Service Definition

The name of the Service is: **Data Services as a Service**

An overview of Data Services as a Service

The Data Services service enables data stored in legacy/operational systems, in the cloud and on-premise, to be exposed as a web service and publish via the Cloud as an API. It is the basis for enablement of e-Government allowing collaboration and re-use of the existing systems/processes and the orchestration of new services/business processes. APIs are exposed in a controlled and secure manner. Data access is provided without ever having to open up databases to networks or the cloud. One of the key initiatives that can be supported is open data - making public sector data accessible on the internet through a secure and manageable process.

Data Services support a number of business use cases including:

- Open Data
- Enterprise Application Integration
- Master Data Management
- Operational Intelligence
- Content Management
- Mobile and Web Applications

Data Services is architected for the cloud including following capabilities:

- Full multi tenancy – allows each department/agency to deploy services/APIs meanwhile maintaining separate records for citizens, employees, properties etc
- Elasticity and load balancing – allows computing resources to be dynamically allocated as needed
- Identity management – based on roles and policies to ensure only authorized users have access to the services, API and data sets
- Data services – permit new cloud applications and APIs to pull data from disparate systems in order to deliver the necessary information or complete a transaction
- API management – allows extension of government data/information and services to other government agencies or authorized 3rd parties (e.g. private sector, voluntary sector) who can then deliver own applications or services on the cloud platform
- Monitoring and billing – allows government agencies to track and bill citizens or 3rd parties including the monetization of APIs and common services
- Governance – ensures data services, new processes and APIs comply with government laws, policies and procedures
Data services enable queries, composite queries, nested queries and full data sets to be published using open APIs. Update transactions are also supported. Data sources supported include relational databases (eg Oracle, DB2), web services, comma separated files, Excel files, Google spreadsheets, LDAP, ODS, RDF, Hadoop, Cassandra and MongoDB.

Supported secure API interfaces are WS-* style web services (SOAP) or an API (REST). Supported data formats are XML and JSON. Drivers are supported for multiple languages/environments including Java, Javascript, .NET, Perl, PHP, Ruby and Scala. APIs supported for SOAP/XML and REST/JSON. Many standards are supported including WSDL, MTOM, XOP, SOAP with Attachments, WS-Trust, WSSecureConnection, WSSecurityPolicy, WS-ReliableMessaging, WS-Policy, WS-BPEL and WS-HT.

Data transformations are supported using XSLT and XQuery. An ETL process is supported to load data from external data sources into the embedded MongoDB NoSQL database. This stores data in a native JSON format and enables rapid integration and aggregation of the data.

The embedded MongoDB database supports a NoSQL model to deliver highly flexible, high performance and scalable data management platform in support of agile service delivery, mobile application delivery and web application delivery. Key capabilities include:

- Flexible schema – a document oriented model makes it easy to store new content without altering the schema of the database
- Data as documents – the document model can store rich pre-aggregated cubes and joined data (from multiple data sources) supporting high performance and fast response times
- Scale-out storage and processing architecture – the database scales out for query/reads and updates/writes processing; and provides 24x7 uptime leveraging cloud platform
- Memory mapped storage engine – keeps active data sets in memory leaving rest of the data on cloud storage until it is accessed; delivers fast response times for dashboards and drill-downs
- In-database aggregation – use of aggregation framework and map-reduce enables dynamic roll-ups and aggregations of the data sets on demand
- Atomic updates – new content can be added to the database in real-time without requiring schema changes
- Large object storage – large objects can be stored such as audio, video, images and large documents in the same database including workflow processing for content management processes
Deployment Models

How can the service be deployed?

The Data Services service can be deployed on private or public cloud such as Amazon EC2. On a public cloud there is a single container instance of the multi-tenant service and shared cloud database.

By partnering with Memset Ltd we offer the Data Services service on a secure private cloud IL2 accredited for G-Cloud. On the secure private cloud deployment each client (organisation) deployment will have a dedicated VM (container) and a dedicated database. A dedicated secure private container with multi-tenant service can be provided for a single organization.

Networks to which the service is connected (directly)?

The Data Services service is accessible from a browser with an internet SSL connection.

Other networks (PSN, JANET, GSI) can be enabled by working with our partner Memset Ltd.

API access available documented and supported?

SOAP and REST APIs are available.

Open standards documented and supported?

Data interchange is supported through XML and JSON formats. Data Services can be published as REST and SOAP services supporting a number of WS-* standards. Business process can be constructed using WS-BPEL and WS-HT standards.

Open source used and documented?

