Beyond CRIS: A research and higher education information system in Poland

Aldona Tomczyńska[1*] Sylwia Ostrowska[1]
Jarosław Protasiewicz[1†] and Emil Podwysocki[1]
[1] The National Information Processing Institute, Poland

aldona.tomczynska@opi.org.pl, sylwia.ostrowska@opi.org.pl
jaroslaw.protasiewicz@opi.org.pl, emil.podwysocki@opi.org.pl

Abstract

This article presents a recently implemented Polish information technology system that provides official data on science and higher education. The RAD-on is an upgraded analytical platform of the CRIS and HEMIS systems that exist in other countries. The article outlines RAD-on's applications in decision-making processes at research, governmental, and nongovernmental institutions. The special role and the innovation of the RAD-on system for the science and higher education sector are illustrated using examples of its practical use by universities, government ministries, foundations, and journalists.

The article also describes trends in the use of data in decision-making processes, reviews the CRIS and HEMIS nationwide systems in use across the world, and compares them to RAD-on to discuss the latter's unique features. It also reflects on RAD-on's functionalities and how they are used in practice by the stakeholders of the science and higher education system. The article concludes by discussing the challenges of open access to government data and the development plans for RAD-on.

[*] https://orcid.org/0000-0002-0832-8081
[†] https://orcid.org/0000-0002-9204-921X

1 Introduction

The Current Research Information System (CRIS) is an information system that stores, manages, and exchanges contextual metadata for research activity [‡]. The Higher Education Information Management System (HEMIS) is a data system that collects, monitors, manages, analyses, and disseminates information about education inputs, processes, and outcomes[§].

The majority of such systems are developed for specific institutions or their associates. This article presents RAD-on: Reports, Analyses and Data on Higher Education and Science, a nationwide IT system that was first implemented in Poland in 2019. The system is an analytical platform that provides official data on science and higher education in Poland. It is available to the public at radon.nauka.gov.pl in the Polish and English languages. In the last 12 months it had 114 thousand unique users and almost two million sessions. RAD-on represents the next stage in the development of the CRIS and HEMIS systems. This article outlines RAD-on's applications in decision-making processes at research institutions, as well as governmental and nongovernmental institutions. The special role and innovation of RAD-on in the science and higher education sector are illustrated using examples of its practical use by universities, government ministries, nongovernmental organisations (NGOs), and journalists.

The article comprises eight sections. Section 1 presents general information. Section 2 outlines the objectives of RAD-on, and section 3 describes them against the backdrop of trends in the use of data in decision-making processes. Sections 4 and 5 focus on RAD-on's functionalities and the technologies that have been used in its development. Section 6 reviews the CRIS and HEMIS nationwide systems, which offer access to data and analytical tools, and compares them to RAD-on to demonstrate the innovative features of the latter. Section 7 discusses how the stakeholders of the science and higher education systems use RAD-on's functionalities in practice. Section 8 reflects on RAD-on's development and maintenance challenges, and discusses the platform's future plans.

2 Project goals

The National Information Processing Institute (OPI PIB) offers IT tools that are necessary in the management of science and higher education processes, including those related to distribution of research and education grants, and to the national research evaluation exercise. Since 2011, OPI PIB, the Polish Ministry of Science and Education, and research institutions have been engaged in the development of various data collection systems, which operate according to official legal regulations. Due to the majority of research institutions being obligated to upload data to the systems on a regular basis, the information stored by such systems is the most up-to-date and reliable. The main problem was that the data was nonpublic and distributed between separate systems, which prevented its comprehensive analysis. Preparing reports for ministries, journalists, and researchers was laborious and time-consuming, as well as a cause of high system load. Technical and methodological integration of separate systems was low.

RAD-on was implemented to make the data collected by OPI PIB available to universities and institutional authorities that are responsible for research and development, public administration staff, representatives of NGOs, scientists, journalists and other groups of individuals and entities interested in science and higher education, such as innovative enterprises. RAD-on was designed to:

1. ensure open and fast access to up-to-date and reliable data on research and development, science, and higher education,

[‡] https://www.eunis.org/wp-content/uploads/2016/03/cris-report-ED.pdf
[§] https://elibrary.worldbank.org/doi/10.1596/978-1-4648-1099-2_ch1

2. support decision-making processes by providing IT tools that analyse and interpret available data.

