Diskover

SOLUTION BRIEF | INSIGHTS DASHBOARD | LIFE SCIENCE

Every file has a value. Now you can see it.

See what is irreplaceable, what a pipeline can rebuild, and what closed studies still cost to store.



RESULTS WITH DISKOVER AT A CANCER RESEARCH CENTER

$542K reclaimed within days of the first index

0.5 PB of temporary data reclaimed

27M pathology files indexed

At a glance

KEY CHALLENGES

Labs keep instrument data, sequencing output, and results on many systems. Pipelines leave copies and temporary files behind, and it is hard to know what is safe to remove. Storage grows, and the bill has no clear owner.

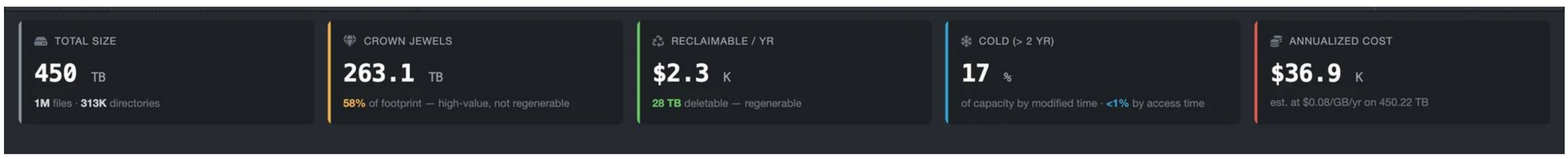
THE SOLUTION

The Insights dashboard groups storage by its role in research, from primary instrument data to pipeline scratch. It reads metadata only, never file content. It shows what each group is worth, what is waste, what it costs, and which fix comes first.

WHY IT MATTERS

Protect primary instrument data, clear what a pipeline can rebuild, and move closed studies to lower-cost storage before you buy more. Each fix shows its size, yearly cost, and score gain, so you can prove the savings.

See how healthy your storage is





Screenshots show Diskover demo data.

FIVE KEY NUMBERS

Total size, crown jewels, reclaimable spend per year, cold data, and annual cost. The storage cost rate is a default that you change in Insights settings.

A SCORE OUT OF 100

Hygiene, waste, and risk combine into one efficiency score. The demo estate scores 83, rated Good. The score is a guide, not an audit.

READY AFTER THE FIRST SCAN

No tags or plugins to set up. Statistics update after each scan, and administrators choose the Life Sciences profile in settings.

About Diskover

Diskover helps enterprises structure unstructured data across cloud, on-prem, and edge environments, making it easier to search, govern, optimize, and deliver file data for compliance, analytics, and AI-ready workflows.

Contact

diskoverdata.com
sales@diskoverdata.com
(800) 560-5853
San Jose, CA

Sort your research data by what it is worth

Workflow classification groups storage by its role in research, gives each group a value tier, and shows how old it is. Primary instrument data is a crown jewel. Derived files, scratch, and logs can be rebuilt. Groups are found from metadata only. For example, whole slide images are .svs, .ndpi, or .mrxs files in folders such as /pathology/ or /slides/.



Screenshots show Diskover demo data.

CROWN JEWELS FIRST

Primary instrument data fills 78% of the demo estate: slide scans, sequencing, imaging, and mass spec. These are the files to protect before any cleanup.

DATA STILL TO REVIEW

66 TB, or 15% of the demo estate, matches no group yet. Review it with its owners before you keep or clear it.

REBUILD IT, DO NOT KEEP IT

Derived and intermediate files take 27.6 TB in the demo. A pipeline can rebuild them, so they do not need the same protection as raw data.

