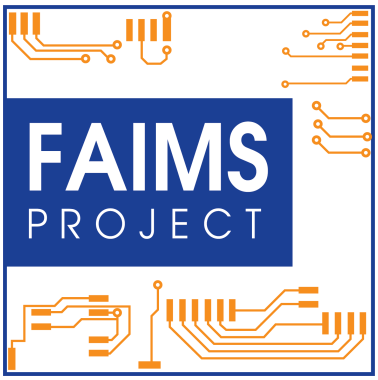
FAIMS Stocktaking Workshop Working Groups, a Report

25 November 2013

Web Version 1.0

Adela Sobotkova

This report summarizes the aims, organization and outputs of the Working Groups created by the Federated Archaeological Information Management Systems (FAIMS) project during the Stocktaking Workshop in August 2012. The Working Groups comprised the community-driven part of the Stocktaking Workshop and produced lists of requirements, recommendations and feedback on the planned development of the FAIMS digital infrastructure for Australian archaeology. While the lists of questions for the Working groups were developed by the FAIMS Leadership team, the responses reproduced here represent an aggregate of notes and reports submitted by the Working group participants.



http://www.fedarch.org

CONTRIBUTORS: Australian Heritage Management Service Pty Ltd nsw ★ Australian Cultural Heritage Management Pty Ltd vic ★ Archaeological Map of Bulgaria (bg) ★ University of California Berkeley (us) ★ University of Canberra ★ Flinders University ★ Godden Mackay Logan ★ Indiana University of South Bend (us) ★ James Cook University ★ University of Leicester (uk) ★ Macquarie University ★ Michigan State University (us) ★ University of New England ★ University of Queensland ★ Snappy Gum Heritage Services ★ Southern Cross University ★ University of Sydney ★ Vanderbilt University (us) ★ University of Western Australia

SUPPORTERS: Archaeology Data Service ★ Digital Antiquity ★ University of Texas, Austin (us) ★ Australasian Institute for Maritime Archaeology ★ Australian Archaeological Association ★ Australasian Society for Historical Archaeology ★ ACT Heritage Project ★ ACT Heritage Unit ★ NSW Heritage Branch ★ NSW Office of Environment and Heritage ★ Monash University ★ University of Melbourne ★ University of Sydney Library ★ Earth Imprints Consulting

nectar UNSW AUSTRALIA LA TROBE UNIVERSITY INTERSECT VeRSI THE UNIVERSITY OF SYDNEY

Contents

1 Introduction

Federated Archaeological Information Management Systems (FAIMS, see www.fedarch.org) project is an Australian initiative funded by the National eResearch Collaboration Tools and Resources (NeCTAR, see www.nectar.gov.au) program to develop an ecosystem for managing archaeological data from their digital creation to archiving. From the beginning the FAIMS project has focused on developing three main components:

1. Mobile application for archaeological data recording;
2. Web application for data processing, and;
3. Digital repository dedicated to the curation of archaeological data.

These tools were developed either from scratch using an open source platform or by reusing an existing code. The Stocktaking Workshop was organized in August 2012 (after the launch of the FAIMS project in June 2012) to solicit input about development directions for the FAIMS project and to inform decisions about structure of the project, technologies, and modules to be supported. One of its aims was to gauge the level of convergence on data standards across academia, industry and state agencies as well as across different archaeological subdisciplines (survey, excavation, artefact analysis and sciences). Originally the Workshop was envisioned to have 35-40 participants with 10 international participants discussing over two days. The response to the call, however, was overwhelming and the duration of Workshop was extended to four days from Thursday 16 August to Sunday 19 August. The Workshop was free to attend for registered participants. The webpage for the Workshop was created using OCS and linked to the FAIMS webpage. Registration, as well as the communication and information related to the Workshop were channeled through this webpage.

1.1 Workshop Structure

In the end the Workshop attracted 80 attendees from all over Australia as well as overseas. The numbers stayed high on Thursday and Friday 16-17 August, and dropped gradually over the weekend. The least attended session on Sunday had 40 participants. Participants included members of the archaeological community, both academic and public, industry and state agency officials as well as all FAIMS developers and NeCTAR representatives. The Workshop was structured as a stocktaking event with Plenary Sessions driving the agenda of smaller afternoon discussion Working Groups and an occasional Focus Group. The Plenary Session speakers presented successful overseas initiatives in digital archaeology and broadened the horizons of the FAIMS debates. Plenary Sessions were recorded and filmed. The records can be found on the FAIMS project webpage[1] as well as in the Agreed Standards Report[2].

After the inspiration provided by the morning presentations of foreign and domestic case studies, the FAIMS stakeholders were broken up in small common themed Working Groups, where they discussed recordkeeping standards, requirements for the mobile systems, federation strategies and other specific

[1] *https://www.fedarch.org/wordpress/category/workshop/workshop-presentations/*

[2] *https://www.fedarch.org/wordpress/mission/project-reports/faims_del4/*

issues that would inform the FAIMS project development. The Working Groups were further complemented by Focus Groups, in which select Working Groups or other subgroups of stakeholders discussed topics under the guidance of an experienced Focus Group facilitator. The results of these Focus Group discussions can be found in the *FAIMS Stocktaking Workshop Focus Groups, a Report*[3].

1.2 Working Group Structure and Membership

Working Groups were intended to generate the basis for the requirements of the FAIMS project. Each group was assigned a specific topic, with a number of questions to guide their discussion. Working Group membership was on an ad-hoc basis (self-selected) and fluctuated throughout the Workshop. The only permanent participant was the Working Group Leader, whose leadership was solicited before the Workshop and who committed to stay on for the whole duration of the group. The Leader was responsible for ensuring that the results of the discussion were reported back daily during evening wrap-up plenary session. The Leader received the information package (see ***Appendix 1***) which provided the Working Group with purpose and questions. Individual Working Group participants were encouraged to continue in the interaction online and post the minutes of their discussions on Google Groups set up for that purpose by FAIMS. A Wi-Fi connection was provided free of charge for all participants at the Workshop. Seven Working Groups were established:

1. Archaeological Survey,
2. Excavation,
3. Artefact recording (which eventually split into)
    1. Ceramic, and
    2. Lithic recording,
4. Archaeological Sciences,
5. Federation,
6. Sustainability, and
7. Sensitive Data.

Four of these groups (Artefacts, Survey, Excavation and Archaeological Sciences) were supplemented by a Focus Group element. The Working Group meetings were scheduled on Friday and Saturday afternoons and Sunday morning. Each Working Group had the opportunity to meet multiple times during the Workshop upon request. Group Leaders could decide whether to call an additional meeting after reading a 15 minute report on daily activities at the evening wrap-up. During this reporting meeting they informed the other stakeholders of the Working Group's progress and map it against the list of original questions.

[3] ***http://hdl.handle.net/102.100.100/11391***

2 Working Group Themes and Outcomes

This section includes a list the original questions for each Working Group, organized alphabetically. Notes typed up and submitted by individual Working Group participants to the FAIMS leadership are appended under Participants' Reflections. The notes vary from direct answers to our questions, mapped vocabularies and minimum threshold vocabularies, to broader discussions of best practices in archaeological work and documentation. This report represents the most complete summary of the Working Group documentation. Group participants have been anonymised for the purpose of this report.

