InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Title: **General Study 09 Final Report: Digital Recordkeeping Practices of GIS Archaeologists Worldwide: Results of a Web-based Survey**

Status: Final (public)

Version: 5.2

Submission Date: June 2006

Release: February 2007

Author: The InterPARES 2 Project

Writer(s): Randy Preston,
School of Library, Archival and Information Studies
The University of British Columbia

Project Unit: Focus 2

URL: http://www.interpares.org/display_file.cfm?doc=ip2_gs09_final_report.pdf

Version History:

Date	Version	Notes Regarding Modification	Changes by
1 Jun 2004	1_0	First Working Draft	Randy Preston
30 Jun 2004	1_1	First Working Draft	Randy Preston
13 Sep 2004	2_0	Second Working Draft	Randy Preston
21 Sep 2004	2_1	Second Working Draft	Randy Preston
23 Feb 2005	2_2	Second Working Draft	Randy Preston
16 Mar 2005	2_3	Second Working Draft	Kevin Glick
3 May 2006	3_0	Third Working Draft	Randy Preston
12 May 2006	3_1	Third Working Draft	Randy Preston
17 May 2006	3_2	Third Working Draft	Randy Preston
22 May 2006	3_3	Third Working Draft	Randy Preston
29 May 2006	3_4	Third Working Draft	Randy Preston
5 Jun 2006	4_0	Final Working Draft	Randy Preston
8 Jun 2006	4_1	Final Working Draft	Randy Preston
14 Jun 2006	4_2	Final Working Draft	Randy Preston
16 Jun 2006	5_0	Final Distribution Draft	Randy Preston
30 Jun 2006	5_1	Final Distribution Draft	Randy Preston
15 Feb 2007	5_2	Final (public)	Randy Preston

TABLE OF CONTENTS

LIST OF FIGURES

1.0 EXECUTIVE SUMMARY

1.1 Introduction

This report summarizes the results of a 40-question, Web-based survey of the digital recordkeeping practices of GIS archaeologists worldwide, conducted in March/April 2004 under the auspices of the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2 Project.

1.2 Background

Now in its final year, InterPARES 2 is a 6-year-long, international collaborative research project funded in part by Canada's Social Science and Humanities Research Council's Multidisciplinary Collaborative Research Initiative (MCRI) programme, and the National Historical Publications and Records Commission (NHPRC) and the National Science Foundation (NSF) of the United States. The goal of the InterPARES 2 Project is to develop effective preservation strategies for digital records to help ensure that society's digitally recorded memory will be accessible to future generations.

One key component of InterPARES 2 research methodology involves conducting general studies, such as the survey reported on here, of artistic, scientific and government activities that produce dynamic, interactive or experiential digital records. The primary purpose of these general studies is to assist each of the project's research units in achieving its objectives by providing supplemental data, in addition to the project's more specific case studies, about the creation and use of digital dynamic, interactive or experiential records, including their purpose(s), their phases and component actions, their byproducts and structures, their contexts, as well as their overall technological environment.

This general study survey was initially developed and administered as part of InterPARES 2 Case Study 14, "Archaeological Records in a Geographical Information System: Research in the American Southwest," led by Richard Pearce-Moses of the State Archives of Arizona. Case Study 14 examined how the records of a GIS system at the *Center for Desert Archaeology* in Arizona are created, maintained and preserved, and the corresponding impact of these processes on the records' authenticity, accuracy and reliability.[1] The goal of the case study is to help answer questions about the nature of digital archaeological records in general, about how the increased reliance on GIS is impacting the archaeological community and, especially, how both of these issues are played out in the recordkeeping habits of archaeologists. Although initially conceived solely as an analytical tool to help assess the representativeness of the Case Study 14 data with respect to the recordkeeping habits of the archaeological community as a whole, a

[1] Within the context of recordkeeping, InterPARES 2 defines these three concepts as follows: (1) Authenticity: The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption; (2) Accuracy: The degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter; (3) Reliability: The trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about, and is established by examining the completeness of the record's form and the amount of control exercised on the process of its creation (InterPARES 2 Project, *Terminology Database Glossary*. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm).

decision was made in September 2005, following a review of the survey's research activities, to extract the survey from Case Study 14 and reclassify it as a distinct general study, General Study 09.

1.3 Purpose and Scope

Within the broader context of InterPARES 2, the survey's general goal was to contribute to the project's overall research objective of developing effective strategies for long-term preservation of accurate, reliable, authentic and accessible digital records in the course of artistic, scientific and e-government activities.

On a more specific level, the goal of the survey was to gather and analyze baseline data about the existing digital recordkeeping knowledge and practices of GIS archaeologists worldwide to help gauge the current level of awareness and understanding within the global archaeological community about: (1) digital preservation issues, (2) digital recordkeeping practices and (3) the potential impact of such practices on the long-term preservation of accurate, reliable, authentic and accessible digital archaeological data and research records.

The survey findings offer a more systemic, baseline perspective on the current state of digital recordkeeping practices within the archaeological community than has hitherto been available. As such, it is hoped that this baseline perspective can serve as a standard against which archaeologists, archivists, collections managers and other key stakeholders can continue to measure and assess the overall effectiveness of continuing developments in the evolution of digital recordkeeping awareness and practices within the archaeological community in the coming years. Ultimately, it is hoped that the survey findings will assist records preservers (e.g., archivists and collections managers) and creators (archaeologists in this case) in better understanding the digital recordkeeping practices of archaeologists and, perhaps more importantly, the potential consequences of these practices on the creation and long-term preservation of accurate, reliable, authentic and accessible digital archaeological data and research records. At the very least, the survey will provide archivists, collections managers and other records preservers with a better understanding of the challenges they can expect to be faced with in the very near future when dealing with the long-term preservation of archaeological data and research records that undoubtedly will find their way into their custody.

1.4 Methodology

The use of a Web-based survey was chosen over more conventional survey options for its ability to reach a wider and, presumably, more representative sample of participants in a relatively easy, quick and cost efficient manner, as well as for its presumed receptiveness within a technologically savvy study group.

The survey consisted of 40 questions, including six 'free-text' questions and 34 either single- or multi-answer, multiple choice questions, each accompanied by a free-text box where participants could qualify, clarify or further comment upon their multiple choice answers, if so desired.

An initial list of prospective participants was compiled primarily from Internet resources, especially the database of GIS Archaeologists initially developed in 1995 by Paul Miller and Ian Johnson at the Sydney University Archaeological Computing Lab and, to a lesser extent, from names extracted from a literature review. Personal invitations to participate in the survey were e-mailed to nearly 900 GIS archaeologists from 69 countries worldwide. Additional invitations were posted to various archaeology- and GIS-related listservs and Internet discussion groups. The survey was available on-line for 32 days in March/April of 2004, from which 117 fully completed and 34 sufficiently completed surveys ultimately were accepted, bringing to 157 the total number of individual survey responses available for analysis.

Statistical correlation tests indicate that participant dropout likely had little, if any, impact on survey results. As well, various informal assessments of survey data *reliability*, including cross-question comparison tests for inconsistencies, suggest there is no reason to suspect that the survey results are unreliable. Due to the lack of an objective measure against which the concepts and issues being addressed in the survey could be assessed, no attempt was made to assess the *validity* of the GIS survey data.

1.5 Results and Conclusions

Analysis of the survey data reveals a number of encouraging and not-so-encouraging insights into the recordkeeping practices of GIS archaeologists. Among the more encouraging insights is a clear indication of a considerable, and growing, level of awareness among GIS archaeologists of the many technical, administrative, professional and societal issues surrounding the long-term preservation of their projects, particularly when compared to the results of an earlier 1998 survey that examined the use of GIS technology in archaeology.[2] A key indicator of this increased awareness is the growing sense of frustration expressed by many participants over the current lack of suitable long-term preservation repositories available to archaeologists, as well as over the continuing absence of any concerted, profession-wide response to these particular issues and concerns.

On the other hand, the survey also reveals that many participants currently engage in idiosyncratic and ad hoc file creation, management, preservation and/or documentation practices that have the potential to seriously compromise the accuracy, reliability, authenticity and accessibility of the files, especially over the long-term. The reluctance (or inability) of many participants to support their recordkeeping practices through the use of formal, systematic and documented procedures has the potential to seriously compromise the accuracy, reliability, authenticity and accessibility of project data and related research records, especially over the long-term. No doubt, some of the more haphazard approaches to GIS file creation, management, preservation and documentation are due, in part, to the fact that GIS archaeologists often have little or no formal GIS training.

[2] K. Gourad (1999), "Geographic Information Systems in Archaeology: A Survey" (unpublished Masters thesis, Hunter College of the City University of New York, Department of Anthropology), 75 pp. Both the thesis and the survey are available at http://khalid.gourad.com/thesis/.

One of the more troubling findings is that fewer than one-in-three participants is concerned enough about the long-term preservation of their projects to factor such concerns into their project planning, design or implementation. Further compounding this situation is a deeply entrenched sense of ambivalence among many of the participants about the importance of preserving their projects in the first place. This ambivalence is being driven by a number of mistaken or misguided assumptions. First, many participants assume that publication alone constitutes sufficient long-term preservation of their research, thus overlooking the critical role that related, unpublished research records play in establishing, supporting and ensuring the long-term integrity and authenticity of the research as a whole. Second, many participants assume that the rapid evolution of GIS tools and research techniques will render most current GIS research results obsolete within a relatively short period, thereby negating the need for long-term preservation of all but the original raw or base data. Finally, there is a perception, especially among those participants working as consultants, that concern for long-term preservation is entirely the client's problem. This perspective exposes a serious lack of understanding of the critical importance, especially when dealing with digital records, of actively addressing and integrating concerns for long-term preservation *throughout all phases* of a project, from initial conception through to completion and preservation.

Although most participants are aware of the importance of good documentation practices to the overall success of long-term preservation efforts, there nevertheless is considerable complacency among many of the participants about actually following through with what they acknowledge to be adequate documentation practices. For many, this complacency is exacerbated by an ongoing struggle to balance project results against documentation practices that, in the absence of standardized policies and procedures, many fear will morph into what at least one participant characterizes as "record keeping mania." This suggests that the real challenge for archivists (and other preservers) is not so much the need to convince archaeologists of the importance of providing (and preserving) an adequate level of project documentation to achieve their long-term preservation goals, as it is the need to highlight the key preservation issues with respect to documentation practices, and to offer archaeologists more practical guidance as to what types of documentation would be most effective to generate and preserve, and at what level(s) of detail.

Recent and encouraging developments are afoot within the archaeological community in the United States to establish a "cyberinfrastructure for archaeology" to facilitate long-term preservation, access and increased cross-project integration of digital archaeological information, especially unpublished data, reports, images and other digital objects, for the benefit of science and society. As the results of this survey clearly attest, however, to have any chance of achieving its goal, this ambitious project must broaden its scope beyond the purely technological issues that, at present, appear to be the project's primary concern, to also address the equally-important non-technical, socio-cultural recordkeeping issues, many of which are highlighted in this survey.

2.0 BACKGROUND

This report summarizes the results of a 40-question, Web-based survey of the digital recordkeeping practices of GIS archaeologists worldwide administered in March 2004. This research was conducted under the auspices of the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2 Project.

2.1 The InterPARES 2 Project

Now in its final year, InterPARES 2 is a 6-year-long, international collaborative research project funded in part by Canada's Social Science and Humanities Research Council's Multidisciplinary Collaborative Research Initiative (MCRI) programme, and the National Historical Publications and Records Commission (NHPRC) and the National Science Foundation (NSF) of the United States.

The immediate aim of InterPARES 2 is to develop and articulate the concepts, principles, criteria and methods that can ensure the creation and maintenance of accurate and reliable digital records and the long-term preservation of authentic digital records. Of particular interest are experiential, interactive and dynamic digital records created in the context of artistic, scientific and government activities. Scholars in the arts and sciences, archivists, artists, scientists, industry specialists and government representatives from around the world are working together to undertake the challenges presented by the manipulability and incompatibility of digital systems, technological obsolescence and media fragility. The ultimate goal of the InterPARES 2 Project is to develop effective preservation strategies for digital records to help ensure that society's digitally recorded memory will be accessible to future generations.

One key component of InterPARES 2 research methodology involves conducting general studies of activities that produce interactive, experiential or dynamic digital records[3] in each of the three areas of interest (artistic, scientific and government). The primary purpose of these general studies is to assist each of the project's research units in achieving its objectives by providing supplemental data, in addition to the project's more specific case studies, about the creation and use of digital dynamic, interactive or experiential records, including their purpose(s), their phases and component actions, their byproducts and structures, their contexts, as well as their overall technological environment.

2.2 Case Study 14

This general study survey was initially conceived, developed and administered as part of InterPARES 2 Case Study 14, "Archaeological Records in a Geographical Information System: Research in the American Southwest," led by Richard Pearce-Moses of the State Archives of

[3] These three types of records are defined by InterPARES2 as follows: (1) Interactive Record: A record with variable content or form that is dependent on user input that is often based on earlier content; (2) Experiential Record: A record produced, used and maintained in an experiential system (i.e., a system that immerses the user in a sensory experience); (3) Dynamic Record: A record that includes content taken from external sources that changes as those external sources change (InterPARES 2 Project, *Terminology Database Glossary*, *supra* note 1).

Arizona. Case Study 14 is examining how the records of a GIS system at the *Center for Desert Archaeology* in Arizona are created, maintained and preserved, and the corresponding impact of these processes on the records' authenticity, accuracy and reliability. The goal of the case study is to help answer questions about the nature of digital archaeological records in general, about how the increased reliance on GIS is impacting the archaeological community and, especially, how both of these issues are played out in the recordkeeping habits of archaeologists.

2.3 General Study 09

The GIS survey presented in this report was initially conceived to help assess the representativeness of the Case Study 14 data with respect to the recordkeeping habits of the archaeological community as a whole. In fact, the initial goal of the survey was to collect data that could be used to help determine whether the digital recordkeeping activities of the GIS archaeologist at the *Center for Desert Archaeology* could in fact be used as a general template for the analysis and development of preservation guidelines for digital records within the broader archaeological community. Following a recent review of the survey's research activities, a decision was made in September 2005 to extract the survey from Case Study 14 and reclassify it as a distinct general study, General Study 09.

Creation, implementation and analysis of the survey was primarily undertaken by InterPARES 2 graduate research assistant, Randy Preston (University of British Columbia), with initial creation assistance from fellow InterPARES 2 graduate research assistant, Erin O'Meara (University of British Columbia),[4] and technical assistance with on-line delivery of the survey from InterPARES 2 Project Technical Co-ordinator, Jean-Pascal Morghese.

[4] Ms. O'Meara also assisted with translations of the Spanish-language responses received from a number of survey participants.

3.0 PURPOSE AND SCOPE

3.1 Survey Rationale

Like practitioners in virtually every other area of scientific research, archaeologists have, over the course of just a few short decades, largely supplanted their traditional analogue-based research methods and materials with digitally-based alternatives. Due in part to the compressed time scale over which this fundamental change has occurred, together with the unprecedented rapidity with which digital technologies have evolved during this same time period, archaeologists have been forced to rapidly transform their research practices, often with unintended and unanticipated consequences. As most practitioners in the scientific community are now acutely aware, the relative ease with which researchers can now capture or create, and subsequently manipulate, research data and other research-related records in digital format, is both a blessing and, from the standpoint of long-term preservation, one of the most enigmatic and pressing recordkeeping challenges of our day. Indeed, in stark contrast to our former analogue-based systems, researchers are now able to capture or generate staggering amounts of numeric, textual, audio and/or video data, information, documents and other digital objects quite literally at the click of a mouse; objects that, in the absence of effective documentation and metadata administration, can quickly evolve into what might best be described as digital detritus.

Although many are quick to pin our current long-term preservation woes on the relative fragility of digital media and the rapid rate at which most digitally-based technologies become obsolescent, these are but two components of what actually is a much more complex and, in many respects, largely behavioural problem. In fact, even if we were able today to overcome the technical problems related to long-term preservation of digital objects, the long-term accessibility of those objects from both an *administrative* and an *intellectual* perspective would still be anything but assured, unless it could be demonstrated that our documentation practices at the time we set those objects aside were in fact adequate to ensure that future users could not only effectively find their way through the vast stores of preserved objects in search of the information they require, but also be able to make sense of that information, especially in terms of the original context in which it was generated, once it was located.

In short, it is argued here that, in light of the immense quantity of digital objects scientists are generating and hoping ultimately to preserve, successful long-term preservation of accurate, reliable, authentic and accessible digital objects turns on the ability of digital object creators and preservers to satisfactorily address not only the purely technical issues of media fragility and systems obsolescence, but also the equally important semi- and non-technical issues related to recordkeeping practices, including object creation and change documentation, content indexing and other metadata management tasks, among others. Inherent in all this is the requirement that digital object creators not only acquire sufficient foreknowledge of effective digital recordkeeping, documentation and preservation practices, but also that they incorporate such knowledge into their procedures for creation, manipulation, maintenance and preservation of the digital objects under their care that will help ensure those objects remain accurate, reliable, authentic and accessible for the long-term.

The purpose of this survey, therefore, was to gather baseline data about the existing digital recordkeeping knowledge and habits of GIS archaeologists to help gauge the current level of awareness and understanding within the archaeological community about: (1) digital preservation issues, (2) digital recordkeeping practices and (3) the potential impact of such practices on the long-term preservation of accurate, reliable, authentic and accessible digital archaeological data and research records.

