

Ontology-based learning environment model of scientific studies

Roman A. Tarasenko¹, Stanislav A. Usenko¹, Yevhenii B. Shapovalov¹, Viktor B. Shapovalov¹, Adrian Paschke² and Iryna M. Savchenko¹

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

Nowadays, there is a wide variety of scientific articles. Due to this fact, it is hard to read and be familiar with all of them. Also, it is hard for a young scientist to understand the complicated terms and methods that are used in a specific research domain. This problem was partially solved by bibliographic management software and other specific software. This article is devoted to the development of an approach for structuration and processing sets of studies using the IT Platform Polyhedron using an ontology-based hierarchical model. In its structure, the ontological graph is complex because it has additional branches from child nodes. The basis of our solution was IMRAD which has been represented in the view of nodes. Those nodes have been connected with specific representations of IMRAD elements. Specific articles have been represented in the view of leaf nodes. That could help to use the taxonomies for the structuration of the articles. Each data block is in the form of separate attributes of the ontological node. The proposed solution allows to obtain structured sets of studies and to separate their characteristics. Thus, the proposed ontology provides the possibility to view all methods, measured parameters, etc. of the studies in a graph node structure and use them to find the studies where they were used.

Keywords

cognitive IT-platform Polyhedron, ontology, ontological tool, scientific studies, scientific reports, learning environments

1. Introduction

Usage of information technologies (IT) in various fields of research activities and the capability of software support in science to automatically classify and structure information, e.g.,

CoSinE 2021: 9th Illia O. Teplytskyi Workshop on Computer Simulation in Education, co-located with the 17th International Conference on ICT in Education, Research, and Industrial Applications: Integration, Harmonization, and Knowledge Transfer (ICTERI 2021), October 1, 2021, Kherson, Ukraine

✉ tarasenko@man.gov.ua (R. A. Tarasenko); sjb@man.gov.ua (Y. B. Shapovalov); svb@man.gov.ua (V. B. Shapovalov); paschke@inf.fu-berlin.de (A. Paschke); SavchenkoI@nas.gov.ua (I. M. Savchenko)

🌐 <https://www.scopus.com/authid/detail.uri?authorId=57211134149> (R. A. Tarasenko);

<https://www.scopus.com/authid/detail.uri?authorId=57224619975> (S. A. Usenko);

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

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

<https://dblp.org/pid/24/2942.html> (A. Paschke);

<https://www.nas.gov.ua/EN/PersonalSite/Statuses/Pages/default.aspx?PersonID=0000020650> (I. M. Savchenko)

🆔 0000-0001-5834-5069 (R. A. Tarasenko); 0000-0002-0440-928X (S. A. Usenko); 0000-0003-3732-9486

(Y. B. Shapovalov); 0000-0001-6315-649X (V. B. Shapovalov); 0000-0003-3156-9040 (A. Paschke); 0000-0002-0273-9496

(I. M. Savchenko)

© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

in publication data, becomes increasingly important. Nowadays, there are huge amounts of research data available that isn't structured, e.g., publications, presentations, etc. It is complicated for young researchers and scientists to use such large amounts of publication data. During the research process, young scientists are looking for, e.g., examples of research methods and parameters. However, this task is challenging at their early stage of the scientific career. Such literature search and analysis problems (e.g. state of art analysis) are challenging for every scientist (including youth, school researchers) during the process of preparing papers and reports. For instance, according to Lens.org, the number of articles on biogas in 2002 was approximately 134, then in 2014, the number grew up to almost 1164, as shown in figure 1.

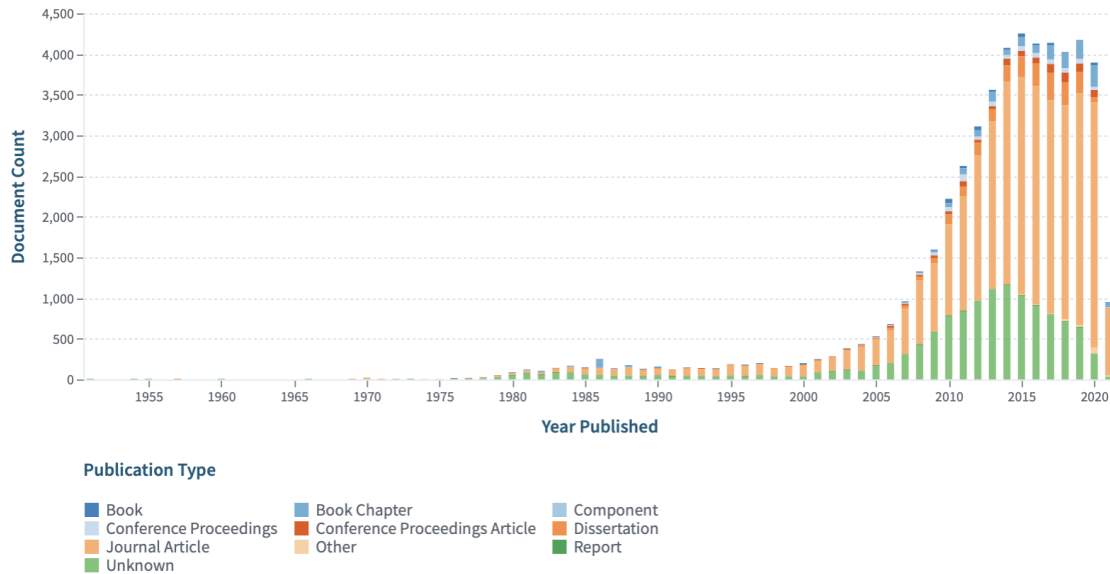


Figure 1: Dynamic of the number of the papers on biogas.

So, it seems relevant to provide a solution that can simplify processing and information/knowledge extraction in scientific publications. There are two hypotheses in our study. The first one is about structuring and digitalization of the data, which can simplify finding the details about the research method. The second one is about structuring the results of previous studies, which can be represented as data of the informational system. Previously, this goal was partially achieved using metadata for data processing. In this paper we further contribute with a semantic ontology and more expressive semantic metadata approach.