2.1 Archaeological Sciences Working Group

2.1.1 Purpose

Formulate the minimum requirements for:

1. facets of scientific data to be captured on the mobile device application
2. functionality of the capture of scientific data on the mobile device application
3. eResearch tools required for searching scientific data
4. eResearch tools required for the analysis of scientific data

The working group should also discuss and where relevant provide direction on any other aspect of the FAIMS project discussed during the course of the Workshop which may explicitly affect the capture, management or analysis of data relevant to archaeological sciences

2.1.2 Topics to Consider

1. Is there a universal artefact record that applies to all types of archaeology (indigenous, classical, historical, maritime)?
2. What attributes do you always record? What attributes do you sometimes omit?
3. What is the minimum "granularity" of the scientific data?
4. What is the smallest atomic (indivisible) observation that can be made about your topic? Is the common record level different from this granularity? Why?

2.1.3 Participants' Answers and Reflections

The Archaeological Sciences Working Group met twice on 17 and 18 August. Membership fluctuated between 5- 8 participants.

2.1.3.1 Themes that need to be covered prior to the creation of any system or module

- Defining the archaeological sciences

  - Classification of the archaeological science as a cluster of disciplines
    1. Essential background?
  - Need to recognise the disciplinary diversity: soil science, geoscience, biological sciences, geochemistry, etc, etc. etc.
  - FAIMS project: taxonomy of archaeological sciences
- Planning archaeological excavation/survey
  - Protocol for archaeologists engaging archaeological science
  - Decision tree for archaeologists/planning
  - Note: archaeological science planning, questions, outcomes, etc. needs to be in alignment with the archaeological project purpose etc.
- Advice and expertise
  - Training, collation of training programs/modules
  - Database of experts, virtual experts, commons, web/discussion/resources, on-line/video linking, advisory service
  - Values/validity/authority
  - Expert versus general skills: tasks that archaeologists can do; when to call in an expert
- Basic archaeological science discipline understandings
  - Collate reference databases for all fields in the archaeological science; international links Identified a priority action potential project
  - Access to archaeological science field texts Identified a priority action potential project
  - Access through portals, wiki model, etc.
  - Validation process
  - Utility of reference list, filtering, decision making
  - Visibility of archaeological science
- Methods standards
  - Methodological design
  - Basic sampling, decision-making, practice, curation
  - National/international methods and taxonomy standards Identified a priority action potential project
  - Access and use of analytical technology
  - Community discussion re understanding of the context/value of standards
  - External database access issues
- Communication
  - Two-way communication archaeology-expert, access to archaeological data, interactivity with archaeological database
  - Standards for reporting
  - Feedback of archaeological science data into other fields, e.g. climate change research

2.1.4 Potential Future Extensions to FAIMS

2.1.4.1 Access to databases, texts, global standards: potential coordination project

Project to support/facilitate dynamic, interactive and web-based informed access to expert archaeological sciences services, reference collections, etc., and allow exposure of specialist collection to the archaeology community.

- Collation of archaeological sciences databases, methods, standards
  - databases for all fields in the archaeological science; international links
  - archaeological science field texts
  - national/international methods and taxonomy standards
- Use FAIMS system to find out what is out there and what is discoverable (cf. Australian National Data Service, Research Data Australia discovery service: fabric for making research collections discoverable online)
- Build architecture for researchers to put material into; access processes (portals, wikis, etc.)
- Support mechanisms to allow intelligent use of the resources.
- Feedback processes: E.g. "How useful?" button.

2.1.4.2 Archaeological Sciences awareness project: advice, training, planning, decision making

Project to provide the dynamic, interactive web-based frame of fostering awareness of the archaeological sciences and their roles in, and relationships with, archaeology, and for assisting archaeologists in making informed decisions regarding access to, and use of, the archaeological sciences.

- Defining the archaeological sciences
  - Classification of the archaeological science as a cluster of disciplines
- Planning archaeological excavation/survey
  - Protocol for engaging archaeological science; decision tree
  - Archaeological science planning, questions, outcomes in alignment with the archaeological project; etc.
- Advice and expertise
  - Training, collation of training programs
  - Database of experts; virtual experts, commons, web/discussion/resources; on-line/video linking, advisory service
  - Values/validity/authority; expert versus general skills: tasks that archaeologists can do; when to call in an expert

- Geo-referencing / GIS focus as a unifying heuristic in FAIMS
- Over and above GIS as an analytical and display tool:
- Location as an archaeological data anchor: other data as variables
- Landscape as a unifying focus
- Benchmarking values and condition
- Archaeological data feeding back into natural sciences
- Consulting link

2.1.5 Workflow in Archaeological Sciences

- Step One
  - Working out what you need; project context; resident expertise
  - FAIMS project extension and advisory service

- Step Two
  - Critical decision path, project planning
  - How? Who? What? to engage in the archaeological sciences
- Step Three
  - Basic field recording forms: Survey v. Excavation
  - Collect examples of currently used forms
- Issues
  - Links between archaeological sciences and archaeology at the planning, operational and analytical levels
  - Record of advice from experts
  - Basic id skills: e.g. is the bone human or cow?
  - Adopting protocols for sampling strategy, data collection, sample/artefact curation?
  - Tensions between observation/collection/analysis: determines purpose and structure of recording
  - Critical collection data, what not to collect … not interpretive labelling
  - Critical analytical pathways … the tyranny of the catalogue …
  - Integration of arch. Sci. sampling into site recording, artefact or sample recording
  - Basic data should be identical to all other site artefact recording
  - Alignment of external/expert labelling with site recording

2.1.6 Information needed for recording

- Basic principles:
  - Alignment with site strategy protocols
  - Archaeological science samples are treated as equivalent to archaeological artefact sampling on site
  - Must be observations, not interpretation, typology, taxonomy …
  - Context/purpose descriptor important
- Requirements:
  - PDA structures; drop down menus; auto-populating; default values determined by site strategy protocols
  - Paper-based capability
- Basics (can be refined using current forms)
  - Two part form: top part: generic for any type of artefact
  - Part One: site/sample basics:
    1. site name/code/square/geolocation
    2. date
    3. weather/conditions
    4. excavator/collector
    5. level/spit
    6. layer/XY(Z) coordinate/feature/screen (mesh size, wet/dry/flotation)
  - Part Two: Material, for all:
    1. For all: align with sampling strategy? (site protocol/other)/ purpose (general, specialist, C14, DNA, etc.)
    2. Shell

3. Animal tissue … articulation Y/N, in situ condition, whole/fragmented
4. Plant matter … All collected, bluk sample, individual, proportion
5. Geological samples: sediment, rock fragments, amount
6. Other

2.1.7 Summary

In their report, the Archaeological Sciences Group outlined the following key ideas:

1. In the recognition that (i) the archaeological sciences are multi-disciplinary, (ii) that the for any specific archaeological project, specific archaeological sciences are relevant, and (iii) that it cannot be expected that archaeologists can be archaeological science experts or practitioners, the core issue is that we need to structure a FAIMS response appropriate to the nature of the engagement between archaeology and the archaeological ssciences.
2. A FAIMS response should be based on the process flow for archaeology engaging the archaeological sciences from planning through to publication, hence important to identify the conditions and requirements for meaningful and informed engagement.
3. Note: the scale of activity is important, e.g. site/project v. discipline; hence more focus is needed on the practice of survey, field data collection, excavation, with lesser focus on post-excavation.

2.2 Artefact Working Group

2.2.1 Purpose

Formulate the minimum requirements for:

- facets of artefact data to be captured on the mobile device application
- functionality of the capture of artefact data on the mobile device application
- eResearch tools required for searching artefact data
- eResearch tools required for the analysis of artefact data
- The working group should also discuss and where relevant provide direction on any other aspect of the FAIMS project discussed during the course of the Workshop which may explicit affect the capture, management or analysis of artefact data.

