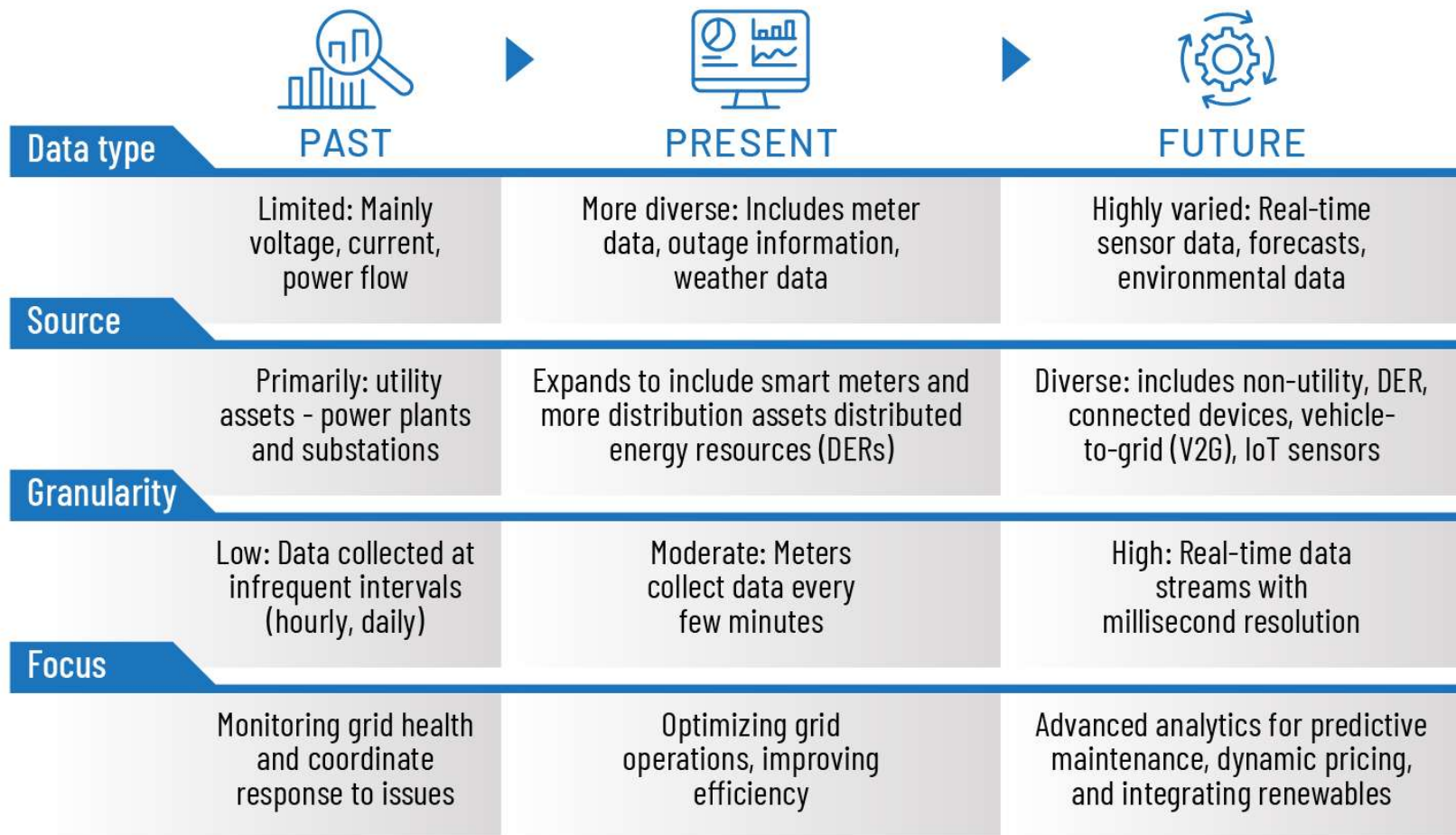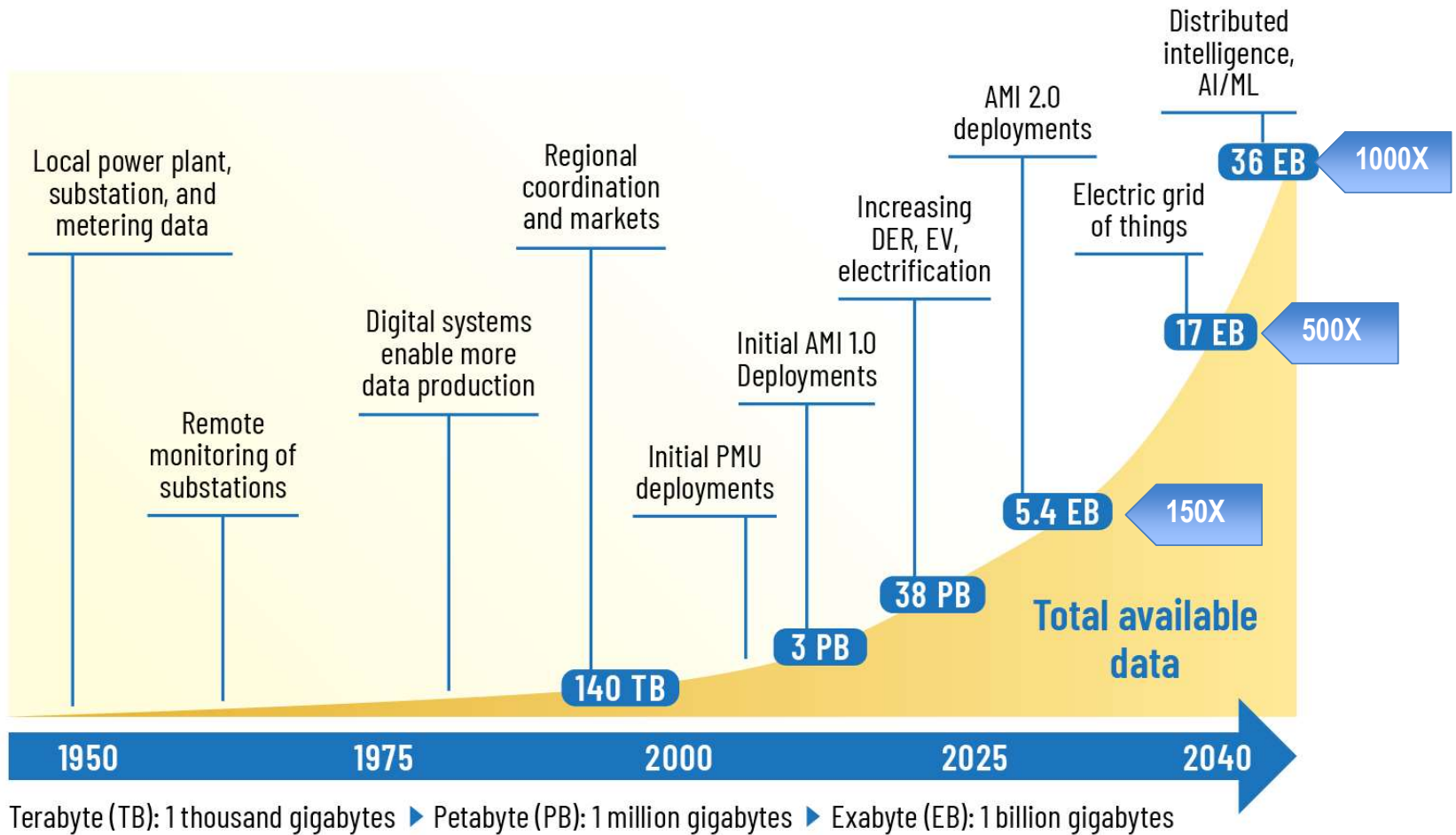# Agenda

