchers, this mechanism will be described in detail. The general view of both works’ “Material and Methods” node is shown in figure 3 [29].

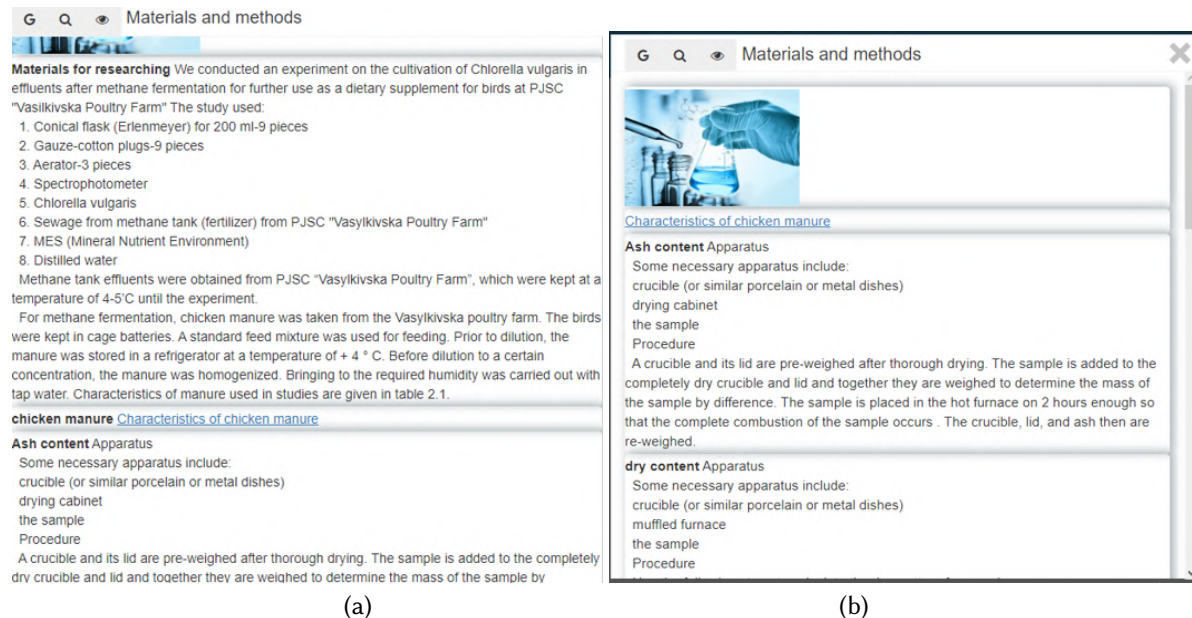


Figure 3: The general view of a) research report “A” b) research report “B” “Materials and methods” node.

For each ontological node that duplicates sections of the research report, and that contain specific indicators after analysing, additional separate leaf nodes with these results have been 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.

5. Information processing of the research report using Polyhedron tools

5.1. Using an audit tool to test a hypothesis

The audit tool [23, 24, 25] 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), and 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 contains some metadata. This ontology also had a simple structure without

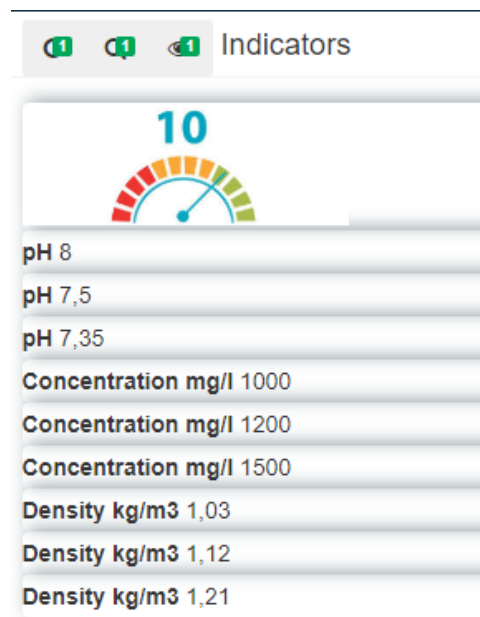


Figure 4: An example of leaf node with indicators after analysing.

