



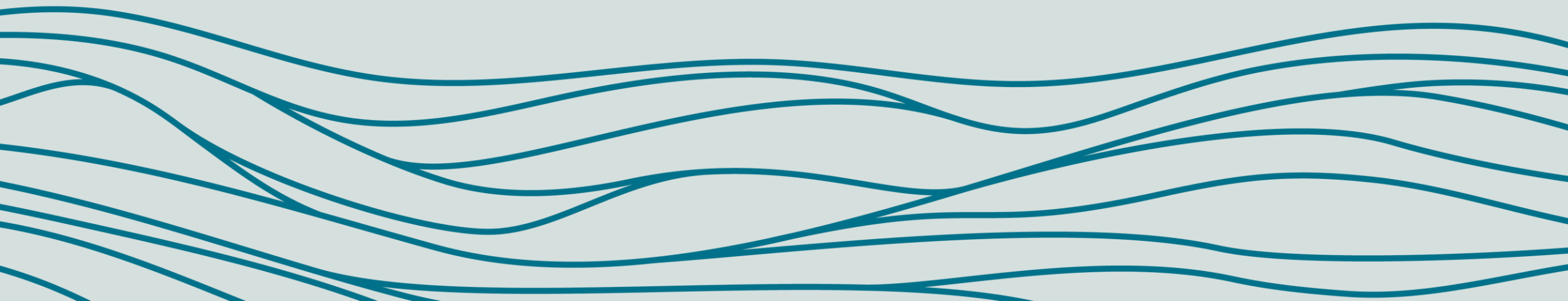
DIVERTING DATA DISASTERS

American Water Works Association – 2024 Meeting
Spokane, Washington

Presented by:

Kelsey Mach, LG, CWRE - *Project Geologist, Aspect Consulting*

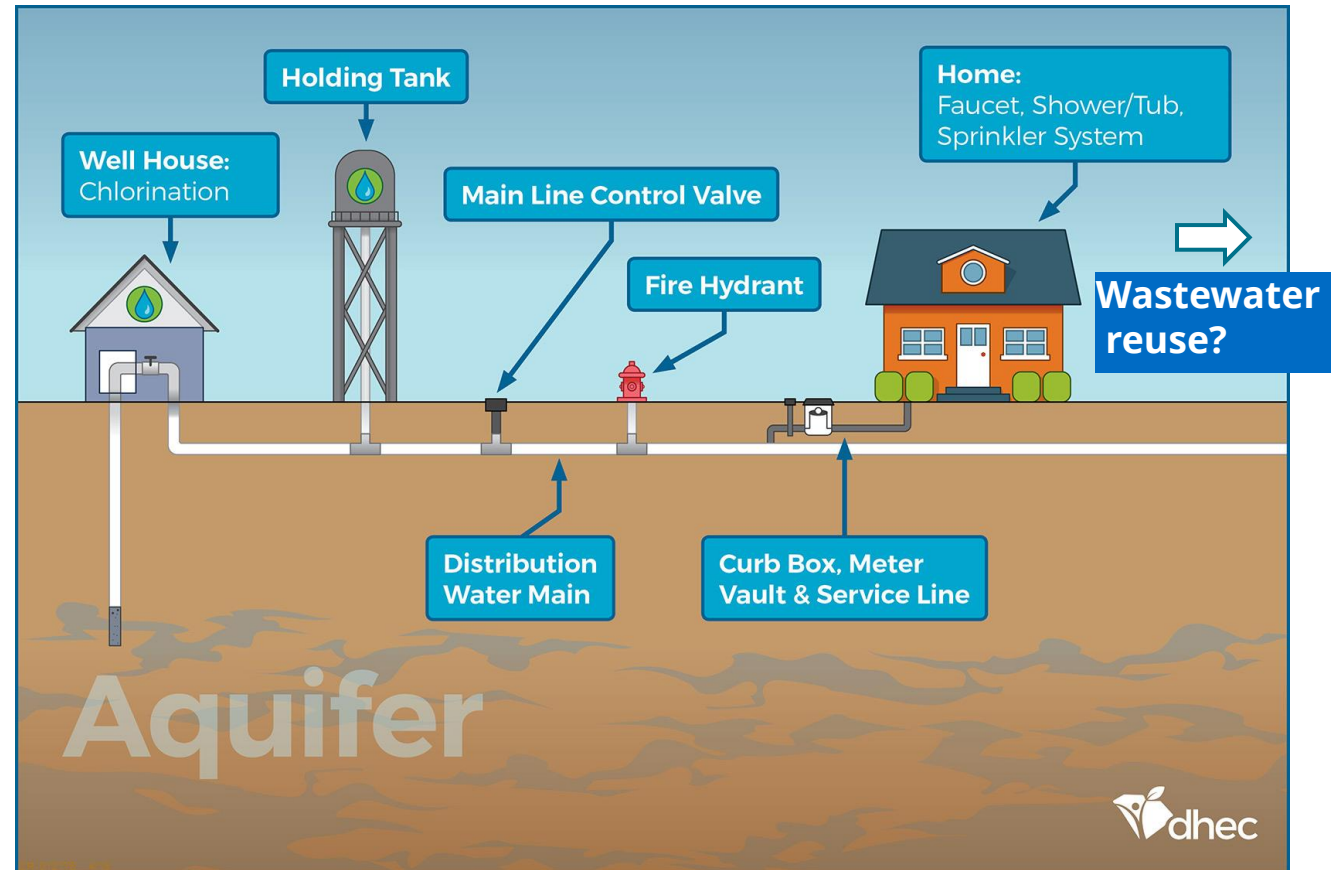
Hanna Winter – *Project Environmental Data Scientist, Aspect Consulting*



