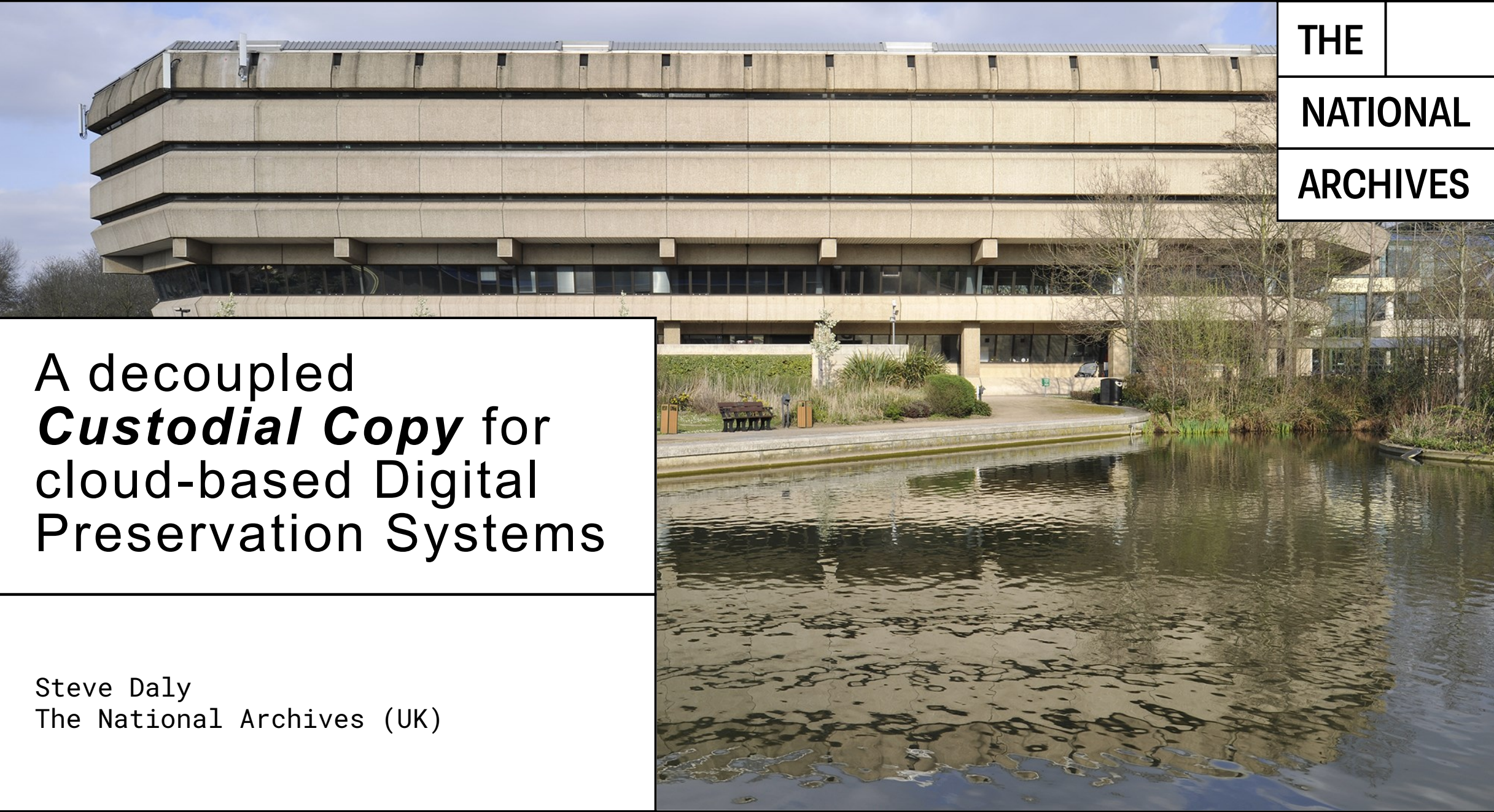


The background of the slide is a photograph of The National Archives building in London. It is a large, multi-story concrete structure with a distinctive stepped, tiered design. The building is situated next to a body of water, which reflects the building's form. In the foreground, there is a paved walkway, some greenery, and a small boat on the water. The sky is overcast.

