

HIIJR Ocean

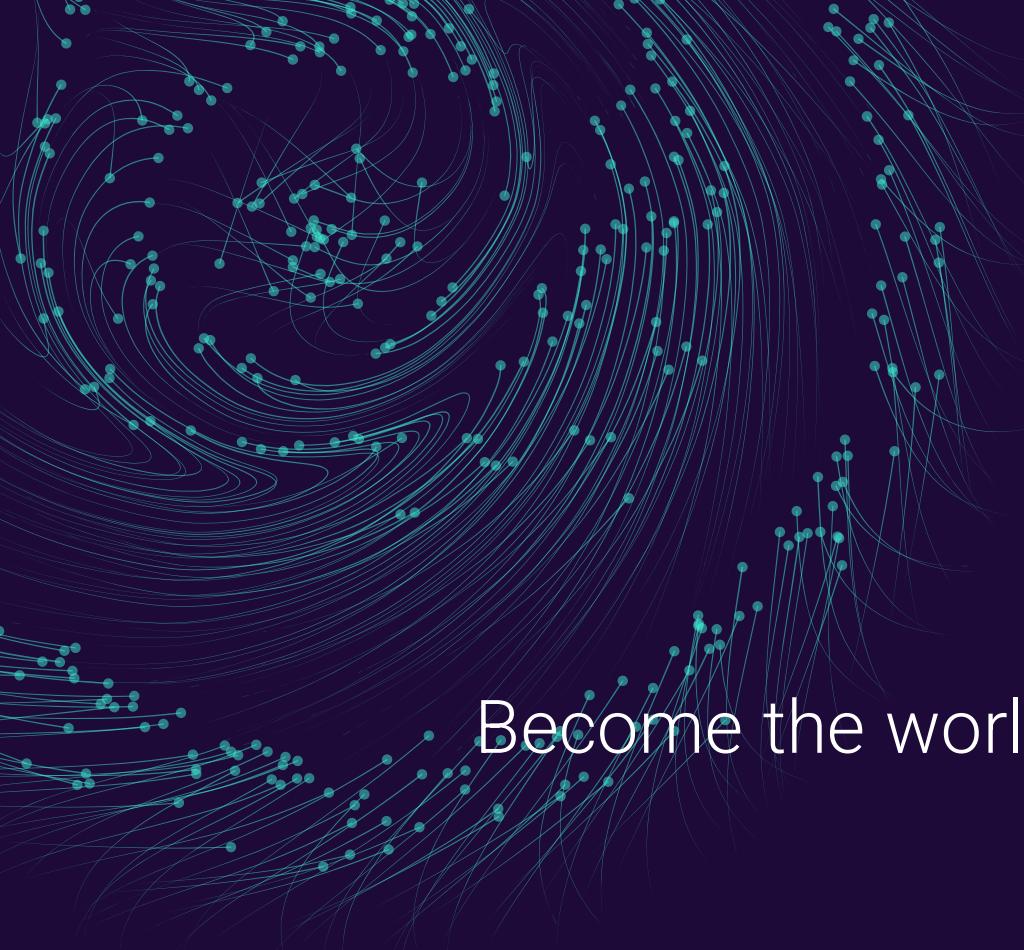
Centre for the Fourth Industrial Revolution



Thomas Li Fredriksen

Lead Solution Architect, HUB Ocean

Based in Stavanger, Norway



HUJR Ocean

Our Vision

Become the world's ocean data collaboration hub

...advancing both ocean health and wealth

HLR Ocean

A non-profit ocean foundation

Founded by



Supporting



United Nations Decade
of Ocean Science

Leading

Ocean Data Action Coalition



In partnership with



Operating



A trusted partner to major global organizations across:

Industry



BAHR
>
accenture



AKER BIOMARINE

MERCURIA

AkerSolutions

Government

WORLD
ECONOMIC
FORUM

Maersk Mc-Kinney Møller Center
for Zero Carbon Shipping

AKER HORIZONS

NOA
The North
Alliance

aramco



IOC DE
International
Oceanographic
Data and Information
Exchange

REV Ocean

SINTEF

NTNU

ØRSTEGES ARKTISKE UNIVERSITET
UIT

VANORA

Global Fishing Watch

Science

Addressing the Ocean Emergency

Ocean Assets Valued at **\$24** trillion

80% of All Life Lives in the Ocean

50% of Earth's oxygen comes from the Ocean

The ocean absorbs **30%** of the world's CO₂

...and **90%** of the excess heat

The Key Stressors on the Ocean

MARINE
POLLUTION

ILLEGAL OR
UNREPORTED
FISHING

OCEAN
WARMING

OCEAN
ACIDIFICATION

OCEAN
HYPOXIA

causing



SEA LEVEL RISE

Flooding coastal communities and drowning wetland habitats



BLEACHING

Warm water coral reefs will be lost if the planet warms by 2°C



TOXIC ALGAE

Larger and more frequent blooms are making animals and people sick



HABITAT LOSS

Lower oxygen levels suffocate animals and shrink habitats



ACIDIFICATION

More acidic water prevents animals from building their shells



FOOD INSECURITY

Disruptions in fisheries affect the marine food web and human food security



Map & manage

100% Science-based Ocean Management by 2030



Renewable energy

40x more by 2050



Sustainable food

6x more by 2050



Green transportation

Zero-emission shipping by 2050



Sustainable investments

& Risk management



Reduction of emissions & pollution

From the world's top sources



Advances in scientific computing

Rapid innovation with more and more open solutions

2003 – Matplotlib

2006 – Numpy

2008 – Pandas

2011 – IPython Notebook

2015 – Tensorflow/Keras

2015 – Project Jupyter

2022 - ChatGPT

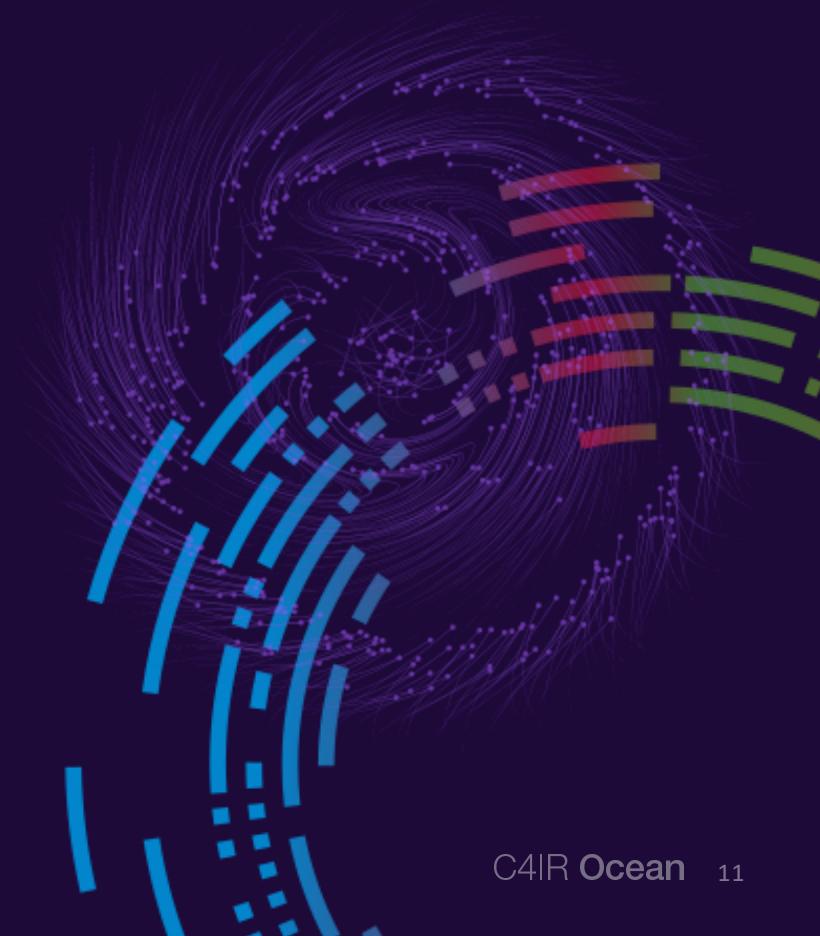
Image credit:

- US Department of Energy - <https://ascr-discovery.org/2007/07/supercomputer-saga/>
- Microsoft - <https://news.microsoft.com/innovation-stories/microsofts-virtual-datacenter-grounds-the-cloud-in-reality/>
- Karsten Madsen from Pexels

Scientists want to spend time on science

... and not cleaning data or setting up complex computers

- Data scientists spend around 40% of their time with data wrangling tasks.
 - New and better tools are helping
- Data volume is a growing obstacle for most research
 - Personal laptops are often not powerful enough to process large amounts of data
 - Operating an on-premise supercomputer is a complex task
- Complex and ambitious research can be completely out of reach to resource-constrained institutions.



Citizen science example



The Great Reef Census is a groundbreaking citizen science effort to survey the Great Barrier Reef. Whether you're on the Reef or on the other side of the world, you can be part of this important conservation initiative.

Every boat, every Citizen, every reef.

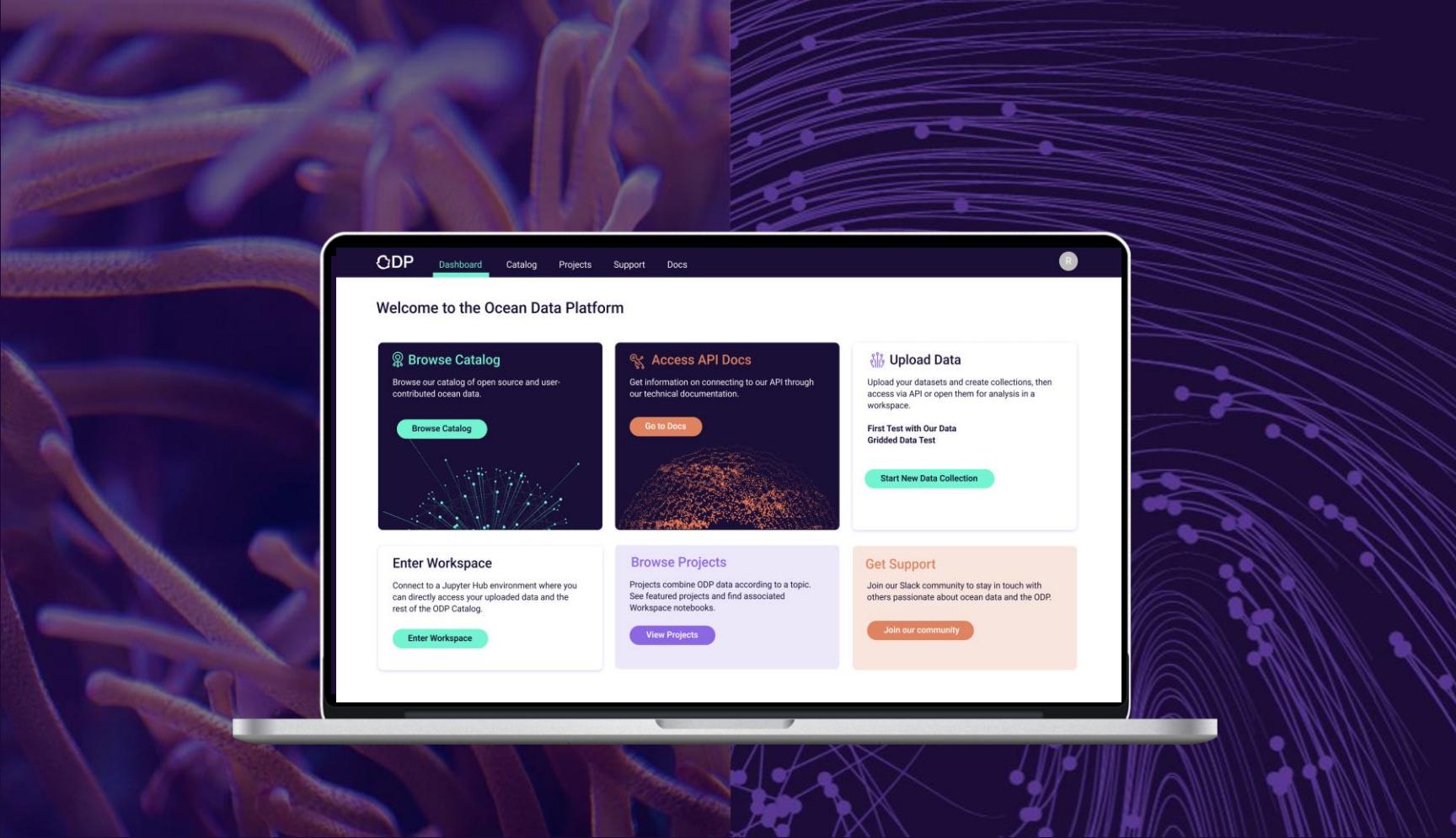
Analyse an image BETA Join us in the water