1.1. Using metadata to provide data management in science publications

To support publication data management, it is relevant to use metadata about each paper. This metadata represents the data about the publication. In this case, relevant information about each specific publication can be represented by the metadata. Metadata can include, e.g., contact information, year of publication, author details, instrument and protocol information, survey tool details and much more [1].

For instance, reference management software maintains a database of articles and creates bibliographies and reference lists for the written works. This software simplifies the record of metadata. There are several popular reference management software, for example, Refworks, Mendeley, EndNote and Zotero [2, 3, 4, 5]. All of these managers can save profiles, build a database of citations, save PDF files and extract metadata from them, import references from library catalogues, websites, and other citation managers [6, 7, 8].

However, these systems only use limited metadata vocabularies without expressive semantic models. For instance, such systems do not support metadata concepts such as “Results”, “Materials and methods”, “References” etc. All these systems do not provide a systematic approach, they are not completely semantically structured and not hierarchical.

1.2. Methods for automatic literature review processing

There are existing different types of metadata information that we can use to structure the articles, for example, by the relation to source, by the function, by the purpose, by the language, by the time:

- By the “Relation to Source”: During this method, the user defines the type of text that can be included in a classification program. A classification that is composed of extracts having exact sentences of a source document is known as an extractive summary. That is the simplest type of classifier.
- By the “Function”: with this method, user can use any helpful and relevant information from source documents, for example, an abstract of a scientific article or the reviewers’ opinion on the quality of work.
- By the “Purpose”: This method structures the article by its purpose or main idea. The user needs to write down the general-purpose or sense of the text in the program by himself.
- By the “Language”: A classifier can be monolingual or multilingual. The monolingual classifier uses only one language and produces an output classification in the same language as the input document whereas, the multilingual classifier uses multiple languages and gives an output classification in one of the languages from the input document.
- By the “Time”: It is possible to arrange the articles by the time of their publication. To do so, the user must enter the publication date of the article in the system, then the system will arrange the article in an appropriate section.

In our opinion, there is a lack of methods that we can use for the structurization of scientific articles. Also, it is necessary to add “By the Results” and “By the methods” methods.

- “By the Results” this method structures the article by its results or conclusions. The user needs to write down the outcomes or verdicts of the text in the program by himself.
- “By the Methods” this method structures the article by the scientific methods or experiments that are used in the essay.

Considering that most of the articles have a typical IMRAD (Introductions, Methods, Results, Abstract, Discussion) structure, it seems advisable to build an algorithm that uses data of many specific articles to create ontological graphs, that can be integrated in specialized educational environments for young scientists.

1.3. Instruments for creation of ontology-based learning environments

A learning environment is a diverse platform where users engage and interact to learn new skills. While learners can learn in various settings, the term typically refers to a digital alternative for the traditional classroom. To improve learning efficiency and adaptability, formalized information resources that provide a high degree of structuring should be used in learning. An ontological approach could support this. The ontological approach provides a holistic and systematic approach to the study of various information sources and a specific subject domain, ensures the conceptualization and taxonomization of terms within the subject area and the existence of relationships between the terms of different subject areas to ensure multidisciplinary. Computer ontologies are one of the effective mechanisms for ensuring a stable digital learning environment.

In recent years significant progress was made in developing ontologies. In this article, an “ontology” is a term that means a software or web system that consists of nodes with data. All ontology nodes are arranged in a specific hierarchical order, often referred to as an ontological tree or ontological graph. The node from which all branches start is called the root node. The other nodes are called subsidiaries.

One of the most perspective solution, in our opinion, is ontologies [9]. For example, we can use hierarchies with multicriteria techniques during the classification of metadata of various articles. Ontologies aim to capture the domain knowledge in a general way and ensure a common understanding of the domain.

IsaViz is a virtual environment for viewing and creating RDF models in the view graphs. IsaViz imports RDF/XML and N-Triples, and exports RDF/XML. Apollo is the program for modelling knowledge systems. Apollo knowledge system base consists of hierarchically organized ontologies that can be inherited from other ontologies. SWOOP contains OWL (Web Ontology Language) validation and offers various. OWL presentation syntax views. At SWOOP, Ontologies can be compared, edited, and combined. Protégé 3.5 is a knowledge-based ontology editor that provides a graphical user interface. It ensures better flexibility for metamodelling, enables the construction of domain ontologies.

1.4. Ontological problems

Nowadays most of the common systems (such as Mendeley, Scopus etc.) provide support for displaying of data but not for comparison and providing search functions. Also, given that articles in the same domain have the same indicators, the metadata of the results can be represented as ontology node attributes and then processed.

Previously, ontological graphs were used to systematize scientific articles [9, 10, 11, 12, 13]. Systematization and structuring in such ontological systems were based on different approaches such as using of scientific article recommendation system [10], a scientific articles tagging system [11], machine learning and automatic summarization [9]. None of the proposed ontological approaches [9, 10, 11, 12] can't provide a decent level of structurization and systematisation.

We have proposed to use the cognitive IT platform Polyhedron [14] for this aim. The core of the Polyhedron system consists of advanced and improved functions of the TODOS IT platform [14] described in previous works. The Polyhedron is a multiagent system that

provides transdisciplinary and interactivity in any study [15]. Besides, cognitive IT-platform Polyhedron contains a different variety of special functions like auditing [15, 16, 17], semantic web, information systematization and ranking [18] transdisciplinary support [19], internal search [20], and have all advantages of ontological interface tools [16, 21]. Due to active states are hyperratio plural partial ordering [19, 22], cognitive IT-platform Polyhedron is an innovative IT technology of ontological management of knowledge and information resources, regardless of the standards of their creation.

The proposed solution can be used with other applications in the field of structuration studies like a virtual educational experiment, provide STEM approach in education [14, 23], using the technology of augmented reality education [24, 25], educational and scientific environments [14, 20].