2.2.2 Scope

For pragmatic purposes, FAIMS has restricted the development of digital artefact recording to two classes of artefacts: ceramics and lithics. It is likely that many of the minimum requirements for functionality, search and analysis of these two classes will be universal across all assemblages. The minimum requirements for data attributes may vary, however, there is potential that the core fields developed by this working group may become the universal model for all classes.

The development of data capture applications within FAIMS is focused on in field recording. While the software and interfaces will be useable on any desktop devices as well as mobile devices, the application is unlikely to provide the complexity required for detailed laboratory recording.

At this stage, the FAIMS team require guidance on the structure of the artefact record, not its contents. That is, we need a minimum list of facets or attribute types, not the vocabularies and taxonomies to populate them. Given the diverse nature of the archaeological community expected to use these tools, these are expected to be developed outside of the project.

2.2.3 Topics to Consider

- Is there a universal artefact record that applies to all types of archaeology (indigenous, classical, historical, maritime)?
- What attributes do you always record? What attributes do you sometimes omit?
- Are there significant differences in the way you record an artefact and the feature or context from which it came? How might that affect the development of a tool for digital capture?
- Do you photograph or sketch artefacts in the field?
- How precise are measurements of artefacts taken in the field?
- How do you access reference data needed to identify and date objects in the field? Should these be integrated in data capture?
- How often do you need to search through artefact data collected in the field (digitally or manually)? Is it critical, or is it really on accessed in the lab?
- Is there more consistency in stratigraphic recording in comparison to artefact recording?
- What visualisation or statistical tools do you rely on to analyse your artefact data?
- How do you find artefact data created by other researchers? Is searching for artefact data more complex than searching for stratigraphic or site data?

2.2.4 Expected Outcomes

- List of essential facets required for recording a ceramic artefact in the field
- List of essential facets required for recording a lithic artefact in the field
- List of functionality requirements for recording artefact data on a mobile device application
- List of functionality requirements for searching artefact data on the FAIMS portal
- Suggestions for analytical tools to be included, and if possible, developed for the FAIMS portal

2.2.5 Participants' Answers and Reflections

The Artefact Working Group met twice on 17 and 18 August. Membership fluctuated between 9-10 participants. Eventually the participants split up among artefact (4) and lithic recording groups (6), whose reflections and minimum sets of attributes are listed below.

2.2.5.1 What do we really want?

- Can you take an image and link it in directly?
- Want system to be customisable.
- Want hierarchical drop down menus.
- Do we want to be able to add vocab in the field? Would some field directors want to stop that.

- Need group record, then individual.
- Want the name of the person to come up automatically–digital signature.
- Flag for uncertainty?
- Want automatic capture of people intervention.
- Want system to capture changes when you upgrade/change records. (versioning)
- Need to be able to move between classes fluidly.
- Comprehensible system for managing of Special Find split bags.
- Need system for inventoried vs non-inventoried.
- Need system which can handle field records as "final records" and records which will be reworked.
- Get mobile device which can calculate minimum/maximum density in the field. (Must record breakage for each artifact.)
- Can you have something that will tell you whether you've reached your minimum sample as you go? (can the device help you assist in making the call on whether or not your sampling strategy is required).
- Running total of square metres excavated, and artifact counts in progress to give you density as you go.
- Georeference every artifact ID.
- Director needs to customize what is required what isn't. Choose whether vocabulary is fixed or not.
- Need to be able to re-label columns
- Would be able to colour code recording situations or anything?
- Want everyone to send in versions of their recording sheets.
- All edits and changes to schema tracked to user.
- Customise label and definition?
- Add your own column (and label and definition)
- Must be able to reuse your schema
- Profile gauge / contour caliper: ability to capture image
- Need to be able to see things in spreadsheet or form view
- Need to be able to search, filter as you go.
- Need to be able to have multiple observations. Multiple people working on the same record (at different times).
- would be good to have a forum for people to discuss the definition or properties of each column–link to datasets that use that column: good basis for comparison and evaluation of standardisation.
- Be able to link or add bibliographic citations to column header or vocabulary item (and link to other similar concepts).

2.2.5.2 Minimum record for Group of Artefacts (e.g. bag, or subgroup)

- Group identifier
- Project identifier
- Site identifier (optional)
- Location (Context or Survey Unit / GPS)
- Date recorded
- Person recorded
- What is made of? (Material or Artefact Class, eg ceramic) (one or more value)
- What specifically is it made of? (Material optional) (one or more value)

- Portion (rims, bases) Optional (one or more value)
- Quantity of pieces? (Quantity)
- Weight (Optional)
- Photograph (Optional)
- Comment (Optional)

2.2.5.3 Minimum record for Individual Ceramic Record

- Artefact ID
- Project ID
- SiteID (optional)
- Location (Context or Survey Unit / GPS)
- Date recorded
- Person recorded
- What is made of? (Material or Artefact Class, eg ceramic) (one or more value)
- What specifically is the body made of? (Paste/Material) (Optional) (one or more value)
- Extensible list of attributes (eg Paste colour, munsell colour, inclusions, direction of inclusions)
- Portion (rims, bases) (one or more value)
- Quantity of pieces? (Quantity)
- Measurements (see below)
- Integrity (includes relationship between sherds and Percentage)
- Linking relationships (ie links with other artefacts, eg conjoin)
- Photograph (Optional)
- Comment (Optional)

2.2.5.4 Measurements

- Need to record different measurements for each artefact (none, one or many).
- Minimum/suggested measurements for different types (eg function/form or skeletal element, eg bottle, scapula, bottle).
- Need measurement unit, device (eg scales/balance, rim template/diam board).
- Need to add a note or comment to each measurement
- Types of measure: body thickness, outer diam, inner diam, height, base, foot ring, rim thickness, shoulder height, capacity estimate, rim section, rim angle, weight, total length, minimum dimension, maximum dimension, rim diam–need to have rim % as well)
- Need to be able to do cross tab analysis of measurements.
- Mobile device can calculate density-low background scatter. But would have to have that beforehand-DIA could do it for each environment. Or go through old reports + look at data that has already been contracted. Find out what "background scatter" is.

2.2.5.5 Most Useful Lithic Recording Attributes

- Raw material types need to have regionally specific drop-down menus:

  - Need to redraw the distribution boundaries of artefact types in Australia by making data accessible for significance assessment
- Technological classification
- Metric types-technological categories precede the typological-distinguished chronologically
- Metrical data
  - Mass
  - Linear measurements (could be taken by laser scanner)
- Look at well collected data: what happens if you drop out attributes?
  - FAIMS portal: conduct multivariate statistics on randomised existing datasets to see what actually matters. Experiment with the robusticity of the data to establish the minimum requirements.
- Cortex-% dorsal cortex
  - Platform variables-cortex is included
  - Not a separate field
- Core rotation (Y/N) or analyse the directions of all the scars?
  - Knowing that the dataset exists is as important as access to the data
- Dorsal scars: not useful to count scars but *none one parallel many
- Stones: we should trust ourselves to tell stones that have inside ourselves
- Retouch is very important analytically, more can wait for next stage
- Overhand removal tool
- Defining background scatters on regional level-using existing data:
  - Raw material drop down menu
  - Technological class:
    1. Core
    2. Flake
    3. FP
    4. Retouch
    5. Broken
  - Metrical
    1. Metrical
    2. Linear measurement
  - Cortex platform surface, % dorsal cortex
  - Core rotation
  - Dorsal scars (0, 1,parallel,many)
  - Retouch (Y-N)
  - Overhang removal
  - Tool
- Analyse existing data and play with attributes

2.2.6 Summary

In sum, the Artefact Group outlined the following concerns:

- Need for a modular design that would accommodate regional and typological differences
- input systems based on recognised, existing standards and practices (producing a library of possible input systems)

- recording based on artefact recording workflow
- smooth image handling
- easy customization
- use of controlled vocabularies with aliases
- comprehensive management of metadata
- maximal automation
- versioning and tracking
- recording of uncertainty (either binary, as a flag, or scaled)

In addition, they recommended that any system be able to manage both group and individual records, produce rudimentary statistics in the field, provide search and query in the field, and connect to a range of external data recording devices (present and near future).