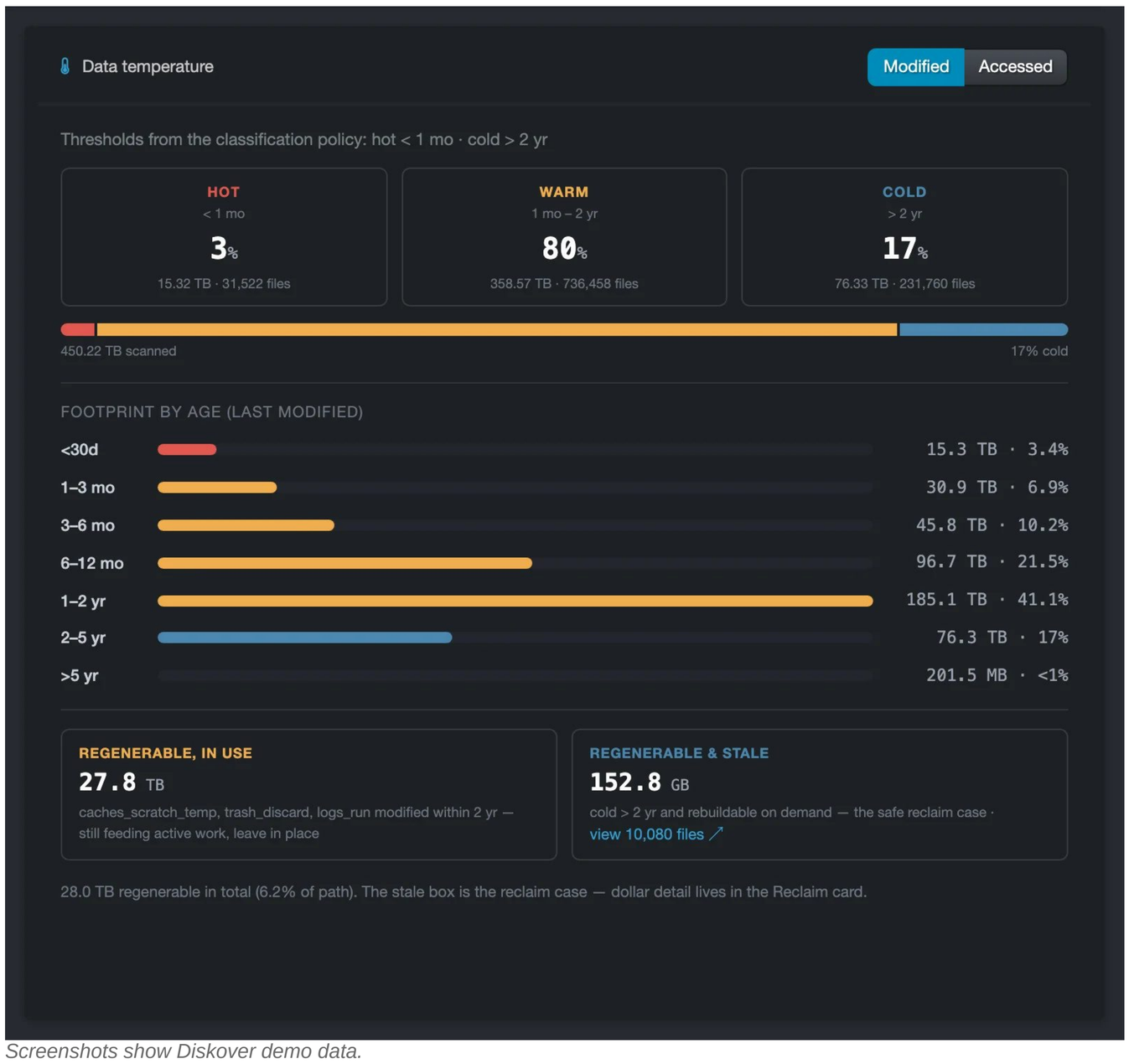
About Diskover

Diskover helps enterprises structure unstructured data across cloud, on-prem, and edge environments, making it easier to search, govern, optimize, and deliver file data for compliance, analytics, and AI-ready workflows.

Contact

diskoverdata.com
sales@diskoverdata.com
(800) 560-5853
San Jose, CA

Find what you can rebuild or remove



Screenshots show Diskover demo data.

Data temperature shows how much research data is idle. Reclaim shows how much of it is waste, and what it costs each year.

HOT, WARM, AND COLD

In the demo, 3% of research data is hot, 80% is warm, and 17% is cold. Administrators set the age limits in Insights settings.

REBUILDABLE FILES, IN USE OR STALE

Derived files and scratch still in use are kept apart from stale ones, so you clear only what is safe to clear.

4.2 TB OF DUPLICATES

In the demo, .gz files alone take 1.71 TB, or 41% of duplicate space. Duplicates show when the duplicates plugin is on.