THE	
NATIONAL	
ARCHIVES	

A decoupled
Custodial Copy for
cloud-based Digital
Preservation Systems

Steve Daly
The National Archives (UK)

The National Archives

The **archive** of the government of England and Wales, holding historic documents such as:

- **Domesday Book** (1086)
- **Magna Carta** (1225)
- **William Shakespeare's will** (1616)

...through to **petabytes of digital files** including digitised and complex modern born-digital records.

Also:

- Official Publisher of **Legislation** and **Gazette**
- UK Government **Web Archive**
- Lead and accredit the **archive sector** in England

...and much more



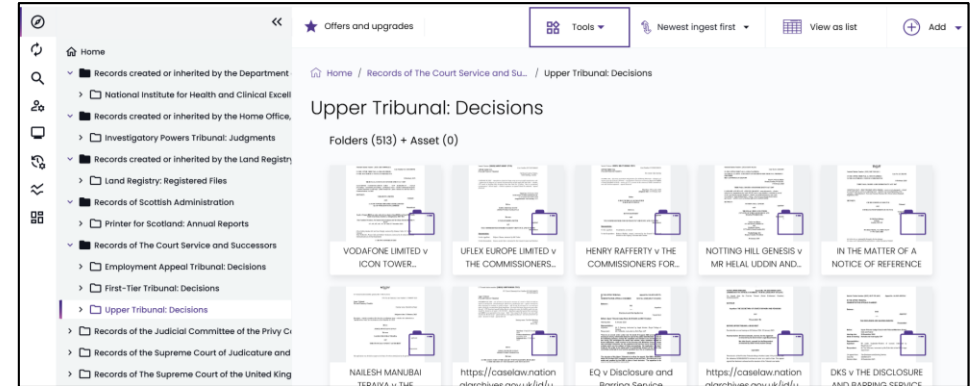
Digital Preservation

Ensuring the long-term life of our digital assets

- **Research, innovation and community-building**
- All work is published as **free open-source** software
- Create tools – **DROID** etc
- Maintain datasets – **PRONOM** etc

Digital Repository Systems

Digital Preservation system (i.e. **DAM**)



Underlying storage is currently **Data Tape Library**



Current - On-site system

Pros

- Fully under own **control**
- Can **customise** as required
- Traditional **security** approaches for on-prem systems



Cons

- Needs **specialist skills** and effort to maintain, support, upgrade, and keep secure
- Can easily get **out of date**
- Sub-optimal **environmental** footprint
- **Cloud connectivity** can be complex, while maintaining security
- **Capital investment** required for growth

