

# Open access to data about higher education and science. Case study of the RAD-on platform in Poland

Jarosław Protasiewicz<sup>\*</sup>, Emil Podwysocki, Sylwia Ostrowska,  
Aldona Tomczyńska<sup>†</sup>

The National Information Processing Institute, al. Niepodległości 188b, 00-608, Warsaw, Poland,  
[jaroslaw.protasiewicz@opi.org.pl](mailto:jaroslaw.protasiewicz@opi.org.pl), [emil.podwysocki@opi.org.pl](mailto:emil.podwysocki@opi.org.pl),  
[sylwia.rosiak@opi.org.pl](mailto:sylwia.rosiak@opi.org.pl), [aldona.tomczyńska@opi.org.pl](mailto:aldona.tomczyńska@opi.org.pl)

## Abstract

In the era of data-driven policies, there is a need to ensure that decision-makers and citizens have easy access to the best quality data available. By creating Rad-on: information system with reports, analyses and data on science and higher education in Poland, we tried to overcome barriers to decision-making in the field of research and innovation. The RAD-on system presents open government data on topics that are important for the science and innovation policy such as: publication patterns of scientists and academics in Poland, graduates from Polish universities and their income, and various data on research projects conducted in Poland, financed from national and international funds. RAD-on aims at ensuring transparency and participation through interaction with its users and a methodologically coherent approach to data. In this research paper, we describe a complex architecture of the RAD-on system and present its reporting capabilities. While doing so, we focus on the development process, which took into account recent studies on usability of open data. We address the problem of the users' understanding of the possibilities of the data platform and our efforts to reduce risks of data being wrongly interpreted and used in decision-making processes.

## 1 RAD-on as an open and integrated data platform

In the era of data-driven policies, there is a need to ensure that decision-makers, citizens and other stakeholders have access to non-privacy-restricted and non-confidential data, produced with public money (Ruijter et al., 2020). By creating Rad-on: information system with reports, analyses and data on science and higher education<sup>‡</sup> (Michajłowicz et al., 2018; Protasiewicz et al. 2019), we tried to overcome barriers to decision-making in the field of research and innovation. In this research paper, we examine the 'reports and analysis' component of the RAD-on by showing its capabilities and potential in reducing the risk of data being wrongly interpreted and used in the decision-making processes. We present the platform's functionalities and potential in creating statistics that are in line with good practices of opening government data (OGD). Furthermore, we focus on best practices in creating OGD platforms that aim at increasing user's satisfaction from data exploration (see Weerakkody et al. 2016).

---

<sup>\*</sup> <https://orcid.org/0000-0002-9204-921X>

<sup>†</sup> <https://orcid.org/0000-0002-0832-8081>

<sup>‡</sup> Website: <https://radon.nauka.gov.pl/> (only in Polish).

The RAD-on presents OGD about science and higher education which come from multiple separate databases administrated by the National Information Processing Institute<sup>§</sup>, e.g. POL-on<sup>\*\*</sup>, a database which gathers information about all scientific institutions and students in Poland; Polish Scientific Bibliography<sup>††</sup>, which stores data about research papers of the Polish academics; and the Polish Graduate Tracking System<sup>‡‡</sup> (ELA), which monitors professional careers of graduates of higher education institutions (HEIs).

All of the systems are integrated through RAD-on to provide the highest quality of data, as measured through the so-called data quality metrics (see Azeroual et al. 2018). Most of data are directly entered into the database by individual HEIs and, in the case of the ELA system, matched with data provided by the Social Insurance Institution. These systems are unique on an international scale. POL-on alone is the biggest public IT system in terms of data size and range in Poland, and, to our knowledge, has no match if compared to similar systems elsewhere in Europe.

Thanks to RAD-on many different analytical questions can be answered: how many higher education institutions are there in Poland? What is their educational offer? How many students are enrolled at each higher education institution? How many of them are foreigners? Which courses are the most popular among student? Which of them give the best chances at the labour market? What salaries might be expected after graduation of a specific university or course? Each analytical report is complimented by methodological and expert comment. Although business intelligence and learning analytics systems for higher education have been built in other countries for similar purposes (e.g. Khatibi et al. 2020, see also Samuelsen et al. 2019), their actual scope is different. Best to our knowledge, the RAD-on system is unique in terms of data accuracy, scale and thematic differentiation.

