Technical Foundations and Applications of Trustworthy AI Systems

Otmane Azeroual
https://orcid.org/0000-0002-5225-389X
University of Hagen, Germany

IGI Global
Scientific Publishing
Publishing Tomorrow's Research Today

Vice President of Editorial — Melissa Wagner
Director of Acquisitions — Mikaela Felty
Director of Book Development — Jocelynn Hessler
Production Manager — Mike Brehm
Cover Design — Jose Rosado

Published in the United States of America by
IGI Global Scientific Publishing
701 East Chocolate Avenue
Hershey, PA, 17033, USA
Tel: 717-533-8845 | Fax: 717-533-7115
Website: https://www.igi-global.com E-mail: cust@igi-global.com

Copyright © 2026 by IGI Global Scientific Publishing. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Use of this publication to train generative artificial intelligence (AI) technologies is expressly prohibited. The publisher reserves all rights to license its use for generative AI training and machine learning model development.

Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global Scientific Publishing of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

(CIP Data Pending)
ISBN13: 9798260014332
Isbn13Softcover: 9798260014349
EISBN13: 9798260014356
British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material.
The views expressed in this book are those of the authors, but not necessarily of the publisher.
This book contains information sourced from authentic and highly regarded references, with reasonable efforts made to ensure the reliability of the data and information presented. The authors, editors, and publisher believe the information in this book to be accurate and true as of the date of publication. Every effort has been made to trace and credit the copyright holders of all materials included. However, the authors, editors, and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. Should any copyright material be found unacknowledged, please inform the publisher so that corrections may be made in future reprints.

Chapter 8
The Ecosystem of State Support for Science in Slovakia and Scientific Data

Danica Zendulková
http://orcid.org/0000-0003-2487-0177
Slovak Centre of Scientific and Technical Information, Slovakia

Gabriela Gavurníková
http://orcid.org/0000-0001-7096-6378
Slovak Centre of Scientific and Technical Information, Slovakia

Ján Turňa
http://orcid.org/0000-0001-5813-566X
Slovak Centre of Scientific and Technical Information, Slovakia

ABSTRACT

The chapter examines the science and research ecosystem in Slovakia, including its constituent elements, actors, processes, and instruments, with particular emphasis on the organization of research information and research data management. Credibility and transparency constitute fundamental pillars of the contemporary scientific ecosystem. In a context where science increasingly depends on open data, interdisciplinary collaboration, and public funding, these principles acquire critical significance for the effective functioning of research and innovation. The first section explores national science policy and its key principles such as open science, research ethics, and gender equality in science and research. The second section focuses on the prerequisites for data-driven policymaking in the field of science and research.

DOI: 10.4018/979-8-2600-1433-2.ch008

Copyright © 2026, IGI Global Scientific Publishing. Copying or distributing in print or electronic forms without written permission of IGI Global Scientific Publishing is prohibited. Use of this chapter to train generative artificial intelligence (AI) technologies is expressly prohibited. The publisher reserves all rights to license its use for generative AI training and machine learning model development.

It addresses standards and models for data recording within information systems supporting scientific activity, the processing of research data, and the integration of such systems as essential conditions for producing high-quality and trustworthy data.

INTRODUCTION

The science and research ecosystem is understood as the set of institutions, actors, processes, and instruments that collectively ensure the functioning of scientific research activities. Credibility and transparency constitute fundamental pillars of the modern scientific ecosystem. In an environment where science increasingly relies on open data, interdisciplinary collaboration, and public funding, these principles acquire critical importance for the effective operation of research and innovation.

The ecosystem of scientific information and its governance is currently regarded as essential for ensuring that science functions effectively and serves its intended purpose (Altman & Cohen, 2022). According to these authors, fundamental principles and values also include fairness and equity, as well as adherence **to** information ethics and sustainability.

In Slovakia, the ecosystem of state support for science and scientific data comprises several interconnected levels: from policy-makers, through funding and evaluation bodies, to institutions, researchers, and research infrastructures. This ecosystem is gradually being shaped by reforms aimed at increasing openness, fairness, and accountability in the allocation of financial resources and in making research outputs accessible to the public.

Transparency in science encompasses not only the accessibility of data, methodologies, and research results, but also clearly defined and verifiable procedures for decision-making on research funding. The credibility of the system, in turn, depends on its integrity, the professional quality of evaluation processes, and independence from political or institutional influences. Together, these elements form the foundation of an environment in which researchers, institutions, and the broader public can trust that public resources are used efficiently and in accordance with the public interest.

In the Slovak context, issues of credibility and transparency relate particularly to the functioning of grant agencies, the management of scientific data, open access, and the implementation of open science principles. However, a robust ecosystem is needed not only where research outputs are openly shared, but also where transparency permeates the research process from the outset, enabling more rigorous and collaborative research (Thibault et al., 2023).

The aim of this section is to characterize the fundamental components of state support for science and of the processes of collecting, processing, preserving, and

disseminating scientific information and data in Slovakia, and to analyse the extent to which trust and transparency are embedded within the existing ecosystem.

STATE SCIENCE POLICY

State science policy in Slovakia may be defined as the set of objectives, principles, and instruments through which the state directs and supports research, development, and innovation. It encompasses the creation of legislative and financial instruments, the coordination of activities among various public bodies, and the funding of research and development.

The state science policy of the Slovak Republic focuses primarily on:

- the development of research, development, and innovation (RDI) as key determinants of competitiveness;
- the support of excellent science and applied research;
- the interconnection of the academic sector, the business community, and public administration;
- the development of human resources in science;
- the internationalization of Slovak research (cooperation with the EU and OECD);
- the transfer of knowledge into practice (technology transfer, start-ups, patents).

In addition to the reorganisation of the system, a central priority is increasing public investment in research, development, and innovation in line with the national strategy, as well as simplifying procedures—specifically public procurement, auditing, and financial control in the area of research, development, and innovation.

The governance system for science and research in Slovakia has a multi-level structure, which will be examined in greater detail in the following paragraphs.

Model of State Support for Research and Development

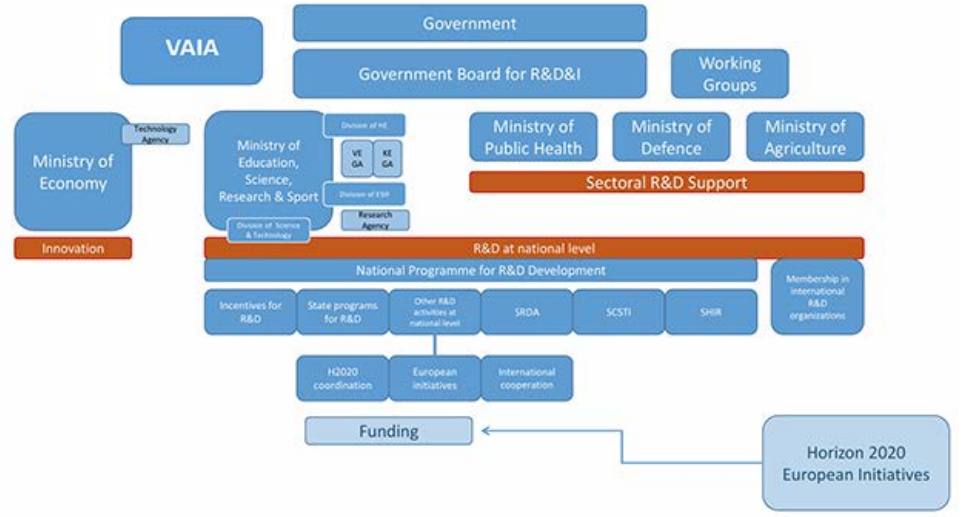
The model of state science policy in Slovakia represents the system through which the state governs, supports, and advances science, research, and innovation. This system is coordinated by the state through a range of institutions, strategic documents, and financial instruments. It combines direct and indirect support from the state budget, complemented by European funding sources and private investment. The model is grounded in the principles of strategic management (setting priorities and objectives), programme-based financing (support through grants, projects, and

programmes), competitiveness and quality evaluation, and cooperation between the academic and business sectors.

Its purpose is to ensure the interconnection of strategic decision-making, coordination, implementation, and evaluation of scientific activities.

Figure 1 provides a graphical representation of the model of state support.

Figure 1. Main stakeholders of the R&D ecosystem in Slovakia



At the strategic level, which determines the overall direction, priorities, and objectives of state science and innovation policy, the central authorities of public administration operate. These include, above all, the Government of the Slovak Republic, the Government Council for Research, Development and Innovation, and the Ministry of Education, Research, Development and Youth of the Slovak Republic (MŠVVaM SR), which has historically been responsible for state science policy (Ministry of Education, Research, Development and Youth of the Slovak Republic, 2024a). A gradual shift is currently taking place, with responsibility for strategy formation and coordination moving from the ministerial level to the supra-ministerial level, specifically to the Office of the Government of the Slovak Republic.

In 2021, the Research and Innovation Authority (VAIA) was established within the Office of the Government as a coordination body tasked with reforming state support for science, research, and the innovation ecosystem. VAIA was created in response to the need to strategically guide and coordinate research and innovation policies in areas that do not clearly fall within the competences of existing ministries. It functions as the executive body and secretariat of the Government Council for Science, Technology, and Innovation.

Since 2025, VAIA has been incorporated into the Office of the Deputy Prime Minister of the Slovak Republic for the Recovery Plan and the Knowledge Economy. Its primary role is the implementation of the *National Strategy for Research, Development and Innovation 2030*, as well as the coordination and management of the Government Council for Science, Technology, and Innovation, whose competences were transferred from the Ministry of Education. In addition to introducing international expert evaluation of research and innovation, VAIA is responsible for developing talent policies and designing a new model of governance and funding for research and development under the Government Council for Science, Technology, and Innovation. However, the process of transferring competences has not yet been completed or fully embedded in legislation. Planned steps include drafting a new Act on Research, Development and Innovation and reorganising and consolidating ministries, agencies, and institutions responsible for policymaking and support in R&D&I.

