Finding the Line of Best Fit:

A Case Study in Cloud Transfer with the Sanborn Maps Collection
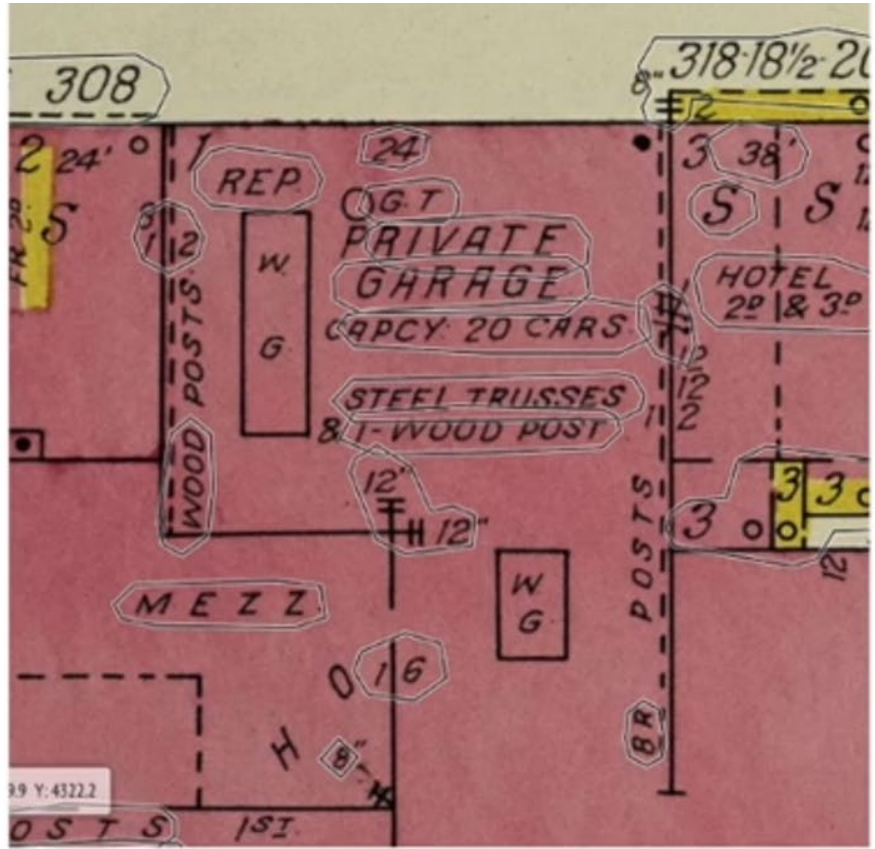
Jaime Mears, LC Labs

Alice Goldfarb, LC Labs

Rachel Trent, LC Geography and Map Division

Sanborn Maps + Machine Reading Maps

Yorkshire CLXXIV.6 Revised: 1907, Published: 1909. National Library of Scotland
Processed with mapKurator

Sanborn Fire Insurance Map from New Orleans, Orleans Parish, Louisiana.
Processed with mapKurator

How do you get 64 TBs of maps (A) at the LC to a research team in Minnesota and the UK (B) using cloud infrastructure while conforming to agency security requirements and a researcher grant deadline??

A

B

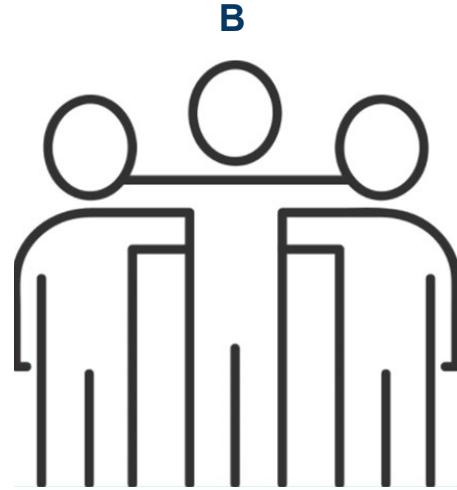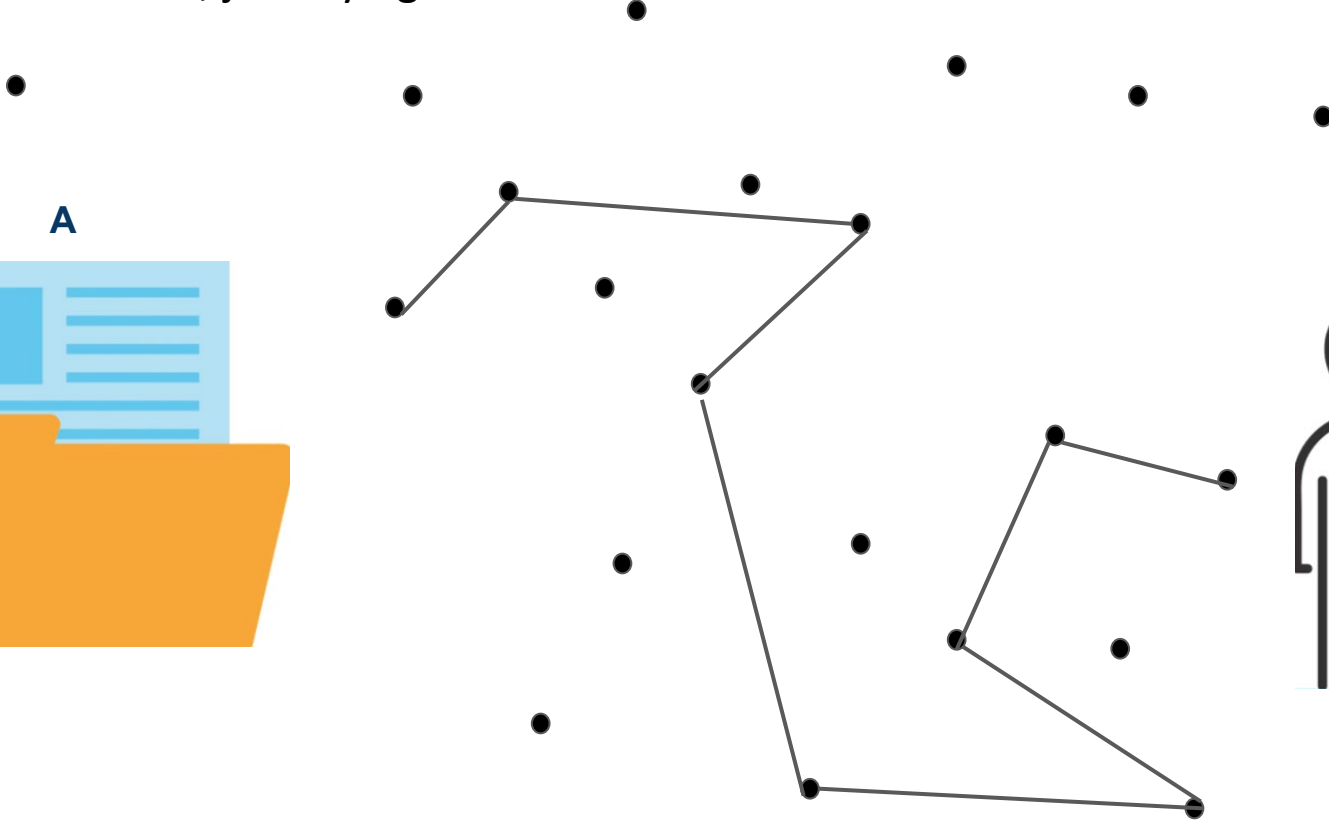You'd think it'd be something like this, a straight line from A to B.

