International Journal of Population Data Science

Journal Website: www.ijpds.org

Understanding the barriers researchers face in accessing sensitive data across multiple trusted research environments in one study

Jano van Hemert[1], Emily Symmons[2], Jennifer Muir[3], Kara Moraw[1], Chris Wood[1], Ioanna Lampaki[1], Jennifer Muir[3], Kara Moraw[1], Chris Wood[1], Ioanna Lampaki[1], Ryan Bremner[3], Kostas Kavoussanakis[1], Carmen Amador[5], Sophie Elfar[2], Roger Halliday[3], and Carole Morris[4]

[1]University of Edinburgh, Edinburgh, United Kingdom
[2]Office for National Statistics, Titchfield, United Kingdom
[3]Research Data Scotland, Edinburgh, United Kingdom
[4]Public Health Scotland, Edinburgh, United Kingdom
[5]National Records of Scotland, Edinburgh, United Kingdom

Objectives

Accredited UK researchers can access anonymised data about UK citizens through dozens of Trusted Research Environments (TRE). Each TRE specialises on national or regional data and specific disciplines such as health care and economic data. We investigate the barriers researchers face who require data combined across these specialised TREs.

Method

We sent a survey to approximately 6300 researchers that interacted with Research Data Scotland or ONS IDS, and to the Health Data Research UK mailing list. We asked what metadata catalogues they use to identify sensitive data for their studies. We included questions to elicit feedback on the features and types of metadata researchers use already and would like to access. In parallel, we designed and are running a study on sensitive data from two national TREs: the Public Health Scotland's National Safe Haven and the Office for National Statistics' Integrated Data Service. We recorded process steps and identified discrepancies between processes from the TREs from start to finish.

Results

We collected responses from 28 researchers on the survey. The top three missing features are for each variable the amount of missing data, its range and its summary statistics (e.g., median). In the study across two national TREs we recorded discrepancies in the following aspects of the service and information governance models: approval processes for the study, requirements for approval of the study, accreditation processes for the researchers, cost models, agreements with data controllers on the use of data and linking and indexing approaches. These discrepancies meant one TRE would allow something another would not. Both TREs are accredited under the Digital Economy Act 2017 to provide access to personal data from the public sector.

Conclusions

To enable research on federated data researchers will need publicly available metadata that provides statistical and quality properties of sensitive data. To reduce time and effort to gain federated access, TREs need to agree to align or unify their processes, notably the process to approve studies.

https://doi.org/10.23889/ijpds.v10i3.3242