sciendo

ICTE Journal

2018, 7(2): 39–45
DOI: 10.2478/ijicte-2018-0008

IS OPEN ACCESS TO RESEARCH DATA A STRATEGIC PRIORITY OF CZECH UNIVERSITIES?

Jakub Novotný[1]

[1]Department of Systems Analysis, Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic {novotnyj@vse.cz}

ABSTRACT

Open access to research data is one of the key themes of current science development concepts and relevant R & D strategies at least in Europe. A systemic change in the modus operandi of science and research should lead to so-called Open Science. The presented paper questions the extent to which the Open Science concept is reflected in the strategies of Czech universities. The paper first describes basic idea of Open Access to Research Data including principles of „FAIR data" as one of the key assumption of it. After a brief characterization of the Czech university sector, the results of the empirical analysis of the inclusion of the Open Access to Research Data concept in the current strategic plans of the Czech universities are presented. The conclusion of the paper is then an evaluation of the results, which reveal an underestimation of the Open Science concept in the current strategic plans of the Czech universities.

KEYWORDS

Research data, open access, Czech universities, open science, FAIR data.

1 INTRODUCTION

The rapid development of information and communication technologies in last few decades affected all economic and social sectors and nearly all human activities. Academic sector is not an exclusion in this context. Science and research activities rely on many ways on ICT and in same time research generate lot of digital content. This digitalization brings new challenges and one of nowadays trend is open access and so called Open Science. So academic institutions looking to the future reflect this trend in their strategic plans and implement it to their publication and data operational agenda.

This paper focus on Czech university sector in this context with the question on how Czech universities reflect this trends in their strategic plans. There will be summed up basic concepts of Open Science and Open Data at first and then there will be provided characterization of Czech university sector. The core contribution of the paper is to provide results of empirical analysis of strategic plans of Czech universities resulting in a discussion of nowadays state and future outlook.

2 OPEN SCIENCE AND OPEN DATA

Science and its individual disciplines are evolving. There is also a change in the way of its operation and the source and mechanism of science funding. The number of researchers and the number of scientific journals and other publishing platforms are increasing. There is also a growing number of research institutions and university students. Technological development (which in itself is the result of scientific knowledge) offers new tools for research and dissemination and publication of results. Digital technologies offer a faster and cheaper way of presenting results than before. The society's relationship to science and

the development of science is also changing. Science and its strategic development have become part of wider political concepts.

In the last decade we can therefore meet the concept of Open Science (or earlier Science 2.0) with three main attributes (EC 2014):

1. A significant increase of scientific production, open research and remote collaboration and online (open) access to scientific information.
2. An emergence of data-intensive science by availability of large-scale datasets (petabytes) and by high performance computing.
3. An increase in the number of actors in science.

Open Science is therefore a systematic change in the modus of operandi of research activities and is affecting the research cycle and all of its stakeholders. Research process in open form is shown in the following figure.