Images **79.2k** Analysed **75.3k**

Survey sites **4,081** People **7,352**

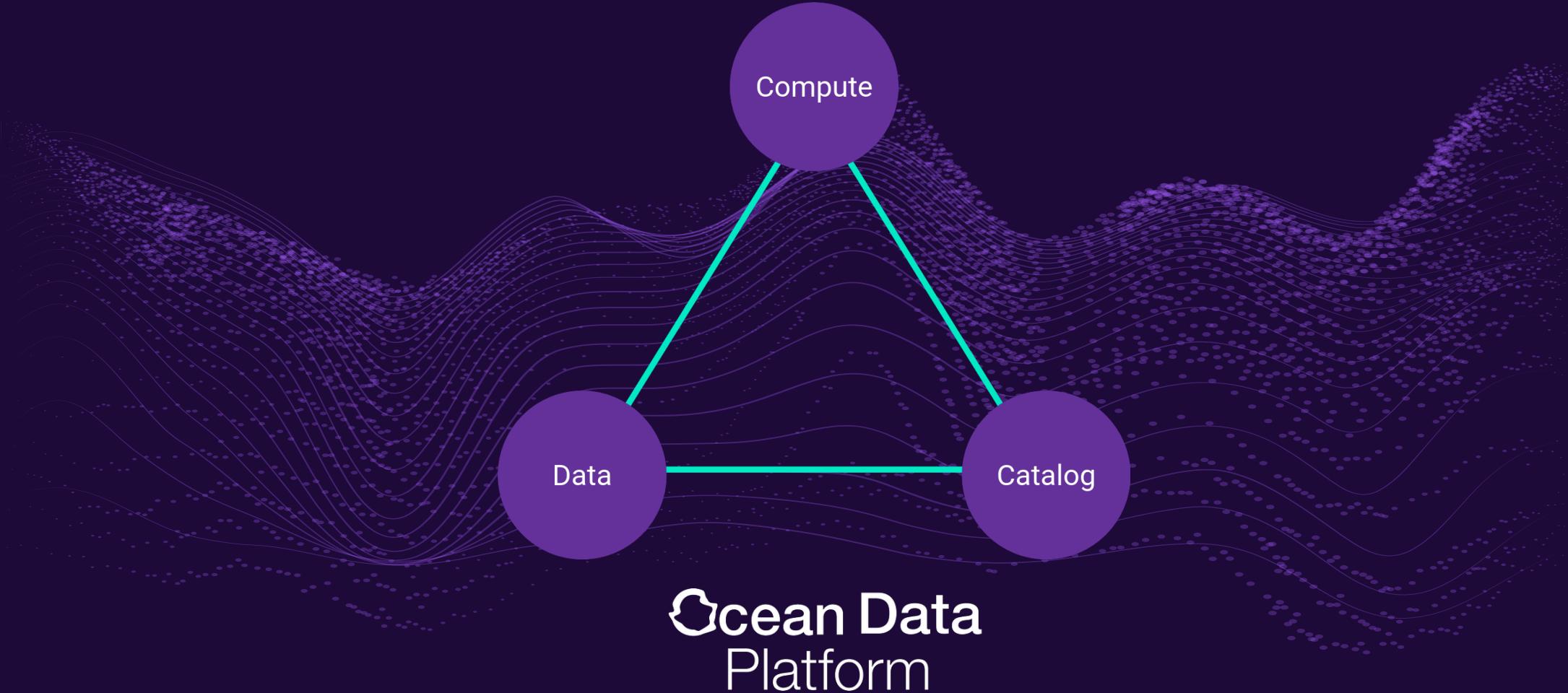


Our Solution:

