Operationalising National Data at Scale Why Mainstream Tech Won at OS

Michael Gordon
Digital Consultant, Former OS Data Hub Lead

The Fallacy of "Just Wrap It"

Many organisations are working with a legacy of data and services whilst trying to modernise.

The “easy” approach is to simply wrap legacy data structures in the latest OGC standards.

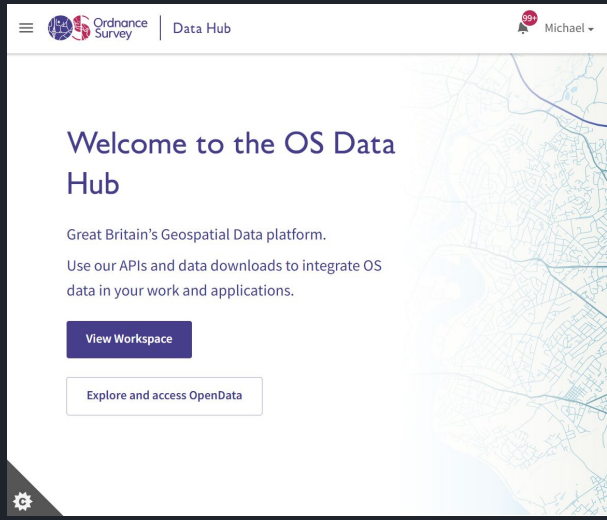
The Result: A standards-compliant mess

Interoperability is not a technical solution or specification.
It is a product decision.

Business Model → User Needs → Tech

A Product-Led Approach

- **Core Constraints:** Policy & Business Model
- **Direction:** Customer Outcomes and User Needs
- **Tech & Data:** Driven by the user needs



We built the storefront first, before we worried about the specific packaging of the goods.

The marketplace is an instantiation of your business model

The User Gap: Deep Discovery

- **GIS Background:**
  Loved precision, rich attribution, geo specific formats

- **Mainstream Tech or Data Background:**
  Valued currency, simplicity, ease of integration

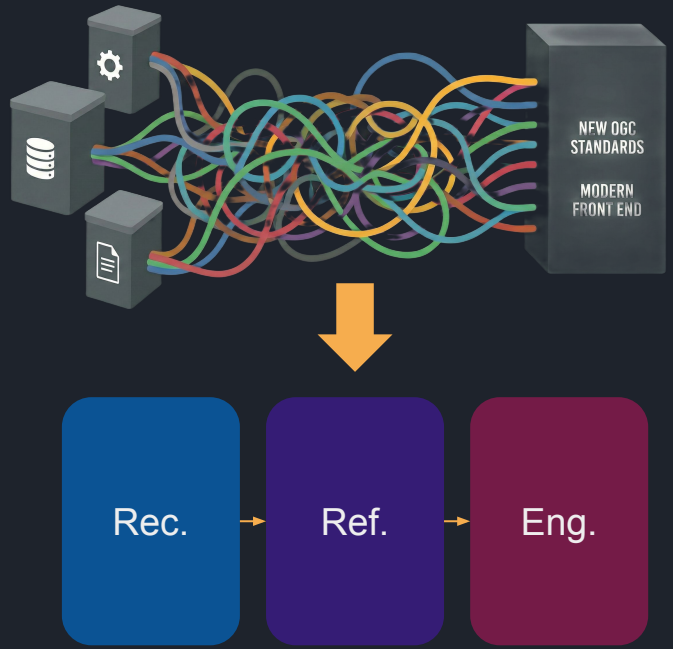
They didn't want data that was 6 weeks old but perfectly structured.

They wanted data that was 1 day old and easy to integrate

Decoupling the mess

- Create a clear separation of concerns

- The characteristics of the product required by user needs drove the architecture:
  - Near real time -> stream processing
  - Must be easy to change -> Multiple schema support
  - Customisable data orders -> Event based subscriptions for custom data recipes



- *The OS Data Hub focused on discoverability, security, authentication, accessibility, monetisation.....+ geo download & APIs*

Mainstream Tech + Geo Specialism

90%+ Mainstream Stack

- Front-end: React (SPA)
- Micro-services architecture
- Identity: Azure IAM
- API Management: Google Apigee
- Compute: Azure Apps and Functions
- Storage: Azure Blob, CosmoDB, Postgres

