Metadata as a Rosetta Stone for Managing Both Data and Storage Resources

Autonomous Data Management

STRONGBOX DATA SOLUTIONS

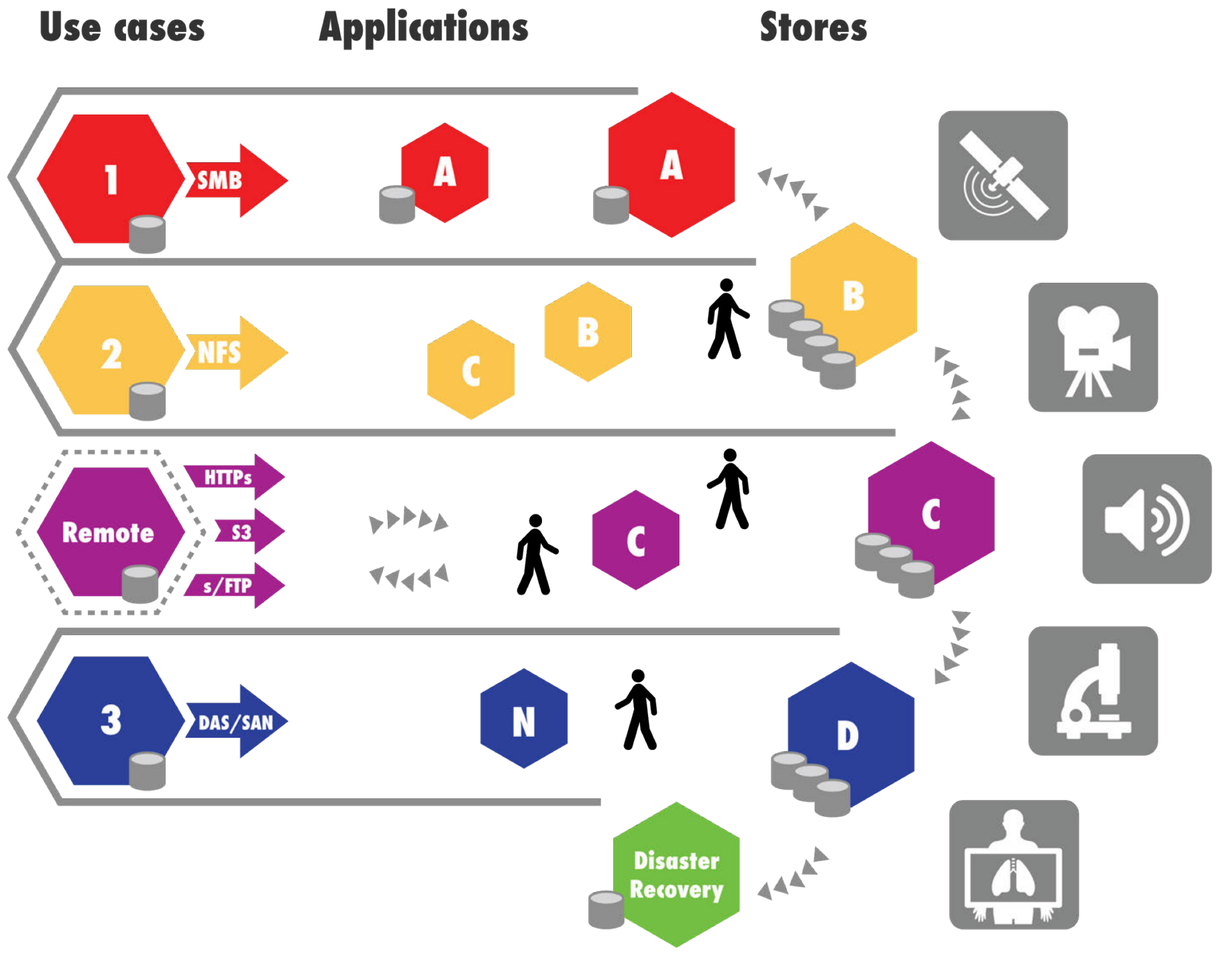
Exponential Data Growth → Growing Challenges

- Storage problem
  - What class of storage is best for current needs?
  - What about tomorrow? More tiers? New types?
  - Is there a one-size fits all storage?
  - What happens when storage needs to be replaced?
- Storage utilization problem
- Backup/protection problem
- Archive problem

STRONGBOX DATA SOLUTIONS

There Is No One-Size-Fits-All Storage Solution

- The inevitable result is:
  - Storage/vendor silos
  - Increased manual processes
  - Reduced efficiency
  - Difficult collaboration
  - Painful data migrations…
  - Increased costs



Heterogeneous Storage Environments the New Normal

Most storage environments consist of more than one vendor solution.

Cloud & multi-cloud is increasingly a part of the equation.

Most Common Problem Areas in Heterogeneous Environments

- Classifying unstructured data
- Cross-platform data migration & tiering
- Storage optimization/consolidation
- Seamless integration between on-prem & cloud
- Automating active archives
- Business continuity and DR

Traditional Industry Focus

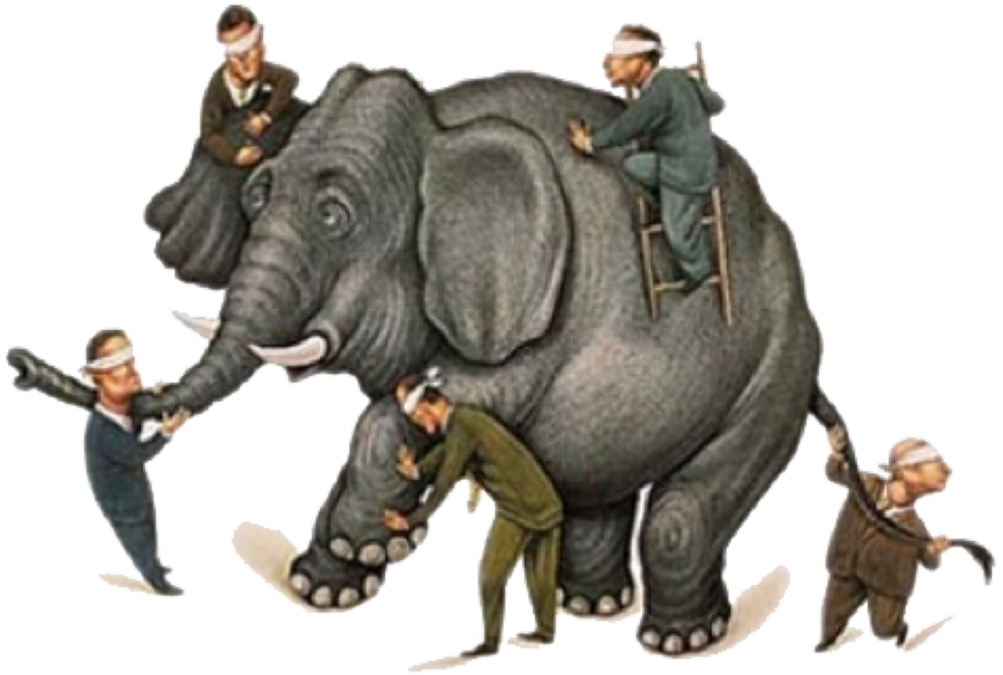
- Managing data growth was typically seen as a storage problem. Why?
  - Data explosion as a forcing function:
    - My storage is filling up? Who do I call?
- The default approach was to add more storage.
  - Costs grew so alternative storage solutions appeared:
    - HSM, active archives, object storage, cloud options...
- Costs and data volumes still grew so:
  - Data reduction strategies: deduplication solutions, etc.
    - Storage consolidation strategies… etc. etc.
- Bottom Line:
  - A storage-centric approach only solves part of the problem:
    - Storage systems alone are not designed to solve this…

To a Hammer Everything Looks Like a Nail

STRONGBOX DATA SOLUTIONS

Data Management Has Many Meanings…

- Information Lifecycle Management
- Document Management Systems
- Archive, long-term preservation
- Global namespace, or object Index
- HSM, or other storage tiering