To meet these objectives, specific solutions and services were proposed. They were tailored to users' needs that experts at OPI PIB had identified in a study conducted before the system's launch. The study was conducted using the in-depth interview method. The twenty-nine interviewees included representatives of potential groups of RAD-on stakeholders: scholars, representatives of university authorities, senior and junior staff of government ministries and expert bodies responsible for science and education, representatives of research funding institutions, journalists, and representatives of education NGOs. Their feedback helped OPI PIB define the scope of services offered by RAD-on.

3 Project environment: Trends in research and student data management

RAD-on's objectives follow international science trends. Member states of the European Union seek to utilise nonprivate and nonconfidential data in decision-making processes on the largest possible scale (Rieder & Simon, 2016; Ruijer et al., 2020). Open government data (OGD) is a reliable source of information that is provided by public administrations or their agencies, free of charge (Gao et al., 2021). OGD has been in practice in the European Union and the United States for over a decade. Related to OGD is the concept of FAIR (findable, accessible, interoperable, reusable), which was originally introduced in *The FAIR Guiding Principles for scientific data management and stewardship* (Wilkinson et al., 2016). In FAIR, emphasis is placed on the creation of metadata sets and instructions regarding their processing that can be understood by humans and computer programs.

As work on OGD and FAIR progressed, new challenges emerged. The initial phase of the works concentrated on the granting of open access to resources and the establishment of descriptive standards. Currently, emphasis is placed on ensuring database interoperability. Studies (e.g. Dombrowski et al., 2012) suggest that the use of OGD in decision-making processes is linked inextricably to the manner in which the data is made available. Users of IT systems are reluctant to utilise raw data (Weerakkody et al., 2017), because cleaning it requires experience and is time-consuming. For that reason, the crucial element in the creation of OGD infrastructure is information processing. Patricia Huston, Victoria L. Edge and Erica Bernier (2019) also stress that the most considerable advantages can be attained by integrating data from multiple sources that form an OGD system; in other words, the optimal system is one that replaces large, raw datasets with access to interconnected and suitably processed information. Only OGD that is prepared in this manner can be properly interpreted and used directly in analytical processes.

The RAD-on system relies heavily on the OGD, FAIR, and data-driven policy concepts. It integrates data on science and higher education in Poland that originates from autonomous databases. One such database is POL-on (Michajłowicz et al., 2019), a system that gathers data on all students, graduates, PhD candidates, and employees of research institutions in Poland and their scholarly activities. The data is collected in the form of information on projects, investments, and patents, among others. The Polish Scholarly Bibliography (*Polska Bibliografia Naukowa*, PBN) is a national system that gathers metadata of scholarly publications, similar to CRISTIN[**]. The National Repository of Theses (*Ogólnopolskie Repozytorium Pisemnych Prac Dyplomowych,* ORPPD) stores the bachelor's and master's theses of Polish graduates. The Polish Graduate Tracking System (*Ekonomiczne Losy Absolwentów,* ELA)

[**] https://www.cristin.no.

provides data on the earnings of graduates of higher education institutions[††]. Currently, RAD-on offers information from all of the systems described above that is not personal or highly sensitive[‡‡].

4 Functionalities and the development process

The RAD-on project's success hinged on the primary task of designing and implementing a data warehouse and a data exchange model to ensure the integration of the IT systems described in Section 3 of this article. The technologies that were used to attain that goal are described in Section 5. This section focuses on RAD-on functionalities that are available to all internet users via the main portal. At this juncture in the evolution of OGD, it is not only the scope of data that is relevant, but also its format. It is also worth mentioning that RAD-on provides a single source of access to knowledge and data on science and higher education. More importantly, RAD-on offers services tailored to the digital proficiency of its users. Public services are grouped into five modules: Reports, Analyses, Data, User account, and API.

