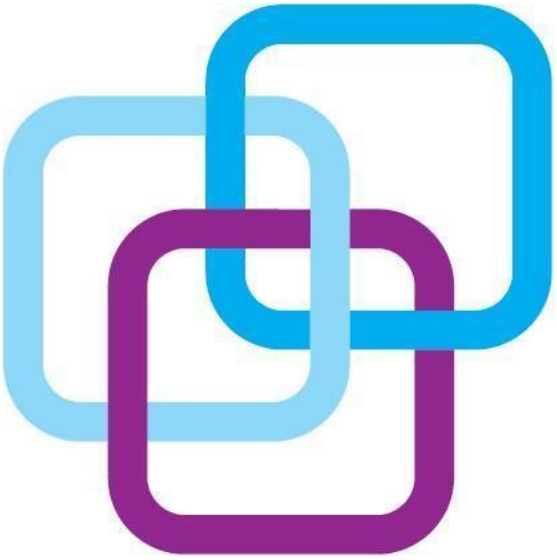
G-Cloud 15 Service Description

CIIM
SAAS

Knowledge Integration Ltd



About us

Knowledge Integration was incorporated in 1999, giving us many years of unrivalled experience in the provision of software services. We have a team of highly skilled, qualified developers and consultants to ensure our services meet and exceed customer requirements. As of 2026 we work in excess of 50 customers across the UK and worldwide with our CIIM solution being used for the publication and management of over 150 million items.

We traditionally focus on the problems of information extraction, description, indexing and searching with skills in ETL, GIS workflows for converting legacy metadata repositories into modern information retrieval and discovery solutions. We use our experience and skills to support a wide range of customers and solution requirements.

What the service is

The CIIM framework is a suite of open-source (AGPL v3) middleware and API components and was originally created in 2009 for the London Museum (formerly Museum of London). Since that time it has been continuously evolving and developing and now provides a reference middleware implementation for the cultural sector, and is in use at a wide range of institutions worldwide.

This highly scalable, fully automated ingest framework comfortably handles over 80 million rich records and seamlessly integrates with 30+ collection, archive, library and digital asset management systems.

CIIM features a schemaless data model and is ingest format agnostic, with a flexible processing workflow and field level data enrichment.

AI integrations, deep zoom/IIIF media and manifest handling, automated publication, media artifact creation/watermarking and optimized search and retrieval across large datasets are only some of the areas in which the CIIM excels. A comprehensive data auditing strategy is also built in, ensuring data integrity across all aspects of ingest, processing and data/media publication.

Request time, codeless, data transformations deliver on-demand Linked Art, LIDO, CIDOC, EDM, Skos, Marc, IIIF and many more. Require a different format, no problem, the required configuration is in your hands.

Who uses CIIM

Although originally developed for the cultural heritage sector - e.g Galleries, Libraries, Archives, Museums and Cultural Aggregators - it is also applicable across a wide range of other sectors wherever making sense of large amounts of disparate data is required.

Our existing customers include:

- Museums - British Museum, National Maritime Museum, ACNMW (National Museum of Wales).
- Archives - The National Archives, Archives Hub and Lloyds Register Foundation.
- Libraries - The British Library, Morgan Library (New York)
- Galleries - The National Gallery, National Gallery of Ireland, Manchester Art Gallery, Whitworth Gallery.
- Aggregators - The Museum Data Service, The Estonian National Cultural aggregator, Archives Hub.

Why choose CIIM

An apparently ‘simple’ problem to solve, CIIM answers the data integration, audit and sharing questions that most other solutions/frameworks haven’t even thought of asking yet.

First and foremost CIIM is not a bespoke solution, but a versatile framework using plugins and extension points to make customisation easy.

It is built on over 16 years of data analysis, use-cases and differing institutional requirements and data models. As an open source product, it continuously evolves based both on user requirements and emerging technologies, with all new features accessible to all users.

CIIM Features

CIIM is a middleware application that connects people to content. It can be used to:

- Deliver a modern, scalable and extensible enterprise architecture information architecture

- Fully automate, API/file format agnostic integrations, cross-system auditing/reporting.
- Simplify your IIIF and media delivery requirements with its automated media processing pipeline
- Model any type of data using its schemaless data model with internal graph representation
- Deliver flexible processing workflows with field-level enrichment, validation and privacy
- Provide a coherent enterprise/semantic search strategy (including Elasticsearch and SOLR)
- Integrate a codeless API building framework (including IIIF, LinkedArt, LIDO, EDM, Marc)
- Merge analytics with your data, to help drive 'insight led' planning
- Deliver a platform for repeatable, flexible and extensible AI integrations/data enrichment incorporating RAG, LLMs, Named Entity recognition, image analysis and audio to text.
- View cross system cross-system data visualisations and data audits via the flexible, accessible user-interface
- Share, search and syndicate your content via the configurable APIs. Providing graduated access for records, fields and media (including delivering different image renditions for different audiences)

Service Interface

The service interface provides a role-based view across all of your data and syndication endpoints. It enables viewing of performance metrics, system activity, auditing and facilitates the management, configuration and scheduling of data integrations. The interface can integrate with an organisation's centralised identity management systems and provide access levels based on memberships set within that system. It has configurable search access points and data partitioning which can be set per-user or cross organisation and user-customisable filtering which can be set at the same granularity. Simple searching and targeted searching using Lucene syntax is also supported.

