**\heading FAIMS Digital Data Survey Report (Old to new)**

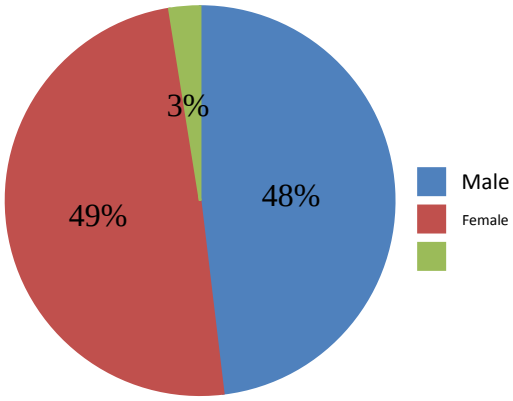
This report presents the results of the FAIMS Digital Data survey, undertaken in July-August 2012. The survey was conceived as an initial and an extensive stocktaking exercise involving FAIMS stakeholders and the broader archaeological community in Australia and overseas. Its purpose was to gather information on the use of technology by archaeologists before the designing of a new archaeological information management infrastructure. Through this collection of information we have implemented as fully as possible the tenet that "stocktaking activities are the precondition to designing a successful tool" \cite{Jones2005}. Although our approach and execution of the survey itself was far from perfect in the details, it substantially improved our development plans and offers a model for similar projects.

\heading **Digital Data Survey**

Digital Data Survey was the first task that FAIMS team undertook after the launch of the project on 5 June 2012. The survey was aimed at academic, consulting, and government archaeologists or related practitioners who produce or use archaeological or cultural heritage data, including students. The focus of the survey was to describe current patterns of information technology use and define key obstacles to increased use of technology within the archaeological community. While its goal was similar, its scope was smaller and design less complex than the "Strategies for Digital Data Survey" organized by the ADS in Britain to assess the creation, archiving, use and re-use of digital data in archaeology \cite{Condron1999}. Our survey also complements the demographic analyses and vocational training survey of the archaeological community in the UK \cite{Aitchison2003, Aitchison2004} and US \cite{Zeder1997} and the survey of archaeological profession in Australia with implications for teaching and learning \cite{Ulm2005a}.

\heading **Survey response rate**

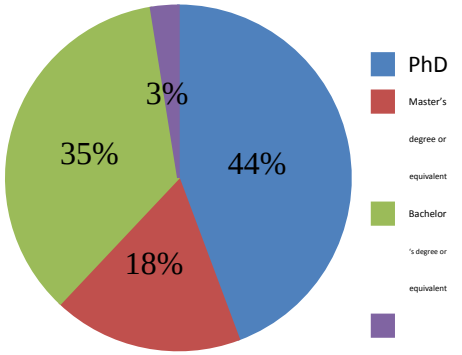
The survey was sent out to the FAIMS email list which at the time included some 120 Australian stakeholder institutions and individuals as well as 40 overseas participants. Contacts on the FAIMS email list included members of Australian archaeological community as well as IT-oriented archaeological networks (e.g. members of the Digital Data Interest Group at the Society for American Archaeology, or affiliates of the Archaeology Data Service in the UK, etc.). AAA and OzArch lists were also used for survey distribution as were other social media. We encouraged our addressees to forward the survey to other colleagues. The survey was open for two weeks during late July-early August 2012 and reminders were sent on Fridays to encourage responses. By 12 August 2012 we collected 128 responses; 81 of them from Australians (3 of those were duplicates), the remaining 39 from overseas participants. The response rate was high thanks to the forwarding mechanism within the surveyed community. We would like to take this opportunity to again thank the participants their time and effort. We present an analysis of the Australian responses here.

