

Novel network analysis of all research requests in Scotland

Stella Telford¹, Rosie Seaman¹, and Jennifer Muir¹¹Research Data Scotland, Edinburgh, United Kingdom

Objectives

Research Data Scotland (RDS) launched the Researcher Access Service (RAS) in 2024. The service was codesigned to meet researcher and data controller needs and maximise the value of and scale administrative linked data in Scotland. Analysis was carried out on all research requests in Scotland since 2015 to identify datasets most commonly asked for in combination to scale the service efficiently.

Method

A dataset that included all research requests in Scotland since 2015 was analysed to identify all unique datasets requested for research in Scotland. Network analysis was then used to understand the relationships and connections between datasets. The networks were visualised, and different parameters were used to better understand connections. Community detection was used to identify the datasets which needed to be provided together to maximise the number of projects which could use the Researcher Access Service pathway.

Results

The work identified that there have been 123 unique datasets requested for research in Scotland since 2015. Of these 123, 114 were not already available through the Researcher Access Service.

The top five most requested datasets that were requested in combination with RAS datasets, alongside the number of times requested are as follows:

1. SBR (Scottish Birth Record) – 131
2. SICSAG (Scottish Intensive Care Society Audit Group) – 84
3. ECOSS (Electronic Communication of Surveillance in Scotland) – 75
4. SCI Diabetes (Diabetes Register) – 70
5. NRS Stillbirths – 64

A number of communities of datasets were identified, for example Education datasets, health datasets, and imaging datasets. Identifying these communities of dataset requests, the need for groups of datasets to be provided together through the RAS became evident.

Conclusion

The work has informed decisions on how best to scale the service. Only when we consider the combinations of datasets being requested and the connections between datasets, are we able to design a modern, scalable and sustainable service.



Temporary page!

\LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because \LaTeX now knows how many pages to expect for this document.