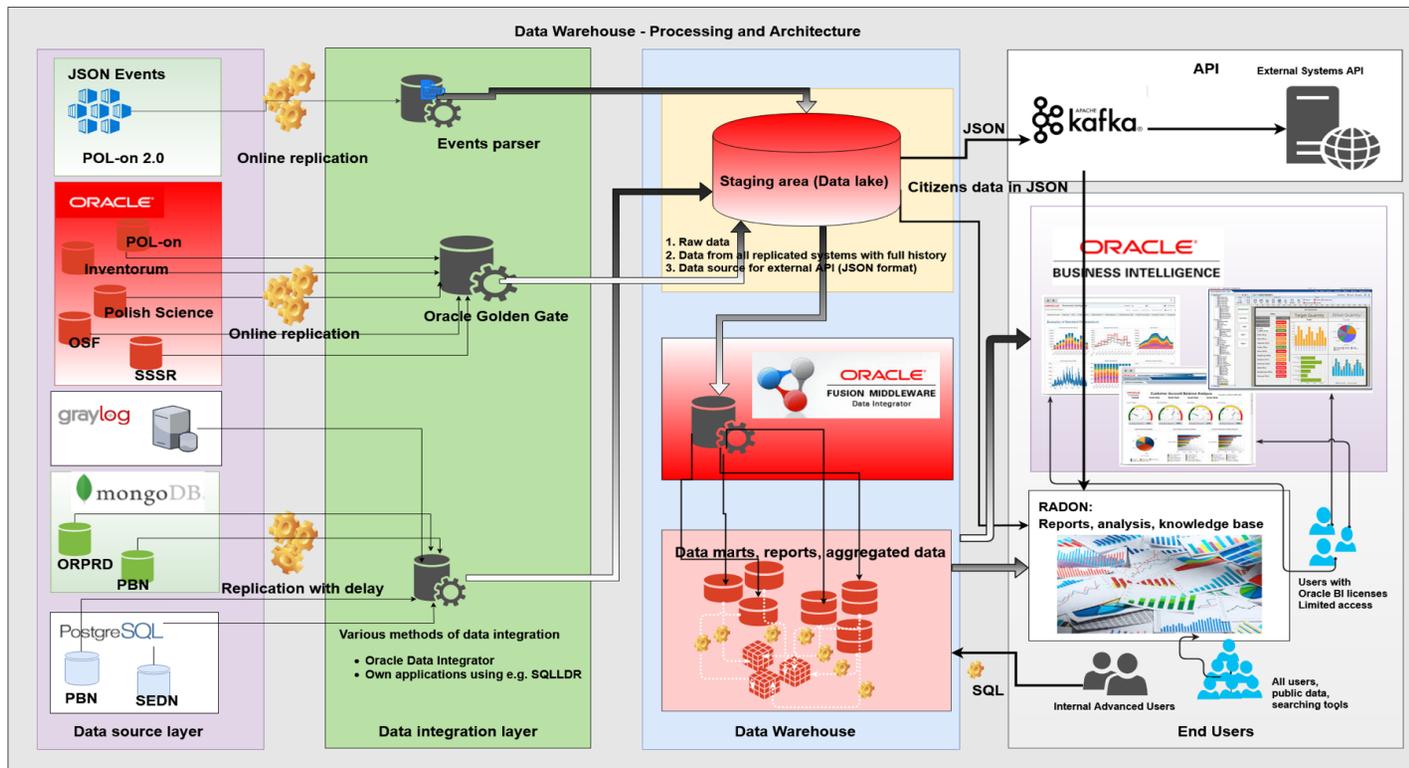
---

<sup>§</sup> Website: <https://www.opi.org.pl> (English version available).

<sup>\*\*</sup> Website: <https://polon.nauka.gov.pl> (English version available).

<sup>††</sup> Website: <https://pbn.nauka.gov.pl/> (only in Polish).

<sup>‡‡</sup> Website: <https://ela.nauka.gov.pl> (English version available).



**Figure 1. A simplified architecture of the RAD-on**

Abbreviations and detailed information on data source layer: POL-on 2.0 – system of information about science and higher education, Inventorum – system that facilitates science-business cooperation, Polish Science – data about Polish scientists and their research, OSF – system for management of funds for science, SSSR – support system for selection of reviewers, ORPPD – national repository of written diploma theses, PBN – database of Polish scientific publications, SEDN – System for Evaluation of Scientific Achievements in Poland.

## 2 Open data in RAD-on: development of dynamic reports

In order to create the most useful tool for analyzing data about science and higher education, we conducted 24 in-depth interviews with stakeholders of RAD-on: representatives of the Polish government and non-governmental organizations, research financing organizations (RFOs), scientific journalists, and employees of universities in Poland (academics and supportive personnel). This qualitative study gave us a comprehensive insight into what functionalities were required by users. Our interviewees pointed out to three main problems related to statistics about science and higher education, such as: dispersion of data into nine separate databases, low reliability of data, lack of methodological unity. We addressed all of these issues in the RAD-on's reports section. The current version of RAD-on takes into account recent studies on open data and its usability (Ruijter et al., 2020).

RAD-on presents data analysis of research about artificial intelligence conducted in Poland and abroad, and other, more general statistics about science and higher education. We plan to further develop our platform in order to offer insight into more complex issues of the science policy such as: publication patterns of scientists and academics in Poland, graduates from each specialization and their income, patents and licenses obtained by employees of scientific institutions in Poland, and various data on research projects that are conducted in Poland and financed from national and international funds. RAD-on will therefore open the largest set of data on these topics and may serve as an example of a nation-wide source of knowledge about the sector of science.