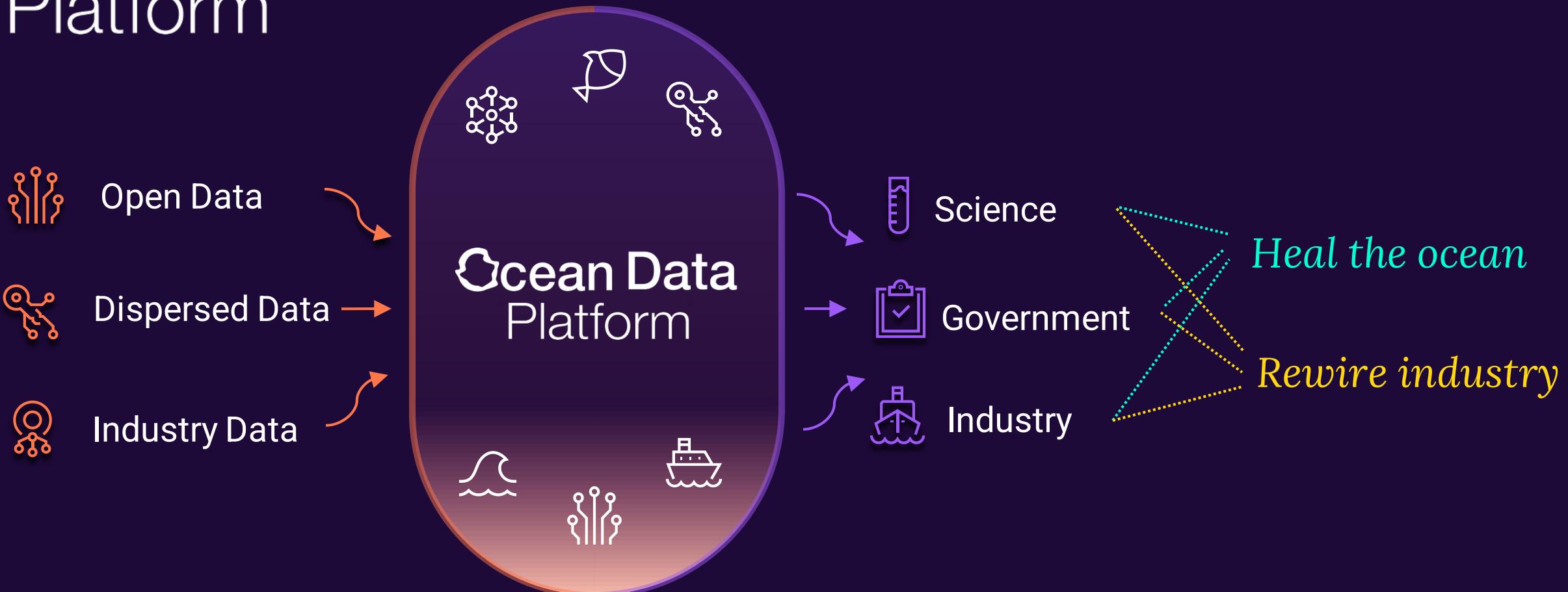


Ocean Data Platform

A *one stop shop* for ocean data



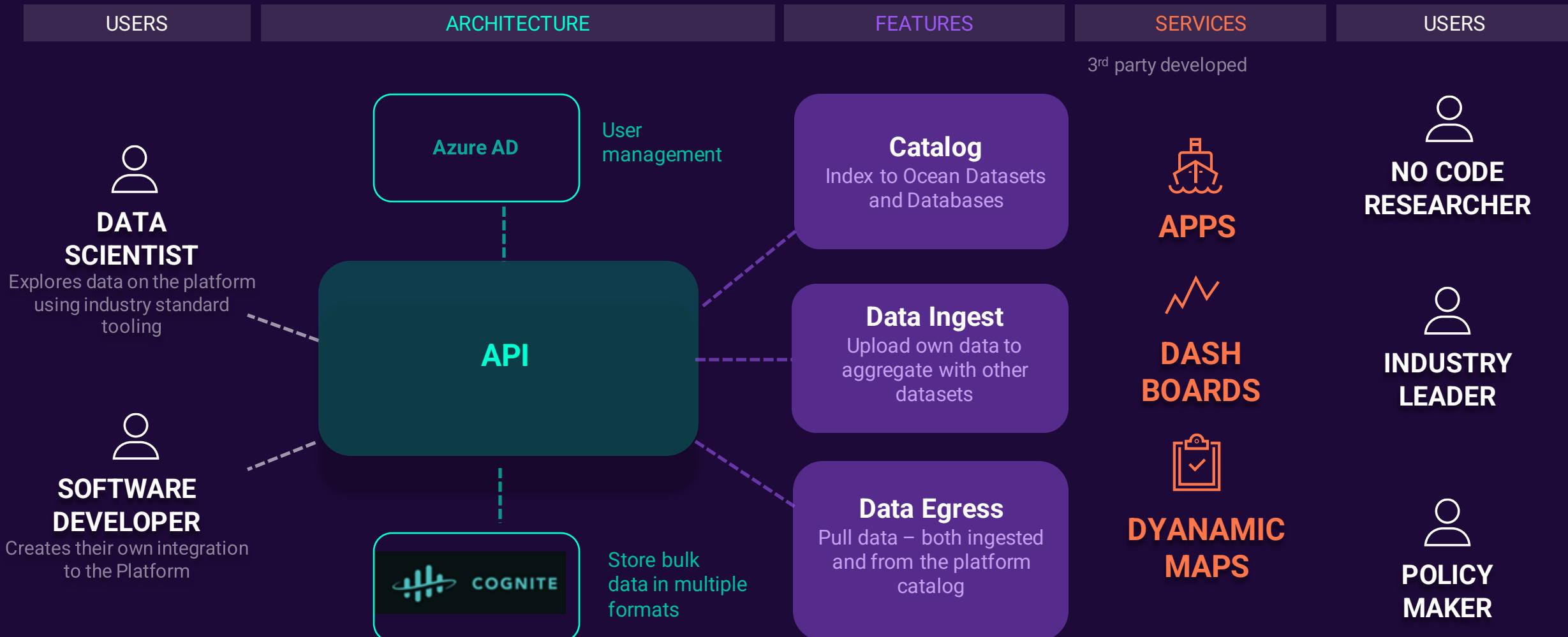
Ocean Data Platform



HIIIR Ocean

Uses & Architecture

Ocean Data Platform



A cloud-based approach

A composite image illustrating the cloud-based approach. It features a smartphone on the left, a laptop in the center, and a large globe on the right, all displaying data analysis and visualization software. The smartphone screen shows a file explorer with a folder named 'World Ocean Database'. The laptop screen displays a Jupyter Notebook interface with code and two maps: one showing a world map with colored dots and another showing a global map with contour lines and dots. Red arrows point from the text descriptions above to specific elements on the screens.

Cloud-based environment with computing power

Direct access to open source data such as the World Ocean Database

Built-in functions for fast exploration of the data

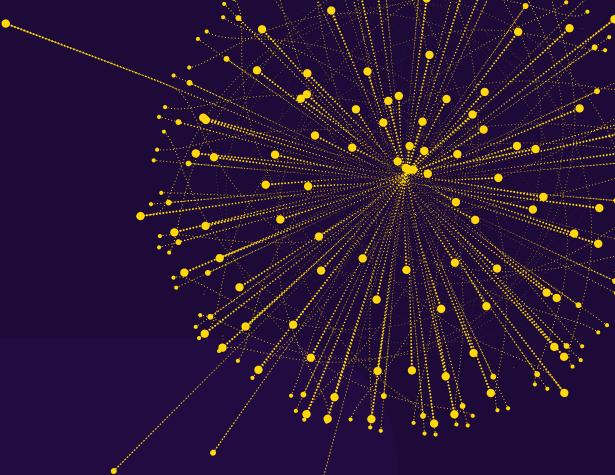
A geospatial platform

Built, operated and maintained by

HU IR Ocean

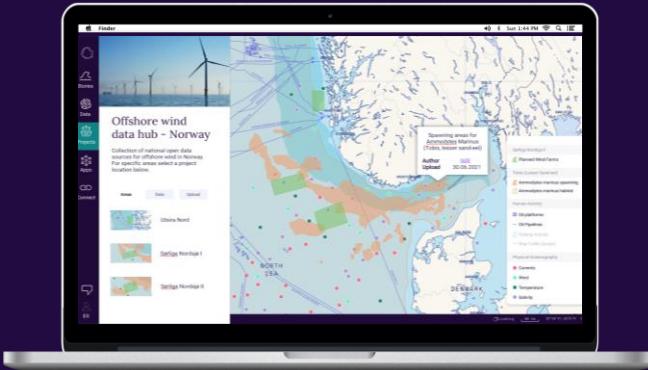
PLATFORM FEATURES

- Self-service publishing of data with DOI
- Data from multiple sources and disciplines in one place
- Full data lineage
- Access to data through APIs
- Data served in uniform formats
- Computational power and a direct connection to data through a JupyterHub environment
 - Pre-installed packages
 - Built-in functions for analysis and figures
 - Low barrier to start coding in Python



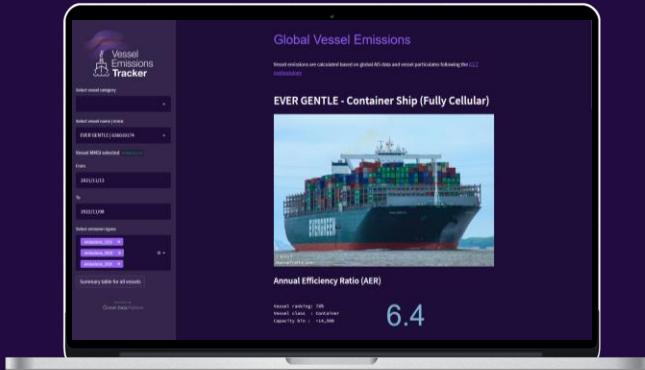
Sample Use Cases Relevant for Industry

OFFSHORE WIND



Environmental impact assessment
of potential site locations

SHIPPING



CO2-emissions estimate for
voyages.
Time spent in marine protected
areas.

