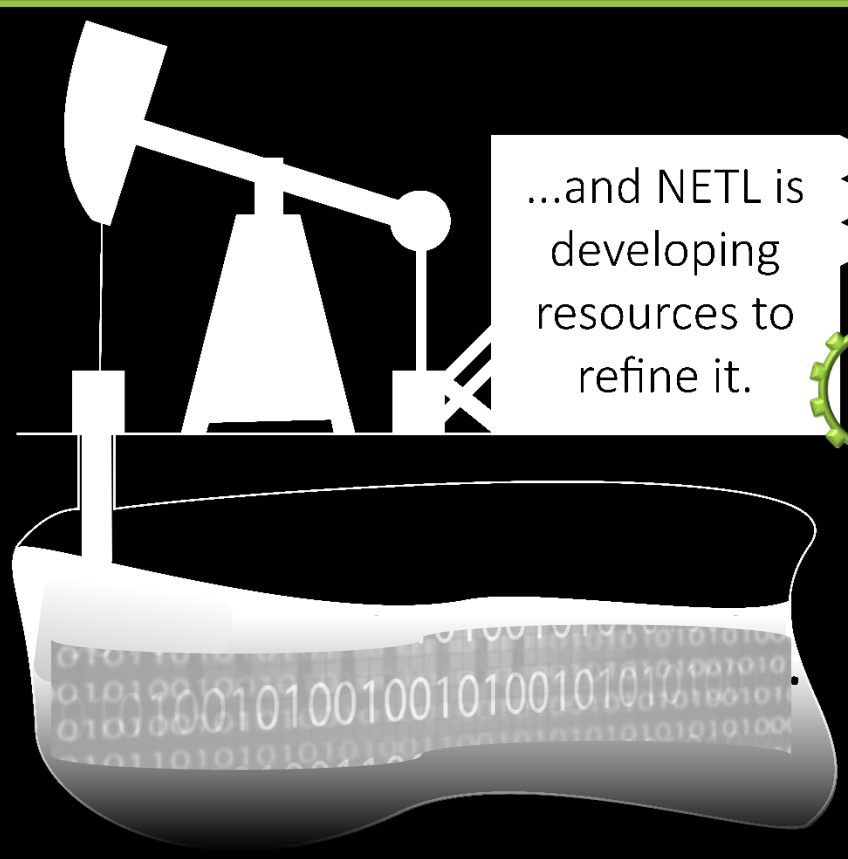
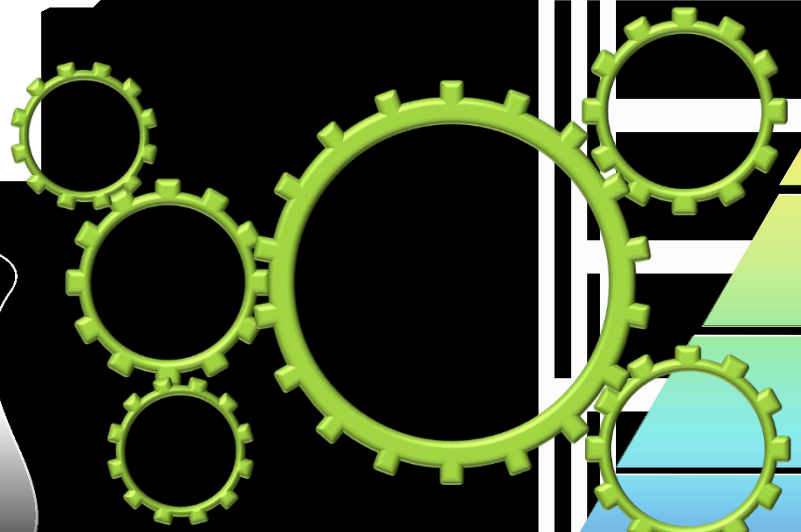
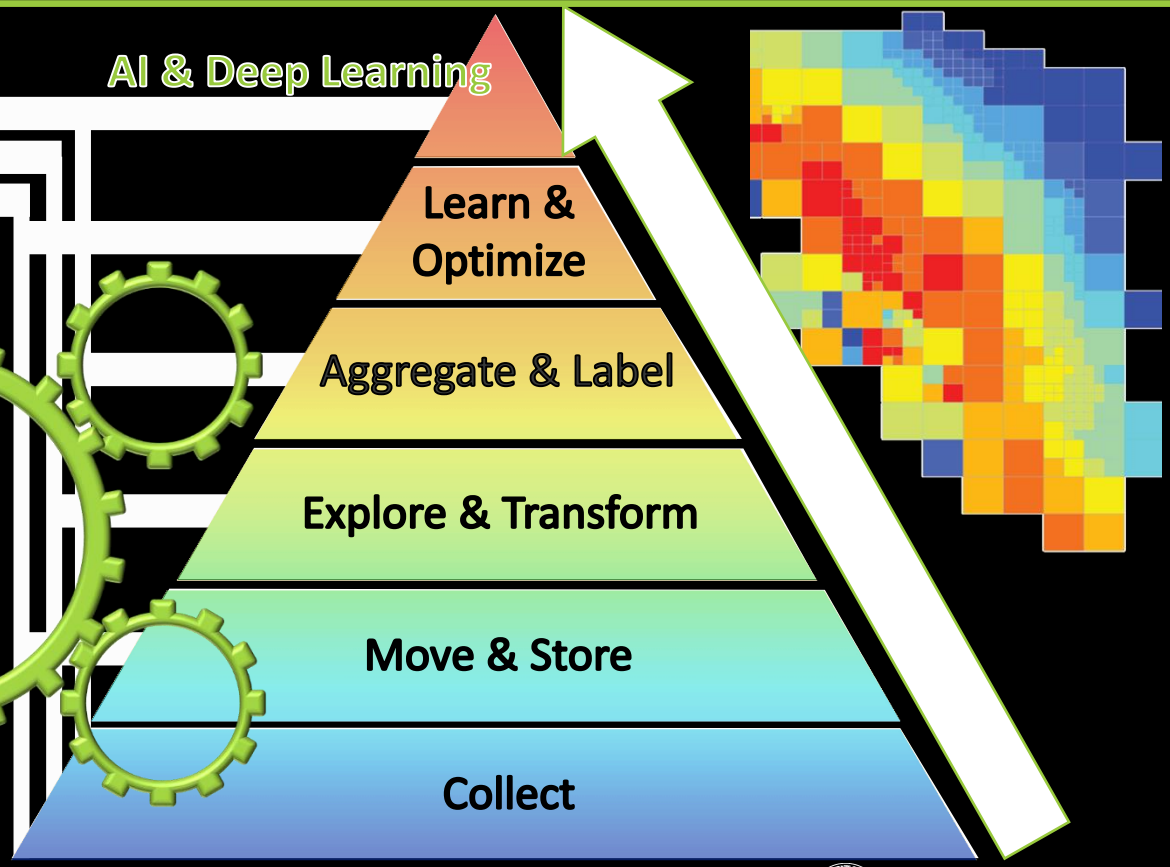
Data is the new oil...