The Reports module is similar to the data dashboards that are offered by some private firms. Reports ensure access to aggregate data on higher education institutions, students, graduates, academic staff, and research conducted both in Poland and internationally. They include tables, visualisations, commentaries, and an option to download datasets in the XLSX, CSV formats, or graphs in PNG format. The Reports module is an analytical platform that helps users discover nationwide trends, compare higher education institutions, and monitor phenomena that influence the science and business sectors. The module is user-friendly for those who lack experience with business intelligence tools, and is designed chiefly for journalists, scholars, and civil servants.

The Analyses module is a repository of reliable studies on science and higher education prepared by scholars at OPI PIB that is based on the data from the RAD-on source systems and selected external sources. Similarly to the Reports module, the Analyses module is a source of knowledge used in decision-making processes.

The Data module contains official public registers with information on research institutions, scholarly and artistic activities, academic staff, and promotion procedures in science. They are popular with representatives of public administration who seek nonaggregated data, such as comprehensive lists of academic teachers employed in Poland or exhaustive lists of university programmes. Data can be displayed or downloaded in various formats, such as XLSX and CSV.

User accounts are used to browse data that is gathered in RAD-on and its source systems. Registered users can update and verify their data to ensure that RAD-on provides the highest quality of information.

The API service guarantees machine access to data that has been properly prepared. The service is used by higher education institutions, enterprises, and entities that operate in the science sector and possess their own IT systems.

RAD-on was developed using agile and waterfall project management methods. Given the extended timeline of the project, a flexible approach to the final product was essential to accommodate the evolving needs of users. To accurately evaluate the quality of services provided by RAD-on, four comprehensive usability tests were conducted after each crucial stage of the system's development. The tests were based on in-depth interviews. Distinct groups of system stakeholders, which had been predefined during the initial needs assessment, were identified. To ensure that the system meets the highest accessibility standards, it was essential to consider the needs of elderly users and users with visual impairments. The usability studies first focused on the RAD-on system mock-up, and then on

[‡‡] According to Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

selected services and on the graphical layout. Each study was followed by a report whose recommendations modified the scope or form of the services. This resulted in significant alterations to the development timeline, yet ultimately led to higher user satisfaction ratings. To monitor the ratings, a dedicated survey that is available on the system website is used. The results of the survey provide information on: 1) overall satisfaction with the reports, analyses, and integrated access to the data; 2) satisfaction with the data's quality; and 3) satisfaction with the machine data processing services. The results are discussed in Section 8.

5 Technology

The RAD-on IT ecosystem lacked a crucial component that could integrate and organise the data generated by individual systems effectively. Due to the absence of a central data source, reporting had to rely directly on production systems. This resulted in nonstandardised processes of exchanging and integrating data between systems, which, in turn, often led to unnecessary workloads for the production systems. To address this issue, we introduced an integration layer, which consisted of a data exchange model (DEM) and a data warehouse.

The DEM is responsible for the integration of source systems and for the provision of data for machine services (such as the RAD-on public API), or for services that are available through the RAD-on portal (such as public data comparisons). DEM participates indirectly in the implementation of all system services, which is why choosing the right technology was a key technical decision from the outset of the project. After conducting a thorough analysis, we decided to depart from the conventional service-oriented architecture approach and utilise the Apache Kafka message queue system to address our communication needs. Apache Kafka combines two data exchange models between distributed systems. Data is made available in the form of an ordered message queue, which enables all records to be read by any number of recipients. Apache Kafka integrates a queue-based communication model with a subscription one. A significant advantage of Apache Kafka is its scalability: it has proved to be a reliable solution for large distributed systems that are available as open-source software.

The data warehouse was designed using two methods: 1) the classic star schema, which is utilised as a data source for reports published on RAD-on, and 2) a data vault, which is applied to accommodate frequent changes in data structures and to work with documents—including those in JSON format. A dedicated data warehouse combines, deduplicates and aggregates data from all of the domain systems described in the fourth section (Figure 1: The data source layer, the data integration layer, and the data warehouse). RAD-on's data warehouse is organised into: 1) the staging area—a space in which source data from domain systems is stored in unprocessed form; 2) the tools for integrating the data warehouse with the domain systems—each of which is adapted to the database technology of the domain system; and 3) aggregated data and data marts, which contain various data models used by business intelligence tools and other IT applications for end users.

