An academic events sub-system of the URIS and its ontology representation to improve scientific usability and motivation of scientists in terms of European integration

Yevhenii B. Shapovalov[1,2], Viktor B. Shapovalov[1,2], Alla G. Zharinova[2], Sergiy S. Zharinov[3], Iryna O. Tsybenko[2] and Oleksiy S. Krasovskiy[3]

[1]*National Center "Junior Academy of Science of Ukraine", 38-44 Degtyarivska Str., Kyiv, 04119, Ukraine*
[2]*The State Scientific and Technical Library of Ukraine, 180 Antonovicha Str., Kyiv, 03150, Ukraine*
[3]*Ukrainian Scientific Center for the Development of Information Technologies, 180 Antonovicha Str., Kyiv, 03150, Ukraine*

Abstract
Edge computing is a modern approach that may be considered as collecting and processing data near the source of its generation with further cloud computing. Therefore, the data processing that collected distributed and further collected at the central level may be considered an edge-based approach. NARCIS, SICRIS, and Research.fi systems exist to store and process scientific data in different countries of Europe, but in Ukraine, such a system is absent. The study foresees using the data collected in different instructions with its processing and collecting in a central database related to academic events. The list of relevant data stored in the academic events system and case diagram of the proposed system is developed. The essential EU legislation documents that should be considered are named. The list of systems that are proposed to provide interoperability with is investigated and described. Models' of receiving data, URIS as the main component of the decentralized approach in science, data exchange and interaction of proposed data base were illustrated and described.

Keywords
distributed data aggregation, scientific data, academic events, aconferences, URIS, UML, data model, interoperability

1. Introduction

The problem of data processing and structurization is actual nowadays [1, 2]. Edge computing can be defined as the model that optimizes cloud computing systems by processing data close to its source at the edge of the network [3]. For sure, the primary definition of edge computing is an aggregation of the vast amount of the data received from IoT (devices/sensors) in edges and

doors-2023: 3rd Edge Computing Workshop, April 7, 2023, Zhytomyr, Ukraine
sjb@man.gov.ua (Y. B. Shapovalov); svb@man.gov.ua (V. B. Shapovalov); zh.alla0812@gmail.com (. A. G. Zharinova); serhii.zharinov@gmail.com (S. S. Zharinov); tsyb2012@ukr.net (I. O. Tsybenko); aleksey_ks@ukr.net (O. S. Krasovskiy)
0000-0003-3732-9486 (Y. B. Shapovalov); 0000-0001-6315-649X (V. B. Shapovalov); 0000-0003-3959-1074 (. A. G. Zharinova); 0000-0003-3568-8127 (S. S. Zharinov); 0000-0002-6408-9011 (I. O. Tsybenko); 0000-0003-1622-2655 (O. S. Krasovskiy)

then its additional processing on the cloud. Edge computing is to provide services and performs calculations at the edge of the network and data generation [4].

However, in some terms, edge computing may be considered as "Decentralized cloud and low-latency computing" as noted in [3]. The necessity of such decentralization is caused by requests to decrease data processing latency [3]. Therefore, one of the vital signs of edge computing is using geographically distributed applications [3, 5]. In [6] edge computing is defined similarly as "near the edge of the network or the source of the data, an open platform that integrates core capabilities such as networking, computing, storage, applications, and provides edge intelligent services nearby to meet the industry agility key requirements in connection, real-time business, data optimization, application intelligence, security, and privacy". Edge computing has decentralized cloud architecture [7]. It enables data processing closer to the network's edge where the data is generated. Therefore, in those terms, edge computing's essential idea may be not using IoT, but geographical distribution of the data, its aggregation and decreasing of latency.

Those features are essential to provide in fields where data is significantly distributed, such as city management [8, 9, 10] and healthcare [10]. However, we think that geographical distribution of the data, its aggregation, and decreasing of latency is pretty essential in the field of science where data generated by scientists and administrators and inputted in computers in edges that should be aggregated and further processed in a central cloud server.

Nowadays, some systems devoted to systemizing information in the sciences exist. In Dutch, NARCIS provides structured research information with information from OAI repositories (publication and other scientific results), websites, and news pages of research institutes [11]. Some more developed systems operate in Slovenia and Finland. In Slovenia, SICRIS stores data on researchers, research groups, projects, programmers, and organizations [12]. Research.fi included Finnish research publications, research data, research projects, open research calls, infrastructures, and organizations [13]. However, in Ukraine, a significant part of the science data is still shared in not machine-readable form, and it is hard to process.

In addition, the data generation is distributed throughout the whole territory of Ukraine. Obviously, to overcome this problem, there should be a single access point to information about research conducted in Ukraine. The information would be presented in a clear and understandable form and be available for re-use by both people and computer programs.