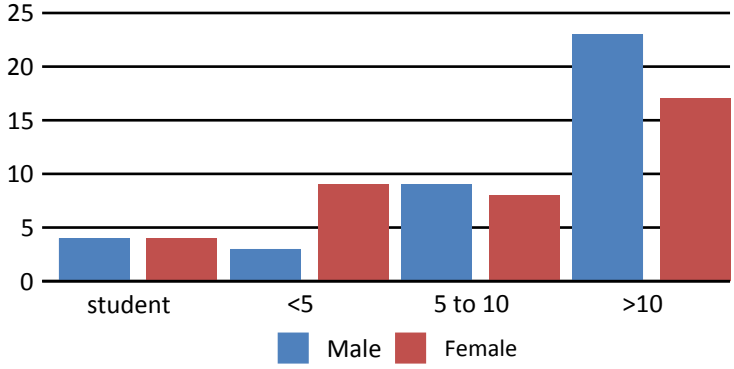



**\heading Profile of Respondents**

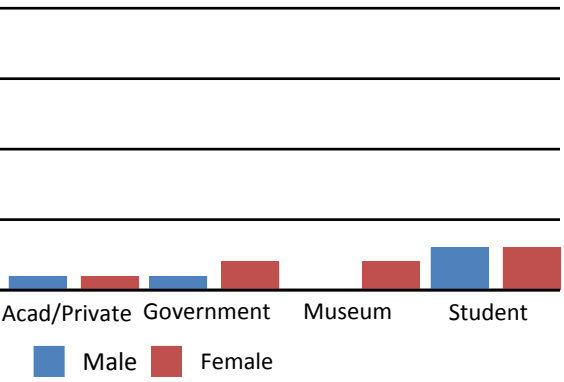
The demographic composition of the FAIMS survey sample is consistent with the profile of the Australian archaeological community presented by Ulm, Nichols and Dalley (2005). Men and women are almost equally represented in the survey, with a ratio of 49:48% (with two participants abstaining).

In terms of geography, most Australian states were well represented in the survey; only the Northern territory and ACT failed to produce any respondents (14 participants did not declare state of origin, which leaves a large margin of error for any geographic analysis). When assessing education, the methodology in the survey was admittedly flawed. The list of options included only Bachelor's, Master's, and Ph.D. degrees, omitting Honours. Overall, 37% of participants declared a Bachelor's degree or equivalent, 3% Honours (volunteered), 17% Master's, and 43% Ph.D. The final results, acknowledging the under-representation of Honours graduates (who likely selected "Bachelor's degree"), is not dissimilar from the qualifications reported in Ulm, Nichols and Dalley (2005, 17, fig.7), where Honours degrees and doctorates form the two largest groups.




The time spent working in the archaeological profession (rather than age) was considered an important factor to assess the proportion of senior and decision-making respondents versus younger practitioners. We broke the time in the profession in three groups: less than five years, five to 10 years, and over 10 years. Our sample favoured more experienced archaeological practitioners, with 55% having spent over 10 years in the profession, 21% in the profession between 5 and 10 years, 15% in the profession for less than 5 years, while only 9% of the sample were students of archaeology (see nearby chart, where gender and experience are mapped against one another).




When asked about primary workplace, 41% of respondents classified themselves as primarily 'academic' and 37% as 'consulting' archaeologists. All other workplaces were less well represented in the survey: government archaeologists occupied a fraction of 4%, museum employees 3% and students comprised 8%. 3% of respondents identified themselves as straddling the private and academic sector. The prevalence of academic over private sector employees differs markedly from Ulm, Nichols and Dalley (2005). In their survey of 'Australian archaeology in profile' 25% of archaeologists were employed in academia and 48% in the private sector (2005, 14). The high ratio of academic respondents to Digital Data survey may be the result of our limited address list (which included many consulting archaeologists, but overrepresented those based at universities), or access limitations on part of the consulting archaeologists, who frequently work in remote locations (http://ads.ahds.ac.uk/project/strategies/4.html#4.2). Our response rate is consistent with the *Strategies for Digital Data* survey, organized by the Archaeology Data Service in the UK, where "the worst response rates were from field archaeology units, the best from national bodies" (Richards, Robinson and Wise 1998, 5.3).

All in all, the limited distribution list and short circulation time for the survey (largely due to the late initiation of the project and the lengthy ethics clearance process) limited the sample size (81 respondents). By comparison, "Australian Archaeology in Profile" garnered some 301 respondents#. Additionally, the technological focus of the survey likely had a self-selective effect, again limiting the pool of respondents. Nevertheless, the survey represents between one-tenth and one-sixth of the archaeologists in Australia#, and likely captures a reasonable snapshot of the FAIMS project's initial target audience of technology-minded archaeologists, including those in more senior, decision-making positions.