3.2 Target Population

This survey targeted archaeologists who currently use, or who have in the past used, a GIS in the course of their archaeological activities and research. Although the survey was only administered in English, concerted efforts nevertheless were made[5] to encourage and accommodate participation from non-native English language archaeologists in hopes of obtaining input from as wide and diverse a cross-section of the international archaeological community as possible, within the constraints imposed by a mono-lingual survey.

[5] Including, for example, translation of survey responses received in Spanish and distribution of invitations to participate in the survey being sent to individuals irrespective of their geographic location and/or presumed nationality.

4.0 METHODOLOGY

4.1 About the Survey

Of the numerous approaches available for conducting a survey, a Web-based survey was chosen for two primary reasons. First, Web-based surveys are one of the easiest, fastest and most cost efficient ways to reach a wide, and presumably more representative, sample of GIS archaeologists worldwide. Second, given the topic of the survey, which presupposes a relatively high degree of computer literacy and technical expertise, it was presumed that most GIS archaeologists had access to, and experience with, the Internet as a research tool and would therefore be receptive to a Web-based survey format.

4.1.1 Data Acquisition Strategy

As noted above, this survey targeted archaeologists who currently use, or who have in the past used, a GIS in the course of their archaeological activities and research. A list of prospective candidates was compiled primarily from Internet resources, especially the database of GIS Archaeologists initially developed in 1995 by Paul Miller and Ian Johnson at the Sydney University Archaeological Computing Lab.[6] The database is an international directory, arranged alphabetically by country, of archaeologists specializing in GIS applications. As of January 2004, the database contained just over 1000 individual listings. However, a large percentage (>45%) of these included "duplicate" entries (e.g., the same individual listed more than once under different name variants or e-mail addresses) and individuals whose listed e-mail contact no longer was valid. Potential candidates also were compiled from a cursory literature review, a process hampered by the fact that not all sources included e-mail contact information. In total, 896 individual e-mails were sent to archaeologists in 69 countries inviting them to participate in the survey. A second "reminder" e-mail was sent out three weeks later. Of the 896 e-mails sent, 350 ultimately were returned as undeliverable. A similar invitation and follow-up reminder also were posted on several archaeology- and/or GIS-related Internet listservs.

4.2.2 Survey Design

As used here, survey design refers to a survey's physical layout and flow. Because survey design has been shown to have an effect on the quality of participants' responses,[7] it is an important variable to consider when assessing the results of any survey. Some researchers argue, for

[6] P. Miller and I. Johnson (2005), "Archaeologists using GIS," Archaeological Computing Laboratory, University of Sydney. Available at http://acl.arts.usyd.edu.au/~scripts/contacts/list_generator.cgi?gis_arch_long.txt. Note: When last checked on 23 Feb 2005, this URL was no longer valid. The current URL is: http://www.acl.arts.usyd.edu.au/index.php?option=com_content&task=view&id=96&Itemid=120. Unfortunately, according to information provided at this new URL, the list of GIS archaeologists is currently not available. However, a copy of the list, current as of 3 Feb 2004, is accessible via the Internet Archive Wayback Machine at http://web.archive.org/web/20040203091815/http://acl.arts.usyd.edu.au/~scripts/contacts/list_generator.cgi?gis_arch_long.txt.

[7] See, for example: D. A. Dillman and D. K. Bowker (2001), "The Web Questionnaire Challenge to Survey Methodologists," in U. D. Reips and M. Bosnjak (eds.), *Dimensions of Internet Science* (Lengerich: Pabst Science Publishers), pp. 159–178; and D. A. Dillman, R. D. Tortora and D. K. Bowker (1998), "Influence of Plain vs. Fancy Design on Response Rates for Web Surveys," in *The 1998 Proceedings of Section on Survey Research Methods* (Dallas, Texas: American Statistical Association). PDF version available at http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf.

example, that Web-based surveys should be presented in conventional formats mirroring those normally used in self-administered paper survey questionnaires,[8] while others suggest that the uniqueness of the Internet medium, with its special interactive design capabilities, visual elements, etc., invite, if not require, special handling of survey design.[9]

One of the key issues of discussion with respect to survey Web design is the use of one- vs. multiple-page design. One-page design is characterized by static, plain HTML (or equivalent) forms in which all survey questions are arranged, often one after the other, on a single HTML page. With this design, participants are able to view the entire survey by scrolling up and down as necessary. In most cases there is no active interaction (such as answer-dependent skip or jump actions) with the participant during survey completion. Such surveys are, essentially, just digital versions of their paper counterparts.

Multiple-page design is characterized by server-side HTML forms capable of various interactive controls such as automatic question skipping and jumping, conditional branching, response validation, time measuring, etc. In this approach, related questions may be grouped together on multiple pages, or, as was done for the GIS survey, each question may be presented on a separate page. Both designs have their advantages and disadvantages.[10]

4.2.3 Survey Content

The survey consisted of 40 questions, two of which were optional.[11] Six of the questions were of the 'free-text' variety, where participants were able to provide open-ended responses. The remaining 34 questions were either single- or multi-answer, multiple choice questions. In addition, each multiple choice question was accompanied by a free-text box where participants could qualify, clarify or further comment upon their multiple choice answers, if so desired. The survey was organized into seven sections as follows:

A. Introduction (1 question)
B. GIS Experience/Background (7 questions)
C. File Management/Documentation Procedures (11 questions)
D. Digital Preservation Practices (13 questions)
E. Data Input/Output Practices (2 questions)
F. Record Quality, Reliability and Authenticity Issues (5 questions)
G. General Comments (1 question)

---

8 See, for example: D. A. Dillman (2000), *Tailored Design of Mail and Other Self-Administered Surveys* (New York, NY: Wiley-Interscience).

9 See, for example: M. P. Couper (2000), "Web Surveys: A Review of Issues and Approaches," *Public Opinion Quarterly* 64(4):464–494.

10 R. L. Clayton and G. S. Werking (1998), "Business Surveys of the Future: The World Wide Web as A Data Collection Methodology," in M. P. Couper, R. P. Baker, J. Bethlehem, C. Z. F. Clark, J. Martin, W. L. Nicholls II and J. M. O'Reilly (eds.), *Computer Assisted Survey Information Collection* (New York, NY: John Wiley and Sons), pp. 543–562; Dillman, *Tailored Design of Mail and Other Self-Administered Surveys*, *supra* note 8; and S. W. Spain (2004), "Top-10 Web Survey Issues and How to Address Them," *iResearch*. Available at http://www.iresearch.com/pages/library/top_10.cfm.

11 While the "Question *X* of 39" reference on the online version of the survey (see Appendix A) suggests there were only 39 questions in the survey, the 39 in this case actually refers to the total number of web *pages* or *screens* presented to the survey participants, not the total number of questions. In fact, the screen for question B5 consisted of two questions (B5a and B5b), bringing to 40 the actual total number of questions in the survey.

Section A consisted of a single, free-text question in which participants were asked to provide a brief definition of GIS, including what they considered to be its most important, significant and/or distinguishing components and functions. The primary purpose for asking this question was to assess what importance, if any, the participants assigned to their role (i.e., the human operator) in the overall geographic information system.

Second B consisted of seven questions. The questions in this section were designed primarily to gather basic background information about the participants and their current level of experience with GIS projects, including how long they have been using GIS, how often they use it, where they use it, etc.

Section C asked eleven questions designed to assess the participants' GIS project and file documentation and management habits. Participants were asked, for example, about their file naming procedures, the ways in which they managed file version control and how often they created non-digital records associated with their digital records. The primary goal of this section was to understand the ways in which the participants dealt with modifications to their data files in terms of if, when and how they typically documented these changes.

Section D was the largest section in the survey, asking thirteen questions designed primarily to gather information about the preservation strategies used by the participants when saving their GIS projects for the long-term, either "in-house" or in a designated repository. In addition, a number of the questions in this section dealt with the issue of metadata, looking especially at how the participants recorded metadata, what, if any, standards they followed and how they integrated metadata into their GIS projects.

Section E consisted of just two questions designed to assess the overall routineness of the procedures the participants used when creating and manipulating the various components and outputs of their GIS projects, and the degree to which those procedures were based on any sort of GIS procedures manual.

Section F asked five questions focusing on data/record reliability, authenticity and accuracy issues related to the participants' GIS projects, including, for example, the security measures used to control access to, and prevent unauthorized modifications to, their project files.

Section G was optional and consisted simply of free-text space for the participants to submit any final thoughts about their GIS recordkeeping activities or experiences that they thought might be pertinent to the survey.

4.2 Risk Analysis

Some researchers suggest that, for the following reasons, Web-based surveys may be more prone to measurement errors than other survey modes. First, because of the relative ease with which they can be created and administered, most Web-based surveys tend to be designed by individuals with little or no training in survey methodology, which often results in poor survey design.[12] The potential effect of this situation must be acknowledged, since, although some background research in survey method and theory was involved in the design of this survey, the survey's author has no formal training in survey methodology. Second, various studies indicate that, on average, individuals accessing information via the Internet read online information more quickly and discriminatingly (i.e., more use of scanning) than they do off-line information.[13] Analysis of the likely effects of this on the results of the GIS survey are presented below in section 4.4 Completion Time).

As with any survey, particularly a Web-based survey of this length and complexity, complications are inevitable. Two potential and three actual problems were encountered during administration of the survey. The first potential problem was that anyone with Internet access, regardless of whether they were an archaeologist, could theoretically have gained access to the survey. In fact, other than gleaning potential candidates from archaeology-related databases and literature sources, no formal or systematic attempt was made to ensure that those individuals invited to participate in the survey were, in fact, archaeologists. As well, other than identifying the survey as a "Survey of Recordkeeping Practices of Archaeologists," and indicating in the Informed Consent Letter—which all participants were to asked to read and required to accept before accessing the survey—that the survey was intended for "archaeologists with past or present first-hand experience in archaeologically-related Geographical Information System (GIS) projects," no other attempt was made to ensure that those individuals who participated in the survey were in fact archaeologists. Considering the highly specialized nature of the survey, it seems highly unlikely that anyone "posing" as an archaeologist with GIS experience would have been able to complete the survey without arousing suspicions. Indeed, as noted below, only 1 of the 195 responses received was confirmed to be from a non-archaeologist (who, despite having considerable GIS experience, was nevertheless excluded from the analysis). The likelihood of individuals outside of the target population accessing the survey was further mitigated by requiring participants to use a log-in ID and password (provided in the invitation e-mails) to enter the survey.

The second potential problem involved the fact that participants were required to complete the survey during a single online session. Because of the nature of the anonymous log-in system used, participants could not return at a later date to a partially completed survey and continue from where they had left off. Also, survey sessions were set to time-out after a specified period of inactivity. Thus, if, as happened in at least one case, participants were temporarily distracted while taking the survey (e.g., by a telephone call), they ran the risk of having their survey session timing-out. Unfortunately, such individuals would be required to log back into the survey and start again from the beginning. As well, whatever answers they had completed in their aborted attempt would still have been recorded in the database. Thus, it is possible that someone who had completed enough of the survey to qualify for inclusion in the analysis before timing-out would be represented more than once if they later completed the survey during a subsequent session.[14] However, a comparison of the answers to certain key questions (e.g., years of experience, current

---

[12] Couper, *supra* note 9.

[13] Manfreda, K. L., Z. Batagelj and V. Vehovar (2002), "Design of Web Survey Questionnaires: Three Basic Experiments," *Journal of Computer-Mediated Communication* 7(3): no pagination. Available at http://www.ascusc.org/jcmc/vol7/issue3/vehovar.html.

[14] For that matter, other than perhaps the time commitment involved, there was nothing to dissuade individuals from completing the survey more than once, should they have been so inclined.

professional affiliation, geographic location and area of research, etc.) between the fully completed surveys and those partially completed surveys included in the analysis (see below), revealed no positive matches.

The first of the three actual problems was that several individuals were unable or unwilling to access the Web-based survey. Reasons given included: (1) unwillingness to accept browser cookies, (2) unreliable Internet connection, (3) browser incompatibility and (4) a temporary server crash on the penultimate day of the survey. This problem was mitigated by developing and distributing to these individuals an MS Word™ version of the Web-based survey that they were instructed to complete at their convenience and return to the author via e-mail.

The second actual problem involved selective input by some participants who chose to ignore (or erroneously/accidentally skipped) certain questions (or parts of questions). This may have been due, in part, to inherent limitations in the online architecture of the survey. Questions were presented in chronological order one screen at a time, with the participant required to click on a "submit" button to advance to the next question. Once submitted, an answer could be reviewed (via the browser's "back" button), but could not be changed, regardless of whether a submitted answer had, in fact, been answered. Thus, if one were to accidentally click the submit button before answering (or fully answering) a question, there was no way for the participant to go back and edit that question. Attempts were made to mitigate the chances of this problem occurring by including a warning message directly above the submit button for each question. Fortunately, this problem does not appear to have occurred with any significance, as only 1.7% of all potentially answerable questions (or parts of answers) went unanswered.[15] A majority of the unanswered questions (65.4%) involved four free-text questions in which participants were asked to provide brief definitions for specific GIS-related concepts. In any event, so as not to skew the results, unanswered questions were excluded from consideration when generating the summary distributions for each question. Thus, the total number of responses used in the calculation of the relative distribution of answers for each question varies, depending on the total number of valid responses received for that question.[16]

The third and final actual problem was by far the most serious, if only for the potential loss of data in which it resulted. Due to a database programming error, 44 participants inadvertently were not given the opportunity to answer 4 questions related specifically to metadata. In particular, participants who answered, "The procedures used to create the file" to question D8,[17] inadvertently were automatically redirected to question D13 upon submitting their question D9 answer, thus circumventing questions D9 – D12. Nevertheless, because the summary percentage distributions for each question are calculated based on the actual number of participants who answered each question, the only real effect of this problem will have been to reduce the sample size for the four questions that were inadvertently skipped (i.e., D9 – D12) in some cases.

[15] This calculation takes into account all partially completed surveys included in the analysis by excluding from consideration all unanswered questions beyond the last answered question.
[16] This number is provided in the caption of the distribution graph for each question as: n=[total number of valid responses].
[17] Question D8 asked: "When recording metadata to document your GIS project files (or groups of files), which of the following types of information do you routinely include?" Participants were asked to select all applicable choices from the list of 11 options presented.

4.3 Response Rate

A total of 195 survey log-ins were recorded during the 32 days that the survey was available on-line. Of these, 43 were rejected due to insufficient completeness and one was rejected because the participant identified him/herself as other than an archaeologist.[18] Of the remaining 151 surveys, 117 were fully completed, while 34 others were sufficiently completed to warrant inclusion in the analyses that follow.[19] In addition, six fully completed MS Word™ version "mail-in" surveys were received, bringing to 157 and 123 the total number of surveys included in the analysis and the total number of fully completed surveys, respectively.

Because general invitations to participate in the survey were posted to several listservs, coupled with the fact that the log-in process was fully anonymous, it is impossible to provide an exact calculation of the rate of response. Assuming that all 157 responses were generated solely from the 546 "successful"[20] e-mail invitations sent to individual archaeologists yields a maximum theoretical response rate of 28.8%. The actual response rate undoubtedly is lower. Nevertheless, given the length and complexity of the survey, together with the fact that is was only open for 32 days (2 April – 3 May), this is considered a better than average response rate.[21] By comparison, for his 8-month-long, 1998 Web-based survey of GIS archaeologists, which, according to the opening instructions was designed to take only "about 5 minutes" to complete (versus the ca. 30 minutes for the current survey), Gourad sent out over 2000 individual e-mail invitations, in addition to invitations posted to various listservs. In the end, 140 acceptable responses were received, yielding a maximum theoretical response rate of less than 7%.[22]

4.4 Completion Time

The mean completion time for the 123 participants who fully completed the on-line version of the survey was just under 30 minutes (0:27:11). The fastest and slowest completion times were 0:5:52 and 2:07:36, respectively. The observed variations in actual completion times can, in part, be attributed to procedural variables such as: (1) how many questions automatically were skipped,[23] (2) how many of the free-text questions were answered (and to what extent) and (3) how many of the answers to the 34 multiple choice questions were accompanied by comments in the "comments" text box provided with each question. For example, a review of the answers provided by the 'fastest' participant reveals that 11 of the 39 questions automatically were

[18] Four of the 195 log-ins involved individuals who, for whatever reason, cycled through every question in the survey without answering any. At least one of these appears to have involved a participant who completed the survey and then immediately logged in again and cycled through the survey a second time without answering any questions. Another 20 log-ins involved individuals who cycled through some of the survey questions without answering any of them. Except where noted, all non-response log-ins have been excluded from the analyses in this report.

[19] As used in this report, a "fully completed" survey is one in which the participant answered all required questions. A "sufficiently completed" survey is one in which the participant either skipped some required questions and/or dropped out prior to the last required question (i.e., question number 39, or F5 as it was designated in the survey), but which was adjudged to contain enough data to warrant inclusion in the summary analyses presented in this report.

[20] i.e., the 896 originally sent out, minus the 350 returned as undeliverable.

[21] For discussion of the potential effect of response rate and its cousin, dropout rate (i.e., the percentage of respondents who terminate their participation in a survey prematurely), on the reliability of the survey, see section *4.5.1 Reliability*.

[22] Gourad, *supra* note 22, p. 75.

[23] The applicability of a number of the questions in the survey was based on a participant's answer to a previous question. In cases where a participant's answer to a question rendered a subsequent question, or group of questions, "not applicable," the participant was automatically redirected to the next applicable question in the survey.

skipped, only one of the free-text questions was answered, and none of the participant's answers were accompanied by additional comments. In contrast, the 'slowest' participant automatically skipped only three questions, answered all free-text questions in considerable detail, and provided substantial comments with almost all answers (26 out of a possible 31).