Therefore, currently, there is a problem related to data collection of the central database and its usage in local machines to solve some local issues. Therefore, it seems relevant to substantiate the functionality, and the main types of metadata should be stored in the academic events sub-system. Therefore, this research is aimed to substantiate the functionality and required data to develop the academic events sub-system.

2. Methods

The analysis of existing systems related to science was provided and considered. The National Ukrainian Research Information System as a concept was described and substantiated as distributed computing system to give a general understanding of the system being developed and the digitalization of science. The data that is relevant to store in the academic events sub-system is described. Use case diagram is developed using Draw.io tool to create UML diagrams and represented to provide an understanding of role model of academic events sub-system. Models of receiving data, URIS as the main component of the decentralized approach in science, data exchange and interaction of proposed data base were illustrated and described.

3. Results

3.1. URIS as the basis of scientific data

The National Ukrainian Research Information System ("URIS") is developing to consolidate service, store, structure, and disseminate up-to-date information on all, studies, scientists, instructions, projects, and other scientific data related to Ukraine [14]. URIS foresees using multiple times the data inputted by users once agreed with the law of Ukraine "On public registers". The system is developed to provide accessibility, interoperability, and allow re-use and facilitate information retrieval following the FAIR (Findability, Accessibility, Interoperability, Reusability) principles and EU Directive 2019/1024.

URIS is designed to deliver crucial scientific information to users. It provides searching, viewing, and exporting to users' data on national science and scientometric indicators of scientists and institutions. This is calculated using data that is updated from local servers of scientific institutions that are edges and sent to the central service. The data receiving model is shown in figure 1. Also, the system can aggregate data taken from commercial and open-data sources requiring decentralization (figure 2).

URIS uses national and international permanent identifiers to identify scientists, publications, and institutions and build accurate relationships between them, which allows to achieve the reliability of the data and ensures its re-use. International identifiers ORCID ID, DOI, ROR provide a connection between publication, affiliation, scientist, and organization. The use of national identifiers EDRPOU (National State Registry of Ukrainian Enterprises and Organizations), RNOKPP (Taxpayer registration card registration number) allows identifying persons and legal entities. URIS provides a link between the national and international identifiers related to scientists and institutions, which allows us to ensure a single model and the data's completeness.

Using an electronic digital signature to identify users will be used in URIS, reducing the likelihood of entering unreliable information. The issue of edge computing, in this case, is the decentralization of tasks and calculations that are performed by individual URIS sub-systems or external systems. Thus, URIS imports data from the national systems of the EDRPOU (Unified State Register of Enterprises and Organizations of Ukraine), EDEBO (State electronic database on education), NRAT (National repository of academic texts), international open databases Crossref, DataCite, ROR, international commercial databases Scops, WoS and others, ensuring the completeness of data of Ukrainian science.

Therefore, a model has been developed that collects data in URIS into a complex model containing information about scientists, institutions, publications, and projects and allows a separate sub-system to receive part of the data from the complex API. These sub-systems can be a part of URIS and work on one server or distributed among different servers, serviced by different technical and administrative teams. This way, a balance between decentralization and

