

Building a Cognitive Data Management Practice

(And Why Doing So is Suddenly So Important)

It's About Your Data, Not Your Storage.

The  **cernity** Suite Powered by
 **FUJIFILM** &  **STRONGBOX**
DATA SOLUTIONS

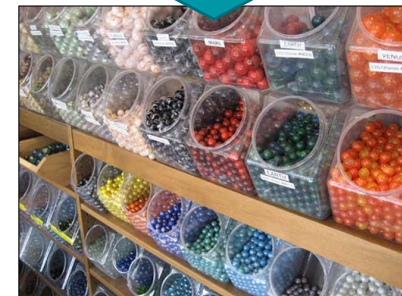
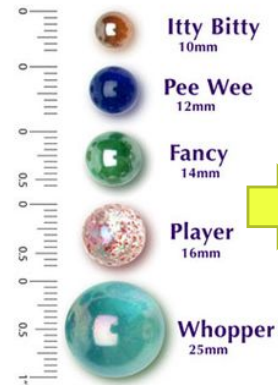
David Cerf

dc erf@Strongboxdata.com

What is Cognitive Data Management?

(And Why Should I Care?)

- Will it help contain storage costs?
- Will it reduce risk?
- Will it improve productivity?
- Will it help me derive new insights from my data?



How Did We Get Here?

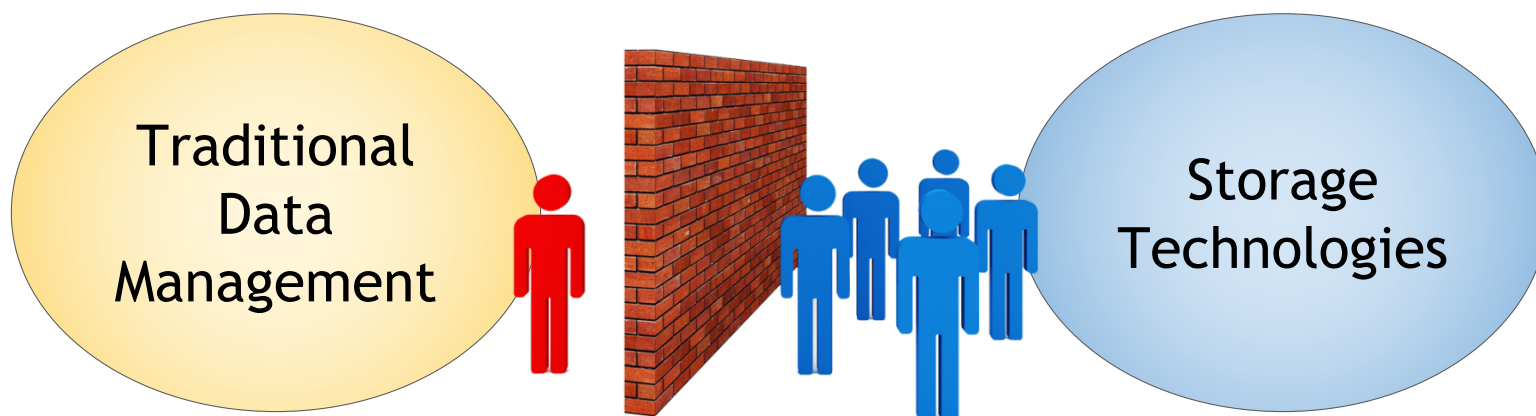
- Managing data growth was typically seen as a storage problem. Why?
 - **Data explosion as a forcing function:**
 - My storage is filling up? Who do I call?
- Traditional approach was to add more storage.
 - **Costs grew so new storage solutions appeared:**
 - HSM, active archives, object storage, cloud options...
- Costs and data volumes still grew so:
 - **Data reduction strategies: deduplication solutions, etc.**
 - Consolidation strategies... etc. etc.
- Bottom Line:
 - **Storage is stupid -- It is not Data Aware**
 - Storage is not qualified to manage data alone.

**To a Hammer
Everything Looks
Like a Nail**



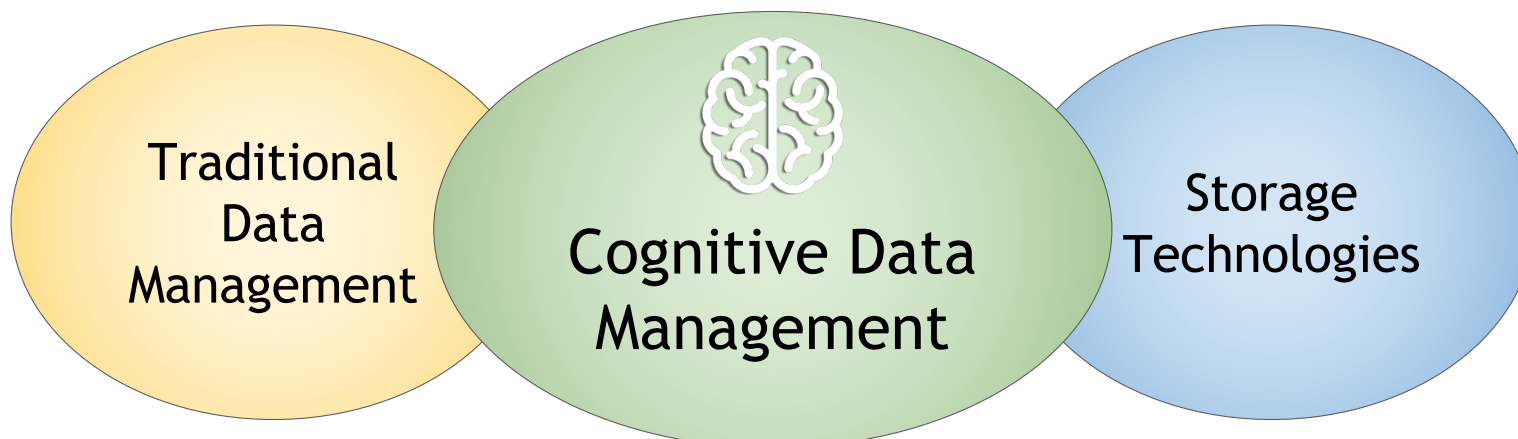
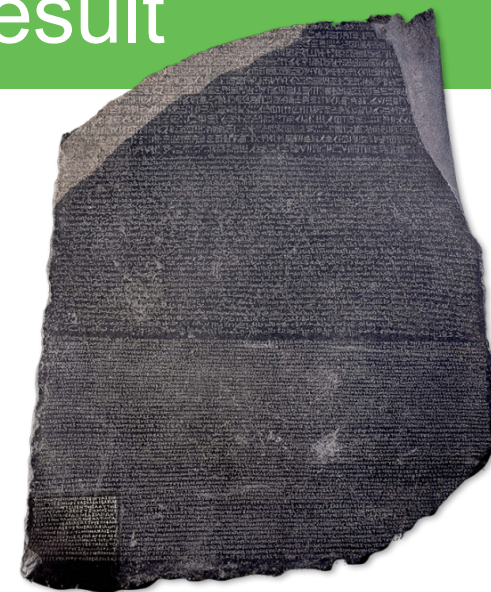
We've Reached a Tipping Point