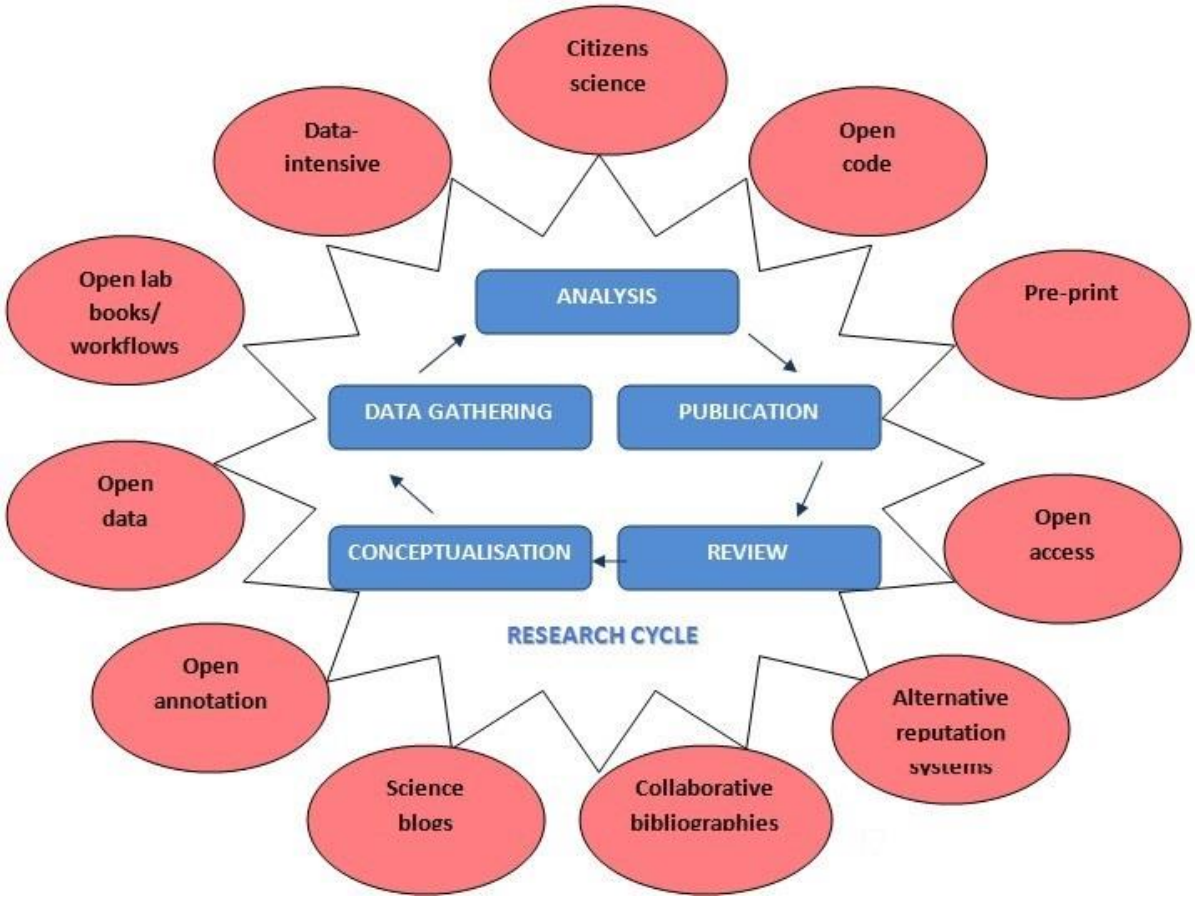

Figure 1 Open Science trends – Source (EC 2014)

Although there are many Open Science trends, open access is the most discussed and elaborate one, at least in terms of science policy within the European Union. According to (EC 2017) "*open access refers to the practice of providing online access to scientific information that is free of charge to the end-user and reusable*" and scientific information are divided to categories: peer-reviewed scientific research articles and research data. It is very important how access is defined. Access is not only "*the right to read, download and print – but also the right to copy, distribute, search, link, crawl and mine.*"

Research data in this context are data (statistics, results of experiments, measurements, observations, survey results, interview recordings and images) in digital form allowing users to freely access, mine, exploit,

reproduce and disseminate them. Open access to scientific publications and research data will according to European science policy (especially through Horizon 2020) improved quality of results, encourage collaboration, avoid duplication of effort, speed up innovation and involve citizens and society to science.

3 FAIR DATA

A much more detailed specification of research data features within the Open Data concept is the so-called FAIR Data principle. The basic document dealing with FAIR Data is the Guidelines on FAIR Data Management, which specifically addresses the recommendations for the Horizon 2020 R & D beneficiary or the participants involved in the Open Research Data Pilot, but its impact on the scientific community is wider and touches the issue of openly accessible scientific data in general. The guide does not detail the principle of FAIR data. It contains only an initial indication that it helps the beneficiaries to make their research data findable, accessible, interoperable and reusable (FAIR) and also states in the annex that the research data should comply with the FAIR principles, and refers to FORCE11 and a published article in Nature (Wilkinson 2016) for further details.

So let's look at the FAIR data concept. The principles are not only related to the data itself (in a strict definition), but also to the research procedures, algorithms and tools that lead to the production of such data. In the basic breakdown, there are 15 principles or recommendations that research data should meet:

To be Findable:

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata (defined by R1 below)

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles

I3. (meta)data include qualified references to other (meta)data

To be Reusable:

R1. meta(data) are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with detailed provenance

R1.3. (meta)data meet domain-relevant community standards

These policies should not serve as a standard or specification, nor does it address the technical implementation of the data produced and stored. Their intention is to assist scientific institutions and research teams in deciding on a specific way of realizing the digital outputs of their research so that these outputs can be searched, accessed, involved in further research, and further exploited within the scientific

community (and not only). It is in fact an explanation of the scientific-methodological requirements for digital outputs of scientific work and the scientific and methodological assumption of machine evaluation and mining of research data.