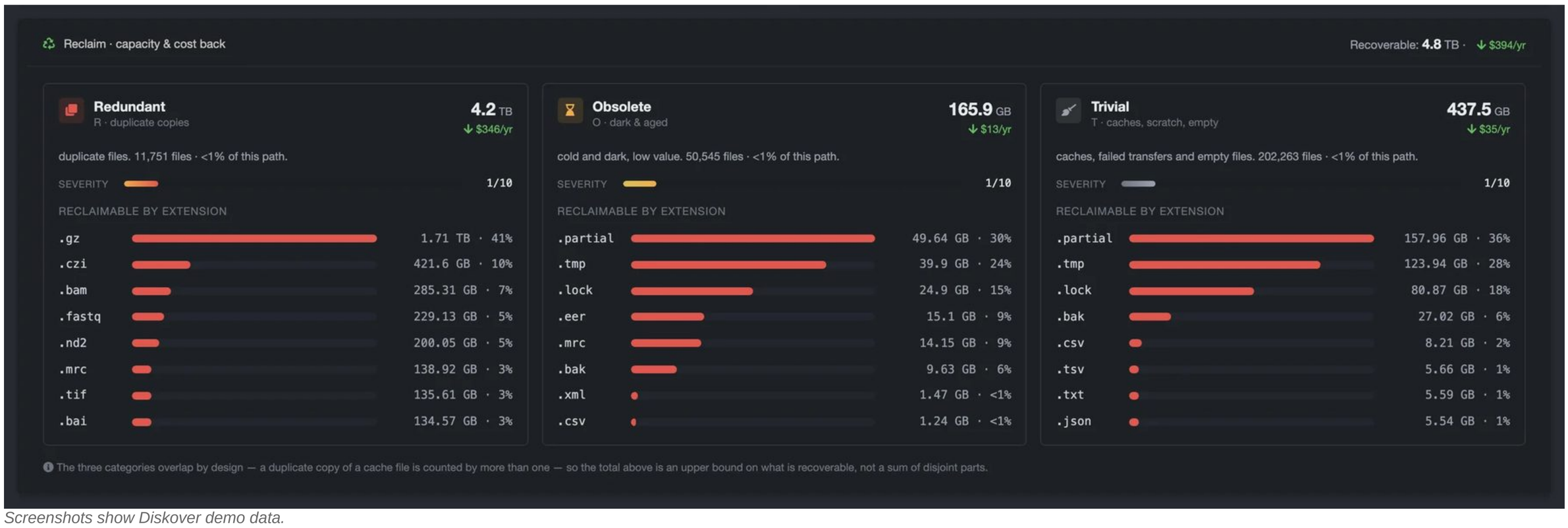
MODIFIED OR ACCESSED

Switch between last modified and last accessed dates, where your storage records them. In the demo, 17% of research data is cold by modified date, but less than 1% by access date. Check before you move old files.

WHERE THE BULK SITS

In the demo, 41% of research data (185 TB) was last changed 1 to 2 years ago. It is warm today and turns cold within a year if it stays unchanged.

See what you can reclaim



Screenshots show Diskover demo data.

REDUNDANT, OBSOLETE, AND TRIVIAL

Reclaim sorts wasted space into three groups and shows the size, yearly cost, and top file types for each. In the demo, 4.8 TB is recoverable, about $394 a year. The groups can overlap, so the total is an upper bound.

About Diskover

Diskover helps enterprises structure unstructured data across cloud, on-prem, and edge environments, making it easier to search, govern, optimize, and deliver file data for compliance, analytics, and AI-ready workflows.

