Housekeeping

- This event is being recorded; Event will be available On Demand after the event at the main training page

  https://clu-in.org/conf/itrc/EDM-1/

- If you have technical difficulties, please use the Q&A Pod to request technical support
- Need confirmation of your participation today?
  - Fill out the online feedback form and check box for confirmation email and certificate

Advancing Environmental Solutions

Environmental Data Management Best Practices (EDM-1, 2022)

Data Exchange

Sponsored by: Interstate Technology and Regulatory Council (www.itrcweb.org)
Hosted by: US EPA Clean Up Information Network (www.clu-in.org)

ECOS ERIS ENVIRONMENTAL RESEARCH INSTITUTE OF THE STATES

ITRC – Shaping the Future of Regulatory Acceptance

- Host Organization
- Network - All 50 states, PR, DC
- Federal Partners
  - DOE
  - DOD
  - EPA
- ITRC Industry Affiliates Program
- Academia
- Community Stakeholders

- Disclaimer
  - https://edm-1.itrcweb.org/about-itrc/#disclaimer
- Partially funded by the US government
  - ITRC nor US government warranty material
  - ITRC nor US government endorse specific products
- ITRC materials available for your use – see usage policy

itrcweb.org
facebook.com/itrcweb
@ITRCWEB
linkedin.com/company/itrc

EDM Trainers

Olga Stewart
Geosyntec Consultants
ostewart@geosyntec.com

Tori Ward
Woodard & Curran, Inc.
vward@woodardcurran.com

Evan Anway
ERM, Inc.
Evan.Anway@ERM.com

Dreas Nielsen
Integral Consulting Inc.
dnielsen@integral-corp.com

Michael Ginsbach
Minnesota Pollution Control Agency
michael.ginsbach@state.mn.us

Dan McCarthy
EarthSoft
dan.mccarthy@earthsoft.com

EDM Training Catalog

Overview Video

ROUND TABLE #1

(Archived)
Data Management Planning and Field Data Collection

ROUND TABLE #2

April 6, 2023
Data Exchange

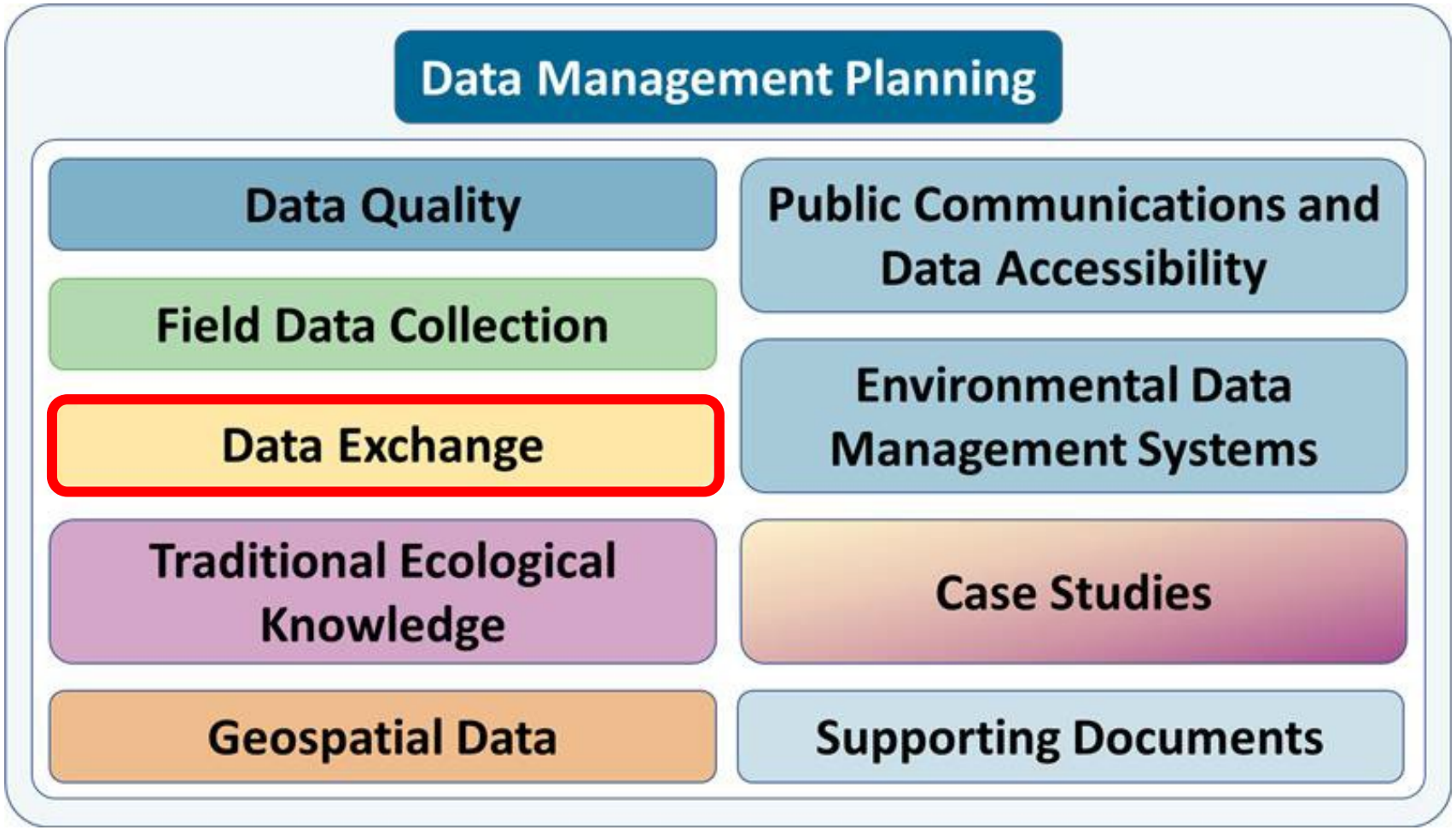
ROUND TABLE #3

May 2, 2023
Data Quality

Discussion

Traditional Ecological Knowledge

Environmental Data Management Best Practices



https://edm-1.itrcweb.org/

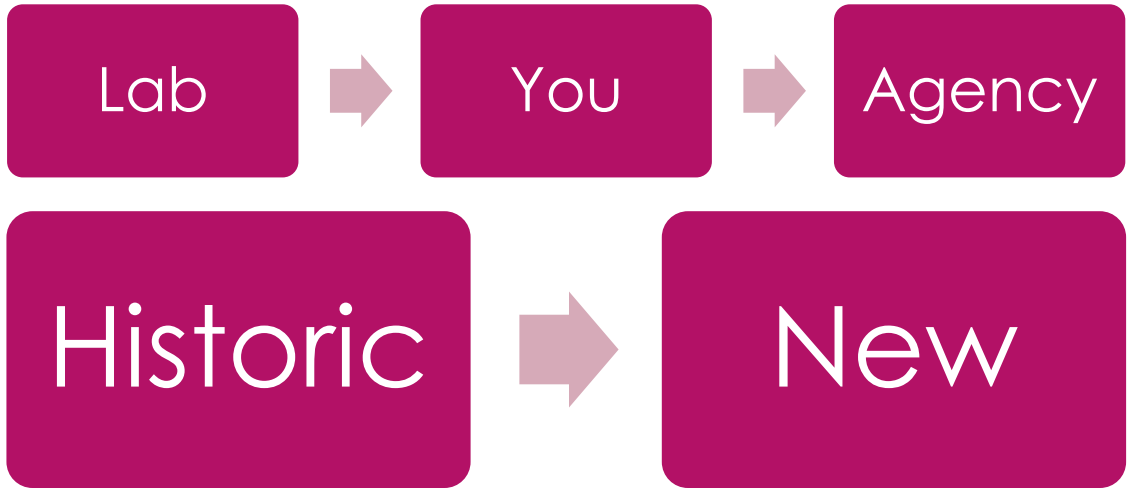
Questions

Please use the Q&A Pod to ask questions for the Expert Panelists.