**\heading Digital competence self-assessment**

Given the purpose of FAIMS project to develop a suite of applications for archaeologists, we attempted to assess user adoption rates. In the survey we asked the respondents to rate their digital literacy. Majority of respondents (45.6%) rated their digital literacy as "average". The next largest group (36.7%) claimed "above average" digital skills, while 11.4% asserted "expert" competence in digital technologies. Only 6.3% reported "below average" competence in using digital technologies. As the terms "average" and "expert" can be highly subjective without a referent we asked respondents to further characterize their digital skills and provided a range of scenarios (see table below). 19% of respondents reported the use of simple software (Office), 50.6% characterized themselves as competent users of more complex software (GIS and databases) and 17.7% reported database building skills. Some 12.7% claimed scripting and programming skills. These characterizations align well with the subjective ratings, matching expert status to scripting and programming. The characterizations further show that the "average" assessment of skills overlaps substantially with the statement: "I can competently use basic Office software".

| **Rate your digital competence** | | **Characterize your digital skills** | |
|---|---|---|---|
| Below average | 6.3% | I can competently use basic office software | 19.0% |
| Average | 45.6% | I can use more complex software (database and GIS software) | 50.6% |
| Above average | 36.7% | I can build databases | 17.7% |
| Excellent - Expert | 11.4% | I can do basic scripting | 7.6% |
| | | I am a programmer | 5.1% |
| | 100.0 % | | 100.0 % |

Overall, more than 90% of the respondents reported average or better competence with digital technologies, and 80% of respondents characterized themselves as users of more complex software. While some of these self-assessments may be subjective and result of sample bias to tech-savvy respondents, their value lies in reporting positive attitude to technology. The results give us hope as they reveal a sizeable group of potential users of the FAIMS infrastructure, who see themselves on the left half of Rogers's bell curve as "innovators", "early adopters", or "early majority" with regard to the use of technology\cite{Rogers1995}. Such a technological self-assessment supports for the idea that a receptive initial target audience for FAIMS applications exists in Australia.

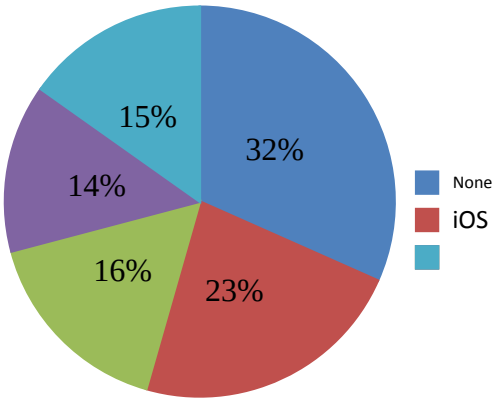
**Software platforms and devices used by archaeologists**

The ease with which archaeologists may adopt a new digital tool partially relies on their current use of specific operating systems and devices[1]. The use of platforms provides an indication of user expectations (in terms of software design, intuitiveness, commercialization, etc.) and the variety of tools used to create archaeological data.

All (100%) respondents use a desktop or laptop computer for their archaeological work. A majority use Windows (69%), followed by Mac (20%),

[1] The use of programmes in archaeology has already been surveyed in Richards 1998, and

with the remainder using a combination of commercial and open source operating systems (11%).




No one reported the exclusive use of an open source operating system (e.g., Linux). When asked about the deployment of mobile devices for archaeological fieldwork, 68% of respondents reported the use of some mobile device. A balanced variety of mobile operating systems were reported, starting with iOS (23%), Windows (16%), Android (14%), or a combination (15%) of the above. In short, the respondents seem well used to Windows (and Microsoft products) as well as Mac desktop environments, which bodes well for desktop applications. A wide variety of mobile environments is currently in use, offering equal chances of success (with appropriate training and support) to the FAIMS development for Android.

**Recording practices and programs commonly used by archaeologists**

The knowledge of specific recording practices and programs currently used by archaeologists forms an important part of stocktaking before the design of a generalized tool for the creation and sharind of digital data. The survey asked for details of the programmes used to create or process digital datasets. The table below lists four to five most popular packages for each category of recording.

Specifically, respondents were asked what spreadsheets, databases, and Geographic Information Systems (GIS) they used for their data. Spreadsheets proved almost universal, with 98% of respondents reporting their use. Spreadsheet users overwhelmingly favour MS Excel (98%). The remaining 2% of respondent use open source software (e.g., Open Office Calc or Google Spreadsheet) or Numbers (an iOS spreadsheet). Databases are also widely deployed by respondents (87%), the majority of whom use MS Access (67%, including 54% using it alone and 13% in combination with spreadsheets), followed by Filemaker (19%), open source databases like MySQL (10%), and ARCGIS geodatabases (4%). GIS consitutes the third most widely used class of software, deployed by 80% of respondents, amongst whom ESRI ARCGIS remains the most popular (65%), with the remainder using MapInfo (15%), open source GIS software (8%), Manifold (6%), and a combination of commercial and open source GIS (6%).