- **Electric Sector Data Context**
    - 12 min Presentation, 20 min Discussion
- **Sector Strategy & Secure Data Portal**
    - 12 min Presentation, 20 min Discussion
- **Data Use & Protection**
    - 12 min Presentation, 20 min Discussion
- **Conclusion & Discussion**
    - 4 min Presentation, 20 min Discussion

Advanced Grid
Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

Electric Sector Data Context

The Evolving Role of Grid Data (Operational Perspective)

Data type	PAST	PRESENT	FUTURE
	Limited: Mainly voltage, current, power flow	More diverse: Includes meter data, outage information, weather data	Highly varied: Real-time sensor data, forecasts, environmental data
Source	Primarily: utility assets - power plants and substations	Expands to include smart meters and more distribution assets distributed energy resources (DERs)	Diverse: includes non-utility, DER, connected devices, vehicle-to-grid (V2G), IoT sensors
Granularity	Low: Data collected at infrequent intervals (hourly, daily)	Moderate: Meters collect data every few minutes	High: Real-time data streams with millisecond resolution
Focus	Monitoring grid health and coordinate response to issues	Optimizing grid operations, improving efficiency	Advanced analytics for predictive maintenance, dynamic pricing, and integrating renewables

Annual Grid Data Production Through The Years

Distributed intelligence, AI/ML — **36 EB** — 1000X

AMI 2.0 deployments

Electric grid of things — **17 EB** — 500X

Local power plant, substation, and metering data

Regional coordination and markets

Increasing DER, EV, electrification

Digital systems enable more data production

Initial AMI 1.0 Deployments

5.4 EB — 150X

Remote monitoring of substations

Initial PMU deployments

38 PB

3 PB

Total available data

140 TB

1950   1975   2000   2025   2040

Terabyte (TB): 1 thousand gigabytes ▶ Petabyte (PB): 1 million gigabytes ▶ Exabyte (EB): 1 billion gigabytes

Grid Data by Location

Terabyte (TB): 1 thousand gigabytes ▶ Petabyte (PB): 1 million gigabytes ▶ Exabyte (EB): 1 billion gigabytes

Legend: ■ 2030 ■ 2040

Location	2030	2040
Generation	103 PB	1.6 EB
Transmission		
Substation	873 PB	7.9 EB
Distribution		
Metering	1.8 EB	8.4 EB
Distributed generation	1.7 EB	5.5 EB
Ev, chargers, internet of things	1 EB	12 EB

Electric Sector Data Strategy Goals

Address Interoperability challenges
Data standardization

Enable Data standardization and interoperability solutions/platforms

Develop functional Data ecosystems with flywheel effects – Data Spaces/Open Data Portals/Platform

- Data discovery
- Data lineage
- Data governance
- Collaboration
- Glossary
- Open APIs

Define and document data strategy for the Electric Sector

Initial whitepaper on data strategy created

Privacy preserving techniques to make more data more open and available to the right people in a secure manner at the right place and at the right time
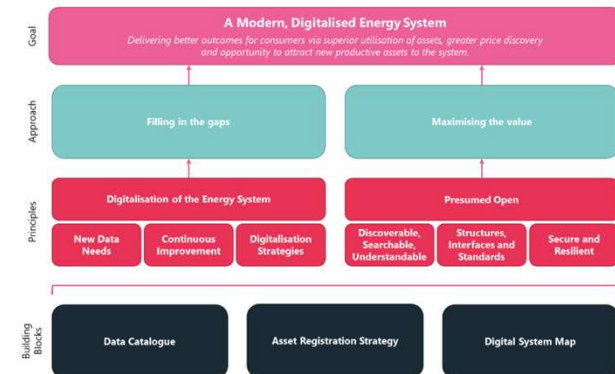
Advanced Grid Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

