The Complexities of Open Data

Hector Dominguez
Open Data Coordinator
Smart City PDX
What is the Open Data Program

City of Portland has established an Open Data program through City Council Ordinance No. 188356 that commits the City to the publication, open access, and widespread sharing of data collected and generated by the City, and by private sector companies, non-profit organizations, academia, and other parties working on behalf of the City.

But not all data is created the same!
Data at the Digital Age

This current era whereby social, economic and political activities are dependent on information and communication technologies.
Privacy Paradox

How we fear threats on one hand and yet voluntarily give away data elsewhere. Researchers call this conundrum the "privacy paradox".

"People are worried about getting taken advantage of."
Big Data

Big data is data sets that are so voluminous and complex that traditional data-processing application software are inadequate. Big data includes collection, storage, analysis, search, sharing, transfer, visualization, querying, updating, information privacy.
Digital Equity

Human rights principles, such as privacy, freedom of expression, and democracy, must guide the use of the information and digital services, mainly in government and social services.

No all information is the same for everyone. Open and timely access of certain information may harm or benefit specific groups of people.
Equity metrics

**Who/Recipients** - People, organizations and technologies who are the senders, recipients, and subjects of information.

**What/Information** - Information type, data fields being used, transmitted or processed

**How/Transformation principles** - Constraints and processing through the flow of information.

**Why** - Purpose of the use of information

**Where** - Foreseen effects due to the use or processing of this information
Determinants of Equity

Source: King County, WA
Ethics and Data

**Ethics** provides a set of standards for behavior that helps us decide how we ought to act in a range of situations. Ethics is all about making choices, and about providing reasons why we should make these choices.
## Ethics and Data

<table>
<thead>
<tr>
<th>Consequentialist</th>
<th>Duty</th>
<th>Virtue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deliberative process</strong></td>
<td>What kind of outcomes should I produce (or try to produce)?</td>
<td>What are my obligations in this situation, and what are the things I should never do?</td>
</tr>
<tr>
<td><strong>Focus</strong></td>
<td>Directs attention to the future effects of an action, for all people who will be directly or indirectly affected by the action.</td>
<td>Directs attention to the duties that exist prior to the situation and determines obligations.</td>
</tr>
<tr>
<td><strong>Definition of Ethical Conduct</strong></td>
<td>Ethical conduct is the action that will achieve the best consequences.</td>
<td>Ethical conduct involves always doing the right thing: never failing to do one's duty.</td>
</tr>
<tr>
<td><strong>Motivation</strong></td>
<td><strong>Aim is to produce the most good.</strong></td>
<td><strong>Aim is to perform the right action.</strong></td>
</tr>
</tbody>
</table>
Ethics and Data - Consequentialism

hypothesical Consequentialism/Utilitarianism

- Maximum happiness for maximum number of people for maximum amount of time.

Example:- If in a village of 100 people, 99 are happy and 1 is miserable; then according to utilitarianism approach the situation in the village is RIGHT as maximum people are happy.
"Deontology...is an approach to ethics that focuses on the rightness or wrongness of actions themselves, as opposed to the rightness or wrongness of the consequences of those actions"

This theory gives 4 parameters for determining the ethicality of an act:
1. **Universal** – If you can universalize an act and still it is right, then it is ethical.
2. **Self** – Are you ready to accept the same act on yourself.
3. **Means** – Any person/resource should not be used/abused as the means to an end.
4. **Long-term societal impact** – Will the society become better in the long term if this act is practised?
Ethics and Data

Example: If in a village of 100 people, 99 are happy and 1 is miserable; then according to utilitarianism approach the situation in the village is **RIGHT** as maximum people are happy.

- But according to deontology theory, the situation is **WRONG** because:
  1. **Universal** – If you can universalize an act and still it is right, then it is ethical. ✓
  2. **Self** – Are you ready to accept the same act on yourself. ✗
  3. **Means** – Any person should not be used as the means to an end. ✗
  4. **Long-term societal impact** – Will the society become better in the long term if this act is practised? ✓
Principles of Trustable Technology

Privacy & Data Practices
Is it designed using state of the art data practices, and respectful of user rights?

Transparency
Is it made clear to users what the device does and how data might be used?

Security
Is it designed and built using state of the art security practices and safeguards?

Stability
How robust is the device and how long of a life cycle can a consumer reasonably expect?

Openness
How open are both the device and the manufacturer’s processes? Is open data used or generated?
Open Data Principles

Public
Accessible
Described
Reusable

Complete
Timely
Managed Post-Release
Open-by-Preference
Open Data maturity model

1. Status Quo
   Individual spreadsheets and PDFs.

2. The Catalog Phase
   Basic catalog of downloadable files.

3. Basic Interactive Experience
   Interactive data tables and basic visualizations.

4. Full Open Data
   Consumer-friendly websites, automated data publishing.

5. Open Data Ecosystem
   Real-time data, industry exchanges, internal data collection hubs.

According to Socrata, Inc.
Open and Accessible Data

Policy/Standards
- Metadata
- Cataloging Standard
- Privacy
- Licensing
- Data terms for contracting/procurement
- Equity

Technical
- Implementation Policy/standards
- PUDL

Data Inventory
- Use Cases
  - Business layer
- Public Records Request
- Performance Indicators
- Resource intense data
- Value Proposition
City Open Data Governability structure

- Privacy Team
- Ethics & Equity Team
- Community or City
- Data Trustees
- Policy
- Technical
- Executive Sponsorship
- Cybersecurity Team
- Legal Team
- Product Dev Team
- City teams
Information Governance Framework

Auditing and assessing compliance

Application development, deployment, maintenance and end-of-life

Rapid data insights for better business decisions

IT Security and Protection

Document/data production for litigation and regulatory investigations

Privacy framework: Policies, procedures and processes

Data policies, procedures and processes:
- Data quality
- Data management - lifecycle stewardship
- Roles and accessibility

Records and archiving policies: Records retention and disposal schedules

Based on 2017 Sibenco Pty, Ltd
Privacy and Information Protection Principles

- Transparency and accountability
- Full lifecycle stewardship
- Equitable data management
- Ethical and non-discriminating use of data
- Data openness
- Automated decision systems
- Data utility
Enabling City with Real Use Cases
Taming the wild elephant

Our Data journey in 2019