<table>
<thead>
<tr>
<th>Function/Purpose</th>
<th>Open Source Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Debian Linux</td>
</tr>
<tr>
<td>Application Server</td>
<td>WS02 Application Server (Apache)</td>
</tr>
<tr>
<td>NoSQL Database</td>
<td>Cassandra and MongoDB</td>
</tr>
<tr>
<td>ETL</td>
<td>Pentaho Data Integrator (Kettle)</td>
</tr>
<tr>
<td>Java Libraries</td>
<td>ExtJS, JQuery, D3.JS</td>
</tr>
<tr>
<td>Data Services</td>
<td>WSO2 Data Services Server</td>
</tr>
<tr>
<td>Platform as Service</td>
<td>WSO2 Stratos</td>
</tr>
<tr>
<td>Business Process</td>
<td>WSO2 Business Process Server</td>
</tr>
</tbody>
</table>

Service Management

Technical boundaries/interfaces of the service documented?

Commercial In Confidence.
Copyright 2012 Cloud Intelligence Services Ltd. All Rights Reserved
Data Services as a Service for the G-Cloud

The Data Services service interfaces are fully documented for both the role of administrators and end-users.

Services available to other suppliers so they can use them to provide services to government?

The Data Services service is available to any other suppliers who wish to leverage the G-Cloud framework.

On-boarding process e.g moving to the service?

An administrator account is created for the organization. The administrator can then create end user accounts. Additionally the on-boarding is provided as a managed service.

The SLA for on-boarding is 24 hours.

Off-boarding process e.g moving off the service?

The administrator can remove the end user accounts and the data. Additionally the off-boarding is provided as a managed service.

The SLA for off-boarding is 72 hours.

Data extraction/removal criteria met?

The criteria are met in accordance with the G-Cloud ITT glossary definition.

Data processing and storage locations defined?

For the G-Cloud the IaaS option available is Memset (located in Guilford, UK).

Outside of the G-Cloud framework a public cloud option Amazon EC2 (Dublin, Ireland or Virginia, USA) can be provided.

For an IL2 compliant and UK sovereign G-Cloud solution the preferred approach is to use Memset IaaS

Data location option can be defined by user?

This is not possible. The organization purchasing through G-Cloud would only have the option of the Memset cloud which is managed in a UK sovereign Tier-3 data centre.

Data centre tier?

Tier-3 UK data centre through our partner Memset Ltd.

Support boundaries/interfaces documented?

This is defined in the support guide.
Data Services as a Service for the G-Cloud

Service roadmap provided?

The next release of the Data Services service is scheduled for Q4 2012 and will include:

- Advanced visualizations for end-users based on the D3.JS data driven documents.
- The service should be IL2 assurance certified providing information assurance to store data to the PROTECT level.
- Additional roadmap from WSO2 (Data Services Server, API Manager and Stratos PaaS) and 10gen (MongoDB) open source projects
- Data Services Server 3.0.0 release
- Impala support

Performance attributes defined and documented?

Performance uptime and response time are documented in the support guide. Database and VM servers are constantly monitored.

Backup & Disaster Recovery?

The instance is backed up with nightly snapshots. Databases are constantly monitored.

Is a support service provided and documented?

The support service is documented in the support guide.

Real-Time management information available?

Management information is provided in a monthly report. Real-time monitoring of the data services server and the NoSQL database is provided through API and web pages.

Self service provisioning/de-provisioning?

This is partially supported. The service itself is provisioned by vendor. But the end users can be added and removed by the client’s administrator.

Time for provisioning/de-provisioning documented?

24 hours for on-boarding and 72 hours for off-boarding. Administrators can add end-users in less than a minute.

Service desk can be used by 3rd party suppliers?

The service can be used by any 3rd party suppliers in the G-Cloud framework.

Commercial In Confidence.
Copyright 2012 Cloud Intelligence Services Ltd. All Rights Reserved
Data Services as a Service for the G-Cloud

Commercial

Unit based pricing model?

The service pricing is £4,000 per server per month. Excluding VAT.

The database storage is up to 250 GB of data with additional data £200/month for 100GB. Excluding VAT.

Aggregated billing?

Billing options can be provided for organizations with multiple departments.

Minimum Contract/Billing Period?

12 months at £4,000/month for a total £48,000 contract value. Excluding VAT.

Free Option?

A free option is not standard offering.

Trial Option?

A trial option can be provided on public cloud (shared container, multi-tenant). The fee would be £1000/month. Minimum term for a trial service subscription is 3 months £3,000. Excluding VAT.

Termination Costs?