- The data explosion has reached **critical mass of unmanageability**.
- Budget pressures and growing storage needs are becoming **overwhelmed by incompatibilities & storage complexity, rising costs**.
- Users need to find the data, understand what it is, and focus on using it.
 - **Users should not have to manage storage to use their data.**



Bridging Two Worlds for One Result

- The answer is in the metadata.
 - Metadata is the “Rosetta Stone” that bridges data AND storage resource management.
- A CDM strategy leverages metadata to be both Storage Aware & Data Aware.
 - The metadata itself drives how data is discovered, utilized, organized, stored, & protected.



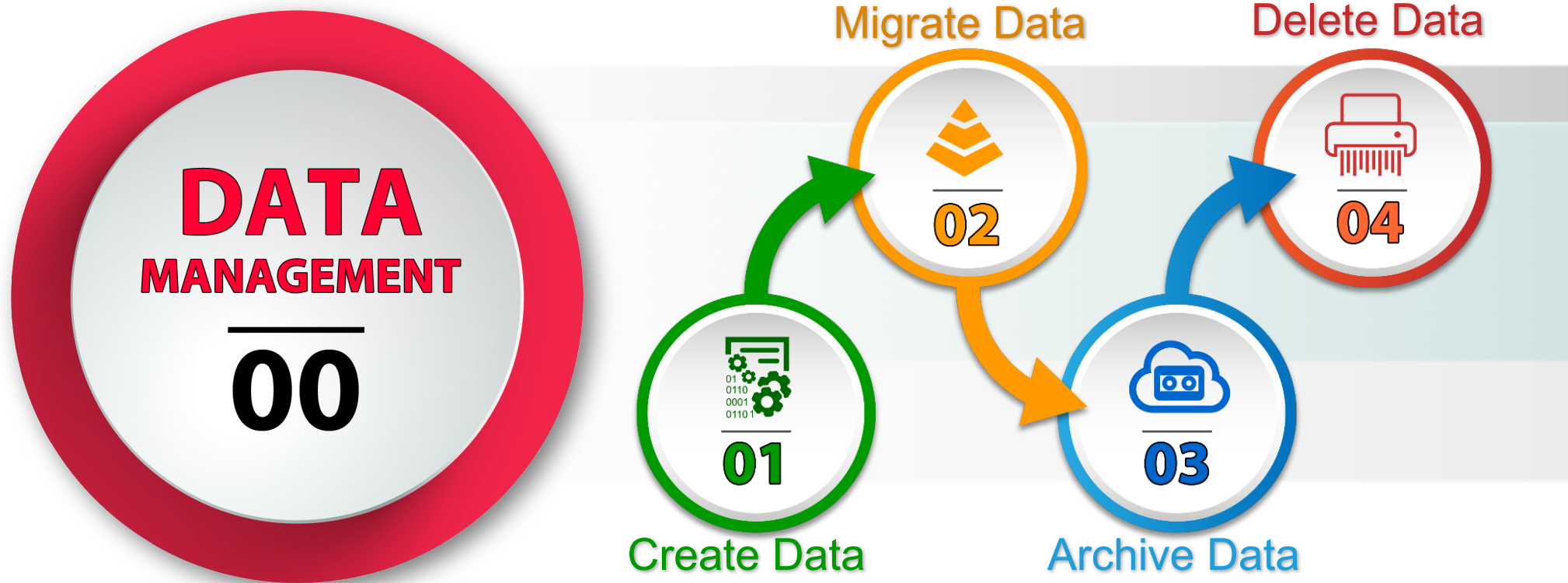
The Problem: Multiple Incompatible Metadata Types

Metadata is literally “Data about the Data”

- **File system metadata:** File name, age, size, owner, etc.
 - Like looking at the cover of a book.
- **Rich file header metadata:** Instrument settings, geospatial info, etc.
 - Like a book’s table of contents and intro section.
- **External metadata:** Information in external databases, DAMs, catalogs, etc.
 - Like reviews, analyses, or other info about the book.
- **User-created metadata:** Project info, workflow triggers, etc...
 - Like notes one might keep about the book...