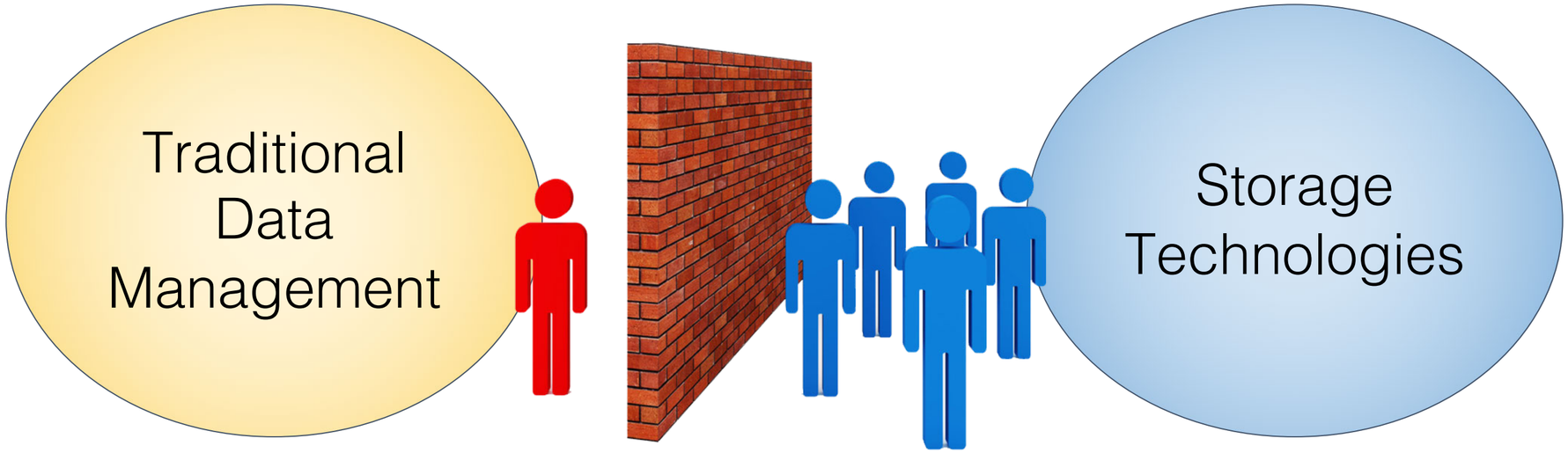


Traditional Data Management Solutions also only address part of the problem…

We've Reached a Tipping Point

Four Constants:

1. Create Data
2. Migrate Data
3. Archive Data
4. Delete Data



Traditional Strategies are Insufficient on Their Own

It’s All About the Metadata

File system metadata →

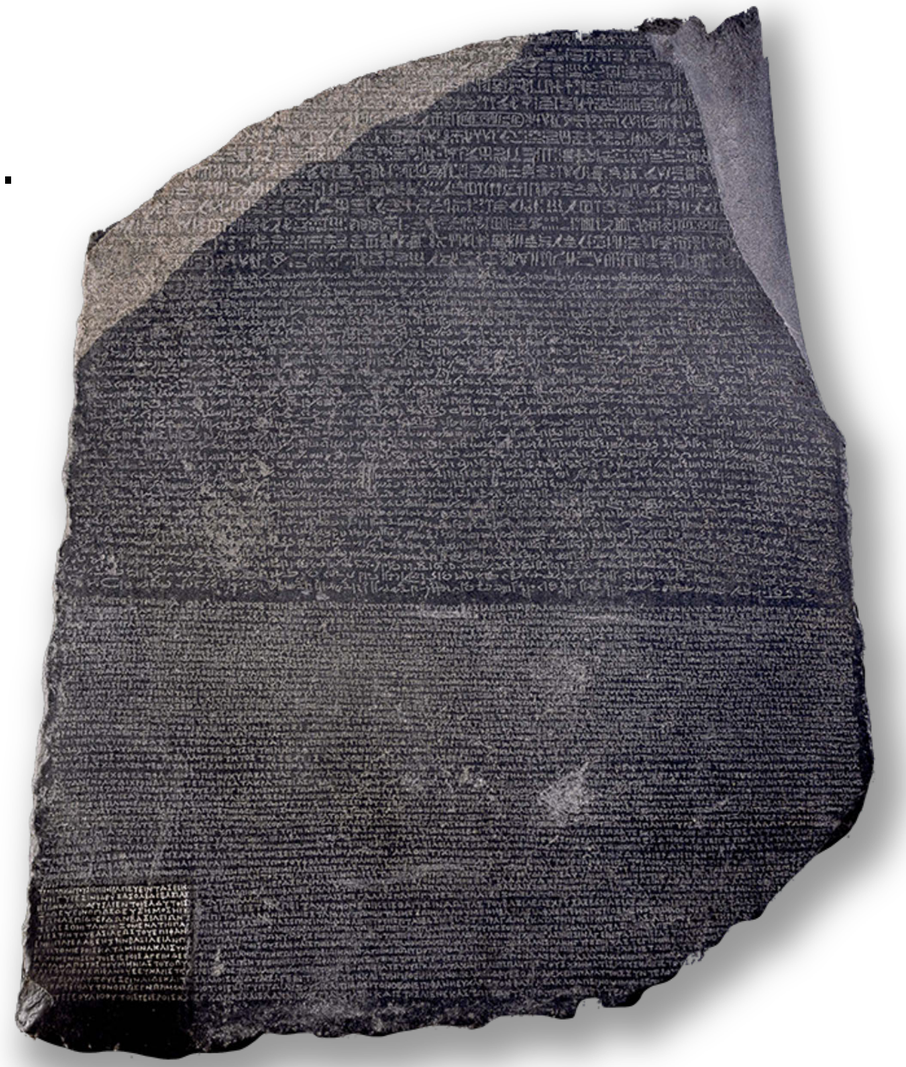
Rich application metadata →

External metadata →

User-created metadata →

Aggregated Metadata is the Rosetta Stone

- Each metadata type has different value.
- Just as data gets siloed, information about the data gets siloed.
- Correlating, and normalizing multiple metadata types empowers data-centric solutions to help address storage problems.