Concerning vocabularies and data standards, the Artefact Group listed two sets of minimum recording requirements and requested a discussion forum to discuss the properties of each attribute.

2.3 Excavation Working Group

2.3.1 Purpose

Formulate the minimum requirements for:

- Excavation data attributes to be deployed across the system (captured through mobile device applications and used to facilitate the production of comparable datasets).
- Functionality of the capture of excavation data on the mobile device application.
- eResearch tools required for searching excavation data.
- eResearch tools required for the analysis of excavation data.
- The working group should also discuss and provide direction on any other aspect of the FAIMS project relevant to the capture, management or analysis of excavation data.

2.3.2 Scope

While the requirements for data attributes may vary from project to project, we hope that this working group may agree on some minimum requirements for excavation that may be widely applicable. FAIMS data capture application development is focused on the digital creation of data in the field on mobile devices (e.g., Android phones and tablets), but the project also envisions the development of companion desktop software. Applications will be modular and extensible. To ensure core functionality and encourage the creation of compatible datasets across as wide a range of projects as possible, we ask this group to focus on a minimum threshold of attributes that should be recorded under all (or almost all) circumstances during excavation.

2.3.3 Topics to Consider

- What are the 10 most important components of excavation record? Why?
- What is the desired granularity of your data? What is the achievable granularity of your data?

- What is the smallest spatial or conceptual entity that you currently record? That you would like to be able to record digitally?
- What data do you need to quantify (e.g., deposit composition is 3% gravel and 3% sand)?
- What in your data is a qualitative observation (e.g., the matrix is highly compacted)?
- What in your data is a measured, objective observation (qualitative or quantitative, e.g., "this stratigraphic unit yielded 10kg of pottery", "the color of the soil in the matrix is Munsel 7YR")?
- What in your data is interpretive (e.g., "this wall is late prehistoric")?
- Is there any "data" that doesn't neatly fit in the above categories?
- When discussing the above questions, did the group use the same vocabulary to refer to the same concepts? What similar concepts were mapped to different words? What different concepts were mapped to the same words?
- Is there a significant difference in recording between "simple" single-phase excavations and "complex" multi-layer excavations, enough so that we need separate systems for each?

2.3.4 Expected Outcomes

- List of essential attributes required for recording excavation data in the field.
- List of functionality requirements for recording excavation data on a mobile device application.
- List of functionality requirements for searching excavation data on the FAIMS portal.
- Suggestions regarding workflow or approach - what would the process of recording a stratigraphic unit in the field look like?
- Suggestions for analytical tools to be included within or developed by the FAIMS portal.
- Should controlled vocabularies be used throughout (and if so, global or local)?

2.3.5 Participants' Reflections

The Excavation Working Group met twice on 17 and 18 August. Membership fluctuated between 4- 5 participants. This group was extremely small, which was interesting in itself - perhaps this says something about:

- the level of interest people have in excavation
- the fact that, in Australia, the vast majority of archaeology, especially in the consulting context, involves surface survey
- while a lot of people participate in excavations, far fewer of those people are involved in project management or the research design of the excavation, so possibly not many people felt confident in contributing to this topic as perhaps they didn't think they adequately "experienced" or "qualified" enough to do so.

When we considered this topic, we used "excavation" to include any kind of excavation, including shovel test pits (ie with minimal control and detail), open area excavation, deep excavations, shallow excavations, hand excavation, equipment excavation. There are probably other types of excavation that we haven't listed here - one that immediately springs to mind is excavation in underwater contexts, however as we didn't have a maritime archaeologist in our group we're not sure we've included everything here but we are hopefully that we've at least captured some essentials!

2.3.6 Essential Attributes for Excavation Recording

In our experiences, there were certain things that we agreed should always be recorded during an excavation in the field, though not all of the below will be recorded depending on the method of excavation (and note these are not in any particular order, though they sort of go from broad to more specific):

- Site name and or unique identifier code
- Site location
- Date of that particular component of the excavation (ie a specific day of that bit of excavation, not just a date range for the more general field season)
- Name of the specific person undertaking the excavation (and potentially also the person taking notes/sieving/sorting)
- Details of the datum/TBM (see comments below with respect to what this is)
- Method of excavation (eg hand, shovel, machine etc)
- Sieve size(s) used (if any)
- Provenience of any plotted artefacts/samples etc (ie an X,Y and Z coordinate) (and obviously this may not be possible in certain types of excavations, such as shovel test pits)
- Sediment matrix, which should include the texture, munsell colour, pH, inclusions (and ideally the same person should be recording this for the whole excavation to ensure consistency but obviously that's not feasible for many situations - this is why it is probably important to ensure that it is clear on the recording form if someone other than the named excavator is doing the pH testing or sediment descriptions etc)
- Depths of the excavation units (or whatever they are being called, eg spits, see below) - it should also be noted somewhere whether arbitrary or cultural stratigraphy is being followed
- Start and end levels of each excavation unit (though this might be recorded automatically into the total station/dumpy etc), most people tend to keep it in written format somewhere as well, as a built-in redundancy measure in case of equipment failure
- Volume and/or weight of sediment removed from each excavation unit (or whatever they are being called)
- Codes/numbers of photographs taken that relate to that particular component of the excavation)
- Artefacts (usually just the general categories that were encountered, eg charcoal, pottery, glass, lithics), associated with some sort of bag label information/inventory so it can be cross-checked
- Sketch of the feature/unit etc, potentially including plan and/or section views)
- Descriptive comments about the excavation (which should be broken into two sections at least, a description-theoretically the "factual" observations- and then the interpretation-which is the more fanciful, subjective thoughts).

2.3.7 Other Needed Information

Other than the "excavation forms" that most people tend to use there are also typically other items that are "created" during an excavation that need to be archived/managed. The following list is not intended to be all inclusive, but gives an idea of the kinds of things we were thinking:

- photographs
- profile/section drawings (to scale)

- sketches/mud maps/plans (not to scale)
- site plant (to scale)
- total station data (or equivalent, eg dumpy, written list of coordinates)
- inventory of photographs taken
- inventory of recovered materials (bags)

2.3.8 Additional Goodies that we'd like

- really rugged tool to handle Australian extremes (dirt, rain, heat, sun glare etc)
- have a LONG battery life
- it has to be affordable for both small and large user groups (eg sole traders on restricted budgets who can't afford large buy in or set up costs)
- an in-built low and high powered microscope
- a built-in total station
- must have highly accurate and precise positioning ability (ground-based radiowave GPS not satellite technology currently being developed might be the way to go here)
- digital voice recorder

2.3.9 Summary

In their report, the Excavation Group reiterated some of the suggestions of the Artefact Group, and added the following concerns:

- excavation unit needs to be flexible and adjustable (shovel test to stratigraphy)
- spatial recording (simultaneous recording of xyz coordinates for single finds and 3D excavation units)
- incorporate vector sketches and raster imagery (photographs, site plans, drawings, etc.)
- allow audio-recording of voice notes (not necessarily with text-to-speech capability)
- have a way to indicate the "level of confidence" (for want of a better phrase) that the user/data entry person has with respect to any of the fields.
- have drop down tables/calendars/autofill functions for dealing with repetitive data such as the site name, excavator name etc.
- indicate that there is potentially some "fuzziness" about how reliable information (e.g. about Munsell colors) is. Perhaps the most obvious "fuzzy" attribute would be the descriptive "interpretation" field.