3 HIGHER EDUCATION SECTOR IN THE CZECH REPUBLIC

There are 64 higher education institutions (universities and colleges) in the Czech Republic. The data are for August 2018 and concern institutions based in the Czech Republic, not the branches of foreign universities (these are 16). According to Act no. 111/1998 coll. (amended and consolidated) on Higher Education Institutions and on Amendments and Supplements to Some Other Acts (the Higher Education Act) higher education institutions may be public, private or state institutions. And the act distinguishes also two types of higher education institutions: university or a non-university type. Higher education institutions of the non-university type provide Bachelor's degree programs and may also carry out Master's degree programs as well as creative activities. Universities than may provide all types of degree programs as well as related scholarly, scientific, research, development, innovation, artistic and other creative activities.

Of the number of 64 higher education institutions, there are 35 non-university and 29 university type. Most universities are public, 2 are state and 3 private. In contrast, most non-university higher education institutions are private and only two are public. The breakdown by type and arrangement is clearly shown in the figure below.

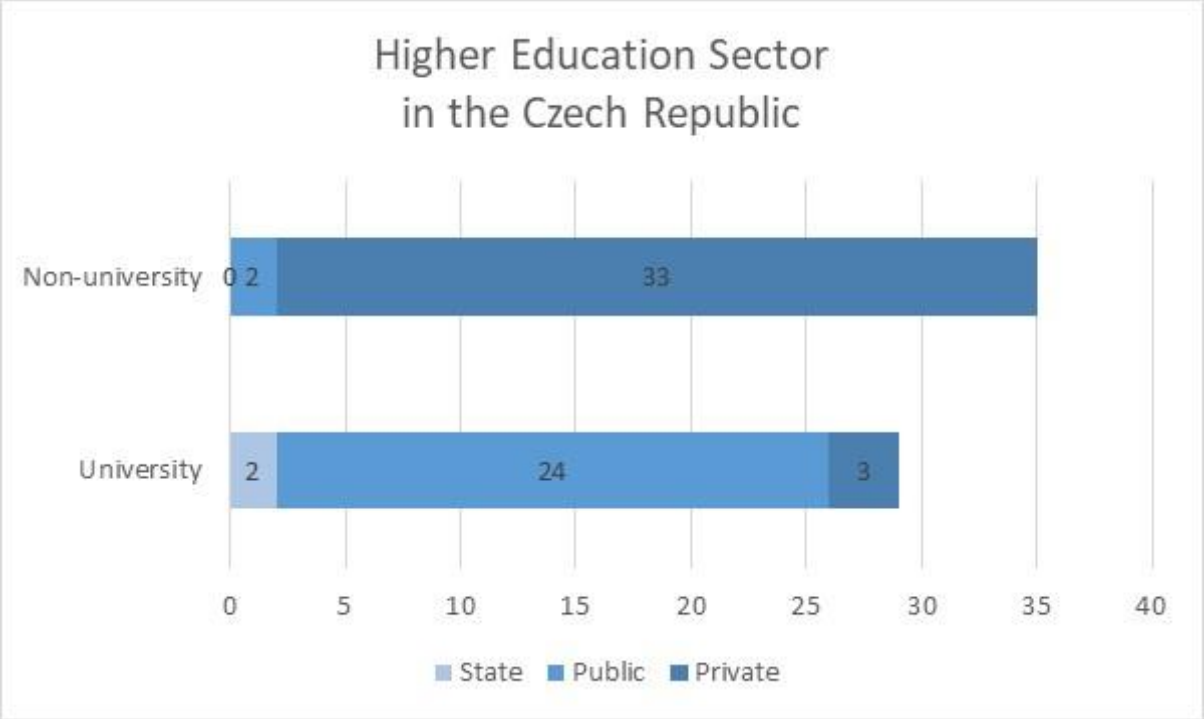

Figure 2 Higher education sector in the Czech Republic (data source: Register of Higher Education Institutions, Czech Republic)

Given that the non-university sector realizes mainly professional bachelor study programs, we will deal exclusively with the university sector in the following text. There are in our focus (open) research data and, from the point of view of higher education, universities are by law the main carriers of research and development and exclusive providers of Ph.D. study programs.

