# Assessment Report on the second Release of FAIMS Repository

FAIMS Deliverable #12D

prepared by: Adela Sobotkova

This report summarises the state of the FAIMS Repository v0.2 released on 19 July 2013, after 14 weeks of development.

## Background:

The FAIMS Repository is a customisation of the Australian Historical Archaeological Depository (AHAD), which is in turn a customisation of the Digital Archaeological Record (tDAR), a large existing open source code base developed by the Arizona State University in the United States under the auspices of Digital Antiquity.

The FAIMS project is expanding AHAD to create a digital repository for documents, datasets, images and sensory data produced by archaeological research in Australia or collected by Australian archaeologists working abroad.

In January 2013, Digital Antiquity revealed a major new release of tDAR (version: In Situ) with a rebuilt search engine, bug report, ability to handle deposit fees, and a whole new look and feel.

## Glossary

AAF -- Australian Access Federation

AHAD -- Australian Historical Archaeological Depository

NeCTAR Cloud -- computational infrastructure provided by the partnership of NeCTAR, Research Data Storage Initiative (RDSI) and Australian National Data Service (ANDS).

## The Aims of Second Release

The requirements and desired features of the second release of the FAIMS Repository combine items deferred from the First Release and those specified by the Software Development Plan (SDP) for this release, including:

* the import of data from the mobile device (SDP Iteration No. 1, has been deferred from Release 1)
* optional map obfuscation (SDP Iteration No. 5)
* segregation of records and resources from each subdomain of archaeology (eg historical vs indigenous or maritime), to enable researchers to quickly find the resources of relevance to their field of research (SDP Iteration No. 7)
* single page display of row-level data within datasets (SDP Iteration No. 8)
* export of data in machine readable format (SDP Iteration No. 9)
* ingest of spatial data and display of metadata (SDP Iteration No. 10)

This release was also to include SDP Iteration No. 6 ‘Upgrade of existing records from AHAD’ which is a content import, and does not require testing of new features. This is not ‘new functionality’ but it populates the Repository with real data that will assist with future testing of performance and useability of the Repository.

Additionally, the AAF and other NeCTAR cloud related requirements that were partially delivered during the first release remained a general development goal during this release.

The second release was originally scheduled for 1 June 2013.

## Changes to the Second Release

In the course of preparation for testing the second Repository release, however, it became clear that owing to problems with implementation, communication and management it would not be possible to deliver the milestone by 1 June. VeRSI assigned a new project manager on 6 May in order to streamline development. Given the change in team, we agreed to the delay of this milestone to 5 July and defer another features, on the condition that all deferred functionality would be delivered for the final release. In the end the Repository release was ready for testing on 19 July.

The interim descoping resulted in (1) the deferral of three features originally scheduled for the second release (spatial data upload, map obfuscation, DOI/UUID for data rows), (2) temporary descoping of one feature (mobile import, see next section for details) and (3) continued deferral of one external dependency from the first release (AAF login) as it continues to be unsupportable. These features are listed below with explanation for their deferment:

* Upload of spatial data as a new data type (SDP Item 10, Story ID 14). The Digital Antiquity team is currently introducing spatial data as a Resource Type within tDAR and will be testing this new feature. Once it is completed, we will evaluate its utility for the FAIMS Repository. If no changes are required and/or testing by the tDAR team is sufficient, no work will be required by VeRSI and no separate FAIMS testing will be necessary. It is envisaged that this functionality will be used to import the spatial component of the data from the mobile application.
* Optional map obfuscation (SDP Item 5). This will allow users to make the site location visible on an overview map if this is explicitly enabled for a resource. There has been insufficient time to complete this task for this release.
* DOI for data rows (SDP Item 12a). This will allow for the creation and storage of Digital Object Identifiers for individual artefacts, stratigraphic contexts or any other information at the data row level. This has proven to be more difficult than originally anticipated.
* AAF Login (SDP Item 2) was evaluated and it was found that it currently does not support a user model that is required by both the FAIMS project and the existing tDAR source base. As an externally dependant feature, further input has been requested from NeCTAR to find a suitable solution.
* Upgrade of AHAD data (not a new functionality).

## 0.2 Release - New Features

As a result of the descoping, the second release of the Repository included three originally scheduled features for new development (single-page display, segregation of records and XML export of data), one newly requested feature (email notification), and a descoped version of the mobile import deferred from the first release:

* Email notification of newly registered users (additionally requested feature)
* Import of data from FAIMS mobile application as a “tarball” , which is a zipped file whose contents are not searchable beyond what it in the metadata. This is an interim compromise to the original requirement, which calls for the import of the mobile data in an immediately searchable form (SDP Item 1, Story ID 19, 7, deferred from the first release and full feature devolved to final release).[[1]](#footnote-0)
* Single page display of row-level data within the data sets (SDP Item 8, Story ID 15).
* Segregation of records and resources from each subdomain of archaeology (e.g. historical, indigenous or maritime), to enable researchers to quickly find the resources of relevance to their field of research (SDP Item 7). Rather than coming up with our own FAIMS typology of the archaeological subdomains, we chose to implement this search using the Field of Research (FoR) codes, familiar and in wide use among Australian archaeologists.
* Export of data in machine readable format (SDP Item 9).