4.5 Survey Data Characteristics

A useful survey is one that allows researchers to satisfactorily draw conclusions, formulate theories, or otherwise make claims about the generalizability of its data. Surveys with low reliability and/or validity are difficult to interpret and impossible (or at least pointless) to generalize. Therefore, for a survey to be useful, it first must be demonstrated that the survey's data are both reliable and valid.

4.5.1 Reliability

In the context of survey design, survey reliability is a measure of the extent to which a survey will provide the same results when administered repeatedly. Although survey reliability is affected by various factors, the three most important from the researcher's point of view include: (1) the length of the survey (especially in terms of completion time and its effect on response rate), (2) the quality of its questions and (3) its fit to the group being measured.

In a seminal 1973 article in which he summarized the state of knowledge concerning the relationship between survey length and response rate (primarily with respect to mail surveys), Berdie concluded:

> Common sense suggests that the shorter the questionnaire, the more likely a high response rate, and persons studying questionnaire efficiency have tended to accept this belief in spite of little empirical evidence to support it .... Surprisingly few studies actually have examined correlations between length of questionnaires and rate of response, and those studies that have done so generally have yielded confusing results.[24]

Although the nature of survey vehicles has evolved considerably since the early 1970s, especially with respect to telephone survey techniques and the introduction of the Internet, a more recent review of the literature confirms that there has been remarkably little corresponding evolution in our understanding of the relationship between survey length, response rate and reliability.[25] In fact, this latter review concluded "that the results are still confusing and contradictory, the conclusions are still not clear, and questionnaire designers still aim for shorter questionnaires with little more justification than the logical assumption that longer interviews will result in higher nonresponse."[26] In a widely cited work, Nunnally, for example, suggests that

[24] D. R. Berdie (1973), "Questionnaire Length and Response Rate," *Journal of Applied Statistics* 58(2):278.
[25] K. Bogen (1996), "The Effect of Questionnaire Length on Response Rates: A Review of the Literature," in *Proceedings of the Section on Survey Research Methods, 51st Annual Conference of the American Association for Public Opinion Research, May 16-19, 1996, Salt Lake City, Utah* (Alexandria, VA: American Statistical Association), pp. 1020–1025.
[26] Ibid, p. 1020.

the reliability of a survey is *positively* correlated with the number of questions; the larger the number of questions, the higher the reliability of the survey.[27] Others have argued that, as long as the overall design of a survey is "respondent-friendly," its length can be increased (within limits) without necessarily adversely impacting response rates.[28]

This continuing equivocation stems, in part, from an inconsistency in the way the operational definition of questionnaire length is defined by various authors. Studies that define length as the total number of questions and/or pages often find a negative relationship between questionnaire length and response rates. However, this same relationship tends to be less evident in studies that define length in terms of the duration of a survey in minutes. Complicating matters further is the belief held by many researchers that response rates often are influenced by *subjective* rather than *objective* indices of questionnaire length. Under this scenario, perception of survey length often will be influenced by various subjective factors, most notably the participants' interest in the topic of the survey. Consequently, surveys with higher perceived relevance to the participants typically result in higher response rates regardless of length.

Questionnaire length also is commonly believed to affect data quality. The longer a questionnaire lasts, the more susceptible the participants are to becoming tired, annoyed, bored and/or distracted by external factors. Obviously, this can negatively affect data quality by decreasing the degree of effort and thought that participants are willing to invest in answering the questions. Findings on the effect of questionnaire length on data quality generally suggest the presence of an inverse relationship in which overly long questionnaires negatively effect data quality.[29] Again, this effect appears to be moderated by certain subjective factors, such as level of interest in the questionnaire topic.

Given the infancy of the on-line survey technique, relatively little research into its reliability has been conducted to date. According to Miller, however, recent research into on-line surveys suggests the existence of a positive correlation between survey length (expressed as time in minutes required to complete the survey) and respondent dropout rates (Figure 1). More importantly, this research cautions that dropouts, when viewed as a type of non-response error, can negatively impact the reliability of survey results. Non-response error refers to situations in which the attitudes, opinions and/or behaviours of survey non-respondents (including dropouts) differ in important or significant ways relevant to the survey from survey respondents.[30] This can "result in biased, inflated estimates of concept interest," since "people who drop out of…surveys tend to be less interested in the concept being tested than individuals who do not."[31]

[27] J. C. Nunnally (1978), *Psychometric Theory* (New York: McGraw-Hill).
[28] D. A. Dillman, M. D. Sinclair and J. R. Clark (1993), "Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys," *Public Opinion Quarterly* 57:289–304.
[29] See, for example: B. Burchell and C. Marsh (1992), "The Effect of Questionnaire Length on Survey Response," *Quality and Quantity* 26:233–244; J. G. Helgeson and M. L. Ursic (1994), "The Role of Affective and Cognitive Decision-Making Processes during Questionnaire Completion," *Public Opinion Quarterly* 11(5):493–510; and A. R. Herzog and J. G. Bachman (1981), "Effects of Questionnaire Length on Response Quality," *Public Opinion Quarterly* 45:549–559.
[30] A. Hogg and J. Miller (2003) "Watch Out for Dropouts," *Quirk's Marketing Research Review* July/August. Available at http://www.quirks.com/articles/a2003/20030706.aspx?searchID=14722804&sort=9.
[31] Ibid.

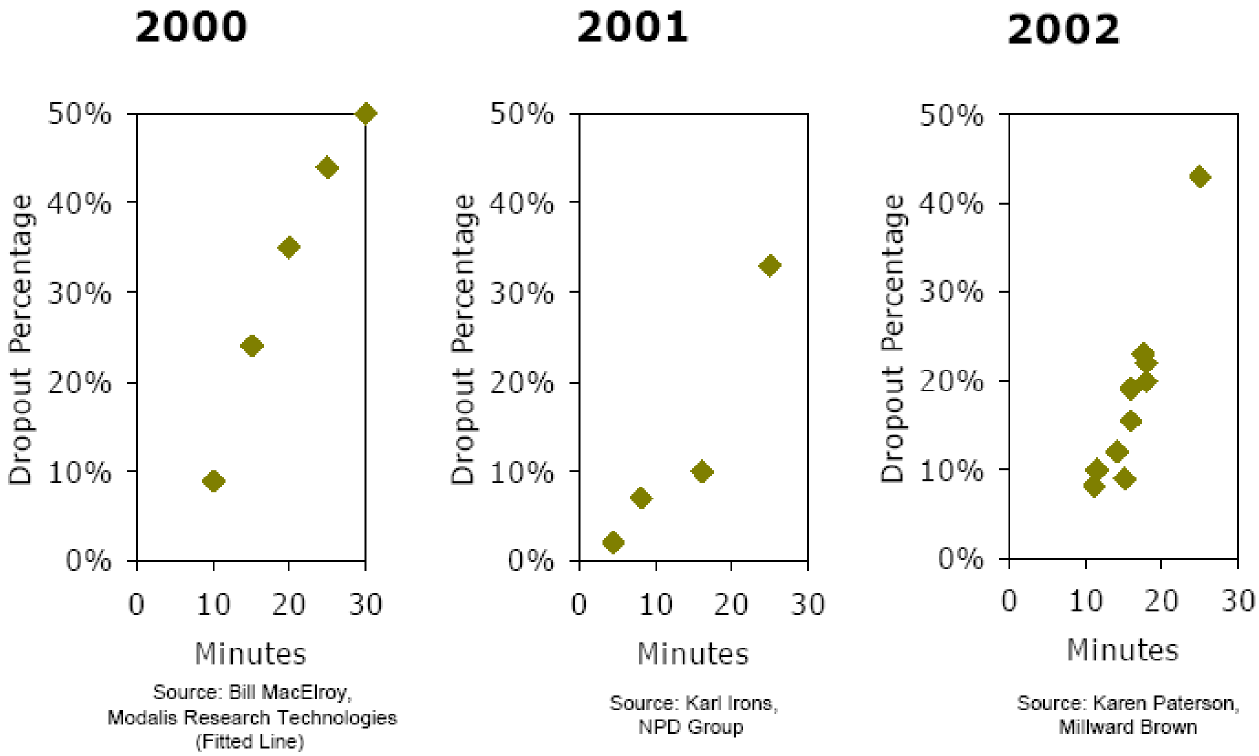


Figure 1. Correlation of dropout rate with length of survey (in minutes) for on-line "concept" surveys.[32]

According to Hogg and Miller, longer Web-based surveys "seem to be more prone to this bias."[33] Of particular concern here is the degree to which non-response error may have affected the generalizability of the survey results. It is important to emphasize, however, that non-response and dropout do not *ipso facto* result in non-response error. In fact, if the relevant attitudes, opinions and/or behaviours of non-respondents and dropouts are similar to those of respondents, there is no non-response error.

4.5.1.1 Survey Dropout

Analysis of the dropout rate associated with the GIS survey shows that the total dropout rate was 27.3%[34] (Figure 2, absolute dropout) and that more than three-quarters (79.5%) of the participants who did drop out did so within the first 20 minutes (Figure 2, cumulative dropout). One of the most commonly cited variables associated with dropout is survey length. Although various factors

[32] Figure taken from: J. Miller (2003), "Online Survey Length: Can Research Findings be Impacted?" *American Marketing Association, 24th Annual Marketing Research Conference, September 14-17, 2003, Los Angeles, CA.*

[33] It is worth emphasizing, however, that these findings are based on what commonly are referred to as consumer "concept test" surveys; surveys which attempt to measure consumer preferences and/or product usage motives. While one may be able to draw certain general parallels between the goals of consumer surveys and the GIS survey, it is argued that the two fundamentally are quite distinct. Consequently, it is questionable whether the methodological research findings of the former can be presumed to apply to the latter.

[34] This figure excludes participants who dropped out of the survey without answering a single question. The total dropout rate for all participant log-ins, including those who dropped out before answering any questions, is 36.2%.

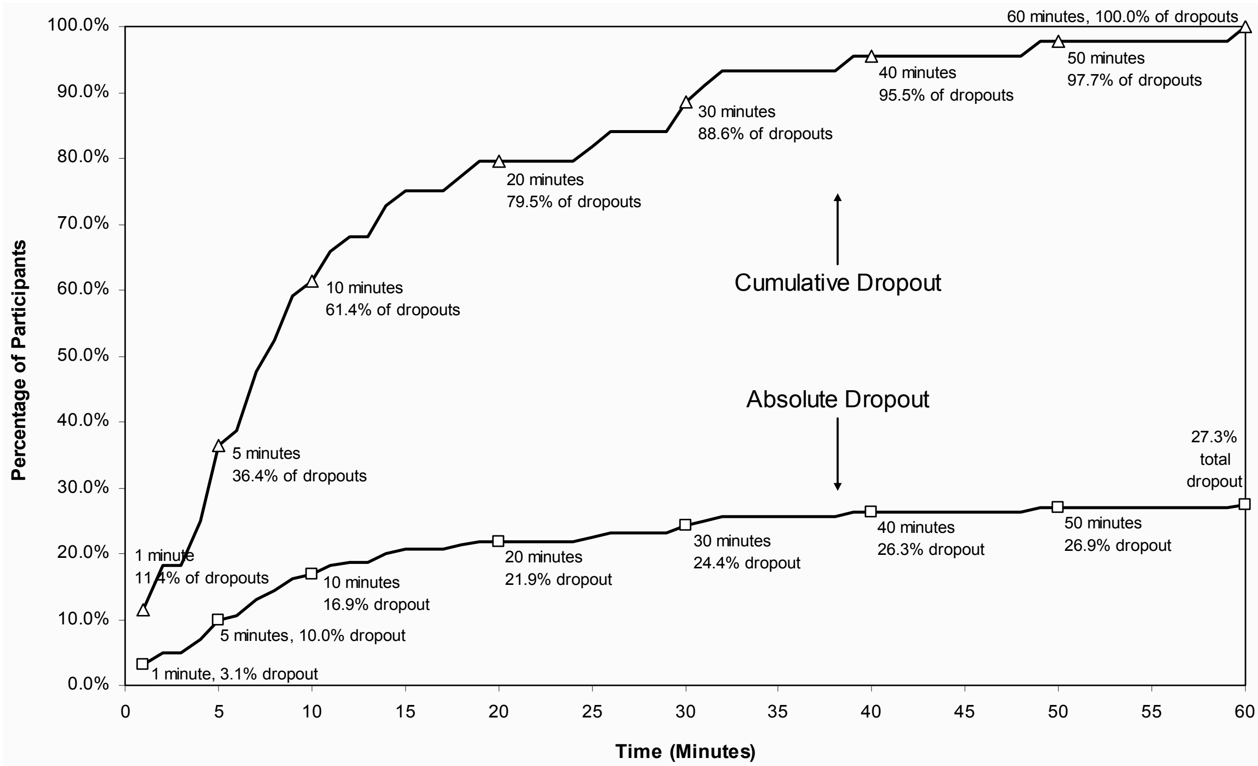


Figure 2. Relationship of participant dropout rates to total time spent on survey (excludes dropouts who did not answer any questions).

influence the length of time it takes to complete a survey, one of the most important (and easiest to quantify) is the total number of questions in the survey. Figure 3 shows the relationship between the total number of questions answered and the dropout rate for the GIS survey. Exactly half (50.0%) of the participants who dropped out did so after answering 15 or fewer of the survey's 40 questions (Figure 3, relative dropout rate line). Three notable increases in the dropout rate are evident at about the 2-4, 11 and 17 question levels (Figure 3, red lines). To assess whether something about the content or structure of the survey questions themselves inadvertently may have helped to concentrate participant dropout at these three parts of the survey, dropout rates by question number (as opposed to number of questions, as is shown in Figure 3) were plotted (Figure 4).[35] As is shown in this figure, these three concentrated dropout areas correspond to question numbers 3-5 (B2-B4 in the survey), 12 (C5 in the survey) and 18 (C11 in the survey).

Questions 3-5 asked participants to indicate the frequency with which they use a GIS, the type(s) of GIS activities with which they are, or have been, involved and the type(s) of GIS system design(s) that they typically use, respectively. All three questions were presented in multiple choice format. Together with Question 2, which asked participants about their total years of ex-

[35] Although subtle, this distinction is necessary since some dropouts automatically skipped certain questions as a result of their answers to previous questions. In fact, depending on how certain questions were answered, participants who dropped out at question 12 could have been asked to answer as many as 12 or as few as 11 questions, while those who dropped out at question 18 could have been asked to answer as many as 18 or as few as 12 questions.

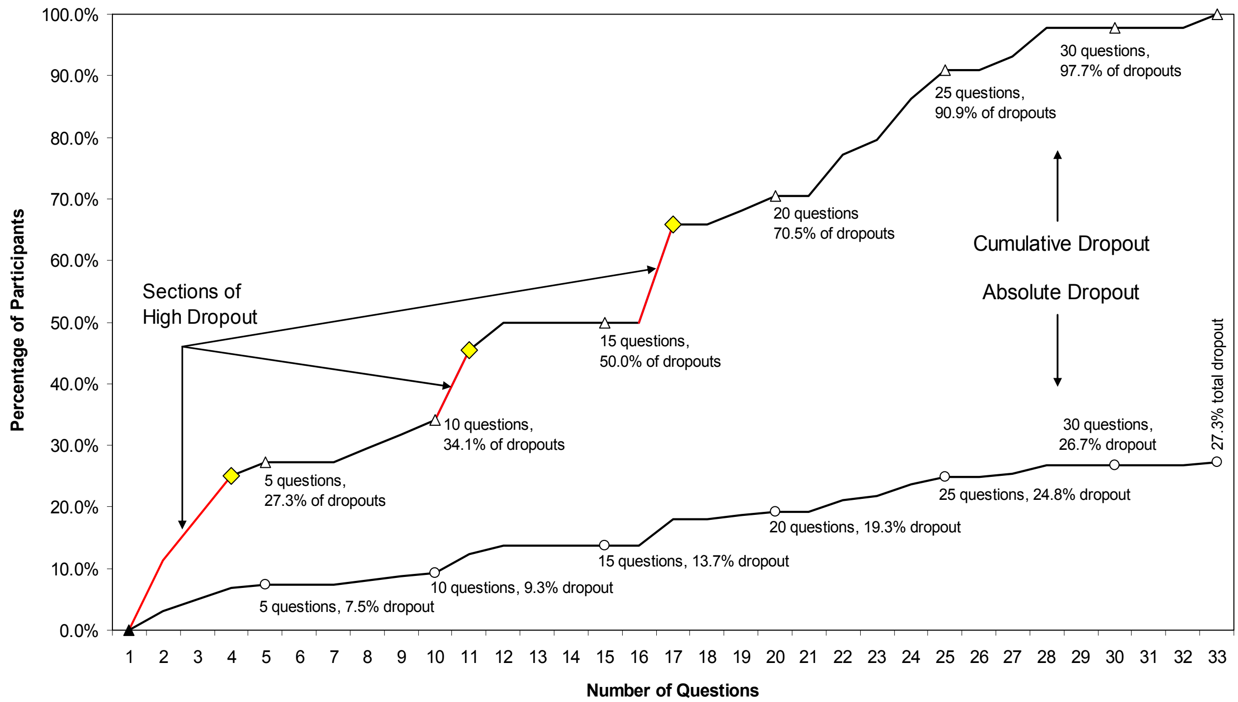


Figure 3. Relationship of participant dropout rates to total number of questions answered at time of dropout (excludes dropouts who did not answer any questions).