2. Materials and methods

2.1. Ontology creation mechanism

The proposed research is based on approach that was proposed before [26], but it provides management and structuring of set of the studies, not single. To create ontologies in cognitive IT-platform Polyhedron, Google Sheets are used to collect and structure the information (see example in figure 2). The sheets with study report data (structure file and numeric/semantic data file) are downloaded and saved in .xls format. The data was separated with expert involving and based on principle that researches of same field have similar input and resulting characteristics. Same is relevant to field of anaerobic digestion [27, 28] that was taken as example to provide structuring and processing using proposed method. The files are loaded to editor.stemua.science to provide graph generation (a part of the cognitive IT platform Polyhedron). After that, the generation of the graph nodes (in .xls) with its characteristics using structure file is carried out. The obtained graphs are saved in .xml format and located in the database. The graphs are filled by semantic and numeric information for ranking or filtering. Ontological nodes are formed using predicate equations.

2.2. Description of the works were used to demonstrate the structuration mechanism

To demonstrate the structuring of the educational study reports, a master thesis of the National University of Food technologies “Development a method for utilization of anaerobic digestion effluent at LLC Vasylykivska Poultry Farm” (in further – report “A”) and a study report prepared for the defence of an educational study in the Junior academy of sciences of Ukraine scientific competition “Development a method for utilization of anaerobic digestion effluent” (in further – report “B”) are used as input data. Two ontologies have been built, the structure of which duplicates the content of the study reports. All numeric data was separated to provide information processing and to provide integration between those works. Their titles have used as the parent node for the ontological graphs. All ontological graphs in the cognitive IT platform Polyhedron have been linked to each other by the mechanism of internal search. Further scientific studies will be linked in the graphs by edges (links) to provide better connection.

| | A | B | C | D | E | F | G | H | I |
|---|----------------------------------|--|--|--|---|--|--|---|---|
| 1 | Scientific reports | Reports on biotechnology | | | | | | | |
| 2 | Reports on biotechnology | Reports on anaerobic digestion | | | | | | | |
| 3 | Reports on anaerobic digestion | Main Metadata | Materials and methods | References | Results | | | | |
| 4 | Results | Final results | Input parameters of the research | | | | | | |
| 5 | Materials and methods | Materials for researching | Methods for Quantitative Analysis | | | | | | |
| 6 | References | Використання хлорелі в сільському господарстві: дис. докт. біол. наук/ Куліщенко Ю.Л., Мельник І.А., Новиков О.М., 2011. | Водорості та гриби; навч. пос. / Костіков І.Ю., Джаган В.В. та ін., 2004 | Екологічна безпека при виробництві продукції птахівництва: дис. докт. біол. наук / Тертична О. В., 2007. | Оценка биомассы Spirulina Platensis (Nordst.) по оптической плотности культуры: дис. докт. биол. наук / Геворгиз Р. Г., 2003. | Ростовые и биохимические характеристики Spirulina Platensis и Chlorella vulgaris при различных условиях минерального питания : дис. докт. / Дробецька И. В., 2001. | Bioconversion of Poultry Droppings for Biogas and Algal Production: дис. докт. біол. наук/ M. Mahadevaswamy, L. V. Venkataraman, 2013. | Removal of nitrate and phosphate from municipal wastewater sludge by Chlorella Vulgaris, Spirulina Platensis and Scenedesmus Quadricauda: дис. докт. біол. наук/ Jalal K.S., Kamaruzzaman B.Y, Akbar J. | |
| 7 | Main Metadata | Object of research | Subject of research | The aim of the research | Practical value of research | Scientific novelty of research | Hypothesis | Keywords | |
| 8 | Input parameters of the research | Temperature mode | Moisture content of substrate | Substrate | | | | | |
| 9 | Materials for researching | Glassware for chemistry lab | Gauze-cotton plugs | Chlorella vulgaris | MES (Mineral Nutrient Environment) | Chicken manure | Spectrophotometer | Effluent from methane tank (fertilizer) from JSC | |

Figure 2: Google sheet with data.

2.3. Ranking tools

Study reports “A” and “B” were also compared with each other by the following criteria: “Short-term economic perspective”, “Long-term economic prospects” using the ranking tool. For creation of the ranking ontologies have been used by the module “Alternative” which is described in previous works [18, 29]. The nodes of the 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.

2.4. Auditing tools

To provide an audit, 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 study, keywords, and other parameters, of the studies 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 presented with the names of the works which need to be audited with the “standard” graph. The parameters of the work used to audit 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 “standard” graph to provide interaction between graphs.

3. Results

3.1. Model of creation ontology to systemizing

As was noted before, IMRAD is widely used to prepare research and science papers. It is possible to provide structuration by the usage of IMRAD components as parent nodes. So, the parent nodes can be represented by Introduction, Methods, Results and Discussion. The discussion part can't be easily structured by an ontology. The most contractionary part that it contains are the analysis and the comparison of the obtained data by the researcher. Specific parts of IMRAD will be used as branch nodes, and the study will be used as a leaf node. So, the general structure of the ontology that structures the research data is represented as:

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

where REP – set of reports, I – sets of Introduction of all study, M – set of Methods of all study, R – set of Results of All study, P – instruments of processing of the results of set of studies discussions.

To provide better systematization we have split the introduction into two different parts – basic metadata and literature review:

$$I = \langle BMD, LR \rangle \quad (2)$$

where BMD – is a set of basic metadata of study, LR – a set of Sources used for Literature Review.

The basic metadata node of the study is linked with the graph's leaf nodes that characterized basic data on the study such as hypothesis, object, subject, practical value, and scientific novelty, etc. So, the nodes of the report's basic metadata of the study can be presented as a further equation:

$$BMD = \langle H, O, S, PV, SN \rangle \quad (3)$$

where H – hypothesis or hypotheses of each specific study; O – object of the study of each specific study; S – subject of the of each specific study; PV – practical value of each specific study; SN – scientific novelty of each specific study.