The main risk associated with such a project is wrong interpretation of data by users of the RAD-on platform. In this paper we further elaborate on our solutions which ensured the delivery of data that is easy to interpret and to reuse in user-made reports.

## References

- Azeroual, O., Saake, G. & Wastl, J. Data measurement in research information systems: metrics for the evaluation of data quality. *Scientometrics* 115, 1271–1290 (2018). <https://doi.org/10.1007/s11192-018-2735-5>
- Michajłowicz M., Niemczyk M., Protasiewicz J., Mroczkowska K. (2018). POL-on: The Information System of Science and Higher Education in Poland, In *EUNIS 2018 Congress Book of Proceedings*.
- Ruijter E, Grimmelikhuijsen S, van den Berg J, Meijer A. Open data work: understanding open data usage from a practice lens. *International Review of Administrative Sciences*. 2020;86(1):3-19. doi:10.1177/0020852317753068
- Protasiewicz J., Rosiak S., Kucharska I., Podwysoki E., Niemczyk M., Błaszczuk Ł., Michajłowicz M. (2019). RAD-on: An integrated System of Services for Science - Online Elections for the Council of Scientific Excellence in Poland, In *EUNIS 20189 Congress Book of Proceedings*.
- Samuelsen, J., Chen, W. & Wasson, B. Integrating multiple data sources for learning analytics—review of literature. *RPTEL* 14, 11 (2019). <https://doi.org/10.1186/s41039-019-0105-4>

Vahid Khatibi, Abbas Keramati, Farid Shirazi, Deployment of a business intelligence model to evaluate Iranian national higher education, *Social Sciences & Humanities Open*, Volume 2, Issue 1, 2020, 100056, ISSN 2590-2911, <https://doi.org/10.1016/j.ssaho.2020.100056>.

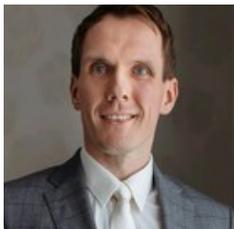
Weerakkody, V., Irani, Z., Kapoor, K. *et al.* Open data and its usability: an empirical view from the Citizen's perspective. *Inf Syst Front* 19, 285–300 (2017). <https://doi.org/10.1007/s10796-016-9679-1>

### 3 AUTHORS' BIOGRAPHIES



**Jarosław Protasiewicz (PhD)** is a director of the National Information Processing Institute and an assistant professor at the same Institute. He received a Ph.D. in computer science at the Systems Research Institute of the Polish Academy of Sciences. His areas of interest include agile project management, software design and development, big data, machine learning, and bio-inspired algorithms.

Email: [jaroslaw.protasiewicz@opi.org.pl](mailto:jaroslaw.protasiewicz@opi.org.pl)



**Emil Podwysocki (MSc)** received a master's degree in telecommunications systems at the Technical University of Lodz. He has 10 years of professional experience related to ETL/ELT, data warehouses and business intelligence. His areas of interest include Oracle technology, big data, business intelligence and data visualizations. Currently, he works as a Database and Business Intelligence Team Leader in the National Information Processing Institute.

Email: [emil.podwysocki@opi.org.pl](mailto:emil.podwysocki@opi.org.pl)

LinkedIn: [www.linkedin.com/in/emil-podwysocki](http://www.linkedin.com/in/emil-podwysocki)



**Sylwia Ostrowska (MSc)**: business analyst and project manager at the National Information Processing Institute. She received a master's degree in Management at the University of Warsaw and completed postgraduate studies in designing information systems at the Warsaw University of Technology. She has professional experience in business analysis and project management in an Agile environment. Currently, she is the project manager of „RAD-on: An integrated System of Services for Science”. Her knowledge in this area is confirmed by following certificates: PRINCE2 Foundation, Professional Scrum Product Owner.

Email: [sylwia.rosiak@opi.org.pl](mailto:sylwia.rosiak@opi.org.pl)

LinkedIn: <http://linkedin.com/in/sylwia-ostrowska-06749b23>



**Aldona Tomczyńska (PhD)**: doctor of philosophy in the field of international political economy. She was awarded her doctorate degree with distinction from the University of Warsaw in 2017. She has worked as a chief researcher at the National Information Processing Institute for over eight years. She conducts research in the area of political economy, innovation, and data science. She also serves as an expert in the field of science and higher education

policy. She coordinated multiple scientific projects, participated in evaluation studies and research and support actions from the Horizon 2020 programmes.

Email: [aldona.tomczynska@opi.org.pl](mailto:aldona.tomczynska@opi.org.pl)

LinkedIn: <http://linkedin.com/in/aldona-tomczyńska-phd-8490b718a>