2.4 Federation Working Group

2.4.1 Scope

FAIMS seeks to explore the possibilities of data federation in archaeology. Federation is defined as automated data interoperability requiring a minimum of manual intervention as opposed to, e.g., loose-coupling, which according to Kansa and Bissel's (2010) definition relies on extensive human

interpretation to reconcile datasets. The degree of federation depends on the level to which semantic and technical interoperability can be achieved.

This panel is about semantic interoperability, and resources that should be made interoperable. We wish to focus on the problem of producing semantically comparable data. The argument has been made that every project has its own research agenda, requiring a unique record customised to its goals. If that is the case, is it counterproductive to develop a master record-keeping protocol?

If minimum record-keeping protocols are desirable, do you think that we can agree on, e.g., 10 things that need to be recorded about any excavation stratigraphic unit, surface survey unit, artefact, etc.?

2.4.2 Topics to Consider

- Are shared record-keeping standards, as well as technical data standards, necessary for the success of FAIMS project?
- If the answer to the above question is affirmative, can FAIMS stakeholders in this room agree on minimum record-keeping protocols?
- Is federation possible without shared record-keeping protocols?
- Is federation a worthwhile goal?
- Is the best achievable goal the loose coupling described by Kansa and Bissel (2010, 44)?
- To what extent is it possible and desirable to separate quantifiable ("objective") and interpretive ("subjective") data? Is there a fundamental difference between a date of a lithic and a length of a lithic?
- How rigid should be the line between subjective and objective recording versus measurement?
- How can we ensure the capture of data that allows re-interpretation?
- If we do have standards, are they regional, chronological, global?
- Can there be a shared set of standards between cultural heritage and research, or archaeology and related disciplines?
- What existing resources need to be supported? Prioritize please resources:
  - Online repositories
  - Field Databases
  - Reference collections
  - Tools
- How do we get these resources (state registries, and others) on board?

2.4.3 Expected Outcomes

FAIMS asks this group to produce answers to all the questions above, and in particular to produce a definitive statement on:

- What level of federation is desirable by FAIMS stakeholders?
- What type of standards should FAIMS embrace?
- What existing resources need to be supported by FAIMS, listed by type and priority?

2.4.4 Participants' Reflections

The Federation Working Group met twice on 17 and 18 August. Membership fluctuated between 5 to 6 participants, mostly developers and technical managers.

2.4.5 Summary

It is critical that all developed recording applications be modular. The applications cannot hard code any records or means of achieving those records. While there is no opposition to an AR application, little thought has been devoted to how it can be used. Overall it is difficult for the archaeological community to break from normal "everything's a paper chart" methodologies. We must not feel constrained by existing systems and approaches and technologies. We need to think to future developments like Google Glass and Augmented Reality devices, at least for inspiration. VeRSI and Brian propose the development and demonstration of a "killer app" for this purpose.

Given modular data requirements, the mobile data standard established was essentially that of an XML templating engine. The application should be able to generate input systems based on XML schema. Furthermore the schema should provide for the ability to interface with external data-collection devices, such that it can collate and absorb the outputs of the external devices.

Non-relational data architectures could serve the community extremely well, as the current relational structure require hacks to get around the limitations of relational databases, hobbling data collection and analysis efforts.

Applications should collect an internal-audit log (engaging cameras, acceleration, position, gait analysis, etc.) to indicate who the data-collector is and as much of their mental/physical state as possible. The program should also have a highly accurate place-time record, as well as linking that to visual records of the site and any appropriate events.

Much User-Centred Design and Scenario Testing will be needed to refine these requirements.

ADS, tDAR, OpenContext, Heurist, and Kora should be considered for the first round of interoperability.

2.5 Sensitive Data Working Group

2.5.1 Purpose

Describe the problem and formulate testable requirements for the various issues surrounding "sensitive" data. What legal, moral, ethical, confidential, political, and social problems can occur with the collection, aggregation, display, and export of the data collected by mobile devices that FAIMS is developing? What do bad actors look for in archaeological data? How can we protect against them?

The working group should also discuss and provide direction on any other aspect of the FAIMS project relevant to the capture, management, or publication of sensitive data.

2.5.2 Scope

FAIMS is aware that many categories of archaeological data collected in the field may be of sensitive, personal and confidential nature. FAIMS is asking this group to define the kinds of data that may be considered sensitive, and to articulate approaches to their recording. FAIMS also requires guidelines and strategies on managing, archiving, sharing and digitally publishing sensitive data.

2.5.3 Topics to Consider

- What kind of sensitive data will we be capturing with mobile devices?
  - What specific components of the data are sensitive? Who should it be protected from?
  - What component of the data should be shared? What level of sharing control should it have in the best of all possible worlds? What could go wrong with that level of sharing control?
  - What makes this data sensitive beyond its individual components?
- What sensitive data is routinely recorded?
- What edge cases should we cover?
- How have previous systems failed you?
- How will people want to interact with this system?
  - Do we want to provide all these methods?
  - What other methods do we want to provide?
- What are the different use cases for adding data, editing data, accessing data, and searching data?
  - What preparatory actions does the system need to support?
- How can the system promote and require external analysts to maintain adequate protection of the data?
- Will people actually use this feature, or will they work around it?
  - How would an analyst actually interact with the data?
  - How will bad actors want to interact with the system?
  - Do we want to presume good faith?
  - Are all actions allowed unless forbidden, or are all actions forbidden unless allowed?
- What general use cases for classification are there?
- What audit capabilities are necessary?
- What aggregations are necessary to prepare data for bulk export?
  - Do these aggregations protect the sensitive data?
  - How could a bad actor de-anonymize these aggregations?
  - Are the detail trade-offs worth the level of protection afforded?
- Is this just complying with statutory regulations without providing true anonymity?
- How would sensitive data be useful to other archaeologists in the field?
  - How would they authenticate themselves?
  - How would they access it?
  - How would bad actors use this as an attack vector?
- How have you used other sensitive data? How did you access it? What sucked about the access method?
- Is there any consistency in rules about sensitive data? What's the most ironic rule "trap" you've experienced in this regard?

2.5.4 Expected Outcomes

- List of (say) 10 types of sensitive data, with a brief statement about what makes each sensitive.
- List of requirements for the recording of sensitive data in the field.
- List of minimum security and publication requirements for archiving and publishing sensitive data.
- List of functionality requirements for searching, publishing and sharing sensitive data on the FAIMS portal.
- Suggestions for tools to be included or developed for the FAIMS portal useful for analysing sensitive data.

2.5.5 Participants' Reflections

The Sensitive Data Working Group met on 18 August. Membership fluctuated between 6 to 7 participants, mostly archaeologists and developers.