OVERVIEW

Data is critical to management of a water system.

Think of what kind of data do you use for your job on a regular basis

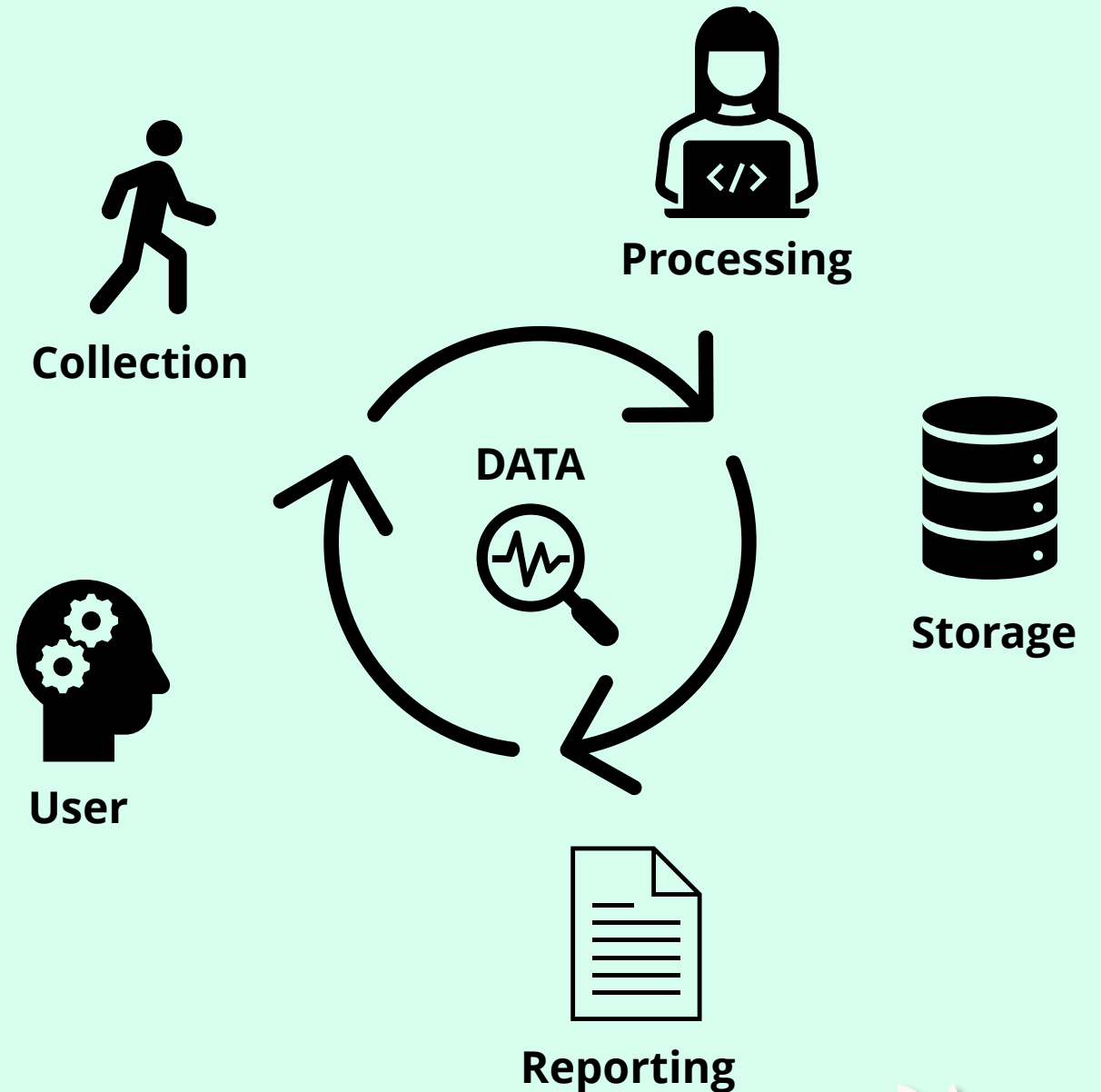


Throughout this presentation we'll go through the **data lifecycle** and discuss **common data pitfalls** and how they can be avoided.

DATA LIFECYCLE

Errors and inefficiencies can be introduced at any step of the data lifecycle.

