earthmover

SC26 WORKSHOP · HIGH PERFORMANCE PYTHON FOR SCIENCE AT SCALE

making science faster

Ryan Abernathey

CEO, Earthmover PBC

available at **slides.earthmover.io**

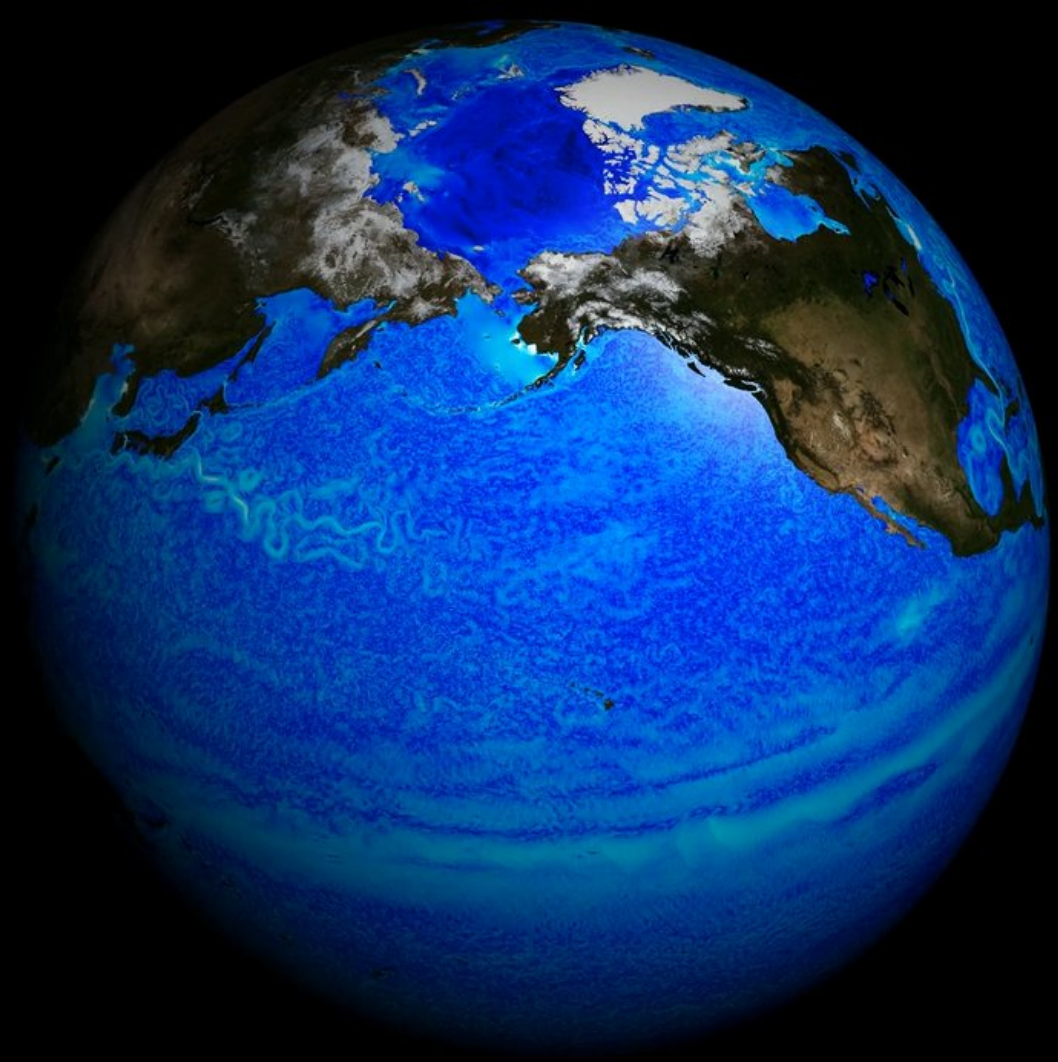
THE DATASET THAT CHANGED MY LIFE

MITgcm LLC4320

CIRCA 2016

Dimitris Menemenlis (NASA JPL)
& Chris Hill (MIT)

Deformation radius: **10 km**
Spatial resolution: **1-2 km**
Grid points: **10^8**
Vertical levels: **100**
One scalar field: **40 GB**
State variables: **10**
Temporal resolution: **1 hour**
Run length: **1 year**
Raw dataset size: **40 PB**
Compressed dataset size: **2 PB**



MY JOURNEY



UNTIL 2023

Earth-System Scientist

- PhD in Climate Physics at MIT
- Obsessed with ocean mesoscale eddies
- Tenured Professor at Columbia University / LDEO



SINCE 2018

Open Source Hacker

- Core contributor to **Xarray**; **Zarr** steering council
- Co-creator of **Pangeo**
- SciPy 2018 talk: **Zarr** for ocean-model data



TODAY

Startup Founder and CEO

- Left the university with my collaborators
- Building **Earthmover**

ABOUT EARTHMOVER

scientists building data infrastructure for science



Ryan Abernathey
CEO & CO-FOUNDER

MIT · COLUMBIA UNIVERSITY



Joe Hamman
CTO & CO-FOUNDER

NCAR · CARBONPLAN



Margaret Francis
CHIEF OPERATING OFFICER

DBT LABS · HEROKU · SALESFORCE



MIT · HARVARD · OXFORD

OCEANOGRAPHY · HYDROLOGY · REMOTE SENSING

CELL BIOLOGY · GENOMICS · PLASMA PHYSICS

PUBLIC BENEFIT CORPORATION · FOUNDED 2022

Our mission: to empower people to use scientific data to solve humanity's greatest challenges.

Our mission is written into our corporate charter.

