Predator Free 2050

Delivering A Central Data Platform

Matthew Hellicar

Applied Innovation

Context

The current state of play within the Predator Free 2050 movement has led to data from projects being stored on a variety of dedicated, customised platforms. This makes it particularly difficult and time consuming to bring together a national view of status and progress towards the goal.

It is proposed that a central platform is created and made available to all projects and developers. This platform needs to be looked after by an independent, unbiased entity.

The architecture and development of that platform needs to be guided by the needs to the users – i.e. the developers, the operational projects, and those advising strategic delivery of the overall programme. A central, independent, unbiased entity should be charged with guiding the evolution of that platform.

Central Data Platform

State of Play: Challenges

- No single data platform has been accepted by all operational projects to rely on for all their field data
- The trap.nz platform has received investment and is in the process of being enhanced. At the time of writing, only a minority of projects and product developers are utilising the APIs provided by trap.nz to publish field data
- Operational projects have chosen to develop and maintain their own custom data management platform, believing their needs to be unique enough to justify the cost and effort
    - This has led to most projects requiring technical IT support and/or data analysis skills
- Operational projects believe that they need to tightly control access to all data to ensure privacy
- Operational projects believe that sharing their data will expose their stakeholders to risk and put their project's reputation at risk
- Operational projects believe they need to handle data sovereignty concerns locally
- Concerns exist across the collective of projects and developers that they are at risk of breaching data sovereignty laws
    - the reality is that, until recently, New Zealand has lacked a viable onshore dedicated cloud infrastructure so all cloud-based solutions in use today risk breaching some of those laws
- No common endpoint exists for product developers to send data to
- No agreed data standard exists for what data should be collected and reported from field devices
- Product developers have defined their own data structures with little or no effort to align these with other developers

- Most product developers have ended up developing their own end-to-end data integration solutions
- Some developers have also chosen to develop their own data management and visualisation platforms
- There is no aligned roadmap for the evolution of the suite of custom solutions developed by either product developers or projects

Opportunities

- New Zealand now has a hyperscale data centre so the option to ensure data does not leave our shores is a reality
- Data Platforms to manage very large sets of complex data exist in many other industries and are considered very mature. An enormous amount of investment has gone into such platforms to allow for the variety of data that needs to be stored and managed
- The data that operational projects need to manage is not unique in any way, as illustrated by the fact that all of it is defined in
- International Data Standards exist and are well defined, allowing for relatively easy compliance to those standards and opening up opportunities for data sharing
- Ethos Environmental have recently produced a thorough and wide-ranging report titled "A Proposal for Measuring and Assessing the Difference We Make: A Pilot Study, Mar 2025". *That* report contains a number of good suggestions relevant to data standards and data processing. All recommendations made therein remain valid. Proposals and recommendations in *this* report are designed to add to those recommendations, not to conflict with any of them.
- Trap.nz already has an API available and making that interface standards-compliant appears to be achievable without too much effort

Ideal Scenario

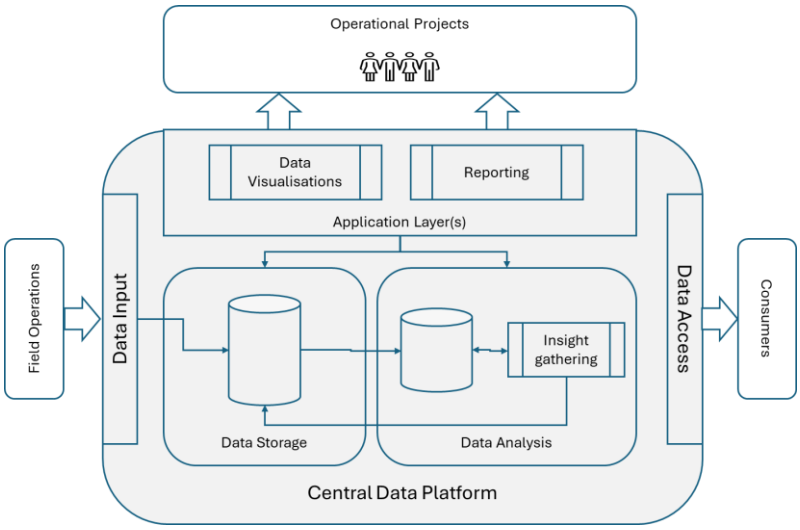
To realise many of the desired efficiencies identified, it is necessary to realise an end-to-end automated data management process. A Central Data Platform is required to enable that.

All options outlined here required a technical application stack (see later) deployed on a scalable and well-managed infrastructure. This in turn demands that a team (virtual or physical) takes on the responsibility of supporting the platform.