NATIONAL ENERGY TECHNOLOGY LABORATORY



...and NETL is developing resources to refine it.

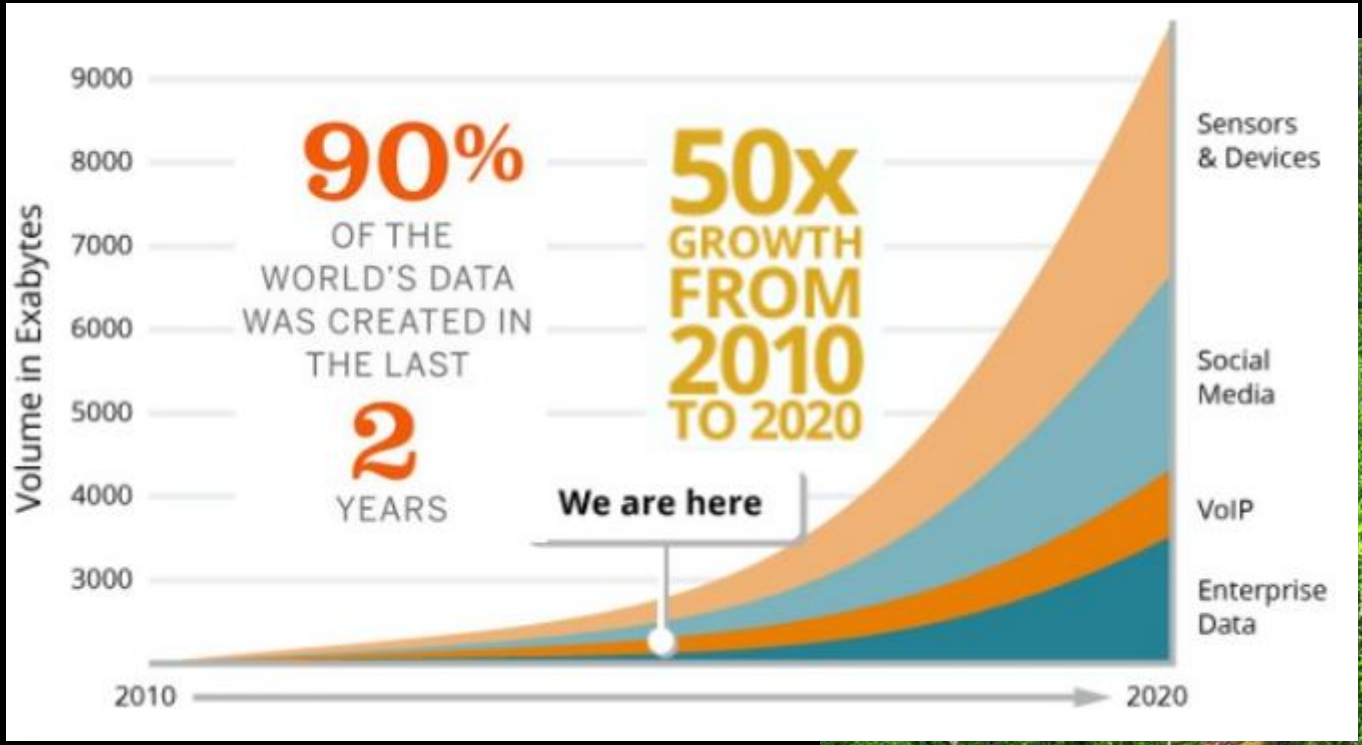


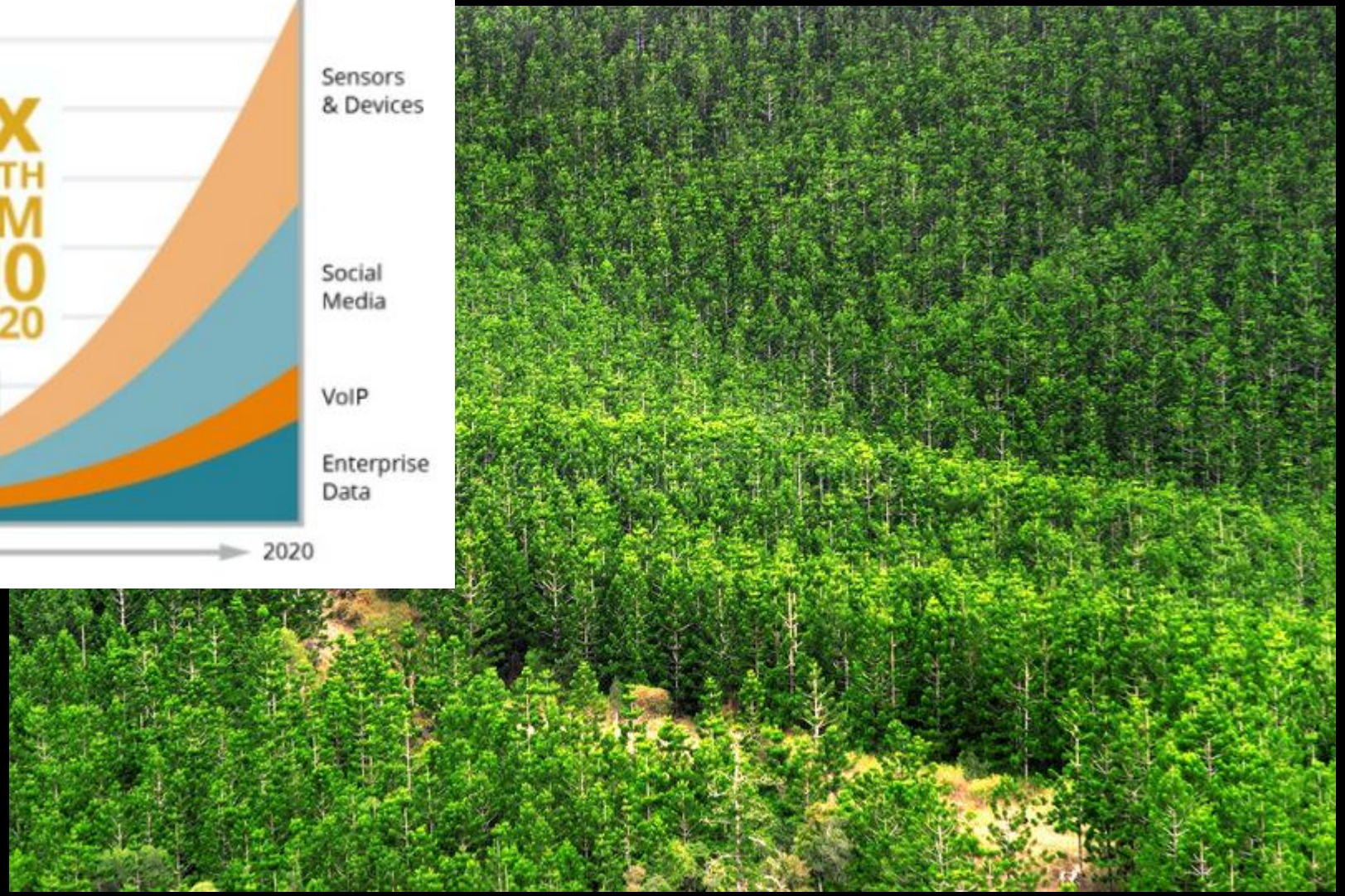


Solutions for Today | Options for Tomorrow

U.S. DEPARTMENT OF ENERGY

As access to open, authoritative data increases science driven analyses face challenges to **efficiently find**, **integrate** and **use** these resources