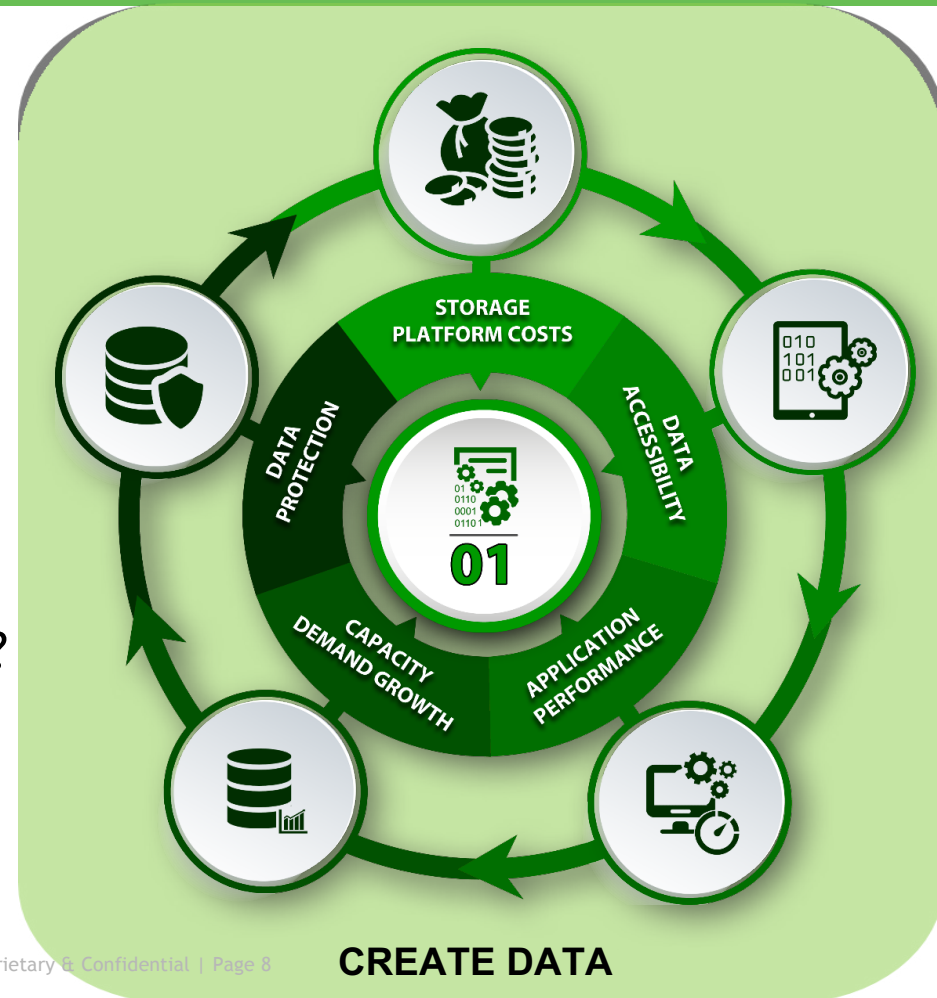
Data Management Involves Four Key Activities...



Create and Store Data... Simple, right?

But...

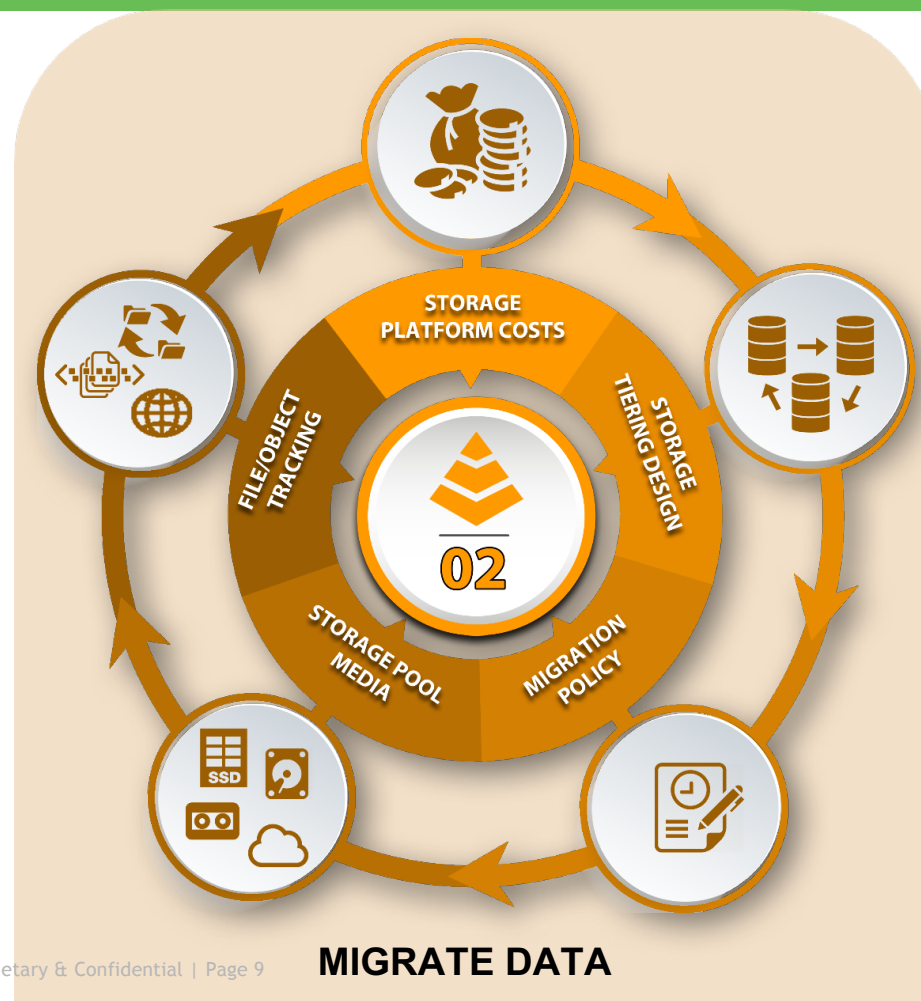
- Which storage platform?
- How about data accessibility/mobility?
- How to minimize latency to optimize app performance?
- How to bend the capacity demand curve?
- How to protect irreplaceable data assets?