OPEN CORE

community open source

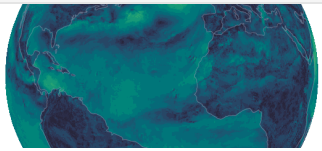
Xarray · Zarr · Icechunk · PANGEO

managed cloud platform

Arraylake, for research institutions, government agencies and companies

DATA-INTENSIVE SCIENCE

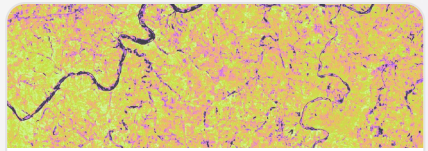
Fields where the ability to interrogate data is the rate-limiting factor



WEATHER & CLIMATE

Predicting extremes and tipping points, weeks to decades ahead

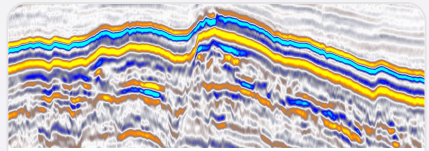
→ Protecting lives, property, food and energy supply



EARTH OBSERVATION

Quantifying the carbon cycle from space

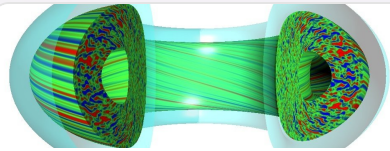
→ Credible carbon accounting that drives emission cuts



GEOPHYSICS

Predicting what lies beneath the subsurface

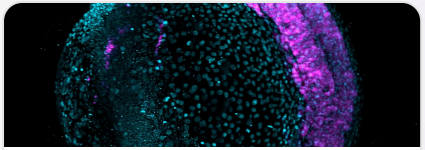
→ More abundant energy and safer resource extraction



FUSION

Achieving sustainable fusion reactions

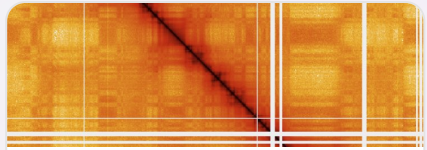
→ Unlimited clean, cheap energy



CELL BIOLOGY

Holistic models of cellular function

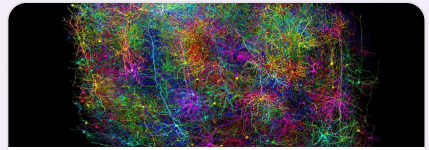
→ Preventing and curing cancer



GENOMICS

Linking genotype to phenotype at population scale

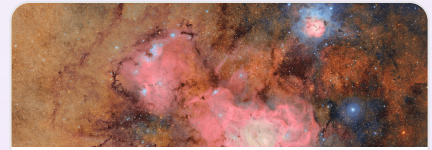
→ Precision medicine for every patient



NEUROSCIENCE

Predicting brain function from structure

→ Preventing and curing mental illness



ASTRONOMY

Measuring and explaining dark energy

→ Understanding the fate of the universe

Images: Allen Institute · NSF–DOE Vera C. Rubin Observatory

THE NETWORK OF KNOWLEDGE

science is the engine of human prosperity

Each discovery is a new node, built on what came before. The rate of advancement is constrained by two things:

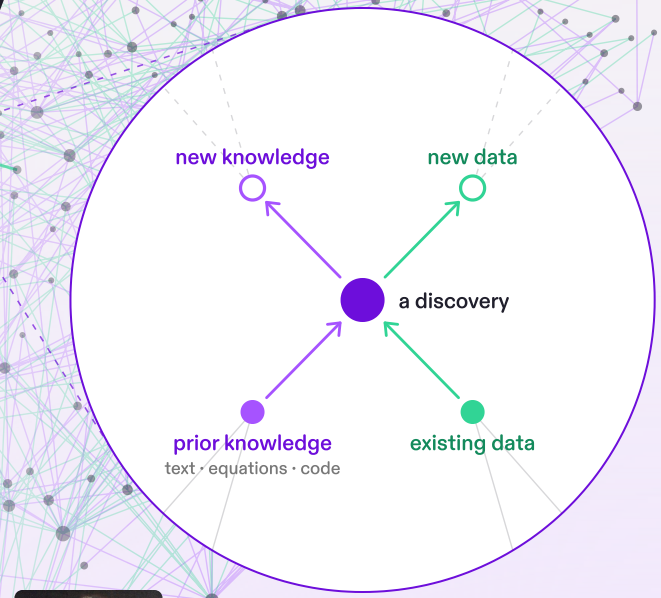
MICRO-SCALE How fast an individual project can be executed

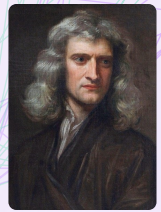
What happens when each PhD student sits down to work every day

MACRO-SCALE How fast knowledge can propagate through the network

Sharing data, code and knowledge: open science

making science faster = accelerating the rate at which we add nodes to this network





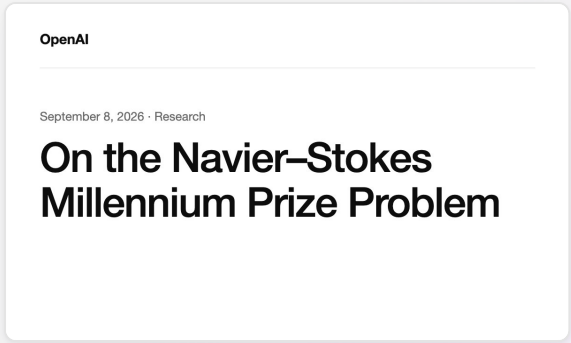
"If I have seen further it is by standing on the shoulders of giants."

Isaac Newton to Robert Hooke, 1675

THE AI DISRUPTION

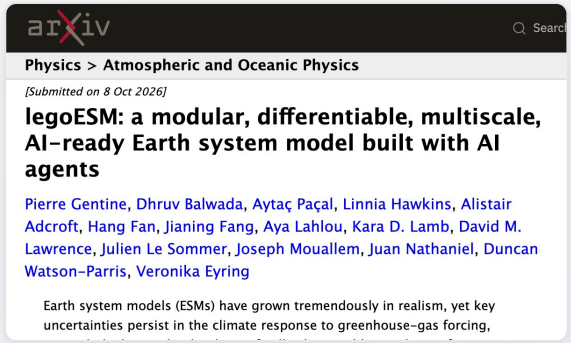
reasoning is free now

Solving the hardest math and theoretical physics problems



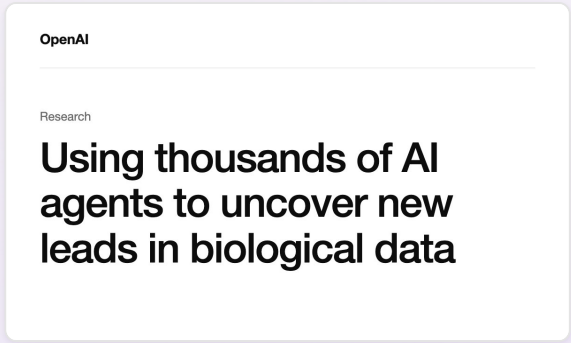
OpenAI · Navier–Stokes Millennium Prize Problem

