How to enable call-for-data in the era of mydata?

Aapo Rista and Hanna Niemi-Hugaerts

*Forum Virium Helsinki*
Opening data

Users contributing and creating data

(Harmonized) APIs

New apps and services

Developer engagement
Engaging citizens in data collection
Case Open311

"Citizens are the ears and the eyes of the city. The department of public works cannot know whether traffic sign has fallen, if no one tells them about it."

Deputy mayor of Helsinki
Pekka Sauri
Engaging citizens in data collection
Case Air quality

Data gathered by officials:
- Motivated by the law and regulations
- Few sensors
- Expensive sensors
- High quality, accurate sensors
- Mostly data about the commons, usually no privacy concerns

Data gathered by private citizens:
- So many sensors
- Inexpensive sensors
- Varying quality, less accurate sensors
- Data is likely to contain sensitive information
City’s data sources - officials vs. citizens

Data gathered by officials:
- Motivated by the law and regulations
- Few sensors
- Expensive sensors
- High quality, accurate sensors
- Mostly data about the commons, usually no privacy concerns

Data gathered by private citizens:
- Motivated by own and/or common needs??
- So many sensors
- Inexpensive sensors
- 3rd party apps, services and hardware
- Varying quality, less accurate sensors
- Data is likely to contain sensitive information, privacy must be taken into account
More accurate and timely data needed

- Public sector organizations could improve their operations, service quality and customer experience if they could engage citizens in data gathering:
  - What are the most popular running paths?
  - What are the smoothest bike routes for commuters?
  - Where are the potholes?
  - What are peoples’ preferences and interests in certain area?
Thoughts about UI
Call-for-data requirements for mydata platform

- **Platform should support data donation but also calls for data**
  - Broadcasting not just available data sources but also need for certain type of data
  - How could the city call for sport facility usage data?

- **Platform should enable dynamic consent**
  - GDPR requirements
  - But how do we ensure the right to be forgotten in cases where data has already become part of statistics or other aggregated data entity?

- **Platform should allow user to limit access to data based on spatial boundaries and/or temporal resolution**
  - E.g. user wants to leave out data related to his home or workplace
  - How would citizens give consent for using their data and define the level of detail that the city can explore and use that data?

- **Platform should enable safe way to store user-related data**
  - Would the donated data be stored in city’s or third parties' data lakes?