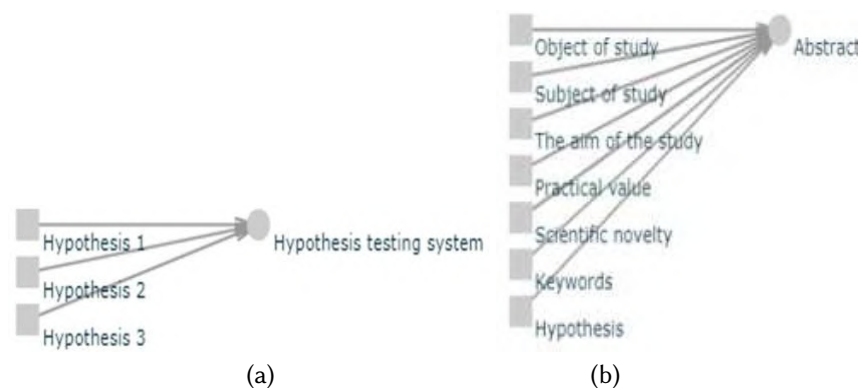


Figure 5: General view of in the taxonomic form the ontology of the comparing” (a) and (b) the ontology of the “standard”.

branches with the parent node, 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.

The system has checked whether the hypothesis is true or false by using the audit function. Those indicators which do not correspond to the standard have been coloured red. Thus, this solution will allow testing the idea of these scientific works and checking other metadata that have already been set by using information from the “Abstract” node (figure 5b).

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 ranking of the results. If there is a large amount of data, the instrument will be helpful 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.

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 be used as a nutrient medium for microalgae Chlorella Vulgaris.

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


Postion	Name	Short economic perspective, points
1	<p>Development a method for utilization of methane tank effluent</p> 	8
2	<p>Development of a rational way for utilization of meta-tank waste at PJSC Vasykivska poultry farm</p> 	6

Figure 7: General view of the ranking result.

5.3. The role of taxonomies of educational studies in centralised informational web-oriented educational environment

5.3.1. Mathematical interoperation of integration of taxonomies of educational studies in centralised informational web-oriented educational environment

As it was shown before, the preleaf nodes (L4) are terms (main ontology elements) that are very promising to use in terms of interoperability with other ontologies of subject areas, for example, with educational programmes that in that term will be the additional instrument for Centralised informational web-oriented educational environment.

So, such connection with the centralised informational web-oriented educational environment concept and ensuring interdisciplinarity is described by formulas. Each ontology is based on the conceptualisation of terms. It means that each ontology is described as a tuple (cortege) of terms from the field it contains:

$$O_i = \langle t_i \rangle \quad (9)$$

So, we can describe ontology of educational programme, ontology of equipment that being used, ontology of method and ontology of educational studies as further:

$$O_1 = t_1, t_2, t_3, t_4, t_5 \quad (10)$$

$$O_2 = t_1, t_3, t_5, t_6 \quad (11)$$

$$O_3 = t_3, t_5, t_6, t_7 \quad (12)$$

$$O_4 = t_1, t_3, t_8, t_9, t_{10}, t_{11}, t_{12} \quad (13)$$

As seen for equations, some ontologies have cross-terms that will provide inoperability with CIWOEE and interdisciplinarity. The terms that are cross-terms will be used by the user to transfer from elements (nodes) of one ontology to aspects of another. For this example, leaf or sub-leaf (in the case of educational studies' ontology; such as specific methods, keywords, objects, etc.), t_1 , t_3 , and t_5 are cross-terms:

$$t_1 \in O_1, O_2, O_4 \quad (14)$$

$$t_3 \in O_1, O_2, O_3, O_4 \quad (15)$$

$$t_5 \in O_1, O_2, O_3 \quad (16)$$

5.3.2. Practical application of ontology-based integration of scientific studies educational programmes of centralised informational web-oriented educational environment

As shown in equations 10-13, the terms of scientific study ontology, such as specific methods, keywords, objects, etc., are used to provide a link with terms of educational programmes. Terms of studies that may be used to link the ontology of scientific studies with the ontology of educational programmes are shown in figure 8.

