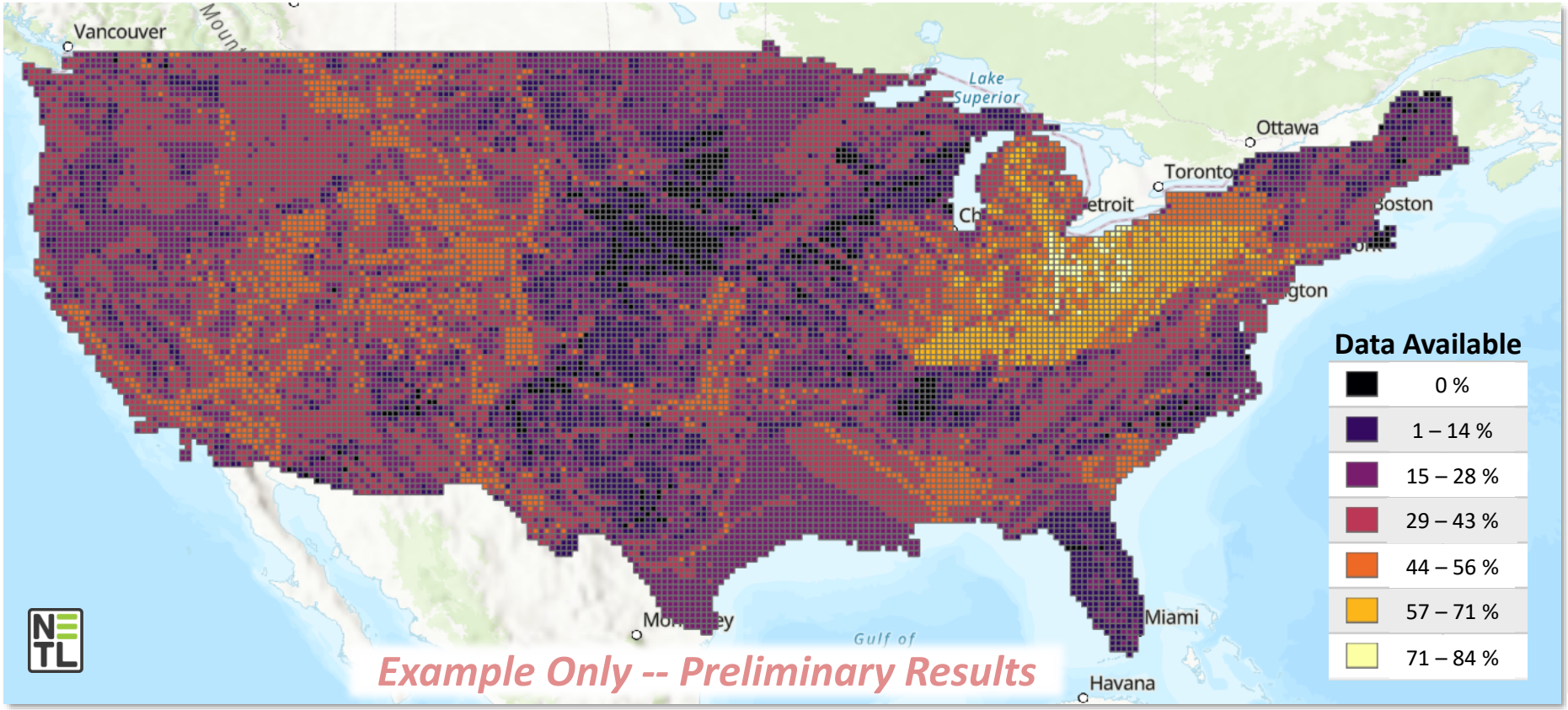
# Where are the Data? Automating a Workflow for Carbon Storage Data Gap Analyses

***Gabe Creason***

*National Energy Technology Laboratory*

***2024 FECM/NETL Carbon Management Research Project Review Meeting***




***August 2024***

# Disclaimer

This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

# Authors and Contact Information


***C. Gabriel Creason,[1] Julia Mulhern,[1,2] Neyda Cordero Rodriguez,[1,2] MacKenzie Mark-Moser,[1] Araceli Lara,[1,3] Jacob Shay,[1,2] Kelly Rose[1]***

***[1]National Energy Technology Laboratory, 1450 Queen Avenue SW, Albany, OR 97321, USA***

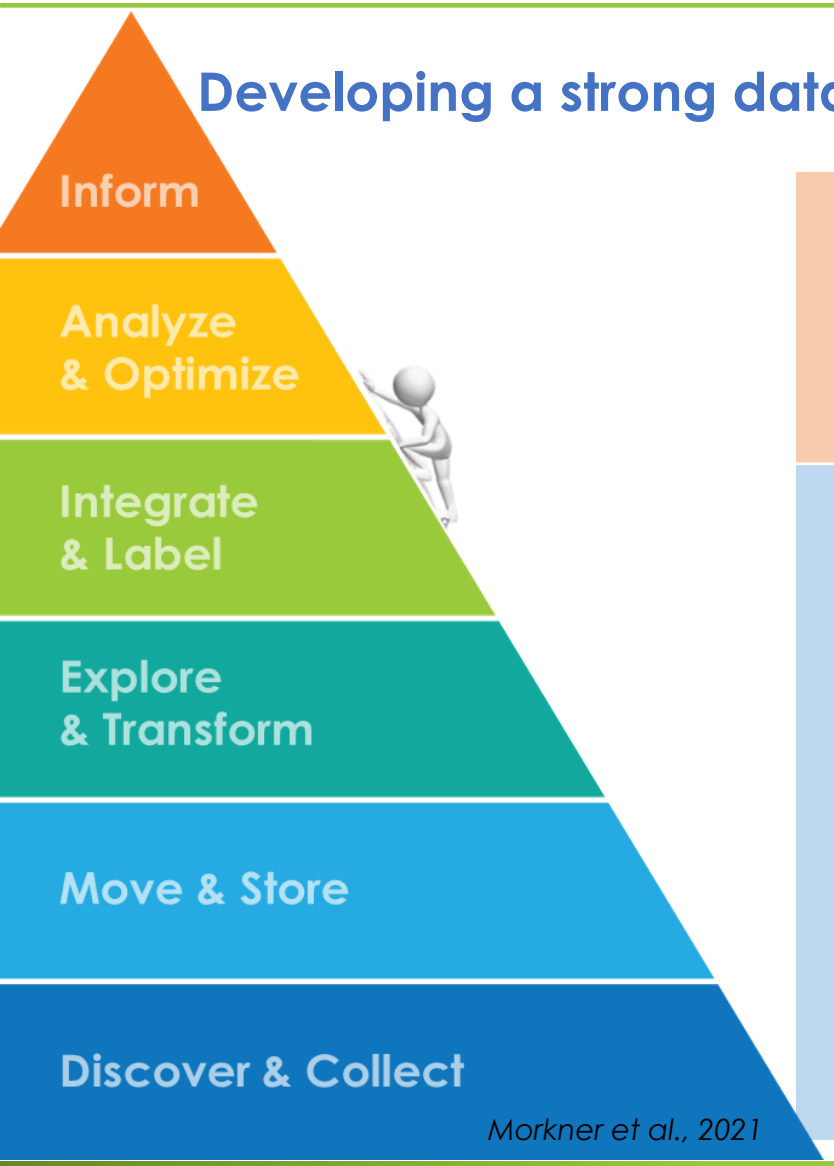
***[2]NETL Support Contractor, 1450 Queen Avenue SW, Albany, OR 97321, USA***