API

All of the features of the system are accessible via an API which is role-based. The service has additional APIs for data delivery, searching, extraction, submission and configuration. All of these APIs can be used to both setup/configure the service and to customise delivery frameworks and data processing pipelines. The mechanisms for configuring the service are split between APIs which are used to setup and configure integrations, schedules, timings and allocated system processing power, and configuration which is used to drive pipeline transformations, publication endpoint configuration, syndications and dissemination formats. As this type of configuration needs to be tracked to avoid introducing errors, it is managed via a tailored CI/CD pipeline which both monitors your configuration and tests/deploys changes to using the Argo API into Kubernetes. The common API tasks are also surfaced via the user interface and this can also be used to schedule, suspend, run on demand and search/filter across all content. This also extends to UI customisations for filters, search access points and data display templates.

Customisation

The service offering can be customised via the user interface and also via our flexible data manipulation language (Proteus) - this allows buyers to configure their own data delivery formats and standards, data transformations, validation rules and processing pipelines. Via the user-interface, end users can set up their own data display templates, search access points and data partitioning. These customisations can be for a specific user only or application wide. Analytics service integrations can also be customised.

More complex customisation can be configured at code level via plugins that use prescriptive interface definitions to be loaded at runtime within the application.

Accessibility

The CIIM administrative interface is WCAG 2.2A compliant.

Levels of backup and restore

The service is backed up/restored using native cloud resources. Levels of backup and restore can be tailored per customer, as this is dependent on the data, response, and backup levels required. This is analysed with the customer and agreed in the customer's specific SLA. The default is:

- Daily retained for 1 week
- Weekly retained for 5 weeks
- Monthly retained for 1 year

Support

Support is available UK working days (Monday- Friday excluding weekends and bank holidays) and working hours (09:00 - 17:00). Our service level agreements (SLAs) include criteria for prioritising issues according to severity and agreed response times for services outside the normal scope of support. This can include minor customisations, consultancy, and training. Support requests can be raised as tickets via Zendesk outside of these hours and will be actioned when working hours resume.

Calls logged on the helpdesk are triaged and initial reply sent within two working hours. Further responses are dependent on nature and tier of the now triaged call, with Critical Tier 1 tickets taking priority

Frontline Support: We have a dedicated front-line support function where staff receive support requests and will respond within agreed service levels. Support requests can be raised via email and directly from the Zendesk portal. Front line support staff have expertise across the full range of our products and are often able to resolve an issue themselves. If this is not the case, the issue is referred to the appropriate development team for a response.

Zendesk and Support Portal: We use Zendesk as our issue management system with users able to log support requests by email or via a simple web portal. The web portal allows customers to track the progress of support requests. The support portal also includes a help centre which contains product-specific user guides and more detailed technical documentation.

User Group: CIIM has a user group. The user group is managed by the user community and involves twice yearly meetings where Knowledge Integration provides information on current and planned developments and users are able to raise any queries and concerns. The user group also provides a useful opportunity for the user community to influence the future development of the product.

Performance

The system has internal performance monitoring and will scale up and down based on load. The system has a centralised 'status monitoring system which can automatically send messages to the support team to pre-empt potential problems. Our system is vulnerability scanned daily to detect issues and inform continuous improvement actions. It is also pen tested annually.

When the systems are configured (as part of the onboarding), metrics are established in the SLA for business-as-usual operating performance, and our system is actively monitored to allow us to analyse where problems may arise from issues both internal and external. Limits are placed on the amount of resource each tenant can utilise guaranteeing that even under heavy load there is no cross instance performance degradation. The overall architecture always has headroom to allow it

to be temporarily 'burstable' to deal with transient heavy load and there are multiple API caches (including cloudflare optionally) to also balance system throughput. If an ongoing resource increase is required for a specific tenant this can be configured and redeployed automatically.

Availability

We guarantee 99.9% uptime for any public facing services. This excludes agreed scheduled downtime and any outages which are the result of issues caused by the underlying cloud provider platform being unavailable.
Uptime is monitored and reviewed as part of service renewal, any failure to hit the agreed availability will be considered and offset against any annual increase in renewal costs.

Resilience/reporting

Our services can be delivered via all the common cloud platforms and are deployed on Kubernetes, typically using Amazon EKS. We specify the numbers of 'replicas' of service components via our automatic configuration management and this replica provides the underlying resilience and auto scaling for handling burstable demand increases, as well as replacing unhealthy services with new healthy replicas. Where cloud platforms are used, their service guarantees cover the uptime of the Kubernetes Platform.