A similar situation is also related to educational programmes. There is an ontology that systemises the data on the knowledge field related to chemistry using schools' educational programmes. It also consists of terms that are presented in the form of nodes. The general view of academic programmes' ontology related to chemistry educational programmes in Ukraine and phrases that may be used to link with the ontology of scientific studies is shown in figure 9.

Therefore, some terms may be related to both educational programmes and scientific studies ontologies. In this case, such links allow using subnodes (keywords, methods, etc.) to find scientific studies related to this term. The exact words to provide interoperability between educational programmes ontology and scientific studies ontologies are shown in figure 10. As can be seen, the terms "methane" and "biogas" are related to both educational programmes and scientific studies and, therefore, they are used to link these ontologies.

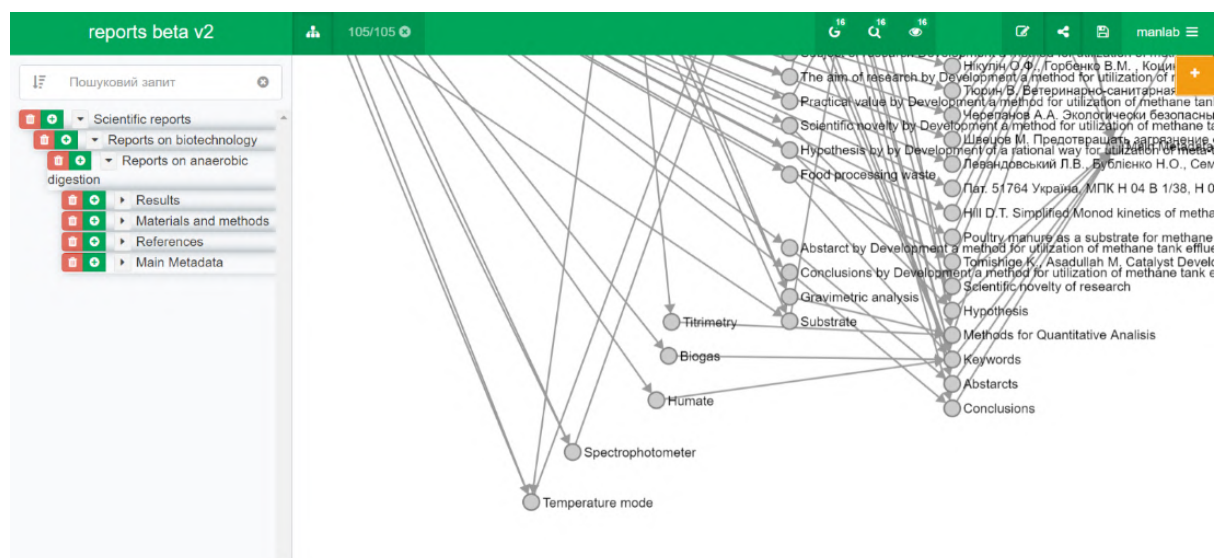


Figure 8: Terms of studies that may be used to link ontology of scientific studies with ontology of educational programmes.