The executive or implementation level consists of grant agencies that provide funding for science and research through grants, programmes, and projects.

The operational level is where research and development are carried out; it comprises the institutions and organisations engaged in scientific research across various sectors.

Another level consists of monitoring, evaluation, control, and support activities, in which research information and the tools for processing it, namely information systems supporting science, play a crucial role.

Strategic Documents

National strategic documents in the field of science policy build upon EU-wide guidelines and methodological frameworks intended for the Member States.

The central document of the science and research ecosystem in Slovakia is the *National Strategy for Research, Development and Innovation* (VAIA, 2023), which outlines the objectives, priorities, and intentions of the Slovak Republic, as well as principles of good governance and effectiveness in the field of research, development and innovation. The strategy is accompanied by an **Action Plan**, which is updated on an ongoing basis. The implementation of the strategy is overseen by the Government Council for Science, Technology and Innovation.

The *Roadmap of Research Infrastructures – SK VI Roadmap 2020–2030* (Ministry of Education, Science, Research and Sport, 2021a) constitutes a key national document for research infrastructures. It maps not only the historical development but also the current status of major public and private research infrastructures in Slovakia.

Since June 2021, Slovakia has implemented the *National Strategy for Open Science 2021–2028* (Ministry of Education, Science, Research and Sport, 2021b), approved by the Government of the Slovak Republic. This strategy defines the steps necessary to implement the principles of open science. Its objective is to create a sustainable model enabling all actors in the Slovak research ecosystem to participate in adopting global open science principles: promoting the dissemination, exchange, reuse, and accessibility of research outputs. In doing so, it aims to strengthen domestic and international scientific collaboration, foster cooperation between the public and private sectors, enhance research infrastructure, support the education of new generations of researchers and university educators, and create conditions for effective strategic governance of research and education. The strategy is implemented through biennial Action Plans.

On 28 August 2024, the Government of the Slovak Republic approved the *Code of Research Integrity and Ethics in Slovakia* (Ministry of Education, Research, Development and Youth, 2024b). This Code is a binding document that establishes ethical requirements and principles for all individuals working in science and research, as well as for organisations involved in research activities. Its aim is to strengthen the quality, credibility, and professionalism of scientific research in Slovakia.

Strategic documents also include Gender Equality Plans (GEPs), which set objectives and measures to promote gender equality within research institutions. Their development is aligned with the activities of the European Commission, which issued the *Gender Equality Strategy 2020–2025* (European Commission, 2020). Since 2022, the existence of a Gender Equality Plan has been a mandatory requirement for participation in the Horizon Europe funding programme.

Legislation

The research, development, and innovation ecosystem in Slovakia is governed by several laws:

- Act No. 172/2005 Coll. on the Organisation of State Support for Research and Development
- Act No. 185/2009 Coll. on Incentives for Research and Development
- Act No. 131/2002 on Universities
- Act No. 133/2002 Coll. on the Slovak Academy of Sciences
- Act No. 243/2017 on Public Research Institutions

The fundamental law is the Act on State Support for Research and Development, which defines the legal framework for supporting science. The Act on Incentives regulates the operation of one of the grant schemes for funding research. The Act on

Universities and the Act on the Slovak Academy of Sciences govern the functioning of key research organisations in Slovakia. The latter law complements the Act on the Slovak Academy of Sciences and defines the transformation of SAS institutes, which prior to its enactment operated as budgetary and contributory organisations, into public research institutions. The significance of this law lies primarily in the regulation of research infrastructure management and financing, enabling these institutions to become competitive within the European Research Area.

Sectoral Structure of Research and Development Actors

The sectoral structure of science and research refers to the way in which entities conducting research and development (R&D) activities are classified according to their founding authority, legal status, and primary mission. This concept enables the analysis and comparison of different components of the research system—for example, the university sector, the government sector, the business sector, and non-profit organisations. The purpose of sectoral classification is to provide an overview of which actors participate in R&D, what resources they use, and what outputs they generate. Such classification is also essential for the development of science and technology policy, research funding, and international statistical comparison following OECD and Eurostat methodologies.

The sectoral structure of Slovak science and research is defined by the Act No. 172/2005 Coll. on State Support for Research and Development as follows:

- **Government sector**, comprising legal entities established by central government authorities;
- **Public research institutions sector**, consisting of institutes of the Slovak Academy of Sciences;
- **Higher education sector**, including public, state, and private universities, as well as legal entities founded by them that carry out research and development;
- **Non-profit sector**, comprising non-profit organisations, civic associations, and associations of legal entities engaged in research and development;
- **Business enterprise sector**, consisting of business organisations conducting research and development as part of their activities.

The most significant actors are the Slovak Academy of Sciences, with its 47 institutes functioning as public research institutions, and universities. In Slovakia, there are 20 public universities, 3 state universities, 10 private universities, and 2 foreign universities.

Other organisations (state, private, or non-profit) declare their involvement in science and research in several ways. First, research and development must be explicitly stated in their founding charter (in the case of state organisations) or in their business activity description (for private or commercial entities).

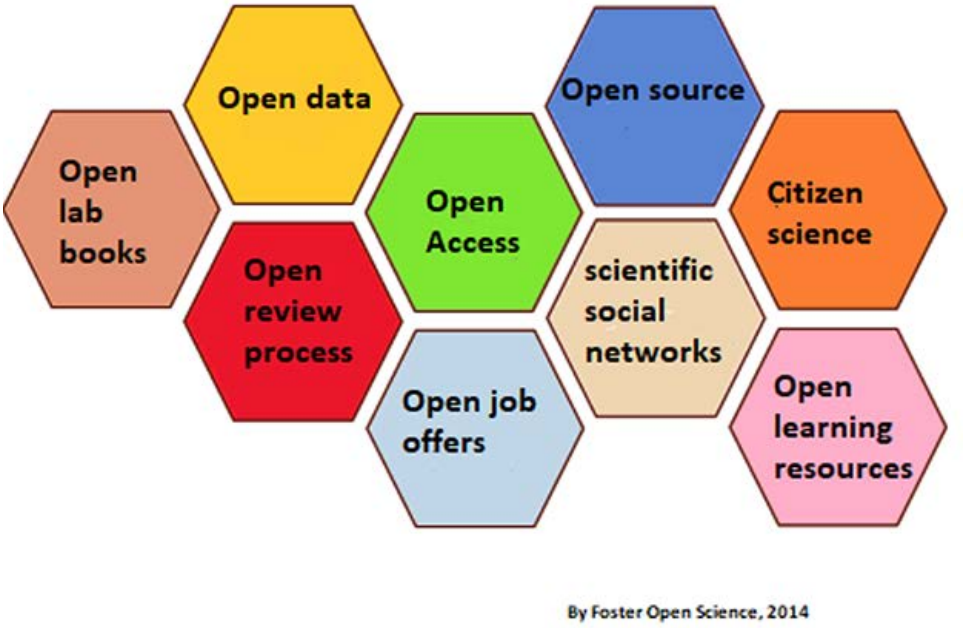
In addition, all such entities must apply for a certificate of eligibility to conduct research and development if they wish to apply for state support. The certificate is issued based on meeting several requirements, including laboratory infrastructure, personnel capacity, previous research activities, and research plans.

This process results in the creation of a database of research organisations within the SK CRIS system, which serves as a reference register for Slovak grant agencies. Before awarding a grant, agencies verify through the database whether the applicant organisation holds a valid certificate of eligibility to conduct R&D. The number of certificate holders has remained relatively stable over time at around 700, the majority of which are business enterprises.

Open Science

Open Science "represents an approach to the research process based on collaboration, sharing, and the open dissemination of knowledge through the optimal use of digital technologies and the application of openness principles in every aspect of the research cycle" (Foster Open Science, 2020). The term *Open Science* is often used as an umbrella concept encompassing various other "movements aimed at removing barriers to the sharing of any outputs, resources, methods, or tools at any stage of the research process" (Foster Open Science, 2020). It also includes numerous related concepts, as illustrated in Figure 2.

Figure 2. Open Science



UNESCO, in its *Recommendation on Open Science* (2021), defines Open Science as an inclusive concept that brings together different movements and practices aimed at making multilingual scientific knowledge openly accessible and reusable for everyone. To enhancing scientific collaboration and information sharing for the benefit of science and society, and opening the processes of knowledge creation, evaluation, and communication to societal actors beyond the traditional scientific community. Open Science spans all scientific disciplines and aspects of scientific practice, including basic and applied sciences, natural and social sciences, and the humanities, and rests on the following key pillars: open scientific knowledge, open scientific infrastructures, scientific communication, open engagement of societal actors, and open dialogue with other knowledge systems.

Open Science promotes transparency, collaboration, and free access to research outputs (Mons, 2019). It increases the efficiency, reliability, and societal responsiveness of scientific research (Bordons et al., 2023).

Individual countries develop Open Science policies, a set of guidelines, rules, regulations, laws, principles, or instructions—intended to implement the values and principles of Open Science in practice. In Slovakia, research institutions have access to the document *Methodological Recommendations for the Development*

and Implementation of Institutional Open Science Policies for Research Institutions when formulating their own Open Science policies.

Open Science also provides an organisational and policy framework for implementing the FAIR principles (Burgelman et al., 2019).

