

DATA ARCHIVING IN SOCIAL SCIENCES FROM THE PERSPECTIVE OF THE CZECH SOCIAL SCIENCE DATA ARCHIVE (ČSDA)

Martin Vávra

Institute of Sociology of the Czech Academy of Sciences

11th Conference on Grey Literature and Repositories

24th October 2018

Archiving of social science data in the Czech Republic as a result of foreign inspirations and international cooperation

- In the presentation, I would like to take a look at a contemporary state of the process of archiving of social science data in the Czech Republic, highlighting the role of influences from abroad and of a modern trend of international cooperation.

Starting points and objectives of the open access to data

"Open access" principles for scientific information also apply to data. These principles have been formulated at an international level - for example, the OECD (2007, Principles and Guidelines for Access to Research Data from Public Funding)

- *"Openness means access on equal terms for the international research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination. Open access to research data from public funding should be easy, timely, user-friendly and preferably Internet-based."*

From the recommendations and the unenforceable principles, we will move to the rules rooted in the rules for providing research with public money – see: EU Horizon 2020, where, the principle of the program says it should be ensured that there will be an open access to the data, created within the research funded by the program.

The meaning of open access to scientific data

- it allows the new conclusions to be based on the previous science results and, thus, supports the cumulative aspect of science and the quality of knowledge
- it facilitates a control of scientific procedures (through the replication of analyzes)
- it promotes cooperation (also at the international level) and reduces the likelihood of duplication of research on the same topic, thereby supporting the efficiency of spending money
- it facilitates involvement of a wider spectrum of actors in the field of science (see the current trend of "citizen science")

The Czech Social Science Data Archive (ČSDA)

- Founded in 1998 under the title Sociological Data Archive
 - Direct inspiration from foreign archives, such as the German Zentral Archive at the GESIS
- Gradual growth, established as a separate department of IS CAS
- In 2011, the name changed to the Czech Social Data Archive - a broader focus on acquisition policy
- At present: more than 800 data sets, mainly from sociological survey research
- ČSDA is not only an archive but also a methodological center
- The basis for further development is incorporation into CESSDA (Consortium of European Social Science Data Archives)

Access to social science data at European level - CESSDA

- CESSDA was originally an informal association of archives
- CESSDA is currently a part of the European Research Infrastructure Consortium (ERIC)
- CESSDA now includes 18 European archives
- CESSDA strives to create a one stop shop to serve as an access point for all data stored in the archives of individual CESSDA members
- CESSDA is committed to the European Open Science Cloud (EOSC) initiative and to the principles of FAIR Data (findability, accessibility, interoperability and reusability)

Standards of archiving

Creating and enforcing transnational standards is a necessary step in the process of integrating data archives

- Metadata standard DDI - a common way of describing data, its starting point was, at the beginning, the Dublin core
- Thesaurus for Social Sciences ELSST - structured keyword dictionary
- CoreTrustSeal Certification System – certification of trusted digital archives

Current challenges for data archiving

Although in the social sciences in the Czech Republic (and generally in all the countries that are CESSDA members) the archiving and access to data are quite good. But there are still things that should be improved:

- the need to improve the culture of data sharing
- the need to raise researchers' awareness of the need to create "data management plans" and generally better work with data
- the need to include new types of data – i.e. social media data

Conclusions

- The basic infrastructure already exists (although, of course, it still needs to be developed, especially at the European level)
- The need to define (of course, only at the level of the state's science policy, but if possible on the basis of the interested actors' debate) what form of open access to scientific data we want and how we would like to support it, and how we will encourage scientists and their institutions to cooperate on that issue.

- Thank you for your attention

- Further comments and questions please send to martin.vavra@soc.cas.cz