The integration of multiple nonheterogeneous data sources required the preparation of a comprehensive data processing procedure (Extract, Transform, Load – ETL). To achieve this, we used various data integration tools and methods. Data from Oracle databases is replicated to the data warehouse in real time, using Oracle Golden Gate. This allowed us to transfer changes in data that are made in production systems efficiently and without overloading the systems. The data replication delay is minimal and does not exceed one second. For other data sources, we use Oracle Data Integrator (ODI) or custom solutions in Python, which are then executed using ODI. We separated the critical processes that require the most up-to-date data from the standard and cyclical reporting ones. The product of the ETL process is clean, organised, and thematically grouped sets of data (data marts). Data structures that form data marts may contain integrated data from multiple systems or areas. The 'golden record', which is prepared as part of the data processing process, is utilised fully by RAD-on, and serves as a reference

for other OPI PIB and external systems, such as those used by higher education institutions. In early 2020, a new component was added to the data warehouse architecture: Oracle APEX, a low-code rapid application development (RAD) platform, which simplified the integration of OPI PIB systems and the data warehouse through the use of fast and easy-to-implement REST API services. The benefits of RAD technology expanded the possibilities of using the data warehouse in communication with users. New, user-friendly interfaces enable users to send their data to the warehouse, eliminating the need for a team of programmers.

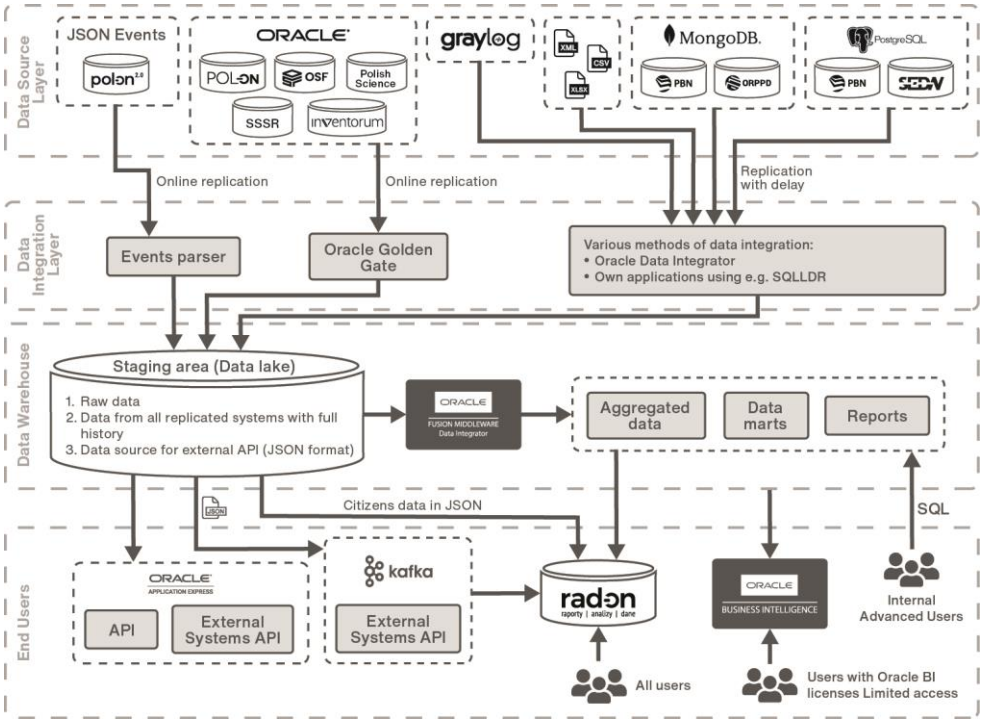

Figure 1: Overall schema of internal architecture of RAD-on

The right data structure, and communication and information exchange between the various elements of RAD-on, as well as between users and RAD-on in the form of a web application, are ensured by REST API. REST API allows for communication between RAD-on services and modules, ensuring that they work together seamlessly. Currently, RAD-on provides data through both the REST API and an API in the form of a graphical interface. RAD-on provides standard data categories that form the foundation of the system's services and modules.