FAIR Principles

Open data are defined as data that anyone can freely use and redistribute, provided that proper attribution and citation are given. They are considered digital information that is available at any time to any user (Wilkinson et al., 2016). Open data are freely accessible without any restrictions. In addition, the FAIR principles focus on the technical and methodological quality of data and their reusability. It is important to emphasize that these principles do not apply only to research data, but also to research information stored in open research infrastructures, including CRIS systems (Hauschke et al., 2021).

The key characteristics of open data are accessibility, reusability, and universal participation without discrimination. In the context of open data, the concept of FAIR data is frequently mentioned. FAIR data are an essential component of Open Science and describe some of the central principles of good data management and open access to research data.

FAIR is an acronym representing four principles that ensure that data are:

Findable

The first step in using data is their findability. Data should be described with metadata that includes information about the context, quality, conditions, and characteristics of the data. This enhances data discovery, reuse, and citation. Metadata should be assigned a persistent identifier, which helps eliminate ambiguity in published data. Metadata should also be registered or indexed in a discoverable source, as identifiers and metadata alone do not guarantee easy findability on the internet.

Accessible

Once a user has found the required data, they need to know how to access it. Metadata should be retrievable using a standard communication protocol that is open and universally implementable. Metadata should remain accessible even if the data itself is no longer available.

Interoperable

Metadata should use a formal, accessible, shared, and broadly applicable language for knowledge representation. To ensure automated findability and interoperability of data files, ontologies, thesauri, and data models should be applied. Metadata should also include references to other metadata. Creating links between metadata sources improves understanding of the data context.

Reusable

The main goal of FAIR principles is to optimize the reuse of data. Therefore, both data and metadata should be well described. Metadata should be richly annotated with precise and relevant attributes to enable reuse in a specific context. Key attributes include clear and accessible data usage licenses and precise provenance information. Data should be organised in standardized formats, supported by established and sustainable formats and controlled vocabularies.

In Slovakia, the importance of applying FAIR principles to research information and the corresponding systems for their management and processing is well recognized. The Current Research Information System (CRIS) is a key part of the research infrastructure, specifically designed to support and operationalize the implementation of FAIR principles (Azeroual et al., 2022).

Ethics in Science

Ethics in science in Slovakia is an essential part of both national science policy and everyday scientific practice. It constitutes a system of rules, principles, and institutions that ensure research is conducted honestly, transparently, and with respect for human rights, the environment, and societal values. Ethical principles also apply to the processing of scientific information and research data (Schöpfel & Azeroual, 2023), including their relationship to the application of Open Science principles (Schöpfel et al., 2022). The importance of these principles is increasing in light of the emerging trend of artificial intelligence utilization (Azeroual et al., 2024).

Scientific ethics concerns researchers' behaviour, the protection of research integrity, respect for the rights of research participants, accountability to society and the environment, and the prevention of research misconduct, including data falsification and plagiarism.

The institutional framework for ethics in science is provided by the Ethics Committee for Research and Development of the Slovak Republic, a national body established under the Ministry of Education, Science, Research, and Sport of the Slovak Republic. The Committee oversees adherence to ethical principles in science

and research, issues opinions on disputed cases, and provides recommendations and methodological guidance. The fundamental document is the Code of Research Integrity and Ethics in Slovakia. Every major research organisation maintains its own ethics committee.

Ethics in science is based on internationally recognized principles such as honesty, responsibility, objectivity, transparency, and respect. Applying these principles safeguards the credibility of science and its contribution to society.

Data Protection

Data protection in science involves compliance with the European General Data Protection Regulation (GDPR) and national legislation, particularly the right to personal data protection, as well as the Act on the Protection of Personal Rights. Scientific research may process personal data only under conditions such as obtaining consent from the data subject or fulfilling tasks in the public interest, while respecting the right to object to data processing.

In Slovakia, the relevant legislation on personal data protection includes:

- Regulation (EU) 2016/679 of the European Parliament and 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, repealing Directive 95/46/EC.
- Act No. 18/2018 Coll. on the Protection of Personal Data, as amended.

Regarding the collection of data on research, development, and innovation within the ecosystem, the legal basis for such data collection is the fulfillment of statutory obligations of the ministry (Act 172/2005 on State Support for Research and Development) and the execution of the Slovak Centre of Scientific and Technical Information (CVTI SR) tasks carried out in the public interest. The public interest is understood as the lawful processing of information on research and development activities funded by state budget resources (Zendulkova & Azeroual, 2022).

Gender Equality

Gender equality in science and research is a key priority of the new European Research Area (ERA). Promoting equality between women and men is a fundamental responsibility of the European Union in all its activities as required by the treaties. Gender equality is a core value of the EU, a fundamental right, and a key

principle of the European Pillar of Social Rights. It is also a crucial condition for an innovative, competitive, and prosperous European economy.

Institutional change strategies implemented through gender equality plans have had a very positive impact in EU countries over the last decade. According to the latest report on progress in the European Research Area, these strategies have acted as a catalyst for improvements in the quality of research and innovation.

In Slovakia, the key document is the *National Strategy for Gender Equality and Equal Opportunities in the Slovak Republic for 2021–2027* (Ministry of Labour, Social Affairs, and Family SR, 2021), which includes a dedicated chapter on gender equality in science, research, and education. This document is followed by institutional gender equality plans, which are mandatory for organisations applying for funding under certain EU programs.

The objectives of gender equality measures in science and research are to:

- Improve the representation of women in research and decision-making bodies.
- Eliminate gender bias in the evaluation of scientific work.
- Support work–life balance for researchers (ERAPortál, 2025).

Monitoring the state of gender equality is carried out both through statistical surveys and based on data from information systems (Zendulková et al., 2022).

Science and research policy defines the strategic direction of a country's scientific development. Its significance lies in effectively focusing research on areas with the greatest societal and economic impact.

A trustworthy and responsible science policy seeks to create transparent and stable funding sources for scientific and innovation activities, and it promotes Slovakia's integration into European and global research spaces. It aims to link science with practice and the economy so that research outcomes contribute to economic growth, new products, and value-added services. It also fosters social responsibility in science, transparency, and ethical research practices.

Data-driven decision-making ensures that policy is based on measurable, verifiable evidence. The availability and quality of data allow for better analyses, predictive modelling, impact monitoring, and decision transparency. A trustworthy, transparent, and responsible science policy must therefore be based on high-quality and accessible data. As Numajiri & Hayashi (2024) emphasize, the availability of open data and their use in analyses is a critical prerequisite not only for any data-driven research but also for credible science policy.

DATA ON SCIENCE AND RESEARCH AND THEIR PROCESSING

Data on science and research constitute a fundamental element of the modern governance of the research and innovation ecosystem. They enable the systematic monitoring of scientific capacities, activities, outputs and impacts, thereby forming the basis for effective policy-making, long-term strategic planning and international benchmarking.

A key actor in the field of science and research data processing in Slovakia is the CVTI SR. As an organisation directly managed by the Ministry of Education, Research, Development and Youth of the Slovak Republic (MŠVVaM SR), CVTI SR serves as the national information centre for science, technology, innovation and education, as well as the country's specialised scientific library. It coordinates and operates interdisciplinary research and development centres and national infrastructures for research, development, innovation and education. Its mission is to support the development of science, technology and education. Its tasks include the operation of research information systems, the management of library and information resources, the provision of information services for the expert community, methodological and analytical work to support governance and evaluation in research, development and higher education, the delivery of projects supporting research and education, and the popularisation of science and technology.

CVTI SR has long been responsible for building, developing and operating several nationwide information systems, tools and registers serving research institutions within the remit of the Ministry. These responsibilities are carried out on the basis of authorisations and delegated mandates. CVTI SR operates these systems within its own ICT infrastructure, the Data Centre for Research and Development.

Competencies of CVTI SR are derived from several tasks:

- Official international cooperation in the development of the European Open Science Cloud (EOSC).
- Delegated responsibility from the Ministry for operating and developing national information systems supporting research and development.
- Delegated responsibility for processing and providing access to data from national information systems and from global scientific databases.
- Cooperation with OpenAIRE in supporting open science and the dissemination of scientific information and data.

The operation of key national information systems for science and research in Slovakia addresses complex and globally relevant challenges: research data processing, support for open science, interoperability of information systems, the 'once-only' principle, open research data publishing, FAIR data management, support for

research evaluation, master data management, central identity management and the provision of ICT infrastructure services.

Among the key information systems and registers operated by CVTI SR are:

CREPČ – Central Register of Publication Activity, a comprehensive bibliographic database of publications produced by Slovak higher education institutions, used for the evaluation of publication activity, statistical reporting for university funding, accreditation processes, periodic evaluation of research, development, artistic and other creative outputs, and for the implementation of performance-based funding.

CREUČ – Central Register of Artistic Activity, a bibliographic database of artistic outputs of Slovak higher education institutions.

SK CRIS – National Research Information System, containing structured data on research projects undertaken by Slovak organisations, their outputs, a map of research infrastructure, a register of research organisations and a database of researchers. It supports capability assessments and statistical surveys of research and development potential and has previously been used to administer selected national grant schemes.

CIP VVI – Central Information Portal for Research, Development and Innovation, a core information, management and oversight tool of national science and technology policy.

CRZP and ANTIPLAG – Central Register of Theses and Dissertations and Antiplagiarism System, which securely stores digital theses and performs automated originality checks.

SCIDAP – National Open Digital Repository for Publications, preserving and providing access to full-text publications produced by Slovak research institutions.

SVD – National Open Digital Repository for Research Data, enabling the secure storage and open dissemination of research datasets.

OAPP – Open Access Publishing Platform, supporting the publication of open access journals and monographs.