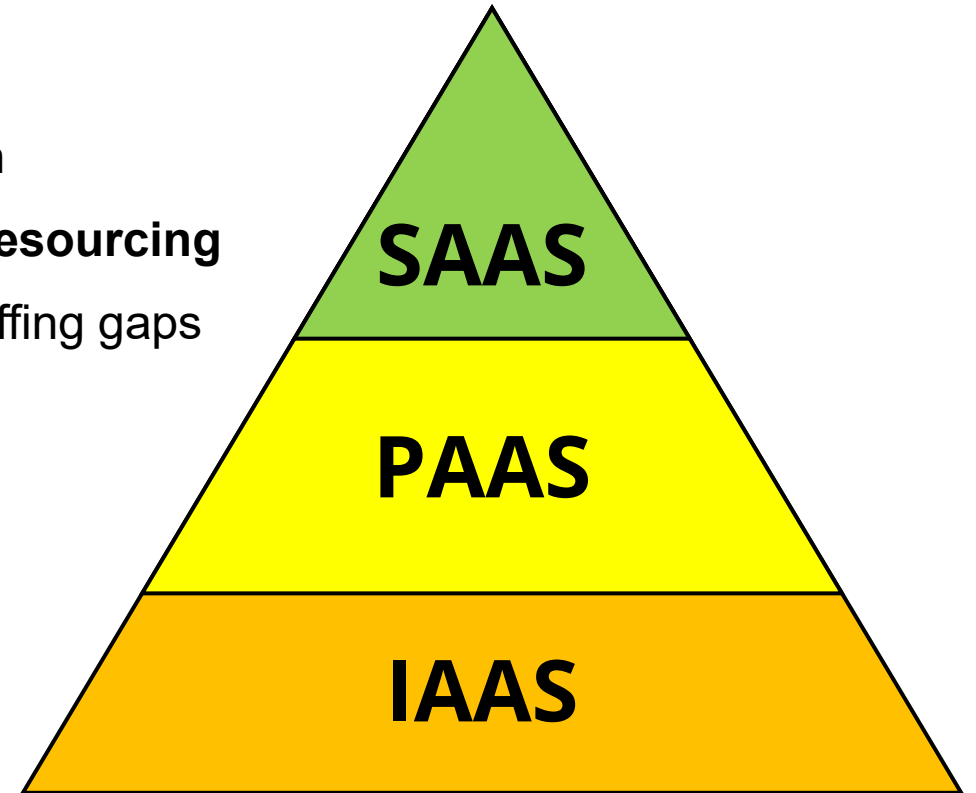
Future - Cloud Managed Service

Pros

- Matches UK Government **Cloud First** “high stack” approach
- No need to maintain infrastructure and software – **simpler resourcing**
- Archive still **functional, safe and secure** during internal staffing gaps
- Easier **scalability** and integration with **cloud services**