A

B

But in our experience, it was more like this. Testing multiple option with shifting requirements, just trying to find a line of best fit.
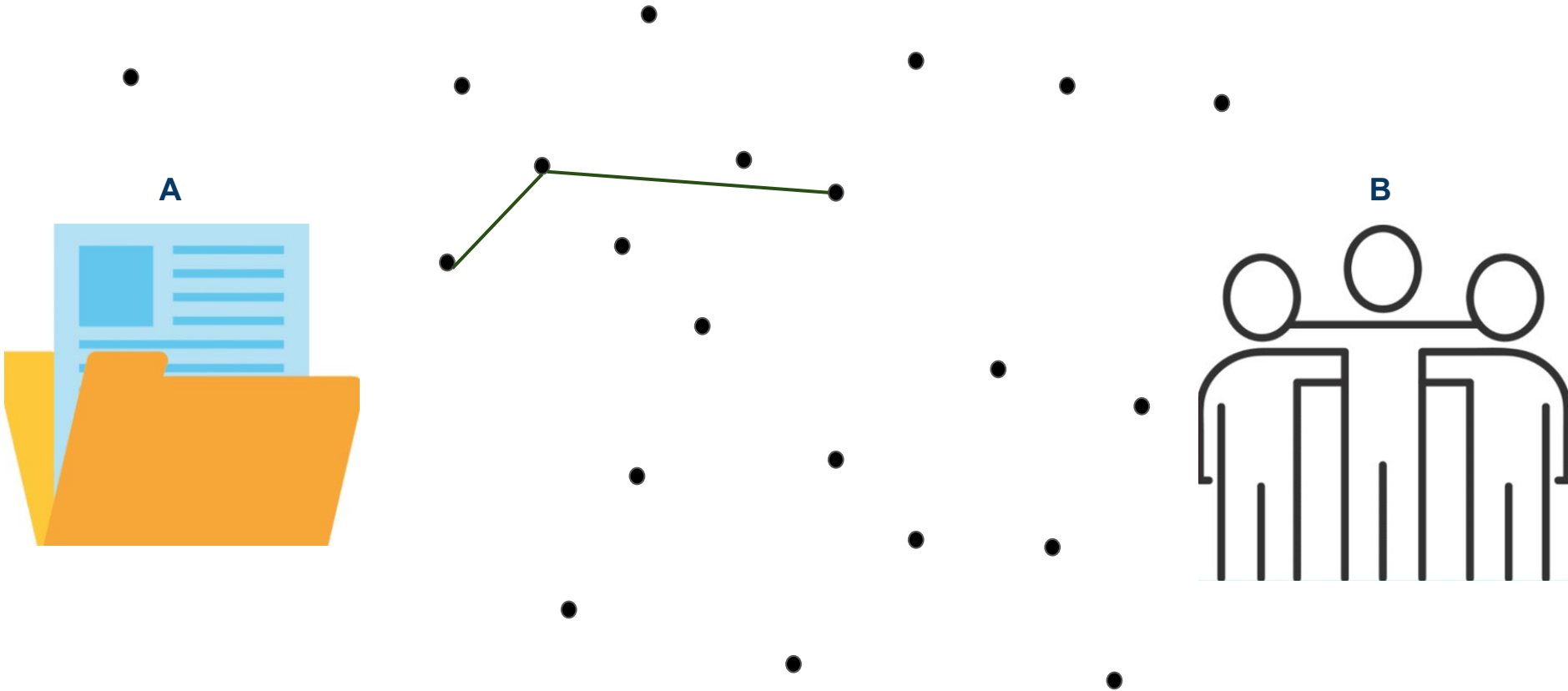
A