Writing complex scientific code in any language



legoESM · Fortran climate models → JAX

Identifying new directions for research



OpenAI · thousands of agents on biological data

This is not just “grunt work.” It’s the work done by **PhD students, postdocs, and PIs.**

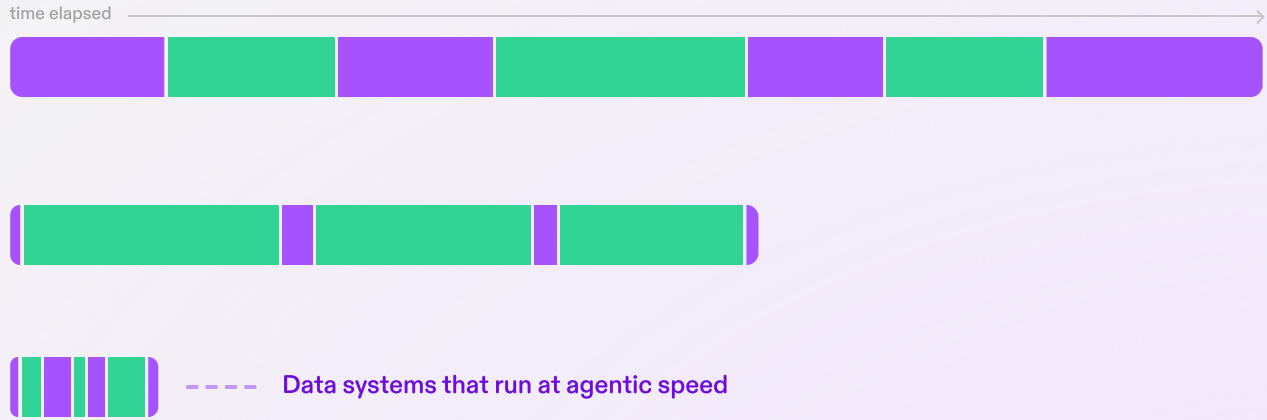
The genie is out of the bottle.

AGENTIC SCIENCE

Data infrastructure is now THE bottleneck

Thinking and writing code

Waiting for data processing



Credit: Ian Cook, CEO of Columnar, came up with this way of visualizing time spent.

THE PROBLEM

scientific data is hard to use!

THREE CORE PROBLEMS



Files and file formats

A zoo of binary formats, tied to domain libraries never built for cloud or AI. One dataset spans millions of files — and you're expected to download them.



Data analysis via “scripts”

Science logic tangled up with physical execution — and most domain scientists aren't distributed-systems programmers.



Fragmented “data portals”

Each one takes expert context to navigate. Sharing large data back with the community is slow and expensive.

> The problem is that [scientific] data does not live in a single database with a single interface. It is a **messy network of roads**, each with its own identifiers, conventions, formats, filtering logic, and degree of programmatic access. Some data are straightforward to access programmatically. Others, not so much.
>
> Anthropic | Agents in biology

THE SOLUTION

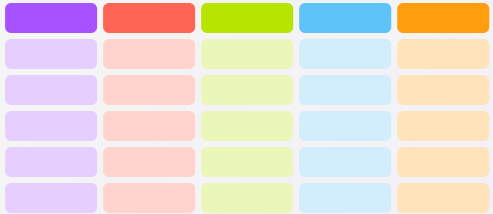
Make scientific data work more like a **database** — while respecting the unique shape and structure of scientific data.

science needs data infrastructure built for tensors

Tables

FOR DISCRETE RECORDS

Events · Customers · Transactions · Inventories



FILE FORMATS

Parquet · Apache Arrow · Iceberg · Delta Lake

ANALYTICAL DATABASES

Snowflake · Databricks · BigQuery · ClickHouse

DuckDB · Apache DataFusion · Polars · Apache Spark

Tensors

FOR CONTINUOUS FIELDS

Satellites · Microscopes · Telescopes · Sensors · Simulations



Earthmover helps maintain Zarr, the open-source cloud-native tensor format.

FILE FORMATS

TIFF · HDF · NetCDF · Zarr

ANALYTICAL DATABASES

? no tensor-native query engine... yet



icechunk

from file archives → to AI-ready data

> “I want `OR_ABI-L2-CMIPF-M6C13_G16_s2021063….nc` ”
>
> — said no one ever.

> “I want **GOES-16** **band 13 over Gulf of Mexico, 2021**”
>
> — someone, probably.

Scientists don’t care about files. They care about **datasets** and **queries**.

Increasingly, so do **AI agents**.

ANALYSIS-READY, CLOUD-OPTIMIZED

from files to arco data cubes

Humans and agents don't want to deal with thousands of individual files. They want ARCO data cubes.

Analysis Ready

> *All timesteps, variables, ensemble members, etc. are organized into a single logical object (data cube).*

Cloud Optimized

> *Applications can efficiently stream data directly from cloud object storage with low latency and high throughput.*

Zarr is the most widely used ARCO format for array data.