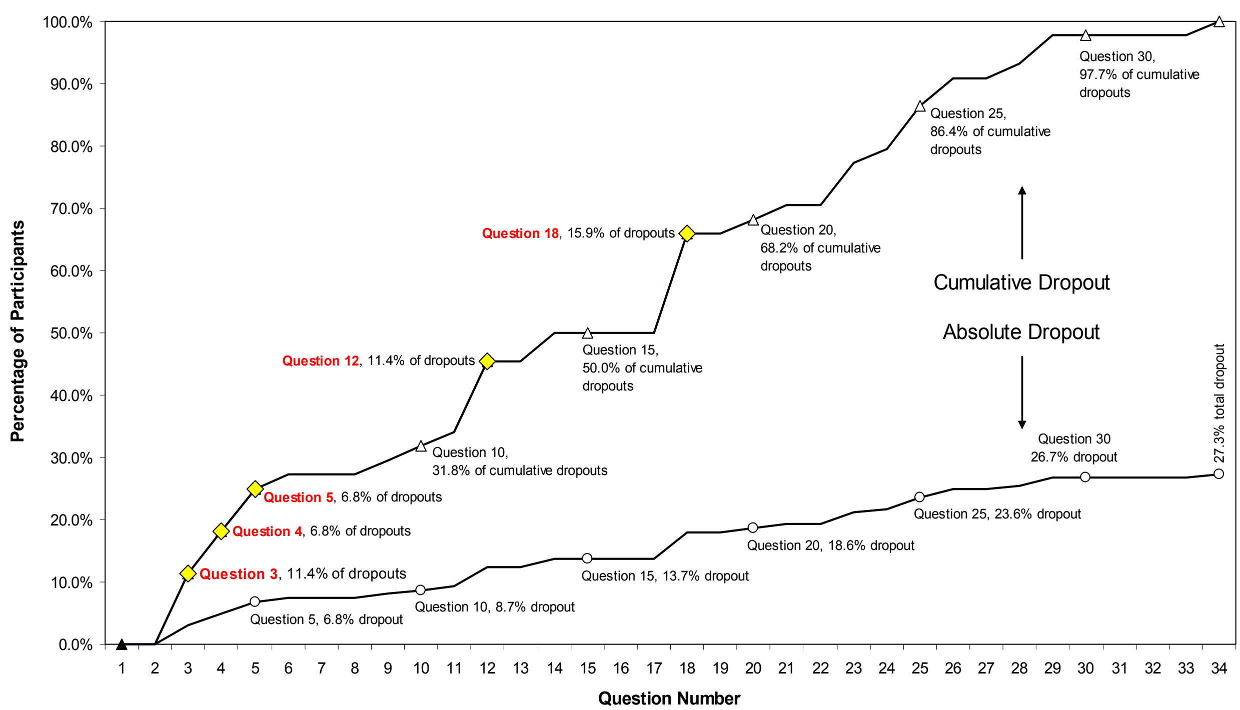


Figure 4. Relationship of participant dropout rates to survey question number at which participants dropped out or stopped answering questions (excludes dropouts who did not answer any questions).

perience using a GIS, this suite of questions was designed to gather background information about a participant's overall experience with a GIS. One possibility for the increased dropout rate associated with this suite of questions is that, taken together, they discouraged some participants, especially those with limited and/or infrequent GIS experience, from continuing with the survey. To test this hypothesis, the experience levels of those participants who dropped out at questions 3, 4 or 5[36] were plotted against those of the non-dropouts (Figure 5). Given the relatively small sample size of question 3-5 dropouts (n=11), it is impossible to draw any concrete conclusions about any relationship between experience level (as expressed in terms of years of experience with a GIS) and the accelerated survey dropout rate noted for questions 3-5. However, as Figure 5 clearly shows, there is a substantially heavier concentration of dropouts in the 0-6 years of experience range (63.7%) than in the 6-10+ years of experience range (27.3%). In contrast, the same distribution among non-dropouts is more evenly split; 55.7% for 0-6 years of experience, versus 42.4% for 6-10+ years of experience. Taken together, these two observations seem to suggest that questions 2-5 did in fact dissuade some participants, especially those with less GIS experience, from continuing with the survey.

Question 12 (C5 in the survey) sought to determine which aspects, if any, of their GIS projects the participants typically documented. A short pick-list of predefined aspects (e.g., the history of

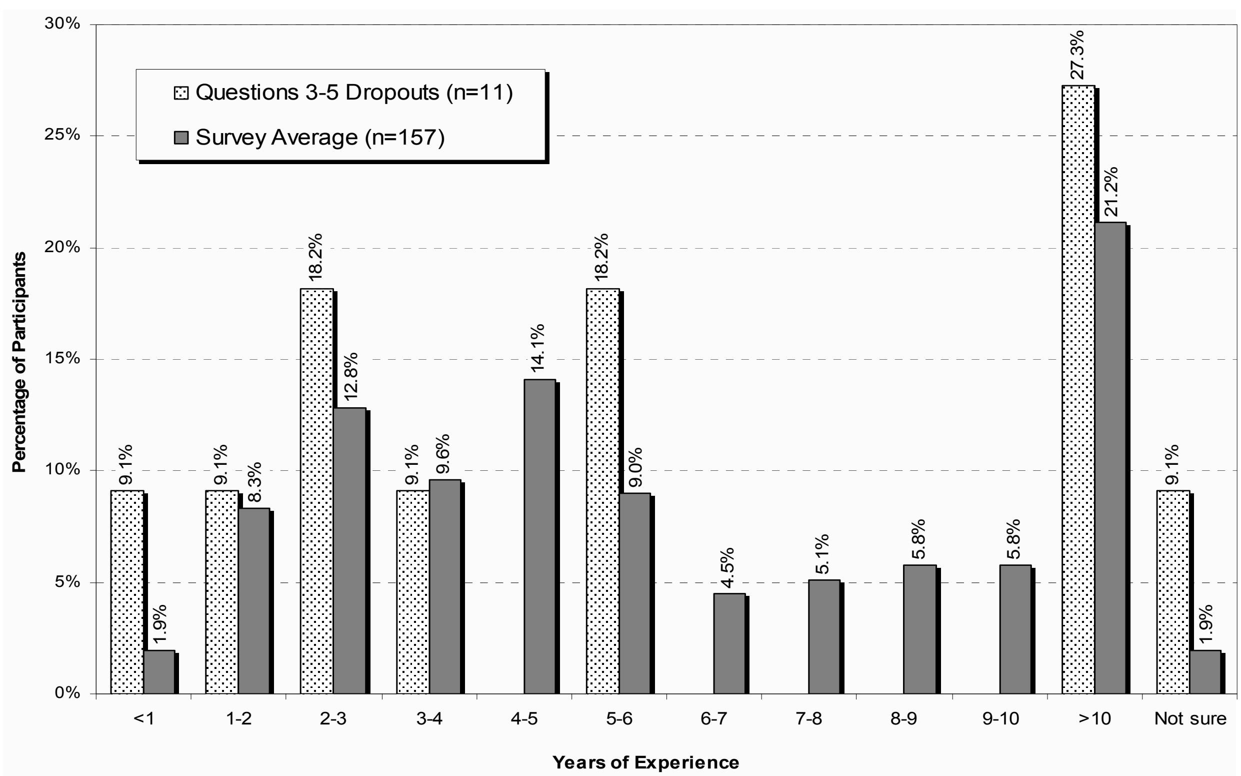


Figure 5. Comparison of experience levels of participants who dropped out of the survey at question 3, 4 or 5, to participants who did not.

[36] To clarify, dropping out at question *X* means that a participant accessed the question but did not answer it and subsequently either terminated the survey at that point by closing their browser, or else continued to access subsequent survey questions without answering them.

modifications made to individual files or datasets, the reasons for modifications, the software used in the project, etc.) was provided and participants were asked to select any and all that applied. To test whether a participant's overall GIS experience and use may have affected the dropout rate for this question, results from questions 2-4 (questions B1-B3 in the survey) for question 12 dropouts were plotted against non-dropouts (Figure 6).

Once again, a small sample size of dropouts (n=5) makes it difficult to draw any definitive conclusions about the relationship between GIS experience level or use and the increased survey dropout rate noted for question 12. In terms of years of experience, there is an even distribution between dropouts with 0-6 years (40%) and those with 6-10+ years (40%). This contrasts slightly with the same distribution among non-dropouts, which is weighted slightly toward participants with 0-6 years of experience (55.7%) over those with 6-10+ years of experience (42.4%). Frequency of use among dropouts is evenly split at 20% for each of the five 'frequency of use' categories. This contrasts rather sharply with the strongly positively skewed distribution (toward more frequent GIS use) noted for non-dropouts.

Perhaps the most interesting results presented in Figure 6 are for the "type of experience" measure. This measure summarizes the results of survey question 4 (B3), which asked participants to indicate with which of 5 general GIS project 'phases' they were, or had in the past been, involved (participants could choose more than one). These 5 phases included: (1) planning, (2) development/design, (3) implementation/data input, (4) data analysis and (5) data preservation. The distribution for this measure among dropouts is noticeably more negatively skewed than is the case for non-dropouts. Here, dropouts tended, overall, to have more experience in GIS project data input, analysis and preservation activities and less experience in GIS project planning and development activities, than did non-dropouts. What this may suggest is that participants involved in project planning and design have a greater awareness of the procedures involved in documenting a GIS project's activities than do participants who are not involved in project planning and design activities. If true, the increase in the dropout rate associated with this question may, in part, be attributable to a sense of inexperience with, or knowledge of, documentary procedures. In other words, some of those who dropped out at this point may have done so because they no longer felt 'qualified' enough to continue taking the survey. On the other hand, some of the dropouts may simply have become disinterested in the survey due to the focus on documentary procedures in the four questions immediately preceding question 12.

Question 18 (C11 in the survey) was a free-text question that asked participants to briefly describe what they considered to be sufficient project documentation and why. Once again, the results from questions 2-4 were plotted, this time for question 18 dropouts and non-dropouts, to test whether a participant's overall GIS experience and use may have been a factor in the increased dropout rate for this question (Figure 7).

The usual cautions regarding a small sample size notwithstanding, there appears, on the whole, to be greater conformity between the distributions of question 18 dropouts and non-dropouts across all three experience measures. The distribution for the "years of experience" measure, for example, is split 57% - 43% between dropouts with 0-6 years and those with 6-10+ years, respectively. This compares quite favourably with the 56% - 42% distribution noted among non-dropouts. These findings suggest that differences in the levels of GIS experience and use among survey participants likely had little, if any, influence on the increased dropout rate associated

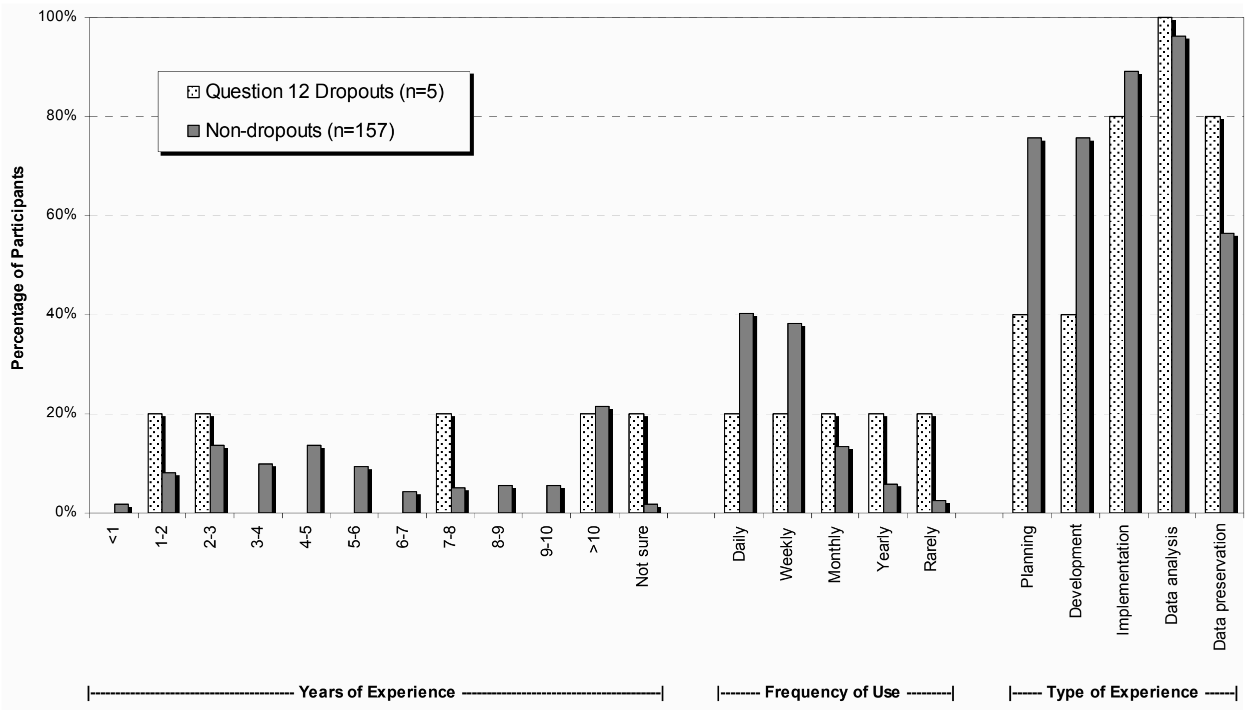


Figure 6. Comparison of various experience-related measures for participants who dropped out of the survey at question 12, to participants who did not.

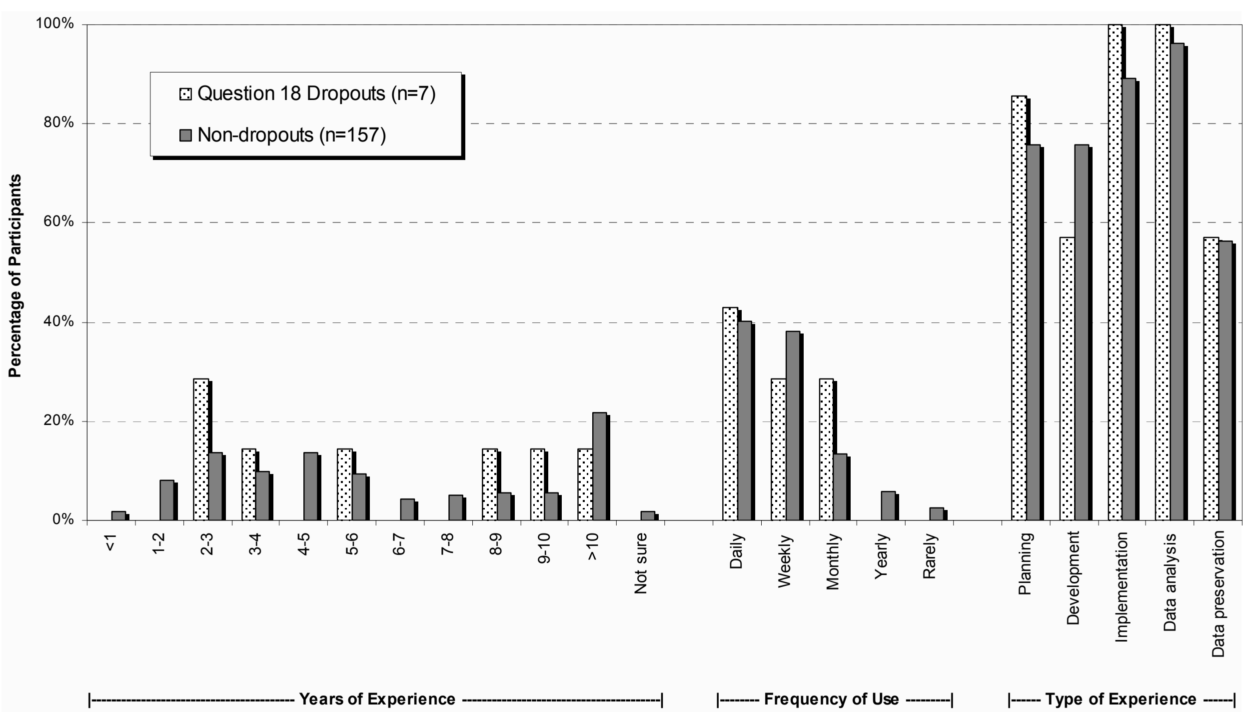


Figure 7. Comparison of various experience-related measures for participants who dropped out of the survey at question 18, to participants who did not.

with question 18. Instead, participants who dropped out at this point in the survey likely did so for reasons other than those related to concerns about their level of GIS experience vis-à-vis their ability to continue to answer the survey questions. Considering the location of the question, at approximately the halfway point in a relatively long survey, coupled with the fact that it was an open-ended free-text question, rather than multiple choice, it is quite possible that general disinterest and/or survey fatigue were the most significant contributing factors for the increased dropout rate at this point in the survey.

In an attempt to assess the potential impact of dropout bias, the answers to survey questions B1-B6 provided by the dropouts and those who completed the survey were analyzed to determine whether there were statistically significant differences between the answers provided by the two groups. Significant differences between these two groups would suggest that participant dropout may have significantly biased the survey results.

The first of the six questions analyzed asked participants to estimate how many years they have been using a GIS. Although the participants who dropped out of the survey appear to have slightly less overall experience (mean ≈ 4-6 years) than those who completed the survey (mean ≈ 5-7 years) (Figure 8), the difference is not significant. In fact, both distributions are significantly correlated (Kendall's tau = 0.508, p=.01). In other words, there is no significant difference in the distribution of GIS experience among the survey dropouts and those who completed the survey.

The second question asked participants to estimate how frequently they use a GIS. As shown in Figure 9, participants who completed the survey appear, on the whole, to use a GIS somewhat more frequently than those who dropped out of the survey. However, rank order analysis of the two distributions indicates that they are, instead, significantly correlated (r_s = 0.872, p <.05), meaning that there is no significant difference in the distribution of the frequency of GIS use among the survey dropouts and those who completed the survey.