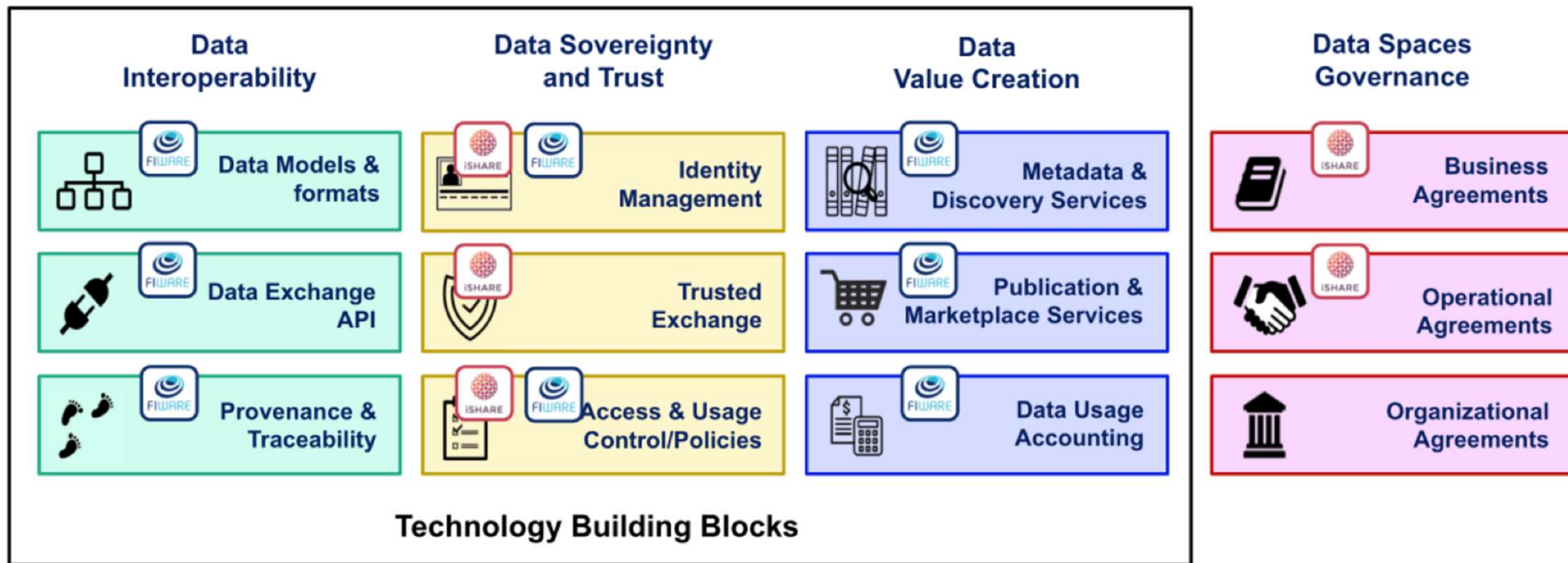## EU Energy Data Strategy

Australia and UK Energy Data Strategy

https://www.datocms-assets.com/32572/1657767015-esb-data-strategy-final-reccomendations-july-2021.pdf
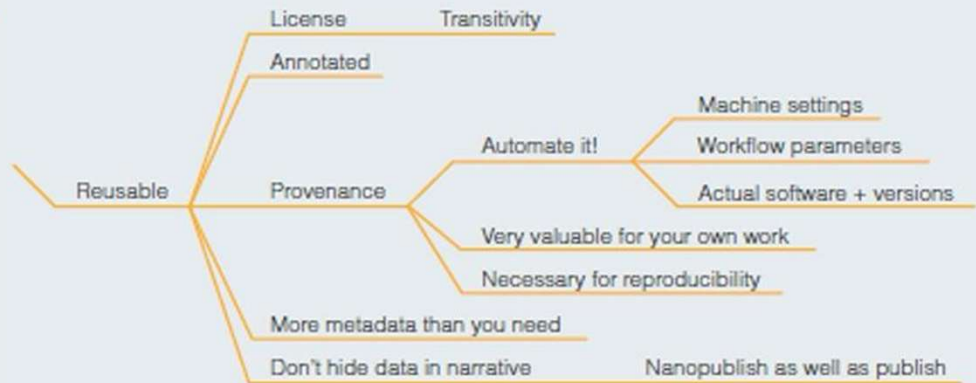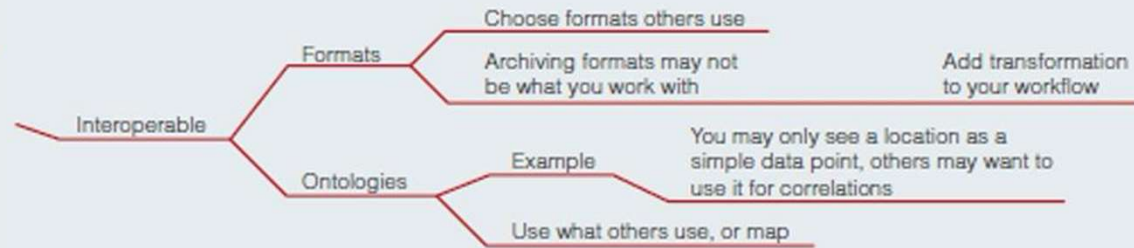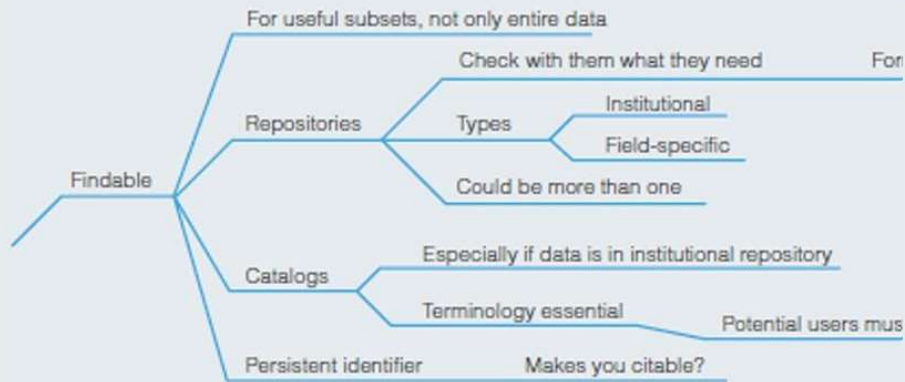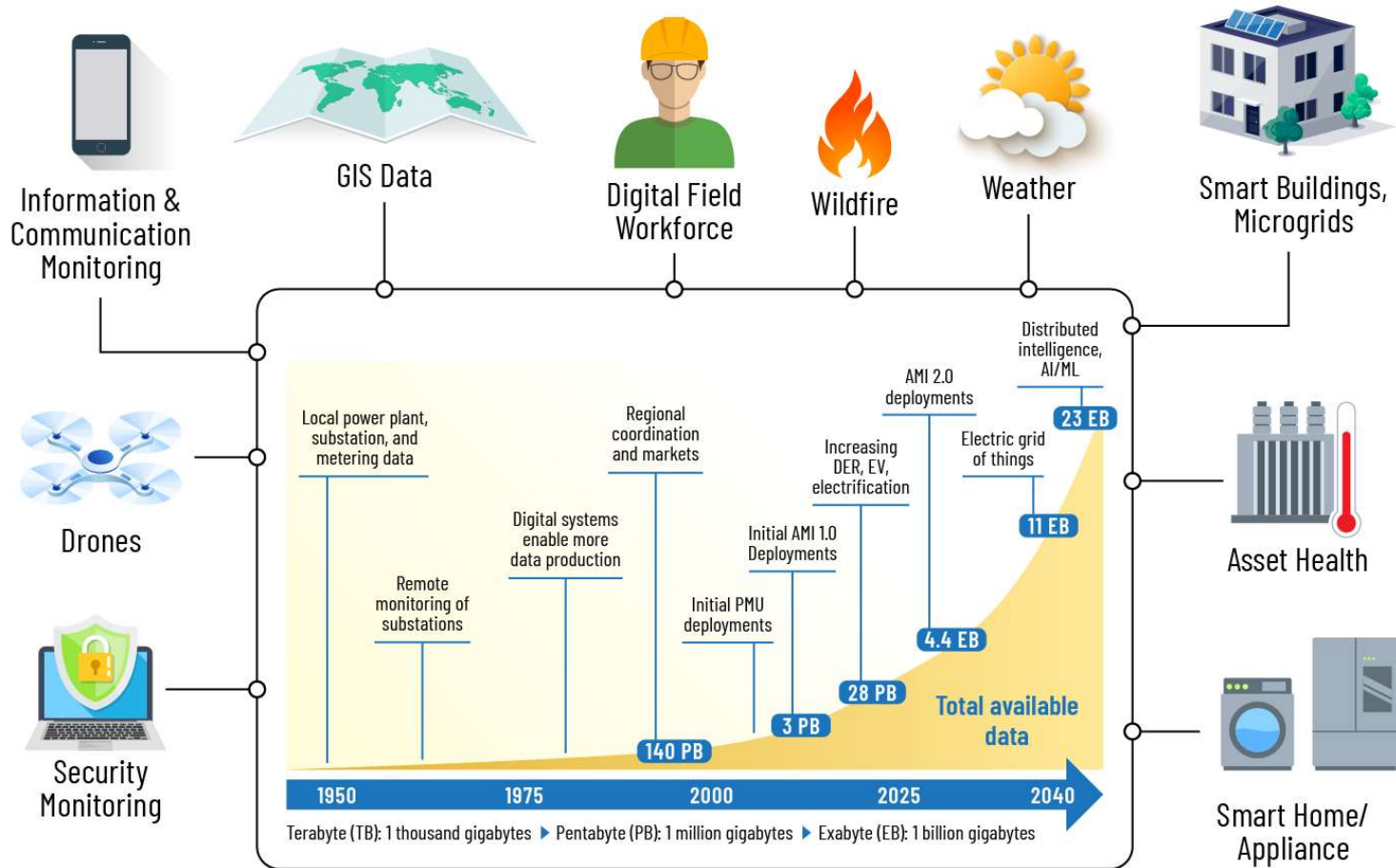
Energy Data spaces

The Energy Data Space can enable the digitalization of the energy transition by providing an architecture to make data available to increase the efficiency in asset and system operation

https://www.opendei.eu/wp-content/uploads/2022/10/OPEN-DEI-Energy-Data-Spaces-EHM-v1.07.pdf

SYNERGY

Societal Data Management according to the FAIR principles

Findable
- For useful subsets, not only entire data
- Repositories
  - Check with them what they need — For...
  - Types
    - Institutional
    - Field-specific
  - Could be more than one
- Catalogs
  - Especially if data is in institutional repository
  - Terminology essential
    - Potential users mus...
- Persistent identifier — Makes you citable?

Accessible
- Quality of the repository
  - Longevity
  - Your own web site may not be a good repository
- Data access committee — Conditions?
- No unlimited embargo period
  - Paper
  - Patent
- Quality of your own center
  - Example: think through what wi... happen if someone deletes the wrong file?
  - Example: What will happen if the whole system fails

Interoperable
- Formats
  - Choose formats others use
  - Archiving formats may not be what you work with — Add transformation to your workflow
- Ontologies
  - Example — You may only see a location as a simple data point, others may want to use it for correlations
  - Use what others use, or map

Reusable
- License — Transitivity
- Annotated
- Provenance
  - Automate it!
    - Machine settings
    - Workflow parameters
    - Actual software + versions
  - Very valuable for your own work
  - Necessary for reproducibility
- More metadata than you need
- Don't hide data in narrative — Nanopublish as well as publish

Big Data? Surely you jest!