FINANCE



Environmental risk assessments
when evaluating new investments

Data Marketplace

Incentivise data sharing

Data Offering

- Data providers can share access to their data for a free or for a price.
- Users can create new data products based and share it.

Data Demand

- Sometimes the data we need does not exist.
- Users can request the data be provided by other users.
- Data collections or new data products.

World Ocean Data, One Place

Below is a sample of data on the Ocean Data Platform. Users will be able to pull and analyse data through API, allowing for powerful fusion of

data. Along with large open-source datasets, industrial data will also be available, allowing more transparent and layered analysis



DATA TYPES

1

Physical
Oceanography

Temperature,
Salinity, Density,
Pressure

2

Chemical
Oceanography

Nitrate, Silicate,
Carbon, Oxygen,
pH

3

Geological
Oceanography

Bathymetry,
Sediments,
Hydrothermal vents

4

Biological
Oceanography

Phytoplankton,
Primary Productivity

5

Marine
Biology

Biodiversity,
Occurrence data
on organisms

6

Human Activity
and Boundaries

Marine regions,
EEZ, wind farms,
platforms,
aquaculture
facilities etc.

DATA Variety

1

Tabular

Physical measurements,
Financial transactions, list of addresses

2

Gridded

Weather, Climate models, interpolated maps

3

Rasterized

Satellite imagery, elevation maps, land cover maps, bathymetry

4

Video

Journey recordings, webcam streams

5

Audio

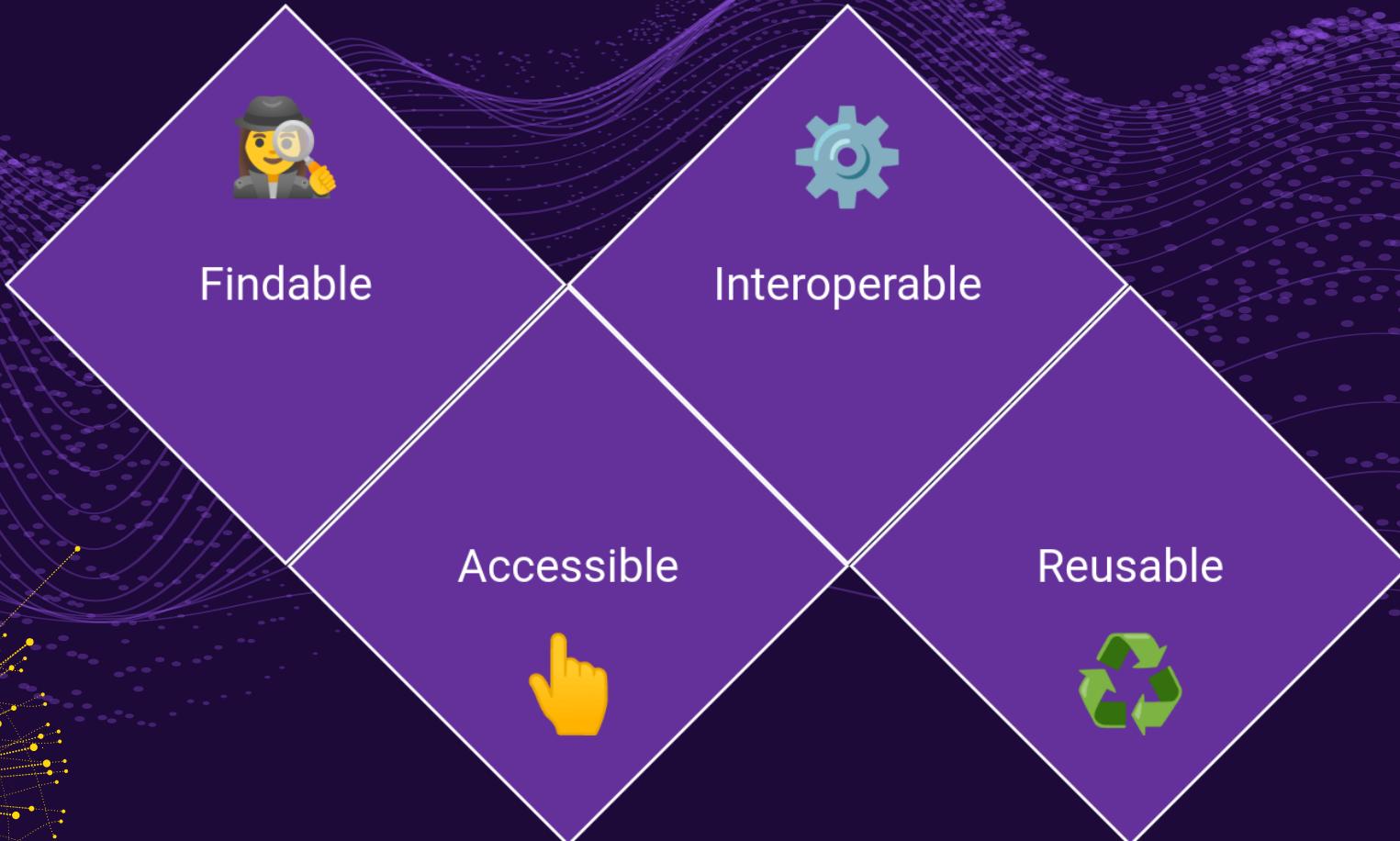
Sound recordings, Hydrophonic recordings, whale song

6

Point clouds

3D Scanning and mapping

What is "F.A.I.R"?



What makes a "F.A.I.R" data platform?

Translating fair principles to design criteria

Leverage metadata

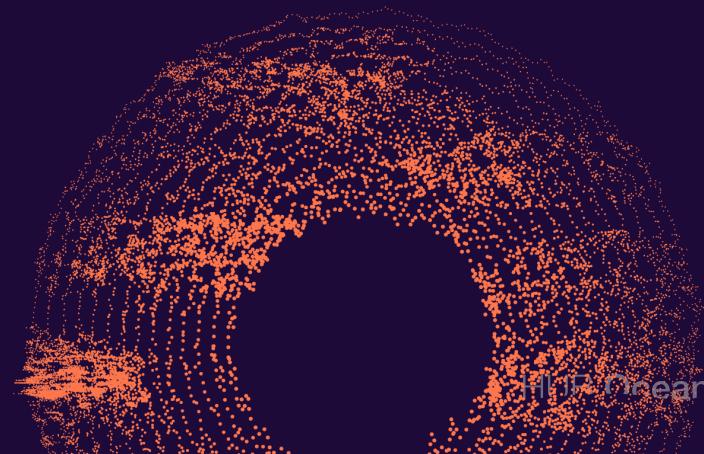
- Define an extensible but simple metadata framework.
- "Borrow" good ideas from others instead of reinventing the wheel.
- Specify an essential set base of metadata but allow users to add on top of that.