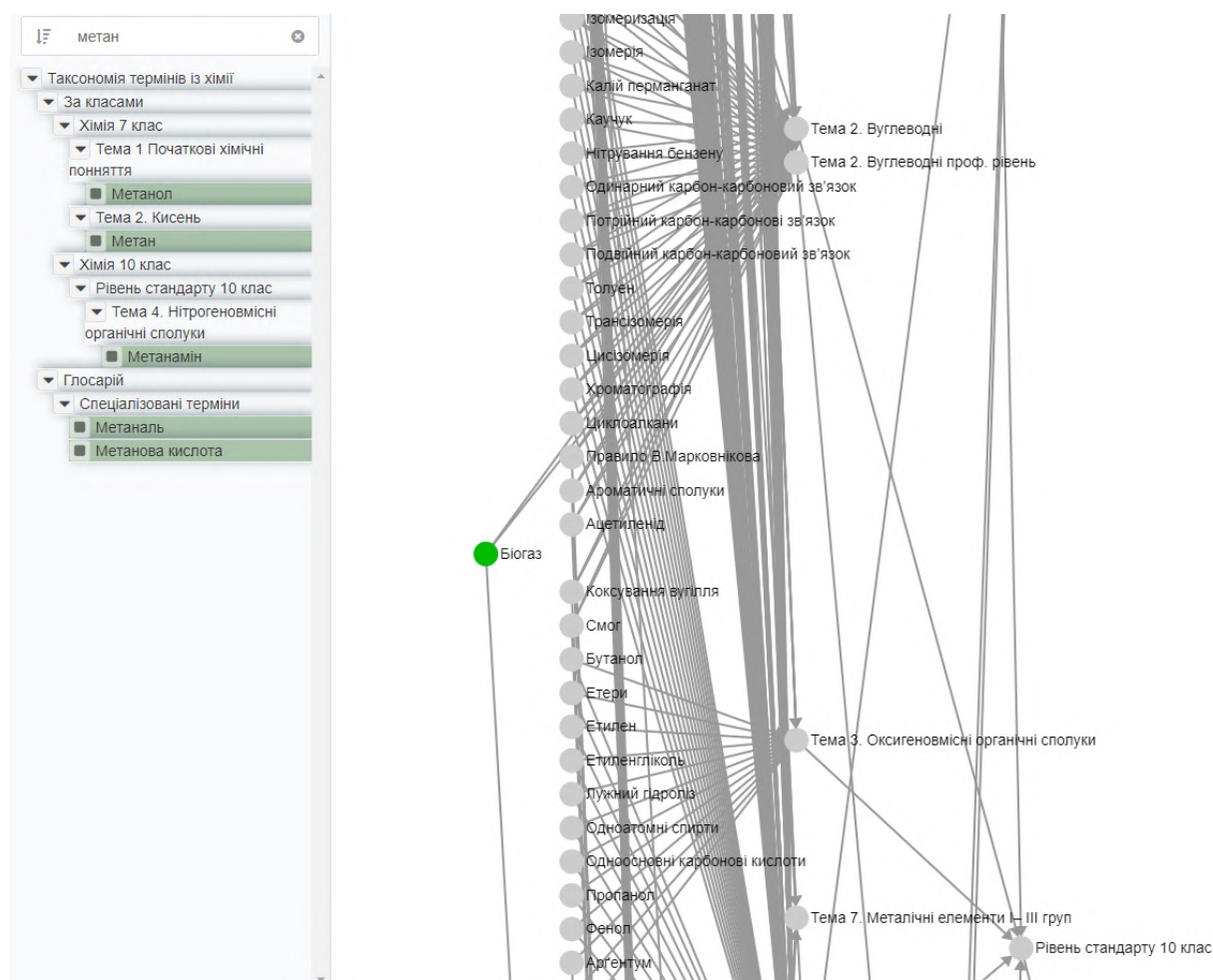


Figure 9: The general view of educational programmes' ontology related to chemistry educational programmes in Ukraine and terms that may be used to link with ontology of scientific studies.