2.5.5.1 Key Messages

- Sensitive data is a complicated issue and it is very difficult to be prescriptive (without knowing the actual application proposed) BUT, the solutions are relatively simple to implement technically (generally speaking)
- The right solution can be complicated but can be resolved with good design (requires good consultation)

2.5.6 Individual points to consider in any database with sensitive data:

- Identify what is the worst that can happen? Consider the worst case scenario- someone hiding data as an abuse of system for example – and then think how to mitigate it.
- Is the extra cost and complication worth it? (when considering every potential sensitive data issue)
- Complexity to meet level of risk?
- What is sensitive? (How do you define it?)
- Who establishes sensitivity? Who can modify it once it has been set?
- Who takes responsibility for a given state of sensitivity?
- All agreed: people determine sensitivity, NOT machines
- Should sensitivity be binary? Or fuzzy?
- Can't be binary as in yes or no, because it depends on the intended use
- But can be binary as in whether someone needs to be consulted (for example)
- If consultation is required, it creates an issue of maintaining the currency of who should be consulted?
- Good outcomes for sensitivity also require education for both the supplier and user of data about what all the possible sensitivities options are
- Rather than defining what data is sensitive, it could instead be better to ensure what the protocol of dealing with sensitive data is
- Thresholds of sensitivity – what is the minimum amount of data that is sensitive? eg. Aboriginal data about burials/religious+pictures/locations are almost always sensitive
- Is it possible to work out a standard process? Probably can't be prescriptive?

- Issues of sensitivity are just about whether data is hidden or not...the whole system needs to embrace it
- One supporting mechanism: maintain log file of managing sensitivity – history of sensitivity decisions.
- Or... log all activity of each user: what they download and when
- Log file would create opportunity to then inform a user that once something is made sensitive (for example if it is downloaded when open, but subsequently made sensitive)
- Important that people are therefore informed that once data is retrieved from the system they take on obligation to manage appropriately – this might be part of code of ethics the user abides when they sign into the system, which would include informing them that data sensitivity may change over time, but their obligations to manage appropriately remain?
- Projects may already have defined agreements in place about data sensitivity (eg. academic projects and ethics approvals) – add ability to upload and attach project protocols to data
- How is sensitive data handled in queries of data: should it be made invisible (but letting the user know whether those omissions take place, and whether that should say what it is or how much)
- Data licensing as another solution: Mukurtu (Indigenous licensing scheme, http://www.mukurtu.org/)- could be defined by supplier, and/or agreement to abide by code of ethics on entering the system
- Don't want to over complicate entering of every piece if data – balance usability
- Ability of people to be able to lodge sensitivity issues identified by the user is required

2.5.7 Summary

In summary, the Sensitive Data Working Group agreed that technical solutions to the sensitive data are fairly straightforward if properly consulted. It explored different ways of handling Sensitive data beyond merely restricting access, hiding or ambiguating location. One of its main outcomes focused on facilitating access to meaningful derived information, while hiding actual sensitive data.

2.6 Survey Working Group

2.6.1 Purpose

Formulate the minimum requirements for:

- Pedestrian surface survey data attributes to be deployed across the system (captured through mobile device applications and used to facilitate the production of comparable datasets).
- Functionality of the capture of survey data on the mobile device application.
- eResearch tools required for searching survey data.
- eResearch tools required for the analysis of survey data.

The working group should also discuss and provide direction on any other aspect of the FAIMS project relevant to the capture, management or analysis of pedestrian surface survey data.

2.6.2 Scope

While the requirements for data attributes may vary from project to project, we hope that this working group may agree on some minimum requirements for pedestrian surface survey that may be widely applicable.

FAIMS data capture application development is focused on the digital creation of data in the field on mobile devices (e.g., Andriod phones and tablets), but the project also envisions the development of companion desktop software. Applications will be modular and extensible. To ensure core functionality and encourage the creation of compatible datasets across as wide a range of projects as possible,

We ask this group to focus on a minimum threshold of attributes that should be recorded under all (or almost all) circumstances during surface survey.

2.6.3 Topics to Consider

- What are the 10 most important components of survey record? Why?
- What is the desired granularity of your data? What is the smallest spatial or conceptual entity that you would like to be able to record digitally?
- What is the achievable granularity of your data? For example, do you only record site-level data, or do you record blanket artefact densities, including absences of artefacts? What is the smallest spatial or conceptual entity that you record?
- What data do you need to quantify (e.g., "this unit has 70% visibility")?
- What in your data is a qualitative observation (e.g., "this field is difficult to cross")?
- What in your data is a measured, objective observation (qualitative or quantitative, e.g., "this total pickup yielded 10kg of pottery", "the color of this ceramic fabric is Munsel 7YR")?
- What in your data is interpretive (e.g., "the pottery in this unit dates to the Roman era")?
- Is there any "data" that doesn't neatly fit in the above categories?
- When discussing the above questions, did the group use the same vocabulary to refer to the same concepts? What similar concepts were mapped to different words? What different concepts were mapped to the same words?

2.6.4 Expected Outcomes

- List of essential attributes required for recording surface survey units in the field.
- List of functionality requirements for recording surface survey unit data on a mobile device application.
- List of functionality requirements for searching surface survey unit data on the FAIMS portal.
- Suggestions regarding workflow or approach - what would the process of recording a survey unit in the field look like?
- Suggestions for analytical tools to be included within or developed by the FAIMS portal.
- Should controlled vocabularies be used throughout (and if so, global or local)?

2.6.5 Participants' Reflections

The Survey Working Group met twice on 17 and 18 August. Membership fluctuated between 10 - 12 participants.

2.6.5.1 Key Messages

- There was general agreement that digital data capture of survey data was essential to future work.
- After description and discussion of a wide variety of survey methodologies including maritime and terrestrial surveys for research and consultancy it was generally agreed that basic commonalities to all archaeological survey that could be supported by a flexible and highly customisable mobile tool.
- While general concepts are shared there is no standard methodology for survey. Each survey methodology should be internally consistent and should be fully documented.

2.6.5.2 Survey Record Structure

Six potential levels of the survey record were identified:

- Project - The identity of the project or job with possible additional identifiers for season, session or other subdivisions. FAIMS will require some sort of project registry to issue unique project codes for use in unit identifiers (see later).
- Survey unit - The primary administrative subdivision of the survey. This can take many forms including transect, arbitrary area, area identified from remote sensing, or an existing area such as a field or land parcel. These will always be expressed as a polygon. Generally survey units are likely to have the same survey conditions, such as visibility, and they are often defined by or confined to such areas, but the tool must be flexible enough to deal with changes of condition or survey methodology within a unit.
- Concentration - An area, often but not necessarily within a survey unit, that typically defines a concentration of artefacts.
- Sample Area - An area, often but not necessarily within a survey unit and/or a concentration that typically defines an arbitrary metric area for quantitative sampling.
- Artefact - Generally a point, not necessarily within a survey unit and/or a concentration that defines the location of an isolated artefact or group of artefacts.
- Survey track(s) - The path(s) taken during survey by surveyors recorded by GNSS. This must include anything from initial scouting and reconnaissance of the survey area to the systematic coverage of an area by a team of walkers.
- Untyped object/note - Some information simply does not belong in an identifiable unit (including general notes and interpretation), this unit type acts as an agnostic container for this type of material.