***[3]Oak Ridge Institute for Science and Education Fellowship, 1450 Queen Avenue SW, Albany, OR 97321, USA***

# Data are the *Energy* for Analysis & Inquiry

NATIONAL ENERGY TECHNOLOGY LABORATORY

## Developing a strong data foundation is key to any program/project's success




*Morkner et al., 2021*

**20%**

**80%**

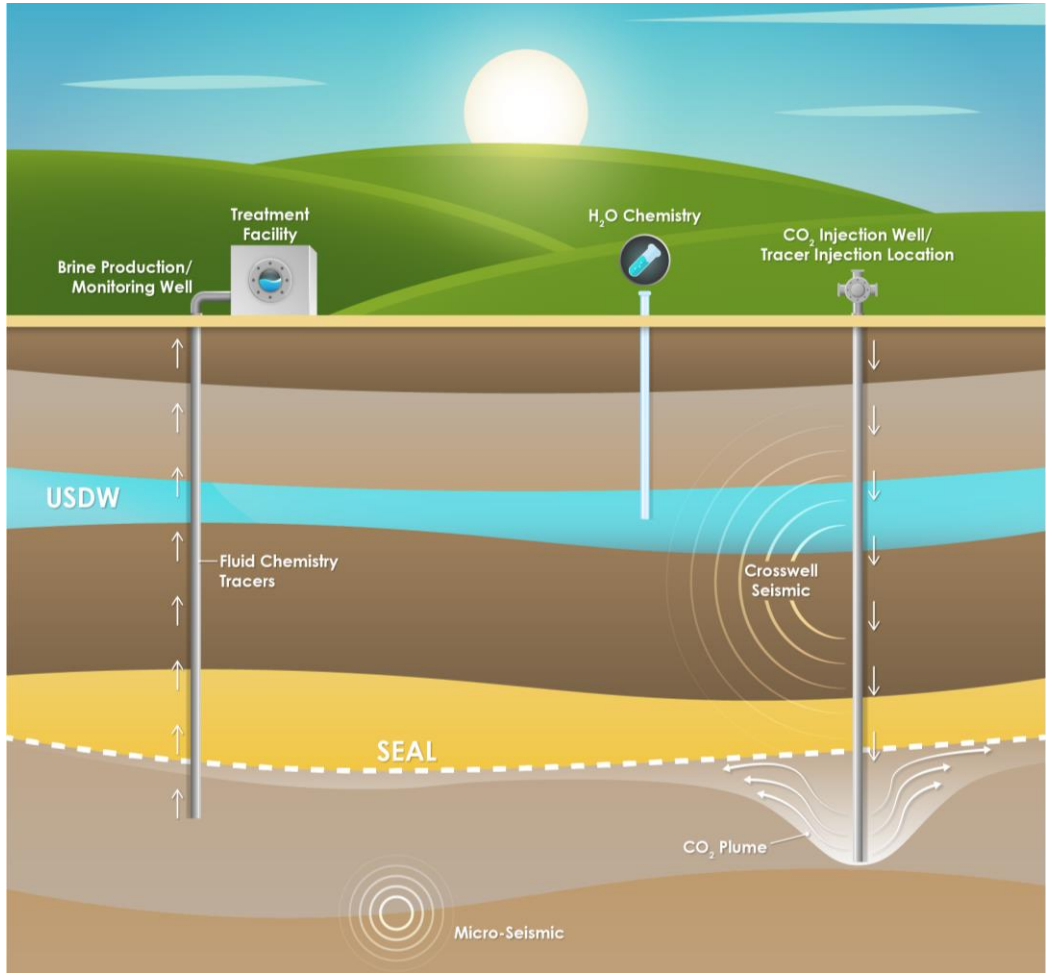
Data-driven teams can spend up to ~80% of their time addressing the bottom components of the "data pyramid"

*Crowdflower 2016*

# A Workflow for Carbon Storage Data Gap Analyses

## Motivation:

- **Need data** to inform $CO_2$ storage resource assessments
- Carbon storage technical viability **determined by more than geological factors**:
  - *environmental*
  - *socio-economic*
  - *infrastructure*




## Overview:

1. **Define an ontology of components** of carbon storage technical viability
2. **Collect and label data** with appropriate components/categories
3. **Spatially assess** the availability of different components/categories