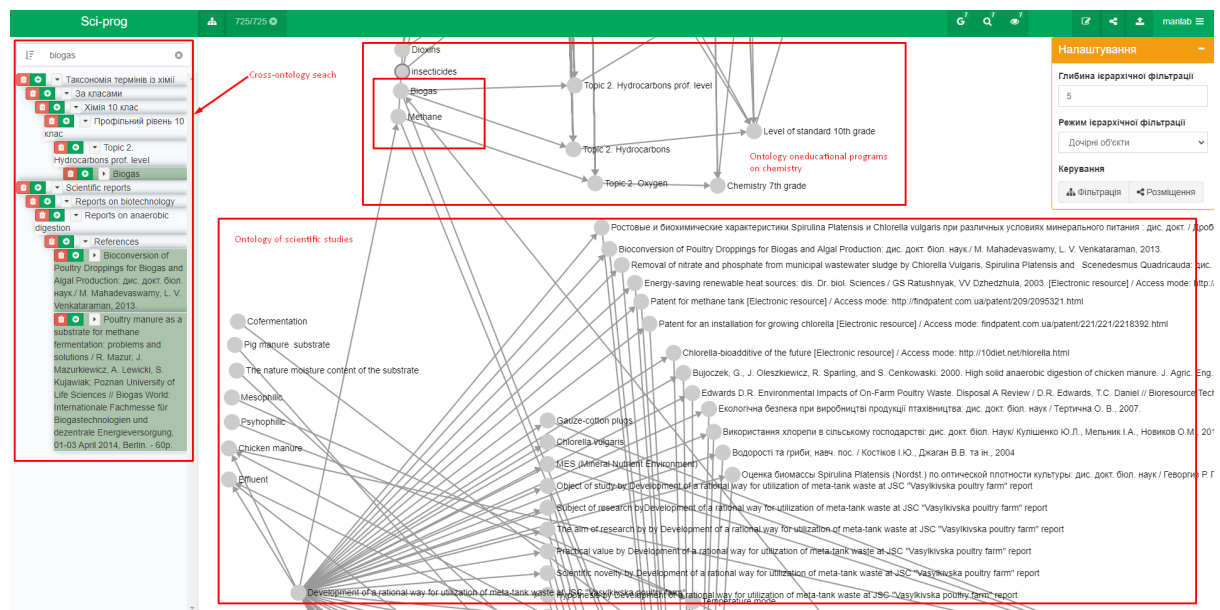


Figure 10: Using of same terms to provide interoperability between educational programmes ontology and scientific studies ontologies.

6. Discussion

The proposed database follows the “Leiden Manifesto for 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 straightforward. 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. It can provide a base assessment of individual researchers’ qualitative judgement of their portfolios. Because all ontological graphs are validated by experts, in this way, it is possible to avoid misplaced concreteness, including false precision, and recognise the systemic effects of all assessments and indicators. In addition, indicators can be scrutinised regularly and updated in the obtained ontological database. Furthermore, the proposed ontology-based research reports can be integrated into a single environment – ontology repositories, as suggested before [39].

The process starts with paper creation. For this stage, we can use various text editors, for example, Word or Google Docs. Then an 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 to the graph. In our case, the IT Platform Polyhedron is used for this. And last but not least, it is possible to use the “Alternative” system, which includes Audit, Filtering, and Ranking instruments. All proposed ontology tools are illustrated in the workflow diagram figure 11.

It is worth mentioning that this methodology of the centralised information web-oriented educational environment of Ukraine has been developed, and with the ontological approach is more systematic now. Educational programmes are essential to the world picture that is given to people during education, so they contain all basic terms that may be used to systematise other fields of human activities, including scientific studies. Like researchers, pupils interested in terms can also use such specific term nodes to continue their studying by investigating the studies conducted.

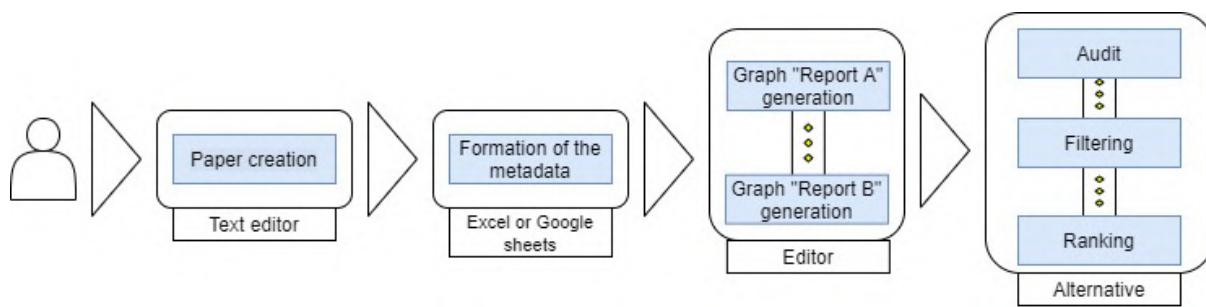


Figure 11: Workflow diagram of the creation of structured ontologies on scientific reports and their processing.