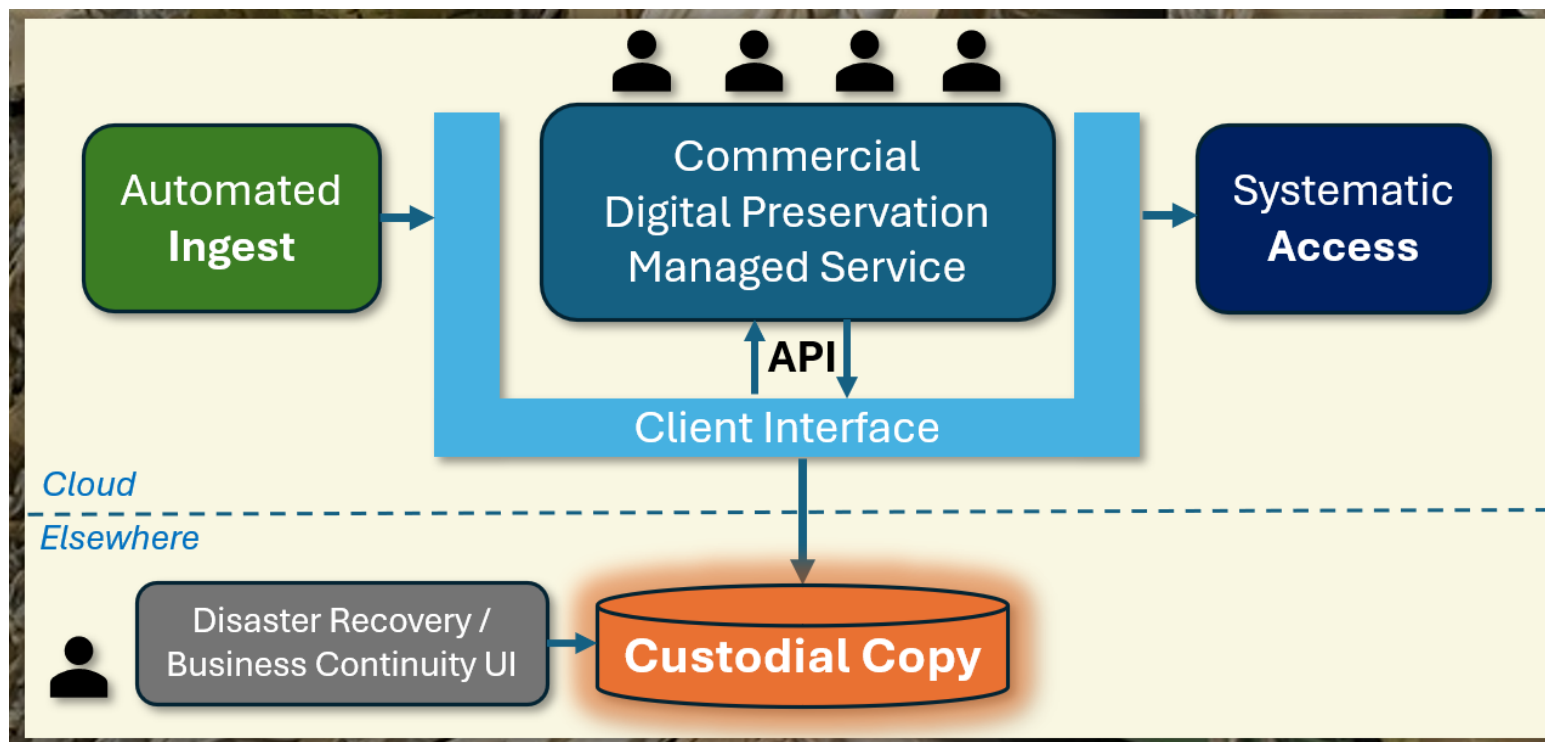
Cons

- Need to manage **supplier risks**
 - Vendor **failure**
 - **Exit planning**, vendor **lock-in**
 - Potential reduction in **Control**
 - Needs different **DR / BusCon** approach



Best of Both Worlds - Custodial Copy

Active management in the cloud



Long-term copy elsewhere

Custodial Copy - Data Structure

Not (just) an IT Backup

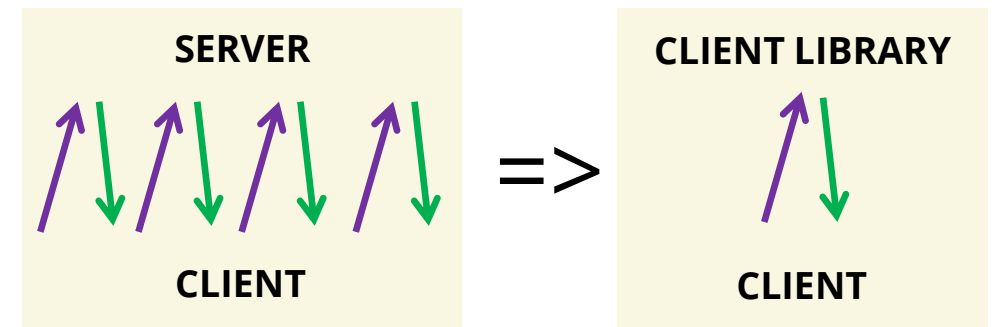
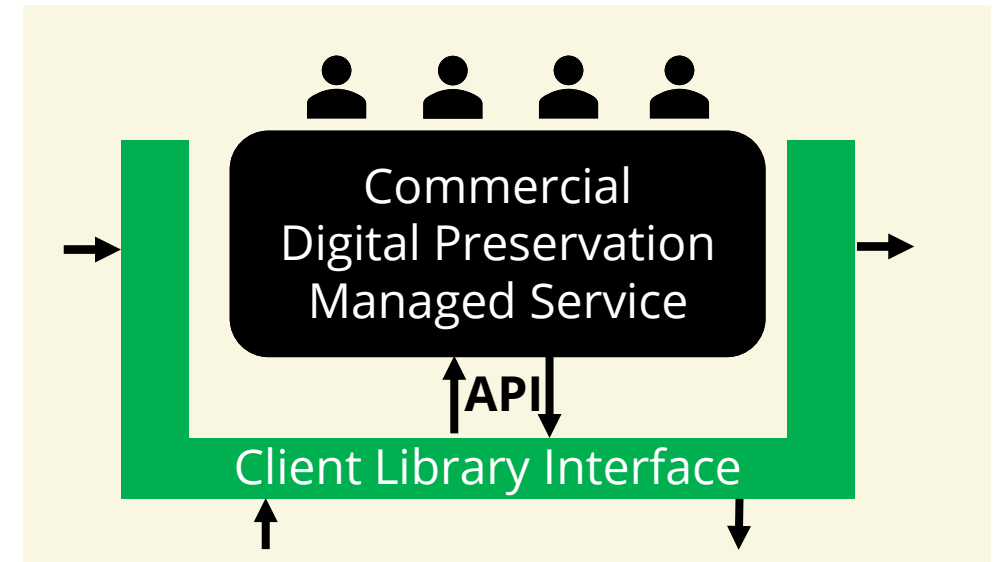
- Structured for **long-term readability**
- **Vendor Agnostic**
- Written using **Oxford Common File Layout** (OCFL)
- Includes **Abstraction, Versioning, Delete Protection**