Also opportunities for Improvement at each step



DATA TYPES

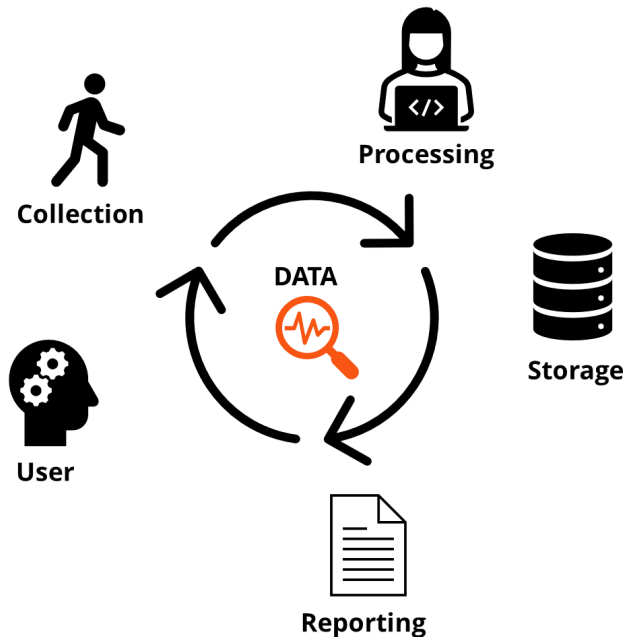


Where does the data come from?

- Production data
- Water use
- Water rights
- Water quality
- Infrastructure specifications

What format is it in?

- Tabular
- Geospatial
- Non-tabular



DATA TYPES



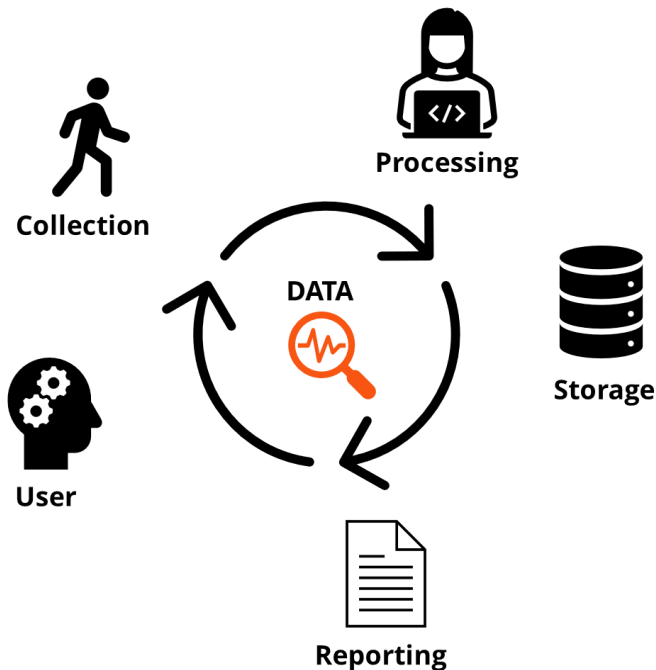
Common Pitfall: Not considering the data type/formatting

"Square peg - round hole"