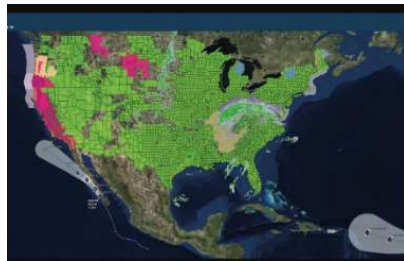# Sector Strategy & Secure Data Portal

Data Portals Today

- **EIA**

- **Smart Meter Texas**

- **EAGLE-I**

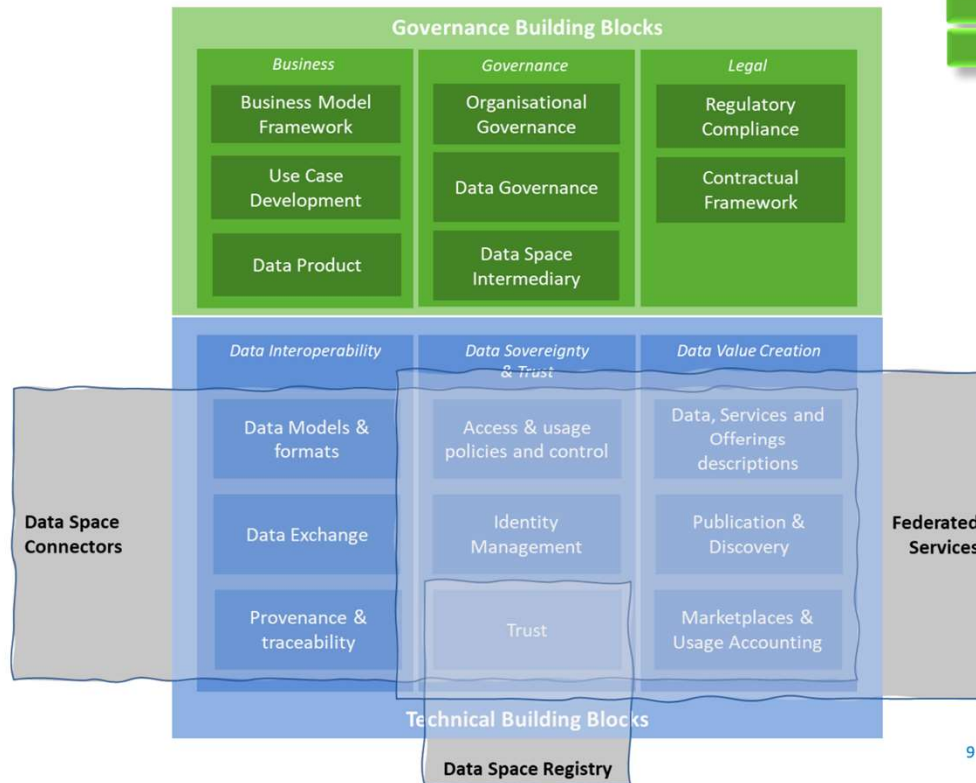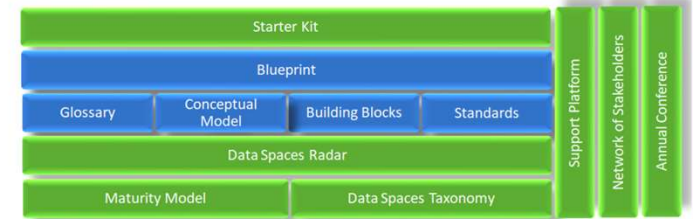- **Wood Macenzie (formerly Genscape)**
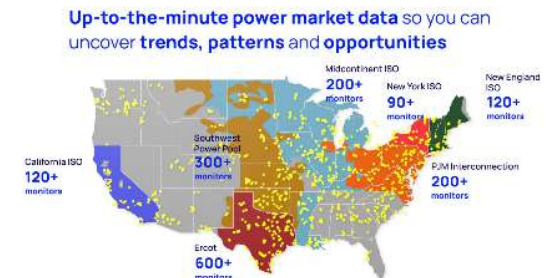
Building Blocks Taxonomy

DSSC Delivery Plan

DATA SPACES SUPPORT CENTRE

Starter Kit				
Blueprint				
Glossary	Conceptual Model	Building Blocks	Standards	
Data Spaces Radar				
Maturity Model		Data Spaces Taxonomy		

Support Platform · Network of Stakeholders · Annual Conference

Governance Building Blocks

Business	Governance	Legal
Business Model Framework	Organisational Governance	Regulatory Compliance
Use Case Development	Data Governance	Contractual Framework
Data Product	Data Space Intermediary	

Technical Building Blocks

Data Space Connectors

Data Interoperability	Data Sovereignty & Trust	Data Value Creation
Data Models & formats	Access & usage policies and control	Data, Services and Offerings descriptions
Data Exchange	Identity Management	Publication & Discovery
Provenance & traceability	Trust	Marketplaces & Usage Accounting

Federated Services

Data Space Registry

9

Develop and Deploy a Data Catalog Solution for Data Discovery and Transparency with Data Governance

Goal: Implement a data catalog solution that provides capabilities like standardization and interoperability, while handling data privacy and confidentiality.