2.6.5.3 Survey Data Identifiers

All data recorded should be attached to an ADNS (archaeological domain name service) number and date-stamped, allowing versions of the data on the lab server (for example) to be synchronised with data on individual mobile devices without risk of overwriting more recent changes. ADNS offers a methodology for providing unique identifiers for all entities (projects, sites, survey units, phases, levels, artefacts, bags, samples etc.) which are studied by various projects in the FAIMS universe.

- Domain ID - Project ID - Device ID - Entity ID - subdivision - subdivision ….
- A concern was voiced that machine-generated IDs will be long and tedious to record, and even more difficult to remember or find physically, where people are used to looking for "Find 57" or "Trench A".
- The compromise suggested is to allow for one (or multiple) user-defined (that is allocated internally by the project) labels to be attached to recorded units or artefacts, most probably at accessioning into a field lab. This will allow the team to create an additional project-specific alpha-numeric code with embedded implied attributes (such as P1, P2 … for pottery, B1, B2 … for bone, T1, T2 … for trenches etc.) which could be used internally and by people studying the material.
- The system should ensure uniqueness of these labels on a per-project basis so that they can be used as human-friendly labelling if desired within a project context but would still be globally unique (ie. project + human-friendly label = unequivocal identification within the FAIMS universe).

2.6.6 Survey Data Management

Important concepts to be considered when scoping and evaluating tool(s).

Tiered data management - There will be typically up to 5 layers of data capture/storage/management devices (a similar structure applies to excavation data):

- Individual mobile devices such as PDA, phone or tablet -
  - Ideally each team member will have a mobile device capable of recording survey track and observations, including photographs (internal camera or derived from a separate camera)
  - Most likely type for many surveys is a GPS-enabled Android tablet with camera, but could also be mobile phone.
  - Device will minimally display map backgrounds, data entry forms and functions for managing peripheral data capture.
  - Captured data should remain on the device even if it is synchronised to level 2, to provide redundant storage in case of accident
- In-field server (optional)
  - This is an optional level which is only likely to be implemented in larger and better resourced teams. It would probably be a portable wireless backpack micro-server
  - In a field survey team this might be carried by the team leader and used to consolidate data wirelessly and continuously from individual team members' mobile devices
  - Team leader would probably access it via a mobile device the same as other members of the team, but would have greater access levels/functions
  - It could allow continuous monitoring of differences between the records of different team members to identify errors on the spot eg. mis-recording of visibility or vegetation.
  - It could allow coordination of survey unit numbers across collectors and accept input streams such as a total station or specialised data collectors
  - Provides redundant storage of data on individual collections which are more likely to bite the dust eg. due to being dropped
- Field base/Laboratory server (optional)
  - Typically a desktop or powerful notebook in the field base, lab or dig house.
  - Used to consolidate data from individual data capture devices for a project.

  - Typically handles only one project, but may have more than one if there are simultaneous survey and excavation projects in the same area.
  - Most probably networked except on small projects, may or may not be internet connected.
  - Needs to run software which can synchronise data with field capture devices, manipulate, generate simple analytical results and backup.
  - Will be used to download spatial and other data (if internet connected), revise recording templates, create packages of data (map/spatial data plus recorded data from prior fieldwork) for download to mobile devices for use in the field
- Research server
  - Typically at the home institution/office of the project, often institutional infrastructure
  - For single-project teams could also be the laboratory server returned to base
  - May handle data from multiple projects
  - Nearly always internet connected and will generally provide feeds and other web services to make information publicly available (or to a specified community)
  - Can do everything the laboratory server can do, plus probably more analysis, mapping/visualisation and backup of data.
  - Can generate repository packages for submission to level 5
- Repository
  - Often an institutional repository such as University Library or a public repository such as ADS.
  - Services many projects
  - Provides indefinite sustainable preservation of the project data
  - Will probably require preparation of deposit packages rather than synchronising directly with level 4
  - May harvest data through standards such has OAI-PMH
  - May expose data as published web pages or feeds

Synchronisation of data in layers 1 - 3 (and possibly 4) should be transparent and seamless; it should not matter if the data for an entity exists in more than one place, only the most recent version will be synchronised.

All changes should be stored in an audio file, allowing a complete history to be generated (and/or the data to be rolled back if needed - development of a rolled-back data service is beyond the scope of this project).

2.6.7 Survey Module Requirements

- Device redundancy - There must be alternate power sources and recharge options for mobile devices. There must be replacement/alternative devices in case of failure or breakage (run on a variety of devices, including laptops as an inconvenient last resort)
- Customisable interface - The tool must allow the user to customise the interface to the requirements of the methodology being used, without recourse to programming. This must include the ability to hide any unwanted fields.
- Flexibility - To cope with the unpredictable nature of survey the tool must be able to move seamlessly between the levels of recording listed above so that any element of the survey record can be generated or modified when required.

- Redundant data - Collection of data such as device ID, time stamping, location even if not central to the survey data record is important in resolving problems and errors. It should be collected and maintained.
- Contextual data - As part of the survey record the tool should allow the collection of contextual data such as information on weather conditions, general visibility etc.
- Peripheral data - In addition to the highly structured survey record the tool should easily allow the collection of peripheral data such as nearby prominent features, views and real estate signs to act as supporting information, checks and to facilitate relocation independent of the GNSS record.
- Individual records - As the performance of individual recorders can have a significant impact on the primary survey record, the identity of each recorder and information about expertise, motivation and performance should be included in the record where the survey methodology supports this.
- Synchronization must be invisible, instantaneous, transparent and seamless, and must allow versioning.

2.6.8 Summary

Survey Working Group discussed the recording needs of surface survey practitioners from the point of view of survey data management, workflow organization and data structures. Technical and conceptual requirements for the potential survey modules were discussed as well as general groups of attributes that form the basis of survey recording. The Survey Working Group was one of the best attended discussions during the Workshop and brought forth a detailed and specific vision of survey recording system.

2.7 Sustainability Working Group

2.7.1 Purpose and Scope

FAIMS seeks to develop an infrastructure that would shepherd archaeological data from its digital creation to processing, archiving and publication. We want to establish the core of an open –source ecosystem for archaeological data management. Such infrastructure will combine tools and repositories, and exceed many of the existing successful digital resources in scope. Given the idiosyncratic data that archaeologists collect, the mobile tools will need to be continually developed and customized to fit the needs of new archaeological objectives. Repositories and databases behind them will require and ongoing maintenance. Maintaining the functionality of the entire infrastructure and sustaining the development of its individual components will be a major challenge. FAIMS would like this group to consider the challenges of long term maintenance and development of the FAIMS infrastructure and its components. FAIMS asks this group to articulate approaches and models to sustain the project.

2.7.2 Topics to Consider

- How do you usually acquire the tools and software for your fieldwork?
- How much does it cost your project to manage your data?
- How often do you encounter incompatible datasets that hinder your research?

- Do you think you could invest more in the tools and less in personnel?
- What are the main differences between academic and contract archaeologists in how they invest in data management?
- Should individual components be funded separately and through different strategies?
  - Consider these strategies:
  - customization fee for mobile apps
  - curation fee for storage or publication
  - subscription to FAIMS portal
  - access fee for use various FAIMS components
  - grant support
- Are you willing to pay for individual components separately? If so, which ones?
- What models of funding are you familiar with or can you identify for an open source archaeological data management eco-system?

2.7.3 Expected Outcomes

- Identify possible funding models for FAIMS project.
- Identify possible funding models for individual FAIMS components.
- What in your opinion is the best strategy for long-term sustainability for FAIMS digital tools?
- Identify "reasonable costs" for individual fees and funding strategies.
- Identify best strategies for different groups of stakeholders to maintain the FAIMS tools.