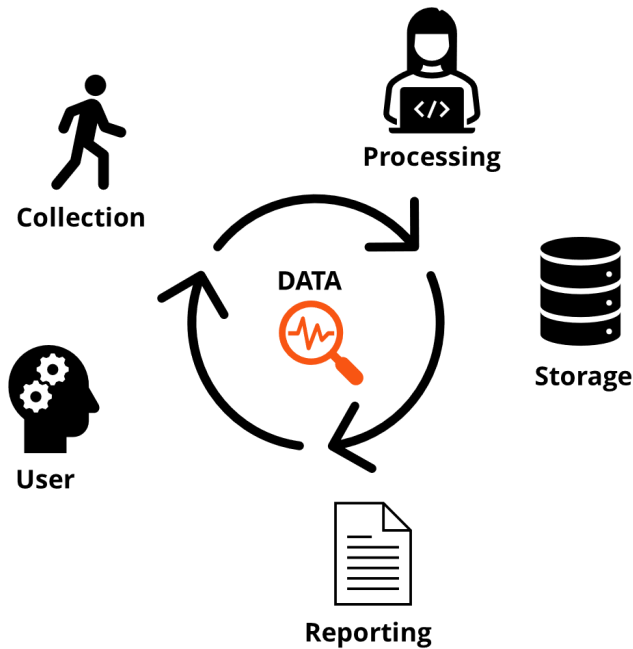
"0001234" becomes "1234"

Potassium (7440-09-7) turns into an impossible date 9/1/7440

✓ **Solution:** Create standard data formats for common uses



DATA TYPES



Common Pitfall: Not including metadata

What is metadata? Data that *explains* the data.

- ✓ **Solution:**
- ✓ **Include metadata into the data collection and documentation process**