Contact

diskoverdata.com
sales@diskoverdata.com
(800) 560-5853
San Jose, CA

Which fix to make first

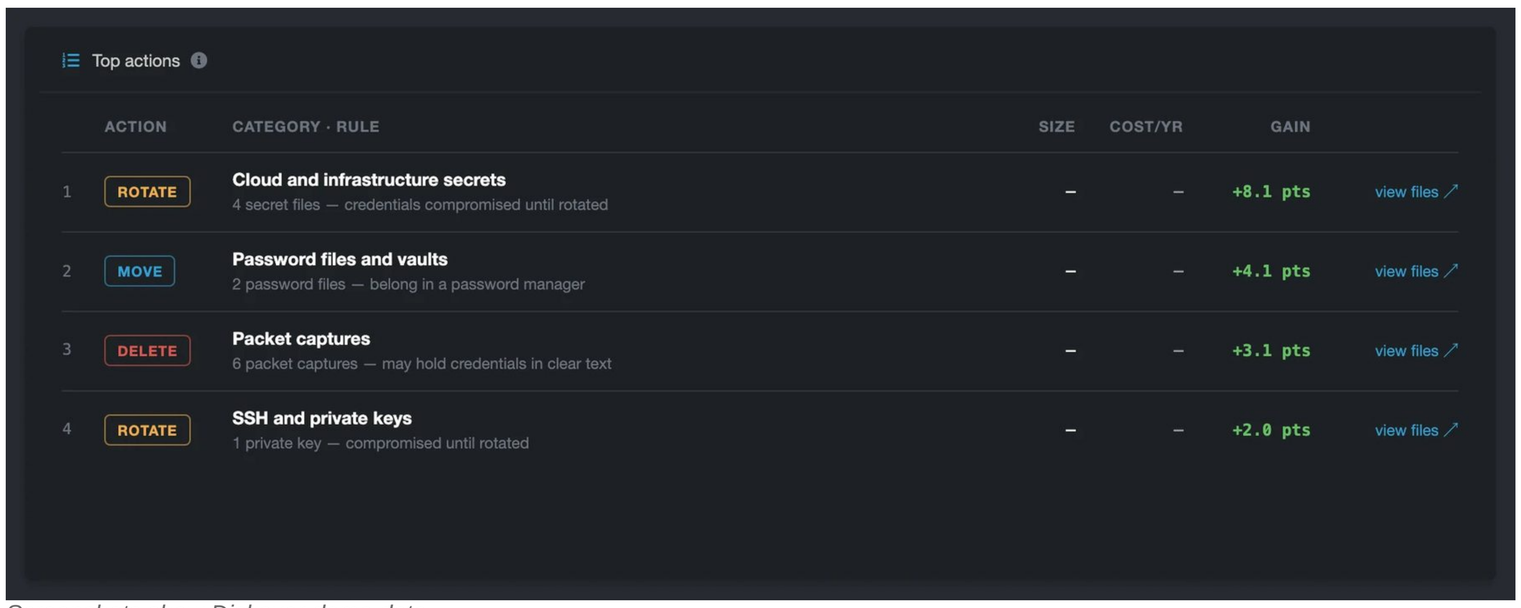
Top actions ranks every fix by what it is worth. Each one (delete, move, or rotate) shows its size, its yearly cost, and the score points it adds. In the demo, the first fix is to rotate exposed cloud secrets.

RISK, NOT ONLY SPACE

Some fixes save no space but cut risk, like exposed secrets and keys.

CHECK THE FILES FIRST

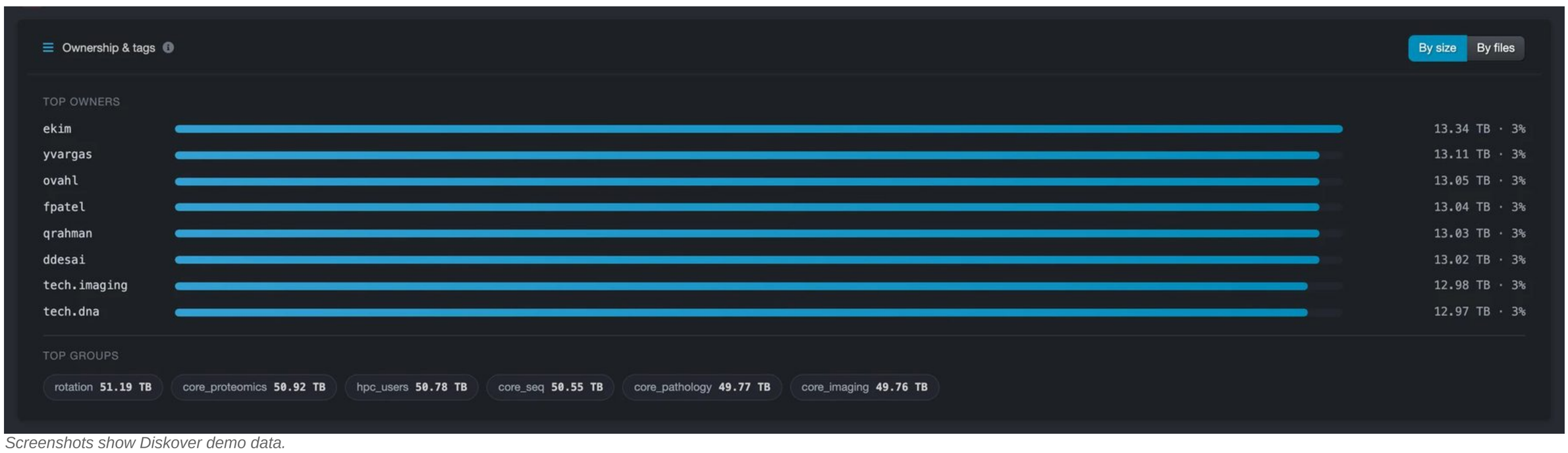
Each fix links to its file list, so you can review before you act.