The third question asked about the participants' involvement in the various phases of their GIS project(s). Although those participants who dropped out of the survey had slightly less overall experience in the planning and development/design phases and slightly more experience in each of the other three phases (implementation/data entry, data analysis and data preservation) than did those who completed the survey (Figure 10), a chi-square test of independence shows that the difference between the two groups is, in fact, not significant, χ^2 (4, N = 624) = 1.01, p >.05. Participants who completed the survey were no more or less likely to have experience in any particular phase of a GIS project than those participants who dropped out of the survey.

The fourth question asked participants to identify the type(s) of GIS system designs (file processing, extended or hybrid) they typically used. As shown in Figure 11, there appear to be no obvious differences in the types of system designs used by the survey dropouts relative to those participants who completed the survey. Indeed, a chi-square likelihood ratio test of independence (Wilks' G^2)[37] confirms that the difference between the two groups is not significant, Wilks' G^2 (4, N = 236) = 0.77, p >.05. Participants who completed the survey were no more or less likely to use any particular type of GIS system design than those participants who dropped out of the survey.

[37] Wilks' G^2 was used because the counts in several of the cells in the contingency table fell below 5.

In the fifth question, participants were asked to indicate the geographic focus of their research. Although graphical comparison of the data (Figure 12) shows some variability between the two groups, a chi-square likelihood ratio test of independence demonstrates that the variability is not significant, Wilks' G^2 (10, N = 179) = 8.14, p >.05. Participants who completed the survey were no more or less likely to be involved in research in any particular geographic area than those participants who dropped out of the survey.

Finally, the six question asked participants to identify their current professional archaeological affiliation(s). As shown in Figure 13, there are some differences between the two groups, with participants who completed the survey apparently more likely than those who did not to be affiliated with public sector cultural resource management (CRM), but less likely to be affiliated with private sector CRM. However, a chi-square likelihood ratio test of independence indicates that these differences are not significant, Wilks' G^2 (4, N = 201) = 2.56, p >.05. Participants who completed the survey were no more or less likely to have any particular professional affiliation than those participants who dropped out of the survey.

In summary, statistical tests indicate that there are no significant differences in the answers to these six questions between the participants who completed the survey and those who did not. Although these tests do not *prove* that there are no significant differences between the two groups, or more particularly, that the survey results have not significantly been biased by the effects of participant dropout, they do seem, nonetheless, to provide strong evidence that such is the case.

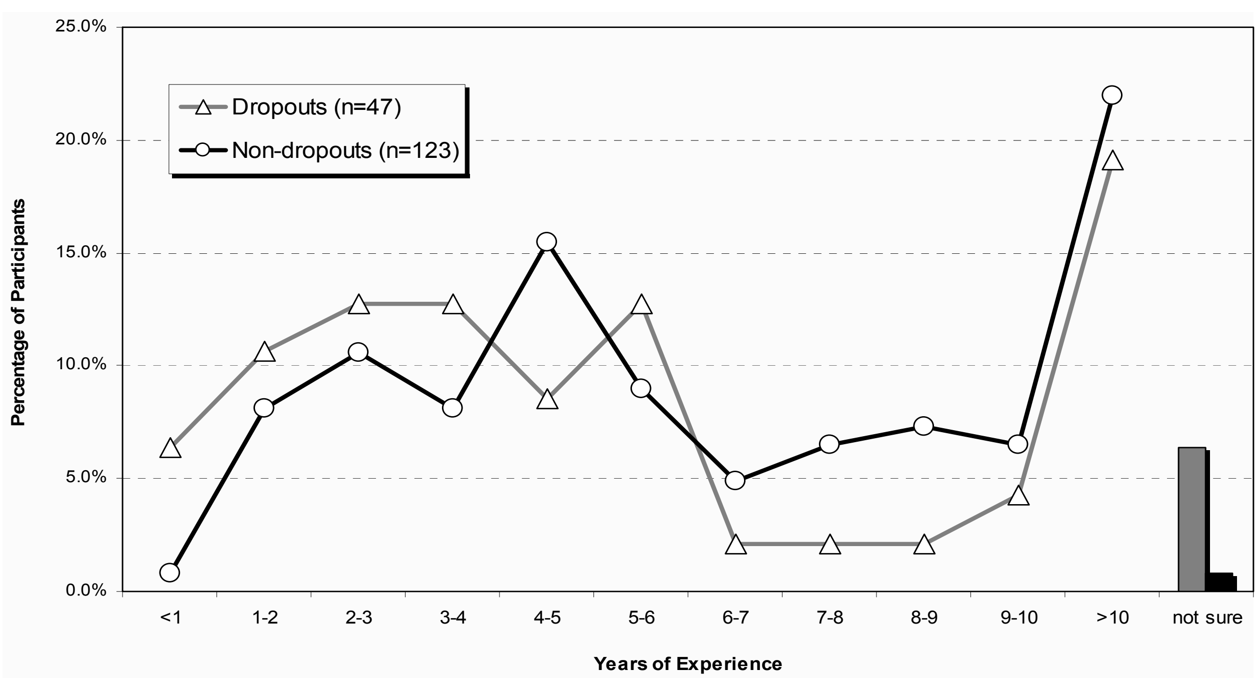


Figure 8. Comparison of the relative percentages of years of GIS experience of participants who completed the survey with those who dropped out.

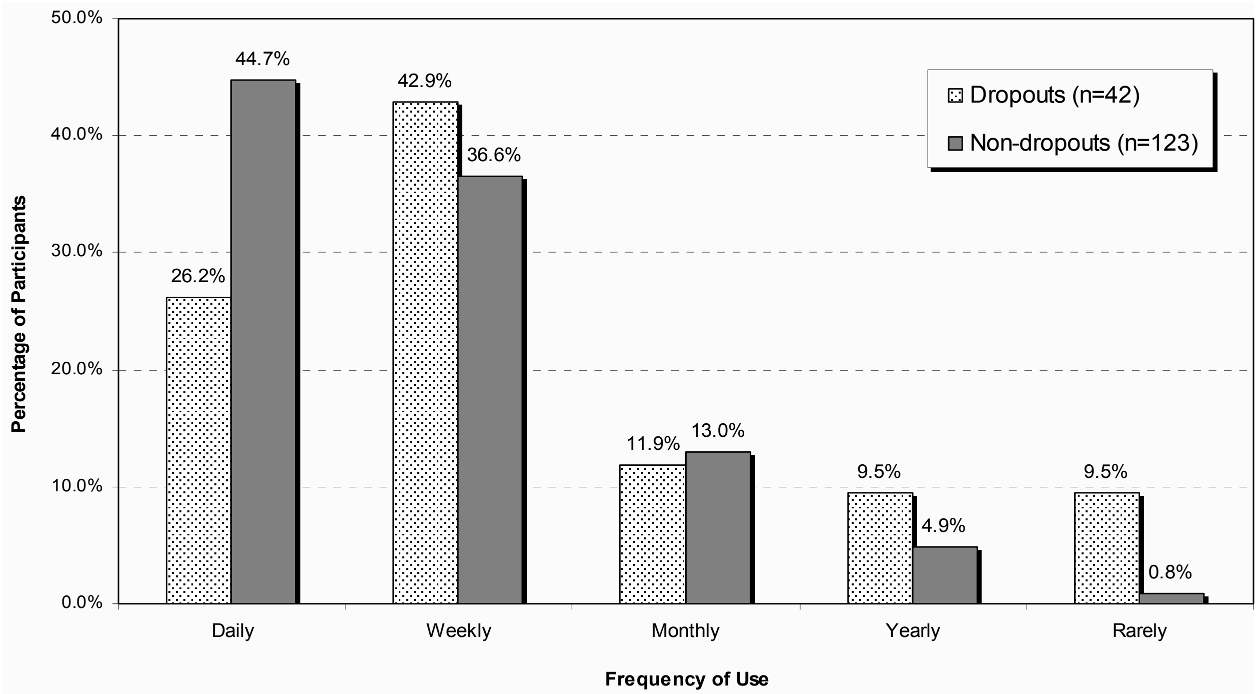


Figure 9. Comparison of the relative percentages of the frequency of GIS use by participants who completed the survey with those who dropped out.

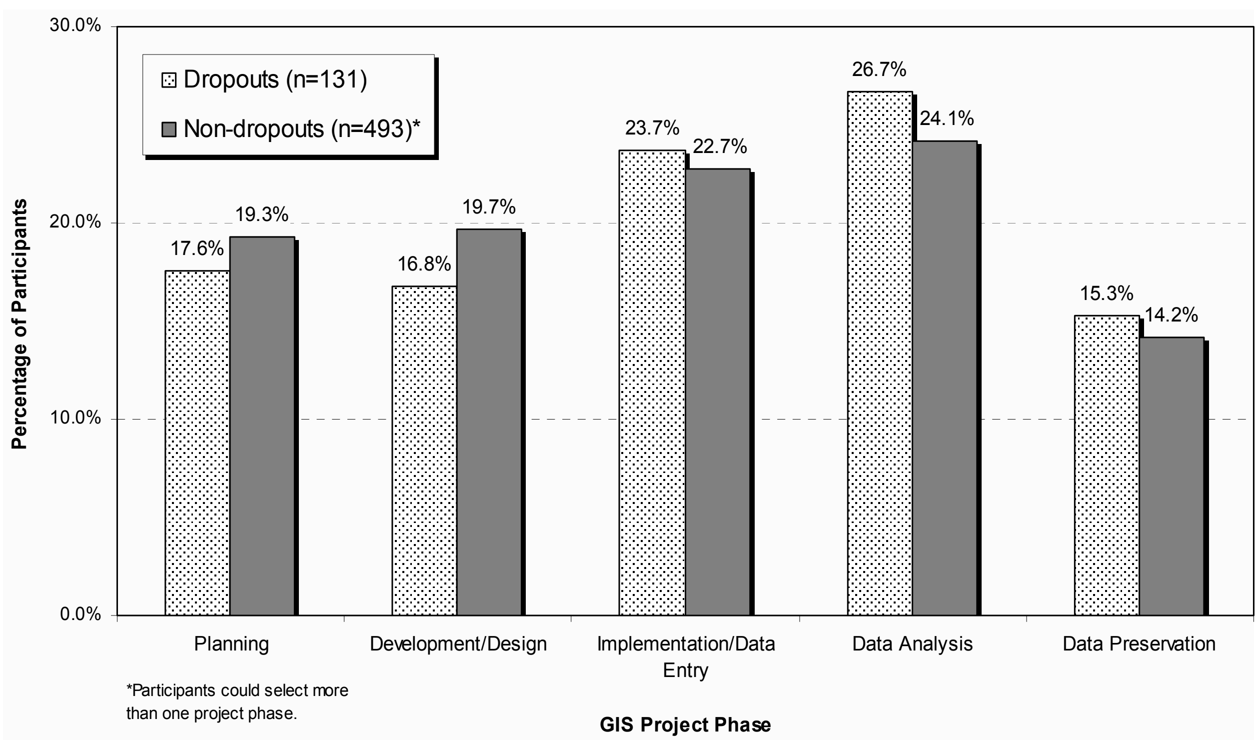


Figure 10. Comparison of the relative percentages of GIS project phase experience of participants who completed the survey with those who dropped out.

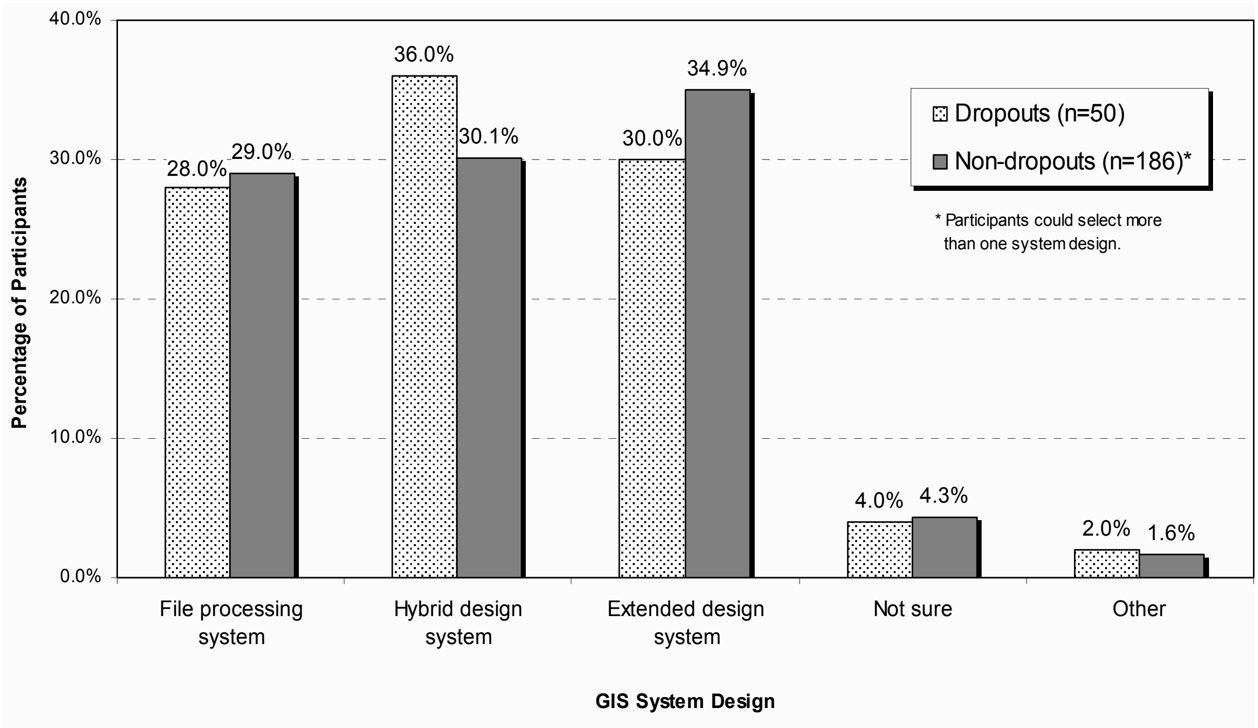


Figure 11. Comparison of the relative percentages of GIS system designs used by participants who completed the survey with those who dropped out.

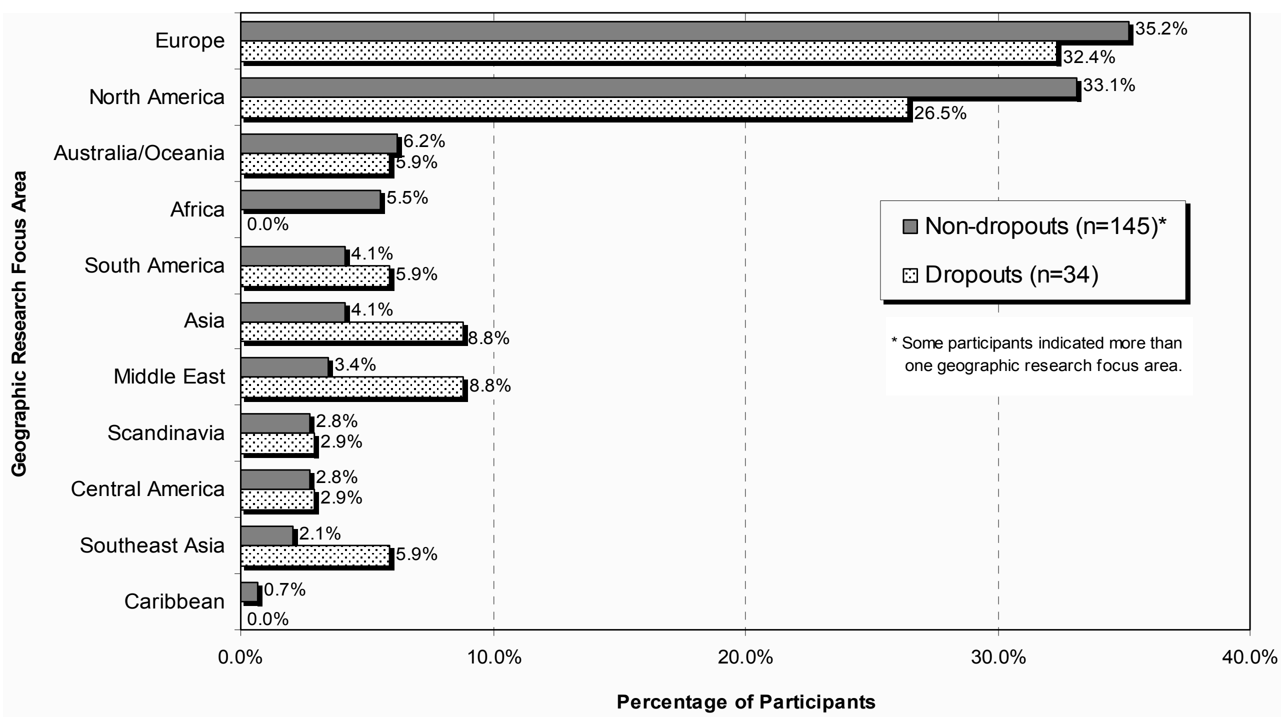


Figure 12. Comparison of the relative percentages of geographic research foci of participants who completed the survey with those who dropped out.

