

International Journal of Population Data Science



Journal Website: www.ijpds.org

The power of linked data for supporting operational decision making in the Criminal Justice System

Andreas Soteriades¹ and James Forster¹

¹Ministry of Justice, London, United Kingdom

Objectives

Timely data access can be a significant problem across the Criminal Justice System as information is often found across multiple databases with no common identifiers. Our digital dashboard links and brings key information together, providing probation practitioners with an integrated, comprehensive view of individual offender data, facilitating timely, effective decision-making.

Methods

We built data pipelines using R and SQL to link data and fed the outputs to a dashboard built with Golem, a state-of-the-art R Shiny framework for building production-grade applications. These pipelines were built to be efficient and scalable and utilised the R package dbplyr and data build tools (dbt), both of which allow data manipulation at scale using SQL behind the scenes. Data were linked with Splink, an open-source algorithm for fast, accurate and scalable probabilistic data linkage. All tasks to run the pipeline were automated using Airflow and deployment of the dashboard is managed by Kubernetes.

Results

This work resulted in the development of a high-quality user-friendly dashboard, which has been successfully deployed and trialled in one Probation Delivery Unit and allows probation practitioners to easily access information they need for supervision of offenders.

Conclusion

The dashboard has demonstrated what can be achieved when data is linked from internal and external sources. By making required data more accessible, operational issues can be reduced with accompanying increases in productivity and user satisfaction.