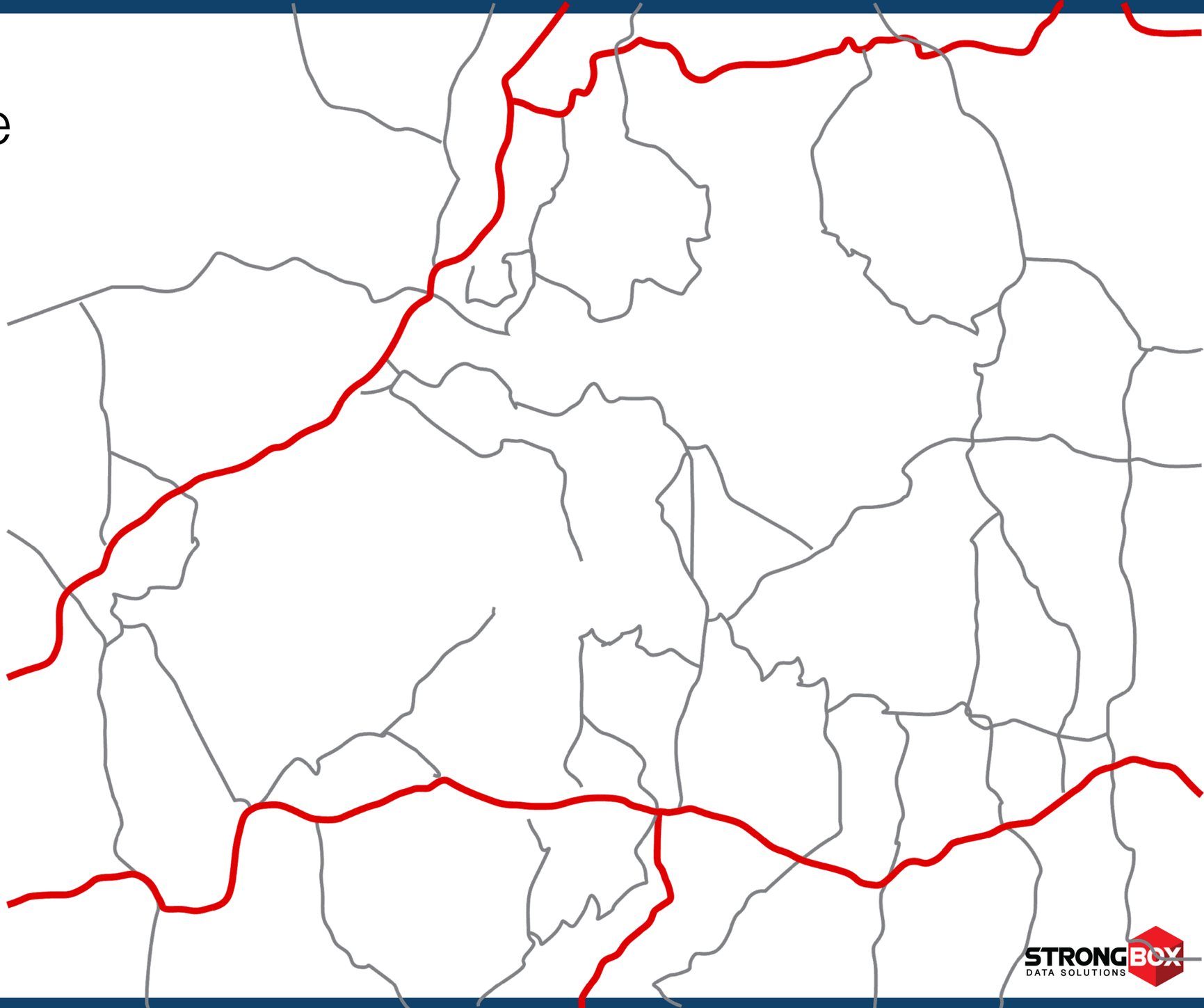


STRONGBOX DATA SOLUTIONS

Each metadata type provides only a part of the picture

Example:

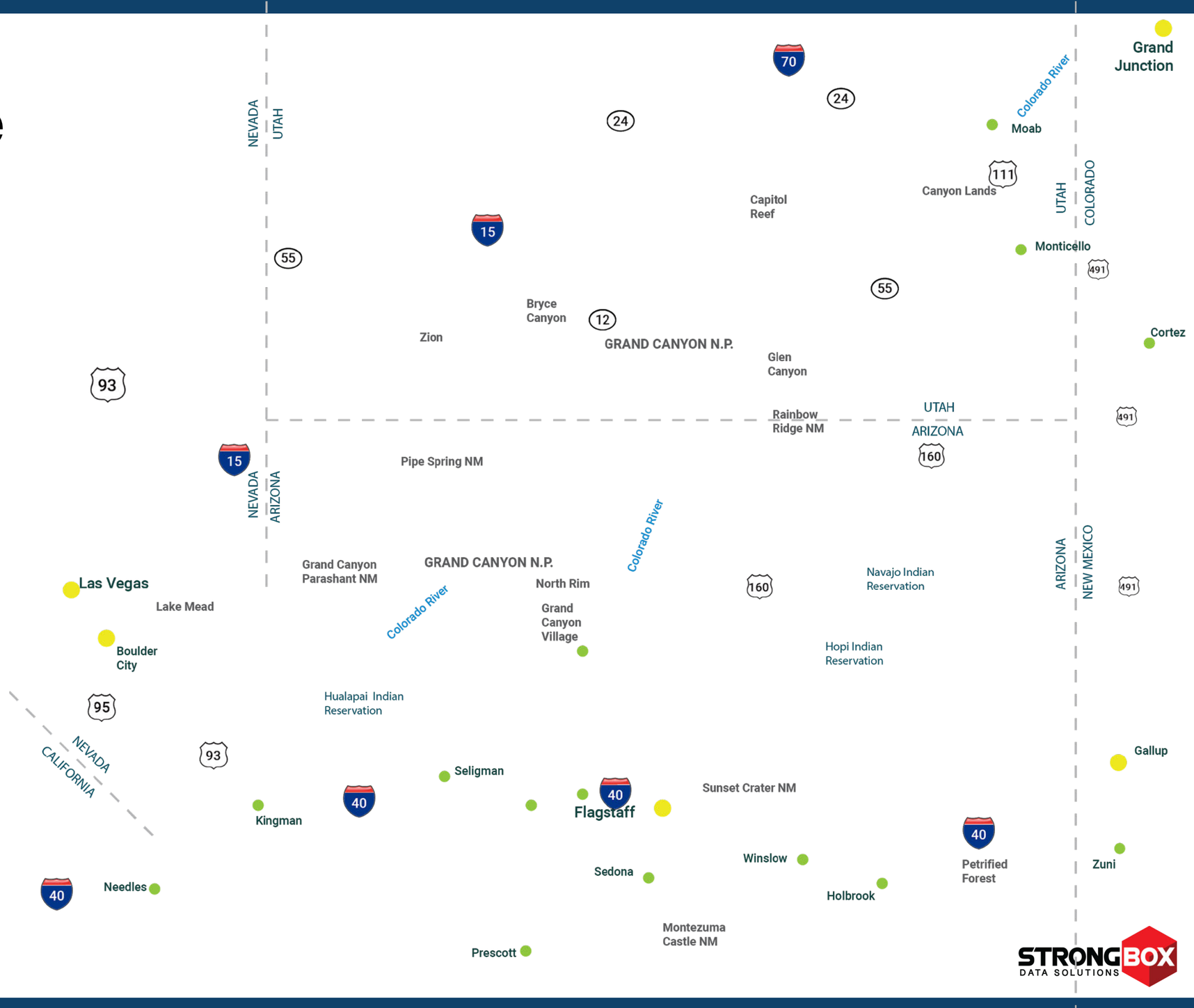
- Layers on a map
- Road lines



Each metadata type provides only a part of the picture

Example:

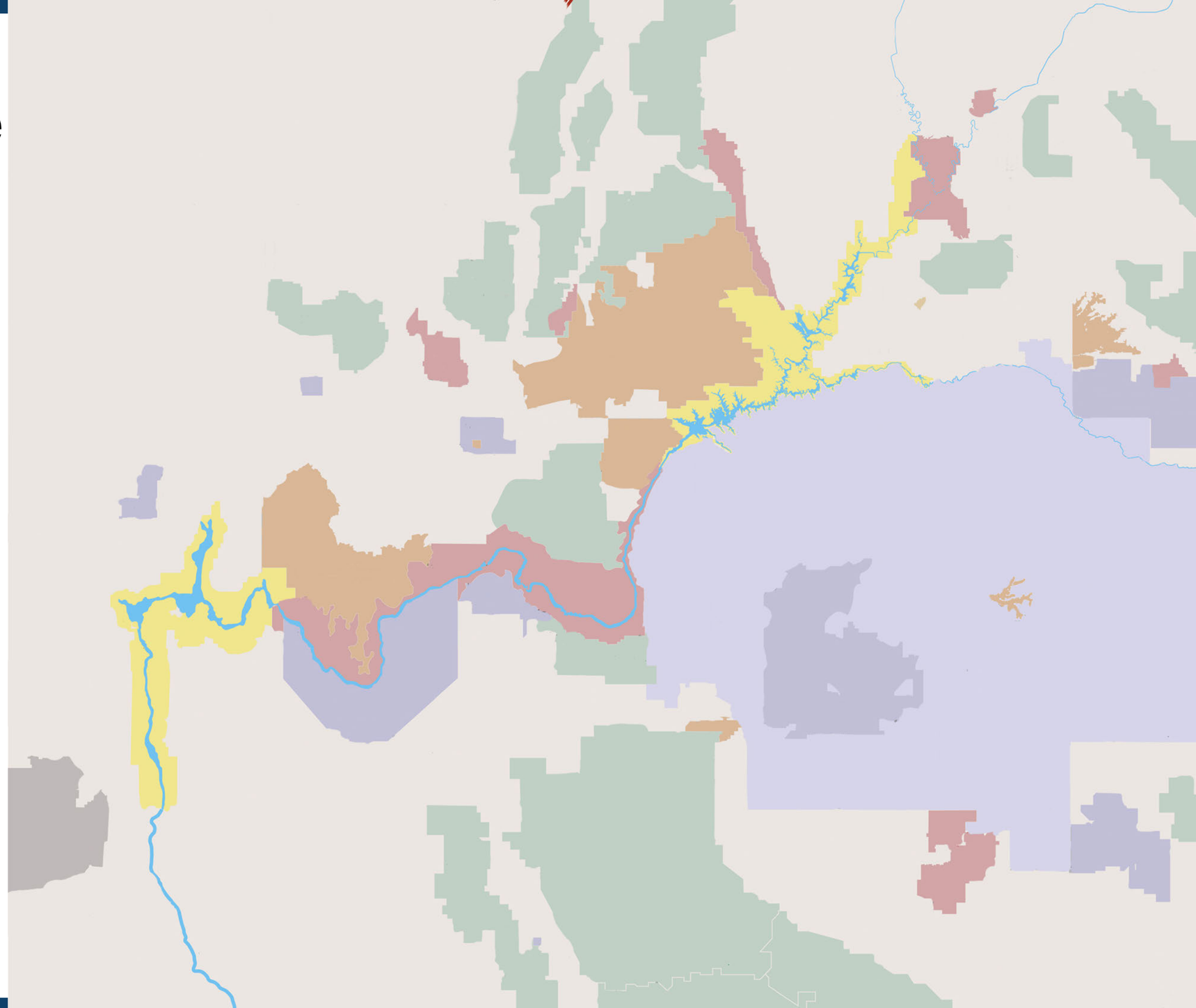
- Layers on a map
- Place names



Each metadata type provides only a part of the picture

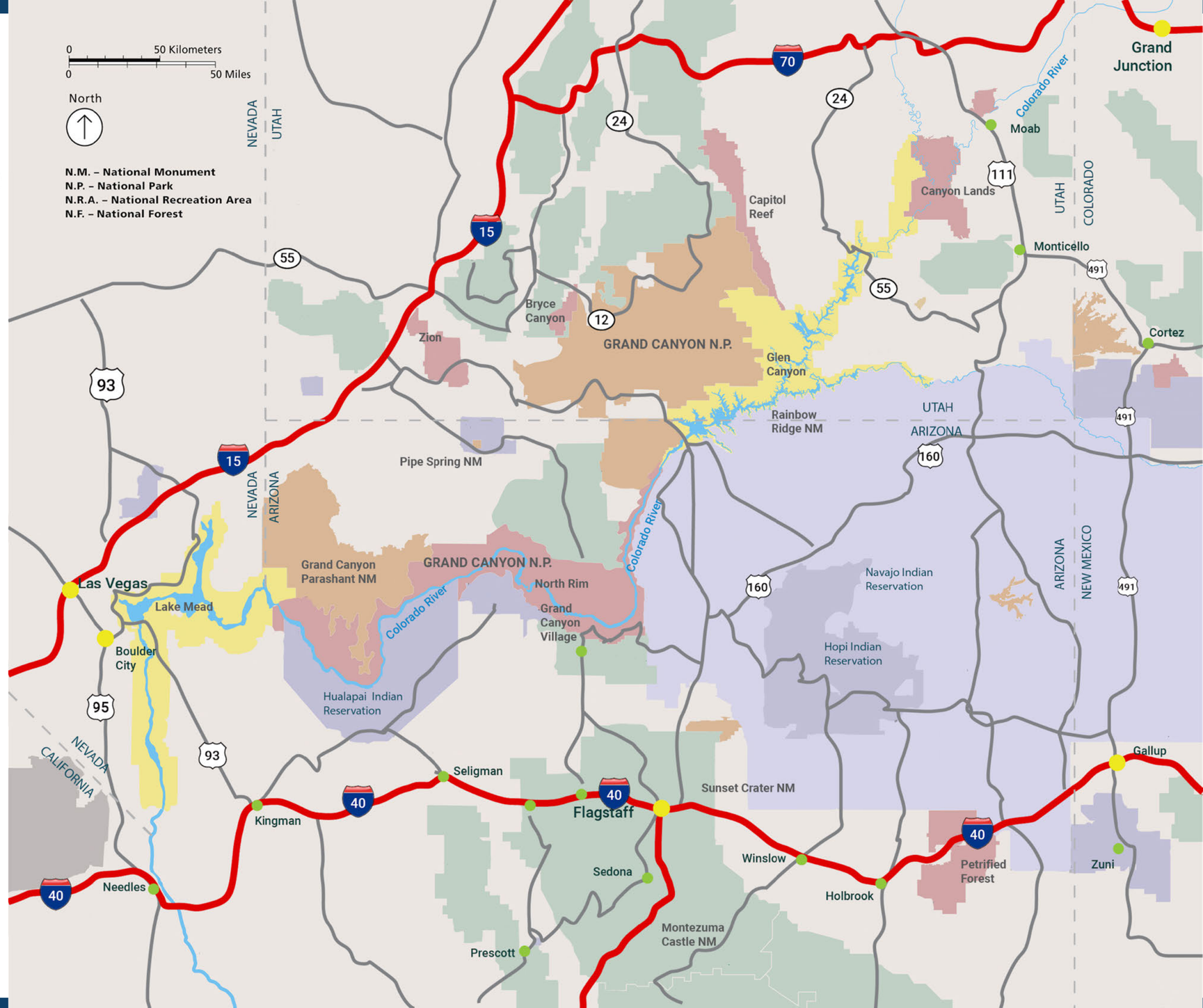
Example:

- Layers on a map
- Terrain data



All sources of information combine to complete the picture

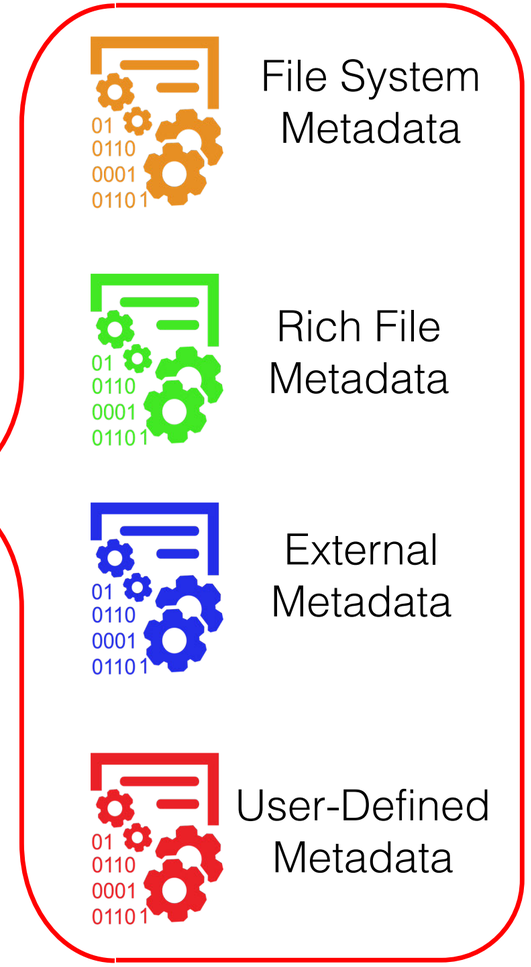
Without a complete map, one doesn't know where to go.