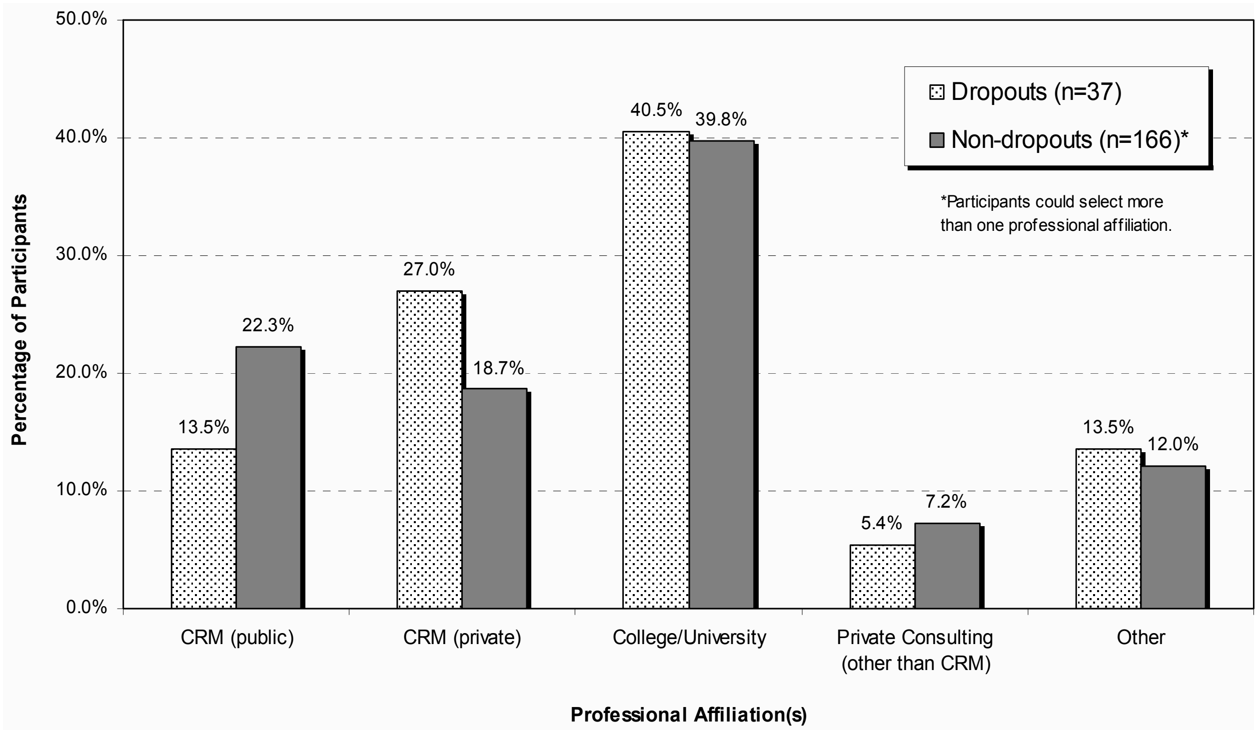


Figure 13. Comparison of the relative percentages of professional affiliations of participants who completed the survey with those who dropped out.

4.5.1.2 Survey Participant Behaviour

For a variety of reasons, some survey participants may provide inconsistent or unreliable responses to survey questions. Participants may, for example, mark their answers randomly (perhaps without even reading the questions), just to get the survey over with. However, considering that participation in the GIS survey was strictly voluntary, and that bored or otherwise disengaged participants could exit the survey at any time simply by closing their Web browser, such a scenario seems highly unlikely here. Inconsistent or unreliable responses can also result from problems reading or understanding the questions. Given that the survey was in English, which undoubtedly was not the mother tongue of many of the participants, this may have been an issue for some of the participants. However, steps were taken during the design of the survey to help mitigate the effect of this issue. For example, following recommendations in the survey design literature, questions were phrased in the most understandable way possible using common GIS terminology whenever possible. As well, key concepts and terms with the potential for being perceived differently by participants on linguistic or cultural grounds were "standardized" via brief definitions and/or examples to help harmonize the responses. Considering that the majority of participants were from English-speaking countries, coupled with the fact that all but one of those who were not provided their free-text answers and supplementary comments in English, suggests that question comprehension (or incomprehension) was not a significant factor. As well, it is important to emphasize that, because comment boxes were provided for each question, participants were able to qualify any answer, if necessary. Another possibility is that participants may consciously give false or insincere responses just for fun, to try to "screw up" the survey. Again, however, this seems unlikely for this survey considering its voluntary nature, together with the fact that it was selectively directed toward researchers with GIS experience. Finally, survey participants may provide biased or unreliable responses simply because they do not feel comfortable divulging information they perceive to be "personal" or "sensitive." However, the potential for any such bias in this case is presumed to have been mitigated by the anonymous nature of the survey.

There are various formal means for assessing the effects of participant behaviour on the reliability of survey data, including *test-retest*, *equivalent-form* and *internal consistency* tests. The *test-retest* approach involves administering the same survey to the same group of participants at two different points in time, with the degree to which both administrations are in agreement serving as a measure of the reliability of the survey. An *equivalent-form* test requires the creation and administration of two different surveys designed to assess identical constructs using differently-worded questions. The degree of agreement, or correlation, between the results of the two surveys serves as a measure of their equivalent-form reliability. *Internal consistency* is a measure of the concordance between two variables that measure the same general attribute. In the context of a survey, internal consistency refers to the degree to which a participant's answers to two related questions are in agreement. For example, poor internal consistency would be indicated in a situation where a participant answered "Yes" to a question that asked, "Do you record metadata?" and "Not applicable" to another question that asked, "What types of information do you include in your metadata?" Unfortunately, assessing the reliability of the GIS survey by any of these means poses problems. The practical and economic considerations inherent in the test-retest and equivalent-form approaches precluded them as viable options. Although more practically and economically feasible, internal consistency tests, such as Cronbach's alpha, require a greater degree of structural consistency to the survey answers (e.g., an identical pool of available answers from which to choose for each question) than exists in the GIS survey.

4.5.1.3 Summary of Survey Reliability

Although no formal assessments of data reliability were applied to the survey results, several informal measures, in addition to the arguments cited above, leave no reason to suspect that the results are unreliable. First, the mean completion time of just under 30 minutes compares favourably with the mean completion time noted during a pilot test of the survey (from which the estimated completion time of 30 minutes was derived). This suggests that the survey participants did in fact take the time to read and process the questions and provide truthful and sincere answers, based on the assumption that someone choosing answers at random is much less likely to take the time to read the questions. Related to this assumption is the fact that all of the participants took the time to answer the free-text questions and/or provide additional comments about their answers to particular questions. Second, the architecture of the on-line survey greatly reduced the likelihood of inconsistent answers (whether intentional or accidental). As noted earlier, the survey questions were presented one at a time, with participants required to click on a "submit" button to advance to the next question. One of the primary reasons for imposing this restriction was to enable contextualized navigation of the survey questions based on participant responses. For example, in question C5, participants were asked, "When documenting your GIS projects, which of the following aspects do you typically document?" Those who answered, "Not applicable (documentation typically is not created for my GIS projects)," automatically were skipped ahead to question C11, bypassing questions C6-C10, all of which dealt with various aspects of project documentation procedures. Third, the survey data for each participant were subjected to a manual test for inconsistencies. The test involved comparing responses to questions C5 and C11. Question C5 asked participants to identify which aspects of their GIS projects they typically documented, while question C11 was a free-text question that asked participants to comment on what they considered to be sufficient documentation and why. Of particular interest were the C11 comments of those participants who indicated in question C5 that documentation typically was not created for their GIS projects. Of those participants who answered "not applicable" to question C5, 40% chose not to answer question C11 (perhaps because they felt it did not apply to them), 28% commented on why they typically do not document their projects, while the remaining 32% discussed, in a hypothetical way, what they thought should be documented under ideal conditions. In other words, there were no apparent inconsistencies between the participants' answers to questions C5 and C11.

4.5.2 Validity

In contrast to reliability, which measures how consistently an instrument, such as a survey, measures whatever it is measuring, validity is concerned with the degree to which an instrument actually measures what it is intended to measure. Thus, in the case of the GIS survey, validity may be defined as the degree to which the survey actually measures or collects data about the digital recordkeeping habits of GIS archaeologists. In this context, the fundamental validity question then becomes, what is the match between the information provided by the survey and what it was intended to show (i.e., the inferences to be drawn from the results about the recordkeeping habits of GIS archaeologists)? It is important to bear in mind the distinction between reliability and validity and how these two measures may or may not influence each other. Although low reliability certainly will limit one's ability to determine validity, high reliability does not guarantee high validity. In other words, reliability is necessary, but not sufficient, for validity.[38] In fact, the validity of a measure depends on the reliability, appropriateness and the completeness of its underlying data. It can be argued that all surveys inherently lack some degree of validity because of their incomplete coverage. As well, it must be acknowledged that in all but the most straightforward and factual of surveys, a survey will generate "subjective" data of varying validity depending, in part, on the formulation of its questions. Even a straightforward, factual survey, despite perhaps resulting in more accurate and "objective" data, nevertheless may still produce data that are selective and unrepresentative of the concept(s) it is attempting to measure.

Just as with reliability, there are various ways of both conceptualising and measuring validity (e.g., face validity, content validity, construct validity, etc.). Regardless the specific concept of validity in question, its assessment invariably involves some degree of professional judgement, a process that is itself always subjective and predisposed to varying validity. In most cases, validity is a much more difficult variable to measure and assess than reliability. The most common approach is to find an independent, objective measure of the survey concept being measured, a measure whose validity already is known or accepted *a priori* (i.e., a "golden

[38] The classic example of this concept is the "relationship" between head size and intelligence. While head size can be reliably measured, it cannot be used as a valid index of intelligence.

standard”), and compare it to the survey responses. Unfortunately, this approach cannot be used to assess the validity of the GIS survey since no such “golden standards” exist for the concepts and issues addressed in the survey. Because of the difficulties of measuring validity in the absence of such standards, no attempt is made here to determine the validity of the GIS survey data. Instead, the issue of validity assessment is presented here merely as a goal to be pursued: (1) by formulating clear survey questions and aiming them accurately at the kinds of information required to address the studied concepts and (2) by defining and describing the studied concepts, in relation to the resulting survey data, as accurately as possible and by comparing the findings to existing GIS recordkeeping data whenever possible.

5.0 SUMMARY OF RESULTS

5.1 Section A: Introduction

This section consisted of a single, introductory, free-text question in which participants were asked to briefly describe what they considered a GIS to be, including what they felt were its most important, significant and/or distinguishing components and functions. As noted above, the primary purpose for asking this question was to assess what importance, if any, the participants assigned to their role (i.e., the human operator) in the overall geographic information system.

The vast majority of participants were in agreement in describing a GIS as a "tool" used for any number of purposes including data integration and database management, spatial-temporal analysis and visualization, iterative mapping, etc. Among its many features, the most important and distinguishing, as cited by a majority of the participants, was its ability to incorporate, manage, manipulate, query and analyze visual and statistical relationships between large and often disparate sets of digital spatial-temporal data to, in the words of one participant, "reveal patterning relevant to research and management goals." The ability of a GIS to produce high-quality outputs of these analyses was also seen as an important function.

Participant's responses were, however, less congruent when describing the actual components that delineate a GIS. For some, software (or an integrated suite of software) was all that was required to delineate a GIS, while for others it was software and data. One participant defined a GIS both as data and as the technology used to manipulate those data; on the one hand describing a GIS as "a set of geographically-based data that, taken all together, make some kind of sense for any specific purpose (i.e., research)," while on the other hand stating that "I could also agree with the idea that [a] GIS is also the technology used to manage and analyze those data." Still others described a GIS as a combination of software, hardware (computers, digitizers, remote sensing equipment, GPSs, etc.) and, in many cases, data. Several participants characterized GIS as more of a methodological approach or a type of analytical technique, placing less emphasis on its software, hardware and data aspects. One participant, for example, described a GIS from a more theoretical perspective, emphasizing that a GIS was "not simply a bit of software," but rather "a state of mind…[drawing] together strands regarding spatial theory from a number of disciplines," resulting in the "implementation of GIS theory using some software and some data." Although this last participant hints at the role of the human investigator as an integral component of a GIS, it is interesting that only 3 of the 151 participants who answered this question actually made any explicit mention of the significance of the human component in the system. One considered a GIS to consist of software, hardware, human resources and data, while emphasizing that "the most important component of a GIS is the human resources managing or utilizing the system." Another similarly described a GIS as consisting "of a computer component, a spatial component, a data component, and most importantly a human component." The third likened GIS to "a tool, like a trowel, a shovel, or a notebook," while cautioning that "like any tool, the user must know how to use it and what the proper uses are."

Although few, if any, of the participants likely would disagree that human operators are a necessary and important component of any GIS, it is interesting that so few included them in their descriptions of GISs and important GIS components. No doubt this is due, in part, to the various ways in which the participants use and interact with their GISs, with some participants more heavily involved in those aspects of their projects in which the human element perhaps is perceived as less significant, such as data entry or the output of maps, relative to those aspects in which the significance of the human element may be more readily apparent, such as GIS planning, development and data analysis. On the other hand, perhaps the human component is seen by most as a "given" and therefore generally not worthy of mention. Then again, the same argument could as easily be made for software, hardware and data, yet each of these elements were specifically identified by a majority of the participants. Perhaps, instead, this under-representation of the human element in GISs points to the presence of a certain degree of disconnect between the actual influence that the participants wield on their GIS projects and their perception of that influence. Several of the participants admitted feeling somewhat overwhelmed by the analytical potential of their GISs, due, primarily, to a lack of formal GIS training. In the words of one participant, "ideally [it is] a hardcore spatial-analytical tool, although in practice it tends to get used mostly for illustrative mapping purposes." Adding credence to this notion of a degree of disconnect between the tool and many of its users are the results of Gourad's 1998 GIS survey in which he noted a fairly high level of discordance between his participants' knowledge of common GIS analytical pitfalls and limitations and his participants' responses to them. In fact, in his concluding remarks, Gourad states that, "the survey results demonstrated that many concepts which directly or indirectly affect the quality of GIS output, are either still unknown to most archaeologists or are not considered serious enough to warrant changes in project designs."[39] If true, this could potentially have ramifications with regard to the broader issue of the long-term preservation of digital GIS records. In fact, it is possible that the more alienation, disconnection and/or intimidation one feels toward one's GIS projects, the less likely one may be to feel the sense of obligation, initiative and/or competence necessary to effectively address the project's long-term preservation requirements.

5.2 Section B: GIS Experience/Background

5.2.1 Geographic Distribution

Participants were asked to indicate the geographic focus their GIS research at whatever scale they felt was most appropriate (e.g., North America, United States, American Southwest, Arizona, etc.) and, optionally, to specify their country of current employment. As summarized in Figure 14, the geographic distribution of survey participants spanned at least 30 countries[40] from 6 continents. Of those who specified their country of employment, the largest percentage (36.6%) were from the United States, followed by the United Kingdom (13.1%), Australia (8.5%) and Canada (4.6%). The overall representation of the remaining 27 countries combined (37.2%) was roughly equivalent to that of the United States. Given the relative concentration of GIS archaeologists in the top 4 countries, coupled with the fact that the survey was in English, these results are not surprising. Moreover, it is interesting to note that while these results are similar to those associated with Gourad's survey with respect to the top 4 countries (although the relative position of Canada and Australia are transposed in the current survey), there are marked differences in the actual distribution of countries represented in the two surveys, with 9 of the 25

[39] Gourad, *supra* note 22, p. 62.
[40] The actual number of countries is slightly higher since the countries of the United Kingdom have been lumped together.

countries in Gourad's survey not represented in the current survey and 14 of the 30 countries in the current survey not represented in Gourad's survey (Figure 15).

As would be expected, the geographic distribution of research foci among the participants generally approximates the geographic distribution of the participants, with the heaviest research concentration occurring in Europe (35.3%), followed closely by North America (31.8%) (Figure 16). At only 5.9%, Australia/Oceania ranks a distant third, only slightly ahead of Southeast Asia and the Middle East (both at 5.3%).

5.2.2 Current Professional Archaeological Affiliation

As shown in Figure 17, just over half (52.2) of the participants are affiliated with a college or university, while a nearly equal number (52.9%) are affiliated in some way with either public or private Cultural Resource Management (CRM). Nearly 15% of the participants are affiliated with various other organizations, including in particular, museums and private research institutes.

5.2.3 Years of Experience

Gourad found that more than 72% of his participants had more than two years of experience with GIS, which he considers to be "a reasonable amount of time to familiarize oneself with the technology" (1999: 38). His highest concentration of participants was in the 4-5 years of experience range, with an overall average experience of >4.2 years. In contrast, more than 89% of the participants in the current survey have more than two years of GIS experience, with an average experience of >6.1 years,[41] and with the highest concentration of participants (21.2%) falling within the >10 years of experience range Figure 18). Rank order analysis of the data from the two surveys indicates that the two distributions are not statistically correlated ($r_s = -0.0357$). This result is not entirely unexpected. Indeed, the shift toward participants with more GIS experience noted in the current survey may simply reflect the fact that six years separates the two surveys, such that any participants from the current survey who may also have participated in Gourad's earlier survey would have substantially more GIS experience. However, the differences in the geographic distribution of the participants involved in both surveys suggests that the difference observed in the distribution of experience levels between the participants in the two surveys is in fact real.

[41] Average experience was estimated by multiplying the total number of participants in each experience level category by the mean of the temporal spread of each category, with the resulting values for each category summed and then divided by the total number of participants. For example, participants in the <1 year category were multiplied by 0.5, those in the 1-2 years category by 1.5, etc. Given the substantial number of individuals in the >10 category, it is highly likely that the resulting average is a conservative estimate. This is so because experience levels were artificially capped at >10 in the survey. Thus, the true experience levels of any individuals in the >10 category with 11 or more years of experience will be underrepresented since such individuals are lumped into the >10 category and multiplied by 10.5. Also, it is noted that the estimate for Gourad's data was further complicated by the fact that his experience categories are not consistently demarcated (e.g., 1-2, 2-3, 4-5, 6-7, etc., vs. 1-2, 2-3, 3-4, 4-5, 5-6, 6-7, etc.) (see Gourad, *supra* note 22, Figure 4.3, p. 39).

