take the “meh” out of metadata

harness the “chi” in archiving

take the “rage” out of storage

put the “tada” in metadata
The Most Sophisticated Platform for Unstructured Data Management

If your files could talk, what could they tell you about themselves?

If your files could listen and obey, what would you tell them to do?

Copyright 2021 Starfish Storage Corporation -- CONFIDENTIAL
Discovery + Execution: A Simple But Powerful Paradigm

If your files could talk, what could they tell you about themselves?

If your files could listen and obey, what would you tell them to do?

**DISCOVERY**

*a data catalog for unstructured data*
- massively scalable - billions of files and objects
- extensible metadata - tags and key-value pairs
- simple and turnkey but suitable for custom integration

**EXECUTION**

*a scale-out data mover & batch processor*
- LINUX or Windows – File and Object
- migrate - replicate - workflow - archive - backup - purge
- easy to customize - runs your code or ours

Copyright 2021 Starfish Storage Corporation -- CONFIDENTIAL
Lather, Rinse, Repeat

There is a feedback loop between the catalog and the jobs engine.

- Query database
- Do some stuff
- Update database

- Query the database
- Do more stuff
- Update database
Knowledge Begets Knowledge

Copyright 2021 Starfish Storage Corporation -- CONFIDENTIAL
The Grand Vision

Governance & Stewardship
Archivists – Curators – Librarians – Knowledge Managers
Data Governors – Auditors – Data Wranglers
Retention / Disposition
Anonymization / Privacy
Compliance (GDPR - Export Controls. . . )
Data integrity assurance
Content classification
Data provenance
Data cataloging

Creation & Consumption
Researchers – Artists – Animators – Engineers – Quants
Core Facilities – Data Scientists
Metadata tagging
Search
Workflow automation
Data provenance
Collaboration
Data reusability
Data transmission

Sys Admins (Storage, Backup, HPC) – IT Directors
Data Movement: Archiving/ Tiering, Migration, Replication, Cloud-bursting, Backup & Restore
Reporting: Capacity Planning, Aging, Cost Accounting, Charge-back / Show-back
Misc Operations: Permission Management, GDPR Compliance, Job Scheduling

Copyright 2021 Starfish Storage Corporation -- CONFIDENTIAL
Starfish is Open, Low Risk and Built for Science

★ Not open source but very open
  – Metadata stored in PostGres and we publish the schema
  – API, CLI, write your own GUI, tap the database if necessary
  – No proprietary encoding of files or objects

★ Starfish is out of the data path
  – It does not slow things down
  – It does not introduce points of failure
  – It does not write proprietary formats into industry standard interface

★ Starfish was built for Research Data Management
  – Our long-term roadmap addresses the challenges of science
  – Forged in the data center at Harvard Med School
  – Battle-tested with just about every modality of life sciences research and user bad behavior

★ Starfish is vendor-independent
  – Works across storage devices, storage types, now and in the future