FOT-Net Data International Workshop

Data Considerations for Deployment Phase

Ram Kandarpa
October 05, 2015

Booz | Allen | Hamilton
100 YEARS

Conversations regarding data re-use should begin during the pre-deployment phase, particularly when drafting data collection plans

▸ Data collection strategies are often developed by team responsible for deployment

▸ Deployment team may not think beyond the immediate use of the data

▸ Considerations for post-deployment data re-use should begin during pre-deployment phase

▸ Considering data re-use during pre-deployment brings to the "forefront" many issues that are currently dealt with post-deployment
  – For example, erroneous data, gaps in data, data reliability, external factors that may influence data, data documentation, data anonymization / sanitization, and data accessibility

▸ Include external team of professionals / researchers to maximize benefit of pre-deployment consideration of data re-use

▸ External team will bring additional perspectives to data collection, to better position the collected data for re-use after the deployment

▸ To maximize effectiveness of this approach, combine perspectives of external team with lessons learned from previous FOT data re-use activities

Applying insights from prior FOTs will facilitate more effective "use" and "re-use" of collected data

▸ Solutions to potential data re-use issues should be built into the deployment design, thus avoiding these issues post deployment

▸ For example:

– **Data Gap Issues**: Often times re-use of data is encouraged by availability of contextual (and complementary) data, which may already be collected without the deployment team spending significant resources

– **Privacy Concerns**: In terms of collected data containing sensitivities, including PII, different levels of data collection strategies may be applied in parallel, with one potentially placing emphasis on complete data collection, while another focusing on collecting non-sensitive data

• For vehicle trajectory data, pre-deployment discussions should consider inclusion of "privacy by design" principles and strategies

– **Data Accessibility**: Developing a data accessibility policy, which facilitates reuse, can inform how data is not only collected but also how it is stored – down to the architectural design level, to allow individuals with different privileges access different features of the data