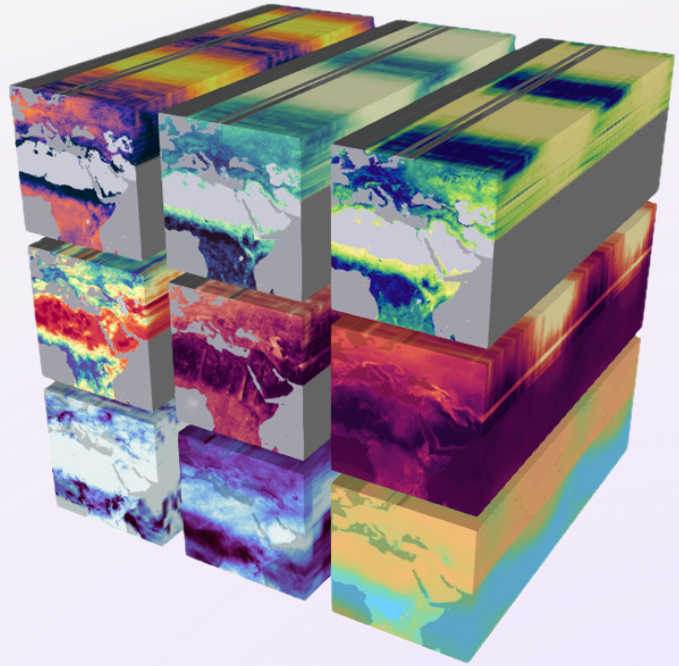


Image credit: Mahecha et al. 2020. doi.org/10.5194/esd-11-201-2020

CORE DATA MODEL

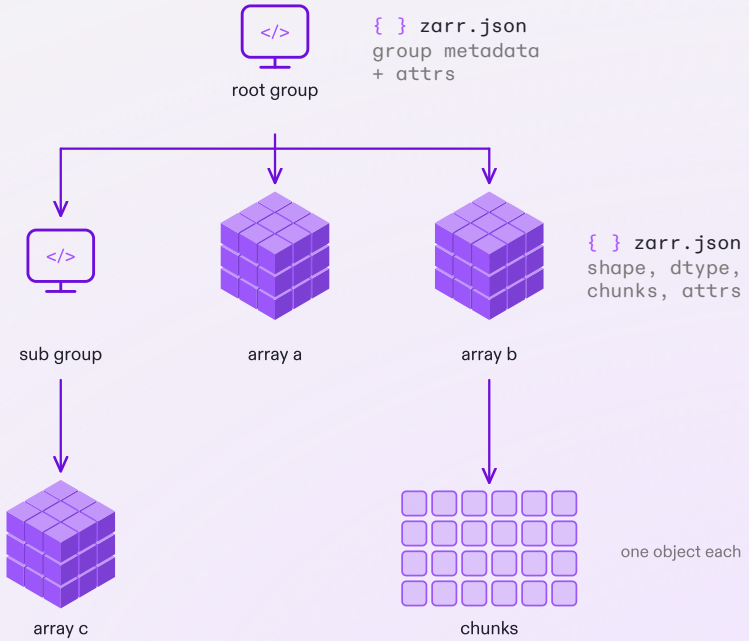
HDF and Zarr

- ✓ **A hierarchical tensor data model** — groups, arrays, metadata. Flexible enough for a huge range of scientific domains.
- ✗ **One file can't hold a petabyte** — so datasets grow an ad-hoc, unstandardized "outer" hierarchy of files.
- ✗ **Built for low-latency POSIX** — not for high-latency cloud object storage.

Zarr

- ✓ **Explodes the HDF5 model into many objects**, separating metadata from data.
- ✓ **Transformative adoption** — genomics, bioimaging, Earth-system data. DeepMind's AI weather forecasts; Copernicus Sentinel missions.

New challenge: consistency at scale → **Icechunk**



WHAT IS ICECHUNK



icechunk

Open-source, high-performance array format and library for scientific data

01

Scale and Speed

Gigabytes to petabytes. Never the bottleneck.

02

Zero-Copy Ingest

HDF5, NetCDF, h5ad, TIFF, ND2, GRIB and more

03

ACID Transactions

Commits are all or nothing. No partial reads.

04

Versioning

Cite exact versions and collaborate on branches

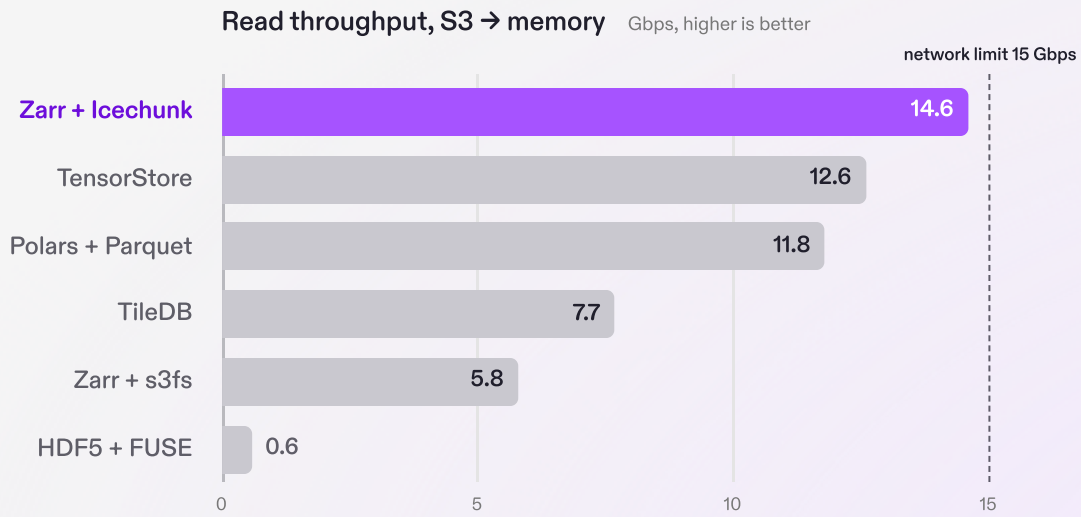
05

No Server Required

Client and format only. No service to run.

ICECHUNK

scale and speed



SATURATES THE NETWORK

37 Gbps reads on a 40 Gbps m8g.24xlarge. Real ERA5 data hits **28 Gbps** on the 15 Gbps box: compression moves half the bytes.

ANY SIZE

GB on a laptop

TB in a bucket

PB referenced in place

Same library at every scale. Reads fetch only the chunks you need, in parallel.

AWS m8g.8xlarge (32 vCPU, 15 Gbps) ↔ S3, us-east-1 · 3.2 GB int64 random array, ZSTD, 12.5 MB chunks · median of 5 runs, no Dask · Icechunk: updated 2026 run · Source: I/O-Maxing Tensors in the Cloud (Earthmover, Nov 2025)

ICECHUNK

zero-copy ingest

Query the archive in place.

Reading TIFF, HDF5, NetCDF or GRIB from the cloud means round trips for headers and offset tables before you find the bytes you want. So people download whole files.

Icechunk parses each file once and records where every chunk's bytes are. A virtual chunk read is **one range request**: cloud-optimized, without duplicating or changing the originals.