The last crucial component of the system architecture is the report builder, which is used to present statistics in the reporting module. The report builder is an analytical platform developed independently by OPI PIB, which enables the institute to use it without incurring licensing fees. The reports engine is a service that supports the generation of HTML pages and output files that contain graphical representations of data. The data is edited by data analysts to create specific reports, which can be published on public and private websites. The data presented in a single report can come from multiple queries that refer to various relational databases and flat files, such as CSV and XLSX. To manage these dependencies, the architecture of the analytical platform is divided into four layers of abstraction: connectors, cache, queries, and reports. Each layer is responsible for managing one type of object. The responsibilities of each layer are as follows:

1) The connector layer is responsible for defining and managing connections to external data sources. In the case of a relational database, a connector is a pool of connections to any database.

2) The cache layer allows for faster access to data. The analytical platform enables the creation of reports and the execution of queries directly using source databases. A local H2 database, which can store copies of data from external sources, eliminates the need for continuous querying of the source servers.

3) The query layer enables the definition and execution of queries. A query object consists of a reference to a connector object, query content (SQL), and an optional set of specific parameters. If a query is parameterised, the values of the defined parameters should be transferred during its execution. This enables the creation of filters that can be used in unique reports.

4) The report layer contains report definitions. Report specification is a complex object for which a special set is defined. The set contains a list of queries to be used in the report, a report parameters list that matches the parameters of previous queries, and a report section list. During the report generation process, specification objects are transformed into result objects that contain data to be presented to end users. The engine (the front-end) transforms the result objects into ready-to-use pages that contain reports with data, which can be presented to users. By using layers and sections, analysts can create various reports and upload ready-to-publish dashboards to any website.

Technology	**Main task**	**Justification of the chosen technology**
Java Enterprise Edition JDK8 Spring Hibernate	Programming tools	A classic tech stack for developing enterprise-class systems
Oracle Enterprise Edition database	Data gathering, aggregating, and sharing	A reliable dataset technology in enterprise-class systems used by the key OPI PIB systems to facilitate data integration
Oracle Data Integrator	Data integration	ETL process management
Oracle Golden Gate		Real-mode data replication from source systems (Oracle)
Oracle Analytics Server	BI tool	Information desktops, self-service reporting, email report distribution
Oracle APEX	Low-code platform	Fast implementation of web applications and RestAPI interfaces
Apache Kafka	Message broker	A free, extremely efficient, and distributed platform to manage the exchange of messages between systems.
Elasticsearch	Data indexing and searching	Implementation of full-text search
Tomcat	Application server	A free application server
Graylog	Application monitoring	A free system that analyses application logs
Swagger	Documentation	Self-documenting Rest API services
Kubernetes	Orchestration	The most popular tool that manages, automates, and scales container applications

Table 1: The main technologies used by RAD-on

Table 1 presents the summary of technologies used for different RAD-on components, including the above-mentioned data analysis tool that provides tailor-made interactive reports displayed on the project's website.

6 Innovativeness

To our knowledge, RAD-on is a unique system of its type. In comparison to systems developed in other European countries, RAD-on stands out in terms of: (1) its comprehensive range of official data—RAD-on integrates information on didactics and research conducted by research institutions in Poland, collected in accordance with legal regulations; (2) its high degree of data disaggregation and interoperability; and (3) its extensive portfolio of services.

Many countries are developing nationwide research information systems that collect data on scholars and their research, and facilitate the sharing of this information between services to support, for example, the equitable distribution of research funds. According to EuroCRIS presentations and publications, such systems have been or are being developed by at least thirty countries[§§]. One of the most recent projects is Research Portal Denmark[***], which was launched in 2022. It contains data on science from various sources, including from research funding institutions, but it lacks data on higher education. Unlike other nationwide EuroCRIS systems, RAD-on provides access to a comprehensive selection of official data: it integrates information on research conducted in Poland, as well as data on the didactic activity of universities.