According to the Higher Education Act each higher institution prepares, discusses in its bodies and with the Ministry of Education, Youth and Sports and publicly publishes the strategic plan for the educational and creative activities of the higher education institution and the annual implementation plan of this strategic plan. The higher level of strategy document is general strategic plan for the educational and creative activities of the Ministry of Education, Youth and Sports. These strategic plans (ministries and

individual institutions) are thus key documents for the development and direction of jungle higher education institutions in the Czech Republic.

4 OPEN RESEARCH DATA AS A TOPIC OF STRATEGIC PLANS OF CZECH UNIVERSITIES

In the first part of this article, we characterized Open Science and Open Research Data. Although the cited documents are mainly related to the program Horizon 2020, their impacts are wider and clearly represent the trend in science and research operation. Before proceeding with further step, it is appropriate to make a turning point to existing analyze of the state of the art of implementation of Open Science concept in the Czech Republic.

The analytical document *Access to and Preservation of Scientific Information in Europe* (EC 2012) states that there are "*very little or no open access to research data policies in place, but some plans in place or under development*." Coordination in this area in the Czech Republic is then left to the university libraries, or their associations, according to the cited document. Association of Libraries of Czech Universities (ALCU) signed in year 2012 Berlin declaration and publish *ALCU Open Access Policy.* Subsequently, on June 14, 2017, the Czech National Strategy for Open Access to Research Information for 2017-2020 has been approved by the Government of the Czech Republic. So the Open Science principles are therefore gradually implemented at a general political level. The question therefore is, how this issue reflect Czech universities in their strategic plans.

As stated above, the strategic plan is the main and basic strategic document of Czech universities. For this purpose, the author of this paper analyzed the strategic plans of all Czech universities. Specifically, these were strategic plans for years 2016 – 2020 obligatory published by each institution. The key searched term was Open Research Data and Open Science in the individual priorities of each institution's strategic plan. The results are illustrated in the following graph.