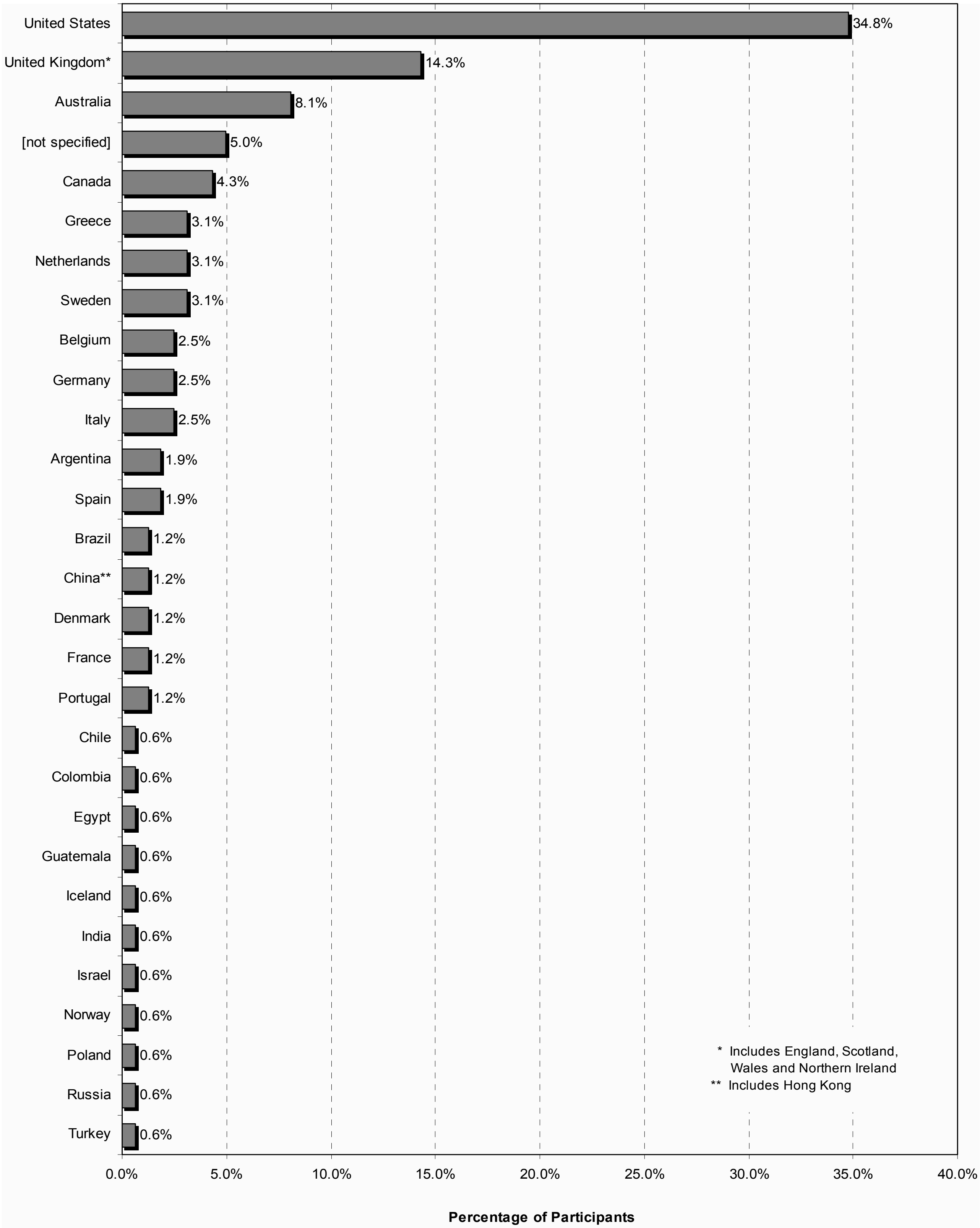


Figure 14. Distribution of survey participants by country (n=157).

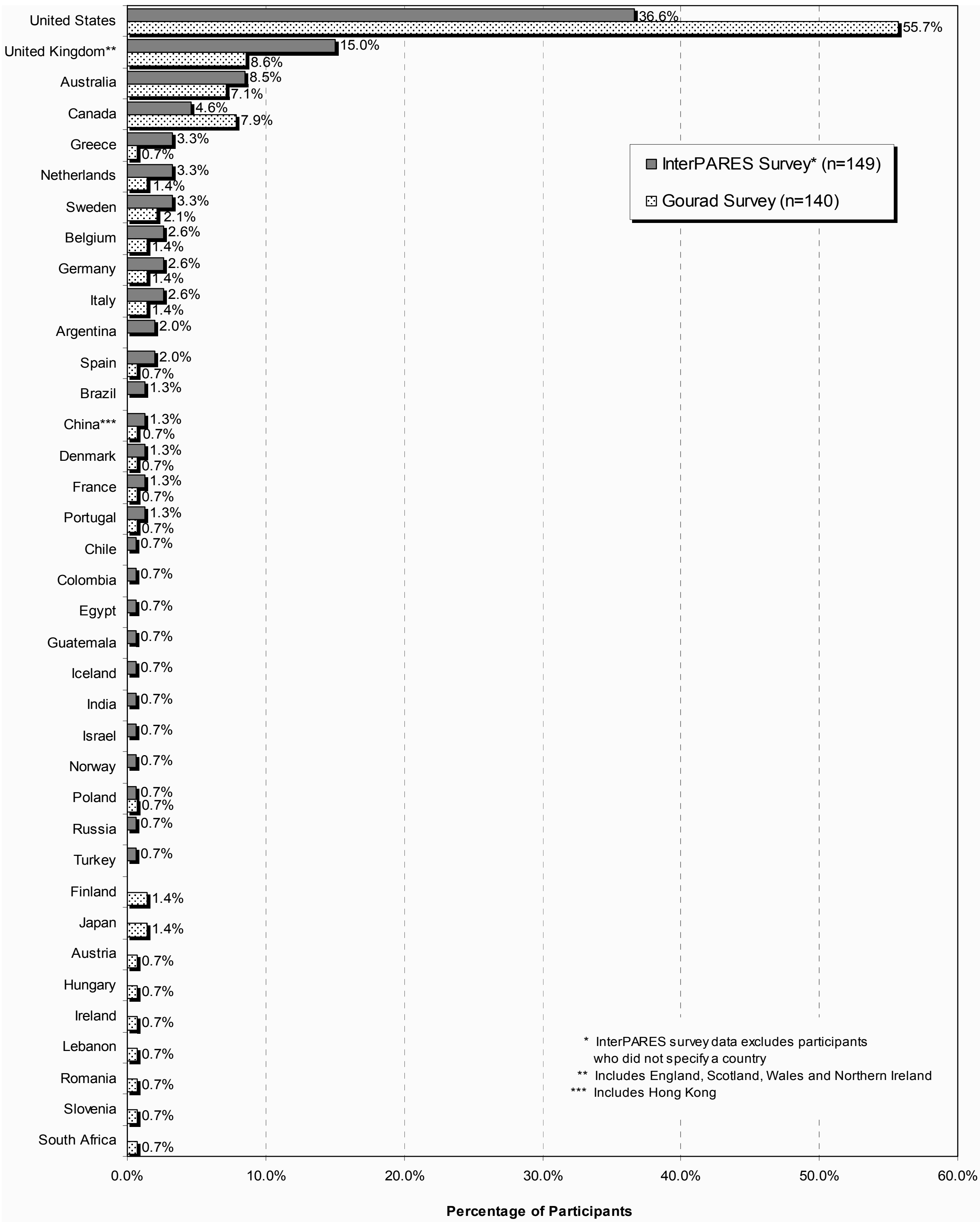


Figure 15. Distribution of survey participants by country: InterPARES GIS survey vs. Gourad's 1998 GIS survey.

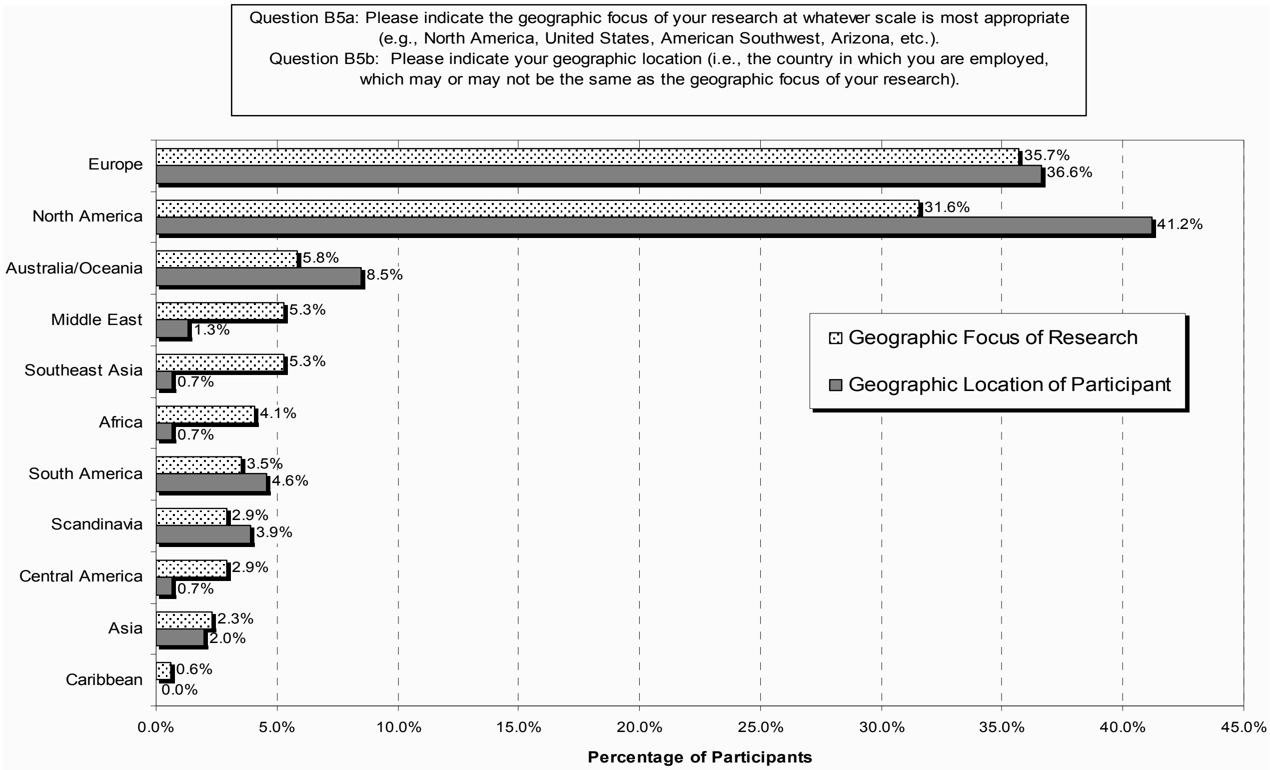


Figure 16. Geographic focus of research vs. geographic location of survey participants (n=149).

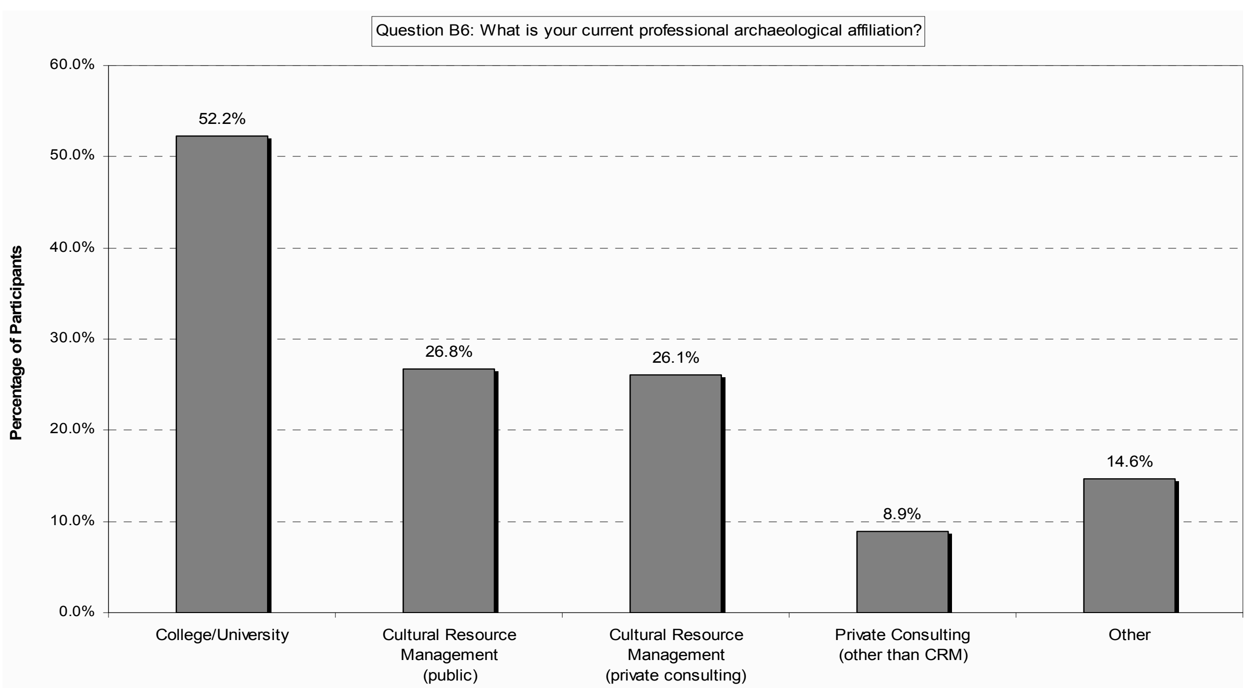


Figure 17. Professional affiliation(s) of survey participants (n=157).

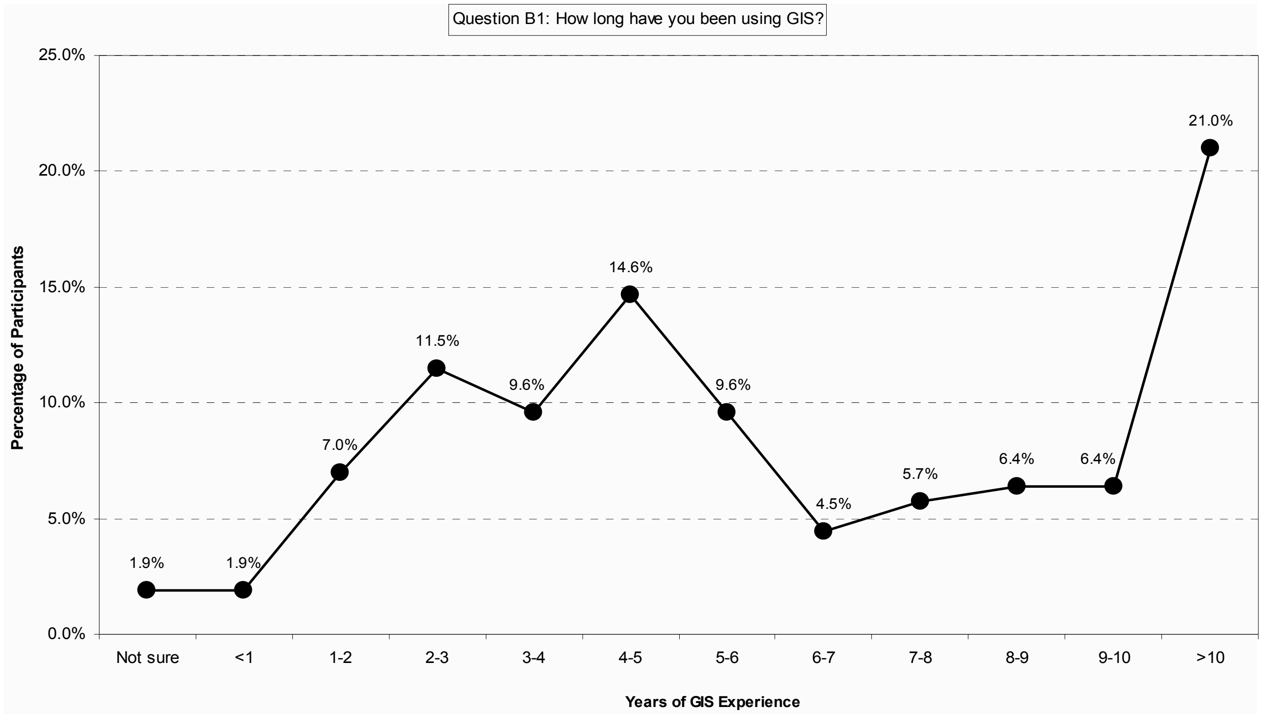


Figure 18. Number of years of GIS experience of survey participants (n=157).

5.2.4 Frequency of Use

The frequency of GIS use among the participants in the current survey was high, with 40.1% and 38.2% using GIS on a daily and weekly basis, respectively (Figure 19). Gourad also noted a high frequency of GIS use among the participants in his survey, with 29.3% and 31.4% using GIS on a daily and weekly basis, respectively.[42] In fact, the frequency of GIS use among the participants in the InterPARES survey shows a significant positive correlation ($r_s = 0.90$) with that noted for the participants in Gourad's survey.

