

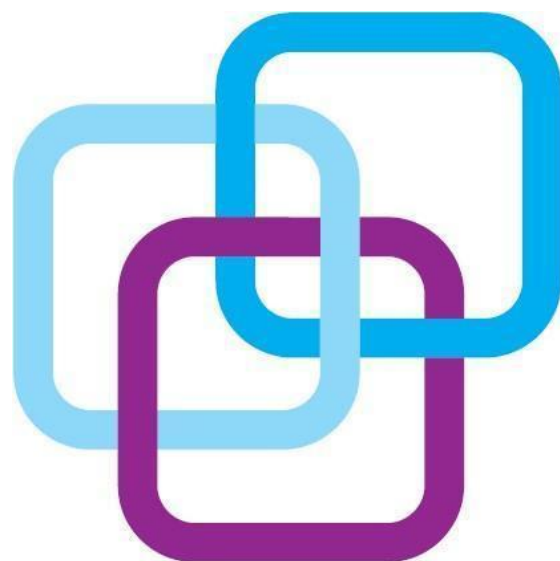
G-Cloud 15

Arches for

Science

Service description

Knowledge Integration Ltd



About Us

Knowledge Integration was incorporated in 1999, giving us many years of unrivalled experience in the provision of bespoke digital solutions. We have a team of skilled, qualified developers and consultants in place to ensure our services meet and exceed our customer's requirements.

Our role as a service provider of the Arches software and member of the community dates back to 2017, where we worked with the City of Lincoln on the first implementation of Arches in the UK. As of 2025, we have worked with 15+ Arches customers worldwide, and have membership in the Arches Governance and contributor status in the open-source developer team.

We traditionally focus on the problems of ETL and data migration from legacy systems into Arches, in addition to developing bespoke customisations and extensions to both individual projects and the core software. We use our experience and skills to support a wide range of customers and solution requirements.

Service Definition Arches for Science

What the service is

Arches is an open-source (AGPL3) data management platform designed for the cultural heritage sector. Created in 2012 with funding from the Getty Conservation Institute and World Monuments Fund, it is utilised worldwide by a wide range of Heritage Inventories, Historic Environment Records, Heritage Science organisations, and Digital Humanities and Research.

It uses advanced data modelling to manage, define and structure data and controlled vocabularies; discovery and visualisation tools to search, report and visualise data, with emphasis on geospatial data; and project/task management tools to manage sophisticated data editing procedures.

Arches for Science (AfS) is a comprehensive, purpose-built open-source data management application designed for heritage scientists, professionals and organisations. As an application of the Arches platform, it incorporates the benefits and features of the core software.

With a focus on the technical examination of art, the software includes tools for analytical image annotation and comparison with an integration with the IIIF Service Manager, graphical visualisation of instrument data, and material sampling. Heritage Science ontology-based data resource models have been designed using the LinkedArt subset of the CIDOC Conceptual Reference Model (CRM). The use of workflows helps streamline and separate out data entry into common tasks, and generate reports through a direct integration with Arches data.

Who uses Arches

Arches is widely and internationally used in the cultural heritage sector, used for Heritage Inventories (including Historic Environment Records), Heritage Science, and Digital Humanities and Research worldwide.

Our customers include:

- Historic Environment Records: The City of Lincoln HER, The East Sussex HER, The Isle of Man HER (Manx National Heritage), Jersey Heritage
- Heritage Inventories: Maritime Asia Heritage Survey
- Conservation Science: Royal Botanic Garden Edinburgh Centre for Middle Eastern Plants

Arches Features

Arches can be used to:

- Deliver a modern, scalable and extensible GIS-focused inventory solution
- Employ advanced data discovery mechanisms (including semantic, temporal, and spatial) and visualisation
- Model your data using ontology-based standards including CIDOC CRM
- Construct and manage controlled vocabularies
- Provide configurable data, user and group access control and support SSO
- Provide built-in geospatial tooling with configurable mapping (including historical maps, overlays, and satellite imagery) and seamless GIS integrations
- Provide a coherent and accessible user-interface supporting internationalisation
- Customise and validate data entry with ontology driven templates
- Migrate transform and distribute your data in bulk
- Provide data management, versioning, auditing and visualisation (reporting, relationships)
- Deliver API services for integrations and the synchronisation and population of data

Arches for Science can be used to:

- Ontology driven resource models derived from LinkedArt and CIDOC CRM standards, designed for Heritage Science
- Curation workflows for streamlining data entry
- Provide analytical image annotation for object sampling
- Visualise, compare and interact with 2D graphical charts of data including instrument data
- Analyse images such as elemental maps using the IIIF image service manager comparison tool
- Provide workflows for generating object report documents generation autopopulated with Arches data
- Integrate with the Arches for Reference and Sample Collections (RaSColls) application

Service Interface

The accessible and access controlled Arches service user interface covers all core functionality of the platform, and supports localisation across multiple languages. Discover and visualise data using semantic, advanced query, and map-based searching. Construct

data models and terminologies using the comprehensive and configurable, ontological (CIDOC CRM) focused resource model designer interface, and the controlled list manager, respectively. Populate and manage resources using the resource editor or extensible workflow stages, and view with configurable reports. New interface pages can be created and modified using standard web framework methods, or extended with custom made or community plugins or “Arches Applications”.