These features of the FAIMS Repository have been tested and reviewed with the help of external testers on 19 July 2013, and approved by the FAIMS Steering Committee delegate on 7 August 2013.

All of the tested functionality passed.

## Issues:

During the second release difficulties emerged with tDAR code and the FAIMS requirements. The VeRSI developer continued to encounter difficulty parsing the mobile data tarball, although eventually (3 July) this was clarified as a misunderstanding (complexities in the resulting datasets were mistaken for errors in the data extraction). Other options are now being investigated to extra data in the formats expected by tDAR. This is one of the challenges of building connections between developing tools and complex software suites designed with different purposes (in tDAR’s case, legacy data). (In contrast, Heurist developers were able to parse and ingest the data after the release of the first tarballs in mid-April.) The difficulties are compounded by the fact that the Repository team is working inter-state and under pressure of a major reorganisation and merger.

Several Repository features are currently under threat of non completion:

1. Map obfuscation—a feature that will provide flexibility in the sensitive data capabilities of tDAR;
2. Left join—a feature that would allow links between artefacts and contexts and other relationships created on the mobile device—may need to be dropped given the fact little time is left until the final release;
3. The implementation of persistent pointers, DOI (Digitial Object Identifier) or UUIDs (Universally Unique Identifiers) appears intractable due to tDAR’s refreshing of the stored data tables (tables being the ‘atomic unit’ of storage in tDAR; FAIMS would like artefact rows to be our ‘atomic unit’);
4. The shapefile ingest, deferred by VeRSI because Digital Antiquity already has that on their development plan;
5. The import of records from AHAD has fallen off the table during the second release and will likely require an extension of the final release to ensure its completion.

## Overall Repository Summary as of Second Release: planned vs. achieved

All in all the status of deliverables scheduled for the first and second releases can be summarized in two tables. The first outlines the fundamental FAIMS objectives for the Repository, and the other lists useful enhancements that we wished to implement in the tDAR code by the second release. VeRSI delivery of these features is listed in the right column.

|  |  |
| --- | --- |
| **FAIMS overarching goals for Repository** | **Delivery status against** |
| Widely and easily accessible online Repository  (NeCTAR Cloud, block-storage) | partially delivered |
| Persistent well archived data (ingest of tarball) | delivered |
| Searchable data (ingest of mobile data) | not delivered |
| Publishable, shareable, and reference-able data (DOI/UUIDs) | not delivered |
| Enhanced data sensitivity (map obfuscation) | not delivered |
| Storing of relationships between eg. records and images (left outer join) | not delivered |

|  |  |
| --- | --- |
| **FAIMS enhancements to tDAR code** | **VeRSI delivery status** |
| Row-level pages (useful mostly when combined with DOIs) | delivered |
| Export of data in machine readable format | delivered |
| Shapefiles/GIS import | not delivered (prepared by DA?) |
| Facetted search (using FoR codes) | delivered |
| Rebranding (FAIMS logo and colours) | delivered |
| Australia-wide easy login (AAF) -- descoped | not delivered/Dropped |
| Automated email notification to new users | delivered |

## Discussion:

The development undertaken on the v0.2 of FAIMS Repository was conducted after major delays to the schedule determined during the project Elaboration Phase. After substantial descoping, the remaining stories passed User Acceptance Testing. Three stories were removed from the UAT Criteria, the DOI/UUID at row level for technical reasons, the Map Obfuscation (Sensitive Data management) because of lack of time, and Spatial Data import because the functionality is being developed by Digital Antiquity. The Mobile Data Import was further descoped into zipped file upload, because of increased technical difficulty and delay, on the condition that full data ingest will be attempted by Final Release. All of the deferred features remain for the Final Release next to the ones originally scheduled for it. External dependencies, such as the AAF, are on hold and will be revisited once NeCTAR provides resources and guidance on the advance.

Unfortunately, the Repository Development continues to fall behind, and there is considerable risk of delayed delivery or failure to deliver of critical-to-FAIMS features.

The communications have improved thanks to the new manager on the VeRSI side of project, but we are uncertain whether the ground lost in the interim can be made up by the project’s end.

## Summary

The first release of the FAIMS Repository (v0.2) offers the following functionality: email notification to new users of the Repository, ability to upload and store tarballs of mobile data from the FAIMS Mobile application as a single zipped file, ability to view row level records on a single page, searchability of the content using the FoR codes, and ability to export data in machine readable format.

This new functionality of the second release of FAIMS Repository was subject to external testing and has been validated to the satisfaction of the Steering Committee.

The development of the Repository has fallen behind schedule and was revised due to the technical and managerial challenges involved. We are re-prioritizing to retain the functionality crucial to the overarching FAIMS goals. We may need to apply for an extension to development in order to ensure that the critical outstanding features make it into the Final Release.

1. I quote here from the Assessment Report on the First Release of FAIMS Repository(#12B): “Support for mobile device data import. In terms of supporting mobile data, it was originally envisaged that the existing ingest mechanisms of tDAR would be used. However, it has become apparent in the course of elaboration and development that this is no longer a satisfactory approach, and significant work will have to be done to bring in the mobile data. This work has been started in the week of 4-8 March. Given this delay, this story will be relegated into second release (1 June 2013) and tested in the associated UAT.” [↑](#footnote-ref-0)