B

Cloud infrastructure presents seemingly a universe of possibilities for hosting and transferring data, and each of these possibilities presents another variable on the journey.
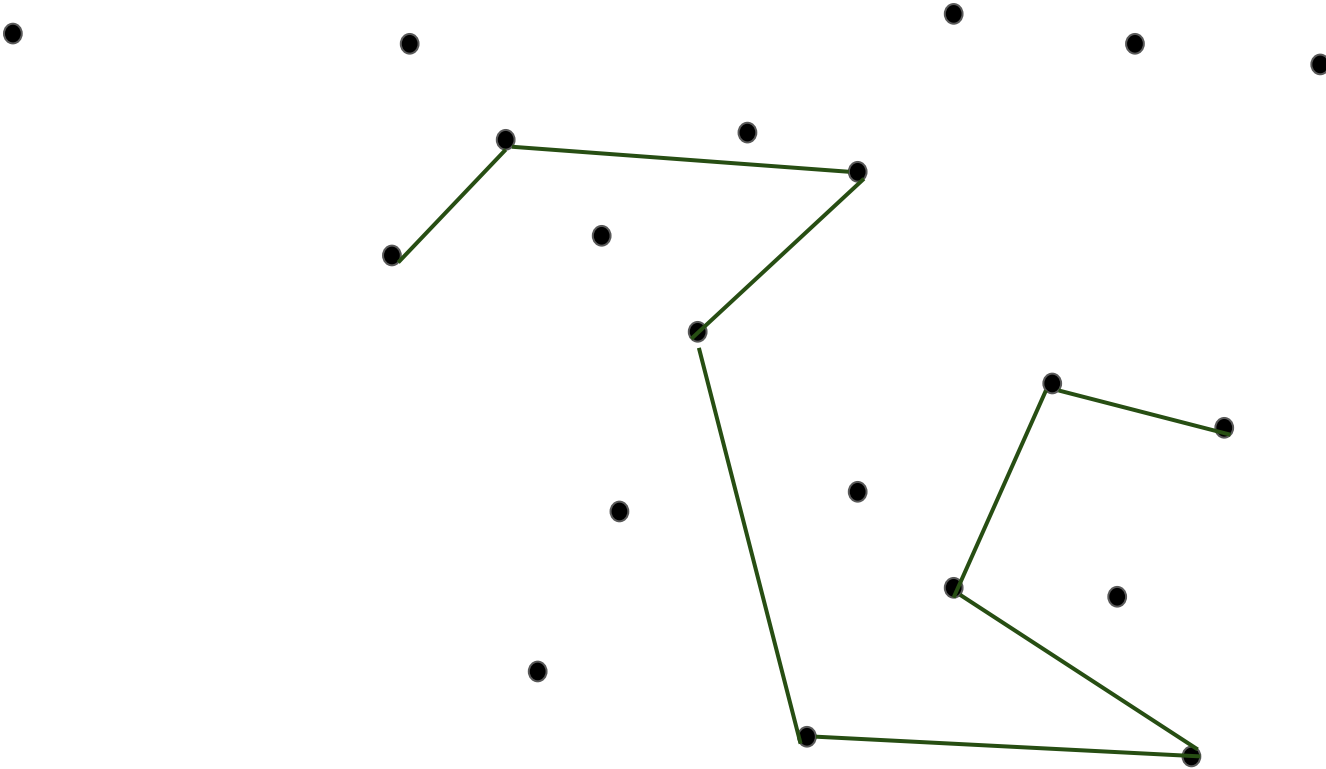
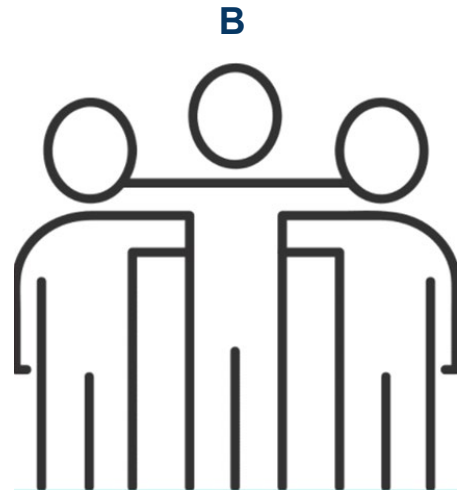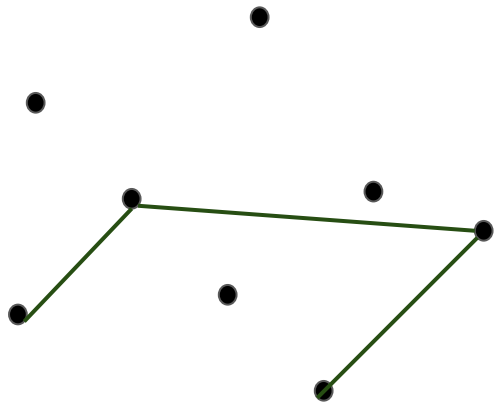And so our story begins - with a cooperative agreement!