If terminated within the 12 month term then the remainder of the annual fees will be billed. If terminated at the end of the term then there is no cost. Trial subscriptions can’t be terminated.

Supplier contract terms/jurisdiction?

Cloud Intelligence Services Ltd and Amtex Solutions Ltd are located in the UK under jurisdiction of English law. The IaaS provider is also located in England.
Data Services as a Service for the G-Cloud

Payment Options?

Purchase Order and BACS are supported

Data Visualization

Authorized users may store datasets?

This is core functionality of Data Services. End users can upload their datasets to the cloud service into a embedded NoSQL database. Or alternatively the access is provided to databases such as MySQL, Oracle, SQL Server, Spreadsheets, Files etc.

Authorised users may make datasets public?

This functionality can be enabled using API Manager publishing data as an API (SOAP/XML or REST/JSON) for the query and export of data sets

Part of an integrated collaborative tool set?

The service is not part of a collaboration tool although API can be integrated into other collaborative tools or workflow processes. Additional support for WS-BPEL and WS-HT can be provided to support collaborative workflow processes.

Publish datasets as visualizations?

This is core functionality

End users may search for datasets?

Search is not part of the current functionality

Types of display?

Some advanced visualization documents include box plots, bubble charts, bullet charts, calendar view, cartogram, chord diagram, dendrogram, force directed graph, circle packing, population pyramid, stacked bars, steamgraph, sunburst, node-link tree, treemap, voronoi diagram and tag clouds.

Maximum dataset size?

Commercial In Confidence.
Copyright 2012 Cloud Intelligence Services Ltd. All Rights Reserved
Data Services as a Service for the G-Cloud

The standard embedded data storage for the service is 250GB although federated access is available to any data set.

Specialist Cloud Services - Generic

*Do you provide vendor specific services?*

Services are provided for the implementation of the Data service and connectivity.

*If the vendor(s) have accreditation are you accredited?*

UK G-Cloud IaaS partner provider Memset Ltd has a number of accreditations and certifications e.g ISO 9001, ISO 14001, ISO 27001. Amtex Systems has ISO 9001 certificate for IT enabled services, business process outsourcing, software testing, data integration and reporting.

*Vendor accreditation?*

We will submit the Data Services Software as a Service for IL2 accreditation in the near future.
Service Levels

Platform IaaS provider partner Memset Ltd has a 99.995% uptime guarantee for networks, servers and storage. 24x7 support is provided for the IaaS and this support is wrapped into the Data Services support levels.

For the Data Services and API Management service production support is provided 24x7x365. Incidents are submitted via e-mail and/or telephone. The target response and resolution times for incidents are:

<table>
<thead>
<tr>
<th>Severity Level</th>
<th>Response Time</th>
<th>Resolution Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 hour</td>
<td>24 hours</td>
</tr>
<tr>
<td>2</td>
<td>4 hours</td>
<td>48 hours</td>
</tr>
<tr>
<td>3</td>
<td>8 hours</td>
<td>72 hours</td>
</tr>
<tr>
<td>4</td>
<td>24 hours</td>
<td>None</td>
</tr>
</tbody>
</table>

The severity levels are defined based on impact on a production usage of the Data Services service in the table below: “Production” is defined as one by end users where a failure of the Data Services service in production will have immediate economic impact on the client’s business operations”

<table>
<thead>
<tr>
<th>Severity Level</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Client’s business operations have been disrupted either through a loss of data or a loss of service.</td>
</tr>
<tr>
<td>2</td>
<td>Client’s business operations can continue in a restricted fashion. Client’s business operations are at risk.</td>
</tr>
<tr>
<td>3</td>
<td>Some aspects of the service are impaired but client can continue to use the service. Client’s business operations are at minimal risk.</td>
</tr>
<tr>
<td>4</td>
<td>General usage questions, requests for new features or documentation issues.</td>
</tr>
</tbody>
</table>
Termination Options

Clients may terminate the production Data Services service with 30 days notice for any reason, subject to full payment of remaining SaaS fees for the 12 month subscription period.

Trial service cannot be terminated. Client will be invoiced the full amount for the 3 month trial in advance.

Vendor can terminate the production Data Services service at the end of the initial 12 month subscription period with 90 days notice for any reason.

Data Restoration/Service Migration

Data restoration is provided by vendor for recovering from backups. Daily snapshots are taken with options to restore from the automated backups.

Service migration can be provided as a service through an export of the database environment or a migration of the complete VM onto another cloud instance.

Customer Responsibilities

The client responsibilities are set out in the support guide.

More Information

More information on the service can be found at www.cloudintelligenceservices.co.uk