7. Conclusions

An ontological approach to scientific work systematisation has been proposed, assuring compatibility. A system for arranging research reports based on digital taxonomies (ontologies) has been created. It allows users to construct node hierarchies utilising the natural structure of the reports. Concrete parameters were added to the nodes as metadata (semantic, numeric, images, and links) to enable Polyhedron tools processing. Ranging and filtering were employed to handle semantic and numerical metadata. The obtained results allow for interchange across various study reports (including educational). The “Leiden Manifesto for Scientometrics” is the acknowledged ontological method.

Further study will improve interoperability across research works by developing a single taxonomy that provides hierarchisation using the same methodologies, literature, and report findings and its processing using both methods suggested in the research and newly developed ones.

For the first time, it presents the concept of integration of scientific studies ontologies with educational programmes that make them more usable for both students and young researchers. The proposed approach aligns with the centralised informational web-oriented educational environment concept.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

- [1] L. Oriokot, W. Buwembo, I. G. Munabi, S. C. Kijjambu, 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 (2011) 250. doi:10.1186/1756-0500-4-250.
- [2] Paperpile, What are the different types of research papers?, 2019. URL: <https://paperpile.com/g/types-of-research-papers/>.
- [3] A. Ramesh Babu, Y. P. Singh, Determinants of research productivity, *Scientometrics* 43 (1998) 309–329. doi:10.1007/BF02457402.
- [4] G. Abramo, C. A. D’Angelo, F. Di Costa, University-industry research collaboration: a model to assess university capability, *Higher Education* 62 (2011) 163–181. doi:10.1007/s10734-010-9372-0.
- [5] M. Pianta, D. Archibugi, Specialization and size of scientific activities: A bibliometric analysis of advanced countries, *Scientometrics* 22 (1991) 341–358. doi:10.1007/BF02019767.
- [6] M. E. J. Newman, The structure of scientific collaboration networks, *Proceedings of the National Academy of Sciences* 98 (2001) 404–409. doi:10.1073/pnas.98.2.404.
- [7] T. Braun, W. Glänzel, A. Schubert, Publication and cooperation patterns of the authors of neuroscience journals, *Scientometrics* 51 (2001) 499–510. doi:10.1023/A:1019643002560.
- [8] S. M. Lawani, Some bibliometric correlates of quality in scientific research, *Scientometrics* 9 (1986) 13–25. doi:10.1007/BF02016604.
- [9] A. A. Khasseh, F. Soheili, H. S. Moghaddam, A. M. Chelak, Intellectual structure of knowledge in iMetrics: A co-word analysis, *Information Processing & Management* 53 (2017) 705–720. doi:10.1016/j.ipm.2017.02.001.