Geo: The thin layer of specialist value on top

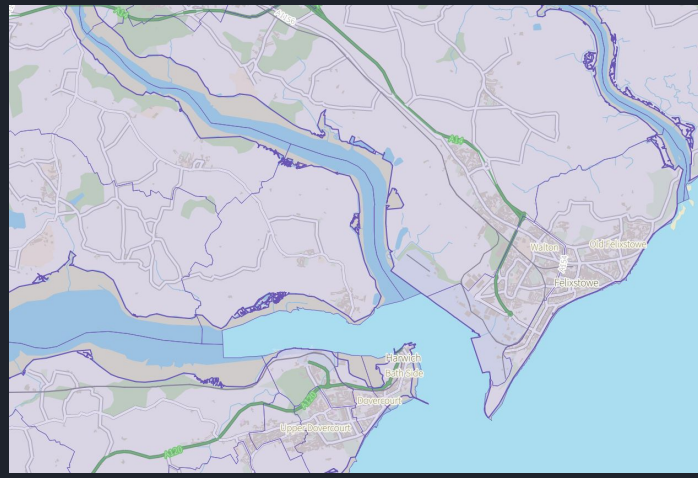
A concrete example - Boundaries and Borders

Context: We were intending to make boundary data available via OGC API - Features as we were doing with other data types.

Problem: Admin boundaries payload size > API Gateway limits

Solution: Understand users - switch to OGC API - Tiles

Work with the mainstream tech, understand users and let standards be constrained by these.

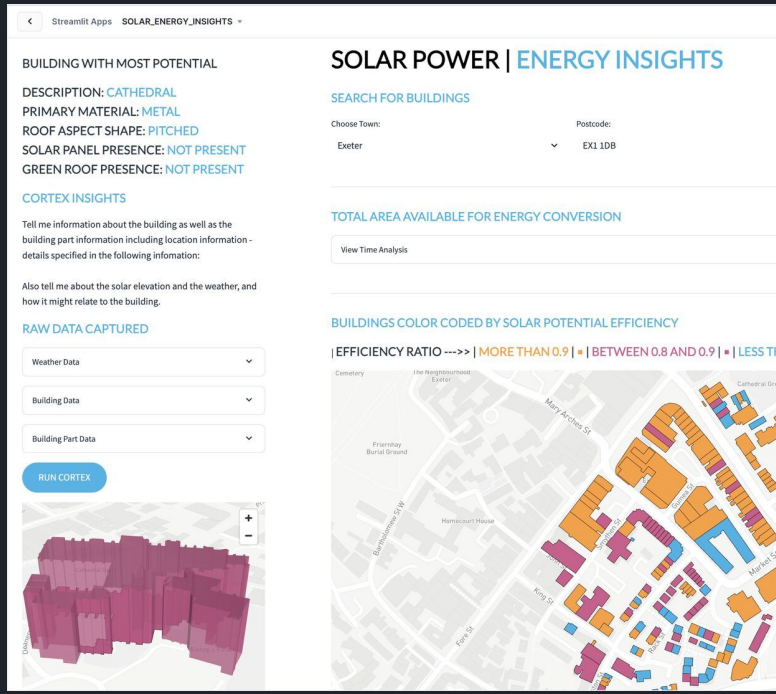


Hyperscaled Data Ecosystems

For AI training or national analytics, users don't want 10 million API calls.

They want data **already in their environment** - and this is increasingly performed directly in hyperscaler platforms.

Emerging Strategy: Provide data natively in hyperscaler platforms using Cloud Native Geospatial formats (such as Parquet).



User-Centric Interoperability

Web & Mobile

OGC APIs

Data Science

Cloud Native Data + OGC APIs

Offline / Field / GIS

OGC GeoPackage

Hyperscalers and AI

Native Integration via CNG

Summary

- Market & Business Models set the constraint.
- Customer outcomes and users "Job to be Done" sets the direction.
- Mainstream Cloud Tech > Geo Tech
- Be Opinionated: Offer the right option for the use case.

What are your thoughts?

We operationalised interoperability at scale not by being purists, but by being pragmatists.

We treated OGC standards as a toolkit to meet our users needs, not a religion.

By focusing on user needs, through the lens of policy and business constraints, we simplified choices in adopting standards and modernising our legacy

Connect with me and keep the conversation going!

Michael Gordon
Digital Consultant