Complete Control of Both Data AND Storage

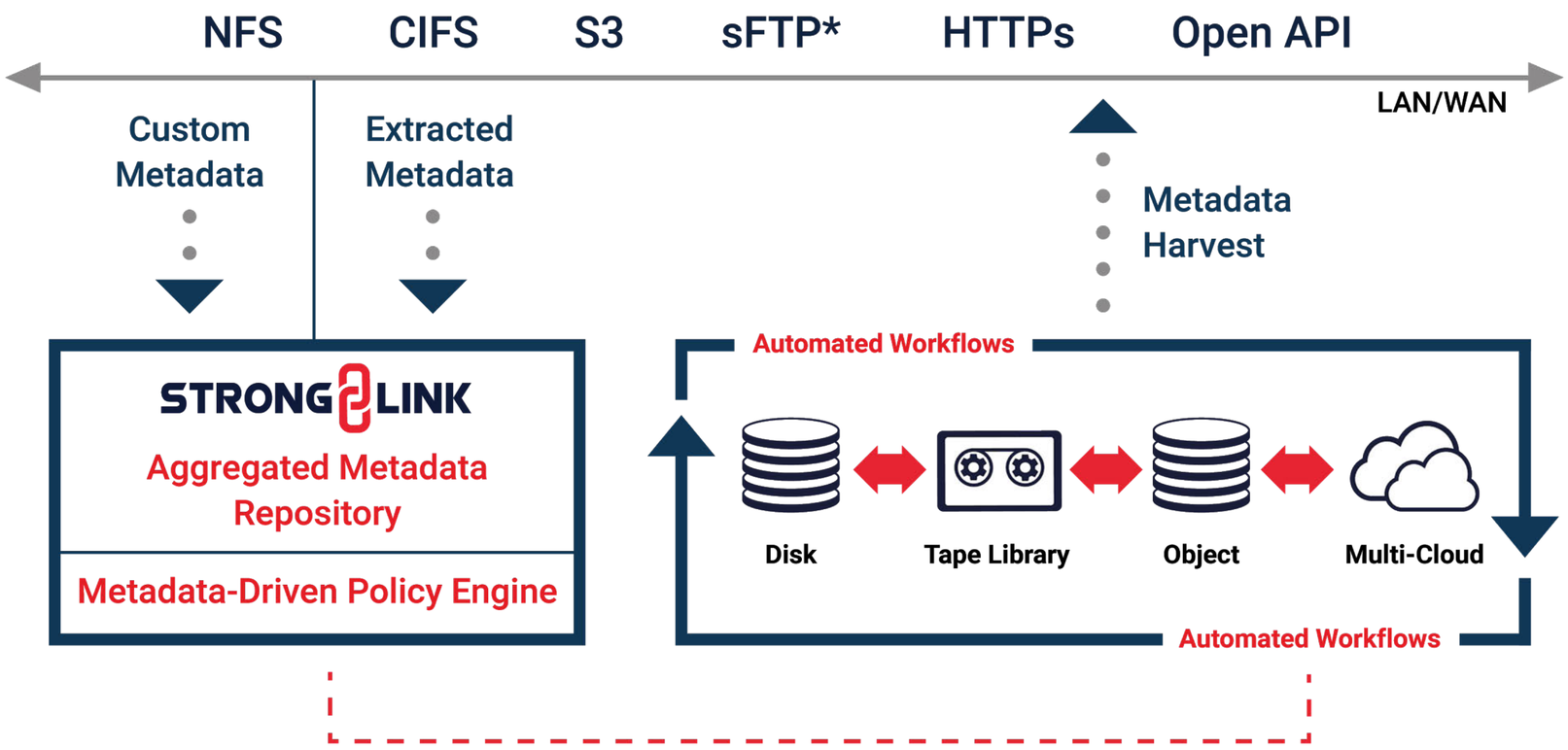
Normalizing metadata to drive policy-based management.

- Automatic data classification & global query
- Bridge heterogeneous, multi-vendor storage.
- Automate policy-based workflows, data migration, data preservation.
- Reduce complexity & costs.



Empower people to better use their data, not waste time wrangling it.

Aggregated Metadata Powers Data Automation

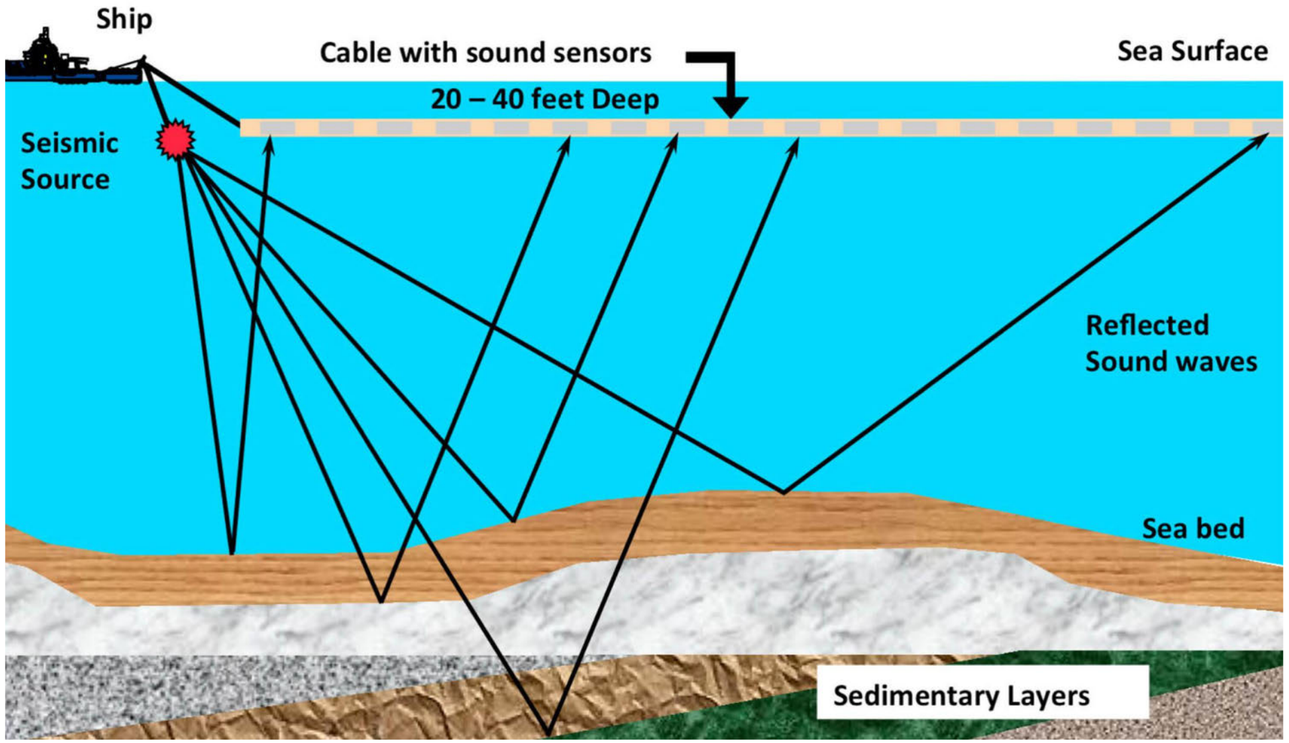


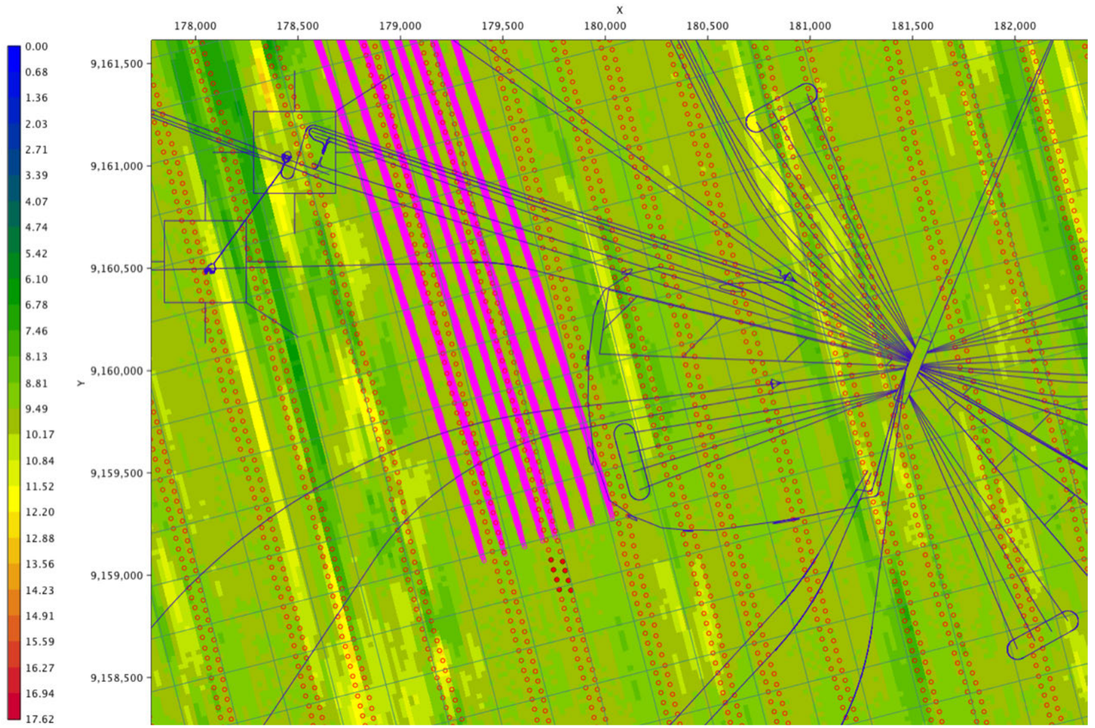
- **Metadata is harvested from multiple data sources**
  - File System metadata
  - Extract rich header metadata
  - 3rd party metadata
  - Users can create their own custom metadata
- Files don't have to be moved to harvest metadata, but file can be migrated at anytime, metadata always stays with the file
- **Open architecture**
  - Normalize metadata w/o transformation
  - Original schemas retained