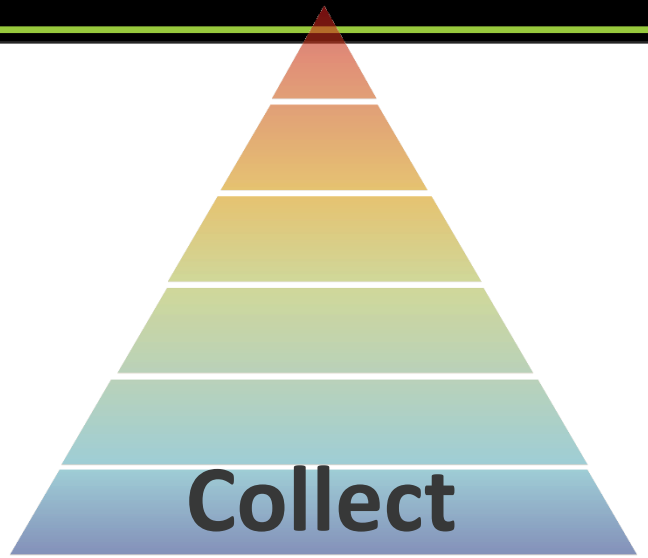




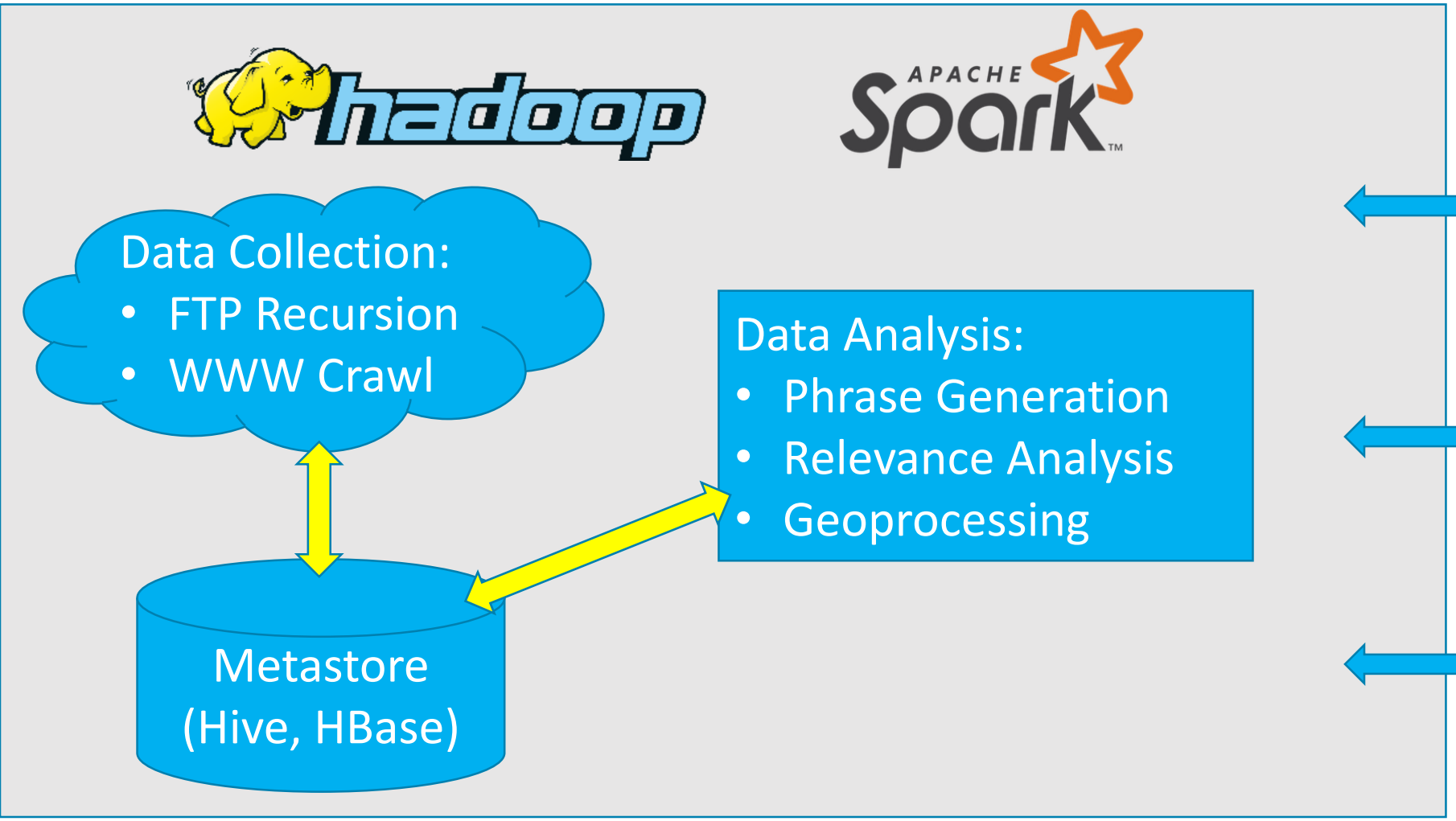
- Volume, variety, and velocity of data online is growing… exponentially
- How will you parse the tree from the forest?