***Where are the data?? Are they useful?***

# 1. Ontology for CS Technical Viability

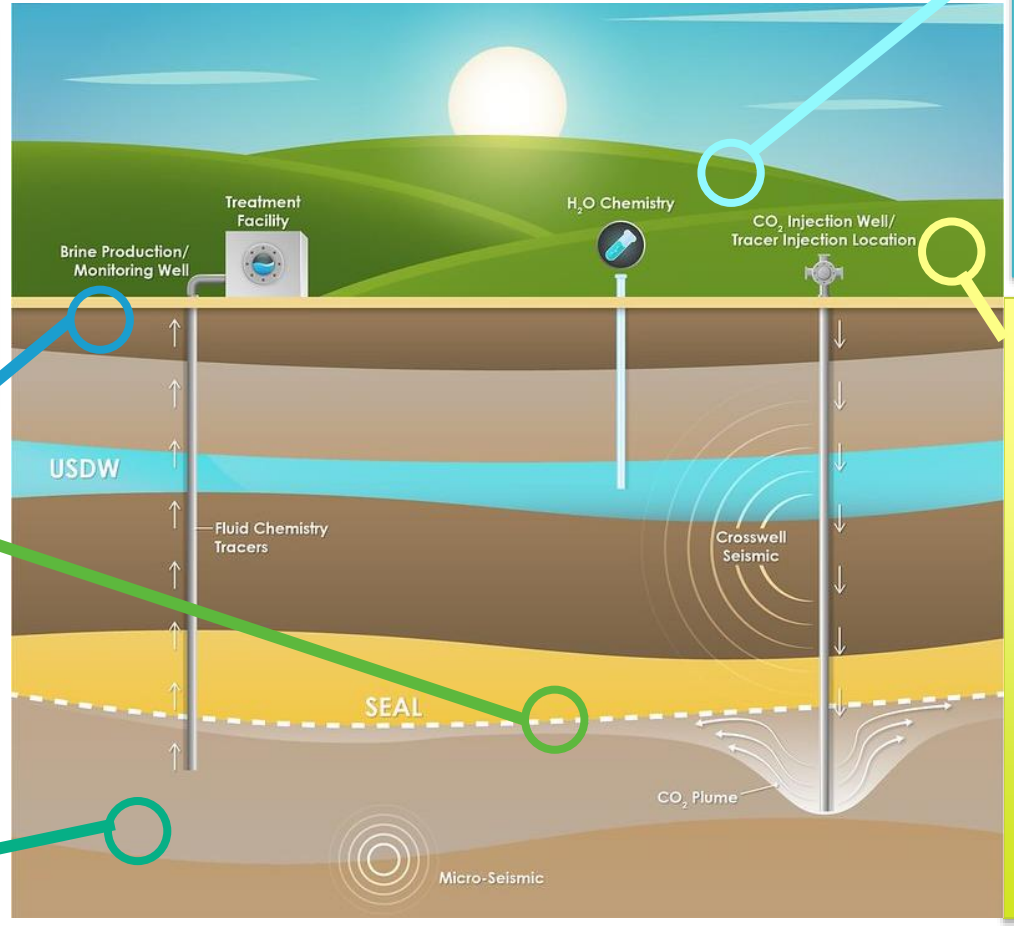
## *5 Categories → 14 Subcategories*

- Reflects **multidisciplinary requirements** of carbon storage projects

**Retention and Geomechanical Risk**
- Seals and Pressure
- Trap
- Faulting

**Hazards**
- Subsurface Hazards
- Surface Hazards

**Reservoir Suitability**
- Reservoir Quality
- Reservoir Geometry
- Reservoir Conditions

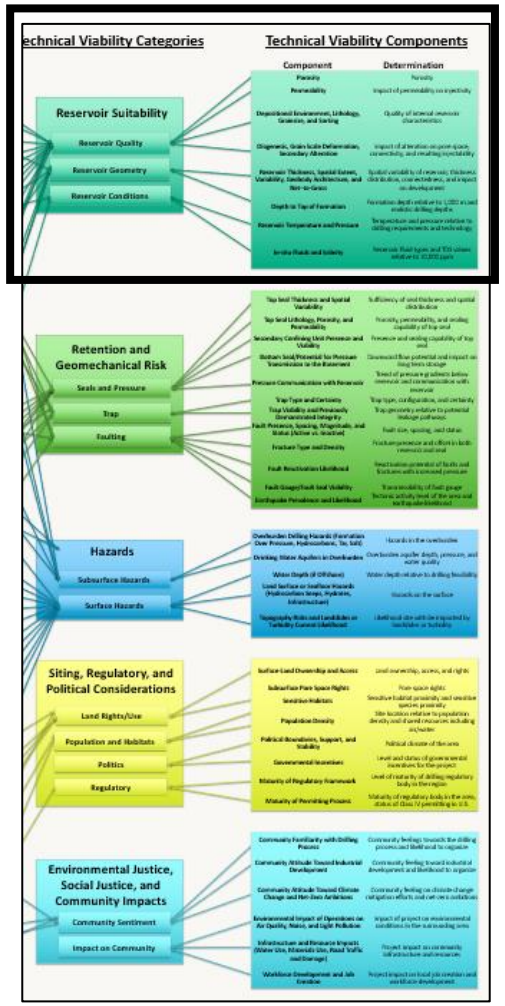


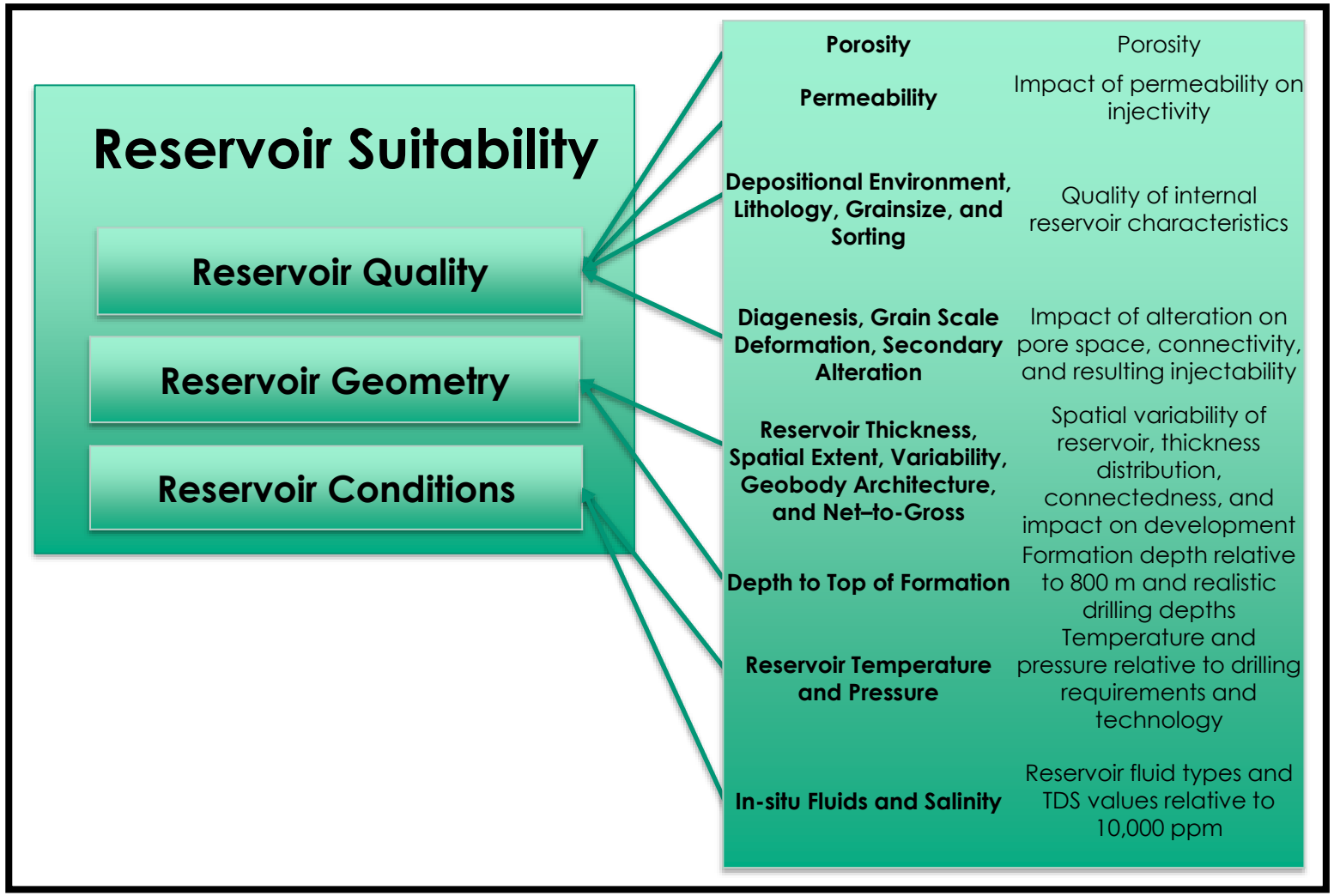

**Environmental Justice, Social Justice, and Community Impacts**
- Community Sentiment
- Impact on Community

**Siting, Regulatory, and Jurisdiction Considerations**
- Land Rights/Use
- Population and Habitats
- Jurisdiction
- Regulatory