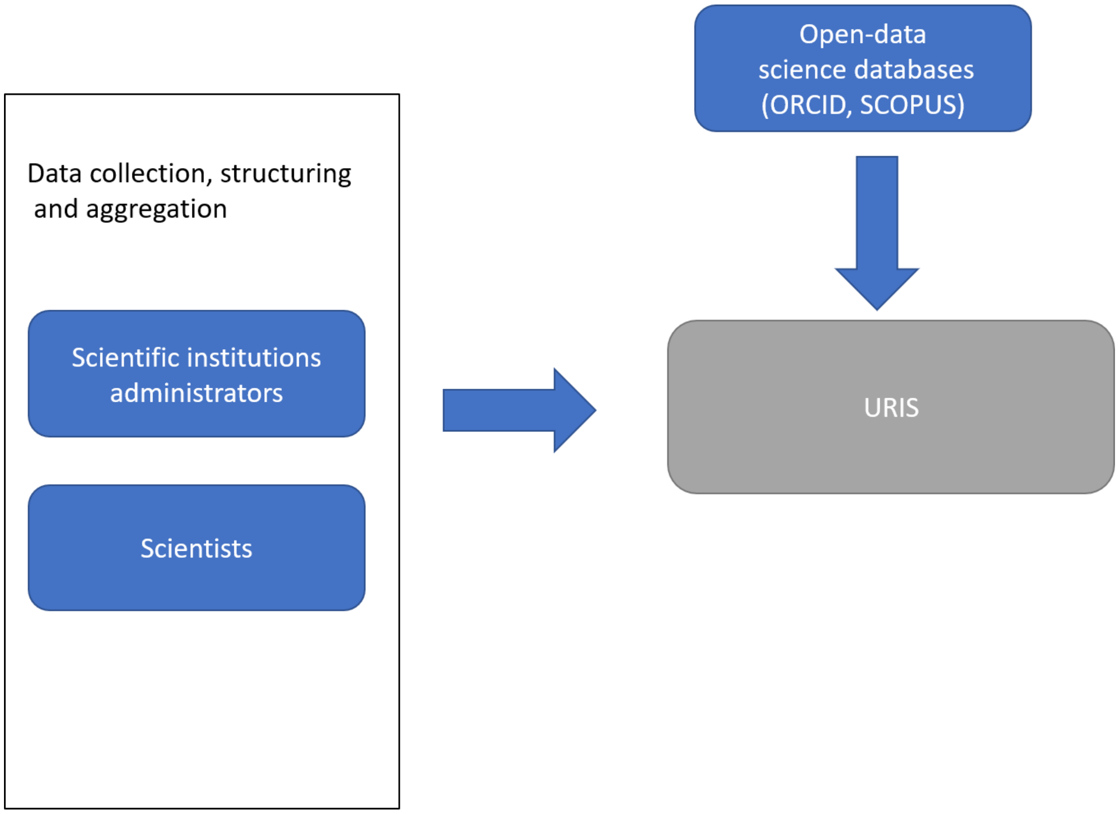

Figure 1: The model of receiving data.

the completeness of data on Ukrainian science is ensured. The model of data exchange of URIS is shown in figure 3.

The successful implementation of the URIS project will primarily help scientists themselves, as researchers will not have to enter the same information several times, which will reduce the burden of preparing applications and reports, and information about current research, potential research partners, and the necessary devices and equipment will be available to all stakeholders. The dashboard approach will be helpful in management at both national and institutional levels.

The main components of the URIS are sub-subsystems (registers) of scientists, institutions, projects, publications, infrastructure, academic events, and projects. The system foresees its own calculations, such as the Open Ukrainian Citation Index (OUCI).

Therefore, URIS can be considered as one of the elements of the edge-based system where users are local institutions and scientists that may interpret data in the way they are required. In the paper, we will substantiate the primary functional and relevant data to store in the sub-subsystem on academic events.

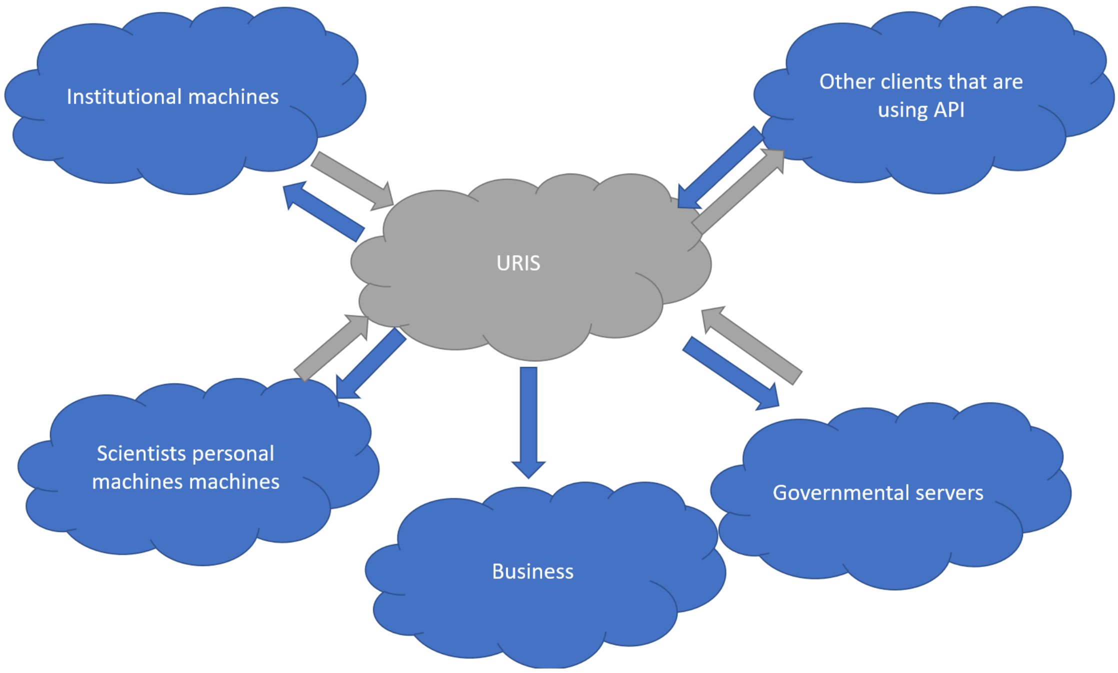

Figure 2: URIS as the main component of the decentralized approach in science.

