Emerging Roles for Optimising Re-Use of Open Government Data (OGD)

Xiao, F., Lyon, L., Zou, N., Gradeck, R.

February 21, 2018
3924 Theft of Property Lost, Mislaid or Delivered by Mistake.

100 Block Smithfield St
Pittsburgh, PA 15222

Central Business District

located at nearest corner to the arrest location.
Background Information

Transportation Data

Environment Data

Crime Data

Housing and Property Data

Others......
Research Questions

- **RQ1:** What are the common missions of open government data centers?

- **RQ2:** What user services and supportive tasks are provided by open government data centers?

- **RQ3:** Which specialist job types are needed to deliver these open government data services?
Methodology
Methodology—Research samples

Three Perspectives:

- Geographical distribution, the city size, and scale
- Maturity
- Familiarity

- NYC OpenData (New York Open Data Center)
- DataSF (open data portal of San Francisco)
- WPRDC (Western Pennsylvania Regional Data Center)
- London Datastore (Greater London open data portal)
Methodology—Content Analysis

Content Analysis

collected, examined and analyzed

Table 1. The OGD official websites

<table>
<thead>
<tr>
<th>OGD portal</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYC OpenData</td>
<td><a href="https://opendata.cityofnewyork.us/">https://opendata.cityofnewyork.us/</a></td>
</tr>
<tr>
<td>DataSF</td>
<td><a href="https://datasf.org/">https://datasf.org/</a></td>
</tr>
<tr>
<td>WPRDC</td>
<td><a href="http://www.wprdc.org/">http://www.wprdc.org/</a></td>
</tr>
<tr>
<td>London Datastore</td>
<td><a href="https://data.london.gov.uk/">https://data.london.gov.uk/</a></td>
</tr>
</tbody>
</table>
Methodology—Content Analysis

Data Collection

Extracted raw data about the mission, services, and the tasks. This step, the coder simply collected relevant data into the three classes:

- Analyzed and Classified the services information
  - Five Categories:
    - Availability
    - Understandability
    - Technical Help
    - Social Engagement (Interactive)
    - Improving User Data Literacy

Data Analysis

- Analyzed and Classified the Supporting Tasks information into same five categories.
Results
Results — Five Categories

**Availability**
be available to anyone who wants to access, download and use them

**Social Engagement (Interactivity)**
Promote citizens participation in the social activities

**Technical Help**
Create APIs or tools to better serve various types of data users to use and apply data

**Understandability**
Be easy to be understood.

**User Data Literacy Improvement**
Enhancing user data skills is a necessary requirement for empowering data use.
RQ1: What are the common missions of open government data centers?

- Engage citizens to use OGD
- Empower use of data
- Easier to find and use
- Free London’s data & present the data in a meaningful way
RQ2: What user services and supportive tasks are provided by open government data centers?

- **Availability**: Platform for publishing data; Search for data; Browse by Category, publisher, new datasets, popular datasets, Download data.
- **Understandability**: Guidelines for using the platform for new users; Guidelines for API documentation for data veterans, Metadata & data dictionary, Show cases (data analysis report).
- **Technical Help**: APIs, Data visualization tools.
- **Social Engagement**: Ask questions, leave comments and request datasets.
- **Social Engagement**: Training to enhance skills in data use, Providing technological training-related information, Help desk.
**Results — Supporting Tasks**

### Understandability
- Provide context for data for making data useful, such as making sense of data.
- Provide Show cases or analysis reports.
- Operationalize the City’s data strategy making data more accessible and useful at the City.

### Social engagement (Interactivity)
- Encourage more publishers to publish data.
- Engage users through activities, training, technical assistance and other events.
- Monitor and respond to comments on public datasets.
- Feedback loops from users to publishers.

### Technical help
- Develop tools to help make data more accessible and improving tools and features.
- Visualize or build apps from the data available on the site.

### User data skills improvement
- Raise awareness of additional resources and data sources.
- Organize help desk to help users.
- Privacy education and consultation

---

**Work Content**

- Clean data
- Publish data
- Update data
- Review data
- Extract, transform and load services
- Make sure a good data quality and management information (such as metadata and data dictionary)
- Maintain the open data platform: software and servers
Discussion
Discussion

MISSON

WORKS & SERVICES

NEW ROLES

- Provide context for data
- Provide Show cases or analysis reports
- ……

Understandability

Data Interpreter

Data Consultant

Data Visual Assistant

Data Skills

- Organize “Help Desk” to help users
- Data literacy-related trainings
- ……

Tools

- Develop tools to help make data more accessible
- Visualize or build apps from data to assist users to use data
- ……
Discussion

Data Interpreter

The goal of this role is ‘to make sense of data for users’. A data interpreter’s specific activities consist of working with data providers to create data guides, collecting and creating data-related blogs and data stories, working independently or with programmers to make maps or other visualizations, and informing data-related policies.