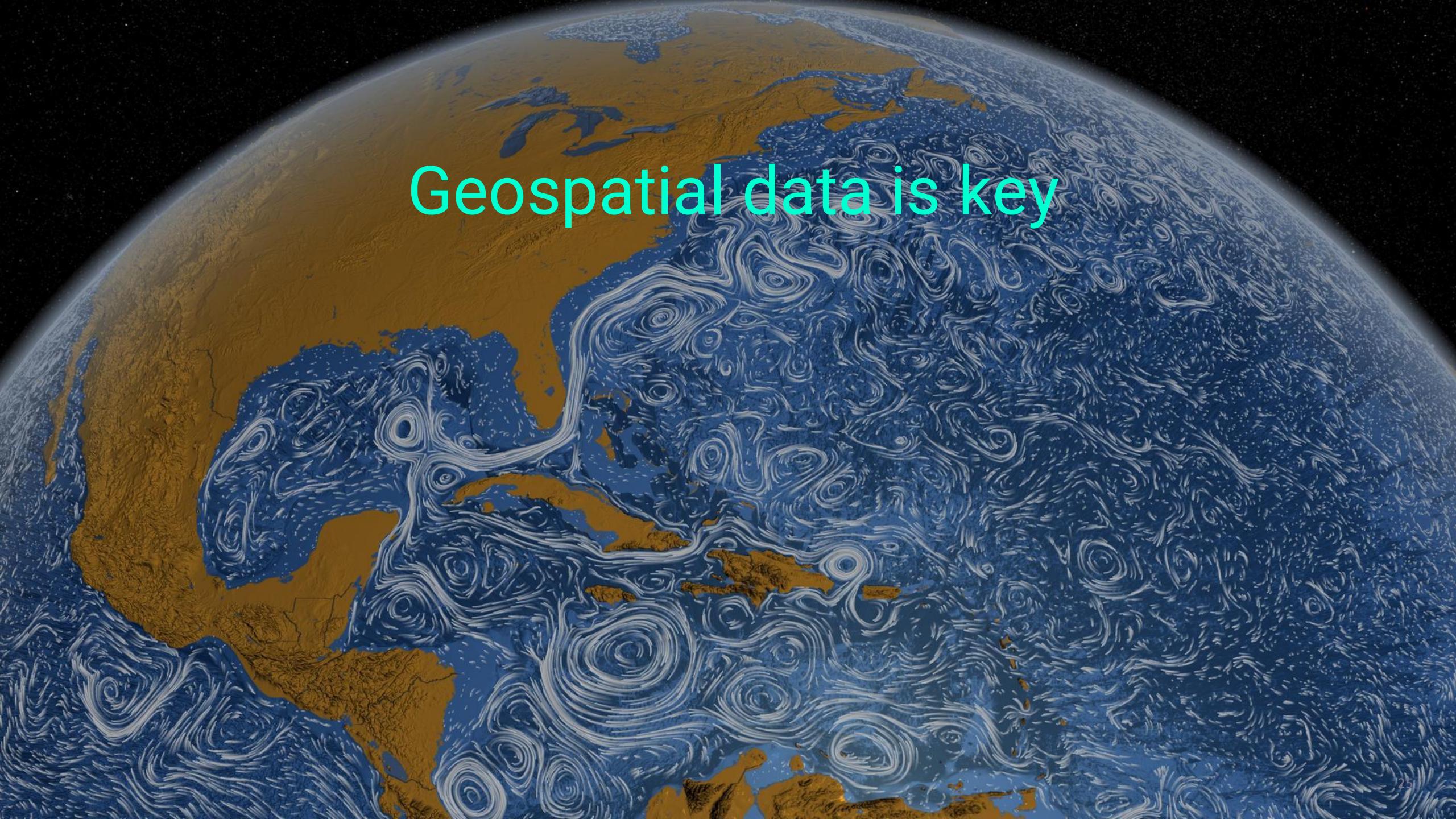
Cater to the user

- Recognize that there are different kind of users and needs
- Provide solutions that solve user problems instead of "our" problems
- Shape the technology stack to provide solutions to both low-code/no-code users and advanced users

Open Source

- Rely on already established frameworks and methodologies
- Share our work with the world
- Collaborate instead of compete



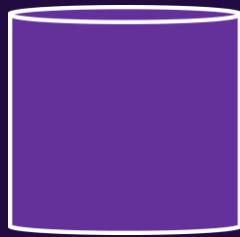
A satellite map of Earth, focusing on the Atlantic Ocean and Europe. The map shows cloud formations as white and grey swirls against a blue background. Landmasses are shown in brown and yellow. The text "Geospatial data is key" is overlaid in the upper right quadrant.