A

B
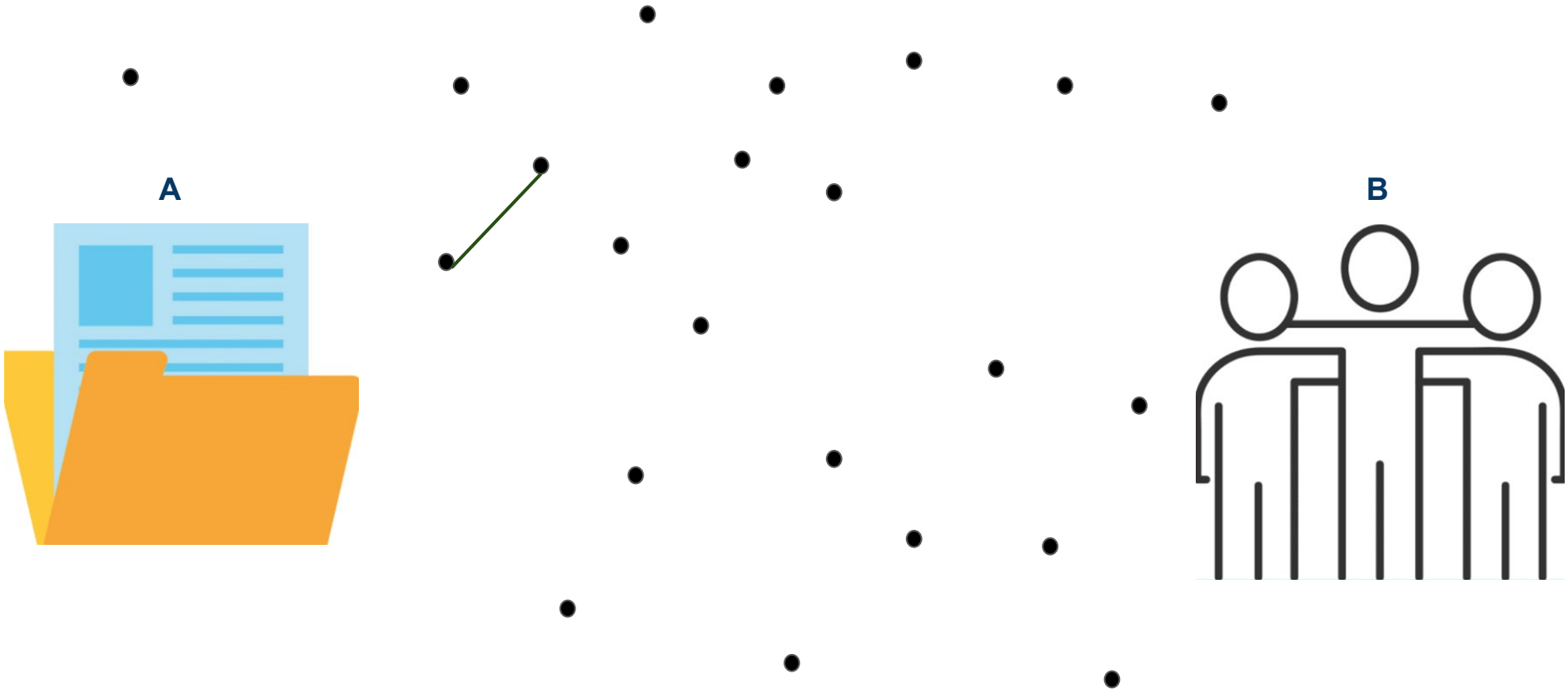
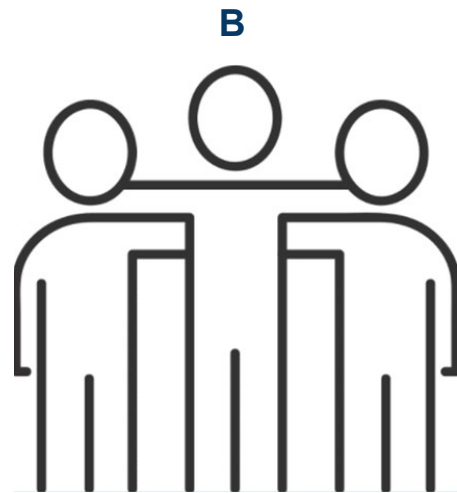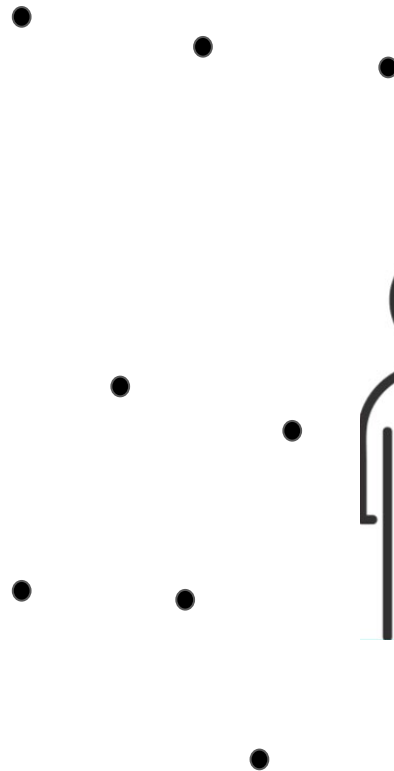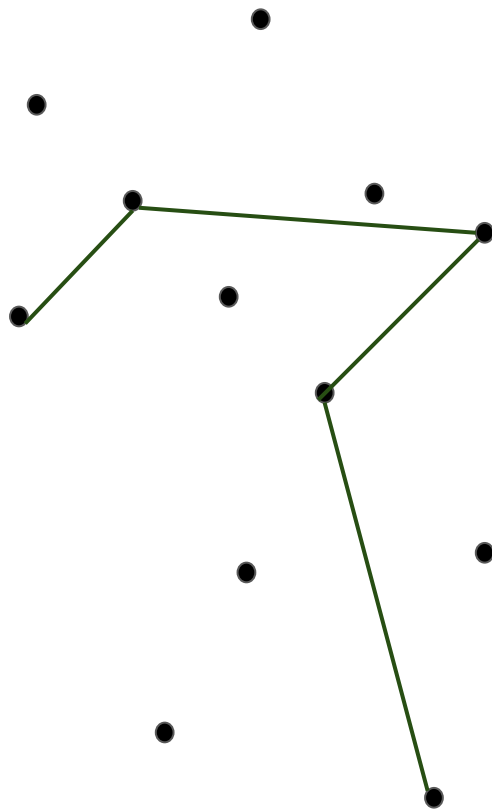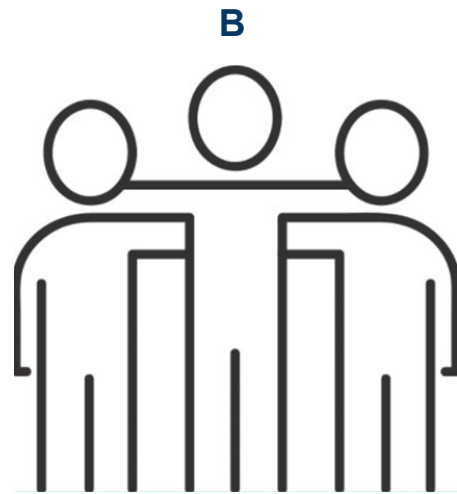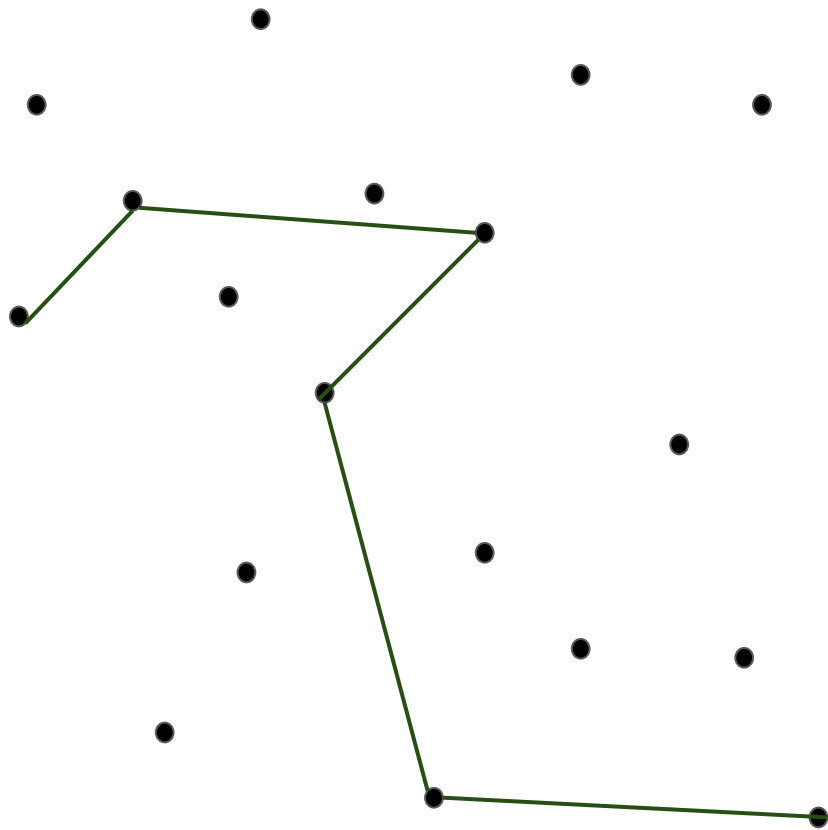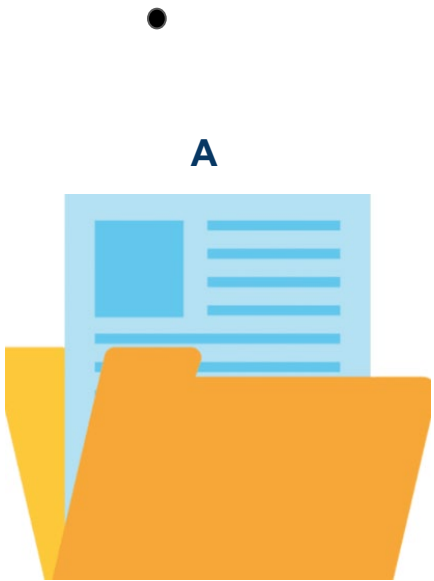Then we learned more, and Azure was only available for a sample of dataset.

