

An Identifier Scheme for the Digitising Scotland Project

Özgür Akgün¹, Ahmad Alsadeeqi², Peter Christen³, Tom Dalton¹, Alan Dearle¹, Chris Dibben⁶, Eilidh Garrett⁴, Alasdair Gray², Graham Kirby¹, Alice Reid⁵

¹ School of Computer Science, University of St Andrews, UK
{ozgur.akgun, tsd4, alan.dearle, graham.kirby}@st-andrews.ac.uk

² Heriot-Watt University, Edinburgh, UK

³ The Australian National University, Canberra, Australia

⁴ University of Essex, UK

⁵ University of Cambridge, UK

⁶ University of Edinburgh, UK

The Digitising Scotland project is having the vital records of Scotland transcribed from images of the original handwritten registration books. Linking the resulting dataset of 24 million vital records covering the lives of 18 million people is a major challenge requiring improved record linkage techniques. Discussions within the Digitising Scotland project have been hampered by each group using their own identification scheme, generally based on the order that the certificates have been processed.

To enable fruitful discussions between the distributed Digitising Scotland team, we required a mechanism for uniquely identifying each individual present on the certificates. From the identifier it should be possible to determine the type of certificate and the role played. We have devised a protocol to generate for any individual on the certificate a unique identifier, without using a computer, by exploiting the National Records of Scotland's registration districts. Importantly, the approach does not rely on the handwritten content of the certificates. The resulting scheme has improved the internal discussions within the project.

This work was presented at the following workshop.

Venue ADRN 2017 - The UK Administrative Data Research Network
 Annual Research Conference
Location The Royal College of Surgeons of Edinburgh
Date June 1-2, 2017
URL <http://www.adrn2017.net/>