- **Current Status:**

- Data portal is available https://openenergyhub.ornl.gov/login/
  - Live for Beta testing on May 20th, 2024.
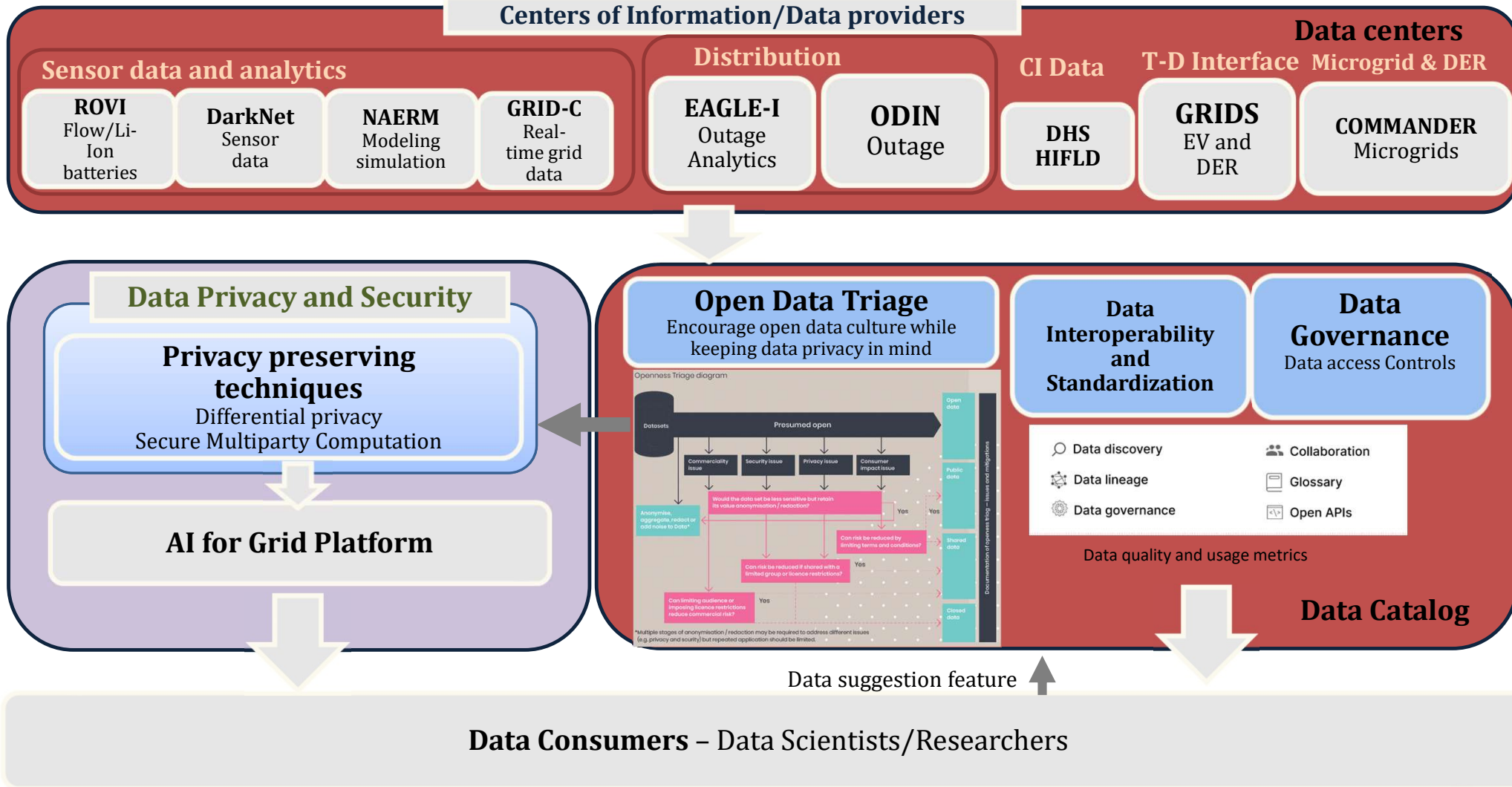 - Added 50+ datasets (data and/or metadata)
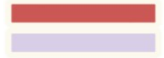 - Linking multiple ORNL project datasets from multiple sponsors

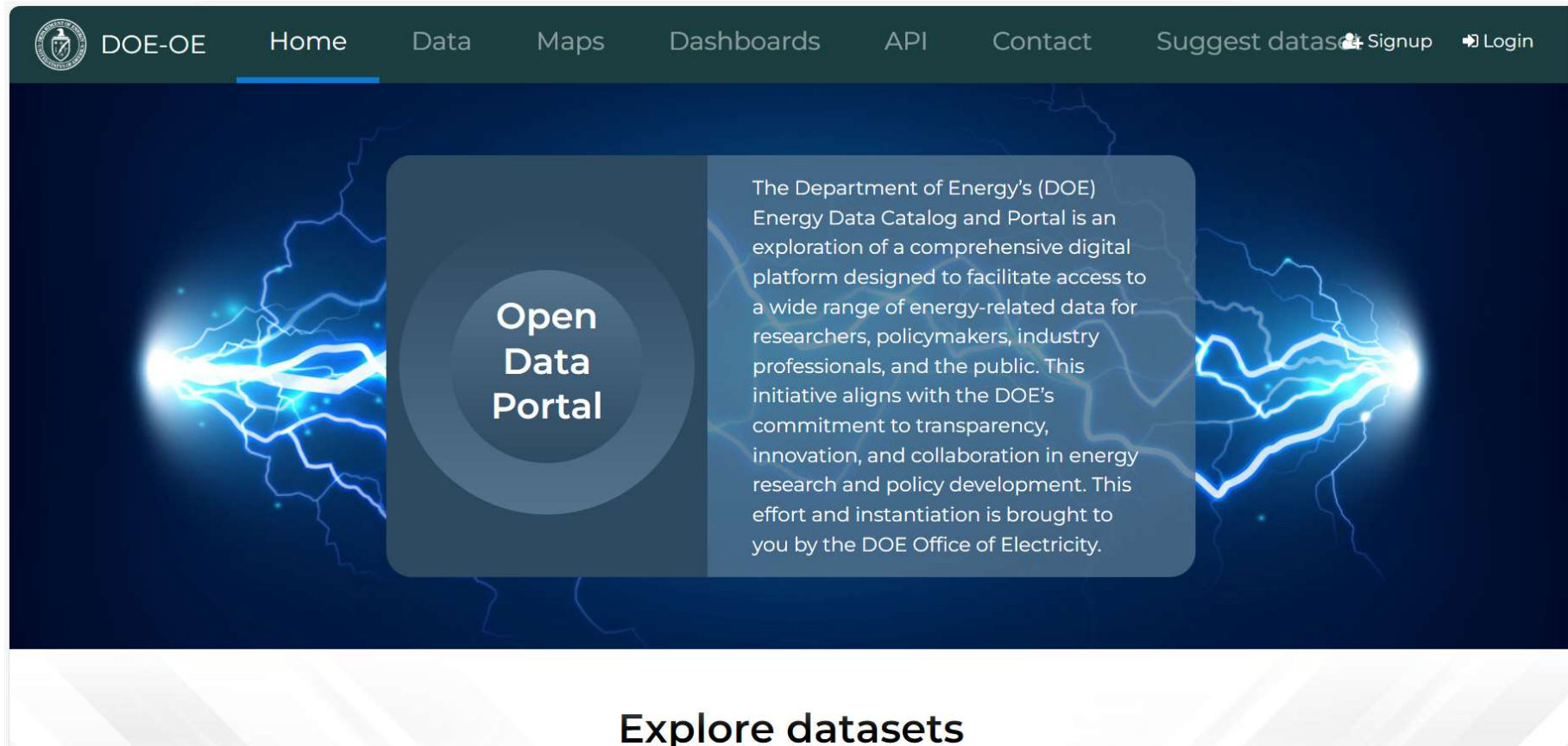# High Level Data Framework for OE projects