Accessibility

The public facing Arches interface is WCAG 2.2 AA compliant

Accessibility of the public-facing Arches interface is of great importance to the open-source project. It is a shared community responsibility and well documented in the Arches documentation. Approved tools and browser extensions are used for testing, including but not limited to:

- Daily retained for 1 week
- WAVE Evaluation Tool
- Lighthouse
- Landmarks
- NVDA screen reader
- W3C HTML Validator
- Contrast checker

Developers are required to run accessibility checks when developing new features or fixes. Any contribution that does not account for accessibility is not contributed back into the software until such tests pass. All customisations to our projects follow the same requirements and guidelines.

API

Many features of the Arches software are accessible via permission-based APIs. The service has APIs for authentication using the secure OAuth 2.0 authentication mechanism, data population/modification for the ability to manage Arches data when integrated with additional solutions or during data migrations and extractions, and paginated searching and data retrieval in JSON or JSON-LD (for linked data) formats. Also included are a number of geospatial endpoints, including spatial view management for retrieving, creating and editing Arches spatial views, and the GeoJSON endpoint for accessing a GeoJSON representation of resource instance data. A history and changelog of collections and resource instance data API is available for use with integrations including the CIIM middleware platform.

Built on a modern and extensible web framework, any missing API functionality can be introduced through plugins or “Arches Applications”

The service is managed via a tailored CI/CD pipeline which both monitors your configuration and tests/deploys changes to using the Argo API into Kubernetes.

Customisation

As an open-source application, any feature or modification can be introduced provided that it aligns with the AGPL3 licence and the project code of conduct.

The service offering can be customised via the user interface in a multitude of ways; buyers can configure data structures and entry fields using the resource model designer, and the resulting public appearance of the data reports templates, plus any specific data customisations such as geospatial styles. Controlled vocabularies used within data can be configured and extended, allowing a buyer to implement their own lists or customise existing ones. A buyer has full control of their data, and possesses the ability to import, export and modify existing data in bulk. Advanced searches can be carefully constructed and saved with images and descriptions for future use or public display.

Further non-user interface customisation can be implemented through written software extensions called “Arches Applications”, or integrations with other components (mapping services) or software.

Levels of backup and restore

The service is backed up/restored using native cloud backup resources. Levels of backup and restore times are tailored per customer, as this is dependent on the data, response, and backup levels required. This is created and agreed in the customer SLA. The default is:

- Daily retained for 1 week
- Weekly retained for 3 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 - 17:30). 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.

Frontline Support: We have a dedicated front-line support function where staff will 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.

Community: Arches is a growing international open-source community funded and supported by the Getty Conservation Institute. Users can engage through a range of methods including the Arches Community Forum, User Interest Groups (UK and US) and regular blog and quarterly bulletins. Developers and technical individuals within the Arches community meet at least once a year in person to discuss and present new developments and interests, and regularly online, through online developer conferences or stand-up calls.

All provide an opportunity to raise questions, concerns and influence the future of the project.

Service Constraints

Support SLA covers business hours of 09:00-17:30 weekdays and excludes weekends and bank holidays. Support requests can be raised as tickets via Zendesk outside of these hours and will be picked up when working hours resume.

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 penetration tested annually.

When the systems are configured (as part of the onboarding), limits are placed on the amount of resource each tenant can utilise. This guarantees 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.

Our services are deployed within a Kubernetes cluster, with cluster management via Argo. 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 '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 and can 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 requirement for using Arches is to have stable internet access, the latest version of popular browsers and, API keys/tokens for accessing mapping or analytical services.

Onboarding

We initially liaise with buyers to investigate their requirements and intentions with the Arches platform, and whether a stock, customisable instance or an existing application with pre-existing data structures (Arches for HERs or Arches for Science) is best placed. The selected service is auto-deployed in a blank state, ready for instant use to buyers. Required data migration projects will commence separately after deployment.

Initial training on the Arches platform is provided by the Arches team for all classes of users 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.

Documentation

Arches has an open community documentation site covering system requirements, how the software works, and various other tutorials and help sections:

<https://arches.readthedocs.io/en/stable/>.

Outage and maintenance management

In the case of an unexpected outage, the system will automatically restart and 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 different formats (including CSV, JSON, SQL, XML) and provided free of charge. Data can also be exported in alternative formats as required. All data will be available at the offboarding stage.

Service usage metrics

The Arches service has data history and auditing analytics, in addition to the ability to integrate with Google Analytics for site-wide metrics.

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

These metrics are available via:

- Through an API
- Real-time dashboards

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 via secure.