Analytical and Statistical Module for Research Evaluation, which provides tools for complex analyses, benchmarking, predictive modelling and evaluation of the quantitative and qualitative performance of Slovak research. It integrates both internal and external data sources including Web of Science, citation databases, SCIDAP, SVD, CRZP, SK CRIS and electronic information resources.

ISS CVTI SR – Integrated Service System, providing ICT infrastructure services, access to integrated applications and tools supporting research and technology transfer, including PRIMO, HAN, ERMS, Mediainfo, the Open Data Platform and the NARCOS national system for respiratory diseases.

Beyond CVTI SR, several other institutions operate relevant information systems within the Slovak research ecosystem, particularly research funding agencies:

The **Slovak Research and Development Agency (APVV)**, established in 2005, is the sole national grant agency providing state funding for research projects, agency programmes and international cooperation initiatives. Its mission is to support high-quality basic and applied research, to contribute to the objectives of national science and technology policy and to stimulate participation in European and international research programmes.

The **Scientific Grant Agency (VEGA)** of the Ministry and the Slovak Academy of Sciences coordinates the evaluation of basic research projects undertaken at universities and research institutes.

The **Cultural and Educational Grant Agency (KEGA)** supports applied research in education, pedagogy and the arts.

In addition, the **ITMS governmental information system** provides a unified platform for recording, processing and monitoring EU structural funds programmes.

Many research organisations also operate their own institutional systems, such as the Slovak Academy of Sciences' databases of publications, projects and personnel, and university systems (e.g., the STU Academic Information System and the University of Žilina Institutional CRIS).

Data processing within these systems is shaped by different legislative frameworks and serves diverse purposes, resulting in heterogeneous processes and varying degrees of obligation for research actors.

Despite this heterogeneity, these data sources form a coherent ecosystem due to the use of international standards, data models and classifications for describing the entities and attributes represented in the databases. These shared principles provide the foundation for the development of a robust and high-quality research information infrastructure.

Metadata on Research and Development

The term *metadata,* that is, data about data, is generally used to describe the data structures in which information about individual objects is stored. In the context of metadata, a variety of data models and standards are employed to enable the integration of heterogeneous databases that contain similar types of content.

Metadata, together with their creation, standardisation and use, are themselves the subject of research. For example, Mayernik (2019) discusses metadata in a broad sense within the sociology of science, examining how researchers document, describe, annotate, organise and manage data in their everyday scientific practice. It is beyond doubt that the use of metadata standards enhances research reproducibility, and not only in computer science (Leipzig et al., 2021). In library and information science, a typical example of metadata is the "catalogue card", or, in contemporary

settings, a database record of a library information system describing a library document (Ukropová, 2002).

Metadata are often reduced to descriptive or administrative elements. In the case of a research project, this includes information such as the project title, project code, start and end dates, funding source, as well as a substantive description of the research in the form of an abstract, keywords and similar descriptors.

A clear overlap with library science emerges particularly in databases of research outputs (publications, patents, products). These outputs are linked to research actors (organisations, researchers) and to research activities (projects).

At this point, it is necessary to highlight standardisation as a mechanism that ensures uniform recording of such information by diverse stakeholders at the international level. The multidisciplinary application of metadata standards is addressed, for instance, by the Committee on Data for Science and Technology (CODATA), which also identifies numerous challenges in this area (Rumble et al., 2019). Traditional metadata standards for research data, used by Datacite, Dataverse Network, Dryad or FigShare, are examined by Farnel and Shiri (2014). However, these classical approaches, which historically focused on the organisation, classification and reuse of resources, are now undergoing rapid transformation with the increasing integration of artificial intelligence technologies (Yang et al., 2025).

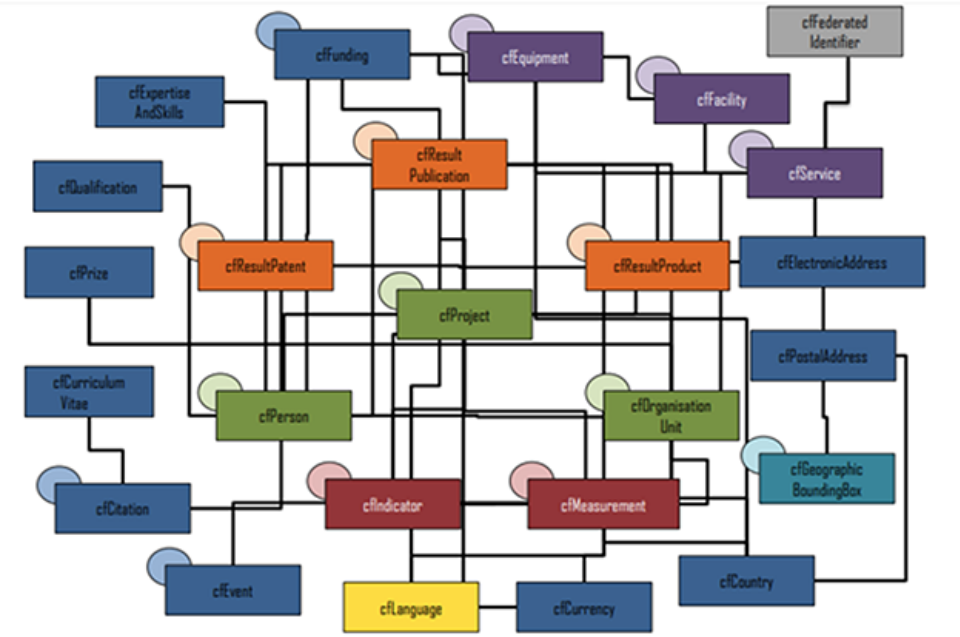
Among the metadata standards used for research and development information in Slovakia, the most important is the specialised CERIF format, alongside Dublin Core, designed for the description of electronic resources and commonly used by memory institutions. The MARC21 format for bibliographic metadata is also widely applied, primarily within library information systems, although it is supported by some research output databases as well.

CERIF (Common European Research Information Format) is an XML data format intended for the organisation of data on science and research. It is recommended by the European Commission as the standard for the automation of research information management (Interoperable Europe, Common European Research Information Format, 2017). At its core, CERIF provides a data model that enables metadata-based representation of research entities, their activities, relationships and outputs.

The elements of CERIF (Figure 3) have a defined structure, semantics and set of relationships, and are organised into hierarchically ordered categories:

1. **Core Entities** (Person, Project, Organisation)
2. **Result Entities** (Publications, Patents, Products)
3. **Second-Level Entities**
4. **Link Entities**

Figure 3. CERIF 1.6 (Dvořák, 2025)



The complete CERIF data model comprises several subsystems, a semantic layer, XML schemas, SQL scripts and implementation examples. An exchange format for metadata CERIF_XML has been developed. The use of the model, which is continuously being updated and expanded, is subject to a Creative Commons licence.

CRIS systems are capable of collecting a wide range of metadata for multiple entities related to research activities. However, as the open-science environment becomes increasingly complex, new opportunities arise for these systems to store even greater volumes of information in a highly interconnected manner, characteristic of the CERIF data model (de Castro, 2024).

Based on more than ten years of experience with the use of the relational CERIF data format, we can conclude that it has proven highly effective. The implementation of metadata within research information systems in Slovakia enables:

- improved discovery and sharing of research information and data;
- thorough documentation of the provenance and context of research;
- assessment of the quality and reproducibility of research outputs;
- compliance with the requirements of open science and those of funding agencies;
- international comparability of research activities;

- participation in international research infrastructures and projects.

Scientific Data, Their Processing and Secure Storage

In Slovakia, scientific data are systematically collected across most scientific disciplines. In the natural sciences, these include genetic and molecular data, monitoring of populations and invasive species, climate and environmental data, as well as geological mapping and various laboratory measurements. Scientific data also encompass data models in artificial intelligence, data from transport and energy systems, economic indicators and meteorological data. A distinct category comprises scientific data in the fields of medical sciences, biomedicine, genetics and bioinformatics.

The processing and secure storage of scientific data in Slovakia are carried out in accordance with the latest European trends and requirements, with a strong emphasis on the FAIR principles. Data governance follows international Open Science principles and legislation on personal data protection. The processing of personal data is regulated by the General Data Protection Regulation (GDPR), with compliance monitored by the Office for Personal Data Protection of the Slovak Republic.

The Ministry of Education, Research, Development and Youth of the Slovak Republic has issued methodological guidelines and is developing specialised infrastructure for data management and archiving. A key instrument for managing data throughout the entire research lifecycle is the Data Management Plan (DMP).

Data are stored in scientific repositories within the SANET network, within the computational capacities of the Slovak Academy of Sciences, in departmental data centres, and in various repositories such as the international ZENODO repository, the Central Data Archive, the Slovakiana archive, the national repositories SCIDAP and SVD, as well as university repositories.

Secure data storage includes the use of appropriate technological solutions and adherence to ethical principles. For secure and long-term storage, digital repositories are used (including technological solutions such as DSpace, EPrints, Invenio, etc.) along with university archives. The Slovak Centre of Scientific and Technical Information (CVTI SR) and the Slovak Academy of Sciences contribute to the development of infrastructure for research data. For example, the SAS operates the Slovak Archive of Social Data (SASD), which serves for non-commercial use and archiving of data from international sociological surveys. Emphasis is placed on protecting identities and maintaining data confidentiality, particularly when working with sensitive personal data.