A

B

Registry of Open Data on AWS?

A

B

Public S3 Bucket?
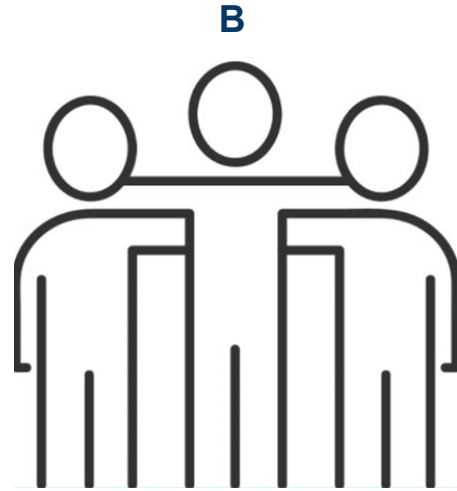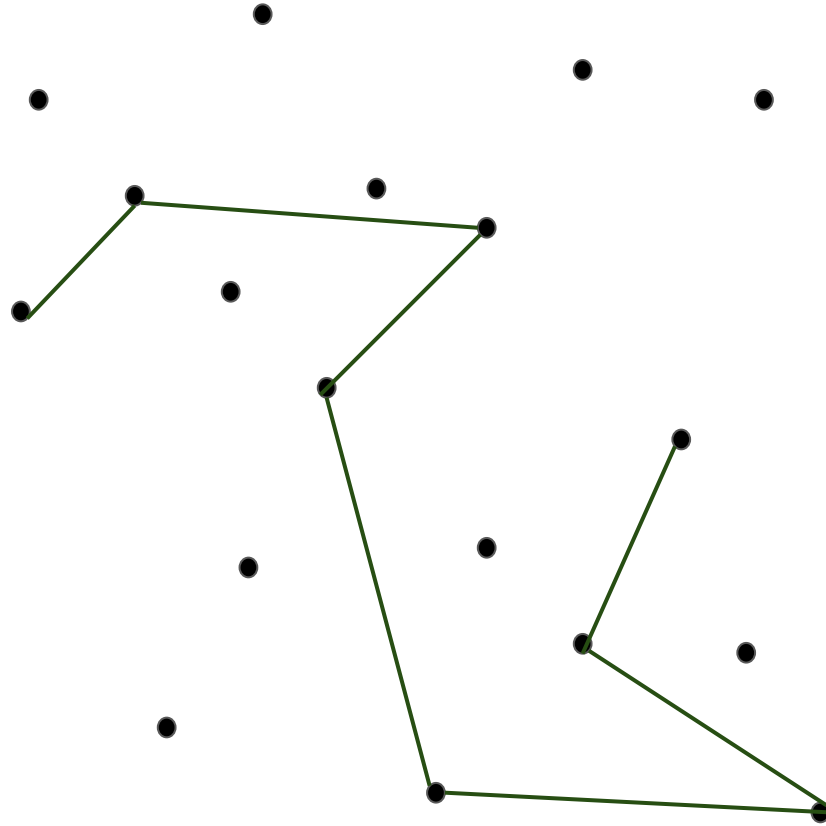
A