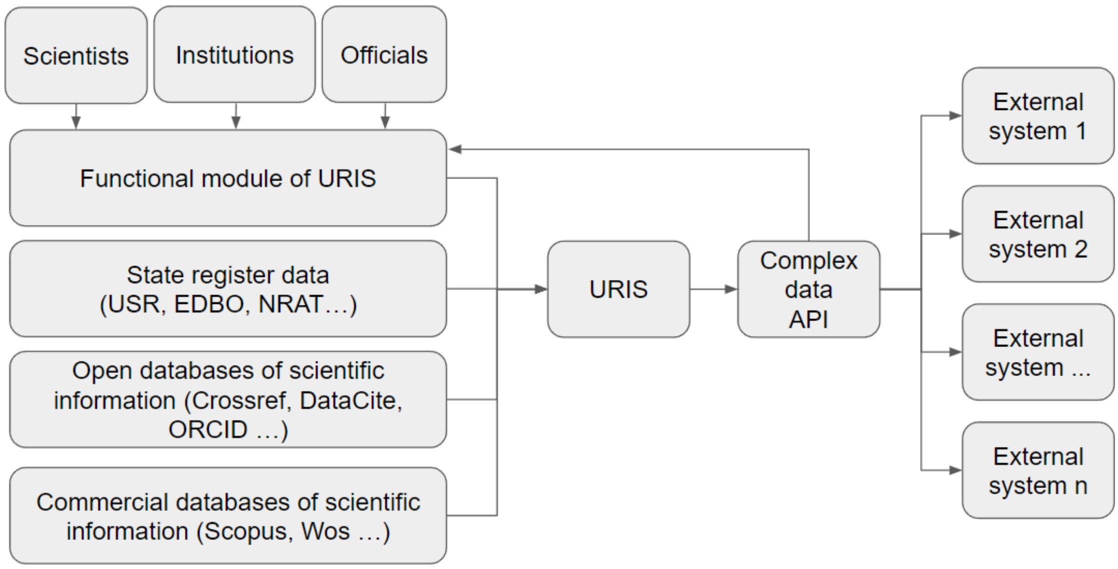

Figure 3: The model of data exchange of URIS.

3.2. Academic event platform as distributed system

The interaction of URIS and the sub-system is shown in the example of the sub-system of academic events. During the development of the submodule, one connector development to

URIS is enough, allowing data submitted by the conference participant to be checked. Using such a connector provide the fill of information about scientists, and affiliated institution, and check publications' citation and indexing. Therefore, spending time developing several connectors, solving organizational and legal problems to receive access to the information.

An ORCID ID is enough to receive the correct name and surname of the participant, his scientific degree, the place where he works, and the projects in which he participates. DOI allows checking whether such an article exists and getting information about the authors, title, abstract, publisher, and other metadata. ROR provides information about the institution to which it refers. The proposed data exchange model between URIS and the sub-module on academic events is shown in figure 4.

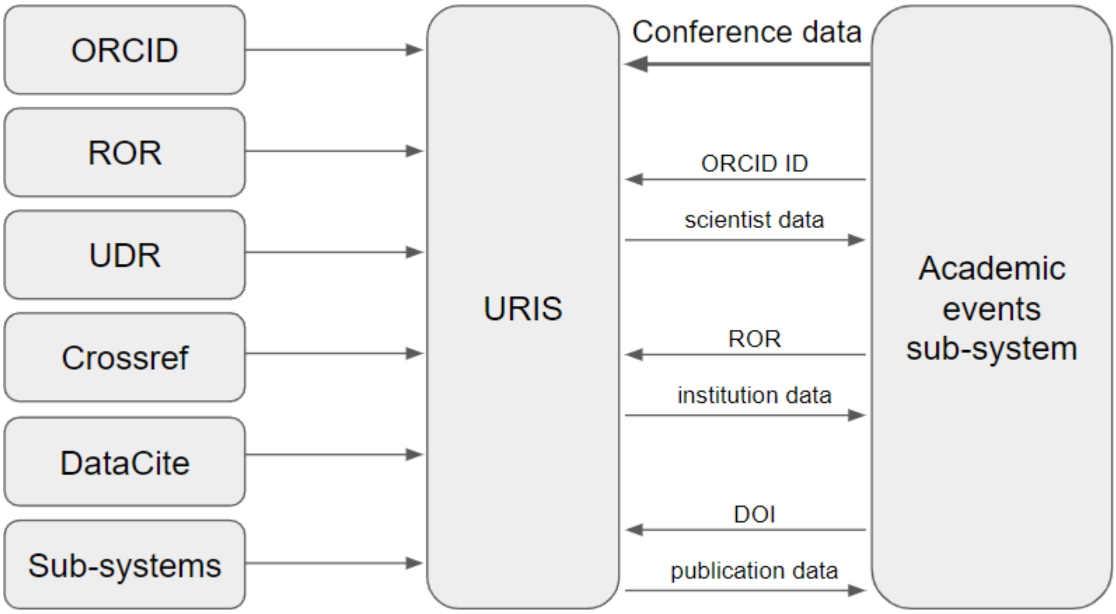

Figure 4: The proposed model of data exchange between URIS and the sub-module on academics events.