GOES-16 ARCHIVE

115 TB → 80 GB

380,000 NetCDF files, as one Icechunk repo

NATIVE READER BILL

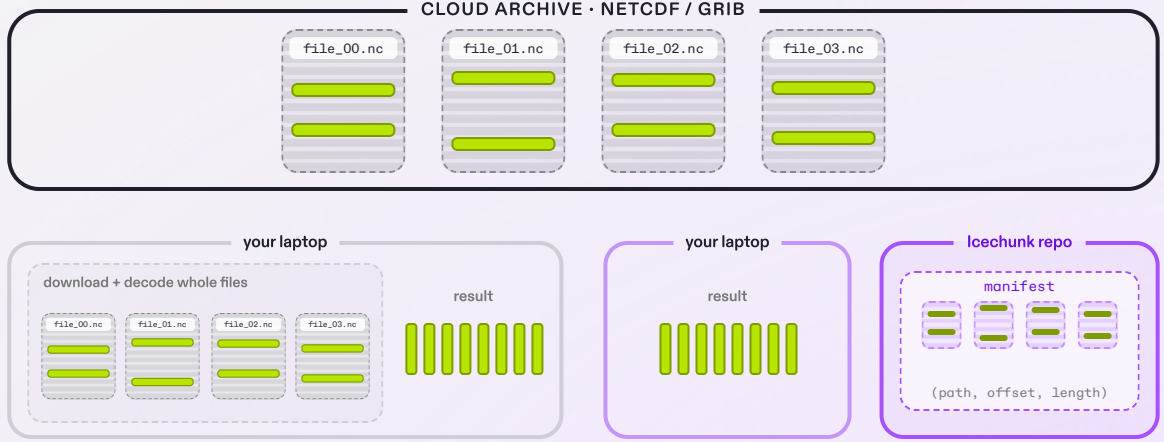
$0.18

16 s · 2.00 GB downloaded

VIRTUAL CHUNKS BILL

$0.01

1 s · 128.0 MB downloaded



100× NASA PILOT · GPM IMERG

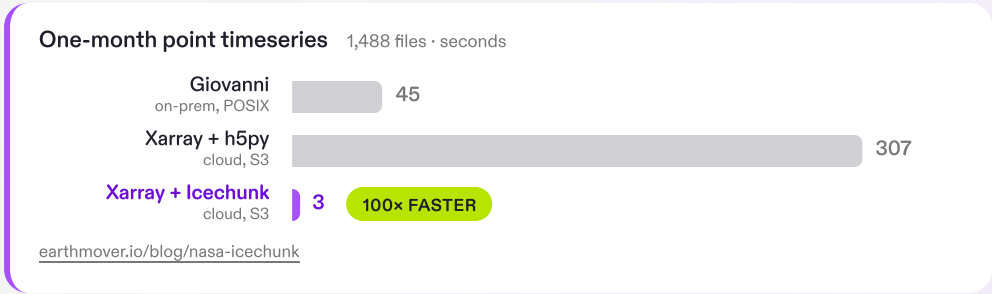
1,488 NetCDF granules virtualized in place on S3: a one-month timeseries in **3 s**, vs 5 min reading the files directly.

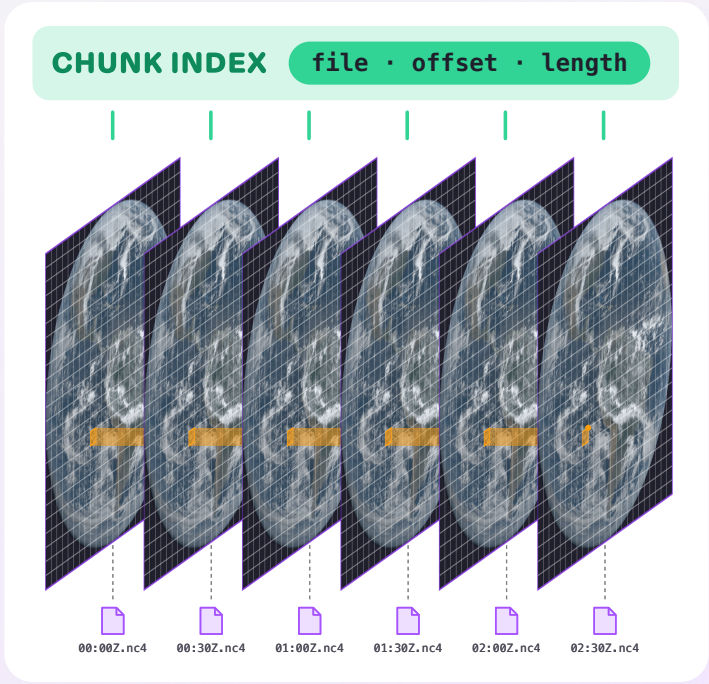
ICECHUNK IN THE WILD

results at NASA

PILOT: GPM IMERG PRECIPITATION

- NASA, Earthmover and Development Seed piloted Icechunk on **GPM IMERG:** a global precipitation snapshot every half hour, **1M+ NetCDF files** in S3
- VirtualiZarr scanned every file and stored a **chunk index** in Icechunk: **no data copied**, no format migration
- The whole archive opens as **one analysis-ready Zarr cube**





ICECHUNK

ACID transactions

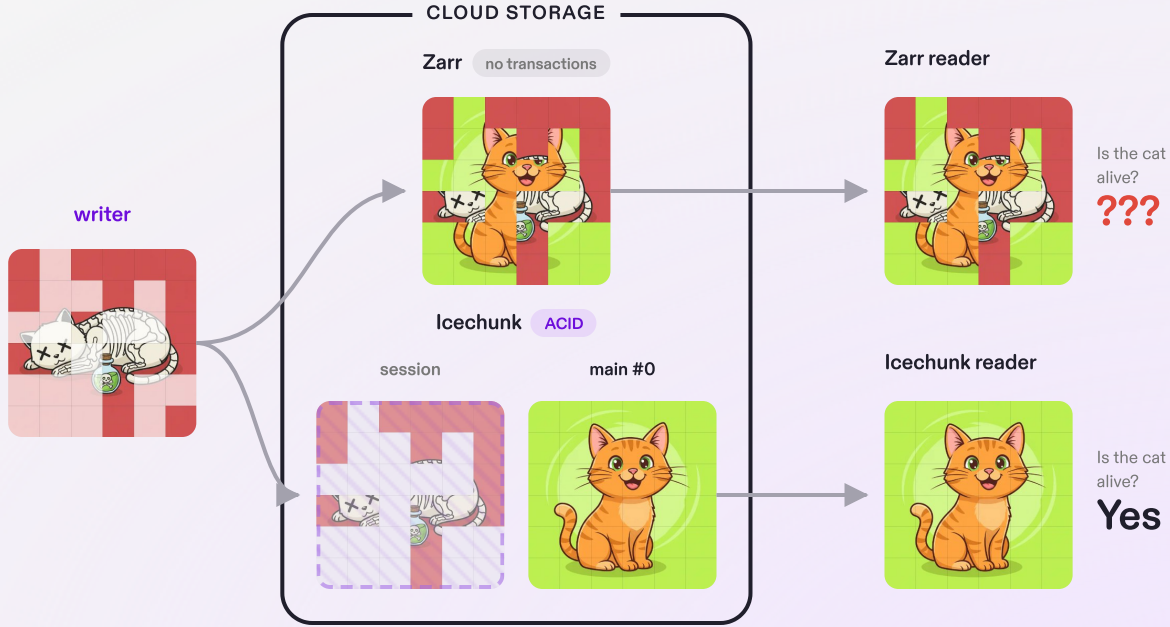
No more Schrödinger's arrays.

