

International Journal of Population Data Science

Journal Website: www.ijpds.org



Pilot Federated Analysis Projects

Simon Thompson¹

¹Swansea University

The Cambrian explosion of Trusted Research Environments (TREs) will inevitably lead to increased data siloing and potentially decrease achievable data science. Federated access to data is a mitigation against this risk. The presentation will cover two major infrastructure demonstrator projects to address this need and start the UK journey towards interconnected TRE's.

TELEPORT was an infrastructure to connect TREs and allow all the data assets to be seen through a single pane of glass while leaving them in situ; the analysis can be performed on each data asset, as well as combining data from multiple TREs to derive knowledge. The analysis is performed within the construct of a pop-up TRE, which is a temporary project-specific TRE. The data governance model is operated in an as-is state, with all egress controls being applied upon extraction of research output. This method has been termed Federated Data.

TREFX is a set of infrastructure based around international GA4GH standards and the RO-CRATE concept to provide a platform to send complex workflows to the TRE's to be performed upon the data assets, with the output post egress checking to be re-packed into the inbound RO-CRATE prior to being returned to the researcher. This method addresses both federated analysis and method reproducibility. This project was extended to cover ADDI workloads for dementia.

Both solutions were jointly developed by Swansea University / SeRP and share common elements. Future work will see both these products being combined into a hybrid approach as well as spin-out products and services.