3.3. Main functions and data stored at academic events sub-system

The primary metadata on academic events should be followed "Direction", "Date of submission", "Dates of the event", "Submission method", "Fees, UAH", "Additional fees, UAH", "Details for payment,", "Contacts", "Event type", "Event description", "Probable indexing", "Publication format specification", "Publisher", "Publication type" "Related events/link", "Event program". In addition, some fields are used from URIS database, which are, organizers' personnel (by user's ID; ORCID ID), author of the publication (by user's ID; ORCID ID), and organizer (ID). Sure, not all of these data may be provided in the final version of the academic events sub-subsystem and may be optimized. However, these are the main types of academic events metadata.

The main functions of the platform are storing data on academic events, noticing users on the updates on the events, automatically sharing the data on the academic events, generation

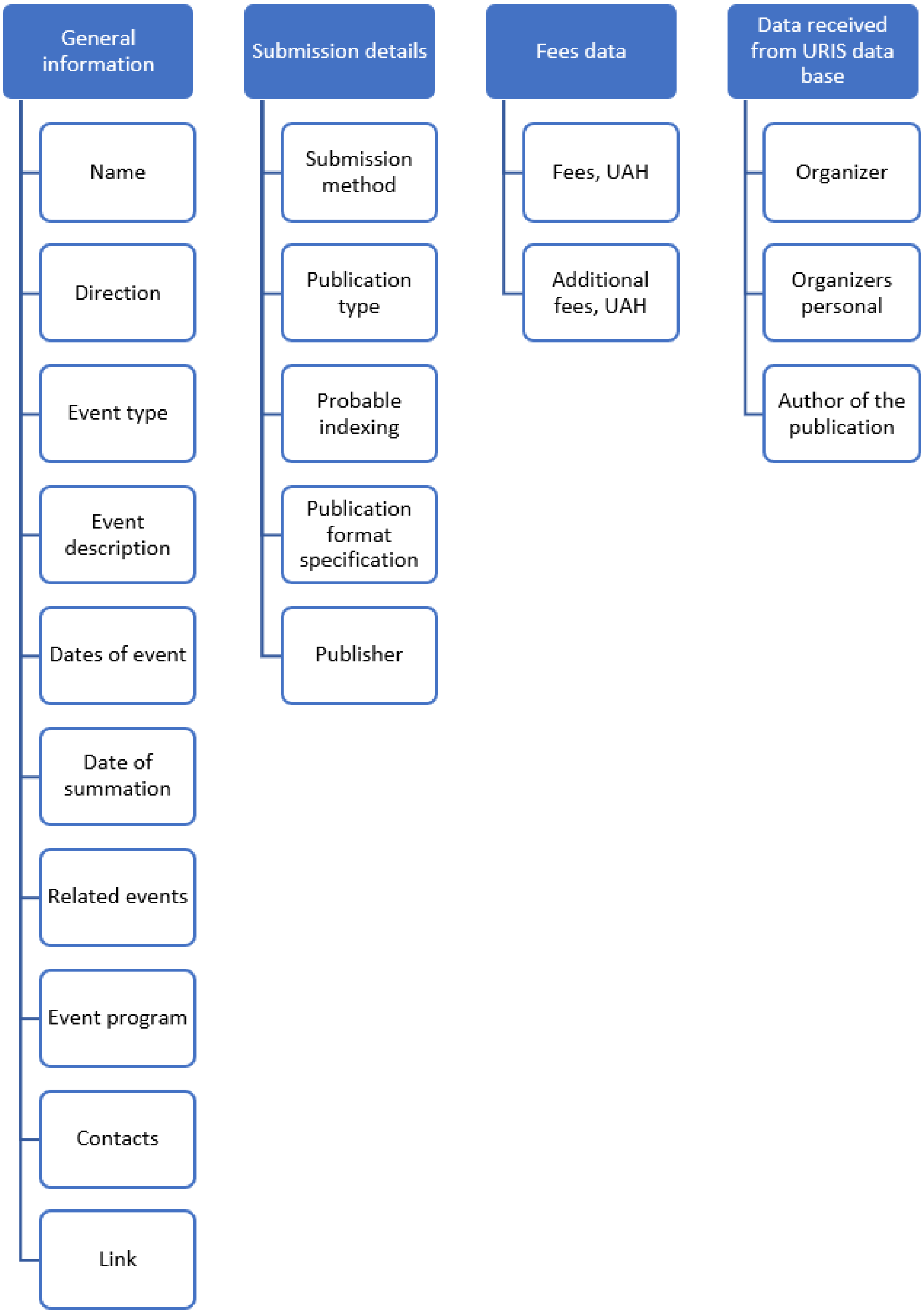

Figure 5: Main metadata of the academic events.