Biomedical data have their own particular characteristics. In the Slovak sectors of biomedicine, genomics and bioinformatics, data processing and storage focus on handling vast amounts of data (so-called BIG DATA) and on the sensitivity of the

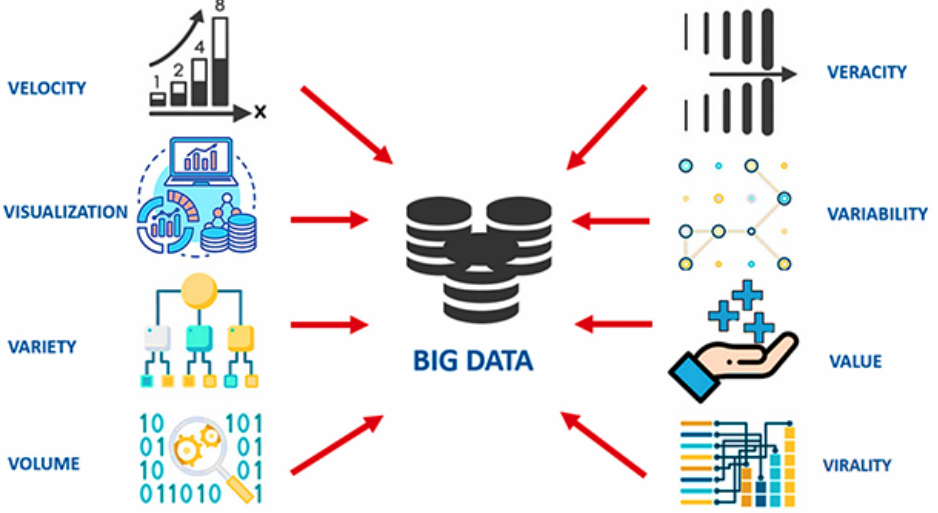
information. Strict ethical and legal standards, as well as specialised infrastructure, play a crucial role. The massive accumulation of extensive molecular and clinical datasets in recent decades has radically transformed personalised medicine and raised high expectations regarding its impact on biomedical research and healthcare. Extremely large collections of biomedical "omics" data (genomics, proteomics, transcriptomics, metabolomics, glycomics, etc.) and clinical data pose major challenges for analysis and interpretation and require the development of new robust algorithms capable of correctly analysing such high-volume datasets while accounting for individual genetic variability—an essential prerequisite for advancing personalised medicine (Hassan et al., 2022).

The processing of biomedical and genomic data demands substantial computational power and adherence to international standards to ensure accuracy and reproducibility. Bioinformatic analyses rely on advanced tools and software for processing sequencing data (e.g., DNA, RNA), genome mapping, variant identification and functional annotation. These processes are carried out on high-performance computing (HPC) infrastructures, where standardisation and data quality are crucial. Strict quality control (QC) and standardised data-processing protocols are implemented in accordance with the requirements of clinical research or diagnostics. From an ethical and legal perspective, working with genetic data requires particular attention to informed consent, as well as anonymisation and/or pseudonymisation to protect individual privacy. Legislation is stringent and follows EU directives.

The definition of "big data" encompasses several dimensions: volume, velocity, variety, value, variability, visualisation, virality and veracity (Figure 4) which describe the enormous scale of structured, semi-structured and unstructured datasets (Bibault, 2020; Normandeau, 2013; Gandomi & Haider, 2015; Diebold et al., 2012).

Given the sensitivity and scale of genomic data, secure storage and underlying infrastructure are critical. Data are stored in specialised, secure digital repositories and biobanks that meet strict security and long-term archiving requirements. In Slovakia, biobanks exist that store not only biological material (e.g., tissue and blood samples) but also associated clinical and genomic data. A crucial aspect is the computational infrastructure for genomic data, which requires a combination of high-performance hardware, specialised software and efficient storage and data-management solutions.

Figure 4. Representation of distinct dimensions of big data



(Hassan, M. et al, 2022)

Among the key components and requirements for processing these data are:

- High-Performance Computing (HPC): Genome analysis is extremely computationally demanding. It requires servers with a high number of CPU cores and often GPU acceleration to speed up tasks such as sequence mapping and variant calling.
- Sufficient Memory (RAM): Some bioinformatics tools—particularly those working with reference genomes—require large amounts of operational memory (often tens to hundreds of gigabytes of RAM).
- Scalable Storage: Genomic data (for example from next-generation sequencing – NGS) generate enormous volumes of data (terabytes to petabytes). The infrastructure must include high-speed, scalable storage systems (such as parallel file systems or cloud storage) with adequate backup and archiving capabilities.
- Network Connectivity: Fast data transfer between storage and compute nodes is critical for minimising processing times.
- Specialised Software and Tools: A range of bioinformatics platforms and analysis tools is used. Examples include platforms such as Galaxy (a web-based analysis platform requiring no command-line interaction) or specific libraries and packages in programming languages such as R or Python, as well as tools like GATK or Samtools.

- Cloud Computing: Many organisations use commercial cloud providers (such as AWS, Microsoft Azure or Google Cloud Platform) for their flexibility, scalability and accessibility of HPC-as-a-Service.
- Data Security and Management: The sensitivity of genomic data requires strict security measures, compliance with ethical and legal requirements, and adherence to established standards (for example within initiatives such as the European Genomic Data Infrastructure – GDI).

Data security is essential due to the sensitive nature of genomic data, which pose a significant risk of re-identification. Key functions and safeguards include:

- Pseudonymisation and Anonymisation: These are standard practices for protecting identity. Pseudonymised data (codes replacing direct identifiers) are still classified as personal data under the GDPR and require additional safeguards. Anonymised data, where re-identification is not possible, may be shared more freely.
- Access Control: Access is restricted to authorised and qualified users. Multi-factor authentication (MFA) and structured access levels are implemented (e.g., tiered access, in which data are processed in a secure environment without direct access to raw data).
- Encryption: Data are encrypted both at rest and in transit. Advanced cryptographic techniques are applied to ensure confidentiality.
- Secure Processing Environments: Analyses are often carried out in specially protected, trusted environments ("secure processing environments") that ensure that data do not leave the controlled infrastructure.
- Auditing and Monitoring: Access logs and data-processing records are regularly reviewed to prevent misuse and to monitor compliance.
- Controlled Access: The scientific community does not receive open access to raw sensitive data. Instead, access is granted in a controlled manner based on approved research projects and after signing agreements on confidentiality and responsible data use.
- International Protocols: Slovak researchers follow international data-sharing protocols and tools, including those developed by the Global Alliance for Genomics and Health (GA4GH).

Academic and research institutions in Slovakia, such as the Slovak Academy of Sciences and selected universities possess the necessary IT infrastructure for processing and storing big data. We can mention Pavol Jozef Šafárik University in Košice, Comenius University in Bratislava, the Comenius University Science Park,

Matej Bel University in Banská Bystrica, the Slovak University of Technology in Bratislava, and the Technical University of Košice.

They are able to ensure access for the scientific and research community through controlled mechanisms and international cooperation, in line with the FAIR principles.

The processing of biomedical and genetic data in Slovakia is subject to strict regulation and is fully harmonised with European Union law. Slovak legislation reflects all major EU directives and regulations in the areas of genetic technologies, biomedicine, personal data protection and the handling of human tissues and cells. The Office for Personal Data Protection of the Slovak Republic plays a central role in supervising compliance of processing activities with the GDPR and national legislation. Slovakia also transposes EU directives into its domestic legal framework and directly applies EU regulations with immediate effect.

1. General Data Protection Regulation (GDPR)

Regulation (EU) 2016/679 is the fundamental pillar of the protection of personal data, including sensitive health and genetic data. Slovak legislation, in particular Act No. 18/2018 Coll. on the Protection of Personal Data, fully implements the requirements of the GDPR. In biomedical research, this means obtaining unambiguous and specific consent from persons whose genetic data are processed with a clear determination of the purpose of use (informed consent), pseudonymisation, anonymization, technical and organizational measures to protect against unauthorized access or misuse.

2. Regulation on the European Health Data Space (EHDS)

Regulation (EU) 2025/327 (adopted in May 2024) creates a specific data environment for the health sector across the EU. It is intended to improve individuals' access to their data (primary use) and at the same time enable the secure cross-border use of anonymised data for research and innovation purposes (secondary use). The National Centre for Health Information (NCHI, 2025) is actively working on the implementation of the EHDS. The regulation will require a significant amendment of the Slovak healthcare system and the establishment of specialised access authorities for health data, thereby standardising and simplifying the sharing of biomedical data at the European level.

3. Legislation on Genetic Technologies and GMOs

Directive 2001/18/EC on the deliberate release into the environment of genetically modified organisms (GMOs) and other related EU regulations (e.g. Regulation

1829/2003 on genetically modified food and feed) regulate the work with GMOs in research and their placing on the market. These directives are transposed into Act No. 151/2002 Coll. on the use of genetic technologies and genetically modified organisms and related decrees (e.g. Decree of the Ministry of the Environment of the Slovak Republic No. 274/2019 Coll.). The Slovak Environmental Inspection (SIŽP) and the Ministry of the Environment of the Slovak Republic supervise compliance with these rules, which are crucial for laboratories working with genetic modifications.