Use Case: FTP Data Mining: Hadoop + EDX

- **Problem:**
  - Need to search data in FTP silos (millions of files, spatial and contextual)
- **Solution:**
  - Index FTP silos using Hadoop



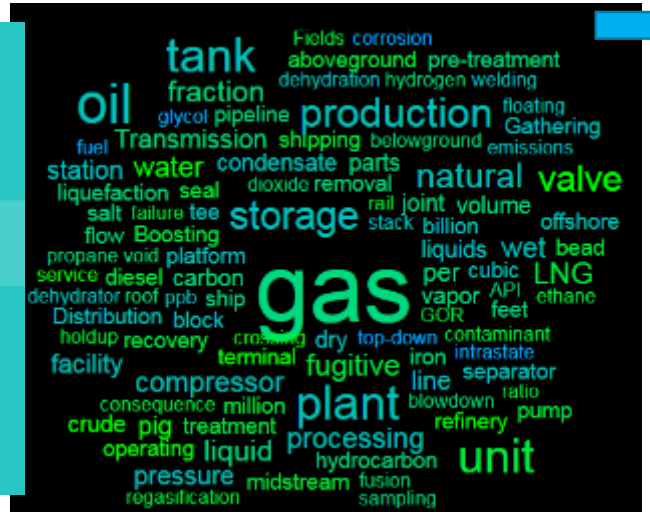
NETL's Big Data Discovery Ecosystem (To Date)