Question Index

What is Data Exchange?

- The process of transferring structured data between parties.



What is Data Exchange?

Data Exchange	Data Migration
Existing data export/import process	NO existing data export/import process
EDD established	NO EDD established
Planned in advance	Conducted as needed
Occur with regular frequency	One-time event
Active data source and target	Data source usually isn't active

Table 3-1 from Data Exchange Overview Fact Sheet

Receiving laboratory data

- Establish EDD format and valid values to be used
- Understand any structural differences between source and target systems
- Document system specifications and transformations or translations required.

Receiving unstructured data

- Understand and document
  - Structural differences
  - Transformations and translations
- Cost and time constraints may require prioritizing data extraction

Negotiating unclear data request

- Cooperate to clearly define requirements
  - Content
  - Format
- Decide who should reformat data

Should I migrate this data?

Cost
Acquisition
Integration
Amount
Complexity
Tools / Tech

Benefits
Usability
Quality
Need

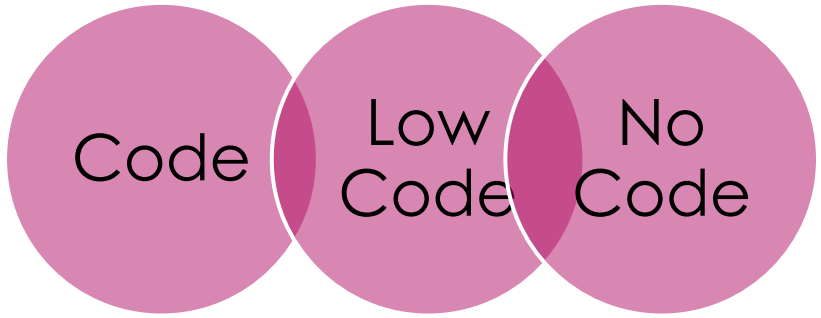
- Involve data users and project managers in discussions
- Data migration does not automatically improve the quality of data

Starting Migration

- Audit or inventory the data
  - See ICEDM white paper: http://www.icedm.net/s/ICEDM-Historical-Data-Migration-Audit-Final.pdf
- Develop a strategy and assign priorities (e.g., based on data uses)
  - Beware: incomplete data can be incorrect data
  - See Data Quality: Best Practices for Environmental Data Quality – EDM (itrcweb.org)