4. Directives on Human Tissues and Cells (HTC)

Directive 2004/23/EC and subsequent technical guidelines set quality and safety standards for the donation, testing, processing, preservation, storage and distribution of human tissues and cells (HoC) intended for clinical use. These will be replaced by the new Regulation (EU) 2024/1938 from August 2027. Within the framework of Slovak legislation, Act No. 578/2004 Coll. on healthcare providers regulates the functioning of tissue facilities and biobanks in accordance with these requirements, thus ensuring safety and ethics when working with human biological material and associated data. Excellent research infrastructure within Slovakia, the biobank complex together with the built Martin Center for Biomedicine and Martin University Hospital represent the potential for state-of-the-art research and clinical testing. Long-term and systematic collection of samples, storage and provision of biological material for scientific research with respect for all the rights of donors contributes to improving the quality of life of the next generation and to the development of personalized medicine. The aim is that the benefit of all activities is primarily for the patient in terms of effective, modern and targeted therapy. (JLF UK, 2023)

Overall, the Slovak legal framework reflects the effort to comply with European standards, balancing the need for innovation and data sharing in research with strict privacy and ethical principles. The security of large-scale genomic data repositories in Slovakia and their availability to the scientific research community are regulated by a strict framework. It combines national legislation with binding European regulations and initiatives such as the GDPR, the EHDS and the 1+MG initiative. The EU flagship initiative "1+ Million Genomes" (1+MG) aims to enable secure access to genomics and corresponding clinical data across Europe in order to support breakthrough research and health policy-making and stimulate personalized healthcare with the potential to improve disease prevention and balance privacy protection with the need for data sharing for the advancement of research and clinical practice. (European Commission, 2025)

Overall, the system is designed to enable scientific progress through data sharing, but only within a strictly controlled and secure environment that protects the rights and privacy of Slovak and EU citizens.

Slovakia is involved in ESFRI infrastructures as part of international cooperation. The European Strategy Forum on Research Infrastructures (ESFRI) aims to build European research and development and support a common approach to science policy-making in research infrastructures in Europe. The ESFRI Forum has prepared a European indicative plan for research infrastructures of pan-European interest, which create a strategic tool for the development of scientific integration in Europe and strengthen their international reach. The last update in 2012 identified 48 research infrastructures (European Strategy Forum on Research Infrastructures, 2015).

The Slovak Republic is involved in the position of a member or a guest member in various ESFRI research infrastructures.

For biomedical research in Slovakia, it is important that the Slovak Republic has become a full member of the pan-European research infrastructure ELIXIR. This step followed Slovakia's previous activities as an observer or prospective member in this and other ESFRI infrastructures (e.g. BBMRI ERIC, EATRIS ERIC). ELIXIR (European Life Science Infrastructure for Biological Information) is a distributed international research infrastructure that brings together resources on life sciences. From all over Europe. These resources include databases, software tools, training materials, cloud storage and supercomputers and expertise. By joining the international infrastructure, Slovakia has committed to building and operating its own national node (ELIXIR SK), which is compatible with the central requirements of ELIXIR and provides services to the Slovak and international research community. An extremely valuable benefit of full membership is to enable Slovak scientists to fully use the resources and services, actively participate in building and managing the infrastructure and collaborate on international projects in the field of biomedicine, bioinformatics and genomics (SK4ERA, 2025).

ELIXIR coordinates these resources to form a unified infrastructure. This infrastructure makes it easier for scientists to find and share data, exchange expertise and agree on best practices. Ultimately, this helps them gain new knowledge about the functioning of living organisms.

CVTI SR, academic institutions and SAS play a key role in the development of the national node and the promotion of the ELIXIR infrastructure in Slovakia. Slovakia's entry into ELIXIR is an important step in strengthening the country's position in the European Research Area and supporting open science and data management in biomedical fields. In Slovakia, the computational infrastructure for biomedical data, including genomic data, is being developed mainly within academic and research institutions and is often linked to funding from EU funds.

For Slovakia, membership in the research infrastructure of EMBL (European Molecular Biology Laboratory) is a significant contribution to research in the life sciences. Slovakia became a full member of EMBL in 2018 after a period of prospective membership that began in 2014. Since then, the Slovak scientific community

has been actively involved in EMBL within its missions, engaging in excellent collaborative research at individual EMBL sites, educating a new generation of Slovak scientists, using new scientific services and technologies, and striving for innovation and research valorisation. We are following a positive trend in the involvement of Slovak scientists in EMBL from the period 2021-2024.

The cooperation between the Slovak scientific community and EMBL continues to grow in the context of the current EMBL program Molecules for Ecosystems. Within this five-year program, EMBL is expanding existing opportunities for its member states to engage, from joint research projects and training modalities to new cutting-edge services. The increase in participation of Slovak scientists observed since the launch of the program proves that these opportunities are actively used by Slovak researchers and that EMBL activities within its missions are of great importance for Slovakia.

Of importance for Slovakia is the cooperation with the excellent bioinformatics institute EMBL-EBI in Hinxton. EMBL-EBI develops and maintains the world's most comprehensive collection of freely available and up-to-date molecular data resources. Created in collaboration with scientists globally, this includes over 40 open databases and hundreds of bioinformatics tools and software, all accessible at EMBL-EBI services to anyone worldwide.

Open life science databases provided by EMBL-EBI form the foundation for much of modern life science research. They save researchers thousands of hours and enable the pursuit of research questions that would be impossible without such comprehensive resources. In 2021, an independent assessment highlighted their critical role in research and development, with 70 percent of participants indicating that not being able to access EMBL-EBI data resources would have a major or severe impact on their work. To capture this value, EMBL has been collecting country-level bibliometric data on publications enabled by EMBL-EBI data resources, showcasing their significance to the scientific community.

In 2024, EMBL-EBI received over 10 million web access requests2 from Slovakia, showcasing that EMBL-EBI data resources are indispensable for Slovak life scientists, both in the academic and industrial sectors. The requests originated from 39,812 unique Slovak IP hosts, each representing a distinct device, such as a computer or a server. Additionally, in 2024, 30 publications3 by authors affiliated with Slovak institutions were enabled by EMBL-EBI data resources. The most frequently cited resources were the European Nucleotide Archive, the Protein Data Bank and UniProt.

The software tools used in Slovakia are typical for the international bioinformatics community, specific commercial platforms vary by workplace.

Open-source tools such as GATK (Genome Analysis Toolkit), samtools, BWA, and various libraries and packages in the R and Python programming languages

predominate. These are used to process and analyse sequencing data obtained by NGS sequencing. (NGS - next generation sequencing).

Some workplaces use web platforms such as Galaxy for simplified analysis without the need for advanced command line skills. To a lesser extent, specialized commercial software solutions for medical data management and analysis are also used.

Slovakia has specific hardware, which is concentrated in larger research centres and universities where HPC centres have been established. Research institutions such as the Slovak Academy of Sciences and technical universities (e.g. TUKE, STU, UMB) and also CVTI SR operate high-performance computing (HPC) clusters for processing computational tasks. These include servers with a high number of CPU cores, large operating memory and high-speed storage. Local servers and storage serve smaller laboratories and healthcare facilities for data collection and temporary processing. The Slovak academic and scientific network SANET provides high-speed connectivity between institutions, which is crucial for the transfer of large volumes of biomedical data.

In the biomedical sector, it is essential to support and develop cloud solutions. Within commercial clouds, global providers such as: "Amazon Web Services" (AWS), "Microsoft Azure", and "Google Cloud Platform" are used for their scalability and flexibility. Currently, there is an effort in Slovakia to develop a state academic cloud infrastructure and provide data services in accordance with EU regulations (e.g. GDPR) for the secure storage and sharing of sensitive medical data.

The projects, which are co-funded by the EU, are aimed at creating digital bio-banks that use cloud technologies to support systemic public research infrastructure and testing of new treatment strategies.

Slovakia is involved in European initiatives such as the "1+ Million Genomes" (1+MG) Initiative and the European Health Data Space (EHDS). These initiatives aim to create a federated and secure infrastructure that will enable cross-border access to popular genomic data in the EU for research purposes, while maintaining high standards of security and privacy.

A prerequisite for international cooperation is a high-quality national infrastructure. Slovak institutions (CVTI SR, SAS) are working on building national nodes (e.g. EOSC-SK), which will connect to the European Open Science Cloud Infrastructure (EOSC) and enable data sharing in accordance with EU standards.

Integration of Data and Systems as a Foundation for High-Quality and Trustworthy Information

In contemporary science and public administration, there is an increasing emphasis on the need for interconnected, interoperable and trustworthy data ecosystems.

Data integration, the coordinated linkage of data from different sources, formats and domains, constitutes a fundamental prerequisite for ensuring the quality, consistency and reliability of information (Turning FAIR into reality, 2018).

Data integration is crucial from several perspectives. In terms of data quality, integration makes it possible to detect errors, duplications and incomplete records across systems. High-quality data must be not only accurate but also consistent and up-to-date. Transparent interconnection of sources simultaneously enhances their traceability, which is important for building trust in their origin and correctness. Data integration also increases the efficiency of research and research management: integrated data systems enable automated processing, analytics and the re-use of data, thereby reducing administrative and scientific redundancy (Wilkinson et al., 2016).

In Slovakia, we have undertaken the implementation of Open Science and the FAIR principles for research data collected not only within SK CRIS, across all systems under our management. We focus on the implementation of unique identifiers, data management, metadata quality control, and the integration of various data sources into broader units with the character of research infrastructures. The aim is to ensure that collected data are used not only at the national level, but also it be made available for international comparison.

To support these integration efforts, the Slovak Centre of Scientific and Technical Information (CVTI SR) has developed the Comprehensive Information System for the Acquisition, Processing, Preservation and Dissemination of Scientific and Bibliometric Data and Publications, and for Providing Access to Tools and Applications Supporting Research and Development (abbreviated as KOMIS).