The main advantages of using such a structure are that some parts of the introduction materials and methods and results (measured parameters) of study (reports) can coincide. A few specific studies that coincide will be linked by nodes (in case of methods and results) or by classes of data (in case of keywords or scientific novelty) due to using the same subnodes of the ontology. Let's represent each work as a set of the Introduction, Methods, Results, and Processing of the data (Discussion):

$$REP_I = \langle I_I, M_I, R_I, P_I \rangle \quad (4)$$

$$REP_{II} = \langle I_{II}, M_{II}, R_{II}, P_{II} \rangle \quad (5)$$

So, these studies can be integrated in a single ontology using IMRAD:

$$O = \langle S_I, S_{II} \rangle = \langle I_I, M_I, R_I, P_I, I_{II}, M_{II}, R_{II}, P_{II} \rangle \quad (6)$$

The same approach will be applied to each element of the IMRAD structure study. Generally, it can be represented as:

$$M = (REP_i) = \sum_i^n M_I \quad (7)$$

where M_i – every separated scientific method.

In a different study, a different set of methods can be used. However, some of them can coincide. Thus, set of methods used in two different study may be represented as:

$$M = (REP_I) = \{M_a, M_b, M_c, M_d\} \quad (8)$$

$$M = (REP_{II}) = \{M_b, M_d, M_f\} \quad (9)$$

And, so, M_b coinciding to both M_I and M_{II} :

$$M_b \in \{M_I, M_{II}\} \quad (10)$$

Therefore, in this case, and M_b can be used as a parent node that connects two different studies. The node M_b itself will contain general theoretic information on it, and node REP_I and REP_{II} will contain information on the specific case of its usage and measured parameters using it.

Similar mechanism can be provided by using specific ontology tools using metadata. For example, there will be a hierarchical approach for representing and usage of keywords:

$$Kw(BMD_i) = \text{Class "Keywords" Type "Array Value"} \{Kw_a, Kw_b, Kw_c, Kw_d\}, \quad (11)$$

where $Kw(BMD_i)$ – node of the basic metadata that integrates all keywords; Kw_i – specific keyword.

Also, as was noted in the introduction, the metadata of each work will be used for filtering the information, and for supporting specific processing functions of the IT solution Polyhedron. Such specifics mechanisms are AUDIT and RANKING. Metadata can be included in each node. For the parent node metadata will be used to represent the general information (for example, essence of the method itself), and the resulting leaf node will contain the specific metadata related to specific study (such as specific results of the study obtained using set methods M; for example, metadata: 5.35, and its class: "Ammonium nitrogen content, g/L"). So, metadata, with the same class, will be processed by using filtering by users request or by ranking using the ranks of the nodes for specific classes (or their set) based on the user's request.

So, the proposed approach uses IMRAD to collect and process the data with ontologies. In this way, the ontologies are constructed not by the specific structure of each work but by the generally accepted IMRAD structure. The parent node will be a specific area set to which the study belongs ($A = \sum_i^n REP_{Ii}$, where A – specific area of set of REP). The A node is linked with I, M, R, P nodes (representing IMRAD). Each IMRAD node is linked with the specific IMRAD type node (such as ammonia determination by Nessler's method (for methods) or "chicken manure" or "glycerine" (for subjects)). And each specific IMRAD type node is linked with leaf nodes of ontology – specific studies where such entities were used.

3.2. Structuring several works simultaneously using an ontological graph

To demonstrate the capabilities of the proposed ontological system, scientific works on anaerobic digestion were chosen. The general view of the resulting graph is shown in figure 3.

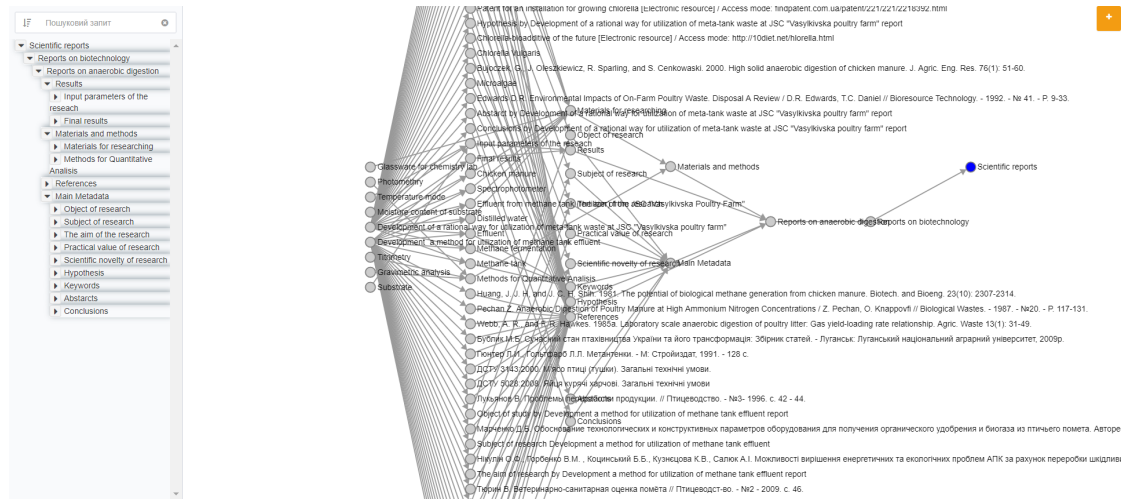


Figure 3: General view of the resulting ontological graph.

The root node of the resulting graph is the “Scientific reports” node. In its structure, the ontological graph is complex because it has additional branches from child nodes. Child nodes are: “Reports on biotechnology” and “Reports on anaerobic digestion.” From the child node “Reports on anaerobic digestion” are going the central sub-leaf nodes that reflect the basic principle of systematization of scientific works: “Results,” “Materials and methods,” “References”. This basic principle is shown in figure 4. A separate node of Main Metadata was also additionally created. This node contains the central metadata: object, subject of study; practical significance, the scientific novelty of study; hypotheses; keywords; abstract, conclusions.