of dashboards on national and instructional levels, and declaration of participation in the academic events by scientists. The primary users of the sub-system are scientists (receiving notifications about new events, declaring participation, submission of publications), institutions administrators (review dashboards, automatization of reporting), government representatives (review dashboards, science management), providers (publish the data on the academic event), administrator of academic event system (check the facts on the academic event, organizations and approve/decline the submission of the academic event data) and other sub-systems of URIS. Therefore, the process is further: providers submit the data on academic events; the administrator checks the data; scientists receive the notification; the event is passing; institutions and government receive data in the form of dashboards, and scientists receive the participation provide in their profiles.

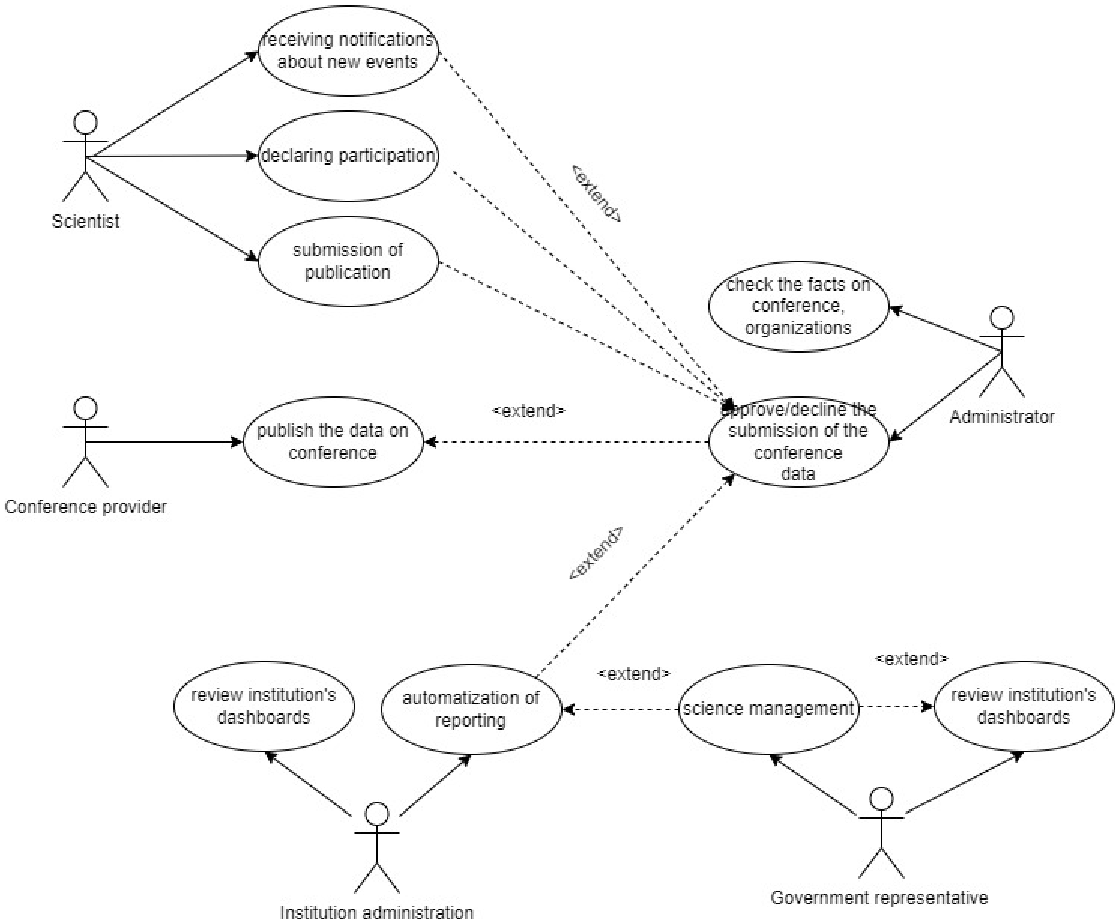

Figure 6: Use case diagram.

4. Discussion

Therefore, the proposed system as part of URIS will be helpful for a scientist. Compared to existing systems such as SICRIS, NARCIS, or Research.fi, the URIS will also collect data on the events that will be useful for scientists and save the scientist from participation in predatory conferences that are pretty widespread in Ukraine.