STRONGBOX DATA SOLUTIONS

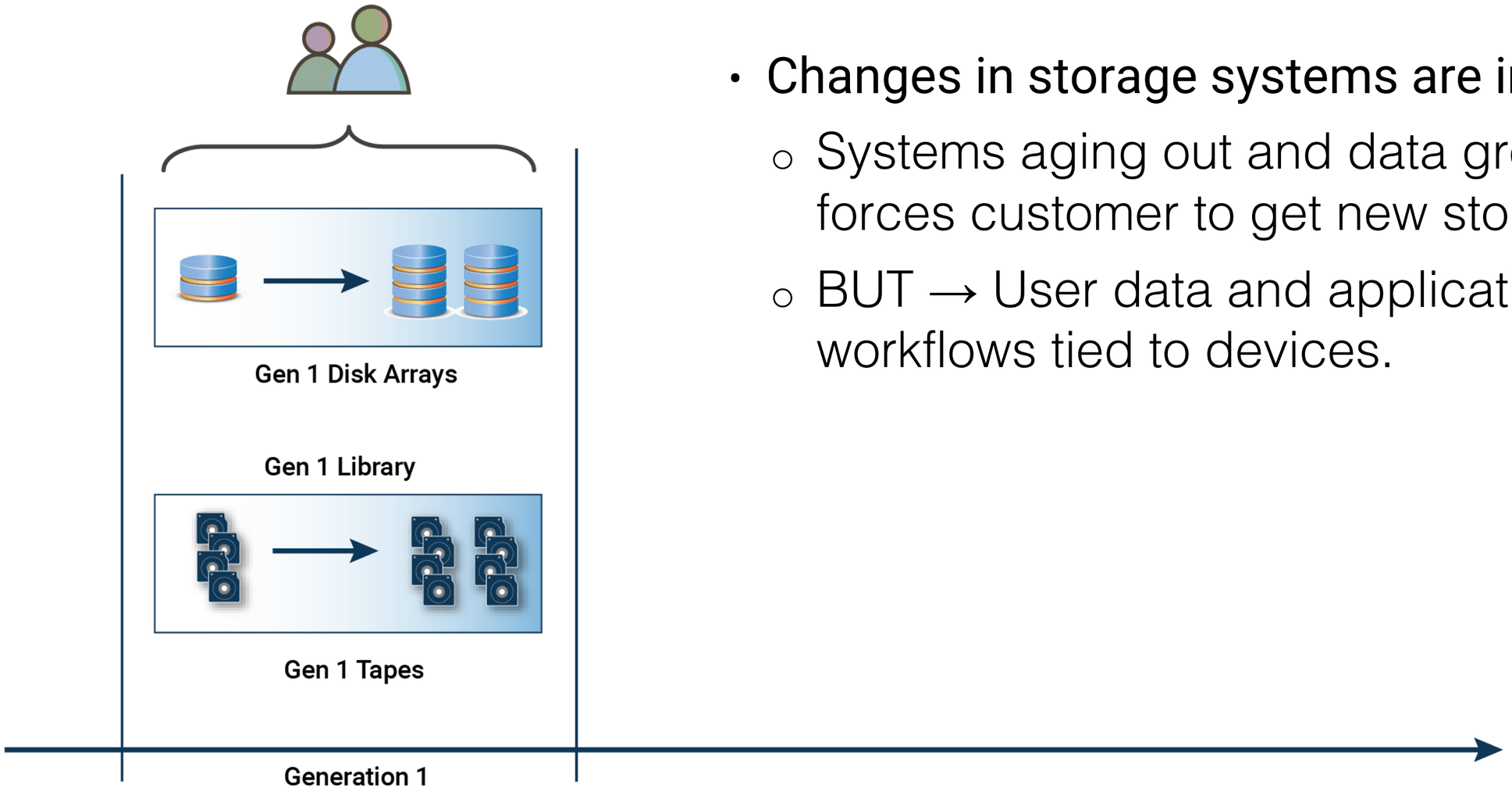
Real-World Use Case

- Oil & Gas company with many PBs of Seismic Data.
  - Data lasts longer than the storage platforms that house it, and users need continual access over generations of infrastructure.



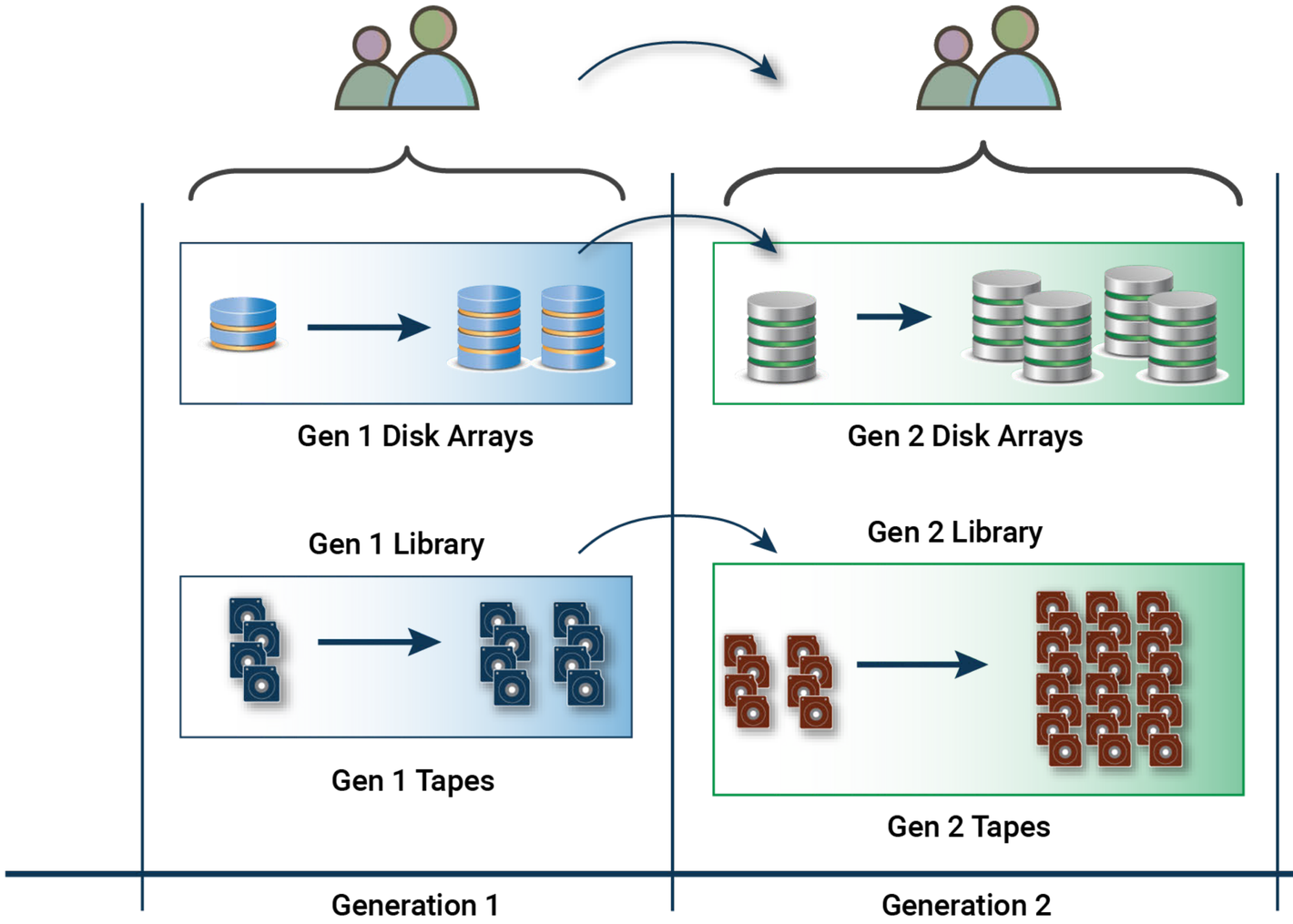


The Problem of Disruptive Data Migration/Tech Refresh



- **Changes in storage systems are inevitable.**
  - Systems aging out and data growth forces customer to get new storage.
  - BUT → User data and application workflows tied to devices.