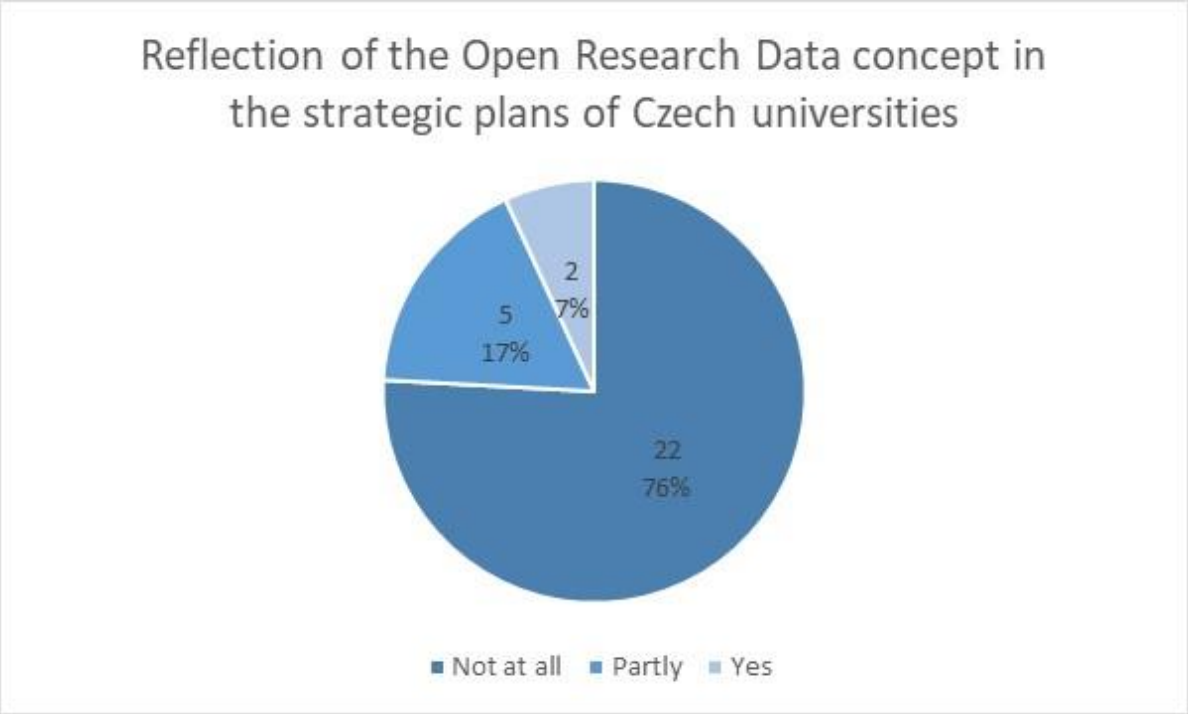

Figure 3 Reflection of the Open Research Data concept in the strategic plans of Czech universities

There is neither explicit nor implicit reflection of Open Science principles in the current strategic plans (i.e. for years 2016 – 2020) of most of the Czech universities. There are 22 universities, which make up more than three quarters of the entire university sector. Five universities partly reflect concept of Open Science in their strategic plan. Specifically, the Academy Of Performing Arts in Prague, Czech Technical University

in Prague, University of South Bohemia in České Budějovice, Palacký University Olomouc and marginally Academy of Arts, Architecture and Design in Prague. Only Masaryk University and Brno University of Technology reflects explicitly Open Data in their strategic plans. In case of strategic plan of Masaryk University there is the statement (MU 2015): "*Developing tools for the efficient storage, processing and retrieval of scientific data (i.e. open research data) as well as relevant university policies*." In case of Brno University of Technology there is than statement: "*to publish Open Data and create knowledge databases*".

It should be noted that there may be documents of lower strategic level regulating the issue of Open Research Data at individual universities. However, this is not a major cross-university priority, resulting from absence in the university's strategic plan.

CONCLUSION

The Open Science concept has been the subject of European Union science policy at least in the last decade and represents a more general trend in science development. The Czech Republic has joined this trend at the government level since 2017. At the level of the Association of Libraries of Czech Universities, this trend has at least been formally reflected in the Czech university sector since 2012. The reality of the strategic plans of the Czech universities is different. Two thirds of Czech universities do not reflect this issue in their strategic plans. Only Masaryk University and Brno University of Technology reflects explicitly Open Data in their strategic plans. It should be added that this state is not due to the fact that the concept of open data has already been satisfactorily implemented in the day-to-day research of Czech universities.

In conclusion, it can be stated that the significance of the Open Data concept is still underestimated by Czech universities. Thus, there is still considerable potential for development in this area. And it is also one of the ways of opening and improving high-school education and research of Czech universities.

ACKNOWLEDGEMENTS

This article was created with the subsidy of the means of institutional support for long-term conceptual development of science and research at Faculty of Informatics and Statistics, University of Economics, Prague.

REFERENCES

Access to and Preservation of Scientific Information in Europe. Report on the implementation of Commission Recommendation. EC 2012. 4890 final.

Act no. 111/1998 coll. (amended and consolidated) on Higher Education Institutions and on Amendments and Supplements to Some Other Acts (the Higher Education Act), Czech Republic

Dlouhodobý záměr VUT 2016-2020 (BUT 2018) from https://www.vutbr.cz/uredni-deska/strategicke-zamery-vut/strategicky-zamer-vut-f76007/dlouhodoby-zamer-vut-2016-2020-d168817/dlouhodoby-zamer-vut-2016-2020-pdf-p157782

Fabián, O. (2013) "Open access in the Czech Republic: an overview", Library Review, Vol. 62 Issue: 4/5, pp. 211-223, https://doi.org/10.1108/LR-09-2012-0096

Guidelines on FAIR Data Management in Horizon 2020. EC 2016. Version 3.0.

Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. EC 2017. Version 3.2.

Guiding principles for findable, accessible, interoperable and re-usable data, publishing version B1.0. from https://www.force11.org/fairprinciples

Masaryk University Strategic Plan 2016–2020. MU 2015. ISBN 978-80-210-8116-1, from https://www.muni.cz/media/3062393/strategic_plan_mu_2016-2020_en.pdf

Public Consultation ‘Science 2.0’: Science in Transition. EC 2014, from http://ec.europa.eu/research/consultations/science-2.0/background.pdf

Wilkinson, M. D. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18