Regarding data on higher education, RAD-on is comparable to systems developed by international organisations, such as Eurostat, and those employed by national statistical offices. In 2014, the European Tertiary Education Register (ETER) was introduced with the goal of providing data on higher education institutions across Europe. ETER is dissimilar to RAD-on because during the data collection and sharing process, it does not rely on integrated databases[†††]. More akin to RAD-on is the system developed by the Higher Education Statistics Agency (HESA) that offers open access to data on higher education in Wales, England, Scotland, and Northern Ireland. The HESA system is constantly modernised and currently provides access to not only data on students, graduates and academic staff, but also to business intelligence tools that enable the data's filtering and visualisation (Burke et al., 2018). Unlike RAD-on, HESA does not support data on science and machine processing services.

The HEMIS and CRIS systems used by independent institutions usually offer multilayer analyses using highly detailed data. A major advantage of these systems is their ability to be tailored to the unique needs of higher education institutions or groups of such institutions. Their most notable drawback is their lack of capability to generate reports that can compare institutions for the purpose of decision-making processes within individual entities and, more broadly, across regions. RAD-on is a nationwide system providing access to a highly comparable data on science and higher education institutions. Users can browse this data, download it in various formats, and use it in their own systems via API.

7 Project impact: Use cases

As RAD-on is a nationwide system, we are able to monitor its impact on the entire sector of science and higher education in Poland. We are particularly interested in the extent to which the project's objectives have been fulfilled.

7.1 Open access to up-to-date and reliable data

RAD-on has successfully integrated 11.45 TB of data on science and higher education from nine distinct source systems. This has facilitated the efficient management of key processes of the Polish

[§§] https://dspacecris.eurocris.org (last accessed: 9 November, 2022).
[***] https://forskningsportal.dk (last accessed: 9 November, 2022).
[†††] https://eter-project.com/

Ministry of Science and Education and research institutions, including the allocation of didactic funds or the provision of data for official reports prepared by Statistics Poland. One of the most notable applications of a data warehouse as an integrator was the national research evaluation exercise in 2022[‡‡‡] (Michajłowicz et al., 2022). The data warehouse collected data from POL-on, PBN, and Web of Science, cleaned it, and transformed it into the target formats. The data was then uploaded to OPI PIB's System for Evaluation of Scientific Achievements (*System Ewaluacji Dorobku Naukowego, SEDN*).

The provision of open access to databases via RAD-on is a gradual process. Between 2020 and 2021, the volume of data shared increased by approximately 25%. This translates into growing popularity for RAD-on. In 2021, documents were viewed or downloaded over 150 million times, which represents a five-fold increase compared to 2020.

Between December 2021 and November 2022, the number of individuals and entities that use RAD-on data in their systems rose significantly. API was most commonly used to download data on research institutions in Poland (105,004,023 downloads), scholars' publications (455,962 downloads), scholarly staff (193,142 downloads), and university programmes (106,414 downloads). As mentioned earlier, between March 2022 and March 2023 RAD-on's website had 114 thousand unique users, mostly from Poland.

RAD-on is used by the Polish National Agency for Academic Exchange (*Narodowa Agencja Wymiany Akademickiej*, NAWA), which is responsible for the international promotion of information on the Polish science and higher education system. NAWA downloads information on higher education institutions in Poland and university programmes. The agency's access to official and regularly updated data enables it to organise the international education recognition process more efficiently. NAWA transmits RAD-on data to representatives of foreign higher education institutions, education recognition centres, and student organisations.

Another institution that uses RAD-on data is the Educational Research Institute (*Instytut Badań Edukacyjnych*, IBE), which conducts interdisciplinary research on the functioning and efficiency of the education system in Poland. With access to machine public data sharing (RAD-on API), the IBE can easily and automatically populate its register of qualifications awarded by higher education entities with data on university programmes.

RAD-on API's users also include foreign entities. As part of the Network4Growth programme[§§§], UNICO downloads data on scholars in Poland, as well as their publications, patents, and projects. The Network4Growth initiative is establishing a network of technology transfer professionals from universities and research institutes in the Visegrad countries (Czechia, Hungary, Poland and Slovakia), as well as Georgia and Armenia.