Figure 4: General view of the main systematizing ontological vertices.

The entire sequence and principle of filling and maintenance of data by users in the received ontology are shown in the Workflow diagram (figure 5).

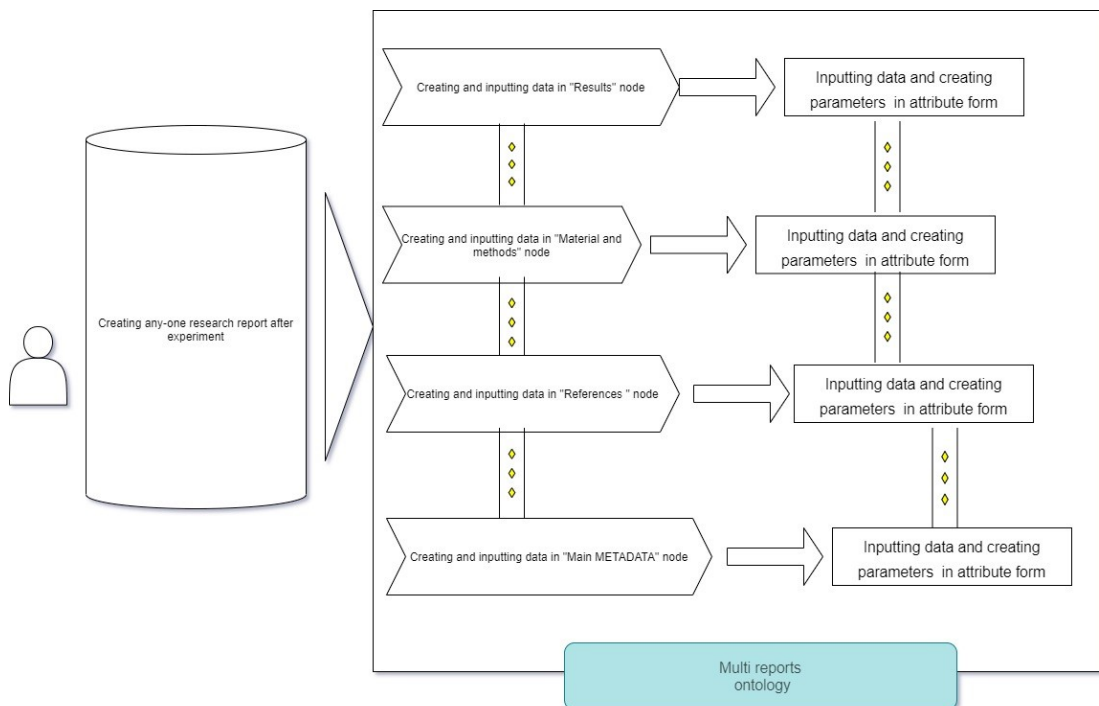


Figure 5: Workflow diagram of filling and maintenance of data by users.

The child nodes of each of these systematizing ontological nodes are the scientific works themselves. Each data block is in the form of separate attributes of the ontological node. This solution allows you to use all the information processing tools of the CIT Polyhedron system. In particular, such tools are both general (for example, filtering) and specialized, such as ranking and auditing. An example of filtering is shown in figure 6.

3.3. Application of ranking mechanism in the structuring of scientific works

All attributes can be used to rank information using the “Alternative module” described in previous works [18, 29]. Each nodes attributes is filled with numeric, textual, and/or mixed types of data. The following attributes are filled with text data: “References”, “Methods for Quantitative Analysis”, “Materials for researching”, “Thermophilic”, “Chicken manure substrate”, “Spectrophotometer parameters of the experiment”, “The actual rate of reproduction of”, “Keywords”, “Glassware”, “Reagents”, “Equipment”, “Object”, “Subject of study”, “The aim of the study”, “Chicken”. Numeric data contains the following attributes: “Initial pH”, “Methane content, % Vol.”, “pH of obtained solid product”, “Ammonium nitrogen concentration, mg/L”, “The concentration of volatile fatty acids (VFA) mg/L”, “The dry matter content, %”, “The ash content, %”. The attribute “The native moisture content of the substrate” contains mixed-type