NATIONAL ENERGY TECHNOLOGY LABORATORY

BIG DATA

For Assistive Discovery



EDX™ Energy Data eXchange

Convert Search Terms & Phrases into an Open O&G Spatial Database

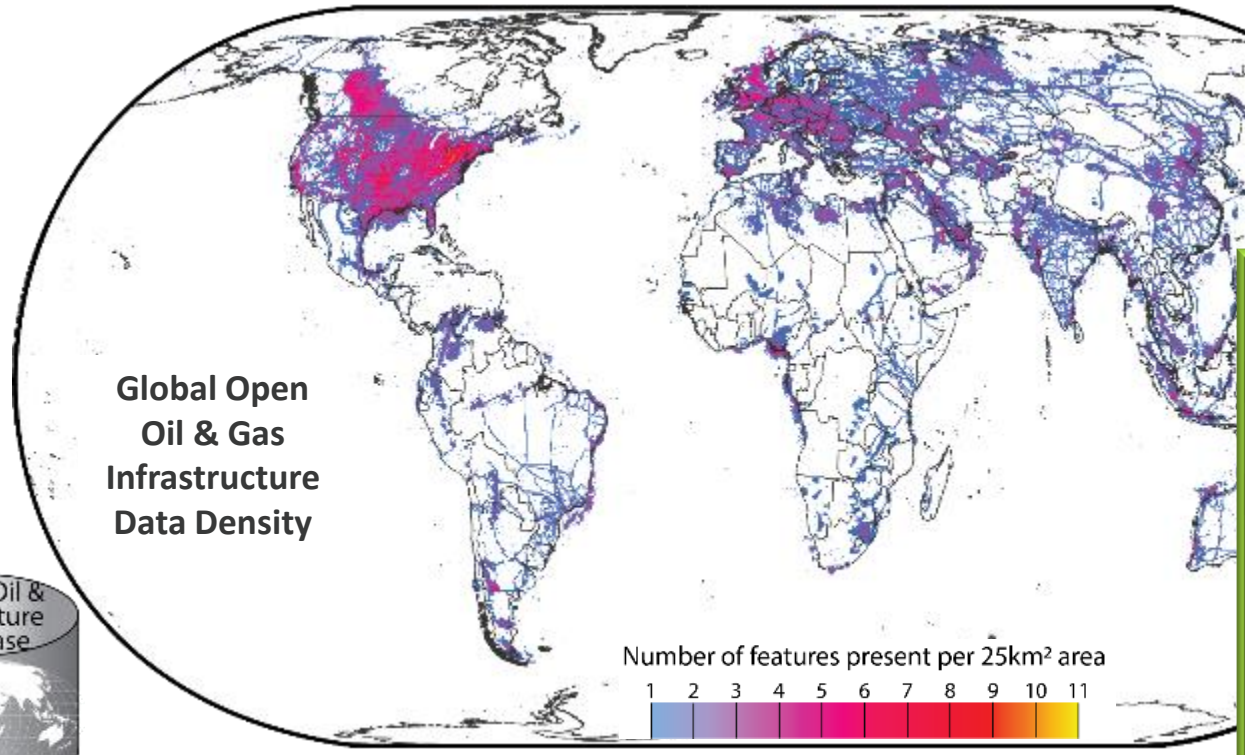
NATIONAL ENERGY TECHNOLOGY LABORATORY

Machine Learning Automated Approach- A tool that scans "seed" resources and identifies relevant keywords, then crawls the web and parses the data for integration

4 months period of performance

Acquisition of disparate data by country, region, & continent totaling:

- **~800 datasets**
- **>4 Millions of features**

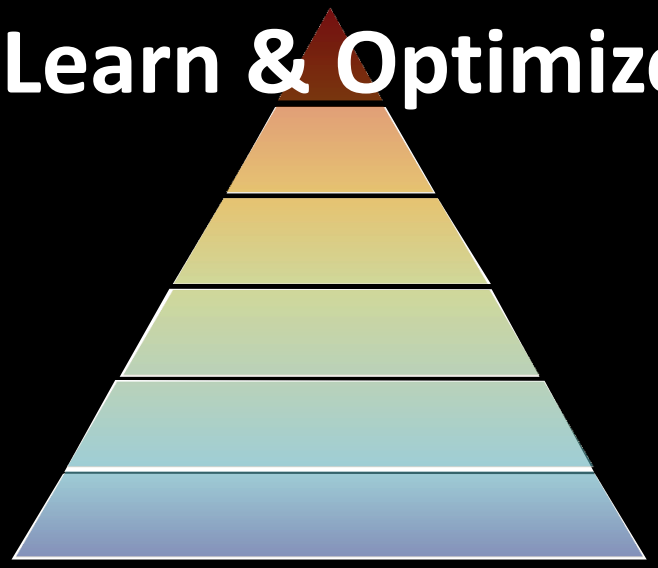


Discovered & integrated *open data sources* of information related to oil & gas infrastructure across the globe