RAD-on data is used in the submission of applications for research grants in the Funding Stream Support System (*Obsługa Strumieni Finansowania*, OSF). OSF is integrated with the RAD-on API service, which provides official data on entities that operate in the science and higher education sectors. Because of this solution, entities that submit applications need not re-enter data that is already available in another public register.

7.2 Support of decision-making processes

Reports and analyses published on the RAD-on platform enable the monitoring of processes in the science and higher education sectors, and reach a wide public, including representatives of ministries and higher education institutions. Its key contributions include periodic analyses of the internationalisation of higher education in Poland prepared for education and business foundations, reports and analyses on the circumstances of graduates on the labour market by sector, and analyses of

[‡‡‡] The process results in the awarding of scholarly categories to all evaluated entities, granting them access to funds and authorisation to run doctoral schools.

[§§§] https://www.v4transfer.com.

the popularity of technology programmes among women. RAD-on data is also used to create national rankings of Polish higher education institutions and university programmes, which help middle school graduates make informed decisions on their future educational paths.

Between November 2021 and January 2023, one item on average was presented in the press or on the radio that mentioned RAD-on data. In that period, approximately 250 items about or references to RAD-on or its data appeared in online media. Journalists discussed statistics regarding foreign students, the percentage of female students who pursued technical programmes, and the accomplishments of scientists in artificial intelligence.

8 Conclusions

To our knowledge RAD-on is currently the largest (in terms of collected and shared data) national IT system in the domain of science and higher education in Europe. The system can be considered an analytical platform that presents official statistics, adheres to the concepts of OGD and FAIR, and contributes to the establishment of data-driven political and managerial processes. RAD-on supplements the IT ecosystem in science and higher education in Poland, which has been developed by OPI PIB for over ten years. RAD-on can inspire the creation of analogous systems in other regions and countries.

RAD-on evolves continuously, and this has resulted in steady increases in data volume and ever-more useful services. Planned works include the creation of scientometric reports and analyses, collaboration with research funding institutions in the field of sharing data on research grants in Poland, promotion of the API's use by universities' internal IT systems and researchers, and further development of analytical tools for data visualisation.

The most considerable challenge we must face is the efficient adaptation of our services to the ever-evolving legal landscape. During the development of the project, the science and higher education system underwent significant transformations. Legal changes to the way data could be collected has disrupted the continuity of its analysis and prevented the automation of some analytical processes.

Other challenges include some that pertain to the UX layer of the analytical platform. Based on the user satisfaction survey, we estimate that one in five users has problems finding the data they seek. In most cases, the data is unavailable due to its nonpublic legal status; in others, it is unavailable because the opening process remains incomplete. The highest-rated aspect of the system is its aesthetic layer (64% positive and 29% neutral opinions), data presentation form, and the website navigation; the lowest-rated is the ease of finding information (50% positive, 28% neutral, and 22% negative). Users frequently highlight the unavailability of certain features in the mobile version of the system, but are satisfied with the growing selection of reports.

RAD-on is designed to benefit all of its users, regardless of their digital literacy or familiarity with the higher education and research landscape. We are confronted with the challenge of striking a balance between providing accurate data descriptions and ensuring that private individuals and journalists can comprehend them. Despite consultations with UX specialists, we continue to lack guidelines on how to design public sector data dashboards (Abduldaem & Gravell, 2019). We must remain aware that users become accustomed to interface layouts and do not appreciate frequent changes to functionalities. We believe that this issue merits further exploration and investigation.

9 References / Citations

Abduldaem, A., & Gravell, A. (2019, February 6). *Principles for the Design and Development of Dashboards: Literature Review*. Retrieved December 14, 2022, from:

https://www.semanticscholar.org/paper/PRINCIPLES-FOR-THE-DESIGN-AND-DEVELOPMENT-OF-REVIEW-Abduldaem-Gravell/7627810cd27c280eedcec24f7b62471a7377eabd.

Burke, S., MacIntyre, R., & Stone, G. (2018). *Library data labs: Using an agile approach to develop library analytics in UK higher education*. Information and Learning Science, 119(1/2), 5–15.