Over Time, Data Must Be Migrated...

Again, questions about...

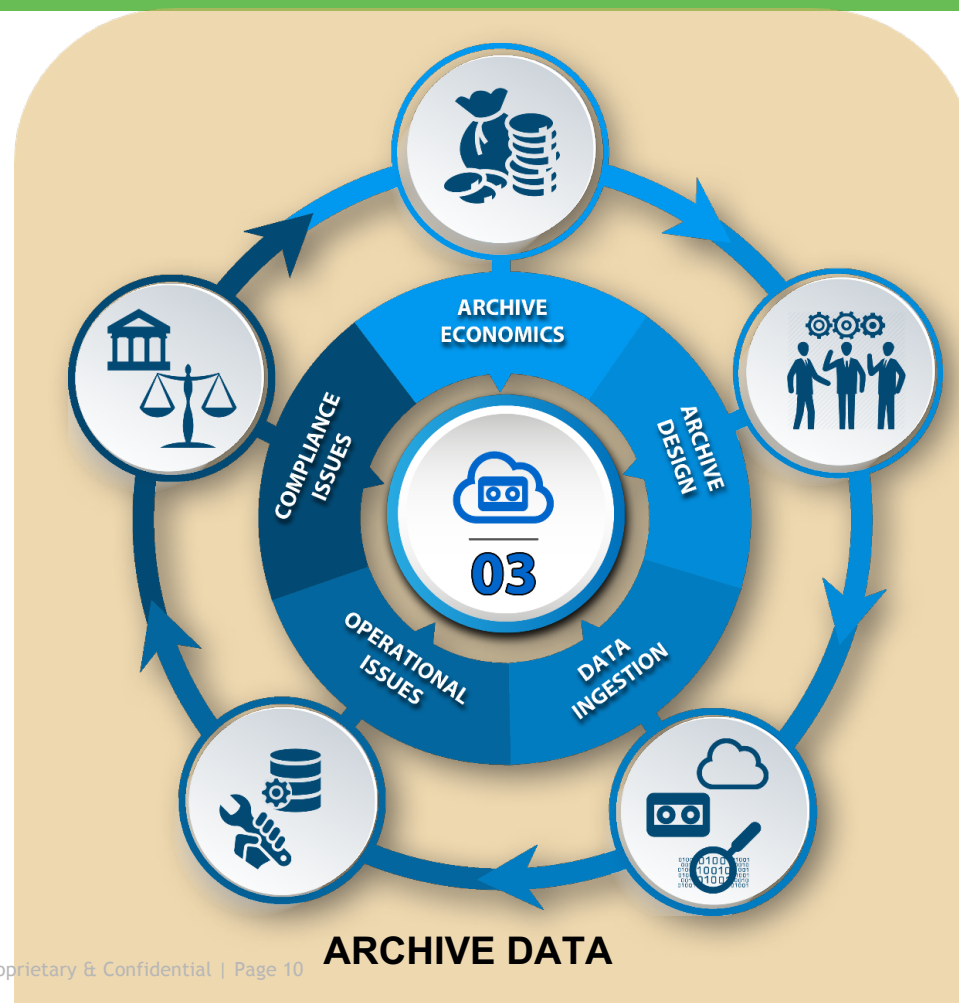
- How much does tiered storage cost?
- How to optimize for latency?
- Who will design policies?
- What storage types should be included?
- How to track files and objects as they move?



Archive Remains Largely Unexplored Territory

Raising even more questions:

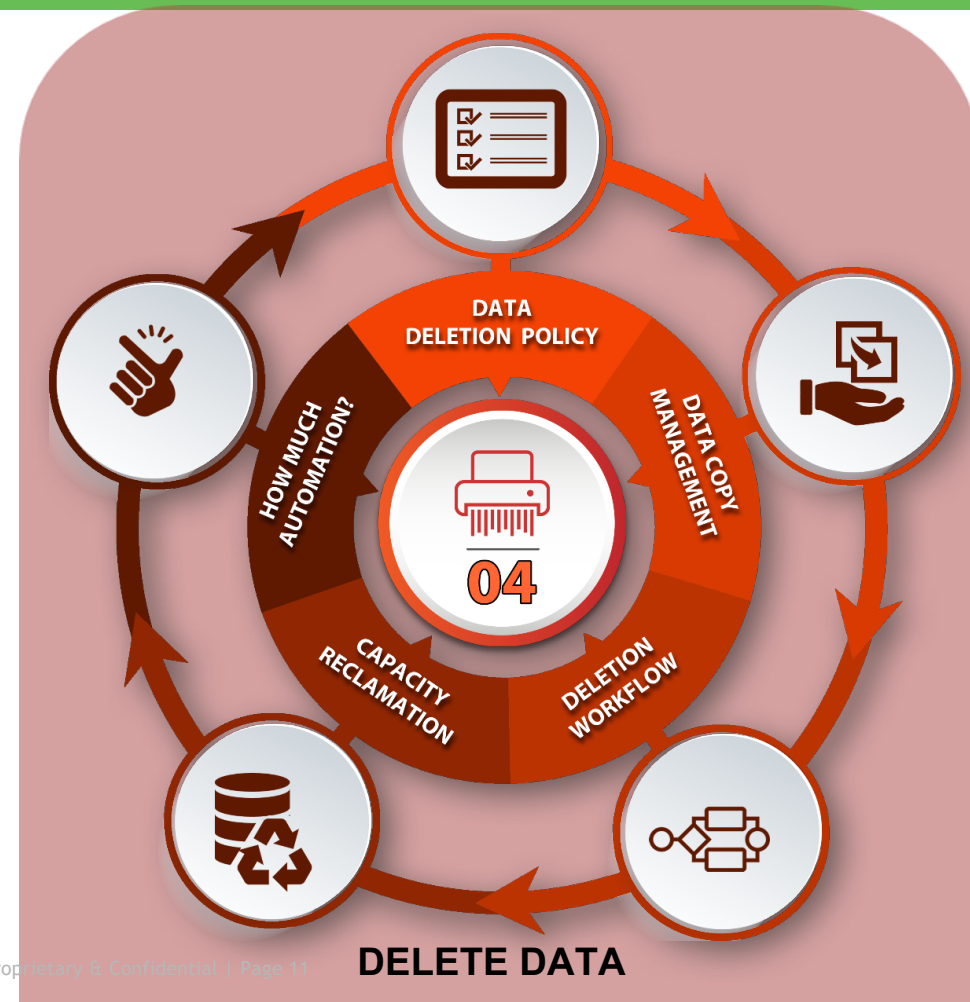
- How expensive is archiving?
→ Is it worth it?
- What is the best design for an archive?
- How do we migrate data into the archive?
- How do we future-proof?
- How do we make archived data easily accessible to users?



