Response to Topic Area 5: Open Sourced Science Geospatial Data Responsibility by Design

Mishandling location data threatens individual privacy in similar ways to ways to mismanagement a person's name in an information system. Geospatial and location tracking data can be recombined with other publically available, open source and big data to expose personal information just as sensitive, and including, health records or social security numbers. Irresponsible location data management can be more dangerous than name alone for vulnerable people. This is particularly serious for those living in areas threatened by conflict or high levels of insecurity.

Feedback from end users, sensors that collect location, participation of citizen scientists, and engagement of underserved and marginalized communities are system design parameters emerging from the first Earth System Observatory (ESO) design conference. Even as these innovations are essential for ESO to be fit for purpose, they may pose new data responsibility challenges for NASA and Open Sourced Science (OSS) researchers and professionals in disciplines that do not typically include research with human subjects or include active participation from people and a diversity of communities.

NASA applies better practice with data it collects in research involving space flight related human subjects, official communications with a variety of partners in private sector and academia, those requesting access to NASA resources and web site users (link to policy notice on next page). Privacy Impact Assessments (PIAs) are undertaken for a variety of new records systems. It is unclear if PIA and other standard good practice has been applied to the types of high frequency two-way interactions implied by engaging end-users, citizen scientists, and the tools like mobile Apps that are tracking locations.

Two areas of protection concern require dedicated investigation. 1) Increasing spatial and temporal resolution of data streams may enable tracking that reveals individual identities and privacy protected health or behavioral information. 2) Apps on mobile devices, wearable technology, and other potential two-way exchange channel technology raise additional privacy concerns again related to tracking and recombination of individuals' digital signatures.

Privacy protections in the United States are a patchwork that differ across States and the type of data collected. This poses a significant design problem for any EOS data system processes that collect, store, analyze or share personal, location and potential geospatial data that can be recombined by domestic partners. Internationally, stricter privacy regulations usually apply. General data Protection Regulations (GDPR) standards in the European Union enshrine two concepts that may be useful to consider for geospatial data responsibility. 'Data protection by default' means that any service needs the user permission to 'opt in' to data collection and that no more data than necessary is collected. Most important for the EOS design, 'Data protection by design' means that a foundational principle at every stage of information system design favors enabling technology and design choices in favor of privacy and security. UN data protection principles serve as another source of standards to enable engagement across international borders.

OSS Geospatial Data Responsibility by Design, if adopted by ESO, would identify better practice and establish a protocol to ensure all data and derivative products from Open Science partners take a minimum set of precautions with location and geospatial data.