It is worth noting that in terms of European integration, it is essential to consider European legislation. The principles of Findability, Accessibility, Interoperability, and Reusability are declared in FAIR. As the EU declared open data and the re-use of public sector information by Directive (EU) 2019/1024, it is vital to develop and provide URIS and its sub-system on academic events that ensure the re-use of scientific data. It is essential to provide General Data Protection Regulation according to Regulation (EU) 2016/679. Adherence and ensuring of points declared seem relevant in the proposed system.

As ontology is a pretty effective tool in structuring scientific data [15, 16, 17], data on academic events may be represented in the form of ontologies that will provide additional structuring. In this case, data on ontologies will be structured by direction, and users will be able to use such structure to separate only by the required direction. It seems relevant to represent such taxonomies in the CIT Polyhedron system [18, 19, 20].

Transparency is crucial to involve and motivate youth to study activities. A lot of studies is demonstrate the importance of motivation in science and education [21, 22, 23, 24] Currently, the untransperency of science may repel youth; therefore, especially for them, it is vital to ensure motivation. It is worth noting that the motivation of youth to do science is acute [23, 25]. Digitalization of science may help by providing equal and transparent conditions [26].

5. Conclusions

The geographical distribution of the data, its aggregation, and decreasing latency are crucial in science. In this field, data generated by scientists and administrators and inputted into the computers in edges should be aggregated and further processed in a central cloud server. The National Ukrainian Research Information System ("URIS") is developing to solve this problem. It foresees developing to consolidate service, store, structure, and disseminate up-to-date information on all studies, scientists, instructions, projects, and other scientific data related to Ukraine. One of the components of URIS is the academic activity sub-system that helps to provide transparency in science and save scientists from participation in predatory. The primary metadata of the academic events includes general information on the event, submission details, fee information, and data obtained from URIS. The proposed user case model foresees the roles of scientists, institutional administrators, government representatives, administrators of academic event systems, and other sub-systems of URIS. It is essential to consider FAIR principles, Directive (EU) 2019/1024, and Regulation (EU) 2016/679. One of the effective ways of representing conference data is taxonomies. Digitalization of science may help by providing equal and transparent conditions.

References

[1] J. Portenoy, J. D. West, Constructing and evaluating automated literature review systems, Scientometrics 125 (2020) 3233–3251. doi:10.1007/s11192-020-03490-w.

[2] M. Popova, L. Globa, R. Novogrudska, Multilevel Ontologies for Big Data Analysis and Processing, Proceedings of International Conference on Applied Innovation in IT 9 (2021) 41–53. doi:10.25673/36583.

[3] J. Taheri, S. Deng (Eds.), Edge Computing: Models, technologies and applications, volume 33 of *IET professional application of computing*, Institution of Engineering and Technology, 2020. doi:10.1049/PBPC033E.

[4] K. Cao, Y. Liu, G. Meng, Q. Sun, An Overview on Edge Computing Research, IEEE Access 8 (2020) 85714–85728. doi:10.1109/ACCESS.2020.2991734.

[5] Z. Zhao, F. Liu, Z. Cai, N. Xiao, Edge Computing: Platforms, Applications and Challenges, Jisuanji Yanjiu yu Fazhan/Computer Research and Development 55 (2018) 327–337. doi:10.7544/issn1000-1239.2018.20170228.

[6] X. Hong, Y. Wang, Edge Computing Technology: Development and Countermeasures, Chinese Journal of Engineering Science 20 (2018) 20. doi:10.15302/j-sscae-2018.02.004.

[7] S. Munirathinam, Chapter Six - Industry 4.0: Industrial Internet of Things (IIOT), in: P. Raj, P. Evangeline (Eds.), The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases, volume 117 of *Advances in Computers*, 2020. doi:10.1016/bs.adcom.2019.10.010.

[8] O. A. Mahmood, A. R. Abdellah, A. Muthanna, A. Koucheryavy, Distributed Edge Computing for Resource Allocation in Smart Cities Based on the IoT, Information 13 (2022) 328. doi:10.3390/info13070328.

[9] L. U. Khan, I. Yaqoob, N. H. Tran, S. M. A. Kazmi, T. N. Dang, C. S. Hong, Edge-Computing-Enabled Smart Cities: A Comprehensive Survey, IEEE Internet of Things Journal 7 (2020) 10200–10232. doi:10.1109/JIOT.2020.2987070.

[10] R. Dave, N. Seliya, N. Siddiqui, The Benefits of Edge Computing in Healthcare, Smart Cities, and IoT, Journal of Computer Sciences and Applications 9 (2021) 23–34. doi:10.12691/jcsa-9-1-3.