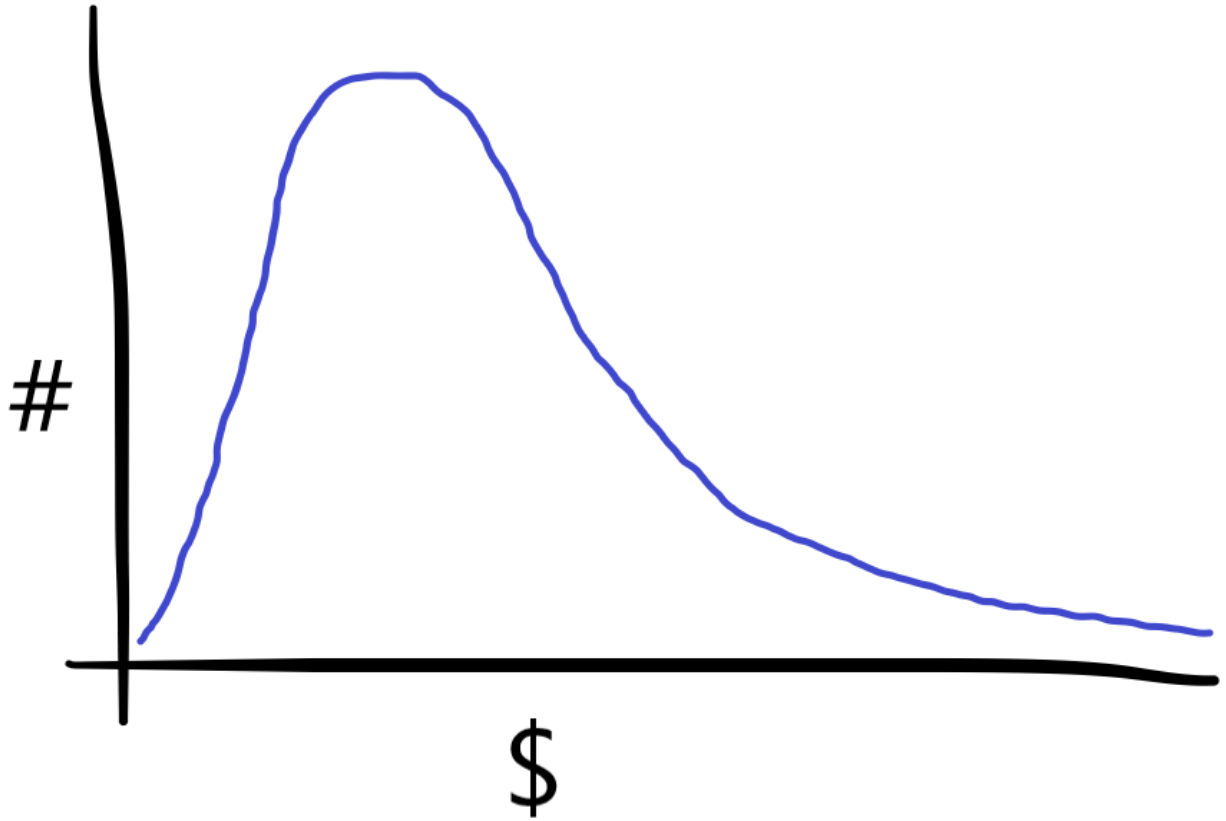
When to Automate

- QA/QC
  - EDD before upload
  - Summary tables/figures of migrated data
- Document decisions, assumptions, and QC
- Automation
  - Reduced cost, increased efficiency
  - Documentation
  - Increased reproducibility
- Software/tools vary by organization