Implemented
In Progress

Centers of Information/Data providers

Data centers

Sensor data and analytics

ROVI
Flow/Li-Ion batteries

DarkNet
Sensor data

NAERM
Modeling simulation

GRID-C
Real-time grid data

Distribution

EAGLE-I
Outage Analytics

ODIN
Outage

CI Data

DHS HIFLD

T-D Interface

GRIDS
EV and DER

Microgrid & DER

COMMANDER
Microgrids

Data Privacy and Security

Privacy preserving techniques
Differential privacy
Secure Multiparty Computation

AI for Grid Platform

Open Data Triage
Encourage open data culture while keeping data privacy in mind

Data Interoperability and Standardization

Data Governance
Data access Controls

Openness Triage diagram

Datasets

Presumed open

Open data

Public data

Commerciality issue

Security issue

Privacy issue

Consumer impact issue

Would the data set be less sensitive but retain its value anonymisation / redaction?

Anonymise, aggregate, redact or add noise to Data*

Can risk be reduced by limiting terms and conditions?

Shared data

Can risk be reduced if shared with a limited group or licence restrictions?

Can limiting audience or imposing licence restrictions reduce commercial risk?

Yes

Closed data

*Multiple stages of anonymisation / redaction may be required to address different issues (e.g. privacy and security) but repeated application should be limited.

Data discovery

Data lineage

Data governance

Collaboration

Glossary

Open APIs

Data quality and usage metrics

Data Catalog

Data suggestion feature

Data Consumers – Data Scientists/Researchers

US Electric Sector Data Strategy – Secure Data Portal Pilot

Metadata Matters!

Metadata

- Commonly defined as "Data about Data."

- NISO: **Structured** information that **describes, explains, locates**, or otherwise makes it easier to **retrieve, use, or manage** information

- Concerned primarily with `citation level (descriptive)` metadata and `rights level (administrative)` metadata as applied to the data portal cards

Current status:

- Ofgem published its [decision on updates](#) to it's Data Best Practices Guidance v1 on August, 7 2023. It sets a bar for data sharing in the wider energy sector and will almost certainly influence guidance in other regulated sectors.

- A particularly interesting update was the specification of **Dublin Core as the minimum metadata standard**, with a requirement for it to be implemented by August 6, 2024.

OAK RIDGE
National Laboratory

Aligning with Standards and Taxonomies

Notes: Use either the Infrastructure Data Taxonomy (IDT) *Chapter 5, Energy*; or the IRENA Energy Taxonomy when constructing subjects, paying particular attention to their respective implementation guidelines. For instance, per the IDT:

In describing a given infrastructure asset, the general convention **is to use the greatest level of detail possible.** If there is insufficient information to identify an asset at a greater level of detail, a less-detailed level may be used.

Example:

`dc:subject` Supercapacitors

`dc:subject` Transmission Line

Taxonomies allow for classification according to a pre-determined system, ensuring consistency and accuracy for information retrieval while aligning with industry best practices.

IDT Path	IDT Description	IDT ID	Cross-sector Reference	NAICS Code
5.1.2.3	**Direct Current (DC) Converter Station** Facilities that convert high voltage power between alternating current (AC) and direct current (DC) for improved efficiency in transmission.	190		221121
5.1.2.4	**Generation Dispatch and Transmission Control Center** Centers that control the dispatching of generators and the high voltage transmission system. Includes generation and transmission supervisory control and data acquisition (SCADA) systems. Includes independent system operator (ISO) facilities.	191		221121
5.1.3	**Electricity Distribution** Facilities used to transmit electricity to final consumers at lower voltages (usually <69 kV).	192		221122
5.1.3.1	**Distribution Line** Facilities that include lines, poles, transformers, and meters.	193		221122
5.1.3.2	**Distribution Substation** Connection points in the distribution system and from the distribution system to the transmission system. May include transformers, capacitors, and switchgear.	194		221122
5.1.3.3	**Distribution Control and Dispatch Center** Centers that control the lower voltage distribution system. Includes distribution supervisory control and data acquisition (SCADA) systems.	195		221122

OAK RIDGE National Laboratory

Data Review and Release/Triage (Pilot)

Goals

- Align with Review and Release (ORNL/DOE) Processes
- Draw from the UK's Open Data Triage (as has NARUC)
- Show the progress of data "openness" throughout the data lifecycle
- Closed and Restricted data sets are also allowable outcomes

Working with Legal and Privacy Office to

- Develop appropriate disclaimers
- Develop sharing policy

Current Excel based calculations (similar to UK)

Planning to operationalize processes up to the DOE level

Advanced Grid Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

Data Use & Protection

Data Users and Uses Grow; Usability Lags

- **Even within a utility, data is not easy to share and easy to use**
- **Operational coordination is increasingly multi-stakeholder**
- **Customer data access, sadly, remains a challenge**
- **Multiple third parties have legitimate grid data needs and data potentially to offer in return**
- **Interoperability remains a critical pursuit, in all futures**