# 1. Ontology for CS Technical Viability

## *5 Categories → 14 Subcategories → 46 Components*



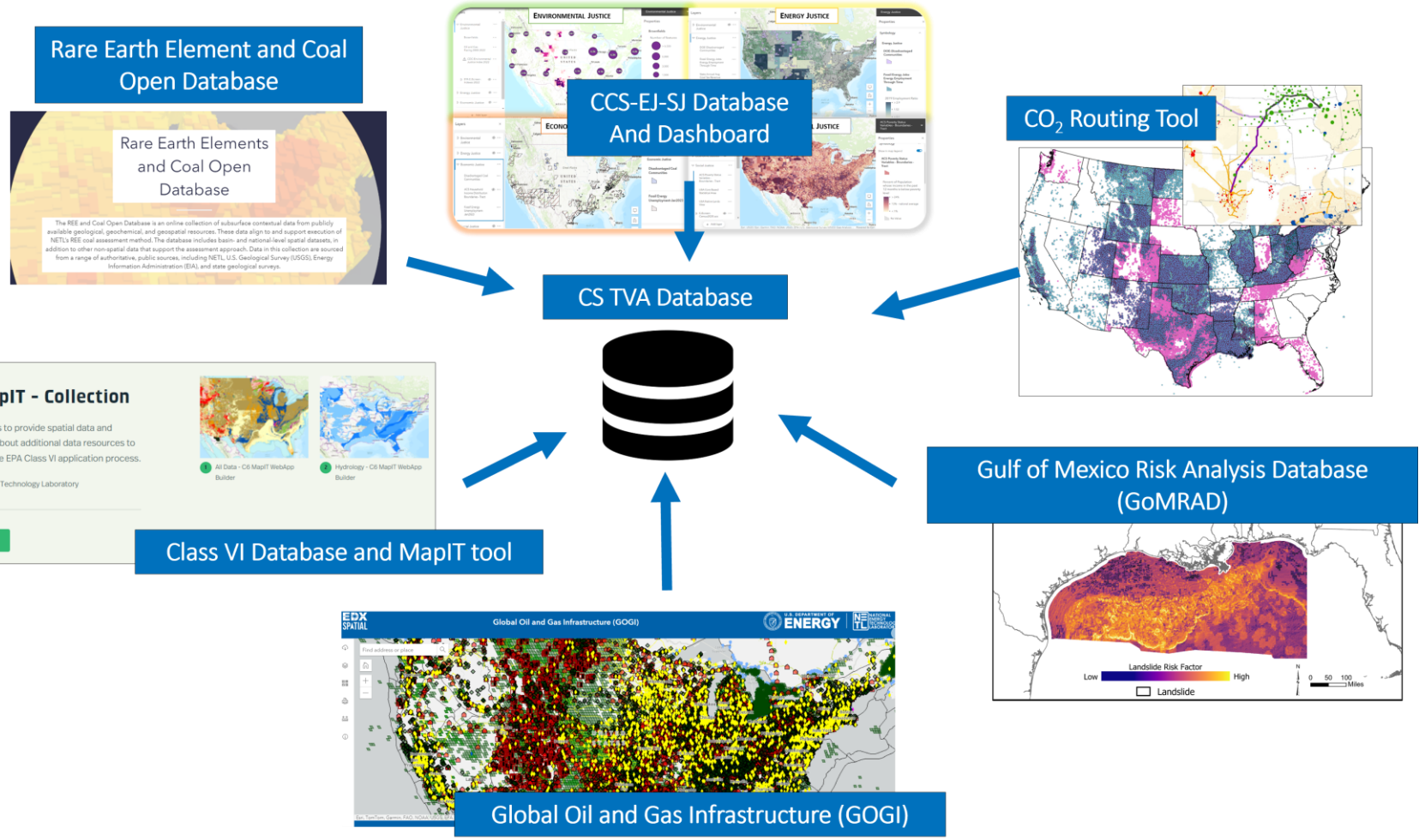
# 2. Collect and Label Data

## *Aggregated public data from across EDX4CCS and other NETL efforts*

- **1400+ files**
- **120 GB**
- **Removed duplicates**

- **Published July 2024**

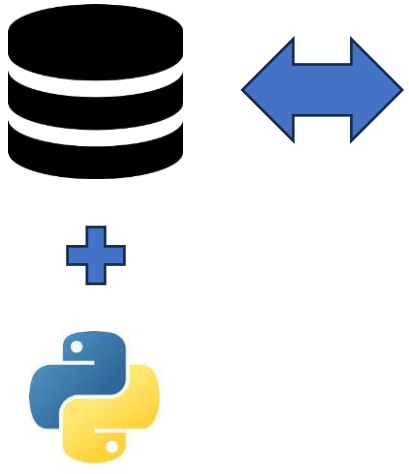
***Mulhern et al., Carbon Storage Technical Viability Approach (CS TVA) Database***
**DOI: 10.18141/1984655**



# 2. Collect and Label Data

## *Defined relationships*

### Database



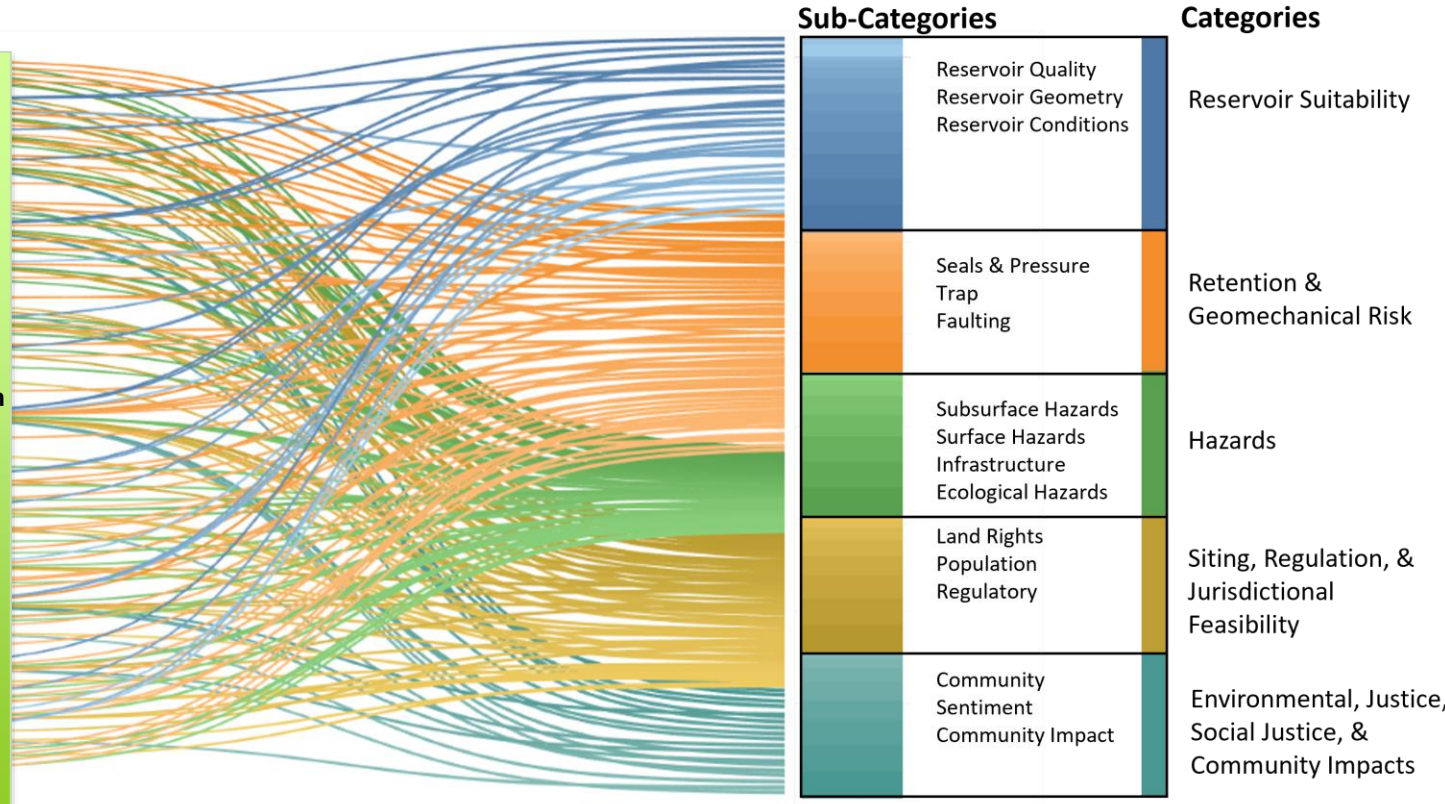
*Developed script to **query database** contents for keywords **to assign data labels***

### Data Types

**Subsurface**

- Porosity
- Permeability, Rel. Permeability, Anisotropy Ratio
- Well Logs, Down-Hole Logs, Log Interp
- Core, Core Logs, Core Samples, Core Analysis
- Petrographic Analysis, Thin Sections
- Paleontology, Micro-Paleontology
- Well Image Logs, Image Log Interpretations
- Secondary Alteration, Diagenetic History, Basin History Model
- Seismic Data, Derived Data Volumes, Horizons & Surfaces, Interpretations (2D/3D)
- Geomodels, Well Correlations, Cross-Sections, Stratigraphic Columns, Sedimentary Facies, Lithology, Geobody Interpretations, V-Shale/Clay
- Reservoir Depth
- Reservoir Thickness
- …
- …
- etc.