And the Deletion of Data at “End of Life”...

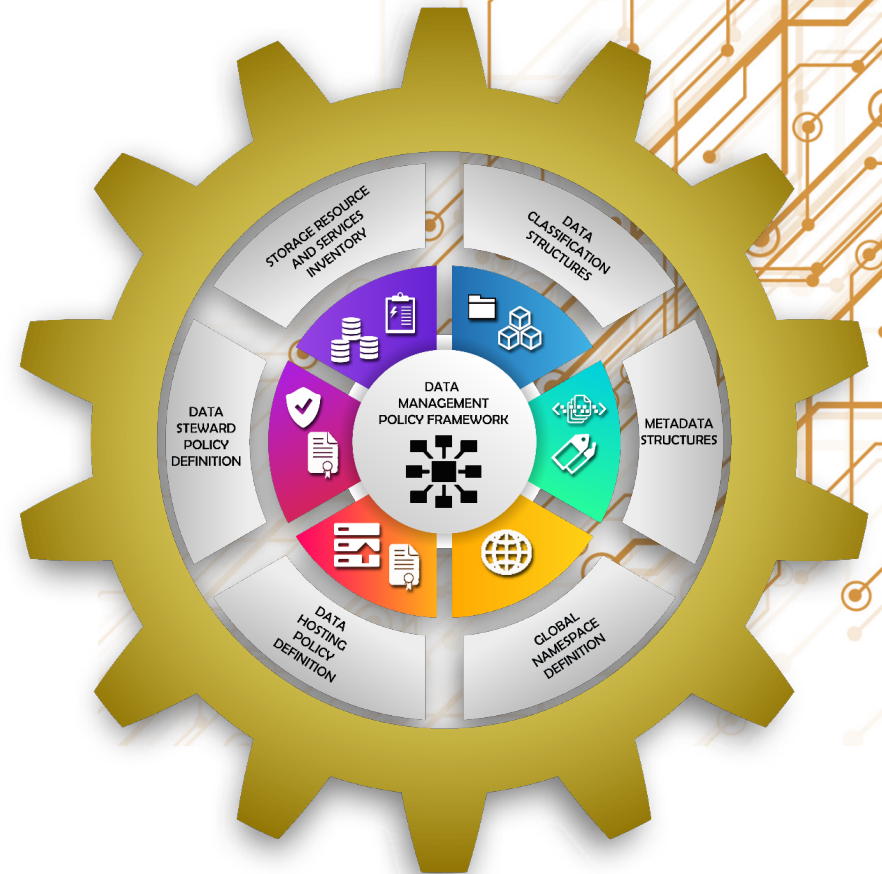
How do I know what to delete?

- Does my deletion policy comply with laws?
- How do I ensure that all copies of data are deleted at EOL?
- How to identify what data to delete and when to delete it?