Majority of respondents use Microsoft products for cataloguing or reporting their data, while ESRI products appear most popular for spatial data processing. Open source or iOS products are in minority. The overwhelming preference for Microsoft is consistent with the ADS survey results from 1998, although our sample shows less overall diversity in the software used. There is a different trend in cataloguing data in MS Excel (98%) among Australian respondents rather than MS Access (55%), which is reversed in the UK.

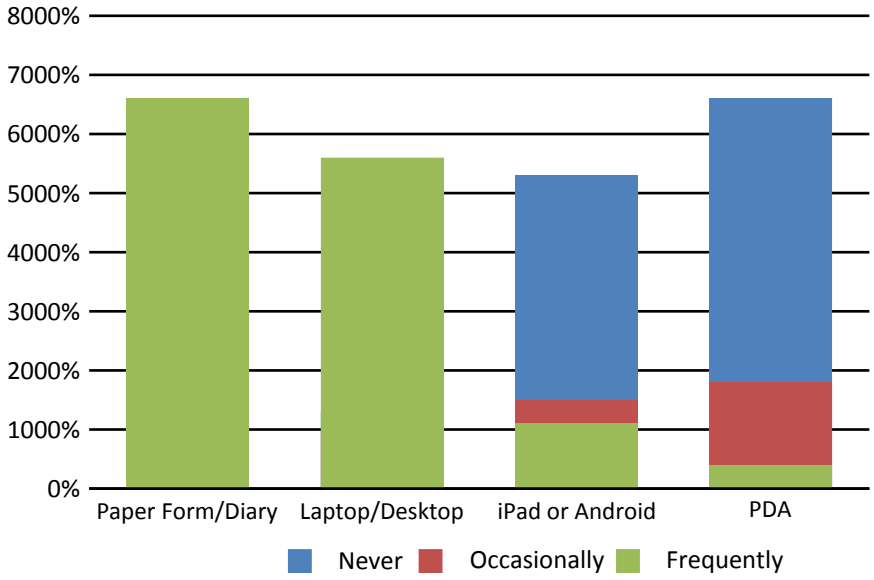
Implications needed. No connection was made between workplace and the use of a particular platform, although this might be interesting. Overall, the respondents capture a wide variety of data in a range of programmes and FAIMS needs to be able to support and archive these types of datasets and accommodate their collection in the mobile platform.

| **Spreadsheets** | | **DB** | | **GIS** | |
|---|---|---|---|---|---|
| None | 2.5% | none | 12.7% | none | 19.0% |
| MS Excel | 94.9% | MS Access | 46.8% | ArcGIS | 51.9% |
| numbers | 1.3% | Filemaker | 16.5% | MapInfo | 12.7% |

| | | | | | |
|---|---|---|---|---|---|
| OSS | 1.3% | MS Access combined with others (Lotus, Filemaker, etc.) | 11.4% | OSS | 6.3% |
| | | OSS | 11.4% | Manifold | 5.1% |
| | | GeoDB | 1.3% | Combination of the above | 5.1% |

On the Persistence of Paper recording for primary data collection

Despite the above-reported ubiquity of computers and widespread use of mobile devices, 84% of respondents reported that they frequently or most frequently use paper forms and journals as their primary recording medium for raw data. Laptop and desktop are used by 70% respondents for data. Only 10% of the respondents claim to have shaken off the use paper and pencil.




The answers to the question of primary data collection medium demonstrate that paper remains the principal recording medium in Australian archaeology. While majority of archaeologists own a mobile device, only some 10-15% use it with any frequency for primary data recording. Furthermore, there is a significant overlap in the use of paper (84%) and desktop/laptop (70%) for primary data recording. This overlap suggests that we have captured the practice of double-entry of primary data, or in other words, the usual workflow during which information in the field is recorded on paper only to get later transcribed into the appropriate spreadsheet or database on the computer in the base or office.