Custodial Copy - Self Describing

- All relevant **metadata harvested** and stored alongside
- Store **software** and documentation in OCFL repo too
- **Self-contained** – can be read without other systems or data
- **Everything you need** to reconstruct, access, or migrate

Custodial Copy - Client Library

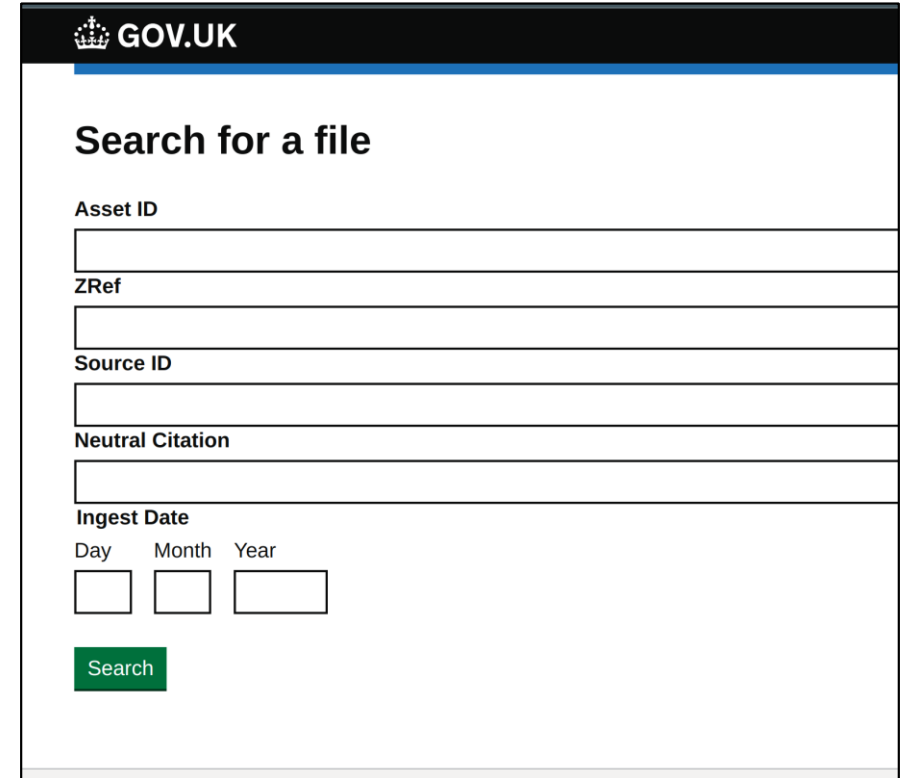
- **Disallow** direct connection
- All **conversations** via client library
- **Not another API Layer** – just a library (like MySQL vs PostgreSQL)
- Simplifies **vendor swap**
- Abstracts **complex conversations** to a higher level