### CS Technical Viability Components




***"One-to-many" relationship** where data types relate to multiple components*

# 3. Spatial Analysis

## *Presence/absence of each dataset*

**Aggregate:**

- Data types
- Data types per component
- Components per category

**Data Type 1**

| | | | |
|---|---|---|---|
| 1 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 |
| 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 |
| 1 | 1 | 1 | 0 |

+

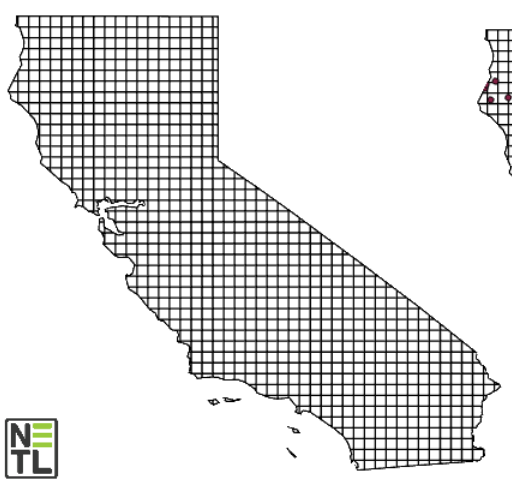
**Data Type 2**

| | | | |
|---|---|---|---|
| 1 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 0 | 0 | 1 | 1 |
| 0 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 |

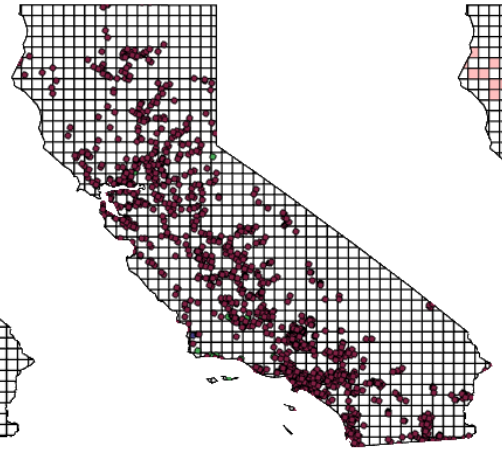
=

**Component A**

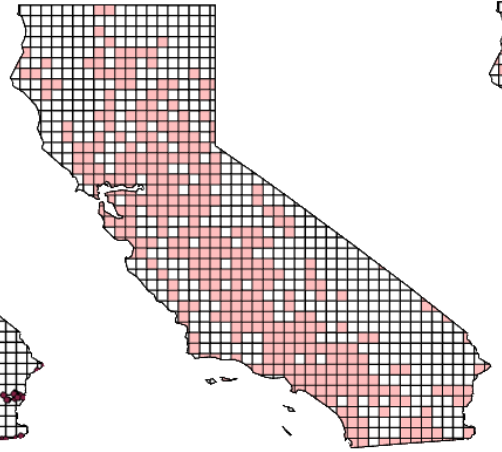
| | | | |
|---|---|---|---|
| 100% | 50% | 0% | 50% |
| 0% | 100% | 100% | 50% |
| 0% | 0% | 100% | 100% |
| 0% | 0% | 100% | 50% |
| 0% | 50% | 100% | 0% |
| 50% | 100% | 100% | 0% |