[11] E. Dijk, C. Baars, A. Hogenaar, M. van Meel, NARCIS: The Gateway to Dutch Scientific Information, in: Digital Spectrum: Integrating Technology and Culture: Supplement to the Proceedings of the 10th International Conference on Electronic Publishing, Data Archiving and Networked Services (DANS), Bansko, Bulgaria, 2006, pp. 49 – 58. URL: https://pure.knaw.nl/ws/portalfiles/portal/86266383/233_elpub2006.content_0.pdf.

[12] L. Curk, Implementation of the Evaluation of Researchers' Bibliographies in Slovenia, Procedia Computer Science 146 (2019) 72–83. doi:10.1016/j.procs.2019.01.082.

[13] J. Nikkanen, H. M. Puuska, Researchers' profiles in the Finnish Research Information Hub, Procedia Computer Science 211 (2022) 206–210. doi:10.1016/j.procs.2022.10.193.

[14] S. Nazarovets, Natsionalna naukovo-informatsiina systema URIS ta pryntsyp yii pobudovy (2020) 1–4. doi:10.5281/zenodo.4038422.

[15] V. V. Prykhodniuk, M. V. Nadutenko, H. M. Potapov, Programmatic system for interactive representation of scientific institution results, Scientific notes of Junior Academy of

Sciences of Ukraine (2022) 91–99. doi:10.51707/2618-0529-2022-24-11.
[16] L. Globa, R. Novogrudskaya, B. Zadoienko, O. Y. Stryzhak, Ontological Model for Scientific Institutions Information Representation, in: 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), 2020, pp. 255–258. doi:10.1109/PICST51311.2020.9467984.
[17] R. A. Tarasenko, S. A. Usenko, Y. B. Shapovalov, V. B. Shapovalov, A. Paschke, I. M. Savchenko, Ontology-based learning environment model of scientific studies, in: V. Ermolayev, A. E. Kiv, S. O. Semerikov, V. N. Soloviev, A. M. Striuk (Eds.), Proceedings of the 9th Illia O. Teplytskyi Workshop on Computer Simulation in Education (CoSinE 2021) co-located with 17th International Conference on ICT in Education, Research, and Industrial Applications: Integration, Harmonization, and Knowledge Transfer (ICTERI 2021), Kherson, Ukraine, October 1, 2021, volume 3083 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2021, pp. 43–58. URL: https://ceur-ws.org/Vol-3083/paper278.pdf.
[18] O. Stryzhak, V. Prykhodniuk, M. Popova, M. Nadutenko, S. Haiko, R. Chepkov, Development of an Oceanographic Databank Based on Ontological Interactive Documents, in: K. Arai (Ed.), Intelligent Computing - Proceedings of the 2021 Computing Conference, Volume 2, SAI 2021, Virtual Event, 15-16 July, 2021, volume 284 of *Lecture Notes in Networks and Systems*, Springer, 2021, pp. 97–114. doi:10.1007/978-3-030-80126-7_8.
[19] 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.
[20] Y. Shapovalov, V. Shapovalov, R. Tarasenko, Z. Bilyk, I. Shapovalova, A. Paschke, F. Andruszkiewicz, Practical application of systemizing expedition research results in the form of taxonomy, Educational Technology Quarterly 2022 (2022) 216–231. doi:10.55056/etq.40.
[21] L. Linnenbrink-Garcia, A. M. Durik, A. M. M. Conley, K. E. Barron, J. M. Tauer, S. A. Karabenick, J. M. Harackiewicz, Measuring situational interest in academic domains, Educational and Psychological Measurement 70 (2010) 647–671. doi:10.1177/0013164409355699.
[22] İ. Dökme, A. Açıksöz, Z. Koyunlu Ünlü, Investigation of STEM fields motivation among female students in science education colleges, International Journal of STEM Education 9 (2022) 8. doi:10.1186/s40594-022-00326-2.
[23] D. Fortus, I. Touitou, Changes to students' motivation to learn science, Disciplinary and Interdisciplinary Science Education Research 3 (2021) 1. doi:10.1186/s43031-020-00029-0.
[24] O. Pursky, A. Selivanova, I. Buchatska, T. Dubovyk, T. Tomashevska, H. Danylchuk, Features of learning motivation in the conditions of coronavirus pandemic (COVID-19), Educational Technology Quarterly 2021 (2021) 375–387. doi:10.55056/etq.31.
[25] S. M. Duisenova, Scientific Motivation of Young Scientists of Higher Educational Institutions (Engaged in Sociological Research), Mediterranean Journal of Social Sciences 6 (2015) 26. doi:10.5901/mjss.2015.v6n6s1p26.
[26] O. Kuzminska, Selecting tools to enhance scholarly communication through the life cycle of scientific research, Educational Technology Quarterly 2021 (2021) 402–414. doi:10.55056/etq.19.