These results demonstrate the potential for the use of mobile application in Australian archaeology : most respondents already have and use a mobile device. Familiarity with the physical objects is helpful and presents one less obstacle to confront. The extent of double entry and use of paper offers a niche and large body of potential users of a mobile application.

**Learning CONTINUE HERE**

Asked about their preferred modes of learning new digital technologies, 92% of respondents selected formal online training, 60% formal in-person training, 40% online forums and blogs, 28% self-directed online learning, 21% and use of manuals, while 12% seek assistance from colleagues (note that respondents could select more than one answer). Responses to this question indicate that

FAIMS cannot rely only on software documentation, wikis, and blogs, but must offer training, at least online and preferably in person as well.

**Information Management needs and desires**

When ranking different types of digital tools by their desirability, archaeologists requested the following applications:

- Laptop and desktop data entry application (93%)
- Data analysis tools (92%)
- Mobile device application for data collection in field (88%)
- Tools facilitating online data publication (79%)
- Tools for automatic submission of data to online registries (72%)
- Mobile device application for labwork (69%)
- Ontology and vocabulary generating tool (48%)

Archaeologists' information management needs, as expressed through this survey, have been combined with the priorities of our initial proposal, NeCTAR requirements, technical, time, and financial constraints, contemporary trends in software development, and the existing landscape of resources available to archaeologists to produce a software development plan. Taken together, these factors have produced a plan that responds to archaeologists' priorities, but does not mirror them exactly (for reasons that will be discussed below). In short, mobile applications for data capture and online archiving and publication platforms (priorities three, four, and six) have become the focus of project development, while existing tools for data entry and analysis (using web rather than desktop applications) are being enhanced to accommodate priorities one and two.

**Data sharing and publication**

Views on data sharing expressed in responses to the survey are very positive. A majority of respondents are open to sharing primary data, with 20% willing to unrestricted distribution of their datasets and 44% willing to share after they have published their work. A quarter of respondents (24%) show more caution and prefer to select with whom they share their data. Relatively few respondents are averse to data sharing, either for personal reasons (5%) or due to external exigencies and trust issues (7%). Overall, 88% of survey respondents are willing to share primary data under at least some circumstances. An overwhelming majority of survey respondents (86%), moreover, affirm the idea of allocating credit, e.g., through the Higher Education Research Data Collection (HERDC) system, for the publication of substantial, peer reviewed primary datasets, with 14% opposed. Taken together, the answers to questions about data sharing and publication indicate significant support for the overarching goal of the project - to produce and disseminate high-quality, compatible, and reusable online datasets - as well as specific initiatives such as investment in an Australian data repository.

**Survey results summary**

## Workshop

In our NeCTAR proposal we requested a major face-to-face meeting with FAIMS stakeholders prior to the actual development. The purpose of FAIMS Stocktaking Workshop was to solicit input about development directions for the FAIMS project and to inform decisions about structure of the project, technologies, modules to be supported and 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). The workshop also served to bring together the members of Steering Committee.

The Workshop was held at UNSW on 16-19 August, 2012 and drew over 70 archaeologists, associated researchers and software developers, from Australia and overseas. The workshop was structured as a stocktaking enterprise with plenary sessions driving the agenda of smaller afternoon discussion workgroups. 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.

After the inspiration provided by the presentation of foreign and domestic case studies, the FAIMS stakeholders were broken up in small common themed workgroups, where they discussed recordkeeping standards, requirements for the mobile systems, federation strategies and other specific issues that would inform the FAIMS project development. The groups created addressed Archaeological Survey, Excavation, Artefacts - split into Ceramic and Lithic recording, Sciences, Federation, Sustainability and Sensitive Data. Each group had the opportunity to meet multiple times during the workshop, and morph the discussion as they saw fit.

Each of the groups was assigned leader who was responsible for ensuring that the results of the discussion were reported back daily during evening wrap-up plenary session. The leader also received an information package that provided some background and outlined goals for each group. Individual workgroup participants were encouraged to post the minutes of their discussions on Google Groups set up for that purpose by FAIMS. Membership in the discussion groups was voluntary.
The discussion group quidelines, as well as plenary session slides and audio files are available at our webpage (<http://www.fedarch.org/wordpress/?cat=10>).

