Opportunity Title: Development of Techniques to Assess Data Aggregation Fellowship
Opportunity Reference Code: ICPD-2024-45

Organization Office of the Director of National Intelligence (ODNI)

Reference Code ICPD-2024-45

How to Apply **Create and release your Profile on Zintellect** – Postdoctoral applicants must create an account and complete a profile in the on-line application system. **Please note: your resume/CV may not exceed 3 pages.**

Complete your application – Enter the rest of the information required for the IC Postdoc Program Research Opportunity. The application itself contains detailed instructions for each one of these components: availability, citizenship, transcripts, dissertation abstract, publication and presentation plan, and information about your Research Advisor co-applicant.

Additional information about the IC Postdoctoral Research Fellowship Program is available on the program website located at: https://orise.orau.gov/icpostdoc/index.html.

If you have questions, send an email to ICPostdoc@orau.org. Please include the reference code for this opportunity in your email.

Application Deadline 2/28/2024 6:00:00 PM Eastern Time Zone

Description **Research Topic Description, including Problem Statement:**

Problem statement: development of a methodology to enable identification and repeatable assessment of risks arising from the aggregation of data sets. This issue is becoming more acute due to the existing volume of published information about national infrastructure produced at the behest of policy makers and regulators, but without adequate consideration of the potential intelligence value to hostile actors and risks associated with these aggregated information sets.

It is recognized that data aggregation arising from combinations of data sets can result in revealing or allowing the inference of information that is not contained in the aggregated data. For data that may be linked to an individual or groups of individuals, it is difficult to measure re-identification or de-anonymization risks that may arise in ways that are both general and meaningful. For data relating to physical assets, it can be difficult to assess what can be inferred about the criticality or sensitivity of the assets and their associated infrastructure.

There have been several publicized examples of anonymized data being de-anonymized enabling the identification or re-identification of individuals and locations. At present there is no published guidance defining how to assess the potential consequences of data aggregation, nor are tools available that allow testing or formal evaluation of combined data sets prior to their publication or disclosure.

While there is some understanding of the issue in respect of personal and travel data, the concept is poorly understood with regards to asset data, particularly relating to infrastructure assets, where factors such proximity, interconnection, and interdependence can create criticalities. A complicating factor with infrastructure data is the need to understand not only the geospatial relationships but also the significance of facilitating the

Opportunity Title: Development of Techniques to Assess Data Aggregation Fellowship
Opportunity Reference Code: ICPD-2024-45

disclosure or inference of links between sensitive or potentially sensitive physical assets/sites and the infrastructure that supports them.

Example Approaches:

Examples of potential data aggregation threats include:

- Identity disclosure – associating individuals with specific records and/or locations which may arise from insufficient de-identification, re-identification by linking data from two or more sets, or from pseudonym reversal.
- Attribute disclosure – identifying an attribute in a dataset held by a specific individual, group of individuals, or by asset(s) with high probability, even if the data associated with the targeted entities are not identified.
- Inferential disclosure - making an inference about an individual, group of individuals, location(s) or asset(s) with high probability, even if the targeted entities were not in the dataset prior to de-identification/anonymization.

The latest draft of NIST SP 800-188 - De-Identifying Government Data Sets (https://doi.org/10.6028/NIST.SP.800-188) provides some background to the issue and an extensive list of references. The proposed research would build upon this to develop methods and, where practical, tools to assist users to identify and address potential aggregation issues.

Key Words: Data aggregation, re-identification, de-anonymization

Qualifications

Postdoc Eligibility

- U.S. citizens only
- Ph.D. in a relevant field must be completed before beginning the appointment and within five years of the appointment start date
- Proposal must be associated with an accredited U.S. university, college, or U.S. government laboratory
- Eligible candidates may only receive one award from the IC Postdoctoral Research Fellowship Program

Research Advisor Eligibility

- Must be an employee of an accredited U.S. university, college or U.S. government laboratory
- Are not required to be U.S. citizens

Eligibility Requirements

- **Citizenship:** U.S. Citizen Only
- **Degree:** Doctoral Degree.
- **Discipline(s):**
  - **Chemistry and Materials Sciences** (12 👁)
  - **Communications and Graphics Design** (2 👁)
  - **Computer, Information, and Data Sciences** (17 👁)
  - **Earth and Geosciences** (21 👁)
  - **Engineering** (27 👁)

zintellect
climb higher

Opportunity Title: Development of Techniques to Assess Data Aggregation Fellowship
Opportunity Reference Code: ICPD-2024-45

- **Environmental and Marine Sciences** (14 👁)
- **Life Health and Medical Sciences** (45 👁)
- **Mathematics and Statistics** (11 👁)
- **Other Non-Science & Engineering** (2 👁)
- **Physics** (16 👁)
- **Science & Engineering-related** (1 👁)
- **Social and Behavioral Sciences** (30 👁)