Where to Begin

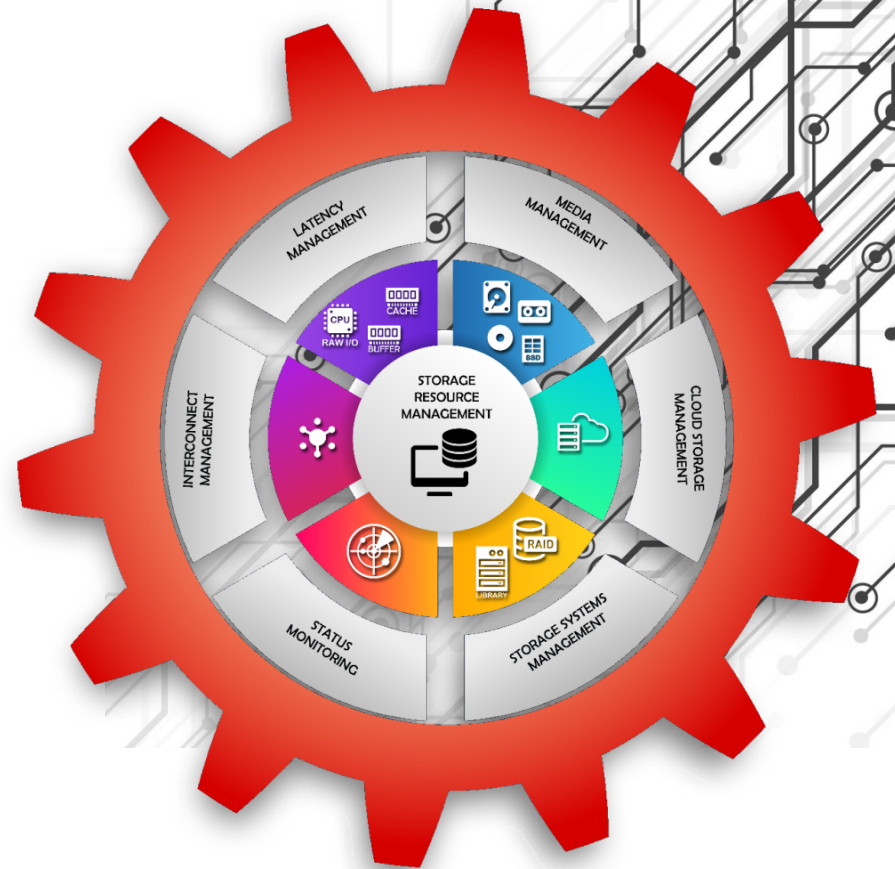
- Management requires a plan or policy for managing data based on:
 - Types and classes of data or subset metadata.
 - Hosting requirements for the data.
 - Service requirements for data: protection, preservation & privacy requirements.
 - Global namespace or other object index needs to be defined.



Add in Storage Resources

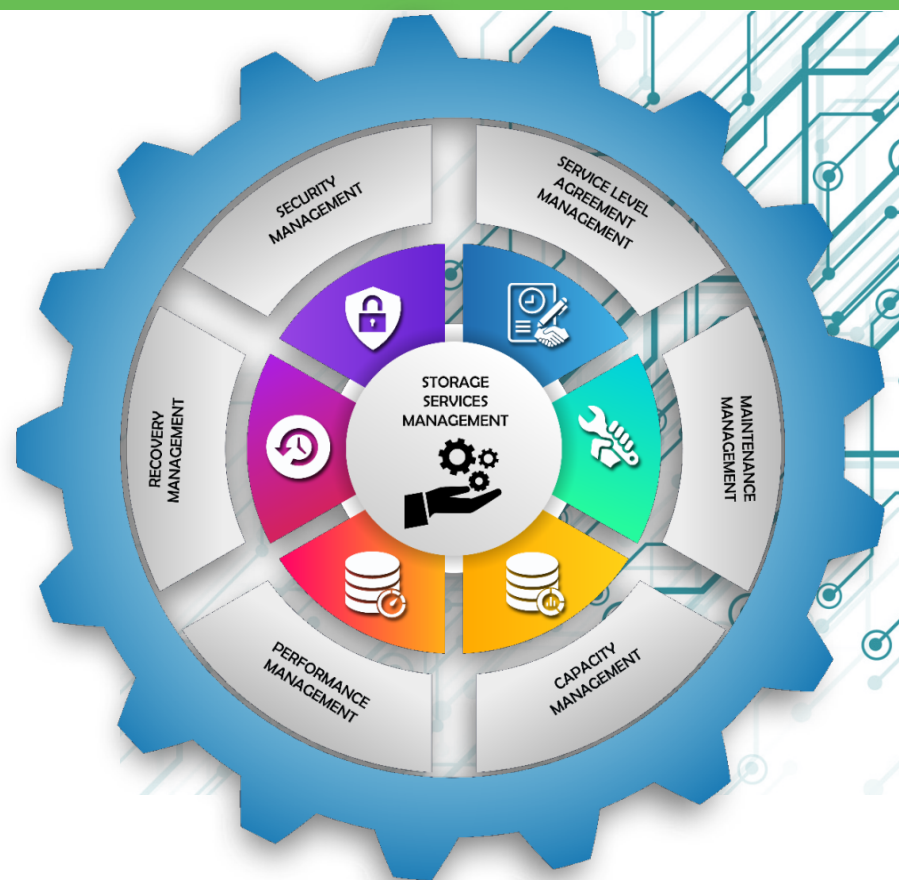
New or existing environments:

- Required storage and network are optimized
- A lot of things to monitor i.e. the Internet of Things (IoT) challenge
- Ongoing & real-time status of storage and network