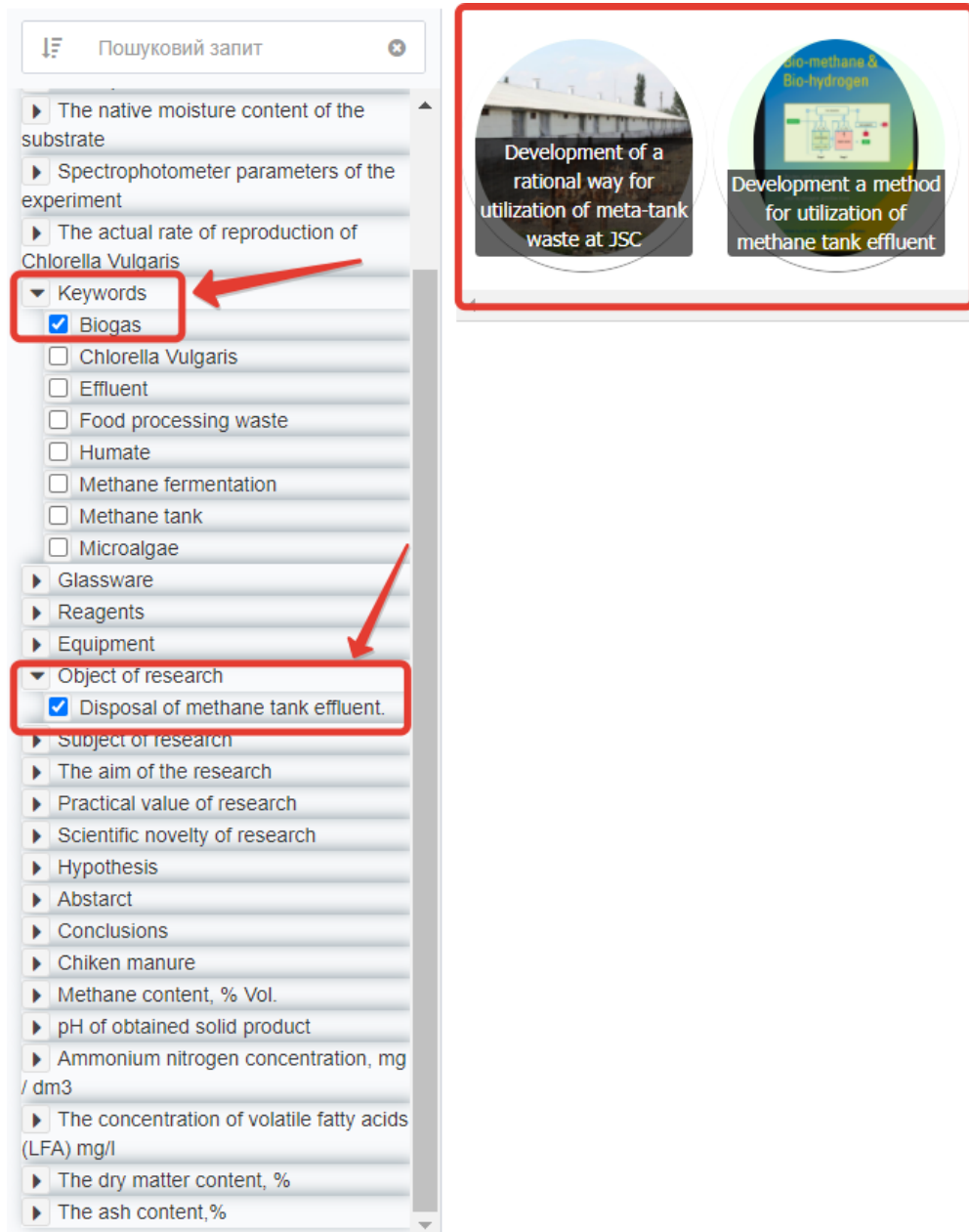


Figure 6: Example of attribute filter.

data, numeric and textual. An example of an incoming data maintenance panel for ranking is shown in figure 7.

For example, there may be a case when user wants to arrange work on the pH. The ranking result is shown in figure 8. Other examples of usage of the Polyhedron IT platform are shown in table 1.

| Назва критерію | Більша шкала | Лінгвістична шкала | Ранжування | Opt. 2 |
|--|--------------|--------------------|------------|--------|
| References | 1 | Мала важкість | 1 | max |
| Methods for Quantitative Analysis | 1 | Мала важкість | 1 | max |
| Materials for research | 1 | Мала важкість | 1 | max |
| Thermophilic | 1 | Мала важкість | 1 | max |
| Chicken manure substrate | 1 | Мала важкість | 1 | max |
| Initial pH | 1 | Абсолютна важкість | 1 | max |
| The native moisture content of the substrate | 1 | Мала важкість | 1 | max |
| Spectrophotometer parameters of the equipment | 1 | Мала важкість | 1 | max |
| The actual rate of reproduction of Clostridia Vulgaria | 1 | Мала важкість | 1 | max |
| Keywords | 1 | Мала важкість | 1 | max |
| Glasses | 1 | Мала важкість | 1 | max |
| Reagents | 1 | Мала важкість | 1 | max |
| Equipment | 1 | Мала важкість | 1 | max |
| Object of research | 1 | Мала важкість | 1 | max |
| Subject of research | 1 | Мала важкість | 1 | max |
| The aim of the research | 1 | Мала важкість | 1 | max |
| Chicken manure | 1 | Мала важкість | 1 | max |
| Methane content, % Vol | 1 | Мала важкість | 1 | max |
| pH of obtained solid product | 1 | Мала важкість | 1 | max |
| Ammonium nitrogen concentration, mg / dm³ | 1 | Мала важкість | 1 | max |
| The concentration of volatile fatty acids (LFA) mg/l | 1 | Мала важкість | 1 | max |
| The dry matter content, % | 1 | Мала важкість | 1 | max |
| The ash content, % | 1 | Мала важкість | 1 | max |

Figure 7: Example of a water panel for data ranking.

| Peirair | Keywords | Methods for Quantitative Analysis | Materials for researching | Equipment | Reagents | Thermophilic | Chicken manure substrate | Initial pH | The native moisture content of the substrate | Glassware | Chicken manure | Spectrophotometer parameters of the experiment | The actual rate of reproduction of Clostridia Vulgaria | Object of research | Subject of research | The aim of the research | Practical value of research | Scientific novelty of research |
|--|---|--|--|---|--|--|--|------------|---|--|--|--|--|-----------------------------------|--|--|--|--|
| Development of a method for utilization of methane tank effluent | Food processing waste Biogas Effluent | Analysis of ammonium nitrogen content with Nessler's reagent | Chicken manure Spectrophotometer Effluent from methane tank JSC "Varykivskia Pivnyy Farm" | Spectrophotometer Crucible (or similar porcelain or metal dishes) Muffled furnace | Distilled water Nessler's reagent HPO4 | The list of scientific reports where was used as this temperature mode for anaerobic | The list of scientific reports where was used as the substrate for | 73.9 | 250 ml volumetric flask List of works have used the substrate with not | 250 ml volumetric flask Conical flask (Erlenmeyer) for 200 ml Pipette ear 1 200 ml | List of works which have used chicken manure as the substrate for methane fermentation | Spectrophotometer parameters from the 1st to the 6th day of the experiment | Rate of reproduction of Clostridia Vulgaria | Disposal of methane tank effluent | Processing of methane tank effluent into humates by the autolysis method | To establish regularities of processing methane tank effluent into humates by the method of methane fermentation | The study indicates the possibility of obtaining humic and tannic acids by the method of this work | 1. Potassium humate was first obtained from methane tank effluent using microalgae |

Figure 8: Example of ranking results.

