1. Why Open Data? (40 mins)

- Intro, why are we here? Greetings, intros. (5 mins) = 5
    - Welcome! This is the first workshop in our Open Data Training Series.
    - So you've probably been hearing about this "open data" thing… that you should do it… that you shouldn't do it… Wondering. If it's right for you-- together in this section we're going to talk through what Open Data is and means, why you should care, and how to talk about Open Data with your colleagues, friends, loved ones and administrators.
    - Let's start with Introductions. Say your name, where you're from, and three words that describe what you do on a daily basis. ~~one thing you'd bring on a desert island, and one challenge you've heard to data sharing.~~
    - (Introduce self)
    - This training was created by Mozilla Science Lab. You may know the name Mozilla, the people that created the Firefox web browser, which is open source. This means that all the code for that project is freely available, and anyone is invited to contribute to work/time/energy to build and improve the product. The Science Lab is another way that Mozilla, through its nonprofit foundation, is championing open and accessible information and encouraging collaboration and participation. Despite our name, at Mozilla Science Lab we don't have an actual laboratory or do any science, but we do innovate to improve research practice and encourage open data and open access to knowledge on the web. More about MSL, resources we offer, and open source and working open later in the [series/day/workshop]
    - In the spirit of participation, of active learning and contribution, we're going to do a series of activities and exercises together today, that you can make relevant to your own work, your own practice and projects. I'm going to talk a bit, hopefully not to much.
- Fun Video (3 mins)
    - We're going to start with a short video that highlights some of the (surprisingly) common issues surrounding closed data.
    - Data Management SNAFU (https://www.youtube.com/watch?v=N2zK3sAtr-4) (0:00- 1:46)
    - There is more to the video but we're going to stop it at this point. I think you get the idea. What are some challenges you've encountered or have heard about when dealing with reusing other people's data?
- Get to a shared definition of open data (2 mins writing, 5 mins sharing) = 7
    - We've bandied about the phrase "open data" a couple times but haven't really defined it. At this point I'd like us all to spend 2 minutes coming up with our own definitions of "open data" and then we'll see what we can come up with together.
    - Q & A what does this mean to you? *interactive exercise

- Open Knowledge Foundation Definition: Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike. (There are reasons to not place those last two requirements on there but we'll get into that later.)

- Why is it important now? Why should I care? (researchers) (librarians) (4 mins talk, 5 mins sharing) =9
  - BIG PICTURE: why should information be freely accessible
  - The concept of open data is not new but it has quickly risen in importance and popularity in the last 5-10 years in large part, I think, because of funder requirements to share data which started showing up at the end of the 1990s. Funders wanted to see an increase in the return on their investments and sharing data for others to reuse was a good way to go about that.
  - There are other reasons too. We've seen plenty of examples lately of scientific studies being retracted due to scientific fraud or honest errors in data analysis. These could have been avoided if more people had a chance to look at the data, review it, discuss it, etc
  - Example of honest error that could have been avoided with open data: Reinhart & Rogoff economics data.
  - Data success & horror stories (share a few) *audience shares too
    - Horror:
      - Reinhart & Rogoff: https://www.washingtonpost.com/news/wonk/wp/2013/04/16/is-the-best-evidence-for-austerity-based-on-an-excel-spreadsheet-error/
      - An error in their Excel spreadsheet made a difference in calculations from -0.1% to 2.2%
      - NASA recording over moon landing
    - Success:
      - There has also been research into impact of open data vs closed.
      - Innovation in open vs not open: https://www.techdirt.com/articles/20130403/09501122561/public-domain-human-genome-project-generated-more-research-more-commercial-activity-than-proprietary-competitor.shtml
      - Two sources of genome sequence data: one open, one patented. Research and innovations from open data set were twice that from the closed data set.

- Challenges & Responses ** EVERYONE (2 mins talking, 5 mins small group role play) = 7
  - Encouraging others to open data isn't without its challenges. And we're going to dig into that with this next exercise. This handout has some common challenges we hear regarding open data. It also includes some possible responses to those challenges.
  - Handout review
  - Role playing exercise

- Why Open & Why Mozilla, what other resources are available (light) (2 mins)
  -
  -

- Q&A and discussion (5 mins)
- Resources:
    - http://www.govtech.com/data/Got-Data-Make-it-Open-Data-with-These-Tips.html

- Disciplinary concerns (add later)