Without transactions, a reader can catch a write halfway: some chunks from before, some from after. An array that never existed.

Atomic commits. A write to any number of arrays lands all at once, or not at all.

Consistent reads. Readers always see one complete commit, even while another process is writing.

Conflict detection. When two writers commit at once, the second is rejected or rebased, never silently overwritten.

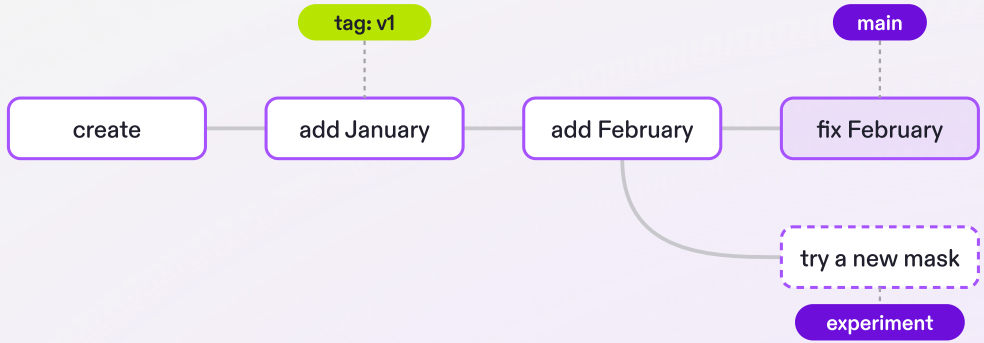


ICECHUNK

data versioning

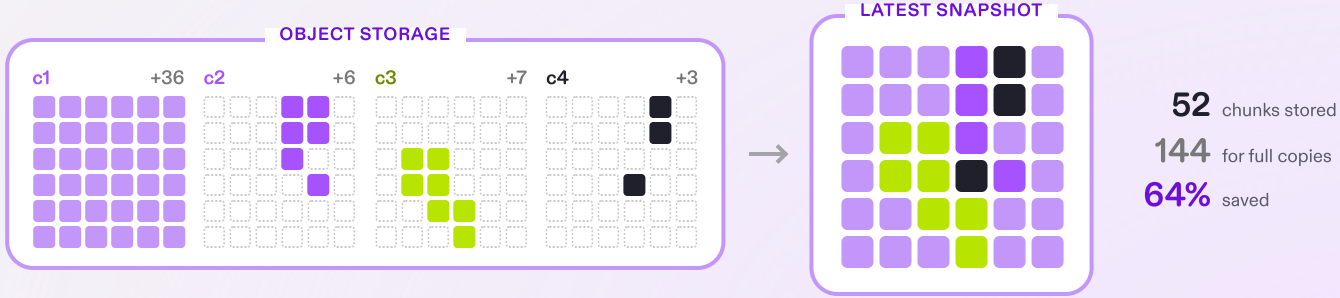
branch, tag, roll back

Every change is a commit. Undo a mistake by moving the branch back, tag a commit to cite the exact data in a paper, or branch to try a change without affecting anyone else.



history without ballooning storage

Keeping history doesn't mean keeping full copies. A commit writes only the chunks that changed; every other chunk points at data already in storage.



ICECHUNK

open source, no server required

The Icechunk library runs inside your code and reads and writes files directly: no server, database or service to run.



Apache 2.0 license
free to use, modify and embed

Open format spec
versioned, public, implementable in any language

github.com/earth-mover/icechunk
developed in the open

ICECHUNK

icechunk adoption

704
GitHub stars

57
contributors

424K
PyPI downloads last month

as of November 2026

SPACE & EARTH OBSERVATION

NASA · JPL Jet Propulsion Laboratory California Institute of Technology · cnes · EUMETSAT · ASF · spire

WEATHER, OCEAN & INTERGOVERNMENTAL

National Weather Service · World Meteorological Organization · NATO OTAN

NATIONAL RESEARCH INSTITUTES

CSIRO · ACCESS National Research Infrastructure · National Oceanography Centre · UK Atomic Energy Authority

AI & COMMERCIAL WEATHER

NVIDIA · dynamical.org · PLANETTE · RENEWCAST WE PREDICT YOUR VALUE · Statkraft

HEALTH & LIFE SCIENCES

dhis2 · ALLEN INSTITUTE · CZ BIOHUB · sgkit · zetta.ai

CLIMATE ANALYTICS, DEFENSE & ENGINEERING

Eagle Rock Analytics · (carbon)plan · Raytheon · Booz | Allen | Hamilton · TETRA TECH



zax

from scripts → to queries

BIG DATA PROCESSING

status quo: data processing scripts

analyze_data.m

```
1  files = dir('data/*.nc');
2  for i = 1:numel(files)
3      sst = ncread(files(i).name, 'sst');
4      anom = sst - mean(sst, 3);
5      result(i) = max(anom(:));
6  end
7  save('result.mat', 'result');
```