Advanced Grid Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

Data Utilization and Sharing is a Hot Topic

- OE's Outage Data Initiative Nationwide

- Model Data exchange at T-D and with the Public

- Grid Transparency (Hosting Capacity etc)

- NARUC Grid Data Sharing effort

- DOE's i2X – Interconnection Innovation Exchange

- DER Registries

Advanced Grid
Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

Data Value Caveats

- **Raw data isn't gold, it's ore**
- **Full fidelity data is hardly the sole source of value, but many claim they need it**
- **Valuable data is usually the result of data cleaning and data fusion, an underappreciated value creation process**
- **Very few parties require or deserve full-fidelity data**
  - Data Dervatives and Data Fusion products are an effective path to value
- **Beware the "Velvet Rope" solution**

Advanced Grid
Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

PET's and Advanced Data Protection Methods Emerge

- **Anonymization**

- **Differential Privacy**

- **Zero Knowledge Proofs**

- **Secure Multi Party Computation**

- **Trusted Execution Environments**

- **Homomorphic Encryption**

- **Federated Learning/Distributed Analytics**

Advanced Grid
Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

Conclusion & Discussion

We can't afford to be tactical...
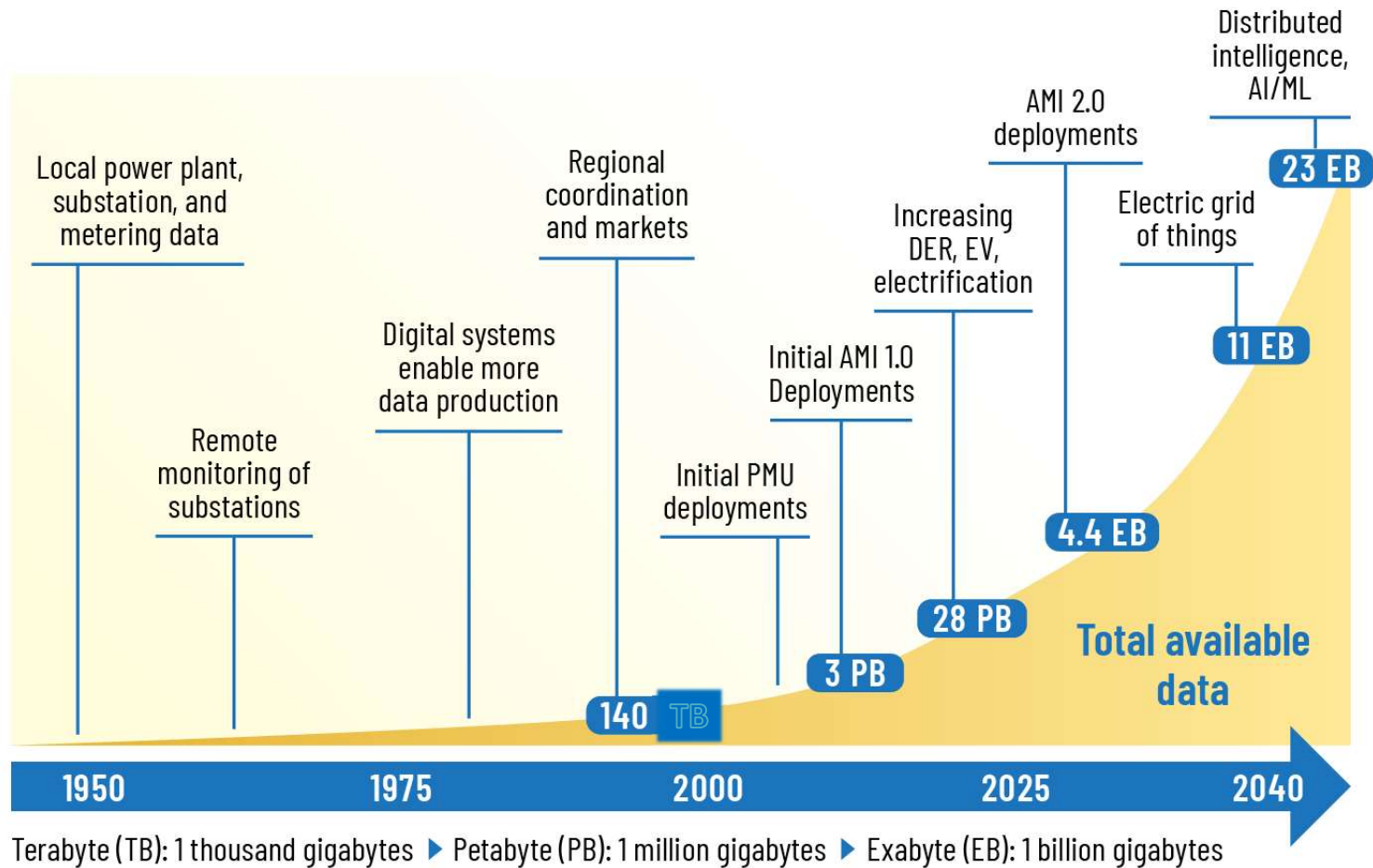
Electric Sector Data Strategy Goals

Address Interoperability challenges
Data standardization

Enable Data standardization and interoperability solutions/platforms

1

Develop functional Data ecosystems with flywheel effects – Data Spaces/Open Data Portals/Platform

2

Data discovery	Collaboration
Data lineage	Glossary
Data governance	Open APIs

Define and document data strategy for the Electric Sector

Initial whitepaper on data strategy created

3

Privacy preserving techniques to <u>make more data more open</u> and available to the right people in a secure manner at the right place and at the right time

4

Advanced Grid Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY

Take Aways

- Its not a "Data Tsunami"
- Who needs Observability, why, and do they need to create it or borrow it?
- Data could be viewed as a workforce
- Non-utility data will grow faster than utility data

Advanced Grid
Research
OFFICE OF ELECTRICITY
US DEPARTMENT OF ENERGY