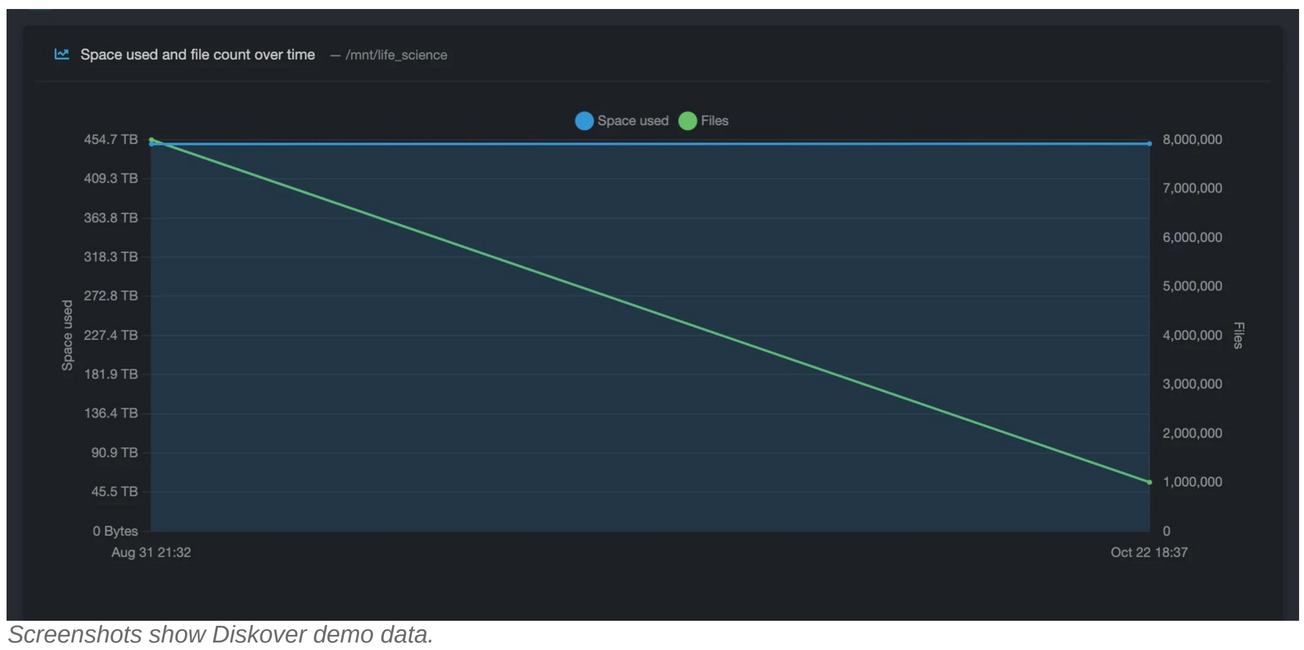
Screenshots show Diskover demo data.

See who owns it, and how fast it grows

Research storage is shared by labs, core facilities, and HPC users. Ownership shows how much space each team holds, and Space used over time shows how fast it grows.



Screenshots show Diskover demo data.



Screenshots show Diskover demo data.

SPACE BY TEAM

In the demo, the largest groups are rotation, core_proteomics, hpc_users, core_seq, core_pathology, and core_imaging, with about 50 TB each.

GROWTH YOU CAN PLAN FOR

Capacity and file count over time show how fast storage grows, so your next purchase is based on facts, not guesses.

GO FURTHER WITH THE AI CONNECTOR

With the optional AI Connector, administrators ask questions about their storage in plain language and get answers from the same data. Every action it proposes waits for human approval.

GET STARTED

Ready to see it on your research data?

Start with a 30-day proof of concept on your own data, and recover your investment in 3 to 6 months.

About Diskover

Diskover helps enterprises structure unstructured data across cloud, on-prem, and edge environments, making it easier to search, govern, optimize, and deliver file data for compliance, analytics, and AI-ready workflows.

Contact

diskoverdata.com
sales@diskoverdata.com
(800) 560-5853
San Jose, CA

Revised November 2026