RUNNING...
file 23 of 96
1 h 34 min elapsed

data/*.nc i = 23





1 of 64 cores busy

SLOW ON TWO LEVELS

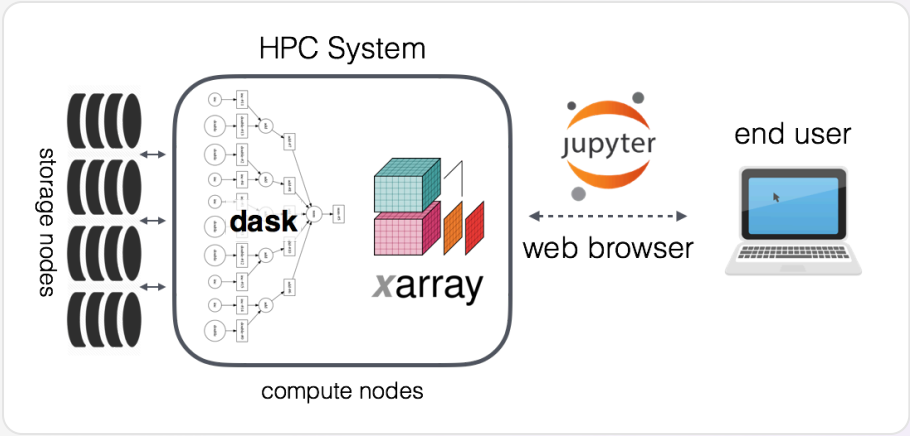
1 **Slow processing code**
Hand-written per file, never optimized.

2 **No parallelism**
One file at a time: the big cluster sits idle.

THE DEEPER PROBLEM The script couples the what (science logic) to the how (the physical execution plan).

BIG DATA PROCESSING

a big step forward: Python, Xarray, Dask and the Pangeo ecosystem



Xarray — labeled N-dimensional arrays for Python: named dimensions, coordinates and metadata on top of NumPy.

Dask — parallel computing for Python: scales NumPy / Pandas analysis from a laptop to a distributed cluster.

Pangeo project · pangeo-data.github.io

ADVANTAGES

- ✓ **Broke free from the centrality of files** — distributed data processing.
- ✓ **Generic, shared tools** instead of everyone building bespoke ones.
- ✓ **Larger data volumes** and more ambitious science.

LIMITATIONS

At scale, still the province of experts.

- ✗ **Manual configuration** — e.g. chunk size strongly drives performance.
- ✗ **Clusters frequently OOM and crash** → expensive iteration or over-provisioning.
- ✗ **Loose coupling between layers** of the stack makes optimization hard or impossible.

ZAX

copying from analytical database engines

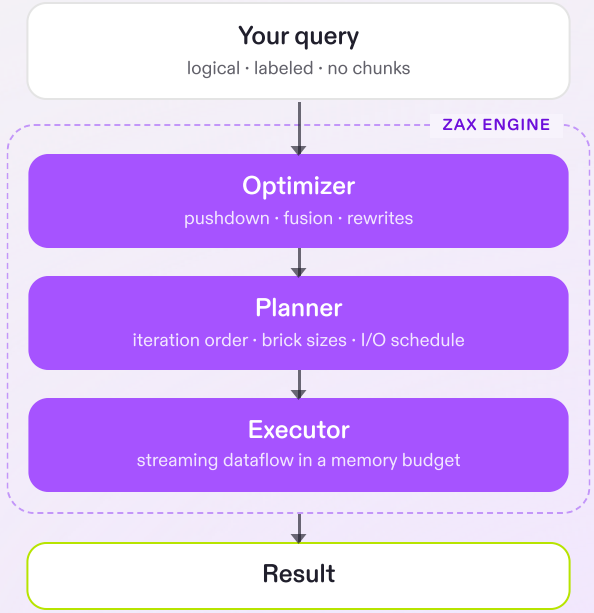
designing a modern query engine for arrays

Declarative	Say *what* to compute over named dimensions	~~science logic tangled with execution~~
Logical level only	No chunk sizes, ever again	~~manual chunk configuration~~
The engine decides	Optimized iteration order, chunk sizes, I/O scheduling	~~loose coupling between layers~~
Predictable	Run within a memory budget, or fail fast	~~clusters that OOM and crash~~
Proven ideas	Planners, optimizers, cost models, views	

INSPIRATION

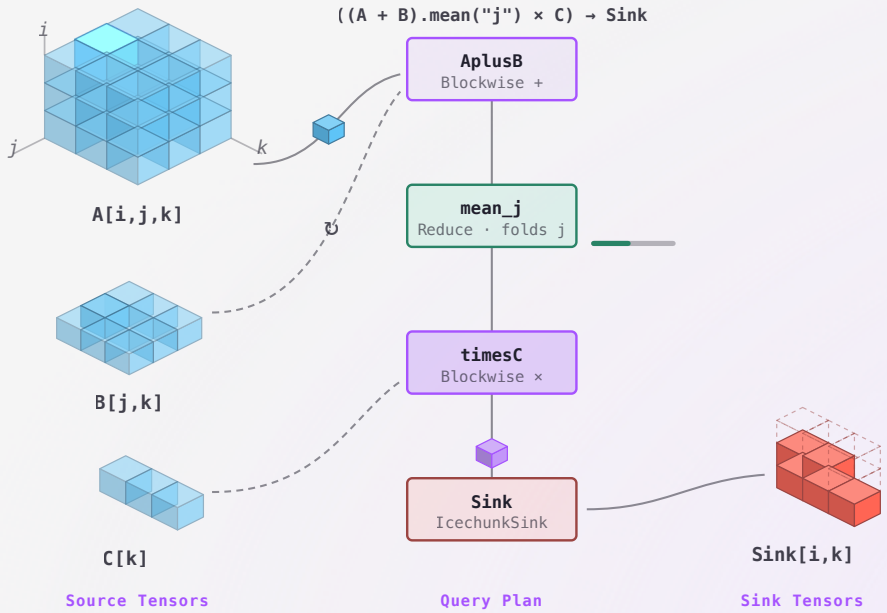
Polars **DuckDB** **DataFusion** **NVIDIA CUDA Tile**

tabular engines that check every box above: Zax is the array analog



ZAX

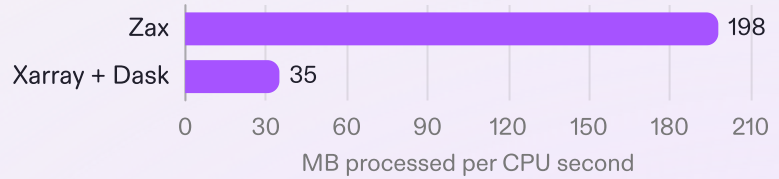
a distributed tensor query engine



DESIGN

A Rust engine that combines the techniques of analytical query planners (**DataFusion**, **Polars**, **DuckDB**) with chunked tensor algebra (**OpenAI Triton**, **NVIDIA CuTe**, **CuTile**). Array expressions are planned as a whole and streamed chunk by chunk, rather than run as a Python task graph (**Xarray** + **Dask**).