Dombrowski, U., Mielke, T., & Engel, C. (2012). *Knowledge Management in Lean Production Systems*. Procedia CIRP, 3, 436–441.

Gao, Y., Janssen, M., & Zhang, C. (2021). *Understanding the evolution of open government data research: Towards open data sustainability and smartness*. International Review of Administrative Sciences, Retrieved December 14, 2022, from: https://doi.org/10.1177/00208523211009955.

Huston, P., Edge, V. L., & Bernier, E. (2019). *Reaping the benefits of Open Data in public health*. Canada Communicable Disease Report, 45(11), 252–256.

Michajłowicz, M., Kozłowski, M., Kowalski, M., Fijałkowski, S., Drogosz, J., Bylina, M., & Furmankowska-Podniesińska, A. (2022). *Scientific activity evaluation in Poland: The IT ecosystem and the optimal selection of achievements.* EPiC Series in Computing, 86, 56–65.

Michajłowicz, M., Niemczyk, M., Protasiewicz, J., & Mroczkowska, K. (2019). *POL-on: The Information System of Science and Higher Education in Poland*. EUNIS 2018 Congress Book of Proceedings.

Ruijer, E., Détienne, F., Baker, M., Groff, J., & Meijer, A. J. (2020). *The Politics of Open Government Data: Understanding Organizational Responses to Pressure for More Transparency*. The American Review of Public Administration, 50(3), 260–274.

Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). *Open data and its usability: An empirical view from the Citizen's perspective*. Information Systems Frontiers, 19(2), 285–300.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(160018). Retrieved December 12, 2022, from: https://doi.org/10.1038/sdata.2016.18.

10 Author biographies

Jarosław Protasiewicz (PhD): experienced researcher, lecturer, and IT professional.

Jarosław acquired his master's degree at the Białystok Technical University, Poland, and he defended his doctoral dissertation at the Systems Research Institute of the Polish Academy of Sciences. Both theses concerned artificial neural networks.

Jarosław's research interests include software design and development, artificial intelligence, and machine learning. His scientific career has long been interwoven with the IT industry. He has extensive IT experience as a software developer, designer, and project manager.

Since 2005, Jarosław has been employed by the National Information Processing Institute (OPI PIB) in Warsaw, Poland, where he initially served as a software developer and designer. Then, as an associate professor, he established and managed the Laboratory of Intelligent Information Systems—the largest laboratory of the institute. Currently, he serves as the head of OPI PIB.

Email: jaroslaw.protasiewicz@opi.org.pl
LinkedIn: https://pl.linkedin.com/in/jarosław-protasiewicz-26749a2a

Aldona Tomczyńska (PhD): is a doctor of philosophy in international political economy. She was awarded her doctorate degree with distinction at the University of Warsaw. She has worked as an assistant professor and data science team leader at the National Information Processing Institute for more than ten years. She conducts research in sociology, and data science.

Email: aldona.tomczynska@opi.org.pl
LinkedIn: http://linkedin.com/in/aldona-tomczyńska-phd-8490b718a

Emil Podwysocki (MSc) obtained a master's degree in telecommunications systems at the Technical University of Lodz in 2009. He has over ten years' professional experience related to ETL/ELT, data warehousing, and business intelligence. His areas of interest include Oracle technology, big data, business intelligence, low-code or no-code technology, and data visualisation. Head of the Laboratory of Databases and Business Analytical Systems at the National Information Processing Institute.

Email: emil.podwysocki@opi.org.pl
LinkedIn: www.linkedin.com/in/emil-podwysocki

Sylwia Ostrowska (MSc): is a business analyst and project manager at the National Information Processing Institute. She obtained a master's degree in Management at the University of Warsaw and completed postgraduate studies in information system design at Warsaw University of Technology. Currently, she is the project manager of RAD-on. She holds PRINCE2 Foundation and Professional Scrum Product Owner certificates.

Email: sylwia.ostrowska@opi.org.pl
LinkedIn: https://www.linkedin.com/in/sylwia-ostrowska-06749b23