U.S. DEPARTMENT OF ENERGY | NATIONAL ENERGY TECHNOLOGY LABORATORY

https://edx.netl.doe.gov/dataset/global-oil-gas-features-database

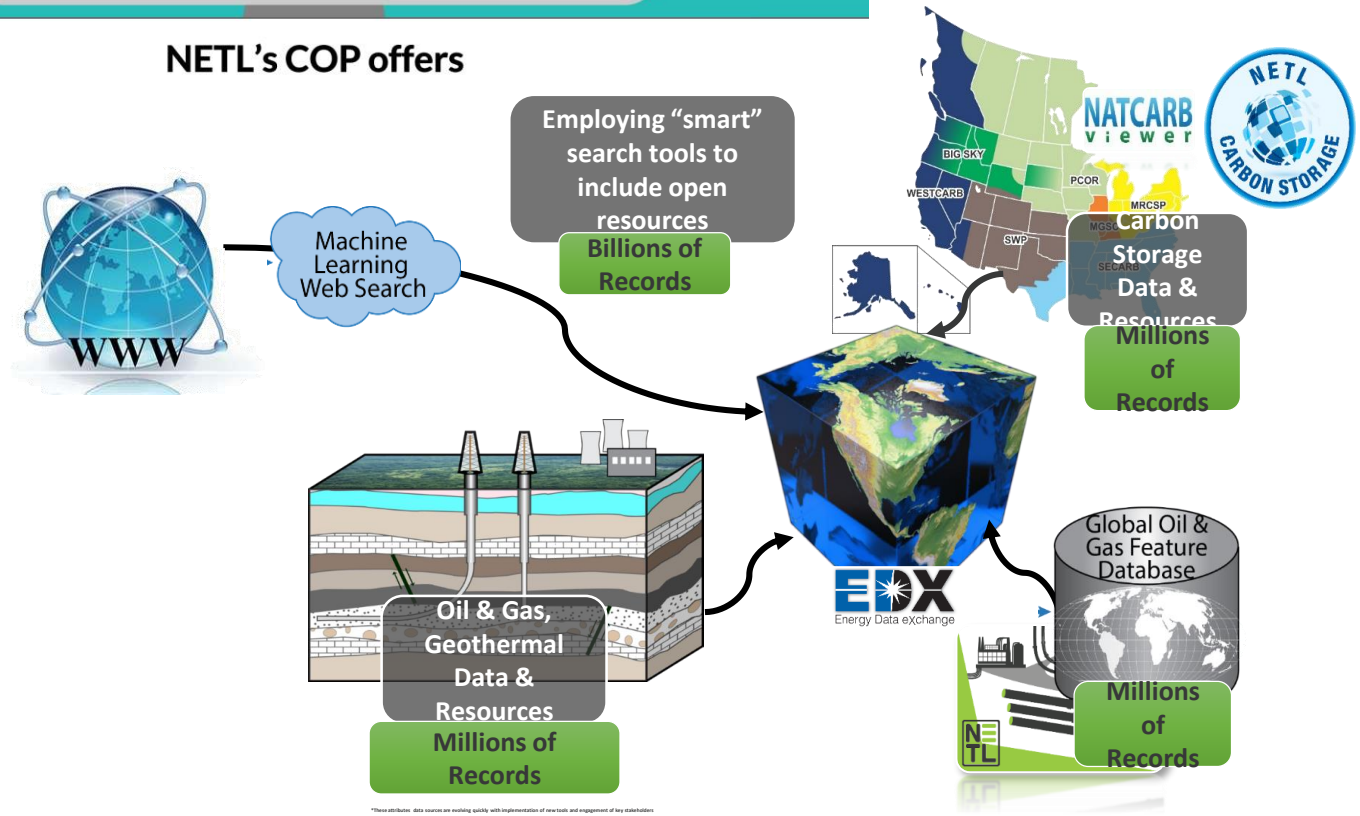
Learn & Optimize



Future Big Data Development & Analysis

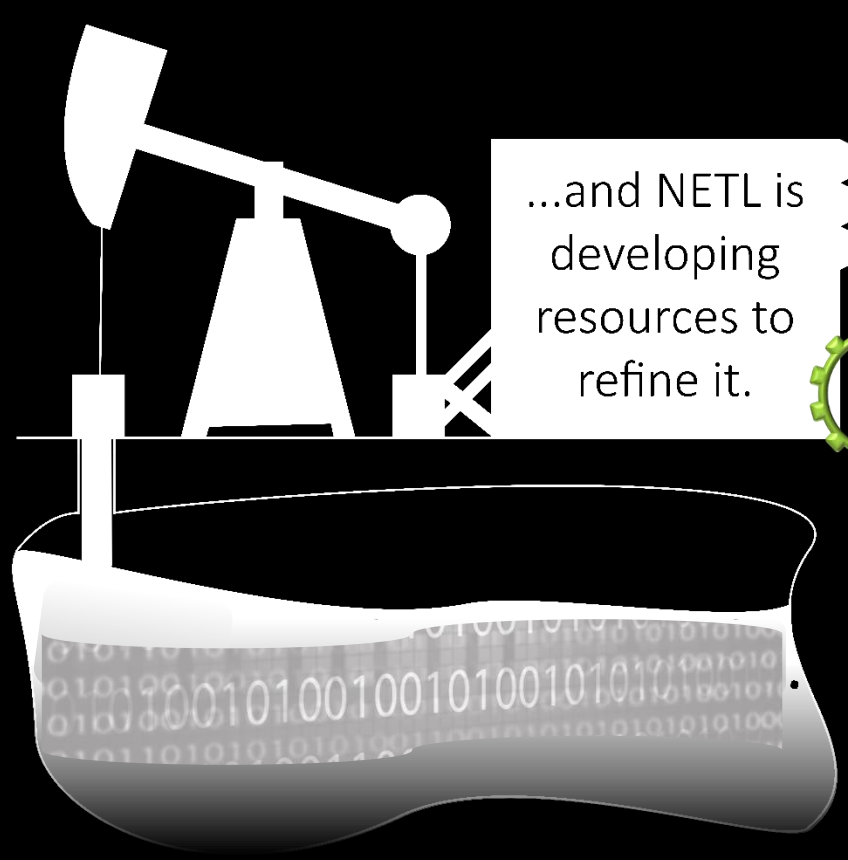
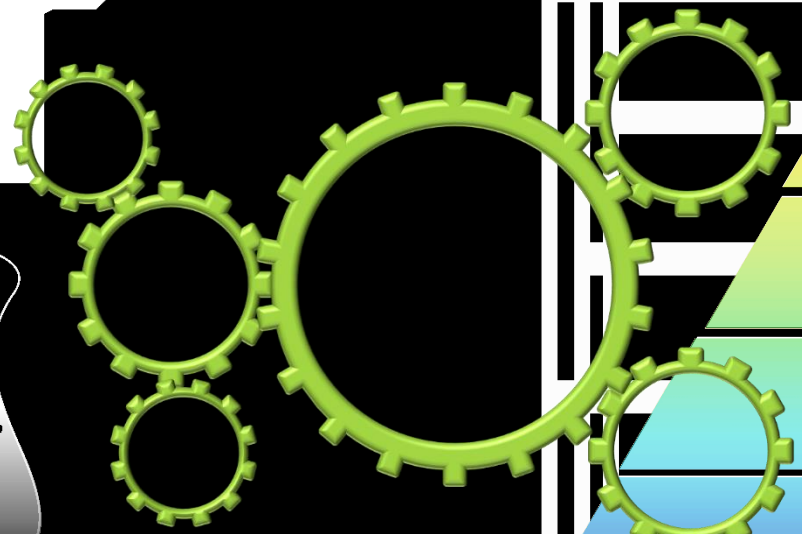
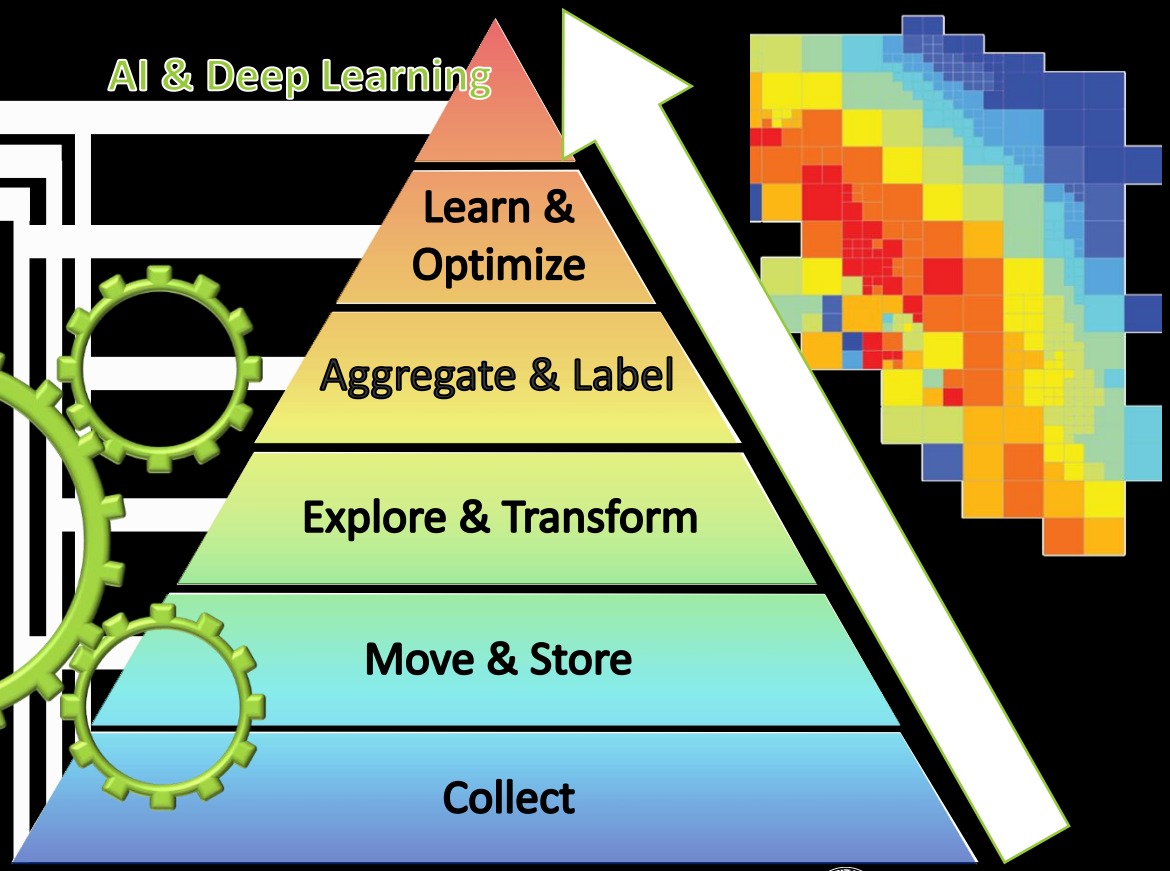
- EDX Cloud Services
- Living Database
- Common Operating Platform for Data Analytics
- Geocube Spatial Data Viewer
- Fuzzy Logic Analytics (SIMPA)
- AWS Development
- Integration with decades of DOE R&D
- Federating Open Source data
- EDX & GeoCube (search and location)
- ID Data gaps in subsurface puzzle

NATIONAL ENERGY TECHNOLOGY LABORATORY



Data is the new oil...

NATIONAL ENERGY TECHNOLOGY LABORATORY



...and NETL is developing resources to refine it.