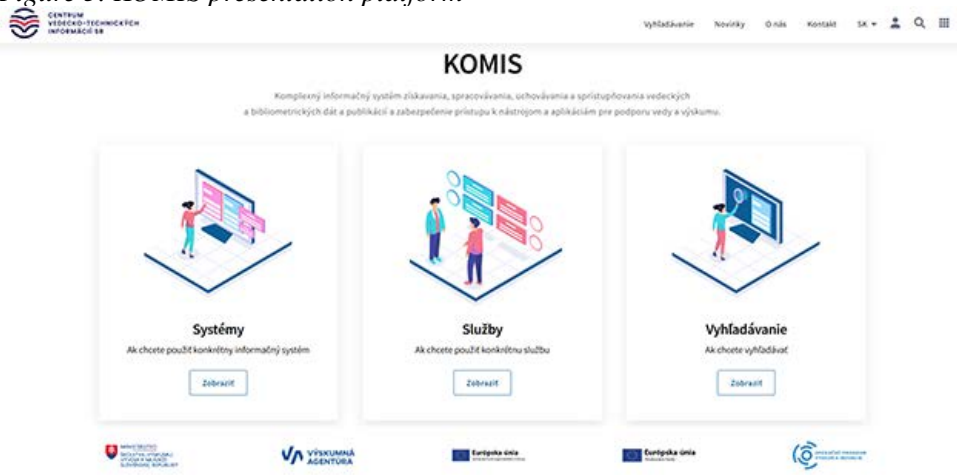
The resulting national research infrastructure aggregates systems supporting research operated by CVTI SR and provides access to services of the Data Centre for Research and Development, such as storage capacity, data processing, and the use of various installed software applications (Software as a Service – SaaS), mainly in the fields of mathematics, statistics and bioinformatics. The solution also includes analytical tools for data processing and a module for searching available data and information resources.

The entry point to the system is a presentation platform (Figure 5) with an implemented Single Sign-On (SSO) authentication process, which allows users to share their credentials easily across various systems and applications. For identity and access management, the open-source tool Keycloak has been deployed. The solution is linked to a local Active Directory identity management system and also enables authentication via an ORCID identifier.

Multiple metadata standards are used within KOMIS. In addition to the CERIF format implemented in SK CRIS, these include mainly DataCite, Dublin Core and MARC21, all processed in JSON. These formats are also supported for data export, particularly from repositories.

KOMIS has defined open API calls using REST-API protocols. These API interfaces are used for communication with external systems as well as for internal communication between the individual components of KOMIS. KOMIS is currently in pilot operation, and its beta version does not yet support full production deployment of all functionalities.

Figure 5. KOMIS presentation platform



As part of strengthening integration objectives, several related functionalities have been implemented in the Slovak Research Information System (SK CRIS), specifically:

- Implementation of researcher ID federation. The system now supports internationally used persistent identifiers (PIDs), specifically ORCID, ResearcherID and Scopus ID, as well as the national identifier ID of a Higher Education Teacher from the register of university employees.
- Integration with the Register of Legal Entities (using a national-level PID for organisations, namely the Identification Number of an Organisation – IČO). The Register of Legal Entities is a reference register. Integration with this register enables SK CRIS to retrieve data on organisations by validating their IČO. This applies the "once only" principle, ensuring that users do not need to re-enter information already available in the national register.
- Extension of recorded research outputs to include DOIs and URLs of the full text of publications.

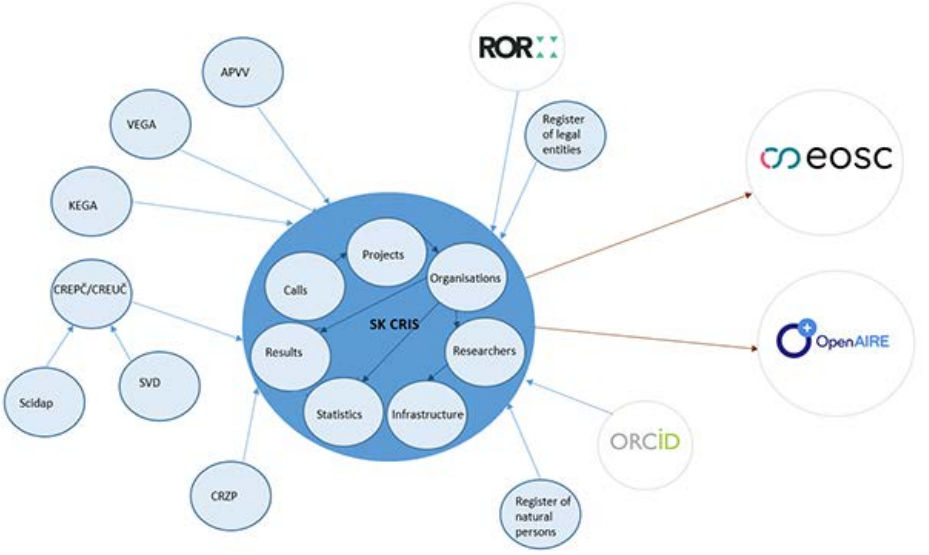
It should be emphasised that researcher identification is already technically implemented in several other research-support systems. However, the provision of an identifier is not yet mandatory, as not all researchers currently possess one of the commonly used persistent identifiers. Full implementation of unique personal identification in practice is contingent upon legislative anchoring.

For research organisations, unique identification has already been put into practice. Slovakia operates a legally mandated Register of Legal Entities, which ensures authoritative identification. Unique identification of Slovak organisations is provided through the national identifier IČO. In addition, Slovak organisations are generally also included in the international Research Organisation Registry (ROR).

Among the controlled vocabularies or thesauri used for the substantive categorisation of recorded objects, the most important are those derived from Slovak legislation: the Classification of Science and Technology Fields and the Classification of Study Fields.

The system integration scheme for research-support systems is presented in Figure 6.

Figure 6. Integration of systems for supporting science in Slovakia



Within international initiatives aimed at the integration of science and research data, Slovakia participates not only in the activities of the international organisation for research information euroCRIS, but primarily in the OpenAIRE project and in the EOSC federation.

Data and systems integration is not merely a technical task, but a strategic pillar of open, trustworthy and reproducible science.

It ensures that data are:

- High-quality – accurate, complete and consistent,
- Trustworthy – with clearly defined provenance and governance,
- Reusable – thanks to shared standards, metadata and interoperable infrastructures.

In a well-integrated data environment, quality and trust do not become additional features, but a natural result of the design of the data ecosystem.

CONCLUSION

Statistical indicators comparing the position of EU countries in science and research place Slovakia only in the lower third.

Slovakia has long exhibited very low expenditure on science and research—both as a share of GDP and in terms of private-sector investment. Other indicators also fall below the EU average, for example the PISA tests for pupils and students, and Slovakia similarly performs worse in innovation outcomes and in linking research with industry.

On the other hand, a new national strategy (VAIA, 2023) has been adopted, defining three fundamental transformations of the research and development ecosystem.

The first concerns investment in a high-quality system. Alongside the necessary increase in funding for research, development and innovation, the mechanisms for allocating and drawing these resources will be reformed.

The second area focuses on talent. The research and innovation ecosystem can flourish only if it is staffed with high-quality, strategically important talent from Slovakia and abroad. The intention is to support the development of domestic talent, attract and retain existing and foreign talent, and purposefully enhance the quality of life and work in Slovak cities where the vast majority of research and innovation resources are concentrated.

The third area introduces a comprehensive framework for setting priorities in the support of research, development and innovation in Slovakia.

The planned reforms also address research information and data. The objective is to strengthen support for open data and open science, and to make data from public administration and research more widely accessible in order to stimulate innovation. To realise the planned measures in the field of open data, it is necessary to translate strategic intentions into concrete actions:

- Legislative support, including the introduction of motivational and/or restrictive measures for entities obliged to provide data.
- Optimisation of data and information collection processes, in line with the *"once only"* principle, with the aim of reducing excessive administrative burden.
- Support for the creation of high-quality metadata, including metadata quality control and the implementation of data curation and data management tasks.
- Implementation of persistent identifiers, international standards and integration tools within modern software solutions developed in accordance with contemporary trends in digitalisation and the use of artificial intelligence tools in science, research and public administration.

The implementation of measures in research policy and in the area of open research information and data will be accompanied by increasing trust among the actors of the research and innovation ecosystem towards the state and, conversely, the trust of the state towards researchers and innovators. A trust-based, transparent, collaborative and stimulating research ecosystem, firmly embedded in the international context, is a prerequisite for successfully addressing the economic and technological challenges of the present.

Funding

This publication was written thanks to the generous support for the project Integration of Genomic Data Resources for National Security in the Human, Food, and Environmental Sectors, Project No. 401101C603

REFERENCES

Altman, M., & Cohen, P. N. (2022). The Scholarly Knowledge Ecosystem: Challenges and Opportunities for the Field of Information. *Frontiers in Research Metrics and Analytics*, *6*, 751553. Advance online publication. DOI: 10.3389/frma.2021.751553 PMID: 35178498

Azeroual, O., & Schöpfel, J., Pölönen, & J. Nikiforova, A. (2022). Putting FAIR Principles in the Context of Research Information: FAIRness for CRIS and CRIS for FAIRness. In *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2022).* SciTePress, 63-71. DOI: 10.5220/0011548700003335

Azeroual, O., Schöpfel, J., Störl, U., & Marušić, A. (2024). Ethical aspects using AI in CRIS. *Procedia Computer Science*, *249*, 150–159. DOI: 10.1016/j.procs.2024.11.058

Bibault, J.-E. (2020). Real-life clinical data mining: Generating hypotheses for evidence-based medicine. *Annals of Translational Medicine*, *8*(4), 69. Advance online publication. DOI: 10.21037/atm.2019.10.99 PMID: 32175362

Bordons, M., González-Albo, B., & Moreno-Solano, L. (2023). Improving our understanding of open access: How it relates to funding, internationality of research and scientific leadership. *Scientometrics*, *128*(8), 4651–4676. DOI: 10.1007/s11192-023-04726-1

Burgelman, J. C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. (2019). Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Frontiers in Big Data*, *2*, 43. Advance online publication. DOI: 10.3389/fdata.2019.00043 PMID: 33693366

Castro, P. (2024). Improving CRIS features to support new Open Access implementation workflows at institutions. *Procedia Computer Science*, *249*, 179–185. DOI: 10.1016/j.procs.2024.11.062

Diebold, F. X., Cheng, X., Diebold, S., Foster, D., Halperin, M., Lohr, S., Mashey, J., Nickolas, T., Pai, M., & Pospiech, M. A. (2012). *Personal Perspective on the Origin (s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline*. CiteSeer.

