

# Service Definition Data Integration & Publishing Platform

**Digital Marketplace:** GCloud13  
**Version:** 1.0 Epimorphics Ltd  
**Date:** 4<sup>th</sup> May 2022  
**Contact:** [gcloud@epimorphics.com](mailto:gcloud@epimorphics.com)

## Data Integration and Publishing Platform - Service Overview

### Summary

Our data integration and publishing platform enables organisations to transform and integrate data from a wide variety of systems and then make that integrated data available for consumption through easy-to-use APIs and bulk downloads. Customised data transformations can be used to ingest data from a wide variety of sources including near-real-time data streams.



### Features

- flexible, high performance data storage
- fully standards-compliant linked data publication platform
- widely used within the UK public-sector
- replicated for fault-tolerance and scalability
- examples include publishing near-real time data
- ingest data from a wide variety of sources
- data quality - validation and data quality integrated with data flows
- data integration - high-performance with large data volumes
- hosted and managed enterprise Data Platform Service
- enable advanced front-end app interfaces beyond the standard UI

### Benefits

- robust, reliable publication of sustainable, trusted and usable 5-star data
- adaptive/flexible platform can be grown to meet changing needs
- deployment flexibility on the cloud or within your own infrastructure
- provide customers with integrated, live updates twenty-four hours a day
- build data integration into your transforming organisation
- bring data from your legacy systems into one place
- publish high-quality data that is actually used
- enable teams to start integrating, managing, curating and using data
- data sharing via flexible APIs and bulk download
- create and deliver services with user needs at their heart

<I wish more of our suppliers were like you guys/>

Digital Lead, at a public sector client

## Introduction to Epimorphics

Epimorphics Ltd is a 100% employee owned company that was formed in 2009 near Bristol in south west England. We are at the forefront of developing standards, tools and applications, helping organisations connect and make more effective use of their data assets through open, standards-based linked data tools and techniques. To find out more about us, our work and our approach see: [www.epimorphics.com/introduction-to-epimorphics](http://www.epimorphics.com/introduction-to-epimorphics).

## Service Description

The **Epimorphics Data Integration and Publishing Platform** is widely used for integrating and linking data for open and non-open data. For example it is used to support the Food Standards Agency ([data.food.gov.uk](http://data.food.gov.uk)) and others.

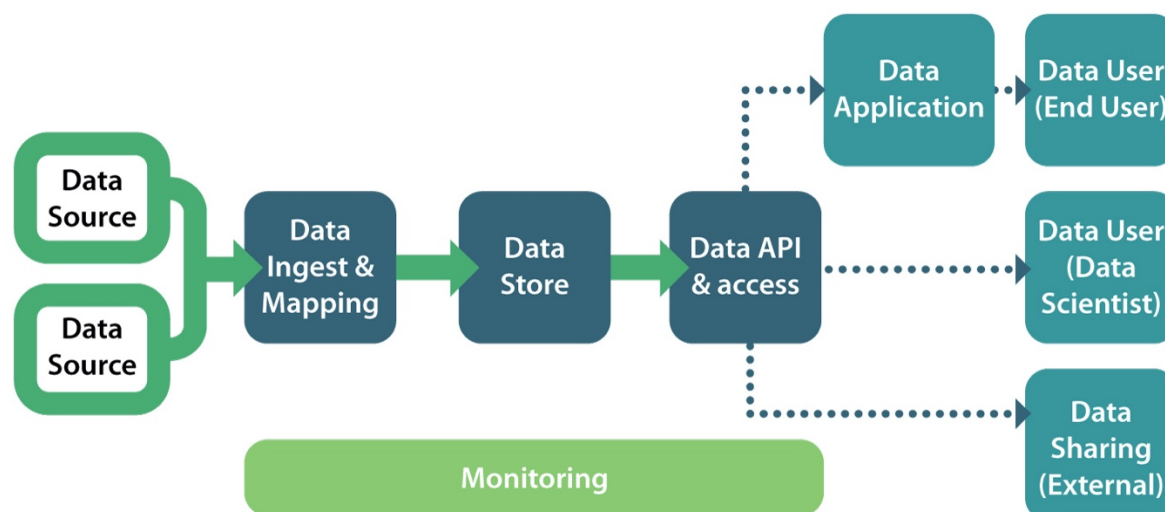
We offer the platform as a fully hosted and managed service for collecting, processing, integrating, merging, publishing and using data. By default we provide a hosted service on top of Amazon Web Services.