COMPUTE EFFICIENCY



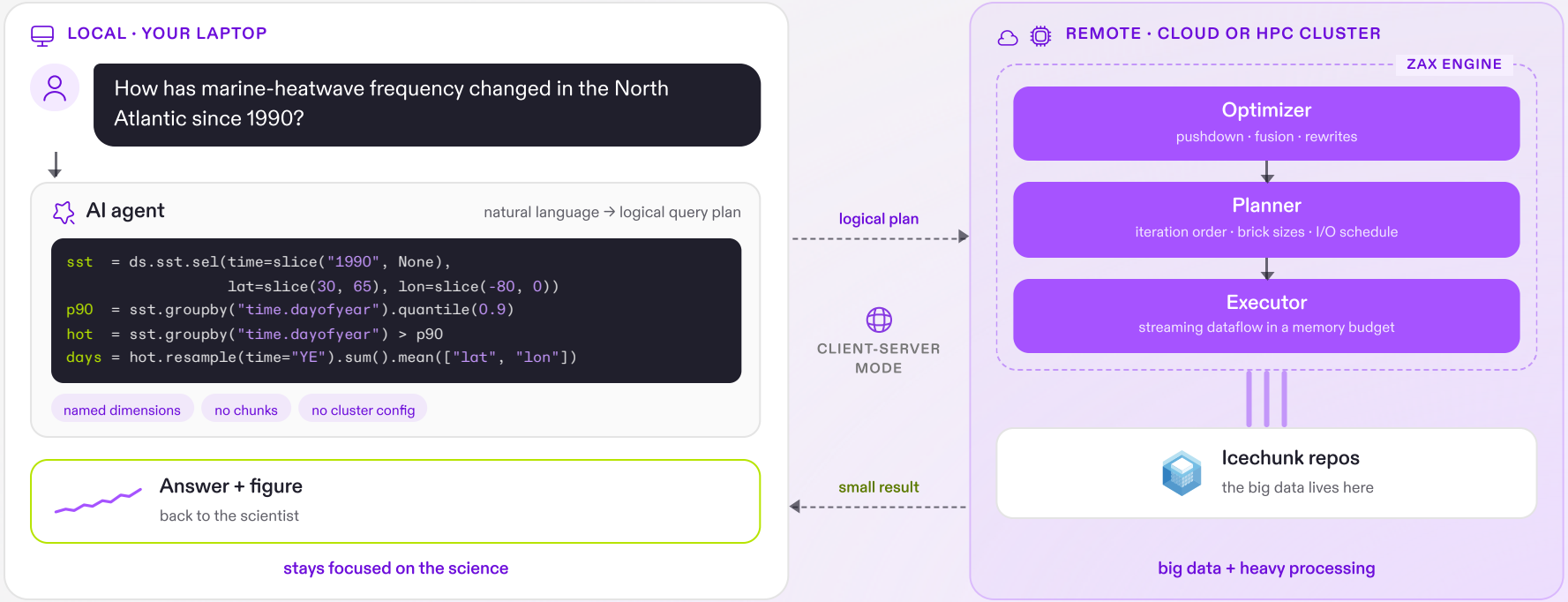
WORKLOADS

- advanced feature engineering
- high-performance streaming dataloading
- complex spatiotemporal joins for multimodal datasets
- ensemble statistics for model evaluation and scoring

ZAX

zax + agents: a perfect match

the agent stays at the science level · zax owns the computational workflow





earthmover

the science data cloud

THE SCIENCE DATA CLOUD

the AI-ready database for natural-world data

INTERFACE

Humans and agents
Same interfaces, same permissions

Humans: web app · CLI · Python

Agents: MCP · skills · CLI · Python

COMPUTE

Compute
Powered by Zax, our tensor query engine

Zax-SQL · tiles · EDR · DAP2 · openEO · Zax Server · soon

Bring Your Own Compute
Read and write the open formats directly

NumPy · xarray · PyTorch · TensorFlow · dask · Polars · Spark · DuckDB · and more

CATALOG

Arraylake
Data management and governance

catalog · access control · teams and SSO · credential vending · Iceberg REST catalog · webhooks · background optimization

Marketplace
Exchange for scientific data

free and paid data · zero-copy subscriptions · private exchanges

OPEN FORMATS

Icechunk
Open-source tensor storage engine

ACID transactions · version control · zero-copy ingest · fast cloud I/O

TENSORS
for continuous fields

Apache Iceberg
Industry-standard open table format

ACID transactions · time travel · any SQL engine

TABLES
for discrete records

Open source, in your own bucket — readable with or without Earthmover.

CLOUD

Bring your own bucket
Or use storage we manage for you

aws · CLOUDFLARE · tigris · VAST · generic S3-compatible

EARTHMOVER IN PRACTICE

national agencies are moving their data to the cloud

NASA

IMVI

Earth-system model data, analysis-ready in NASA's cloud

Served from NASA's own cloud. In the pilot: up to 220× faster time-to-analysis.

MERRA-2 reanalysis

ECCO ocean state

LIS land surface

NOAA

OAR · CEFI

A cloud-native portal for ocean-climate projections

Climate, Ecosystems & Fisheries Initiative; beat the existing server in benchmarks.

CEFI ocean & ecosystems

NWS · CIRRUS

A cloud repository for Weather Service forecasts

Icechunk chosen as a core format for operational forecast data.

NWS forecast data

OCIO

An agency-wide data architecture

NOAA's CIO office is evaluating Earthmover for data use cases across the agency.

NOAA-wide enterprise data

Met Office

MOGREPS-UK

Analysis-ready, cloud-optimized ensemble forecasts

Output from the UK's high-resolution forecast ensemble, as ARCO data.

MOGREPS-UK ensemble

CONCLUSIONS

we're entering a golden age for data-intensive science

THE SCIENTIST'S ROLE IS EVOLVING WITH AGENTS

Less coding
and manual data munging

Faster exploration
of hypotheses with agents

Bigger questions
and better science

BUT DATA INFRASTRUCTURE NEEDS TO KEEP UP

~~file archives~~

AI-ready databases — Icechunk

~~scripts~~

analytical queries — Zax