DATA COLLECTION

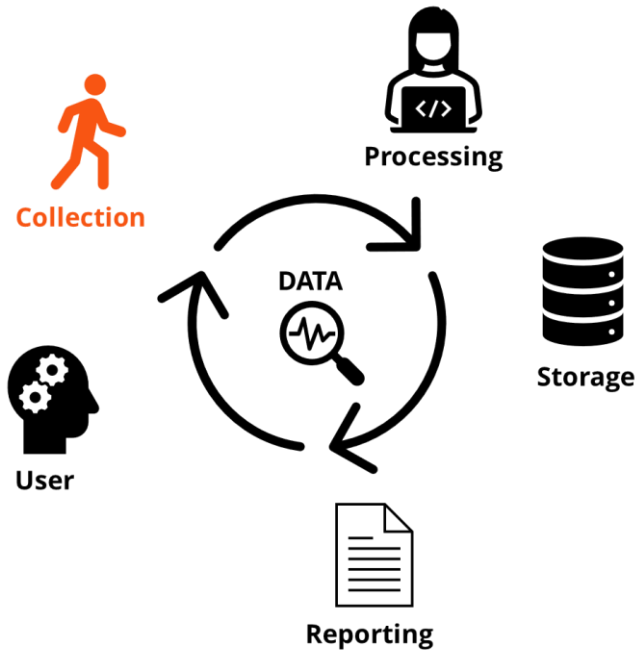


Collection methods

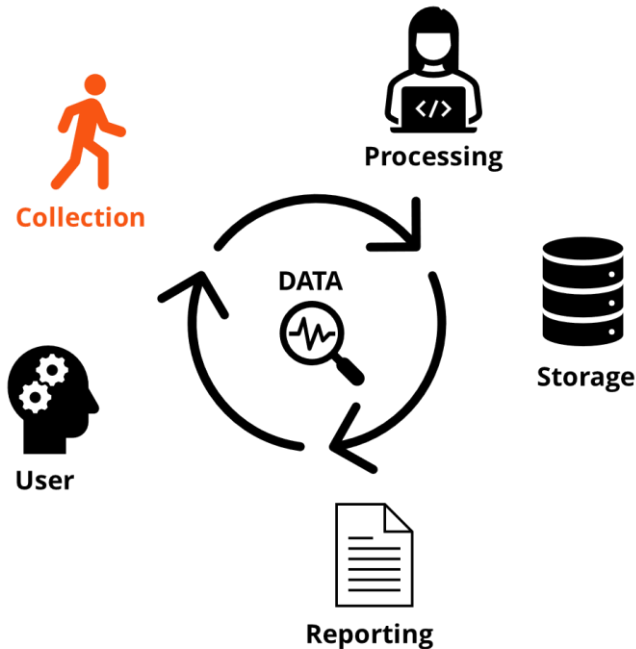
- Forms vs. automated

Collector

- Person vs. programmed



DATA COLLECTION

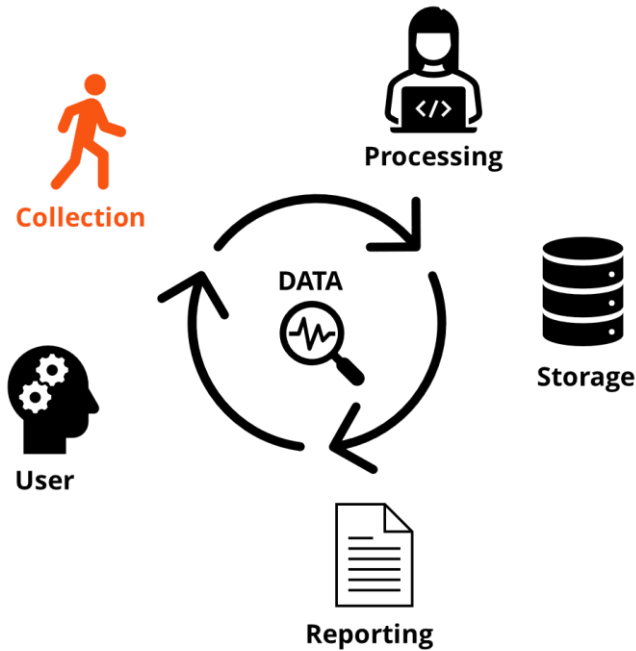


Common Pitfall: No documented procedures

Helps with consistency and understanding what the data means and how to interpret it.

- ✓ **Solution:**
- ✓ **Establish standard operating procedures "SOP" document**

DATA COLLECTION



Common Pitfall:

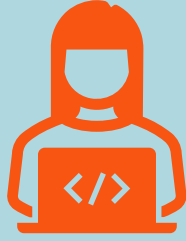
Transcription errors

- Illegible handwriting
- Date/time
- Units

✓ Solution:

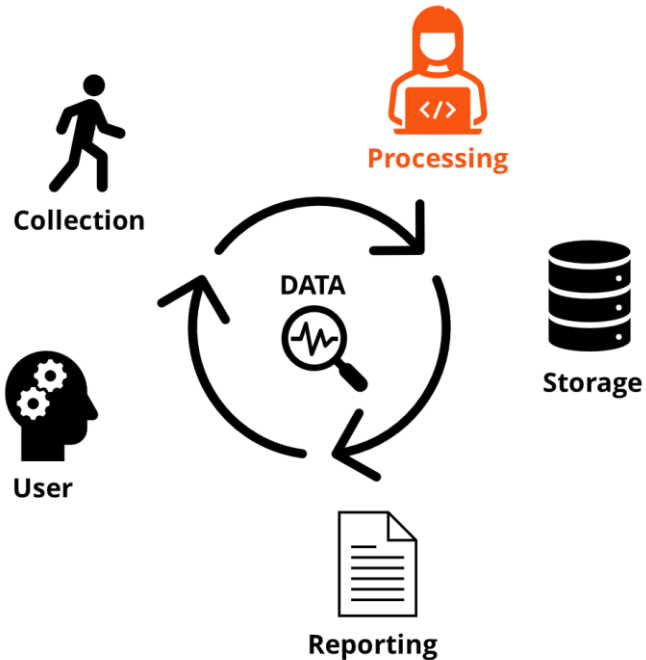
- ✓ Digital forms (Survey123)
- ✓ “Call and Response” method

DATA PROCESSING

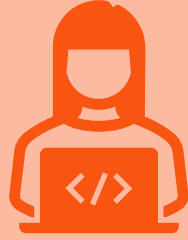


Example of data processing steps:

- Transcription
- Formatting or reformatting
- Combining or relating with other data
- Data validation and QA/QC
- Calculations and conversions