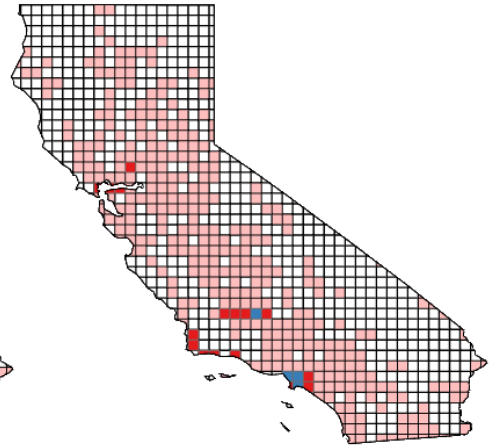
**Empty Grid**



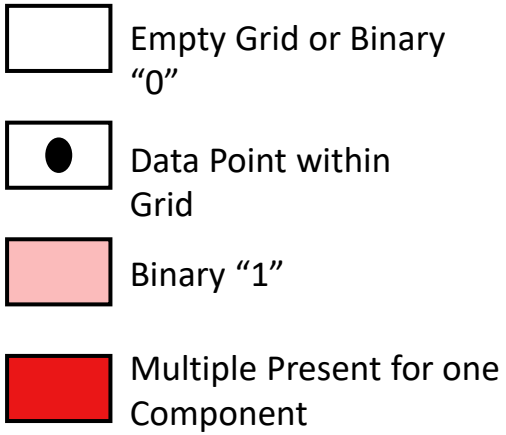
**Power Plants**
**- Data -**



**Power Plants**
**- Results -**



**Power Plants +**
**Processing Plants**



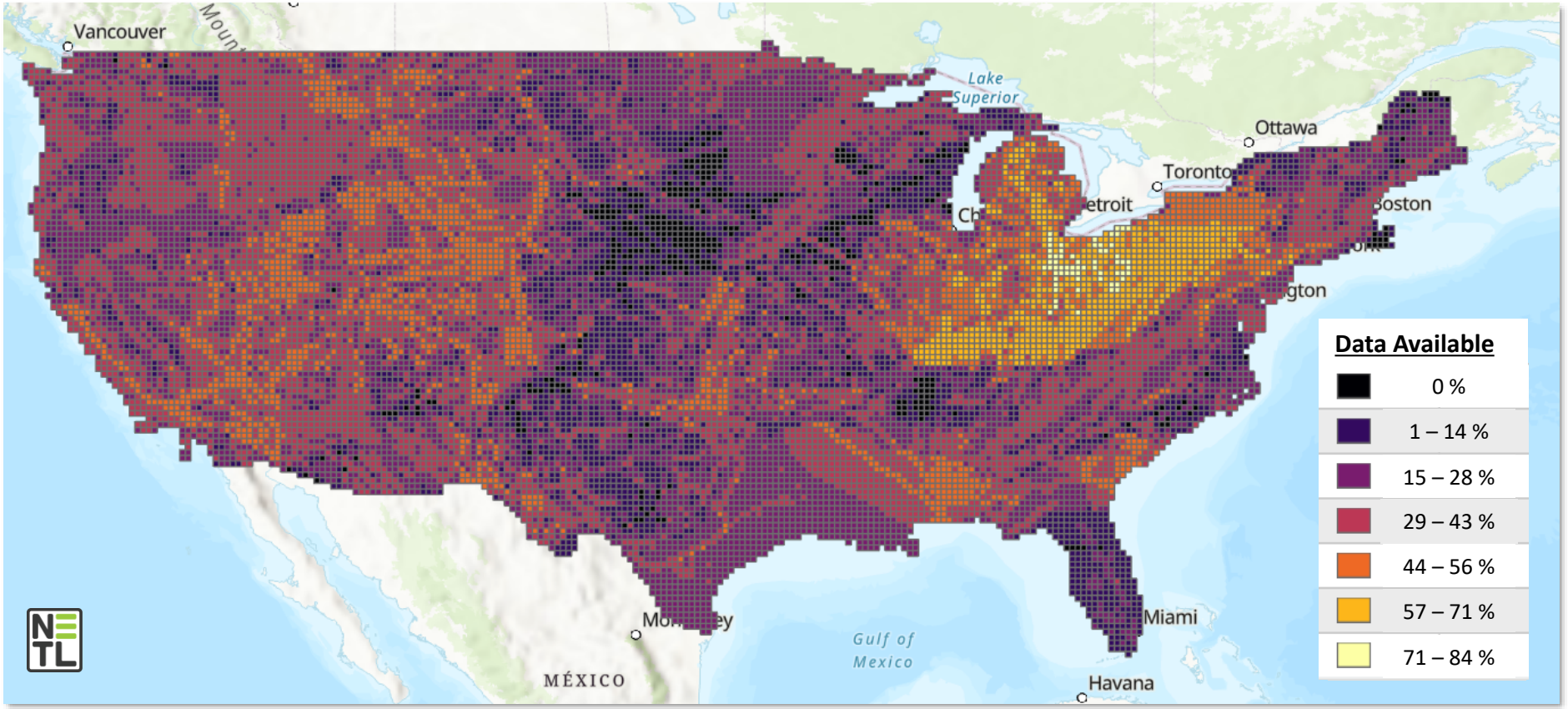
# 3. Spatial Analysis

## *Example Results: Bottom Seal, Induced Seismicity*

- **Category:** Retention and Geomechanical Risk
  - **Sub-category:** Seals and Pressure
    - **Component:** Bottom Seal, Induced Seismicity

**Component Context:** Information about downward flow potential, impact on long term storage, and risk for induced seismicity in crystalline basement.

| Data Types |
|---|
| Seismic Data |
| Pressure Measurements |
| Well Logs |
| Faults - Basement |
| Faults - Undifferentiated |
| Derived 3D Volumes |
| Bedrock Geology |
| Faults - 3D |
| Extent - Basement |
| Depth - Basement |