- [10] L. Kostenko, A. Zhabin, A. Kuznetsov, T. Lukashevich, E. Kukharchuk, T. Simonenko, *Scientometrics : A Tool for Monitoring and Support of Research, Science and Science of Science* (2015) 88–94. URL: http://nbuv.gov.ua/UJRN/NNZ_2015_3_12.
- [11] K. R. Mulla, Identifying and mapping the information science and scientometrics analysis studies in India (2005-2009): A bibliometric study, *Library Philosophy and Practice* (2012) 772. URL: <https://www.researchgate.net/publication/215709587>.
- [12] S. Ravikumar, A. Agrahari, S. N. Singh, Mapping the intellectual structure of scientometrics: a co-word analysis of the journal *Scientometrics* (2005–2010), *Scientometrics* 102 (2015) 929–955. doi:10.1007/s11192-014-1402-8.
- [13] I. S. Pavlovskiy, Using Concepts of Scientific Activity for Semantic Integration of Publications, *Procedia Computer Science* 103 (2017) 370–377. doi:10.1016/j.procs.2017.01.123, XII International Symposium Intelligent Systems 2016, INTELS 2016, 5-7 October 2016, Moscow, Russia.
- [14] B. E. Perron, B. G. Victor, D. R. Hodge, C. P. Salas-Wright, M. G. Vaughn, R. J. Taylor, Laying the foundations for scientometric research: A data science approach, *Research on Social Work Practice* 27 (2017) 802–812. doi:10.1177/1049731515624966.
- [15] M. C. Ramírez, R. A. Rodríguez Devesa, A scientometric look at mathematics education from Scopus database, *The Mathematics Enthusiast* 16 (2019) 37–46. doi:10.54870/1551-3440.1449.
- [16] O. Kinouchi, L. D. Soares, G. C. Cardoso, A simple centrality index for scientific social recognition, *Physica A: Statistical Mechanics and its Applications* 491 (2018) 632–640. doi:10.1016/j.physa.2017.08.072.
- [17] S. Milojević, L. Leydesdorff, Information metrics (*iMetrics*): a research specialty with a socio-cognitive identity?, *Scientometrics* 95 (2013) 141–157. doi:10.1007/s11192-012-0861-z.
- [18] M. Amami, R. Faiz, F. Stella, G. Pasi, A Graph Based Approach to Scientific Paper Recommendation, in: *Proceedings of the International Conference on Web Intelligence, WI '17, Association for Computing Machinery, New York, NY, USA, 2017*, p. 777–782. doi:10.1145/3106426.3106479.
- [19] D. Boughareb, A. Khobizi, R. Boughareb, N. Farah, H. Seridi, A Graph-Based Tag Recommendation for Just Abstracted Scientific Articles Tagging, *International Journal of Cooperative Information Systems* 29 (2020) 2050004. doi:10.1142/S0218843020500045.
- [20] N. Perraudin, Graph-based structures in data science: fundamental limits and applications to machine learning, Ph.D. thesis, École Polytechnique Fédérale De Lausanne, 2017. URL: <https://infoscience.epfl.ch/record/227982?ln=en>.
- [21] D. Parveen, A Graph-based Approach for the Summarization of Scientific Articles, Ph.D. thesis, Ruprecht-Karls-Universität Heidelberg, 2018. doi:10.11588/heidok.00027924.
- [22] R. Alnemr, A. Paschke, C. Meinel, Enabling Reputation Interoperability through Semantic Technologies, in: *Proceedings of the 6th International Conference on Semantic Systems, I-SEMANTICS '10, Association for Computing Machinery, New York, NY, USA, 2010*. doi:10.1145/1839707.1839723.
- [23] O. Y. Stryzhak, V. V. Horborukov, O. V. Franchuk, M. A. Popova, Ontology selection problem and its application to the analysis of limnological systems, *Ecological safety and nature management* 15 (2014) 172–183. URL: http://nbuv.gov.ua/UJRN/ebpk_2014_15_21.
- [24] L. Globa, M. Kovalskyi, O. Stryzhak, Increasing Web Services Discovery Relevancy in the Multi-ontological Environment, in: A. Wiliński, I. E. Fray, J. Pejaś (Eds.), *Soft Computing in Computer and Information Science*, volume 342 of *Advances in Intelligent Systems and Computing*, Springer International Publishing, Cham, 2015, pp. 335–344. doi:10.1007/978-3-319-15147-2_28.
- [25] L. S. Globa, S. Sulima, M. A. Skulysh, S. Dovgyi, O. Stryzhak, Architecture and Operation Algorithms of Mobile Core Network with Virtualization, in: J. H. Ortiz (Ed.), *Mobile Computing*, IntechOpen, Rijeka, 2019. doi:10.5772/intechopen.89608.
- [26] O. Stryzhak, V. Horborukov, V. Prychodniuk, O. Franchuk, R. Chepkov, Decision-making System Based on The Ontology of The Choice Problem, *Journal of Physics: Conference Series* 1828 (2021) 012007. doi:10.1088/1742-6596/1828/1/012007.
- [27] V. B. Shapovalov, Y. B. Shapovalov, Z. I. Bilyk, A. I. Atamas, R. A. Tarasenko, V. V. Tron, Centralized information web-oriented educational environment of ukraine, *CTE Workshop Proceedings* 6