B

Media Shuttle?

A

B

Private S3 Bucket?
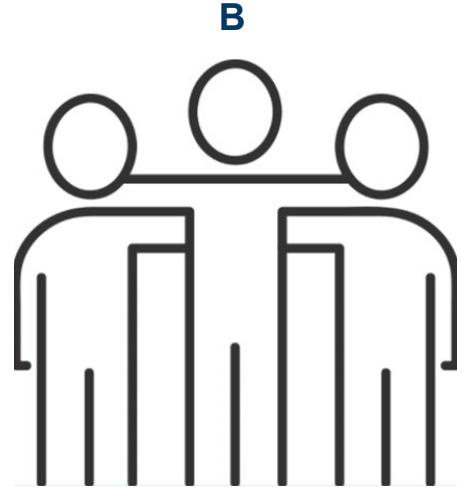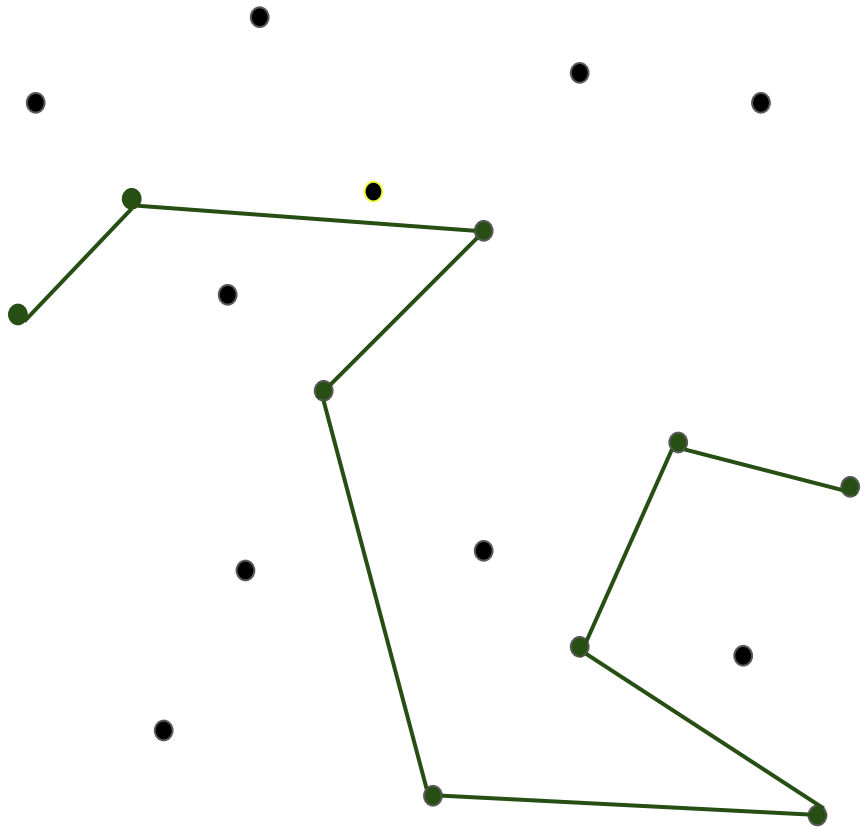
A

