MAKING CLIMATE DATA EASY TO FIND, USE, AND SHARE
Accurate, available, and consistently interoperable climate data is crucial to addressing our global climate crisis. Here is how OS-Climate is using Data Mesh to make climate data easy to find, consume, share, and govern.

October 28, 2022
Eric Broda

Eric Broda is an architect with OS-Climate and is the President of Broda Group Software, a boutique consulting firm that helps companies accelerate their Data Mesh journey.

OS-Climate is Tackling the Climate and ESG Data Challenge
OS-Climate (OS-C) is applying Open Source capabilities to help organizations effectively leverage ESG data. Using this data – in conjunction with Open Source analytical tools – businesses, financial institutions, government, and civil society can drive transparency and insights in the areas of Physical Risk & Resilience, Transition Analysis, and Portfolio Alignment with Net Zero and Adaptation goals. OS-C’s vision is to dramatically boost global capital flows into climate change mitigation and resilience.

OS-C’s mission is to help organizations determine if they are operating in, investing in, or partnering with companies that are aligned with Paris Accord Agreement goals and ESG values.
And to do this, OS-C is building a technical foundation of open source tools to organize and make available climate and ESG data.

But there is a fundamental and insidious obstacle in the way: **trustworthy** climate ESG data is exceedingly difficult to find, consume, share, and govern.

To overcome this obstacle OS-C is on a journey to adopt “Data Mesh” principles with a goal of making ESG and climate data easier to find, consume, share, and govern.

This article describes OS-C’s data mesh journey.

**TRUSTWORTHY CLIMATE DATA IS HARD TO FIND, CONSUME, SHARE, AND GOVERN**

ESG and climate data is the foundational enabler for OS-C. But this data is changing literally every moment, every day, every year:

- Volume of ESG and climate data is exploding, as infrastructure is instrumented to capture emissions data, and as more and more firms choose to publish their emissions data.
- Variability of ESG and climate data is growing exponentially, as more diverse data is made available.
- Value of ESG and climate data is increasing dramatically, as regulators – as well as investors – demand access to this data.

And this situation is complicated by a fundamental challenge: ESG and climate data is hard to find, consume, share, and govern. ESG and Climate data is:

- Distributed across literally thousands of data sources, making it hard to find.
- Unstructured, making it hard to understand and consume.
- Inconsistent, making it hard to share, compare, and aggregate.
OS-C’s goal is to overcome these obstacles and make ESG and climate data easy to find, consume, share, and govern. To achieve this goal, OS-C is:

- Creating an ecosystem of data products – also known as a data mesh – that makes ESG and Climate easy to find, understand, and trust
- Establishing a federated sourcing and governance model, aligned to Data Mesh principles, which combines public data, or “open” data which can be used by anyone, as well as private data which imposes custom contracts between climate data producers and consumers.
- Engineering a “self-serve” data platform, aligned to Data Mesh principles, to unlock ESG and Climate data

**OS-CLIMATE: ESTABLISHING AN ECOSYSTEM OF ESG/CLIMATE DATA PRODUCTS**

Data Mesh is now emerging as a modern approach to addressing large-scale data challenges for large enterprises. But the principles that make Data Mesh so valuable for enterprises also make it an ideal approach to addressing the data challenges OS-C and its ecosystem partners are tackling.

But what is a “Data Mesh”? First and foremost, there is the foundational book, *Data Mesh: Delivering Data-Driven Value at Scale*, by Zhamak Dehghani. And there are a number of great articles available [here](patterns) (patterns), [here](architecture) (architecture), [here](principles) (principles) and [here](lessons learned) (lessons learned) that will provide more detailed technical information about Data Mesh.
For our purposes, we will define Data Mesh as an ecosystem of interacting “Data Products”. A data product is, simplistically, a way of treating data as a product. Each data product:
- Has a clear boundary
- Participates in an ecosystem of producers and consumers
- Uses a platform to make it easy to find, consume, share, and govern data
- Publishes metadata enabling self-service and discovery
- Adheres to a clear set of governance standards while respecting local autonomy

Our reasons for applying data mesh principles are based upon the following rationale.

First, applying Data Mesh principles will make ESG and climate data products easier to find, consume, share, and govern:
- Data Mesh defines data domains – recognized boundaries and locations for trustworthy ESG and climate data – which provides a crucial first step in finding data in our highly distributed and diverse landscape of ESG and climate data
- Data Mesh identifies owners for data domains, which will empower individuals or groups to propose how to make data consumable.
- Data Mesh implements a "self-serve" model, where access to data domains is through standard and consistent mechanisms (for example, via APIs or other common/Open methods) making it easy to share data within and between organizations
- Data Mesh offers a "federated" governance model, that gives data owners the local autonomy to respect regional needs and diverse regulatory demands most effectively while also making it easier federate the governance and trustworthiness of data.
Second, by applying these principles OS-C can establish an ecosystem of data products (also known as a Data Mesh), where each climate data product is discoverable, addressable, interoperable, self-describing, trustworthy, secure.