The role starts with open government data.
Discussion

**Data Consultant**

The goal of this role is ‘to directly assist users to understand the data and teach them technical skills to accurately use the data’. This job requires that the consultant will hold help-desk hours each week to help users who have difficulties with the data, especially from the perspective of technology; organizes meetings to collect information from various groups of people, including the data category needs or the required tools, and then finds the technical solutions to meet their needs.

*This role starts with OGD users.*
Data Visual Assistant

The goal of this role is ‘to assist users to visualize or manipulate data by developing tools that can be directly and easily used by users’. This role’s focus is to develop software tools and apps that can (easily) create data visualizations, such as line charts, bar charts, maps and other infographics. Users can then use the graphical tools and apps to create the specific visualizations they want by simply selecting the parameters and applying them to the whole dataset or data sub-set.

The role starts with tools.
Data interpreter/data consultant roles in each domain, could help to maximize the understanding of complex data and minimize the subsequent misuse of data.

Researchers’ concerns about sharing data
- legal issues
- misuse of data
- incompatible data types

(Tenopir et al., 2011).

Concerns
Risk that data may be misinterpreted due to the complexity of data
Conclusion
Conclusions

- **Three new and specialist OGD roles:**
  - Data interpreter, Data Consultant, Data visual assistant

- The findings of this study may provide valuable indicators for open government data portal managers in developing strategy, planning operational services and in allocating resources for new positions to deliver on such plans.

- The high-level descriptions for the three new specialist roles, together with description of the micro-tasks which they may deliver, provide a good foundation for putative job descriptions for open government data centers to use in the future.
Thanks and Questions
Backup Slides
Open government data (OGD) refers to the governmental data should be available to anyone to access, reuse, and redistribute without copyright restrictions (Kassen, 2013; Veljković et.al., 2014).

OGD aims to promote transparency, accountability and value creation by making government data available to all (OECD, 2018).
Purpose

The purpose of this paper is to propose emerging data roles for open government data (OGD) centers by investigating the missions, services and operational tasks provided by four open government data centers.
Research Questions

- **RQ1**: What are the common missions of open government data centers?

- **RQ2**: What user services and supportive tasks are provided by open government data centers?

- **RQ3**: Which specialist job types are needed to deliver these open government data services?
Challenges of OGD

Data quality concerns

Lack of standards

Understandability of OGD

The complexity of open data platforms

Data quality concerns

Lack of standards

Understandability of OGD

The complexity of open data platforms

The required technological skills of users

(Conradie and Choenni, 2014; Huijboom & Van den Broek, 2011; Productivity Commission, 2017; Zhang et al., 2015)

Weerakkody et al., 2017;
Bayley et al., 2013; Verburg & Neumann, 2011

(Huijboom & Van den Broek, 2011; Martin, 2014; Janssen et al., 2012; Osagie et al., 2017)

(Kapoor, Weerakkody & Sivarajah, 2015; Martin, 2014; Janssen et al., 2012)
Motivations

Fill in the gap of lacking of research on developing the roles within the open government data context

- Literature Review
- Practice situation
Methodology—Data Collection

<table>
<thead>
<tr>
<th>OGD portal</th>
<th>Mission</th>
<th>User service</th>
<th>Support task</th>
</tr>
</thead>
<tbody>
<tr>
<td>DataSF</td>
<td>Our mission is to empower use of data. At DataSF, we seek to transform the way the City works through the use of data. ......</td>
<td>Fundamental: search for data; browse by Category, Publisher, View Types, and Tags; guideline of API documentations; training for enhance skills in data use; data management and process improvement; Metadata; data dictionary; ask questions; leave comments and request datasets</td>
<td>1. Developing a catalog of datasets that can be made public 2. Monitoring and responding to comments on public datasets 3. Monitoring and tracking the rollout of public datasets on the open data portal 4. Ensuring that processes are followed to exclude private, confidential or proprietary data</td>
</tr>
</tbody>
</table>

Table 2: OGD website data collection exemplar.