The platform includes:

- A Linked Data API engine, providing access to the data in several developer-friendly formats (including JSON and CSV) and human-readable web pages
- Customisable text search
- A triple store for storing data as RDF
- A fully SPARQL 1.1-compliant endpoint (made accessible as an optional extra)
- A scale-out, fault tolerant runtime platform
- A data management system, to enable clients to load their own data in source or RDF formats

Optionally we can provide additional upload mechanisms and automation which will integrate with clients' existing workflows to support "business as usual" publication of linked data, including support for near real time data streams.

The platform is customisable and can also host applications running on top of the data.



Our **Data Integration and Publishing** support service provides setup and support services for the data integration and publishing platform, including data modelling and preparation, platform configuration and custom data presentations, as well as migration, testing and ongoing support.

We provide setup and support services for the planning, setup, migration, testing and ongoing support of our **Platform** offerings.

We also offer training courses for people wishing to develop their own skills in linked data publishing (see our **Cloud Support Training service**).

We offer two variants of the platform for procurement via GCloud.

1. The Flexible platform: this includes all the facilities of the platform together with whatever optional extras the customer desires. A combination of the fully flexible platform and our data integration and publishing service will enable the customer to publish and integrate data from a variety of original data formats, with near real-time updates if necessary, as well as have a range of custom data presentations.
2. The Complete Fixed platform: this includes all the core features described in this document, including fault-tolerance and scalability. When combined with our data integration and publishing service this will support a range of custom data presentations.

Our hosting includes support during UK business hours.

*Please note that we are continuing to invest significant development effort into our platform to enable it to handle significantly larger volumes of data, deal with more complex queries more quickly, and to enable new functionality. The overall architecture and component structure described in this section is therefore under constant revision.*

## Optional Additions

### Existing workflow integration

Optionally we can provide additional upload mechanisms which will integrate with a client's existing workflows to support "business as usual" internal or external publication of data. For example we have previously built integrated data processing for the Food Standards Agency, integrating and linking data from multiple legacy business systems into a consolidated data hub. This integration builds upon the automation capabilities provided by our WebOps and Data Management Services.

### URI Reverse Proxy Service

Some clients wish to use their data domain for multiple data services. Typically each data service will be hosted on separate servers and may be run by a variety of organizations or providers.

To support this we can also provide a **reverse proxy service**. All requests to the domain would be routed through the proxy service which would *forward* each request to an appropriate backend service, passing the response back. In this way the data is visible to users through a single domain and they are insulated from the details of, or changes in, the backend service providers.

We have several other optional extras listed in our **pricing document**.

## Service details

As a hosted service, our platform is currently accredited to store and process ILO information only. We are in the process of seeking enhanced accreditation for the future.

All data loaded onto the platform is backed up at the time the data is loaded, so the backup is always an accurate reflection of the data in the system. The replicated nature of the platform means that a hardware failure will not cause data from the running system to be lost. In the event of catastrophic infrastructure failure which takes out all the replicated instances the data will be restored from backup as quickly as possible.

As the system is fully replicated routine maintenance can be carried out without taking the system off-line; there is no need for scheduled maintenance windows when the system is out of service.

We aim for the availability of the system to be 100%. Details of our support services are given in the next section.

## Support

Our hosting support for the full system includes all regular maintenance, monitoring and backups. We will provide reports to the clients on the usage of the system – the precise details of the data reported will be agreed with the client during the setup phase.