Custodial Copy - Proving

In Service Today

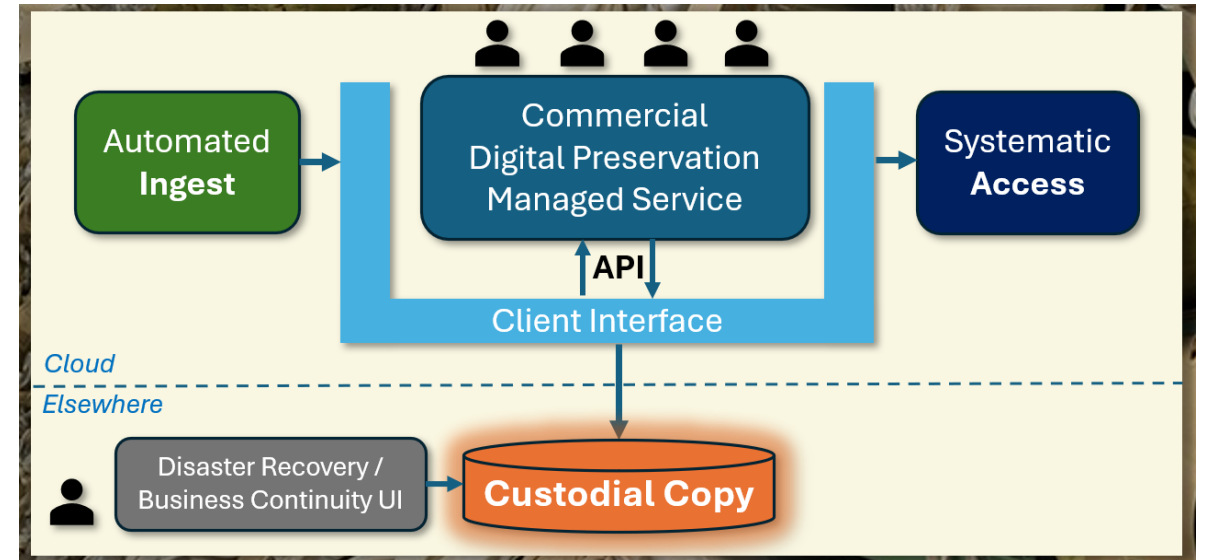
- Custodial Copy currently being written to **on-prem** copy on data tape
- Built to support **multi-cloud** as destination
- Simple Rapid-Developed **GUI** to read OCFL repo
 - Proves **reversability**
 - Provides **DR/BusCon** including for **deleted** files



The screenshot shows a web interface for searching files. At the top is a dark blue header with the GOV.UK logo. Below the header is a white box titled "Search for a file". Inside this box are several input fields: "Asset ID", "ZRef", "Source ID", and "Neutral Citation". Below these is a section for "Ingest Date" with three separate input boxes for "Day", "Month", and "Year". At the bottom of the search box is a green "Search" button.

Custodial Copy - Future Developments

- Intelligent **caching**
- Prove / improve **roundtrip**
- Integrate further **metadata** sources
- **Optimisation** for extremely large file counts
- Demonstrate ability to **swap-out** core product



THE

NATIONAL

ARCHIVES

Thank you



Custodial Copy

Poster: <https://doi.org/10.5281/zenodo.13647420>

Code: <https://github.com/nationalarchives/dr2-custodial-copy>

Steve Daly
The National Archives
Kew, Richmond
TW9 4DU
United Kingdom