

# A Field Guide to Open Data

Alicia Brown

2021-02-01



# Contents

<b>1</b>	<b>Welcome!</b>	<b>5</b>
<b>2</b>	<b>An Open Data Primer</b>	<b>7</b>
2.1	What is Open Data? . . . . .	7
2.2	How is it used? . . . . .	8
2.3	Who uses it? . . . . .	8
<b>3</b>	<b>Use of Open Data</b>	<b>9</b>
3.1	Examples of Usefulness . . . . .	9
3.2	Measuring Use . . . . .	9
3.3	Challenges of Open Data . . . . .	9
<b>4</b>	<b>Imagining Open Data</b>	<b>11</b>
4.1	Sharing . . . . .	11
4.2	Accessing . . . . .	11
4.3	Discovery . . . . .	11
4.4	Improving the Ecosystem . . . . .	12



# Chapter 1

## Welcome!

Thank you for joining me on a journey to explore the use of Open Data in the civic world.

The book is organized into 3 parts -

- Part 1 provides an overview of what Open Data is, how it is used and who uses it.
- Part 2 questions the usefulness of Open Data to its producers, how use is measured and what are the challenges faced by producers and consumers of it.
- Part 3 imagines how information provided by Open Data may be shared, discovered and improved.

I have had the great honor to work with many amazing organizations, agencies and governments to implement technology solutions that enabled the unleashing of data and analytics to power decision making and transparency.<sup>1</sup> The purpose for this book is to synthesize what I have learned working in the field with customers and present examples and ideas to grow the Open Data ecosystem.

---

<sup>1</sup>Brown (2020)



## Chapter 2

# An Open Data Primer

### 2.1 What is Open Data?

“Open data is information or content made freely available to use and redistribute, subject only to the requirement to attribute it to the source.”

— Gartner<sup>1</sup>

#### 2.1.1 Topics

Popular data topics include emergency calls for service, budget, revenue, expenditures, property tax, permits, food inspections, government service calls, lost and found pets, crime incidents, jail bookings, voting registration and election results. When this data includes date, time and location, it can be valuable for performance management and refining staffing for services. When it includes demographic information it makes it possible to analyze differences in data by gender, race and ethnicity, and age and can help inform where there should be more outreach and consider mode of communication that may be most effective.

#### 2.1.2 Types

- Tabular
- Spatial
- Dashboards
- Stories
- Visualizations
- Surveys
- Reports

---

<sup>1</sup>Gartner (2020)

## 2.2 How is it used?

### 2.2.1 Organizational Goals

- Public information
- Transparency
- Fine tune staffing and location
- Equity and performance measurement
- Insights and data mining

### 2.2.2 News

### 2.2.3 Civic

## 2.3 Who uses it?

- Producer organizations
- Community users (civic)
- Academic

### 2.3.1 Personas

There are many people and roles across organizations that require data. Here are some personas based on individuals I have met and worked with to get the data they need to perform their jobs. Please note that the names are made up and intended to reflect any specific gender.

- Alex the Analyst
- Pat the Performance Manager
- the Executive



## Chapter 3

# Use of Open Data

### 3.1 Examples of Usefulness

- Outcomes

### 3.2 Measuring Use

- Utilization
  - Analytics
  - Feedback
1. Who are the real users of this data?
  2. How can we tell who they are?
  3. Are there personas we haven't imagined?
  4. How can we measure actual engagement of these users?
  5. Are they using it the way we thought they were? Ex. Hack to “fix” data in between steps.

### 3.3 Challenges of Open Data

#### 3.3.1 Searching for data

- Paging through results
- Onus of filtering

### 3.3.2 Peeking at data

Nobody wants to download 1 million records of taxi rides, but they do want to know why tips increased after a software update.<sup>1</sup>

---

<sup>1</sup>Wellington (2020)

## Chapter 4

# Imagining Open Data

### 4.1 Sharing

Adding datasets to a platform may benefit power users of data and make it easy consumable by tools and programming languages. However data alone will not advance the knowledge of the community it is intended to serve without also including narrative and insights from the collections of datasets shared to a platform.

- Stories & Narrative
- Live tiles
- Data driven documentation
- Dashboards
- Goals
- Reports
- Info sheets

### 4.2 Accessing

- Tools
- Open source library
- APIs

### 4.3 Discovery

- Metadata
- Harvestable
- Standard schemas

## 4.4 Improving the Ecosystem

- Ownership and stewardship
- Automation
- Data refreshes
- Quality monitoring

# Bibliography

Brown, A. (2020). Mapping 5 years of government data. *msaliciabrown.com*.

Gartner (2020). Gartner glossary: Open data. <https://www.gartner.com>.

Wellington, B. (2020). How software in half of nyc cabs generates \$5.2 million a year in extra tips. *I Quant NY*.