Dvořák, J. (2025). *CERIF Tutorial.* euroCRIS Membership Meeting Leuven, Belgium. http://hdl.handle.net/11366/2727

ERAPortál. (2025). *Rodová rovnosť vo výskume. Politický rámec a základné definície.* [Gender equality in research. Policy framework and basic definitions.] https://eraportal.sk/eraportal/rodova-rovnost-vo-vyskume/politicky-ramec-a-zakladne-definicie/

euroCRIS. (2025) *Directory of Research Information Systems – DRIS.* https://dspacecris.eurocris.org/browse?type=driscountry&order=ASC&rpp=20&authority=824

European Commission. (2020). *Únia rovnosti: stratégia pre rodovú rovnosť na roky 2020 – 2025.* [A Union of Equality: Gender Equality Strategy 2020-2025.] https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-equality-strategy_sk

European Commission. (2025). *European '1+ Million Genomes' initiative*. https://digital-strategy.ec.europa.eu/en/policies/1-million-genomes

European Strategy Forum on Research Infrastructures. (2015). *Európske strategické fórum o výskumných infraštruktúrach (ESFRI roadmap) v interakcii s pripravovanou cestovnou mapou infraštruktúr výskumu a vývoja (SVK roadmap).* [The European Strategy Forum on Research Infrastructures (ESFRI roadmap) in interaction with the upcoming roadmap for research and development infrastructures (SVK roadmap).] https://www.minedu.sk/data/att/8d3/8662.3ee4cb.pdf

Farnel, S., & Shiri, A. (2014). Metadata for Research Data: Current Practices and Trends. *International Conference on Dublin Core and Metadata Applications,* 2014. DOI: 10.23106/dcmi.952136534

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. DOI: 10.1016/j.ijinfomgt.2014.10.007

Hassan, M., Awan, F. M., Naz, A., deAndrés-Galiana, E. J., Alvarez, O., Cernea, A., Fernández-Brillet, L., Fernández-Martínez, J. L., & Kloczkowski, A. (2022). Innovations in Genomics and Big Data Analytics for Personalized Medicine and Health Care: A Review. *International Journal of Molecular Sciences*, *23*(9), 4645. Advance online publication. DOI: 10.3390/ijms23094645 PMID: 35563034

Hauschke, C., Nazarovets, S., Altemeier, F., & Kaliuzhna, N. (2021). Roadmap to FAIR Research Information in Open Infrastructures. *Journal of Library Metadata*, *21*(1–2), 45–61. DOI: 10.1080/19386389.2021.1999156

JLF UK. (2023). *Systémová verejná výskumná infraštruktúra – biobanka pre nádorové a zriedkavé ochorenia.* [Systemic public research infrastructure – biobank for cancer and rare diseases.] https://www.jfmed.uniba.sk/uploads/media/Systemova_verejna_vyskumna_infrastruktura_biobanka_pre_nadorove_a_zriedkave_ochorenia_II_faza.pdf

Leipzig, J., Nüst, D., Hoyt, C. T., Ram, K., & Greenberg, J. (2021). The role of metadata in reproducible computational research. *Patterns (New York, N.Y.)*, *2*(9), 100322. Advance online publication. DOI: 10.1016/j.patter.2021.100322 PMID: 34553169

Mayernik, M. S. (2019). Metadata accounts: Achieving data and evidence in scientific research. *Social Studies of Science*, *49*(5), 732–757. DOI: 10.1177/0306312719863494 PMID: 31354073

Ministry of Education. Science, Research and Sports of SR (2021a). *Cestovná mapa výskumných infraštruktúr - SK VI Roadmap 2020 – 2030.* [Roadmap of research infrastructures - SK RI Roadmap 2020 - 2030.] https://www.minedu.sk/data/att/c28/25884.f33c49.pdf

Ministry of Education. Science, Research and Sports of SR (2021b). *Národná stratégia pre otvorenú vedu na roky 2021-2028.* [National Strategy for Open Science for 2021-2028.] https://www.minedu.sk/data/files/10833_narodna-strategia-pre-otvorenu-vedu-na-roky-2021--2028.pdf

Ministry of Education. Science, Research and Sports of SR (2022). *Metodické odporúčanie na tvorbu a implementáciu inštitucionálnej politiky otvorenej vedy pre inštitúcie vykonávajúce vedu a výskum.* [Methodological recommendation for the creation and implementation of an institutional open science policy for institutions conducting science and research.] https://www.minedu.sk/data/att/62e/33212.73cb3f.pdf

Ministry of Education. Research, Development and Youth of SR (2024a). *Správa o stave výskumu a vývoja v Slovenskej republike a jeho porovnanie so zahraničím za rok 2023.* [Report on the state of research and development in the Slovak Republic and its comparison with abroad for 2023.] https://www.minedu.sk/data/att/108/33300.81834d.pdf

Ministry of Education. Research, Development and Youth of SR (2024 b). *Kódex výskumnej integrity a etiky na Slovensku.* [Code of Research Integrity and Ethics in Slovakia.] https://www.minedu.sk/data/att/155/32501.fe9bf4.pdf

Ministry of Labor. Social Affairs and Family of SR. (2021). *Celoštátna stratégia rovnosti žien a mužov a rovnosti príležitostí v Slovenskej republike na roky 2021 – 2027.* [National Strategy for Equality between Women and Men and Equal Opportunities in the Slovak Republic for 2021-2027.] https://www.employment.gov.sk/files/sk/ministerstvo/spolocny-sekretariat-vyborov/vybor-rodovu-rovnost/dokumenty-udalosti/strategia-rovnosti-z-m-2021-27.pdf

Mons, B. (2019). FAIR Science for Social Machines: Let's Share Metadata Knowlets in the Internet of FAIR Data and Services. *Data Intelligence*, *1*(1), 21–42. DOI: 10.1162/dint_a_00002

NCZI. (2025). *Slovensko napreduje v implementácii EHDS*. [Slovakia is making progress in implementing the EHDS] https://www.nczisk.sk/Aktuality/Pages/Slovensko-napreduje-v-implementacii-EHDS.aspx

Normandeau, K. (2013). *Beyond Volume, Variety and Velocity is the Issue of Big Data Veracity. Inside Big Data.* https://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity

Numajiri, H., & Hayashi, T. (2024). Analysis on open data as a foundation for data-driven research. *Scientometrics*, *129*(10), 6315–6332. DOI: 10.1007/s11192-024-04956-x

Rumble, J.Jr, Broome, J., & Hodson, S. (2019). Building an International Consensus on Multi-Disciplinary Metadata Standards: A CODATA Case History in Nanotechnology. *Data Science Journal*, *18*(1), 12. Advance online publication. DOI: 10.5334/dsj-2019-012

SK4ERA (2025). *Elixir.* https://eraportal.sk/eraportal/oblasti-nasej-podpory/pristup-k-vyskumnym-infrastrukturam/elixir/

Schöpfel, J., & Azeroual, O. (2023). Ethical Issues of the Organization and Management of Research Information. *Communication, technologies et développement, 14*. DOI: 10.4000/ctd.9857

Schöpfel, J., Azeroual, O., & Castro, P. (2022). Research Information Systems and Ethics relating to Open Science. *Procedia Computer Science*, *211*, 36–46. DOI: 10.1016/j.procs.2022.10.174

Thibault, R. T., Amaral, O. B., Argolo, F., Bandrowski, A. E., Davidson, A. R., & Drude, N. I. (2023). Open Science 2.0: Towards a truly collaborative research ecosystem. *PLoS Biology*, *21*(10), e3002362. Advance online publication. DOI: 10.1371/journal.pbio.3002362 PMID: 37856538

Ukropová, D. (2002) Metadáta v slovenských knižniciach. [Metadata in Slovak Libraries] *IT Lib. Informačné technológie a knižnice, 6* (4). https://itlib.cvtisr.sk/clanky/clanek2107/

UNESCO (2021). *UNESCO Recommendation on Open Science*. DOI: 10.54677/MNMH8546

VAIA. (2023). *Národná stratégia výskumu, vývoja a inovácií.* [National Strategy for Research, Development and Innovation.] https://vaia.gov.sk/wp-content/uploads/2023/05/01_Narodna-strategia-vyskumu-vyvoja-a-inovacii_vlastny-material_V2.pdf

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., Bonino da Silva Santos, L., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. Advance online publication. DOI: 10.1038/sdata.2016.18 PMID: 26978244

Yang, W., Fu, R., Amin, M. B., & Kang, B. (2025). The Impact of Modern AI in Metadata Management. *Human-Centric Intelligent Systems*, *5*(3), 323–350. DOI: 10.1007/s44230-025-00106-5

Zendulková, D., & Azeroual, O. (2022). Legal aspects and data protection in relation to the CRIS system. *Procedia Computer Science*, *211*, 17–27. DOI: 10.1016/j.procs.2022.10.172

Zendulková, D., Gavurníková, G., Krivjanska, A., Staňáková, Z., Putalová, A., & Janková, M. (2022). Representation of Women in Slovak Science and Research: An Analysis Based on the CRIS System Data. *Information (Basel)*, *13*(10), 482. DOI: 10.3390/info13100482