We also provide an incident reporting service. The basic service is available during normal business hours (09.00 – 17.30 Mondays – Fridays, excluding public holidays). We provide an email address for incident reporting and will respond to any notification within 4 hours. If an incident results in loss of service we will restore the service within 1 business day; in all other cases we will use reasonable efforts to resolve the incident as quickly as possible.

The replicated nature of our architecture is such that we need not take the system down to perform regular maintenance and system updates.

## Compliance with Open Standards

As a linked-data company we have a passion for open standards. We've been key in designing and developing many of the web standards and vocabularies around linked data. For more information on our open standards work see our website at: [www.epimorphics.com](http://www.epimorphics.com). Specifically, Linked data depends crucially on the correct implementation of the relevant open standards. Our platform fully complies with all the relevant standards, notably:

- RDF syntaxes: RDF/XML, Turtle, N-Triples, JSON-LD
- RDF 1.1 Turtle
- SPARQL 1.1 Query
- SPARQL 1.1 result set formats (XML, JSON, CSV, TSV)
- SPARQL 1.1 Update

## Onboarding and Offboarding

We offer support in getting started with our dedicated setup and support service under Cloud Support. We provide documentation and guides to some of our tools through our website ([www.epimorphics.com](http://www.epimorphics.com)) and our code documentation on Github ([github.com/epimorphics](https://github.com/epimorphics)). In building skills, awareness and supporting the design of data we offer a number of training options.

If no customisation of the web interfaces etc. is required, then we will provide the client access to the data management system so they can have data loaded onto and published by the platform within 5 business days after contracts have been signed. We can provide expedited on-boarding at extra cost if desired.

No user data is collected by the system – the only data stored on the publishing platform is data supplied by the client. On termination of the contract all client data will be securely deleted. During the life of the contract clients can request access to a copy of all the data stored on the system. We have also supported public sector clients in transitioning the technology to their own infrastructure.

## Sustainability

- We regularly assess our full carbon footprint, to quantify, as far as possible, the greenhouse gas emissions associated with the operations of Epimorphics, the services that we run for others and the services that we use.
- We have assessed and been audited for our Scope 1, 2 and 3 emissions, reducing where possible and offsetting where not.
- Our company values promote support for the environment in all aspects of our work. We evaluate contracts we accept to see if a successful outcome to the contract will have a beneficial effect on our environment.

We operate as a Carbon Neutral plus organisation, and publish our footprint details as part of our sustainability statement here: [www.epimorphics.com/sustainability-summary-statement](http://www.epimorphics.com/sustainability-summary-statement)

## Our other Services

### Cloud Software

We have complementary **GCloud Software service offerings:**  
[digitalmarketplace.service.gov.uk/g-cloud/search?q=epimorphics&lot=cloud-software](http://digitalmarketplace.service.gov.uk/g-cloud/search?q=epimorphics&lot=cloud-software)

- **Epimorphics Data Integration & Publishing Platform**
- **Epimorphics Linked Data Publishing Platform**
- **Epimorphics Reference Data Management Platform (Registry)**

### Linked Data Publishing

We have complementary **GCloud Support service offerings:**  
[digitalmarketplace.service.gov.uk/g-cloud/search?q=epimorphics&lot=cloud-support](http://digitalmarketplace.service.gov.uk/g-cloud/search?q=epimorphics&lot=cloud-support)

- **Linked Data Training**
- **Epimorphics Data Architecture & Strategy**
- **Epimorphics Reference Data (Register) Publishing**
- **Epimorphics Linked Data Publishing**
- **Epimorphics Data Integration & Publishing**

## Locations and Contacts

Epimorphics Ltd is headquartered in Portishead, near [Bristol in south west England](#).

### Epimorphics Contacts

G-Cloud [gcloud@epimorphics.com](mailto:gcloud@epimorphics.com)  
or talk to any of the team: <https://www.epimorphics.com/about/> on 01275 399069