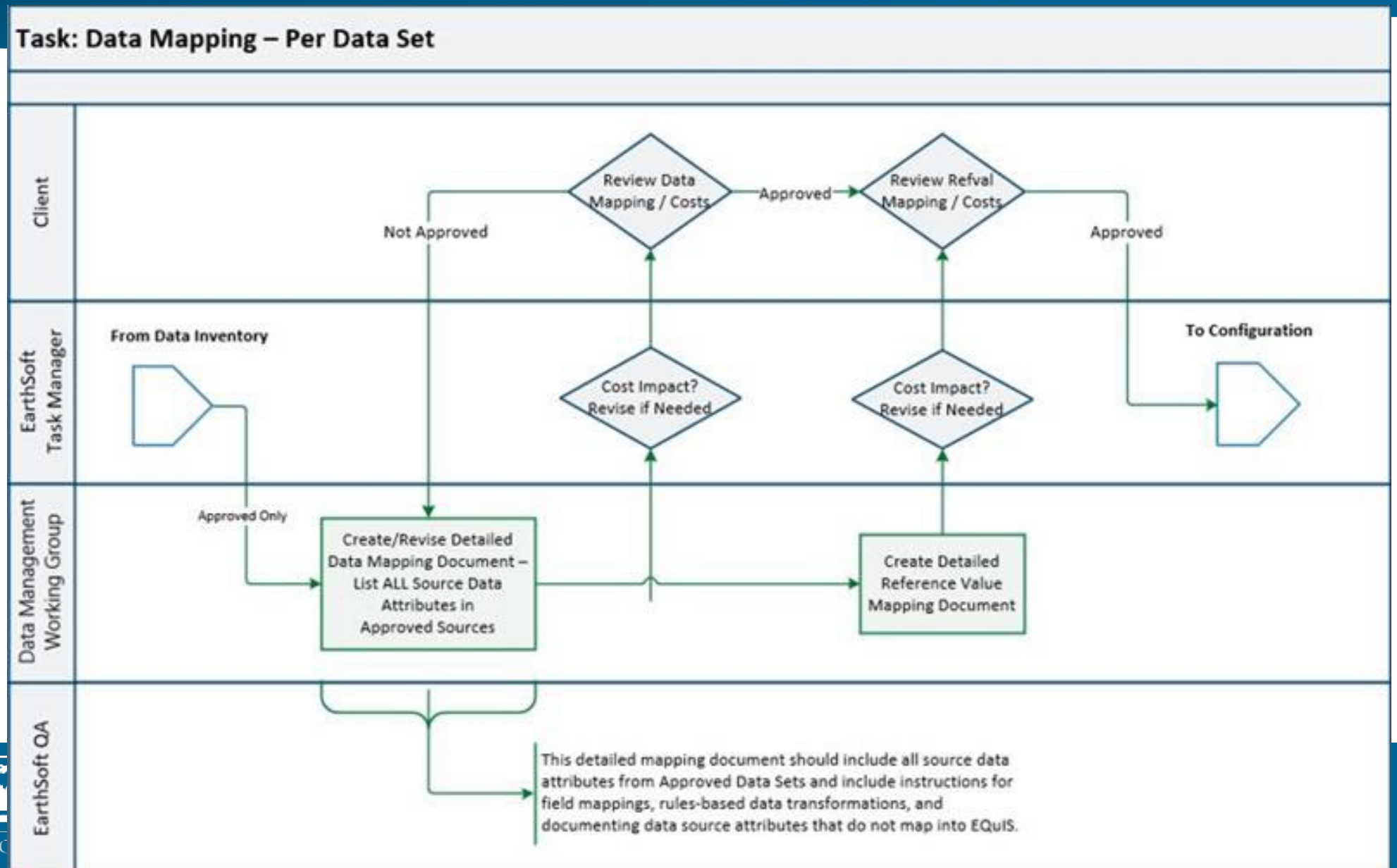


Migration Cost Containment

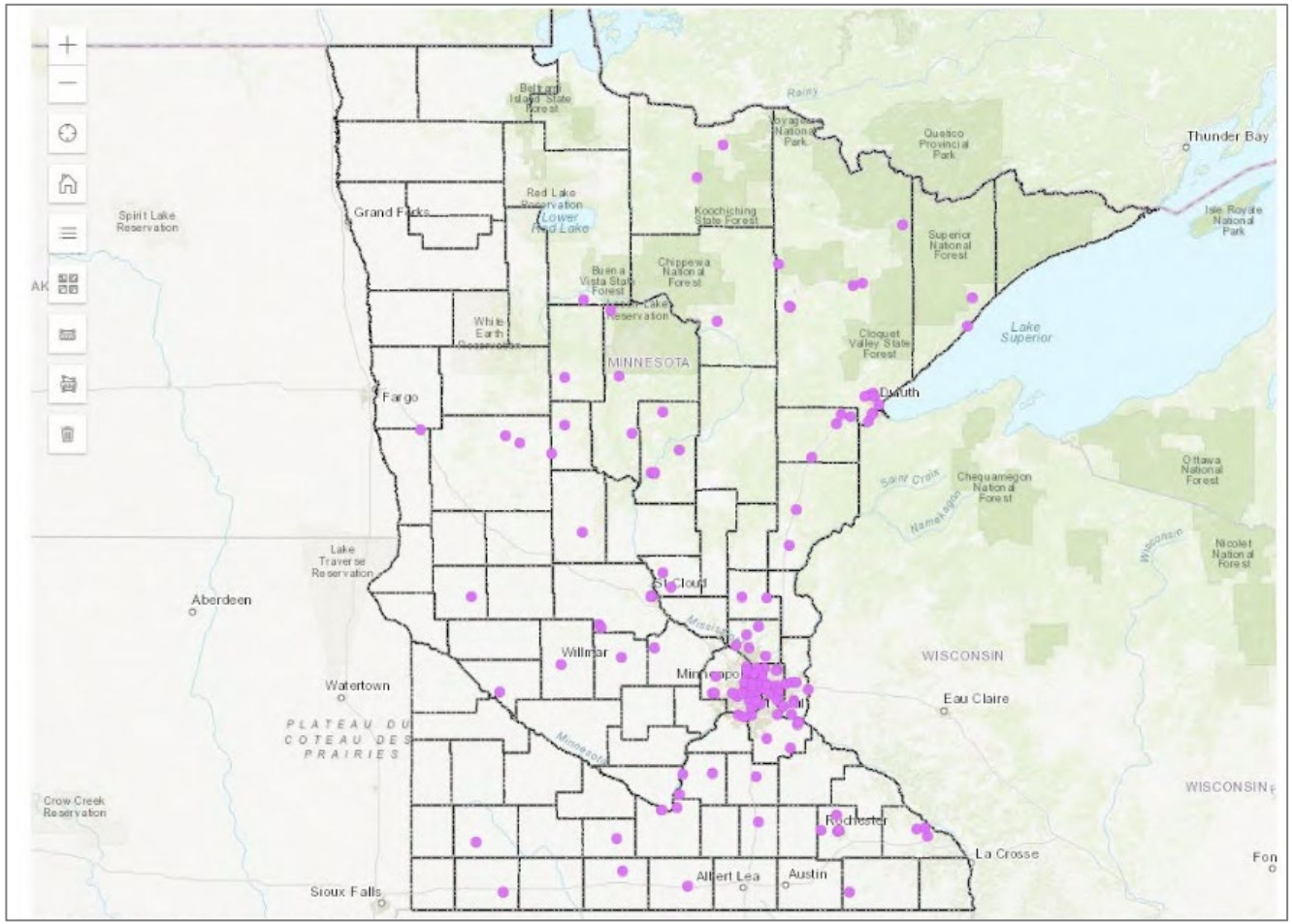
- Migration costs
  - Highly variable w/ long right tail
  - Hard to predict in advance
  - Automate as much as possible
  - Don't muddle costs of correcting historic errors with costs of migration
- Apply traditional project management principles
  - Prioritize order
  - Track progress



Migration Cost Containment



MPCA Historical Data Migration Case Study: Filling Minnesota's Superfund Groundwater Data Accessibility Gap (itrcweb.org)



Links

- Case Study
- Atlas: https://www.pca.state.mn.us/data/minnesota-groundwater-contamination-atlas

Establishing valid values

- Ambiguous or duplicate values can result in problems
- Compilations are available from a variety of sources. See Data Exchange and Valid Values under Supplemental Resources – EDM (itrcweb.org)

The Valid Values fact sheet discussed these points and provides suggestions

Example of Redundant Versions Without Control of Value

Location types:

- monitoring well
- well-monitorign *(Note: misspelling)*
- well-monitoring
- observation well
- wells-monitoring

Valid value refinement

- Cascading effects
- Document and communicate any changes

Case Study: USGS Challenges with secondary use of multi-source water quality monitoring data – EDM (itrcweb.org)

Risks of Data Exchange

Risks

- Liability for incorrect data transformation
- Lack of data accessibility and lack of data security

Mitigation

- QA'd automated processes
- Review by data users and data generators
- Document data changes from migration