- (2019) 246–255. doi:10.55056/cte.383.
- [28] M. Popova, O. Stryzhak, Ontological interface as a means of presenting information resources in the GIS, *Uchenye zapiski Tavricheskogo natsionalnogo universiteta imeni V. I. Vernadskogo. Seriya "Geografiia"* 26(65) (2013) 127–135. URL: <https://tinyurl.com/5fhtrmck>.
- [29] V. Velychko, M. Popova, V. Prykhodniuk, O. Stryzhak, TODOS – IT-platform formation transdisciplinary information environment, *Systems of Arms and Military Equipment* (2017) 10–19. URL: http://nbuv.gov.ua/UJRN/soivt_2017_1_4.
- [30] R. Volckmann, Transdisciplinarity: Basarab Nicolescu Talks with Russ Volckmann, *Integral Review* (2007) 73–90. URL: <https://tinyurl.com/39y6fdm3>.
- [31] B. Nicolescu (Ed.), *Transdisciplinarity: Theory and Practice*, 1 ed., Hampton Press, 2008.
- [32] I. Slipukhina, S. Kuzmenkov, N. Kurilenko, S. Mienailov, H. Sundenko, Virtual Educational Physics Experiment as a Means of Formation of the Scientific Worldview of the Pupils, in: V. Ermolayev, F. Mallet, V. Yakovyna, H. C. Mayr, A. Spivakovsky (Eds.), *Proceedings of the 15th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer. Volume I: Main Conference, Kherson, Ukraine, June 12-15, 2019*, volume 2387 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 318–333. URL: <http://ceur-ws.org/Vol-2387/20190318.pdf>.
- [33] Y. O. Modlo, S. O. Semerikov, E. O. Shmeltzer, Modernization of Professional Training of Electromechanics Bachelors: ICT-based Competence Approach, in: A. E. Kiv, V. N. Soloviev (Eds.), *Proceedings of the 1st International Workshop on Augmented Reality in Education*, Kryvyi Rih, Ukraine, October 2, 2018, volume 2257 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018, pp. 148–172. URL: <https://ceur-ws.org/Vol-2257/paper15.pdf>.
- [34] Y. O. Modlo, S. O. Semerikov, P. P. Nechypurenko, S. L. Bondarevskyi, O. M. Bondarevska, S. T. Tolmachev, The use of mobile Internet devices in the formation of ICT component of bachelors in electromechanics competency in modeling of technical objects, *CTE Workshop Proceedings 6* (2019) 413–428. doi:10.55056/cte.402.
- [35] Z. I. Bilyk, Y. B. Shapovalov, V. B. Shapovalov, A. P. Megalinska, S. O. Zhadan, F. Andruszkiewicz, A. Dołhańczuk-Śródka, P. D. Antonenko, Comparison of Google Lens recognition performance with other plant recognition systems, *Educational Technology Quarterly* 2022 (2022) 328–346. doi:10.55056/etq.433.
- [36] K. Vlasenko, O. Chumak, I. Lovianova, D. Kovalenko, N. Volkova, Methodical requirements for training materials of on-line courses on the platform “Higher school mathematics teacher”, *E3S Web of Conferences* 166 (2020) 10011. doi:10.1051/e3sconf/202016610011.
- [37] Y. B. Shapovalov, V. B. Shapovalov, Z. I. Bilyk, I. M. Shapovalova, Structurization of educational expedition studies in the form of taxonomies, *Educational Dimension* 7 (2022) 130–149. doi:10.31812/educdim.7618.
- [38] S. Zhadan, Y. Shapovalov, R. Tarasenko, A. Salyuk, Development Of An Ammonia Production Method For Carbon-Free Energy Generation, *Eastern-European Journal of Enterprise Technologies* 5 (2021) 66–75. doi:10.15587/1729-4061.2021.243068.
- [39] A. Paschke, R. Schäfermeier, OntoMaven - Maven-Based Ontology Development and Management of Distributed Ontology Repositories, in: G. J. Nalepa, J. Baumeister (Eds.), *Synergies Between Knowledge Engineering and Software Engineering*, volume 626 of *Advances in Intelligent Systems and Computing*, Springer International Publishing, Cham, 2018, pp. 251–273. doi:10.1007/978-3-319-64161-4_12.