Geospatial data is key

How to model Data

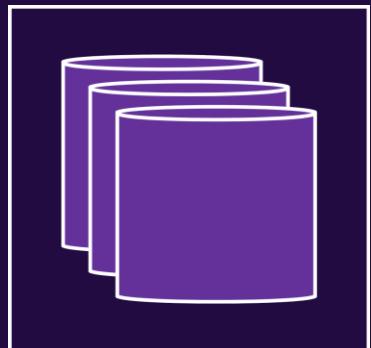


DATASET



A collection of data that is organized or structured in a certain way

DATA COLLECTION



Grouping or aggregation of datasets that have a common theme

OBSERVABLE

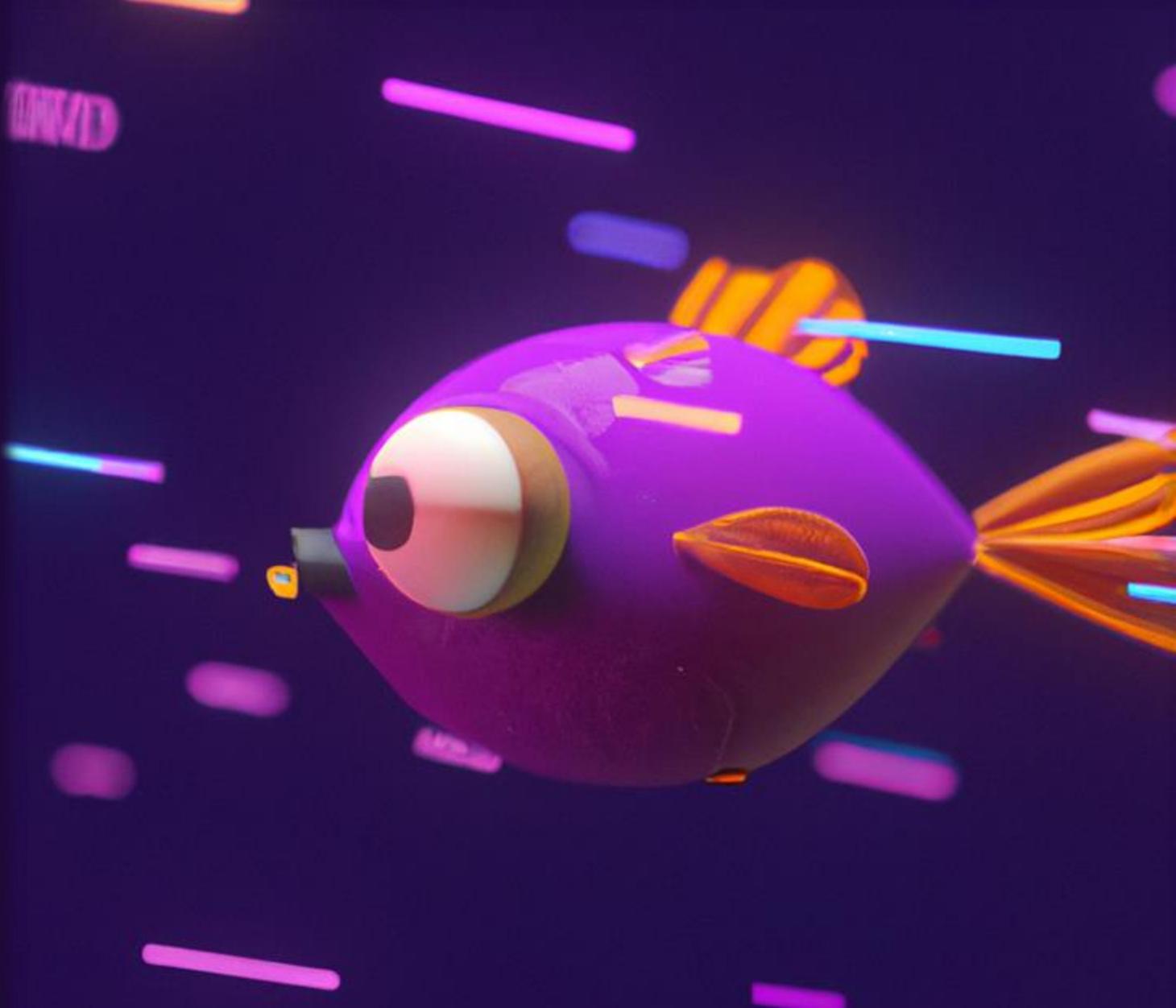


A specific measurement or characteristic derived from a dataset or data collection

Datasets

What metadata do we need?

- What class of data?
 - Tabular/Gridded/etc
- Who is the maintainer?
- How to cite?
- Where to find documentation?
- Licencing?
- How to aid with discoverability?



Datasets

Attributes and traits

- Microsoft has created a nice attributes and traits-framework as part of their CDM
- Define an umbrella-term "dataset attribute":
 - Columns for tabular
 - Dimensions and properties for gridded
- Decorate attributes with "traits":
 - Unit
 - What is measured
 - Coordinate reference system (goespatial)
 - Currency (financial)

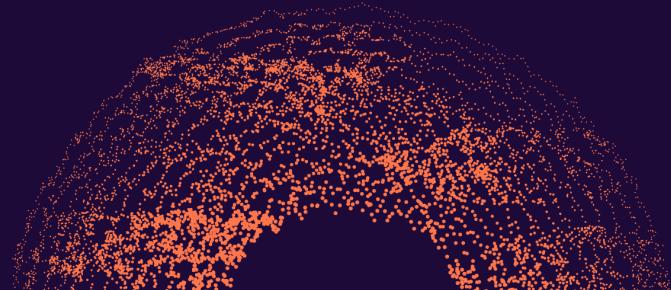


Observables

Examples

Aggregates

- Column statistics
 - Average
 - Variance/Standard Deviation
 - Min/max-values
- Coverage
- Geometrics aggregates
 - Bounding box/coverage
 - Area/circumference



Data Health

- Accuracy
 - The degree to which data correctly describes the real-world object or event being described
- Completeness
 - The extent to which all required data elements are present.
- Consistency
 - The absence of contradictions or discrepancies between different data sources or within the same data source.
- Validity
 - The extent to which data conforms to defined business rules or constraints.



Data Contracts

Formal agreement outlining structure, format, rules and guarantees of how to use a data product.

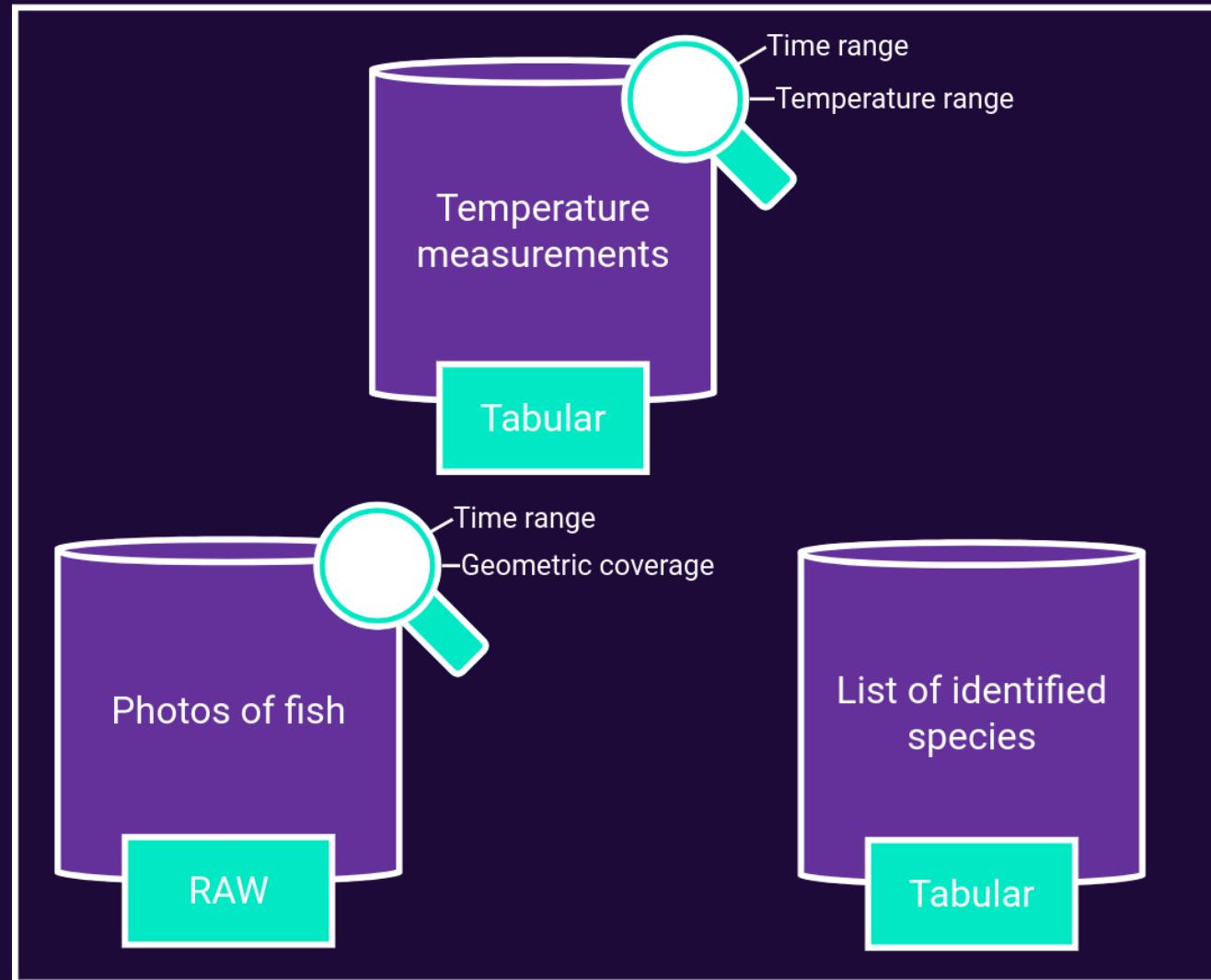
Why?