They are deployed within a Kubernetes cluster, with cluster management via Argo and service management via our own internal monitoring system (Harvey).In addition, our automated monitoring system automatically raises a support ticket when any system anomalies are detected. These anomalies are not limited to actual problems and can be used to pre-empt issues arising. This can include performance slowdowns, processing bottlenecks and API response times. Because these are raised as support tickets they are also visible to the customer when they log in to their support portal.

External monitoring tools are also configured to probe the ‘aliveness’ of deployed applications.

Operational Security

Our core components, implementation code and configuration are managed/ tested within our CI/CD pipeline for all commits, with weekly large scale integration tests as part of this framework, to allow us to capture any potential issues which only surface for large data sets. The potential security impact of any change is always reviewed in terms of changing the threat landscape, attack surface and whether our existing controls mitigate the change. If there is deemed an impact on security, then these

threats are documented and mitigation is put in place. This is always validated by vulnerability and/or penetration testing as required.

We utilise OWASP dependency Check within our CI/CD build pipeline. Scanning of new and existing components happens at verify time as part of the build process. We utilise a central database, updating from the NVD database daily and all our builds use this database. We utilise Intruder.io to check external vulnerabilities daily and to email a report and these are also responded to within a 48 hour window. Both of these services monitor daily CVE updates to provide their analysis. For third party components, we utilise OpenVAS checking against an internal version registry to produce alerts.

Maintenance

Deployment and service upgrades are carried out automatically via CI/CD during a scheduled maintenance window (although there is minimal impact on system availability). They can also be deployed to a mirror system prior to a gateway switch over if required

After Sales Support / Account management

Each installation has a dedicated account manager who will provide both sales and after-sales support. Support will be provided in the form of review meetings (can be face to face) and through the user group, twice yearly as a minimum.

Technical Requirements

The minimum requirements for using CIIM is to have stable internet access, and to use the latest version of the most popular browsers. If any of the systems being integrated require authentication, CIIM will require an API key/token if automatic system integration is required.

Onboarding

We initially liaise with the users to decide which integrations they require, how much data/media they have and what the likely external consumption of the APIs and data will be, in order to allocate an initial CIIM resource profile. These integrations are then configured and auto-deployed in a 'vanilla state' (without customisations). Data is then populated and templates are set up to view the data. At that point any required customisations are defined and implemented (which) depending on complexity may use configuration or require the deployment of bespoke code plugins/configuration. Initial training is provided for all classes of user and specific targeted training (e.g. for internal developers. admin users) can be also included within the package. Initial training is always online but on-site training can be provided at extra cost. Detailed online end-user documentation is provided which

includes tutorials, glossary and 'click through' videos. There is also additional online developer documentation for the transformation and dissemination frameworks.

The service is accessed via its own UI for management, configuration, searching and visualisation and scheduling. The user-interface can be configured by the users themselves. Minimal onboarding is required for the interface.

Training is provided by the user interface team. The interface team has 15+years experience with CIIM. Onboarding and training are completed remotely. We offer an initial two-hour session for this and can provide further training sessions where needed (which can be on-site at additional cost).

Documentation

CIIM has detailed documentation available online covering system requirements, how it works, glossary and a 'how to' section containing click-through videos.

Please see (https://docs.ciim.k-int.com/)

All APIs are documented using OpenAPI/Swagger.

Outage and maintenance management

In the case of an unexpected outage, the system will automatically restart and any in-progress processing will resume from its last position.

Hosting options and locations

Any cloud platform is supported as the system is cloud agnostic. For UK customers all data is hosted in the UK

Offboarding and Access to data (upon exit)

At the end of the contract the service is turned off and ceases to be publicly available. Data export functionality is available at any time during the contract so users can prepare for the end of contract in a controlled manner. For a period of 1 month following shutdown, the service can be reinstated for data export (for example if any problems were discovered with the end of service data exports). There is no charge for data exports even during the month post contract end.

For offboarding, the data in the system can be exported into many different formats (including csv, json, xml) and provided free of charge. Data can also be exported in bespoke formats if required (but this may incur additional cost). All data will be available at the offboarding stage. Exporting and sharing data is an integral part of the platform.

Service usage metrics

Integration and data throughput and API usage metrics. We integrate with multiple analytics services so that any downstream services which use the platform can have their analytics integrated into the data for insight-led planning.

Helpdesk response times, service availability and resource usage are also provided.

These metrics can be made available via

- API
- Real-time dashboards
- Regular reports
- Analytics service integrations

Security

We hold Cyber Essentials Plus and we have a Cyber Security Compliance Group which meets monthly to review and discuss security issues.

Standing agenda points:

- Recording of policy breaches, remediation, staff training
- Continuous policy review
- Upcoming maintenance schedules and action planning
- Infrastructure and internal service upgrades, updates
- Review of standing actions from the previous meeting

All cross system data communication is secure.