The results of the group discussion did not always produce structured technical requirements to the level of detail we had initially envisioned. We did, however, obtain comprehensive general requirements regarding archaeological recording practice and workflow: [footnote: this list comprises ca 10% of the most commonly articulated requirements]

- “it needs to be as easy as paper recording”,
- “accommodate GIS input, output and visualisation of highest possible precision”,
- “it needs to facilitate data streamlining and cleaning”,
- “we need to be able to take paper notes, sketches, add files and annotate them”,

- “we want to control data input, and impose vocabularies”,
- “increase commensurability and consistency”
- “save, reuse, customize, proliferate schemas”
- “work without web connectivity”
- “synchronize data in the field”
- “collect contextual and user data, such as weather conditions, identity of user, etc.”
- “have autofill functions for dealing with repetitive data”
- “indicate the ‘level of confidence’ or ‘level of subjectivity’ that the user/data entry person has with respect to any of the fields”

- capture changes when you upgrade/change records or schema (versioning), track to user
- be able to re-label columns (aliasing)

- “be really rugged to handle Australian extremes (dirt, rain, heat, sun glare etc)”
- “have a LONG battery life”
- “has to be affordable”

While fulfilling all of these requirements is beyond the scope of FAIMS development during 2013, their compilation during the Workshop was extremely useful and informative. The discussion results suggested that different archaeological subdisciplines require customized different data types, precision and vocabularies (sieve dimensions, soil matrix versus shard fabric, etc.). However, they also highlighted general convergence at the level of complexity and flexibility required and underscored the overarching need for a system that is both robust, scalable and customizable. The FAIMS application needs to streamline data entry (default, manual and sensor input, error flagging), facilitate data redundancy, editing and review (track changes, use flags and certainty toggles,) and provides for user customization and interface personalization, while promoting the use of shared vocabularies, best practices, and data management.

**Focus groups results**

Finally, to maximise our information gathering several of the Workshop discussion groups were supplemented by a focus group element, facilitated by an experienced focus group leader. The groups consisted of contract archaeologists, state agency representatives, archaeological scientists, archaeological repository managers, surveyors and excavators. Membership in focus groups was voluntary.
The facilitator asked each group about perceived deficiencies in current archaeological practice, common frustrations with digital medium and needs in terms of collecting, sharing, editing and disseminating archaeological data. These focus groups were recorded, transcribed and analyzed to supplement the information gathered in other fora including the project survey and workgroup discussions.

Focus group discussions zoomed in on the limitations of current archaeological practice, specifically the production of vast amounts of paper records, high cost of digitization and the lack of incentives to share primary archaeological data. Limited technical skills, concerns over publication rights

concerns, sensitivity issues and contractual obligations to investors, and sensitivity issues all contribute to the lack of data sharing. Limited staffing and missing statutory requirements or loose regulatory requirements exacerbate the varied practice across different states and territory in Australia. Most of the participants thought that a comprehensive digital archaeological infrastructure could partially mitigate this situation.

The over the digital medium included fear of hardware failure, annoyance with inflexible software design, threat of catastrophic data loss and high costs associated with hardware and software acquisition and training. Again, the participants saw the benefits of digital medium in facilitating completion, and validation of records, and improving efficiency and redundancy of data, but the plusses and cons were in tight balance here.

The Focus groups more than the discussion groups revealed that archaeological community is one that is fairly conservative in the use of technology and that adopting an unproven system is a highly risky issue, especially among contract archaeologists who are constrained by considerations of cost and efficiency.

When discussing the potential of having a digital archaeological infrastructure, the groups indicated the need for an institutionalized national site registry/clearing house of archaeological data in Australia (or at state level), which would promote best practice through the review and validation of submitted digital data.

Overall the Focus groups provided us with feedback on higher level issues in digital data management, raising topics of missing governance and regulation, questioning sustainabilty of FAIMS infrastructure, and highling the idiosyncracies of archaeological perceptions of the digital medium. All of these caveats, while posing a risk to the success of the FAIMS project, have been invaluable as they have compelled us to reconsider the scope and direction of the FAIMS development, to moderate our expectations and innovative intentions, and to revisit the sustainability plan for the FAIMS infrastructure.

## Digital Data Conclusion

The survey was conceived as an initial and an extensive stocktaking exercise involving FAIMS stakeholders and the broader archaeological community. We invested considerable effort into our stocktaking efforts, and fundamentally revised our project in light of these activities, because we wished to avoid the problems that have undermined other information technology infrastructure projects. As such,