DATA PROCESSING



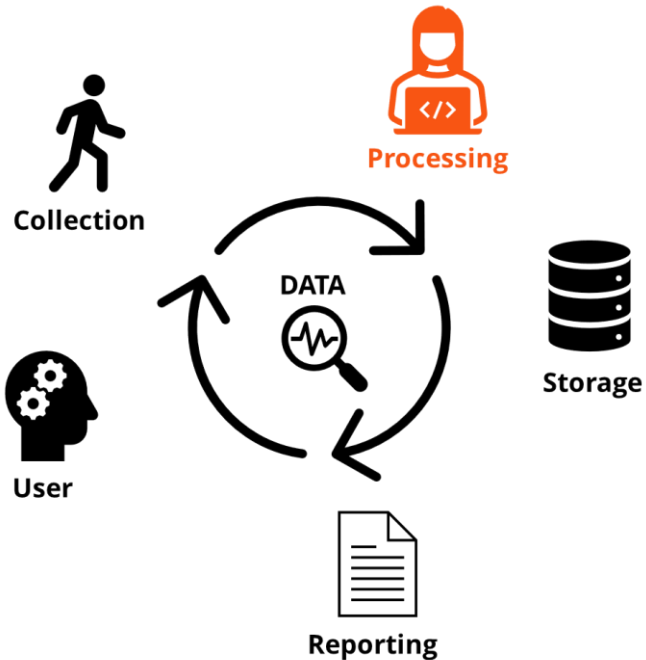
Common Pitfall:

Skipping or minimizing the processing step

- Losing information
- Including outliers

✓ Solutions:

- ✓ QA/QC
- ✓ Data SOP

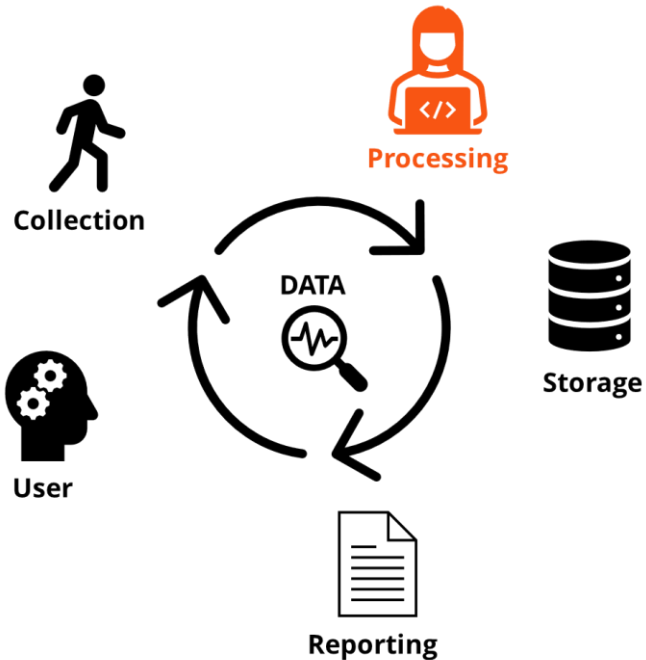


DATA PROCESSING

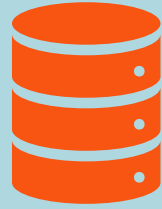


Common Pitfall:
Not tracking data modifications

- ✓ **Solutions:**
- ✓ **Automating data processing and QA/QC procedures**
- ✓ **Data SOP**



DATA STORAGE



Simple to complex –

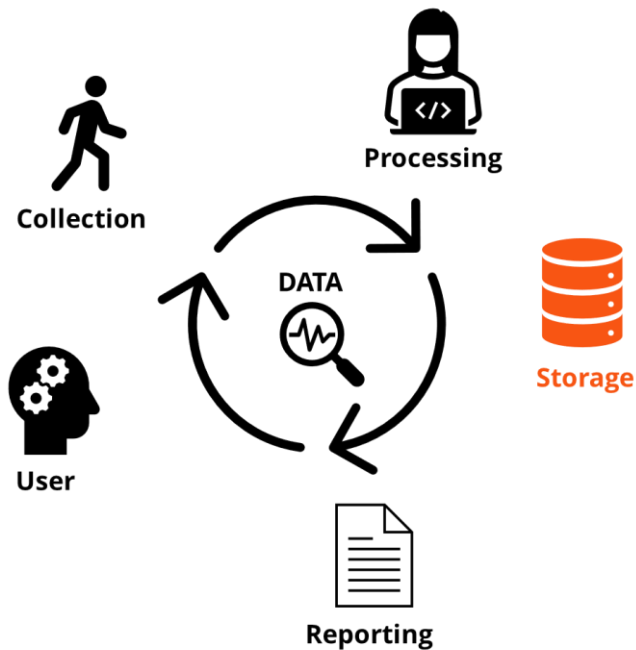
Notebooks, excel spreadsheets, databases

Data and user needs

Security, access, size, speed, quality control, end-use, flexibility and likeliness of change

Databases

Off the shelf options (EQulS) vs. custom (Access, SQL)



DATA STORAGE



Common Pitfalls:

Version control (which is the “final-final”)
Not building in editing restrictions

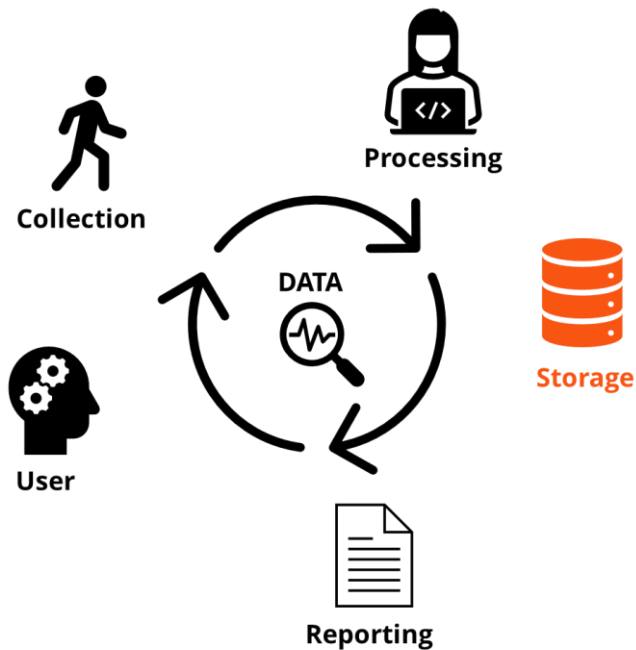
✓ **Solution:** Use of databases vs. spreadsheets

Not updating in accordance with needs

✓ **Solution:** Periodically re-assess data storage needs

Staff turnover

✓ **Solution:** Establish Data SOP and utilize more user-friendly storage tools



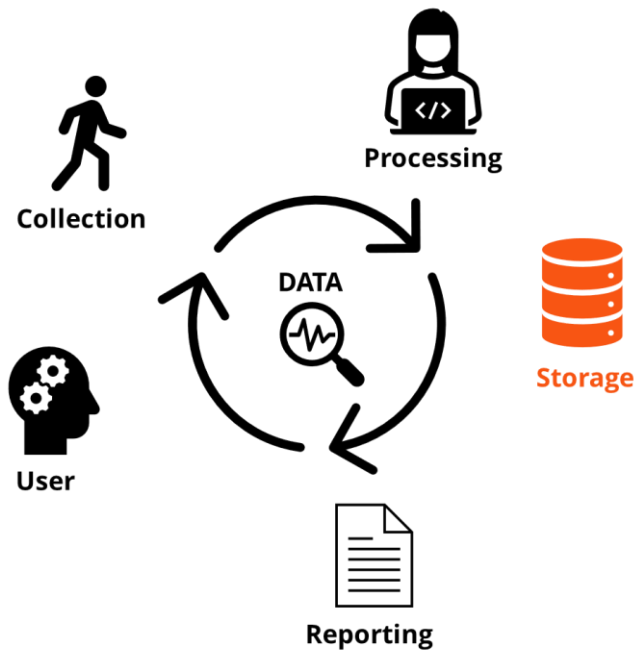
DATA STORAGE



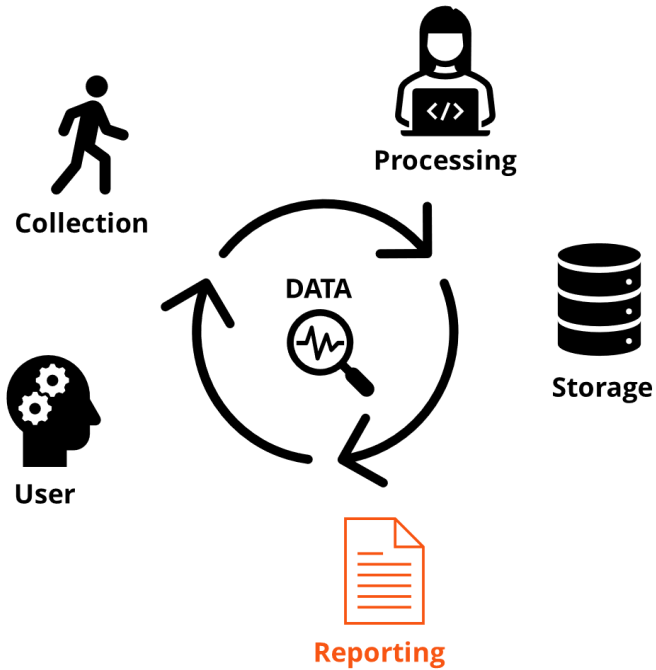
Common Pitfall: Trying to have “one solution fits all”