OS-Climate’s ESG and Climate data product ecosystem will implement Data Mesh capabilities using Open Source components, including:

- A catalog of data that makes it easy to find data
- Common communication services that make it easy to access data
- Distributed “data products” that have clear boundaries and owners to enable rapid decision making
- APIs in each Data Product that make it easy to access and share data

**OS-C’s Ecosystem of ESG/Climate Data Products**

OS-C’s data mesh is an ecosystem of ESG and Climate Data Products, bound together by a common communication backbone and ESG and Climate Data Product Registry.

**OS-C Data Mesh – An Ecosystem of ESG and Climate Data Products**

![OS-C’s Ecosystem of ESG and Climate Data Products](image)

*Figure 3, OS-C’s Ecosystem of ESG and Climate Data Products*

With these data mesh capabilities, data products will make ESG and Climate data easier to find, consume, share, and govern, which will lead to better and faster insights, improved time-to-market for availability of ESG and Climate data, reduced cost and overhead of making ESG and Climate data available and accessible.

**OS-CLIMATE: ESTABLISHING A FEDERATED SOURCING MODEL**

In keeping with Data Mesh federated governance principles, OS-C is establishing a “federated sourcing” model which combines public data, or “open” data which can be used by anyone, as well as private data which imposes custom contracts between climate data producers and consumers.
OS-C’s federated sourcing model will strive to create “shared” data where we establish pre-emptive licenses for limited use, thereby making ESG and Climate data easier to consume and share. We foresee a “climate-data-as-a-service” capability which is by default open, but with clear licenses for access and use of closed data. This approach is modelled upon the practices established by IB1 (Icebreaker 1).

**OS-CLIMATE: CREATING A SELF-SERVE ESG AND CLIMATE PLATFORM**

OS-C is establishing a “self-serve” platform that will make ESG and Climate data easy to find, consume, and share.
Figure 5, Towards a Self-Serve ESG and Climate Data Platform

OS-C foresees an ecosystem of open source products offering “model to microservice” capabilities spanning a variety of key constituencies including business leaders, data scientists, developers, and operations staff. Popular open source tools are integrated into OS-C’s ESG and Climate Data Value Chain, including Spark, Kafka, Jupyter, Trino, and Prometheus, to name a few.

And OS-C’s goal is to make each element of the value chain structured as “self-serve” giving local autonomy and control to the global ESG and Climate data constituency.
To support these capabilities, OS-C will invest in an “ESG/Climate Data Product Factory” with repeatable templates and processes that will make it easy to build, secure, deploy, and manage data products.

It is important to note that OS-Climate is developing these data products as public goods, and these products are not designed to compete with data products of commercial data providers, rather to support commercial innovations by commercial providers such as those of OS-Climate members S&P Global and London Stock Exchange Group.

**CONCLUDING THOUGHTS**

Climate change is a crisis that impacts every person on this wonderful planet Earth. OS-Climate is taking several steps to make it easier for organizations, governments, and non-government organizations to play their part in helping address this crisis.

OS-Climate is a community of practice and action. They welcome contributions big or small, from data engineers, data scientists, software engineers, analysts, subject matter experts, product leads, and project managers. Join us to build, grow, and evolve OS-Climate to deliver better climate outcomes! [Contact OS-Climate](#) or click [here](#) to learn more.

***
ABOUT OS-CLIMATE

OS-Climate (OS-C) is a non-profit, non-competitive organization with an audacious mission of using open source tools to unify the ESG and climate data required to address climate change, one of the grand challenges of our time.

OS-C’s vision is to aggregate the best available data, modeling, computing, and data science worldwide – in accordance with the best scholarship across all relevant academic fields and science-based methodologies – into an AI-enhanced physical-economic model that functions like an operating system, enabling powerful applications for climate-integrated investing in a world where the future will be very different from the past.

OS-C’s goal is to rapidly accelerate the shift of global investment away from relatively GHG-intensive and climate-vulnerable companies, technologies, and infrastructure into mitigation, resilience, and adaptation that is financially sustainable and high-impact — especially in developing countries — as well as to enable design of better policy that effectively engages capital markets in addressing climate change.

OS-Climate – Open Source Tools for ESG & Climate Data

The OS-C platform will accelerate development of scenario-based predictive analytic tools to manage climate-related risk, and investment products to finance climate solutions across every geography, sector, and asset class. The OS-C Open Source Community of Practice and Action – using the highly evolved licensing, structured collaboration, and community management tools of the Linux Foundation – will serve the stakeholder community and enable alignment of effort.
on priority data and modeling needs, focus shared resources on executing those priorities, and accelerate adoption.

OS-C is sponsored by an international community of contributors spanning banks, asset managers, and technology vendors. OS-C is using Open Source technology to develop data and tools required to meet Paris accord climate goals. More information on OS-Climate and its mission, vision, and roadmap can be found [here](#).

Additional OS-Climate resources are listed below:
- [Unlocking Climate-Related Data Through Open Source and Data Mesh Architecture](#) (Vincent Caldeira / Erik Erlandson, ODSC West, 17 November 2021)
- [Powering climate finance action through an open source climate data platform](#) (Vincent Caldeira / Erik Erlandson, Red Hat Summit 2022)

***

All images in this document except where otherwise noted have been created by Eric Broda (the author of this article) or used with permission. Icons used in the images are stock PowerPoint icons and are free from copyrights.

The opinions expressed in this article are mine alone and do not necessarily reflect the views of my clients.