3.4. Application of the audit mechanism in structuring scientific works

Users can also use the specialized audit module described in previous works [15, 16, 17, 29] for all attributes. The graph “Standard” is the ontology itself, containing works that will be supplemented and expanded. This solution will allow users to automatically check whether there is a particular work in the database. Also, this solution will allow checking the hypotheses for compliance with already completed studies. Also, the audit module will allow users to compare an existing METADATA and attributes that are available in ontology at the same time. In particular, these attributes are the materials and methods of the results and the list of sources. Results that do not match the attributes of the “standard” ontology emphasize red. An example

Table 1
Examples of Polyhedron IT platform ranking module

| Name of filter (Ontological attributes) | Priority | Main results (list of ontological nodes) |
|--|----------|---|
| Initial pH (type of data are numbers) | Absolute | Development of a rational way for utilization of meta-tank waste at JSC "Vasylkivska poultry farm", Titrimetry, Methane tank, Development a method for utilization of methane tank effluent Methods for Quantitative Analysis, Materials for researching, Abstract |
| Methane content, % Vol. (type of data are numbers) | Absolute | Development a method for utilization of methane tank effluent, Development of a rational way for utilization of meta-tank waste at JSC "Vasylkivska poultry farm, Methods for Quantitative Analysis, Materials for researching, Abstract |
| The ash content, % | Absolute | Development of a rational way for utilization of meta-tank waste at JSC "Vasylkivska poultry farm, Development a method for utilization of methane tank effluent, Methods for Quantitative Analysis, Materials for researching, Equipment, Reagents, Thermophilic, Chicken manure substrate |
| The dry matter content, % | Absolute | Development a method for utilization of methane tank effluent, Development of a rational way for utilization of meta-tank waste at JSC "Vasylkivska poultry farm Keywords, Methods for Quantitative, Analysis, Materials for researching, Equipment, Thermophilic Chicken manure substrate |

of an audit fragment is shown in figure 9.

reports beta v2
(Audit)

| # | Indicators | Units | DSTU | Development of a rational way for utilization of meta-tank waste at JSC "Vasylkivska poultry farm" | Development a method for utilization of methane tank effluent |
|-----------------|------------|------------|---|---|--|
| Abstract | | | | | |
| 1 | Keywords | Keywords | microalgae | Microalgae | Humate |
| 2 | Hypothesis | Hypothesis | The effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae <i>Chlorella Vulgaris</i> . | The effluent obtained after anaerobic digestion can be used as a nutrient medium for microalgae <i>Chlorella Vulgaris</i> . | The solid fraction of methane fermentation of chicken manure can be recycled by the autocatalytic catalysis method. |

Figure 9: Example of audit results.

4. Discussion and conclusions

We have proposed to use IMRAD as the main approach to structure the articles in the form of a semantic ontology. As a proof of the concept, we have implemented a universal ontological solution that can provide systematization and structuration of any scientific studies. “Polyhedron” is not a reference manager like “Endnote”, “Mendeley”, and it is also not a scientometric database, but in the nearest perspective, it’s possible to convert our technology into a useful analogue. But, the advantages and potential scenarios for usage of our solution have been demonstrated by the example of biogas studies. The potential of using ranking and auditing tools in the obtained ontological database has been also shown. Numeric and semantic characteristics were separated from the main text and used to process by specialized algorithms of IT Platform Polyhedron. For example, users can find studies where specific method was used by both, using the structure and filtering of studies data. The numeric data of studies are processed by the ranking tool that assigns ranks to nodes depending on the value of these numeric characteristics. The created ontology allows to obtain the structured set of studies, separate their characteristics, provide the possibility to view all of the methods, measured parameters in the view of node and use them to find the studies where they were used. The detailed comparison of our ontological solution with most common analogues is presented on the table 2 below.

Table 2
Comparison of “Polyhedron” system with similar analogues

| | Polyhedron* | Mendeley | Endnote | Scopus | Google Scholar |
|---|-------------|----------|---------|---------|----------------|
| Automatic extraction of the information from any added PDFs | Present | Present | Present | Absent | Absent |
| Tags, keywords, or search the full text of most PDFs functions. | Present | Present | Present | Present | Present |
| Ability to cite articles in word/pages | Absent | Present | Present | Absent | Absent |
| Ability to use numeric data of the articles for ranking | Present | Absent | Absent | Absent | Absent |
| Accentuation of important semantic characteristics for management of the wide range of articles | Present | Absent | Absent | Absent | Absent |
| Ability to compare different articles | Present | Absent | Absent | Absent | Absent |
| Visualization of the information | Present | Absent | Absent | Absent | Absent |
| Usage of IMRAD approach to sort articles | Present | Absent | Absent | Absent | Absent |

* proposed approach using CIT “Polyhedron”

As we can see from the table above, our ontological solution has all of the basic functions of the most common software. Our solution can be used as a bibliographic software and as a scientometric database. In addition, our solution could provide such functions as ranking based on specific attributes, ability to compare different articles, visualization of the information in view of an ontological tree or “ontocubes” and usage of IMRAD approach to sort articles.