And Storage Services Management

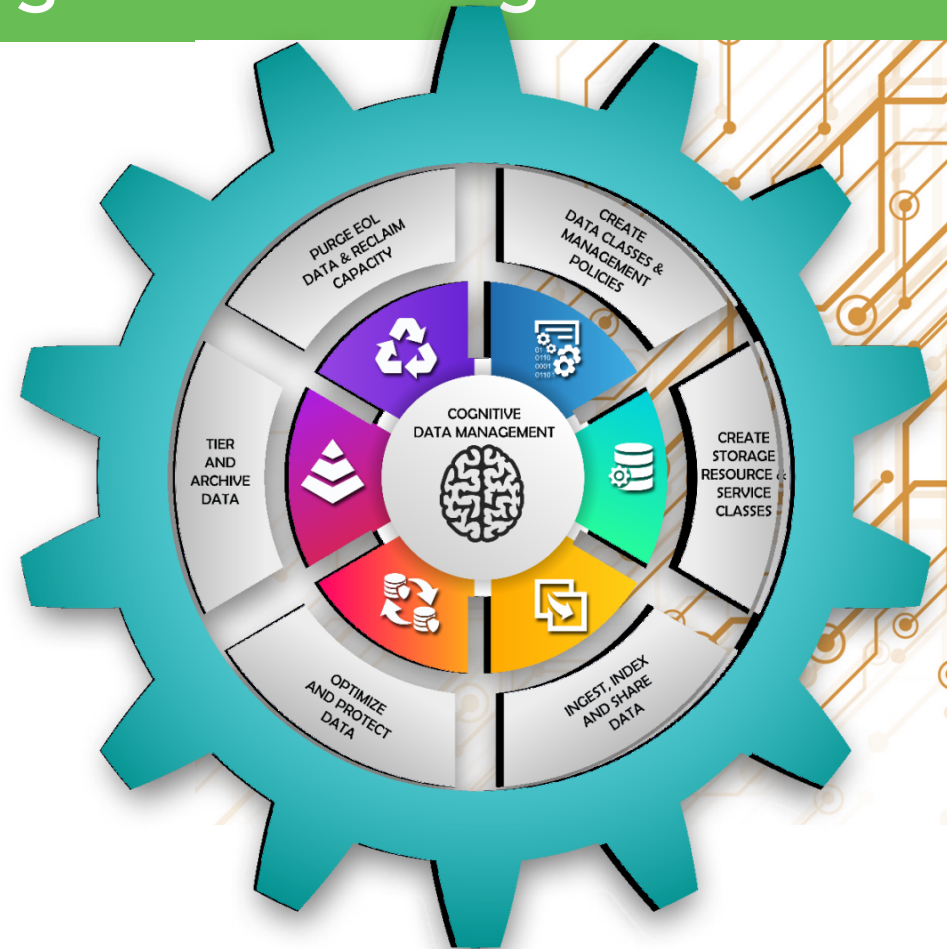
- Storage services that are seamless and fully integrated into holistic global solution including
 - Data protection
 - Data preservation
 - Data deduplication
 - Data security



So, what's this “Cognitive” thing?

“Cognitive Data Management” reflects the demand for autonomous storage, and the need to respond to constant changes in:

- Status of files and objects
- Status of storage services
- Status of storage targets and network
- Evolving use cases and applications

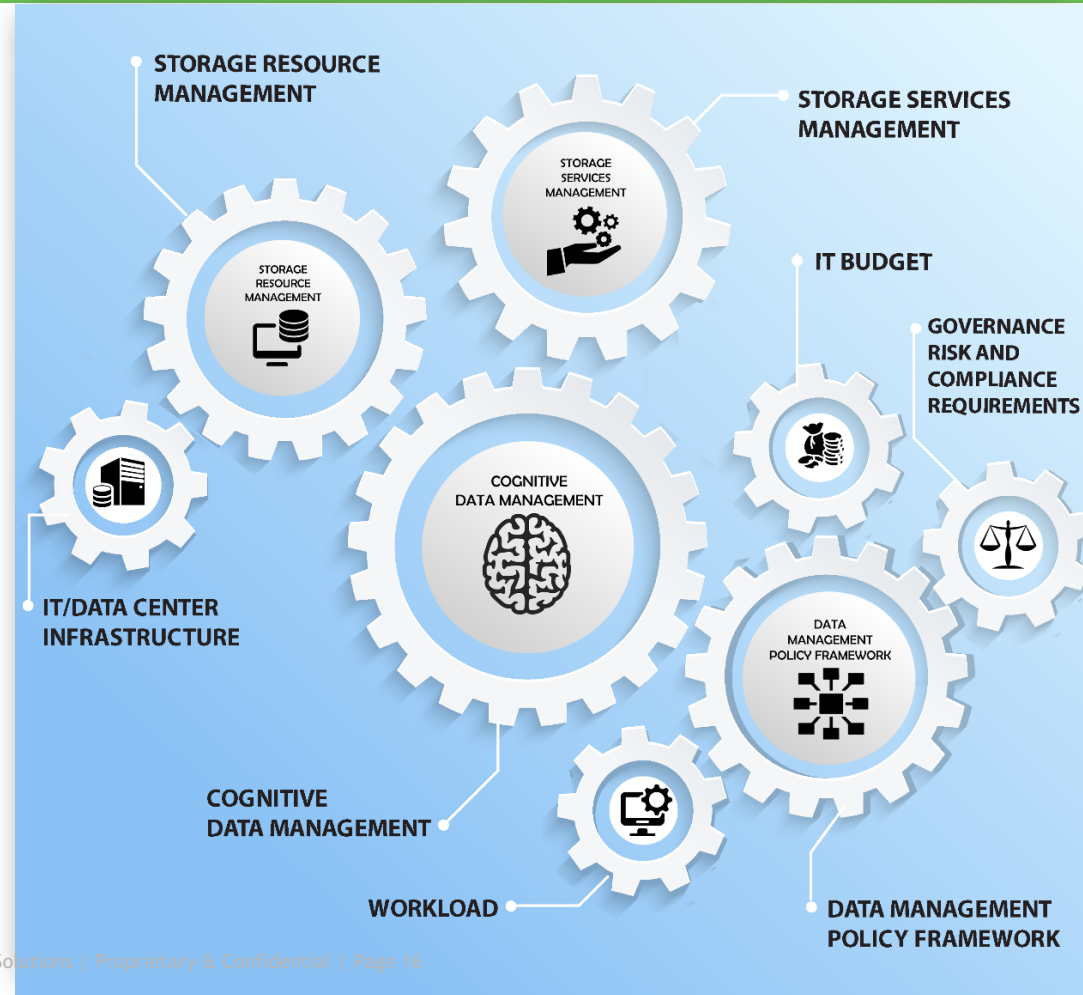


Cognitive Data Management is a Natural Evolution for Data

A single, global solution working across heterogeneous storage.