B
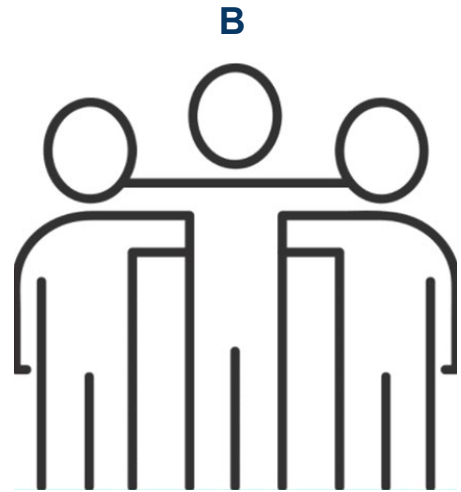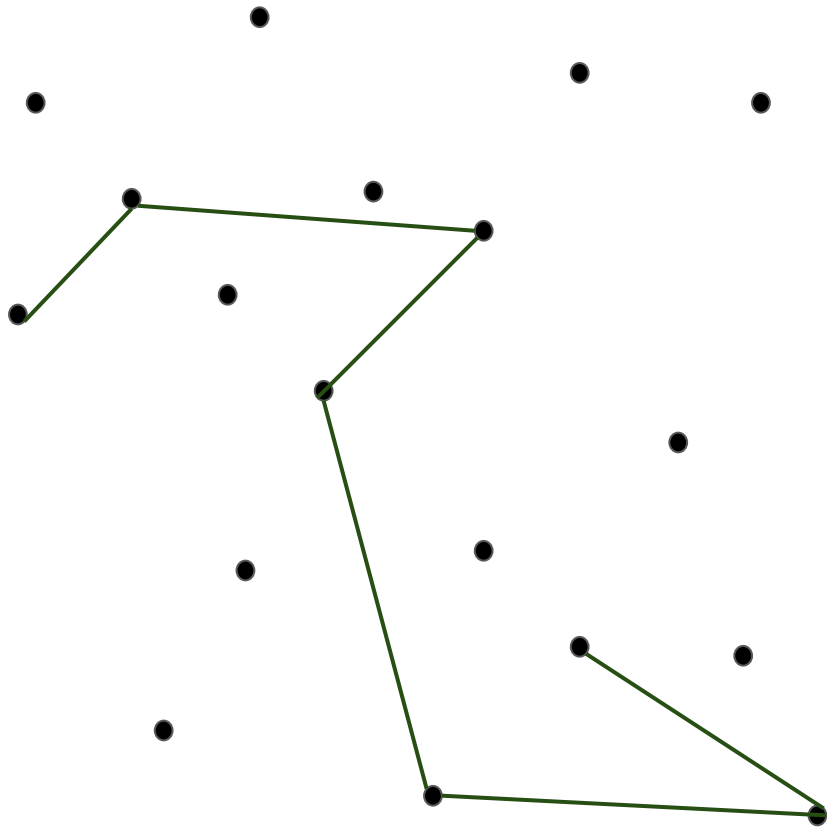
Researcher ability to update bucket policy?

A

B

S3 push to researcher bucket! Success!!!

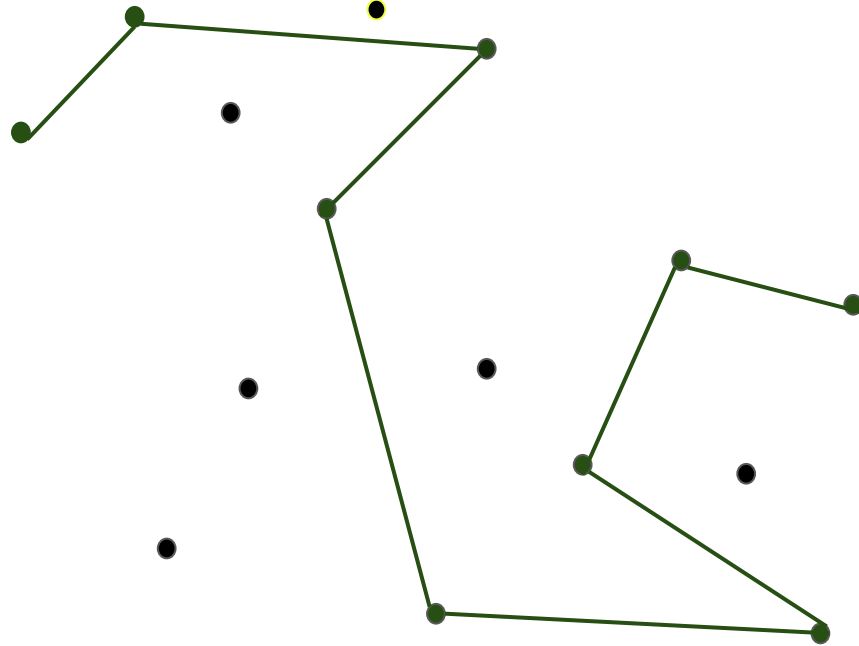# Metadata parallel process

Prepare pick list for
staff to gather images

Metadata
requirements
gathering

Data
dictionary

A

B

Metadata parallel process continued

Prepare pick list for staff to gather images
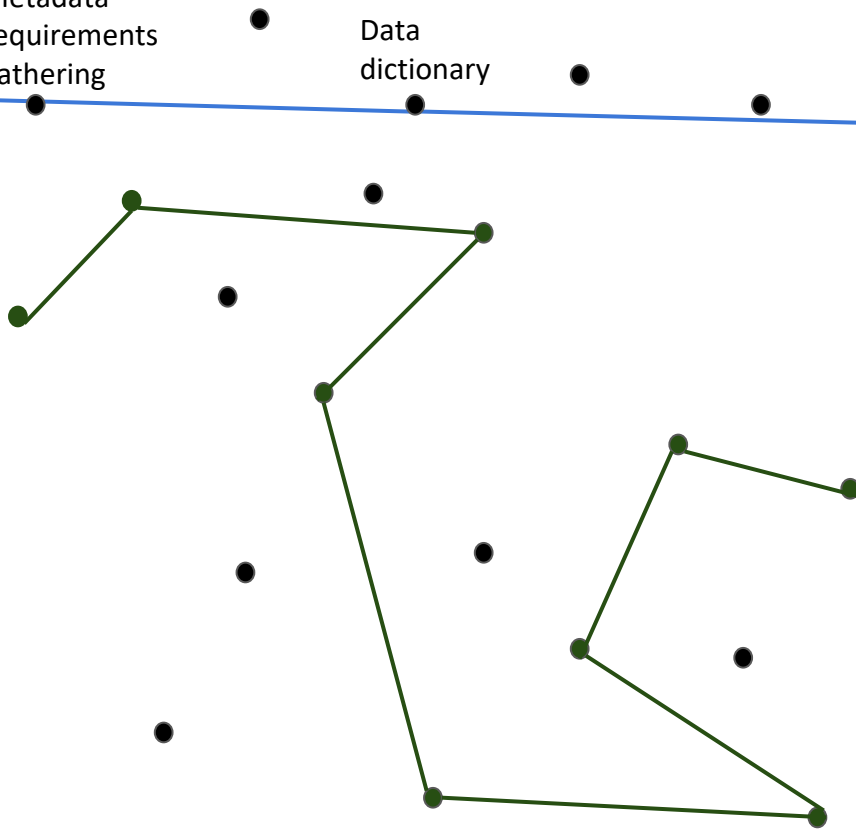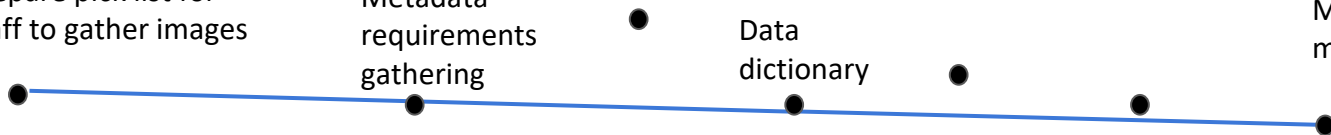
Metadata requirements gathering

