NCRM 2017-2019 - Placement Fellowships

Advertised Placement – Role and Person Specification

Placement Title

The Effect of Inferred Identity on Preserving Privacy, Identification and Disclosure of the Digital Self

Name of receiving organisation	**Government Office for Science and The Alan Turing Institute**
Location of placement	**1 Victoria Street, London, SW1H 0ET**
Desired length of placement	**6 months**
Desired time commitment per week	**Full or part-time (e.g. 0.5FTE)**
Approx. Start Date	**Autumn 2017**

Project Description (max. 400 words)

Please ensure that you explain the relevance of social science to the proposed Fellowship.

The scope and locus of identity has been changing for some time. Thanks to the increasing ubiquity and capability of technology that requires tighter understanding of the human operator, digital identity has become the prevalent, if not sole, factor for many every-day activities. A Foresight report on future identity published by the Government Office for Science [1] reinforces this point stating that; "*Over the next 10 years, people's identities are likely to be significantly affected by several important drivers of change, in particular the rapid pace of developments in technology. The emergence of hyper-connectivity (where people can now be constantly connected online), the spread of social media, and the increase in online personal information, are key factors which will interact to influence identities.*"

Much digital identity is handed over willingly by service users in exchange for services yet there has been no systematic work on what additional aspects of identity can be inferred. Service providers are, presumably within the appropriate legal frameworks, inferring identifying features [2] that serve their purpose (usually commercial) but it is typically not obvious what these additional aspects might be. When the purpose of inferring identity is criminal in nature (e.g. identity theft), the effects of these inferences can be severely detrimental.

At the heart of the Super Identity project [3] is a model that combines physical and digital identity artefacts (specific data points which denote personal identity either directly or indirectly) to provide an interpretation of an individual's identity between the physical and digital domains. This model encodes the relationships between identity artefacts primarily for the purpose of investigative identification but could be repurposed to highlight attribution risk through unintended disclosure [4]. Where one identity artefact is embedded within another (i.e. area code within telephone number), we seek to establish the value of that embedded artefact as highlighted in the results of this study [5].

Suppose a service provider requires a telephone number to create an online account. What identity artefacts are contained within that number? Can you infer the network operator, the age of the number, the location (if it has an area code)? This is an example of an associational unique identifier; a class of formal identifiers. It is therefore not just an understanding of the structure of number (identity artefact) that is critical but its relationship with other artefacts. Only by understanding these inferences can we understand the full exposure of any digital transaction and provide the ability to make an informed choice.

1. Foresight Future Identities (2013) Final Project Report. The Government Office for Science, London.
2. 98 personal data points that Facebook uses to target ads to you. Washington Post, 19/08/2016. https://www.washingtonpost.com/news/the-intersect/wp/2016/08/19/98-personal-data-points-that-facebook-uses-to-target-ads-to-you
3. SID: An Exploration of Super-Identity. (2011 - 2015). EPSRC. http://gow.epsrc.ac.uk/NGBOViewGrant.aspx?GrantRef=EP/J004995/1
4. Elliot, M. Anonymisation, Risk and Privacy. New Approaches to Data Privacy Workshop, December 2016
5. Elliot, M., Mackey, E., O'Shea, S., Tudor, C., & Spicer, K. (2016). End User Licence to Open Government Data? A Simulated Penetration Attack on Two Social Survey Datasets. Journal of Official Statistics, 32(2), 329-348.

Key Deliverables (max. 200 words)

The key output from this project would be a report that contains the following;

- A review of the impact of identity inference on aspects such as privacy, anonymity, pseudonymity and identification via weak indicators.
- A review of existing identity artefact mappings from the Super-Identity Project and the identification of key areas for development to support discovery of online identity exposure.
- An investigation into several every-day online tasks to understand the true exposure of digital identity during these interactions in the digital world.
- An Identification/development of an approach for ultimately reducing the amount of digital identity artefacts required for digital interactions.

Role Description (max. 400 words)

The fellowship will be located at the Alan Turing Institute and will work closely with the Government Office for Science.

Through association with the Alan Turing Institute, the role will involve collaboration and the possibility of leveraging the skills therein; data science, computer science, behavioural science, etc.

There are multiple angles from which to attack this problem and we seek innovative methodologies without having any predispositions towards a particular approach.

We suggest a creative blend of social, computer and data science or at least an appreciation of multiple fields (i.e. statistics). Given that this is an ESRC-funded project, we expect the focus to be more towards social science methodologies to establish the relationships between identity concepts but ideally with an understanding of how computation and formal modelling might play an important role. Some experience of behavioural modelling, logic programming or handling uncertainty would also be useful.

Whilst not tied to a specific or impending policy intervention, the outputs of this work will go on to inform the policy discussions and formulation that government undertakes in this space in the usual manner, including through exposure to the Government Chief Scientific Adviser.

Dissemination of project findings will be subject to confidentiality and shared Intellectual Property agreements made between the Government Office for Science and the NCRM.

The successful candidate will need to go through standard security clearance processes. The level of security needed is Security Check (SC). This can take at least two months.

Research Methods (max. 200 words)

NCRM Fellowships must involve the application or development of advanced research methods. In this section, please clearly detail how the proposed placement will deliver on this objective.

The proposed placement will advance the current thinking on identity management and potentially offer new approaches for managing the exposure of digital identity. This advances the UK Government's goals for Cyber Security and also has potential benefits for all users of digital online services.

Person Specification

Criteria	Essential	Desirable
Qualifications and academic experience	A research background in a scientific discipline (i.e. social, behavioural, data or computer science).	
Knowledge & experience of specific Research Methods		Some experience of behavioural or knowledge modelling, formal modelling, dealing with uncertainty. Some experience of using data-driven methods for determining behavioural insight.
Knowledge & experience working with public sector		An interest in cyber and/or national security is desirable but not essential.
Communicating and influencing	Good oral and written skills.	
Other skills		
Other skills		
Other skills		

Further details

Further details on this specific role can be obtained from:

Name	Ben Tagger
Role	Science and Technology Advisor
Email	ben.tagger@go-science.gsi.gov.uk
Telephone	07860 827 444

Further details on the application process can be obtained from:

Name	Alexandra Frosch
Role	NCRM Centre Manager
Email	a.s.frosch@soton.ac.uk
Telephone	023 8059 7473