- Ensure that data is consumed in a specified way for the duration of the contract.
- Provides guarantees on data consistency and timeliness.
- Reduce uncertainty by eliminating errors and understanding

How?

- Data observables provide metrics on a dataset
- A data contract set valid ranges for these metrics
- Stakeholders are notified when observables report contract terms are breached

Diving Expedition Collection



From the user perspective

Data usage journey

Discover

Search for new data by geographic region, unit, maintainer, etc



Consumer & Process

Data can be downloaded directly or consumed in the ODP workspaces



Share results

Share results by defining a new data product



MAKE DATA SMART 2023



12. OKTOBER 2023



RADISSON BLU SCANDINAVIA - OSLO

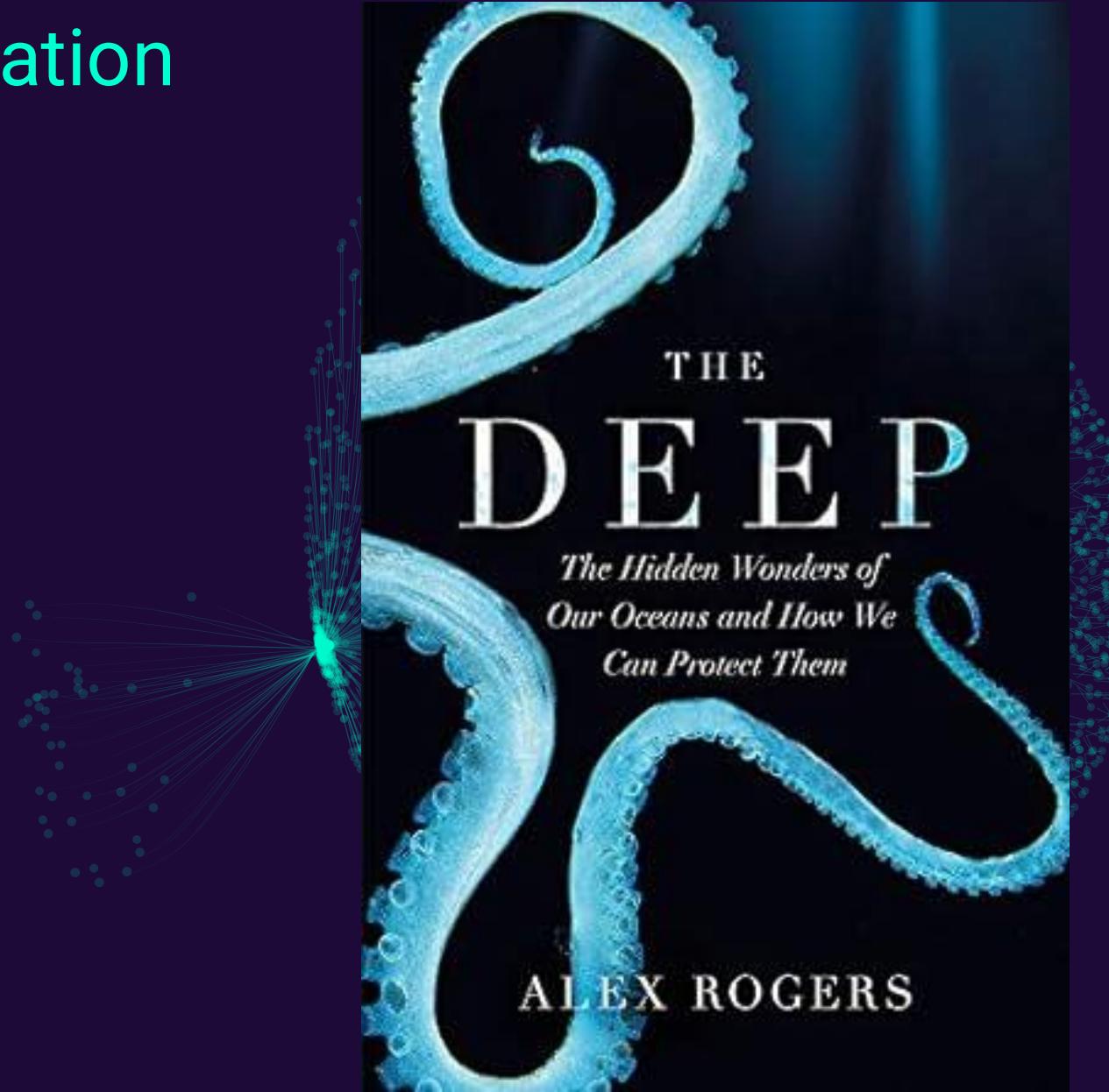


14:15

Mapping a Path to Ocean Sustainability: The Role of Big Geospatial Data and Sensors

Book Recommendation

"The Deep"
by
Alex Rodgers



H4IR Ocean

Centre for the Fourth Industrial Revolution

*Changing the fate of the ocean by unleashing
the power of data, technology, and collaboration*



Email

thomas.fredriksen@oceandata.earth



Website

www.hubocean.earth



LinkedIn

linkedin.com/company/hubocean



Slack

ocean-data-community.slack.com