More than the management of data itself, but the seamless automation of:

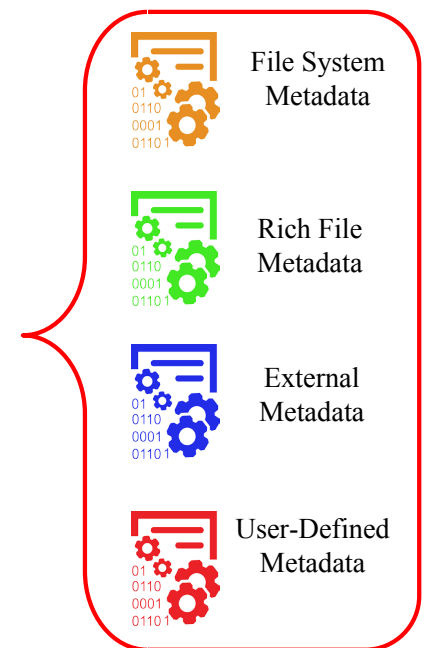
- Storage resources,
- Storage services.
- Powered by metadata and machine learning.



Users Now Have Complete Control of Both Their Data *AND* Storage

Normalizing metadata into a global management framework.

- Bridge incompatible storage/data types.
- Cross-platform global namespace.
- Automate workflows, data migration.
- Break down vendor-imposed storage silos.
- Reduce complexity, & costs.



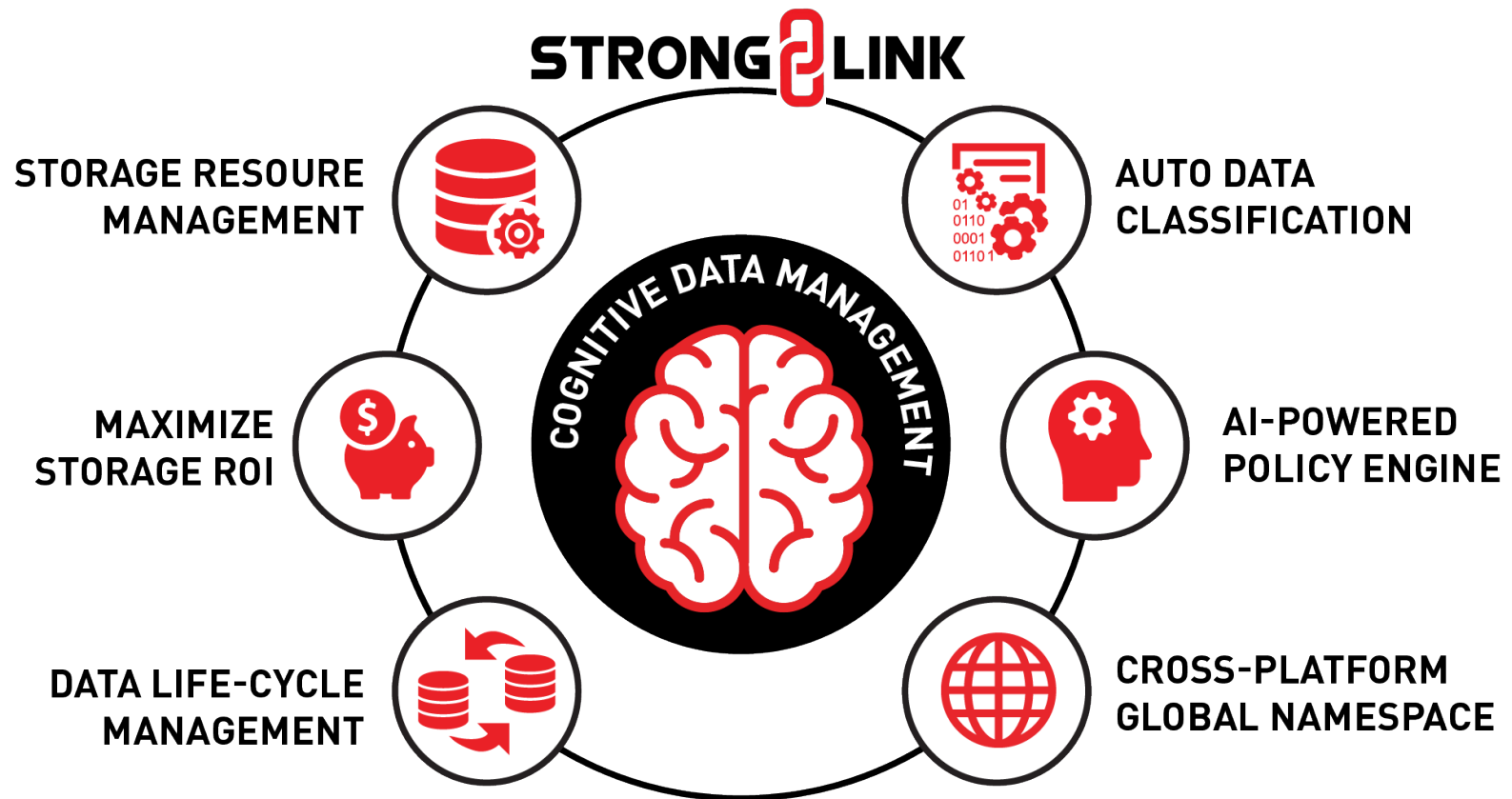
A Cognitive Data Management strategy
enables users to focus on using the data, not wrangling it.



Introducing StrongLink: Part of the **eternity** Suite by



It's About your Data, not your Storage



Questions?

For More Info

Please Visit: dternity.net

Contact:

David Cerf → dcerf@strongboxdata.com



StrongLink is distributed in North America as part of:

The  **dternity suite** by **FUJIFILM** & **STRONGBOX**
DATA SOLUTIONS