Features of a Central Data Platform

The ideal platform will support:

- Data Input: endpoints to allow automated publication of data from devices
- Data Storage: to securely store data aggregated from multiple sources
- Data Analysis: to allow analysis across aggregated data
- Data Visualisations: to provide automated information for consumers (projects)
- Insight Gathering: to derive actionable insights from data
- Reporting: to automate in progress and outcome reporting
- Data Access: endpoints to allow automated external data access (output)



A high-level view of the Central Data Platform

Principles of a Central Data Platform

Principle	Description	Notes
Free at point of use	Uptake of the platform is critical to its long-term success – a subscription-free service should be the target	If the platform needs funding, consider levels of subscription (based on usage & storage) with a free entry-level
Public Endpoints	Input endpoints (APIs) to be public (but with relevant security controls)	Ideally, publishing data to the platform should be free of charge to encourage uptake
De-coupled Data Storage	Master data source not tightly coupled to any application	Utilise pub/sub methodologies to ensure performant access
Standards-based APIs	All input data must comply with published standards	All APIs use data types compliant with GBIF (Darwin Core).
Unique IDs	Unique IDs must be maintained (and enforced by the platform) for devices and projects	Project ID required to map access rights to other platforms such as trap.nz.
User-focused Visualisations	All visualisations designed with end user needs in mind	Simple user-friendly insights with no technical jargon
Offline Analysis	All analysis processing carried out on offline copies of data	Insights/enhanced data can be fed back to master copy
Data Governance	Layered access control in the hands of the data owner	The Data Owner can decide who sees what and at what level of detail
Support Zero	Allow for zero-observation data to be published (e.g. device heartbeats)	Required for future proof-of-absence calculations

Principle	Description	Notes
Scalability & Performance	Elastic infrastructure, high availability, and optimized query performance	The chosen infrastructure is expected to handle most of this
All data available	All data to be available for national visualisations	With relevant obfuscation / anonymization where required
Encryption	In transit and at rest	Ensuring data security
Fully Audited	All access and usage of the platform to be audited	Audits searchable by admins

Implementation Options

The options outlined in this section are not exhaustive but have been chosen with the target of maximising ease of delivery.

Note: some elements of this platform can be considered to already be in progress and/or in place.

Important Guidance Notes:

1. Data Input – Trap.nz has an existing API to allow input and output of data from their source. This API layer is hosted on cloud architecture not guaranteed to be in New Zealand – may not address all Data Sovereignty concerns.
2. Data Storage – Ethos Environmental have been tasked with putting in place a data storage and analysis stack. This stack is considered fit for purpose (as designed) but has not been designed for real-time access from live devices.
3. Data Analysis – the stack delivered by Ethos Environmental (alongside the proposed enhancements) is considered fit for purpose to act as the core of the Data Analysis element of the platform

Layer	Considerations	Options / Next Steps
Data Input	• Public APIs • Security	• EarthRanger Gundi APIs • Enhanced Trap.NZ API
Data Storage	• Hosting • Scalability	NoSQL or other unstructured DB
Data Analysis	• Standalone offline stack	Potentially existing Ethos-delivered platform, co-located with Data Storage for efficient data access
Data Visualisations	• Flexibility • Customisation • Ease of use	• Support multiple solutions but ensure a standards-based visual architecture with extensible with custom views • EarthRanger's Ecoscope extension allows for user-driven customisations
Insight Gathering	• Flexibility • Customisation	Consider Ethos-delivered platform but allow growth and expansion for further analysis and insight tooling in the future

Layer	Considerations	Options / Next Steps
	• Future-proofing	
Reporting	• Flexibility • Customisation • Ease of use	Support multiple solutions but ensure a standards-based visual architecture with extensible with custom views
Data Access	• Public APIs • Security • Scalability for high volumes	• EarthRanger's Gundi APIs • Enhanced Trap.NZ API

Non-Functional Capabilities

In addition to the functional capabilities outlined above, a number of non-functional capabilities of the platform need to be considered.

- Hosting
- Security
- Usability
- Reliability
- Performance
- Availability
- Maintainability
- Scalability
- Supportability
- Auditability

Next Steps

Task	Description / Notes	Priority
Define Support Agreement	• Consider a centralised team, possibly from multiple entities but would need a clear command-and-control structure	1
Co-locate all elements	NZ Hyperscale Data Centre. Approach commercial suppliers to forge preferential agreement for deployment and support	2
Deliver Data Input API	Design, develop, and deploy	3
Deliver Master Database	• NoSQL / other unstructured data store • Integrate with input API	3
Deliver Data Visualisation(s)	Design at least one overall user-focused data visualisation	3
Execute Pilot	Integrate one or more devices with APIs, prove automated end-to-end data management, including	4

Task	Description / Notes	Priority
	first set of non-functionals (accessible, performant, useful)	
Deliver Reporting	Design at least one overall user-focused report Design at least one user-designed report	5

Capabilities Required:

Role	Scope
Enterprise Architect	Define systems, data flows, and non-functionals
Solution Architect (Technical)	Define technical solutions including integration points
Business Data Analyst	Define data requirements, visualisations, insights, reporting
Development	If required, build out of platform