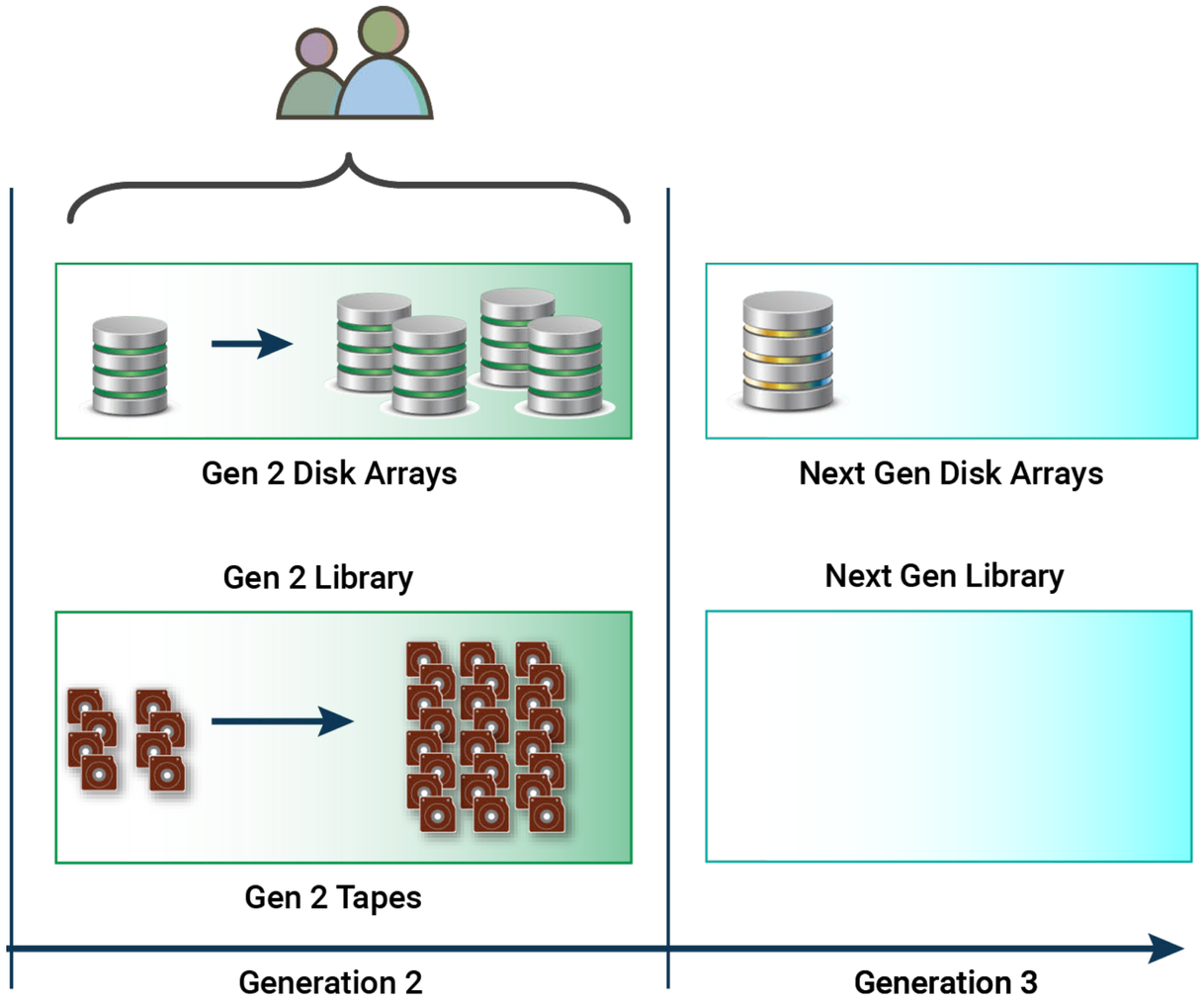
The Problem of Disruptive Data Migration/Tech Refresh



- Cutting over to new systems required wholesale migration.
  - Heavy lifting for IT staff
  - Planned outages, and significant disruption to user access
  - Extremely expensive:
    - Managing cost of both old and new systems during lengthy migration process.

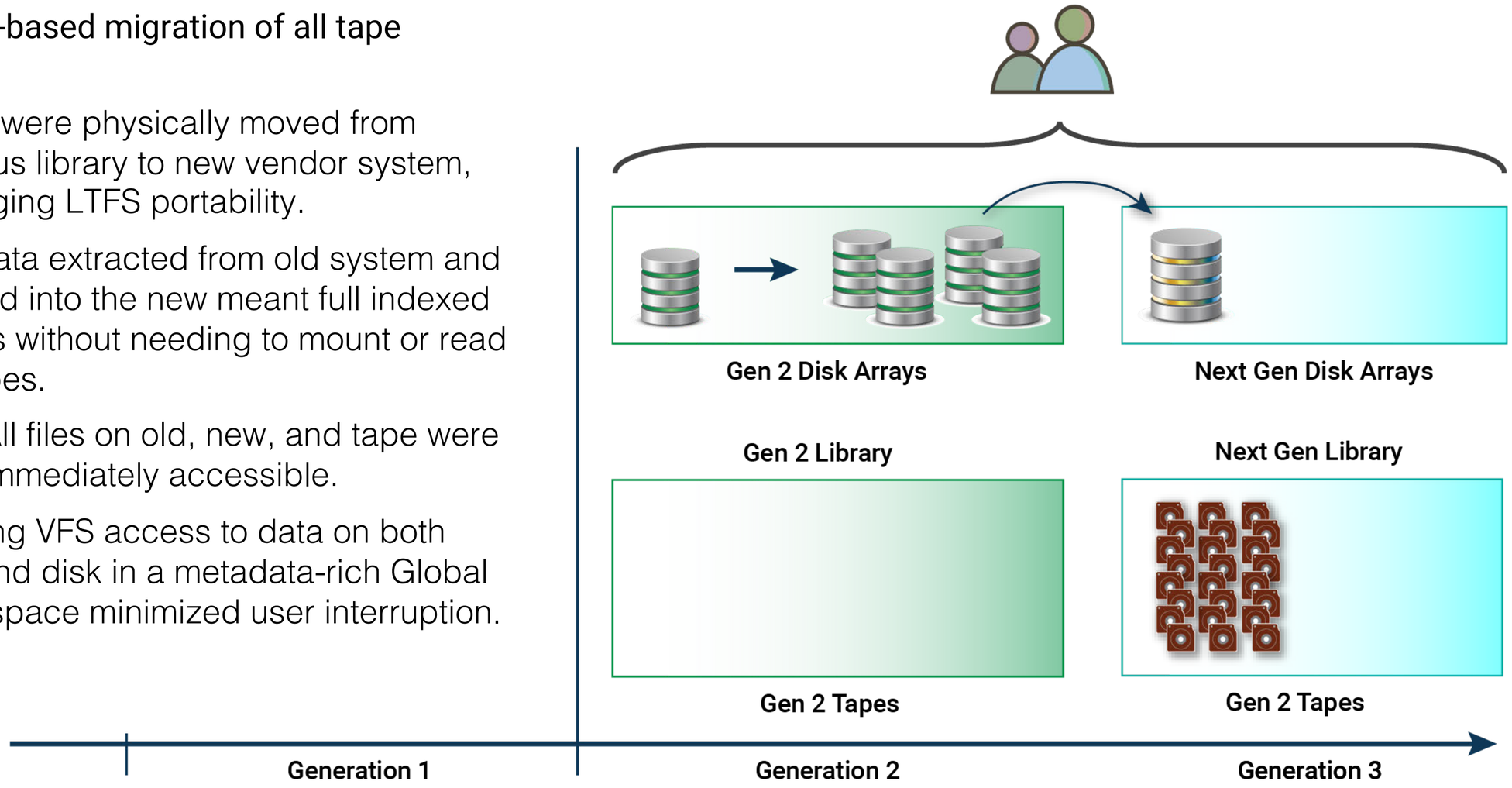
Data Has Grown, and So Did the Problems

- Customer wanted to change storage and library vendors, quickly and with minimal disruption.
- Customer could not afford the time and cost of their previous experience.
  - How migrate to a new system without mounting and reading all the tapes?
  - Can't afford downtime, or to maintain dual systems during migration.
  - Need to quickly reduce DC footprint.