Data dictionary

Merge with transfer manifests

A

B

Storage and Computing Environment

File Format and Resolution

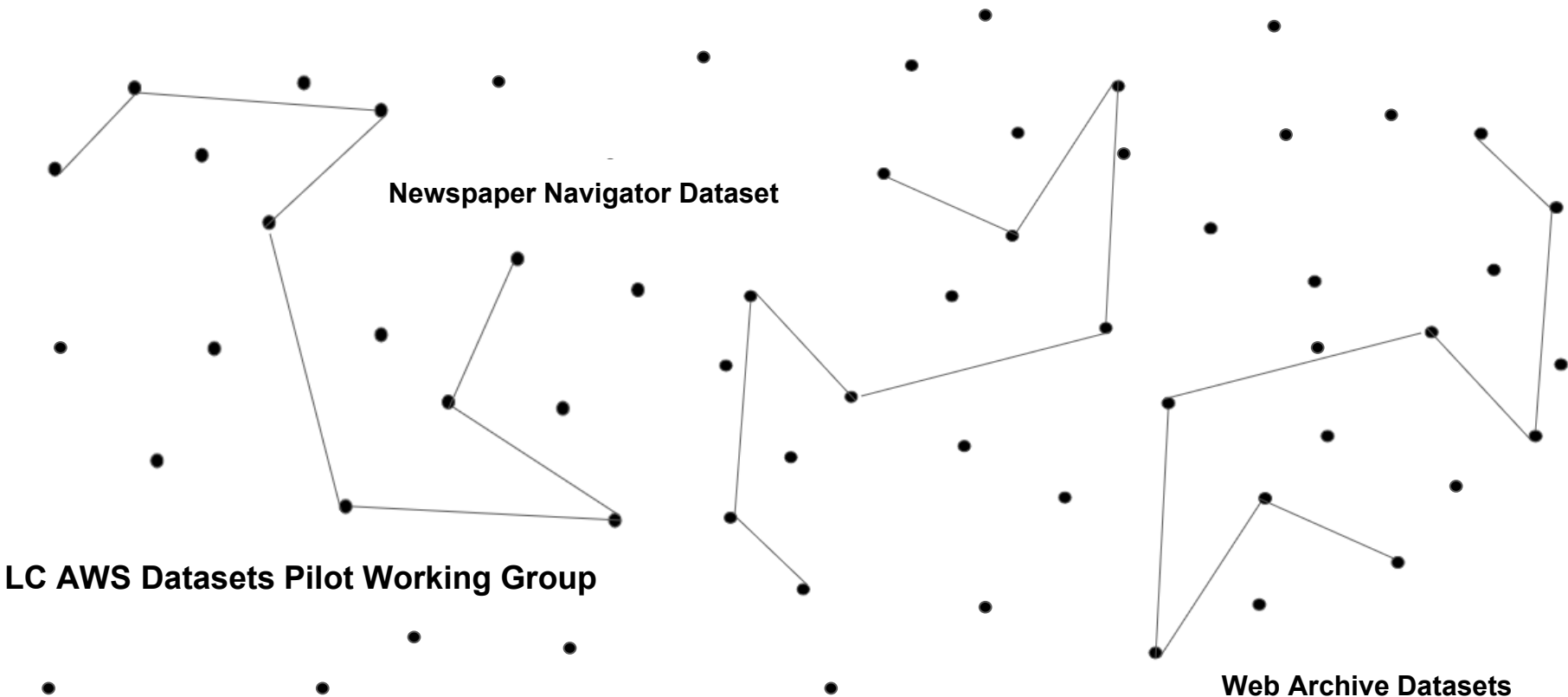IAM Role Authority

Metadata and Provenance

Deadline

Team Members and Roles

Beyond the Case Study

Computing Cultural Heritage in the Cloud Grant

Newspaper Navigator Dataset

LC AWS Datasets Pilot Working Group

Web Archive Datasets

Thank you to the LC team that made this happen!

Carl Watts, IDD
Stephanie Barb, IDD
Chris Adams, ITDD
Jaime Mears, LC Labs
Alice Goldfarb, LC Labs
Rachel Trent, LC Geography and Map Division
Meagan Snow, LC Geography and Map Division
Paulette Hasier, LC Geography and Map Division
Nicole Saylor, LC Labs