All storage types have limitations

- ✓ **Solution:**
- ✓ **multiple data storage types**
- ✓ **custom database**



DATA REPORTING

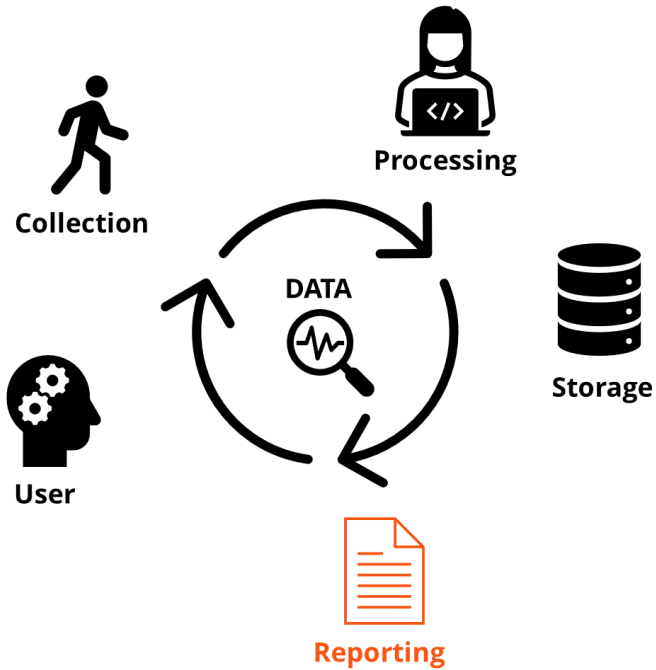


When is a report needed, **what** is the purpose, and **who** is the audience?

Considerations:

- Prepping data for reporting
- Data report format
- Audience Interactivity

DATA REPORTING



Common Pitfalls:

- **Formatting**
- **Missing information**

✓ Solutions:

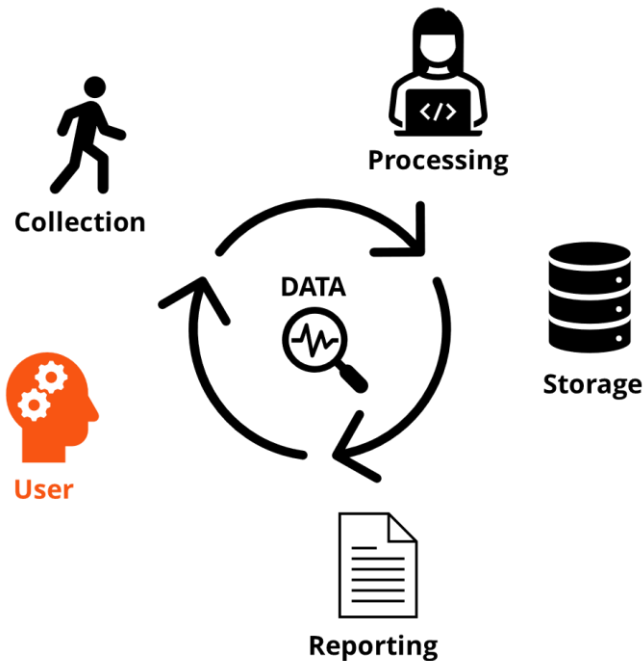
- ✓ **Consider reporting needs during data collection**
- ✓ **Recurring reporting can benefit from automation**

DATA USER



Data user:

High level decision makers vs. detail focused needs



Questions to ask:

- *What level of data does your user need?*
- *What is the goal of the end use?*
- *How will the user interact with the data?*

COMMON THEMES

- ✓ QA/QC at every step
- ✓ Standardization and documentation of procedures are key
- ✓ Automation can reduce error and build in efficiency
- ✓ Data management is not “one-size-fits-all”
- ✓ **Error can easily be compounded!**

SUMMARY

Diverting Data Disasters

- Data can be powerful, but only if we understand it. Specifically:
 - Where it comes from
 - How it was created/modified
 - What it represents
 - And how to access it
- Identify the right tool for the job
- Consider the end use at the beginning of the process
- The data process is iterative, Data SOPs are living documents

**THANK
YOU**

Kelsey Mach – kelsey.mach@aspectconsulting.com

Hanna Winter – hanna.winter@aspectconsulting.com