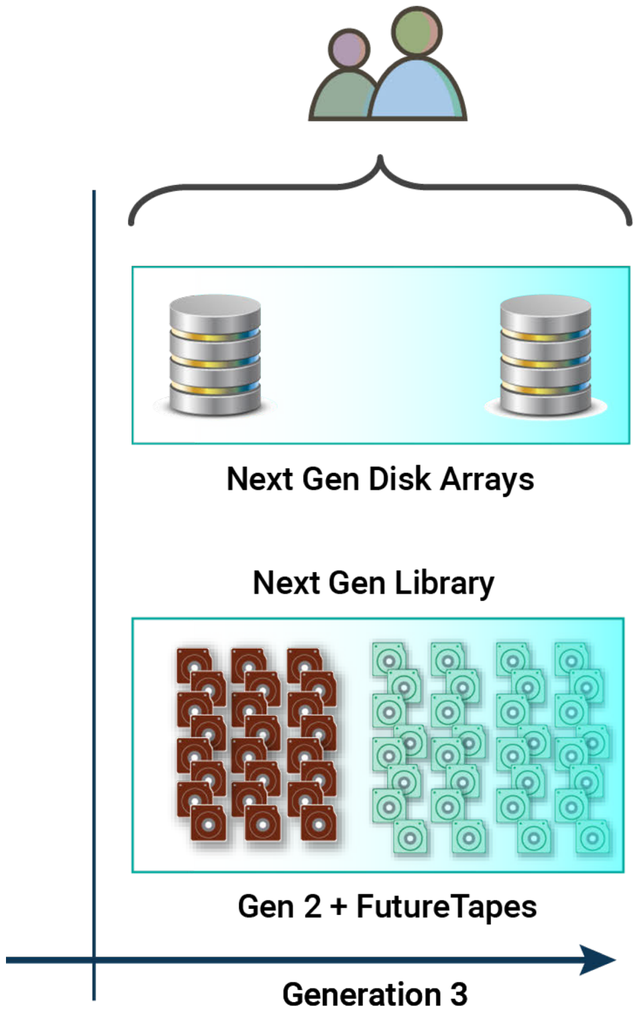
Metadata-driven Seamless Data Migration

- Metadata-based migration of all tape content.
  - Tapes were physically moved from previous library to new vendor system, leveraging LTFS portability.
  - Metadata extracted from old system and inserted into the new meant full indexed access without needing to mount or read the tapes.
    - All files on old, new, and tape were immediately accessible.
  - Creating VFS access to data on both tape and disk in a metadata-rich Global Namespace minimized user interruption.

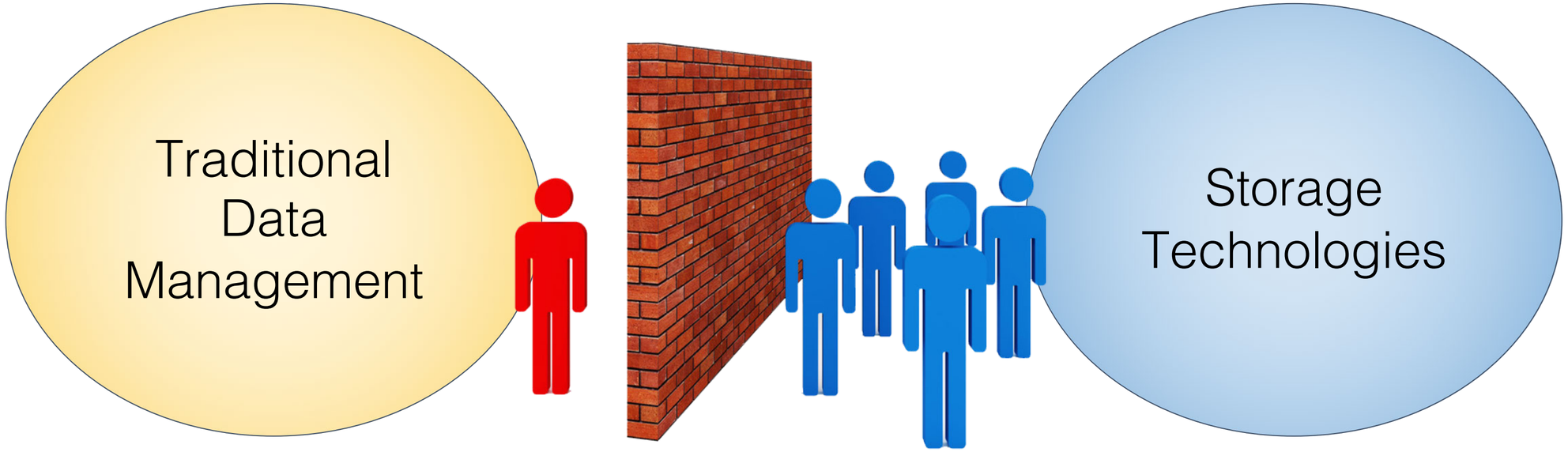


Metadata-Enabled Migration Solved the Problem

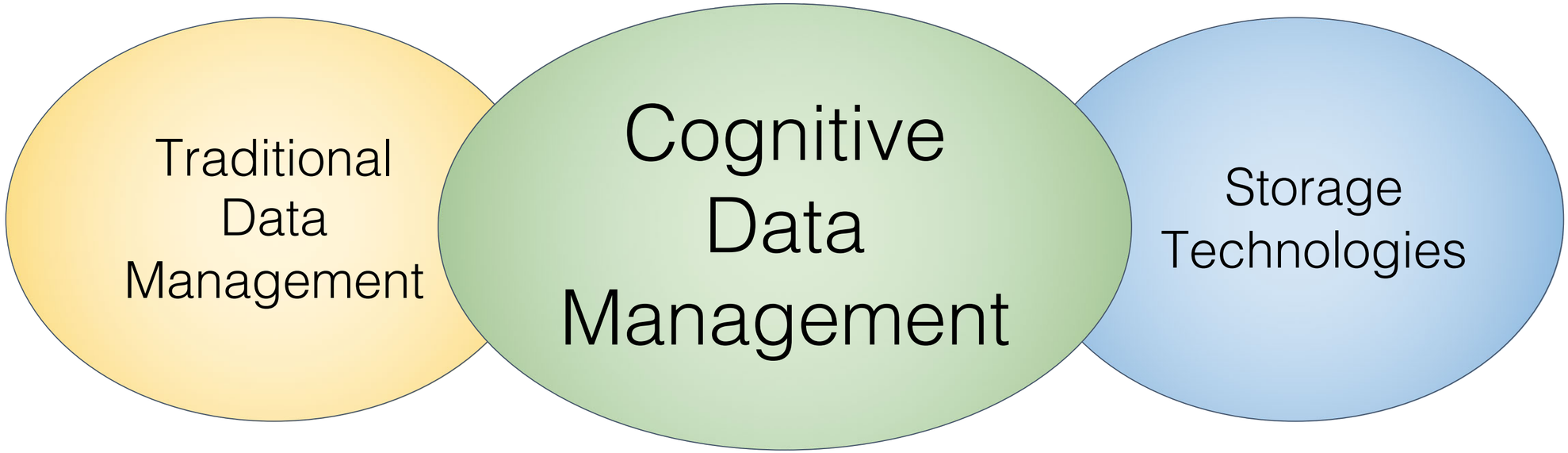
- Old system was retired quickly, minimizing cost, and freeing DC footprint.
- The new environment can scale without limitation.
  - Allows future growth with even incompatible primary storage types when needed, without interrupting user access.
  - Leveraging library vendor's future-proofed support for next gen LTO, the system expands seamlessly.
  - Rich aggregated metadata on all digital assets and storage, enables automation to migrate, tier, create reports, automate workflows, etc.
  - Creates richer, data-driven global control over storage resources and data placement.



This Emerging Need Drives the Market Trend



This Emerging Need Drives the Market Trend



- Global CDM market was valued at $466.8 million in 2017
- Expected to reach $1.47 billion by 2023,
  - Compound Annual Growth Rate (CAGR) of 20.9% during the forecast period.

STRONGBOX DATA SOLUTIONS

STRONGLINK

Automation for
Heterogeneous Storage & Data Management

FUJIFILM

STRONGBOX DATA SOLUTIONS