2.7.4 Ten funding approaches for consideration

- Consider this list a starting point - eliminate, add, and qualify as you see fit.
- Software as a service - a recurring fee for mobile application and other software.
- A customization fee for field recording systems.
- A curation fee for online archiving and publication.
- An access fee for use various FAIMS online components.
- A subscription fee for the FAIMS portal (i.e., a fee charged to libraries, universities, or researchers to access the datasets).
- Pay-per-view for downloads.
- Subscription or hourly rate fees for help-desk support.
- Grant support sought by FAIMS (i.e., applications for further infrastructure grants).
- Expectation or requirement that individual (academic?) users of FAIMS apply for grant support (e.g., on ARC or NSF grants).
- Corporate and/or other institutional sponsorship.

2.7.5 Participants' Reflections

The Sustainability Working Group met once on 18 August. Five participants contributed.

- The problem of sustainability is made complex by open architecture and component structure: different sustainability required for each
- If software is given away free (as requirement of NeCTAR) that limits funding sources

- What is it we are trying to sustain: software or user community?
  - If the project is ultimately about people and behaviour, then sustainability would be about how people change practise over the long term- eg. archiving data, submitting electronically, …
  - In which case it is important to set up the environment properly is crucial
  - But…brief is about software tool/system…so how is this sustainable?
  - How do we discuss sustainability when we are not clear about what the system actually is?
- Sustainability of mobile application + plumbing for system (NOT a repository)
  - You can still consider the sustainability of people- user - community
  - Not many archaeological software tools have proved profitable or sustainable
  - Sustainability of a tool is a bounded question: don't know what it will cost?
  - The plumbing part of system can be on a free model, but mobile application based on it could be for a fee.
  - If the mobile tool is built around providing information to support recording, would it be worth charging for it, and if it does, then is it sustainable
- Fee for service for storage at universities (or University-based cloud) is one mechanism for revenue
  - The project does have everyone talking, so a community has developed, which creates the human part of the federation
- If FAIMS has only ability to build the first part of the system, then the trick will be to ensure that the first component is worth major players (Mining companies or contract archaeologists) who would buy it.
  - If mining companies invest in the system, will archs support it? Some would -some won't.
  - Link with research if the tool is so CRM focused- solution- more research on CRM
- Benefit to research includes the access to grey literature- we don't know what the full research potential of the grey literature is: this is significant in the UK & US
  - Tools for investigating the grey literature: querying etc, is a valuable and potentially chargeable tool
  - Grey literature also important for data quality audit; ensure data quality; having work out on internet is important for sustainability of the discipline
- Another market is land-use planning: better supporting information- but this could also just create more work: if access is streamlined, CHM will utilise and pay for it
  - CHM business levering it under different business models: focus on reviewing data rather than locating it
  - Lot of CHM businesses perhaps haven't got a good mobile tool solution, in which case there is a user base and hence a revenue base
- A possibility is a single mobile tool for collecting data and outputting to any Australian regulatory standard- if this is achieved it would possibly generate a revenue base
  - CHM companies will buy a product that saves them time
  - Need to know how many CHM companies there are in Australia in order to work out a potential revenue base
  - CHM companies would pay for a good service- support from project
  - License for API? Or GPL license- paying for a modifying the API
  - CHM companies paying for cloud storage of data- possible revenue stream
- Start small…one jurisdiction like say ACT- pilot project linked to Sally Brockwell's Act Aboriginal heritage project
  - Training? Is this a revenue base

- There is a big interest from academic researchers, who can fund development under LIEF

2.7.6 Summary

The Sustainability group suggested a wide range of approaches to funding the FAIMS project, combining user-driven approaches to continued seeking of grant funds. One of the key messages was to target an important group of paying stakeholders (mining companies or contract archaeologists) and progress in small steps and closely in sync with the user needs. Large infrastructure with many components poses additional challenges for different business and sustainability plans.

3 Working Group Conclusions

The principal purpose of this report was to summarize the organization and outcomes of FAIMS Working Groups, which discussed best practice(s) in record-keeping and data and metadata standards in various subfields of archaeology at the FAIMS Stocktaking Workshop, held at UNSW in August 2012.

Each Working Group produced a slightly different output. Outputs ranged from highly detailed attribute lists to more generic yet increasingly valuable outline of best practices. Working Group performance exceeded our expectations, especially in having managed to articulate answers to a complex set of questions in short amount of time within a large group of participants.

Most Working Groups found enticing the idea of a digital module that would encourage and promote a common minimum standard. While they acknowledged that archaeological practice can be multi-faceted and produce a variety of standards, they indicated that there is common ground across this range of standards.

Overall, the high-level features seen as important by most participants included:

- All developed systems and applications need to be modular
- Rather than an extensive core data collection system we need a library of data collection tools that are based on existing published standards and existing paper recording systems. Each of these needs be modifiable, extensible and shareable. To the extent possible, these tools need to be mapped against one another (possible because of similarities in vocabularies and ontologies).
- Controlled vocabularies with aliases need to be implemented throughout these systems
- Field–available reference collections and spatially aware applications are in high demand and should be provided. Support, training, and tool tips need to exist for the designed tool features.
- Unique identifiers for each piece of data collected need to be generated with human-readable aliases.
- While we recognize that existing state registry requirements need to be taken into consideration, we have been encouraged by the SC and others to design comprehensive high quality recording systems and promote these to the state registries, mining companies and other stakeholders.
- Systems developed need to accommodate existing and future data collecting devices (total stations, dGPS, digital calipers, etc.). Devices used in field and laboratory need to sync seamlessly and transparently.
- Metadata needs to be comprehensively collected and managed. As much as possible it needs to be collected automatically and derived from existing system components (eg. Login, Timestamps, reference collections).
- In all cases, data creation needs to follow workflows with options to adjust for variations in workflow.

4 Appendix 1: Guidelines for Working Groups Members

- Group Leader
  - The leader will guide discussion, moderate the group, and try to focus the group on defining the problem instead of staking out "solutions."
  - The leader will report on the group progress and observations at the end of each Workshop day.
  - The leader will try to get answers to the questions and objectives in the supplied info sheet.
- Minute taker
  - The minute taker will take notes on the discussion, e-mail them to the group mailing list, and assist the group leader in their presentation of the problem and discussion.
  - The minute taker will take especial note of any jargon differences. One of the goals of the stocktaking is to understand the various vocabularies and necessary Galisonian Trading Zones that have emerged in the data of archaeology. There needs to be no effort to resolve these vocabulary problems, only to note that they exist, to remove and highlight potential future difficulties.
- Group Participants
  - We suggest that individual groups have no more than twelve participating people seated in specific groups. If a group finds itself blessed with an overabundance of members, we urge them to split into sub-groups to consider different aspects of the problem.
  - The FAIMS project exists to federate archaeological data. In order to allow different databases to interact, we must understand the different aspects of data curation and archaeology that influence the problem. These groups will help to articulate those aspects. We expect exploration of the problems, stories about the problem or possible futures, or testable and falsifiable statements that allow us to test if our planned solutions actually solve the problem.
  - We also will be using this Workshop to generate requirements for mobile application for fieldwork. In exactly the same way as data federation, we are not looking for "solutions" but a complete and comprehensive description of the problem and the reality-that-is.
  - We have defined a number of working groups. We would be absolutely delighted for a working group to see an aspect of the problem that we had not anticipated and splinter off into a new domain to explore this previously unforeseen problem.