

Using data to support Troubled Families

Or data matching and analysis with modern tools

Our aim was

To use modern software to support the matching and analysis of data used in the Troubled Families programme

- Data matching – Microsoft Data Quality Services (DQS)
- Analysis – Rapidminer

Matched

- over a million person-address records into almost 600,000 persons living in (current/historical) 97,000 addresses.
- from 20 sources/data sets, various spreadsheets including from the DWP and the Job Centre, others from internal systems (education and social care) and nationally available address lists.

Analysis quickly showed

- characteristics of families we had claimed for and not claimed for
- that a group of families were claimed for more quickly than the others and what the characteristics of these families were

Problems

- Staff changes leading to loss of skills
- GDPR

The question running through the project was whether this was ethical.