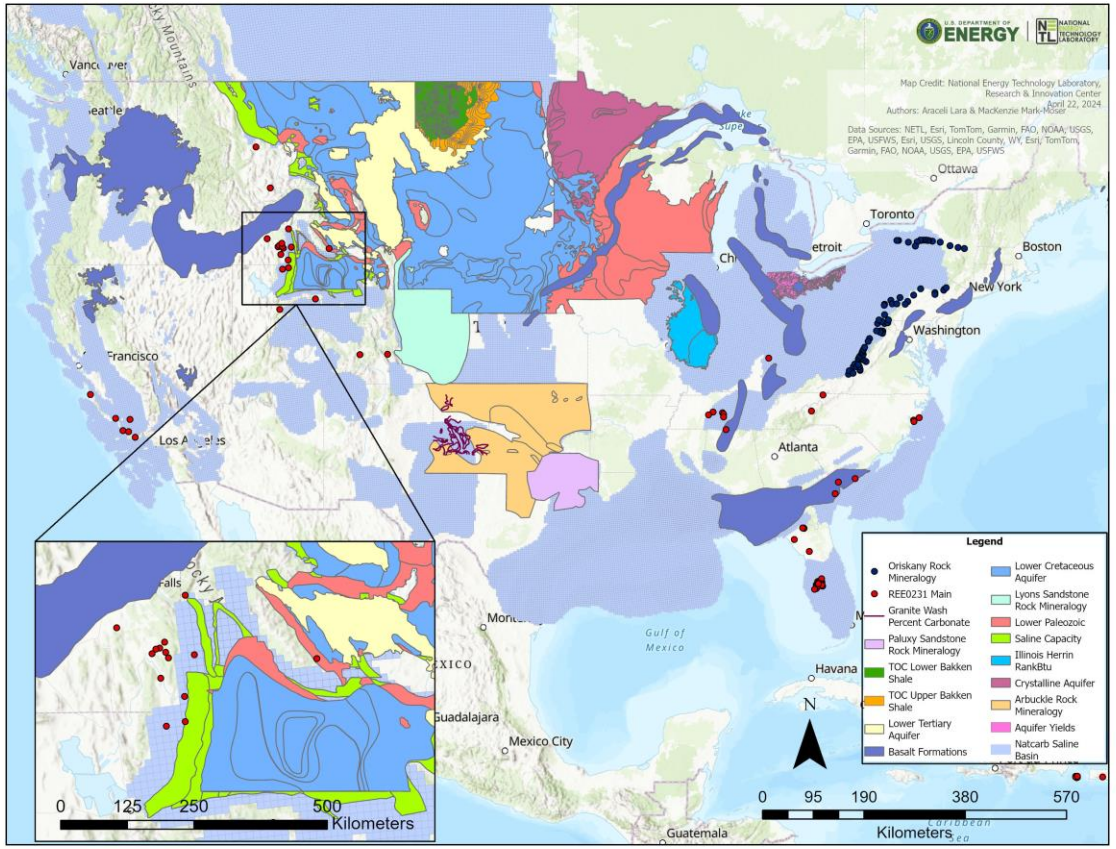


# Next Steps

## *Improved usability and advanced labeling*

1. Develop into **online tool hosted on disCO2ver** platform
   - Improve user accessibility

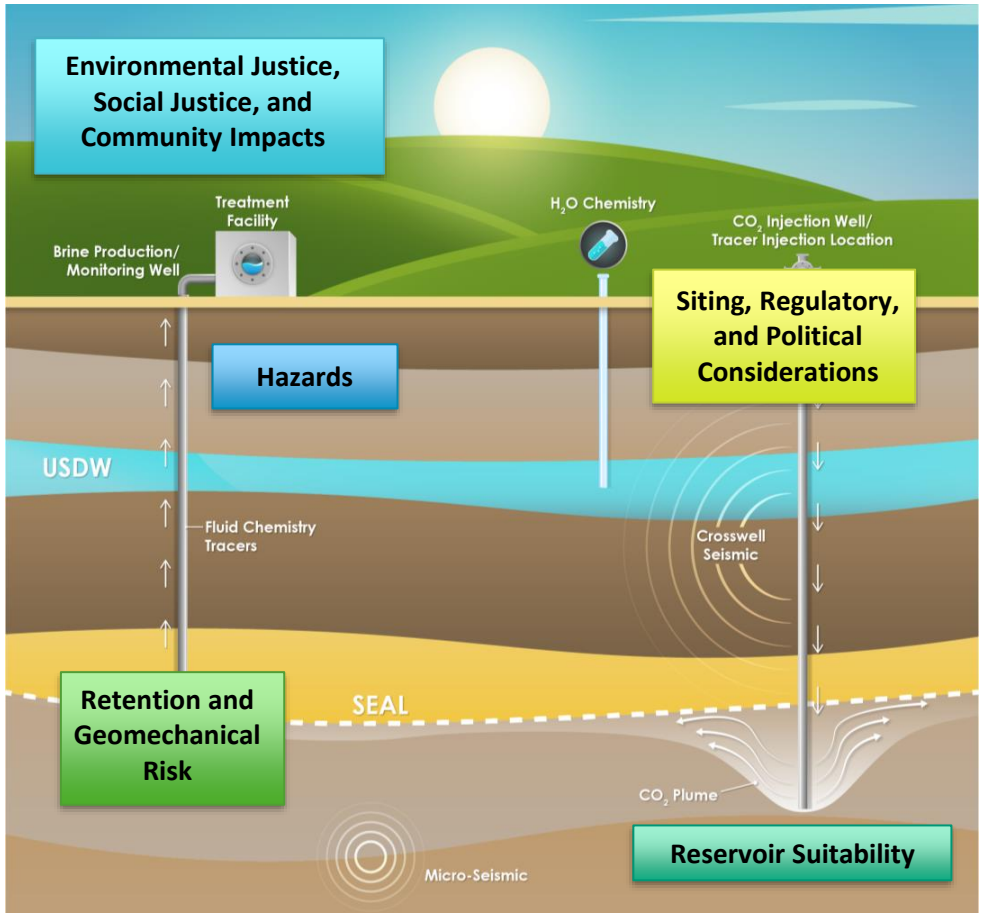
2. **Augment results** using available NETL tools and **NLP/GenAI**
   - **Assigning weights** to different data types
   - **NLP/GenAI assisted classification** of datasets
     - More robust labeling
     - Streamline analysis of new datasets
   - **AI/ML** Subsurface Trend Analysis Tool **for data collection**
     - e.g., Image recognition, image embedding, table extraction




***Where are the data?? Are they useful?***

# Key Takeaways

- Produced a **workflow** for assessing data availability *and* utility for **technically viable carbon storage assessments**
- Considers **multiple factors** beyond reservoir and caprock viability
- **Contextualizes** the available **data**
- **Identifies** areas with **limited data** availability




***Where are the data?? Are they useful?***

# the stats

**54** RIC PRESENTATIONS

**22** POSTERS

**30** TOOL DEMOS

## MONDAY

### Presentations
(10:30AM - 5:25PM)

- 16 disCO2ver presentations

## TUESDAY

### Presentations
(10:30AM - 5:45PM)

- 17 SMART presentations
- 2 disCO2ver presentations
- 2 Geographic focus/tool presentations

### Posters
(5:45PM - 7:45PM)

- 18 CTS Posters
- 2 PSCC Posters
- 1 CDR Poster
- 1 MLEF Poster

### Tool Demos
(5:45PM - 7:45PM)

- 30 Tool Demos
  - SMART
  - NRAP
  - EDX
  - EDX4CCS

## WEDNESDAY

### Presentations
(2:10PM - 4:30PM)

- 3 transport, research, development, and demonstration activities presentations
- 1 transport modeling presentation
- 1 secure storage (basalts/mafic) presentation

## THURSDAY

### Presentations
(10:30AM - 5:20PM)

- 8 NRAP presentations
- 2 NETL RIC Presentations
- 2 Offshore presentations

NETL Research · Carbon Storage Program

NRAP National Risk Assessment Partnership

SMART

EDX disco2ver

https://edx.netl.doe.gov/disco2ver

# Acknowledgments

This work was performed in support of the U.S. Department of Energy's (DOE) Office of Fossil Energy and Carbon Management's EDX4CCS Project, in part, from the Bipartisan Infrastructure Law.

# NETL RESOURCES

VISIT US AT: **www.NETL.DOE.gov**

**@NETL_DOE**

**@NETL_DOE**

**@NationalEnergyTechnologyLaboratory**

POCs
C. Gabe Creason, gabriel.creason@netl.doe.gov
MacKenzie Mark-Moser, mackenzie.mark-moser@netl.doe.gov
Julia Mulhern, Julia.Mulhern@netl.doe.gov
Kelly Rose, Kelly.rose@netl.doe.gov