References

- [1] U. Noyer, D. Beckmann, F. Köster, Semantic technologies and metadata systematisation for evaluating time series in the context of driving experiments, in: 11th International Protégé Conference, 2009, pp. 1–2. URL: <https://protege.stanford.edu/conference/2009/abstracts/P9-Noyer.pdf>.
- [2] S. Kumar Basak, A Comparison of Researcher’s Reference Management Software: Refworks, Mendeley, and EndNote, *Journal of Economics and Behavioral Studies* 6 (2014) 2220–6140. URL: <https://pdfs.semanticscholar.org/1d41/58017608c841c89794fc9fde70e1c061948e.pdf>.
- [3] V. Kumar, D. Sharma, Cloud Computing as a Catalyst in STEM Education, *International Journal of Information and Communication Technology Education* 13 (2017) 38–51. doi:10.4018/IJICTE.2017040104.
- [4] J. Salem, P. Fehrmann, Bibliographic Management Software: A Focus Group Study of the Preferences and Practices of Undergraduate Students, *Public Services Quarterly* 9 (2013) 110–120. doi:10.1080/15228959.2013.785878.
- [5] S. Ram, K. John Paul Anbu, The use of bibliographic management software by Indian library and information science professionals, *Reference Services Review* 42 (2014) 499–513. doi:10.1108/RSR-08-2013-0041.
- [6] C. Ivey, J. Crum, Choosing the right Management tool, *Journal of the Medical Library Association* 106 (2018) 399–403.
- [7] A. Butros, S. Taylor, Managing information: evaluating and selecting citation management software, a look at EndNote, RefWorks, Mendeley and Zotero, *Proceedings of the 36th IAMSLIC Conference: Mar del Plata, Argentina, 17-21 October, 2010* (2011) 53–66. URL: <http://hdl.handle.net/1912/4595>.
- [8] S. Cuschieri, V. Grech, N. Calleja, WASP (Write a Scientific Paper): The use of bibliographic management software, *Early Human Development* 128 (2019) 118–119. doi:10.1016/j.earlhumdev.2018.09.012.
- [9] D. Parveen, A Graph-based Approach for the Summarization of Scientific Articles, Ph.D. thesis, Ruprecht-Karls-Universität Heidelberg, 2018.
- [10] M. Amami, R. Faiz, F. Stella, G. Pasi, A graph based approach to scientific paper recommendation, in: *Proceedings - 2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017, 2017*, pp. 777–782. doi:10.1145/3106426.3106479.
- [11] 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.
- [12] 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>.
- [13] Y. Poulakakis, K. Vassilakis, M. Kalogiannakis, S. Panagiotakis, Ontological modeling of educational resources: a proposed implementation for Greek schools, *Education and Information Technologies* 22 (2017) 1737–1755. doi:10.1007/s10639-016-9511-z.
- [14] Y. B. Shapovalov, V. B. Shapovalov, V. I. Zaselskiy, TODOS as digital science-support environment to provide STEM-education, *CEUR Workshop Proceedings* 2433 (2019) 232–245.

- [15] O. Stryzhak, V. Gorborukov, O. Franchuk, M. Popova, Ontology of the choice problem and its application in the analysis of limnological systems, *Ecological safety and nature management* (2014) 172–183.
- [16] 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*, Springer International Publishing, Cham, 2015, pp. 335–344. doi:10.1007/978-3-319-15147-2_28.
- [17] L. Globa, S. Sulima, M. Skulysh, S. Dovgyi, O. Stryzhak, Architecture and Operation Algorithms of Mobile Core Network with Virtualization, in: J. H. Ortiz (Ed.), *Mobile Computing*, IntechOpen, 2020, pp. 1–22. doi:10.5772/intechopen.89608.
- [18] V. V. Gorborukov, O. Y. Stryzhak, O. V. Franchuk, V. B. Shapovalov, Ontological representation of the ranking alternatives task, *Mathematical modeling in economics* (2018) 49–69.
- [19] B. Nicolescu (Ed.), *Transdisciplinary: Theory and Practice*, Hampton Press, 2008.
- [20] 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, *CEUR Workshop Proceedings 2433* (2019) 246–255.
- [21] D. Martyniuk, M. Falkenthal, N. Karam, A. Paschke, K. Wild, An analysis of ontological entities to represent knowledge on quantum computing algorithms and implementations, *CEUR Workshop Proceedings 2836* (2021) 1–9.
- [22] R. Volckmann, Transdisciplinarity: Basarab Nicolescu Talks with Russ Volckmann, *Integral Review* 4 (2007) 73–90. URL: <https://integral-review.org/transdisciplinarity-basarab-nicolescu-talks-with-russ-volckmann/>.
- [23] Y. B. Shapovalov, V. B. Shapovalov, F. Andruszkiewicz, N. P. Volkova, Analyzing of main trends of STEM education in Ukraine using stemua.science statistics, *CEUR Workshop Proceedings 2643* (2020) 448–461.
- [24] V. B. Shapovalov, Y. B. Shapovalov, Z. I. Bilyk, A. P. Megalinska, I. O. Muzyka, The Google Lens analyzing quality: An analysis of the possibility to use in the educational process, *CEUR Workshop Proceedings 2547* (2020) 117–129.
- [25] Z. I. Bilyk, Y. B. Shapovalov, V. B. Shapovalov, A. P. Megalinska, F. Andruszkiewicz, A. Dolhanczuk-Sródkka, Assessment of mobile phone applications feasibility on plant recognition: comparison with Google Lens AR-app, *CEUR Workshop Proceedings 2731* (2020) 61–78.
- [26] Y. B. Shapovalov, V. B. Shapovalov, R. A. Tarasenko, S. A. Usenko, A. Paschke, A semantic structuring of educational research using ontologies, *CEUR Workshop Proceedings 2879* (2020) 105–123.
- [27] Y. Shapovalov, S. Zhadan, G. Bochmann, A. Salyuk, V. Nykyforov, Dry anaerobic digestion of chicken manure: A review, *Applied Sciences* 10 (2020) 7825.
- [28] V. Ivanov, Y. B. Shapovalov, V. Stabnikov, A. I. Salyuk, O. Stabnikova, M. ul Rajput Haq, Barakatullahb, Z. Ahmed, Iron-containing clay and hematite iron ore in slurry-phase anaerobic digestion of chicken manure, *AIMS Materials Science* 6 (2020) 821–832.
- [29] R. A. Tarasenko, V. B. Shapovalov, S. A. Usenko, Y. B. Shapovalov, I. M. Savchenko, Y. Y. Pashchenko, A. Paschke, Comparison of ontology with non-ontology tools for educational research, *CEUR Workshop Proceedings 2879* (2020) 82–104.