5.2.5 Project Phase Experience

Question B3 sought to discern the nature of the participants' involvement in each of several discrete procedural phases of GIS research, including: A. Planning, B. Development/Design, C. Implementation/Data entry, D. Data analysis and E. Data preservation. In addition to specifying which phase(s) they were, or have been, involved with, participants with experience in two or more phases were asked to indicate with which phases they had the most and the least experience. A couple of the participants expressed some reservations about their ability to distinguish among these five discrete phases within their GIS projects. One participant, for

[42] The reader is cautioned that the value of 41% that Gourad cites in his survey for the frequency of daily GIS use among his survey participants actually is the *raw* value of participants reporting daily use, not the percentage of participants (see Gourad, *supra* note 22, p. 38).

example, commented that, "design is preservation—meaning that it's an integrated part in the design process of a large-scale GIS, rather than [an] add-on or afterthought." Another noted that, "it is impossible to rank the above in a sense as each relies on each other and it makes an assumption that one should have 'GIS Projects' as an independent activity; in a practical case all our projects utilize GIS, it is the project that comes first, the GIS is simply a day to day management tool which offers particular functionality." On the whole, however, the vast majority of participants appear to have been able to satisfactorily distinguish among the five phases within the context of their GIS research projects.

Overall involvement was greater than 75% in each of the procedural phases, with the notable exception of data preservation (56.4%), with the participants identifying the implementation/data entry and data analysis phases as the two phases with which they have had the most experience (Figure 20). Notably, 94.9% of participants have been involved in two or more phases, while nearly half (46.8%) have been involved in all five phases. The considerable breadth of experience indicated by these results is in keeping with the generally high level of temporal GIS experience noted earlier.

Of particular interest here with respect to the issue of the long-term preservation of digital records is the fact that only slightly more than half (56.4%) of the participants have had any experience with the preservation phase of their GIS projects, of whom only 4.5% consider data preservation to be the phase with which they have had the most experience. Even more informative is the fact that 60.2% of the participants who listed data preservation as one of the phases in which they have been involved expressly identified this phase as the one with which they have had the least experience. These are by far the lowest and highest percentages noted for "most" and "least" experience for any of the five phases.

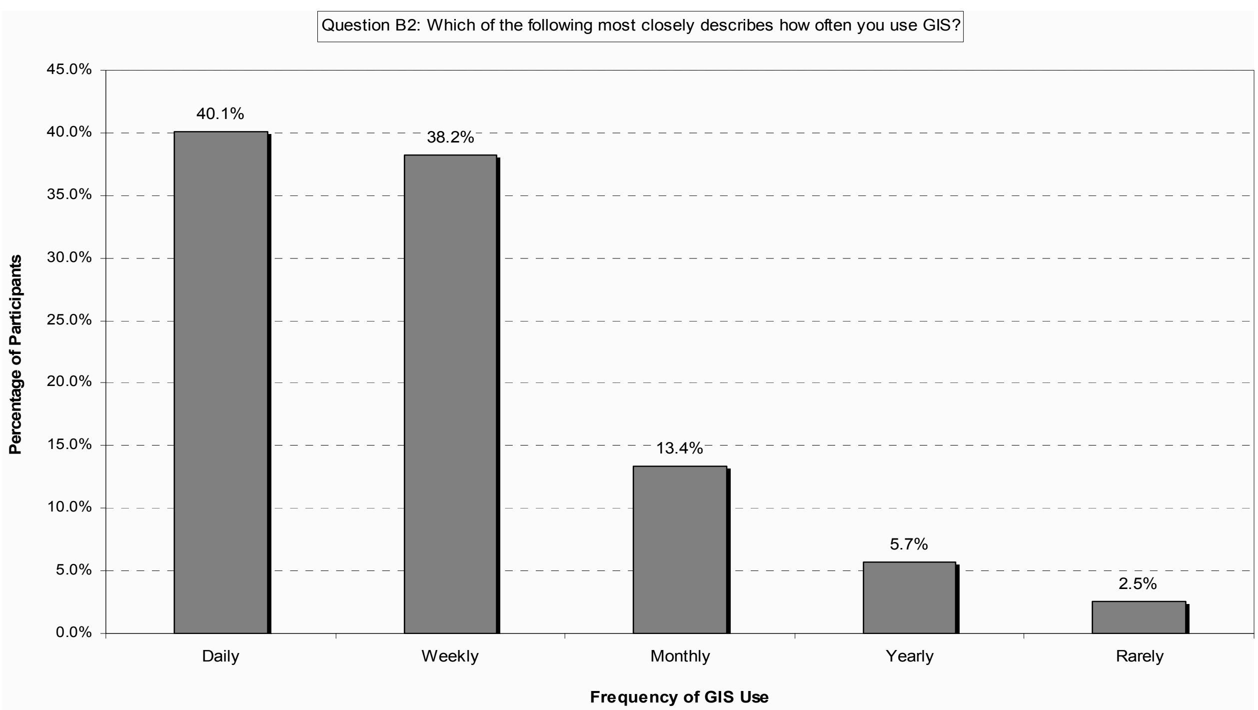


Figure 19. Frequency of GIS use by survey participants (n=157).

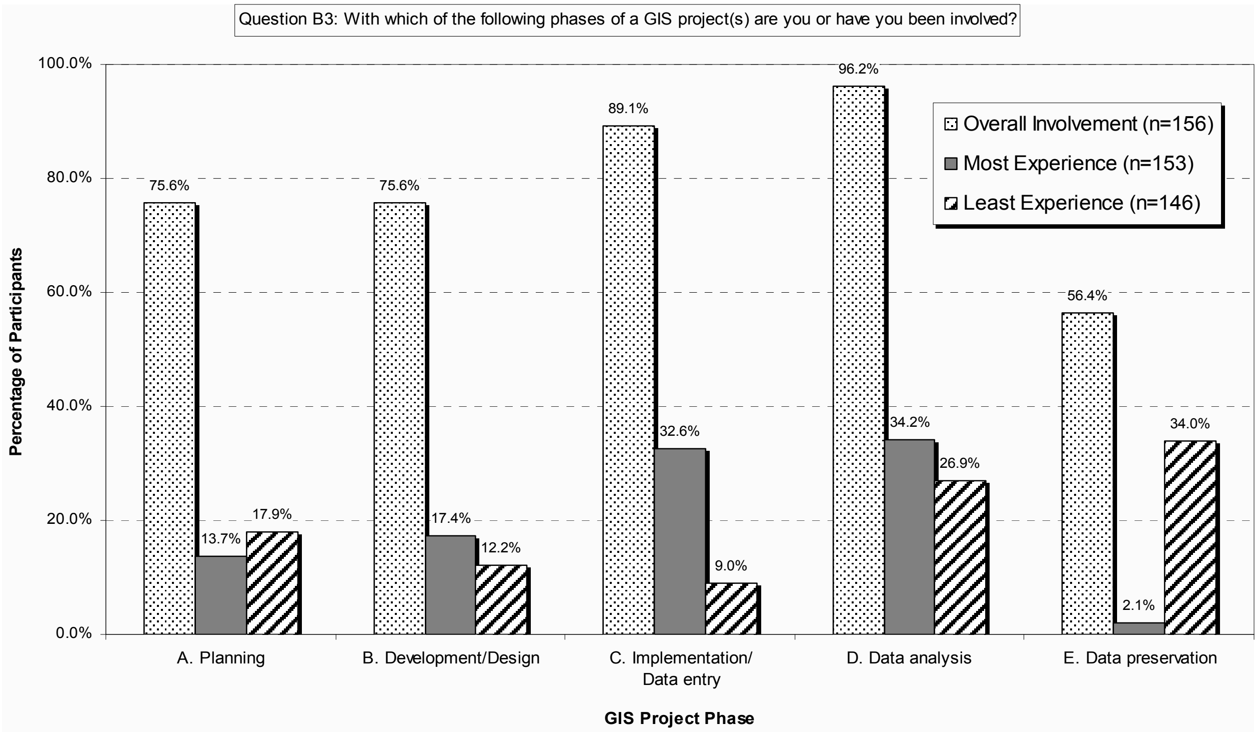


Figure 20. Involvement of survey participants in various key GIS project phases.

5.2.6 GIS Design(s) Used

Question B4 asked participants to specify the type(s) of GIS system designs they typically used. This issue will likely have important consequences with regard to preservation concerns because of the very different ways in which each system design stores, references and manipulates data and files. The three major system designs include: (1) extended design systems, (2) file processing systems and (3) hybrid design systems. In extended design systems, both geographical and attribute data are stored together in a single database. File processing systems, on the other hand, store each data set and function as separate files which are dynamically linked together during analytical processing. Finally, hybrid design systems store the attribute data in a conventional database management system (DBMS) with separate software used to manage the geographical data. As shown in Figure 21, use appears to be fairly evenly divided among the three system designs. It also is worth noting that 36% of the participants indicated that they actively use more than one system design.

5.3 Section C: File Management/Documentation Procedures

5.3.1 File Naming Procedures

Just under half (48.9%) of the participants always or usually use a standardized and/or documented procedure for naming their GIS-related digital files (Figure 22). For most of the remaining participants (31.0%), the diligence with which this task is pursued varies from one project to the next. A number of participants commented that they currently are in the process of deciding upon a standardized file-naming procedure, while others noted that their standardized procedures tend to break down over time, due either to the complexity and size of the project or to the inability to convince fellow colleagues to follow the procedures. As one participant put it, "I always TRY and am almost always forced into non-compliance. My solution has been to build audit tools to fix anomalies after the fact." Still others emphasized that standardized procedures are used only for "files to be retained," or for "crucial data files that are included in the end delivery to a customer," whereas intermediate, single-use and temporary files are, as a general rule, not subjected to standardized naming procedures. Several participants emphasized that they use descriptive naming conventions chosen to convey intuitive meaning rather than using systematic, yet arbitrary identifiers (e.g., sequential file numbers) with no inherent correlation to the contents of the file. Only one participant admitted to never using standardized naming procedures, noting that, "as a researcher, I am the only user of the files, so I have no need to follow a procedure."

5.3.2 File Version Control

As one participant noted, file version control is one of the essential file management tasks required to help ensure effective and accurate rollback when necessary. Despite this, nearly half (43.5%) of all participants indicated that no standardized version control system is used to help distinguish among different versions of the same file (Figure 23). A common justification among several of these participants was that because only one person was responsible for any given project file, file versioning (presumably) was not an important concern. However, while a single-user environment may indeed reduce the potential for file versioning problems, relative to multi-user environments in which more than one person has access and modification privileges to the same files, single-user environments can not be presumed to preclude the problem.

Of those participants who do use some standardized form of file versioning control, the overwhelming majority (73.6%) use some type of file naming convention to distinguish among different versions of the same file (Figure 24), while 17.2% employ two or more distinct control strategies. Several participants commented that their file versioning is programmatically controlled by the database software, typically through the use of time stamps and, occasionally, front end file audit and renaming functions. One participant noted that he/she currently is working to incorporate *ISO 19115 Geographic information—Metadata* standards.

5.3.3 Related Paper Documents

As shown in Figure 25, more than 95% of participants indicated that they possess either the capacity or the responsibility for creating paper documents related to their digital GIS files (Figure 25). However, among these participants, nearly half (48%) indicated that they do so only occasionally (Figure 26). As noted in the comments to Question C3, related paper documents typically include reports, attribute/data file printouts, metadata printouts, flow-charts, line maps and, occasionally, colour maps and satellite imagery. Several participants noted that such printouts serve only as paper "backups," while others indicated that the printouts are actively used for data auditing, review, drafting and analysis purposes. A number of participants noted that

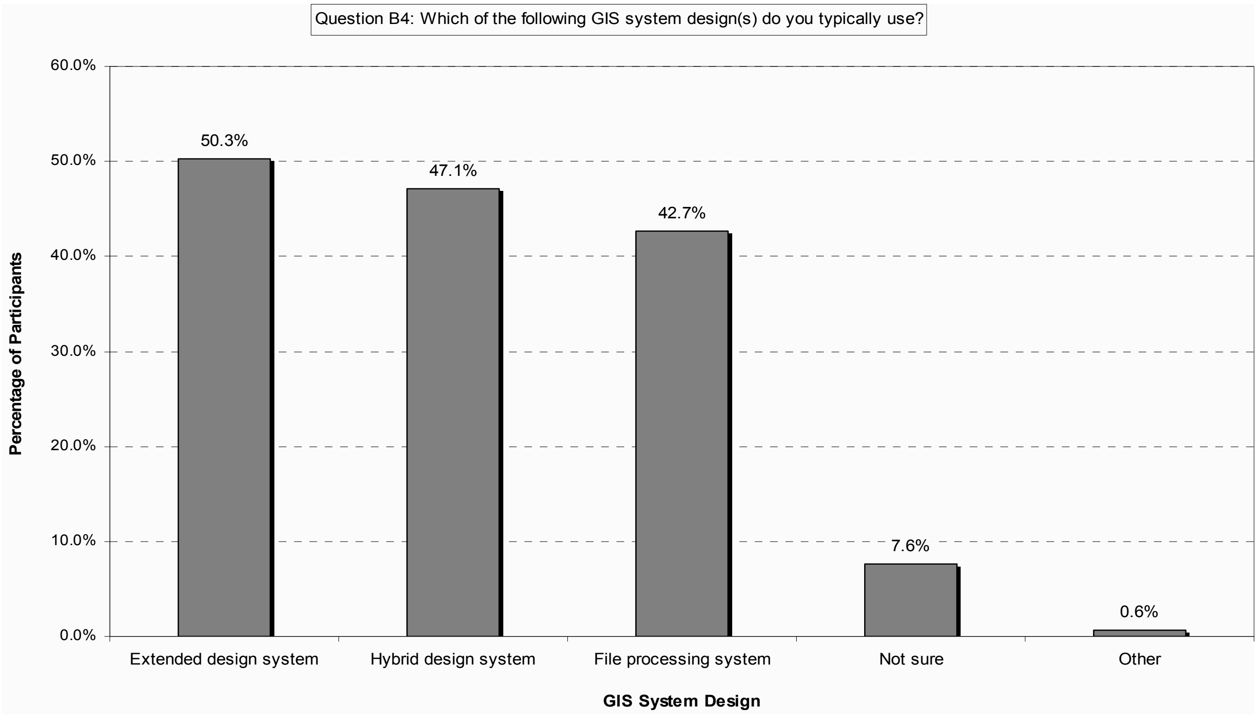


Figure 21. Summary of GIS system design(s) typically used by survey participants (n=157).

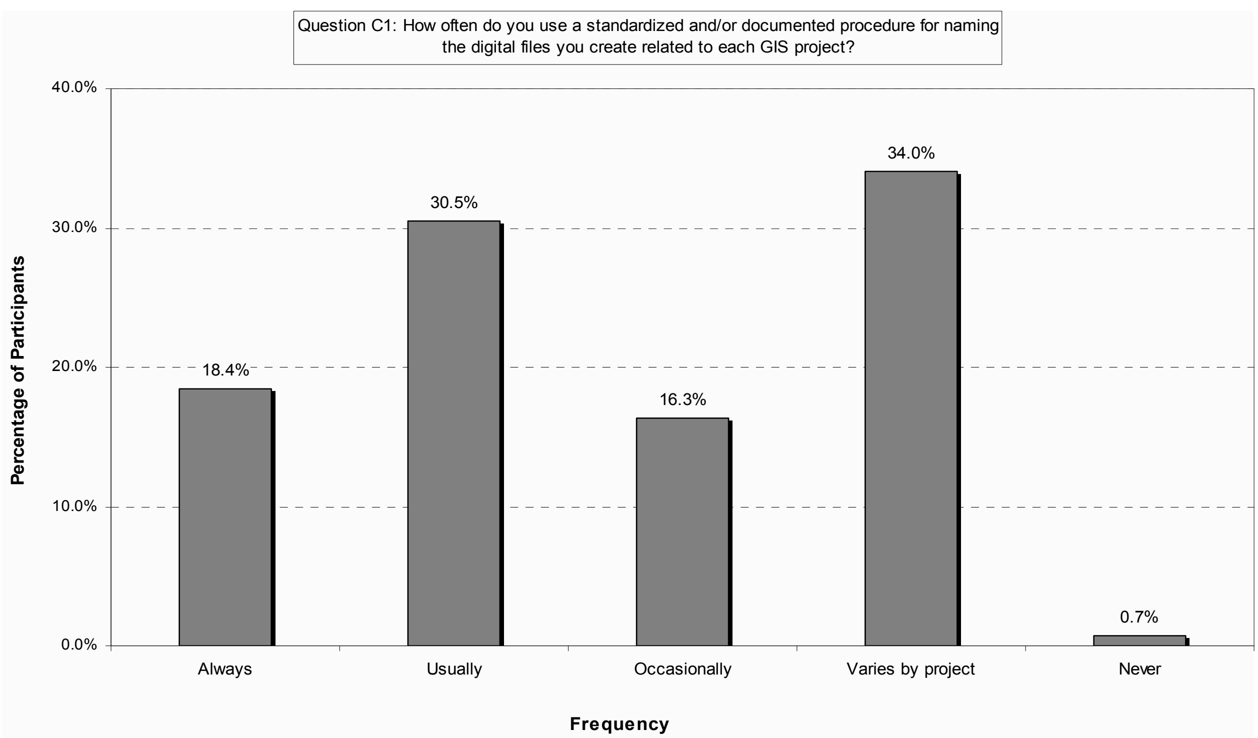


Figure 22. Frequency with which survey participants use standardized procedures for naming their GIS project files ("not applicable" responses excluded) (n=141).

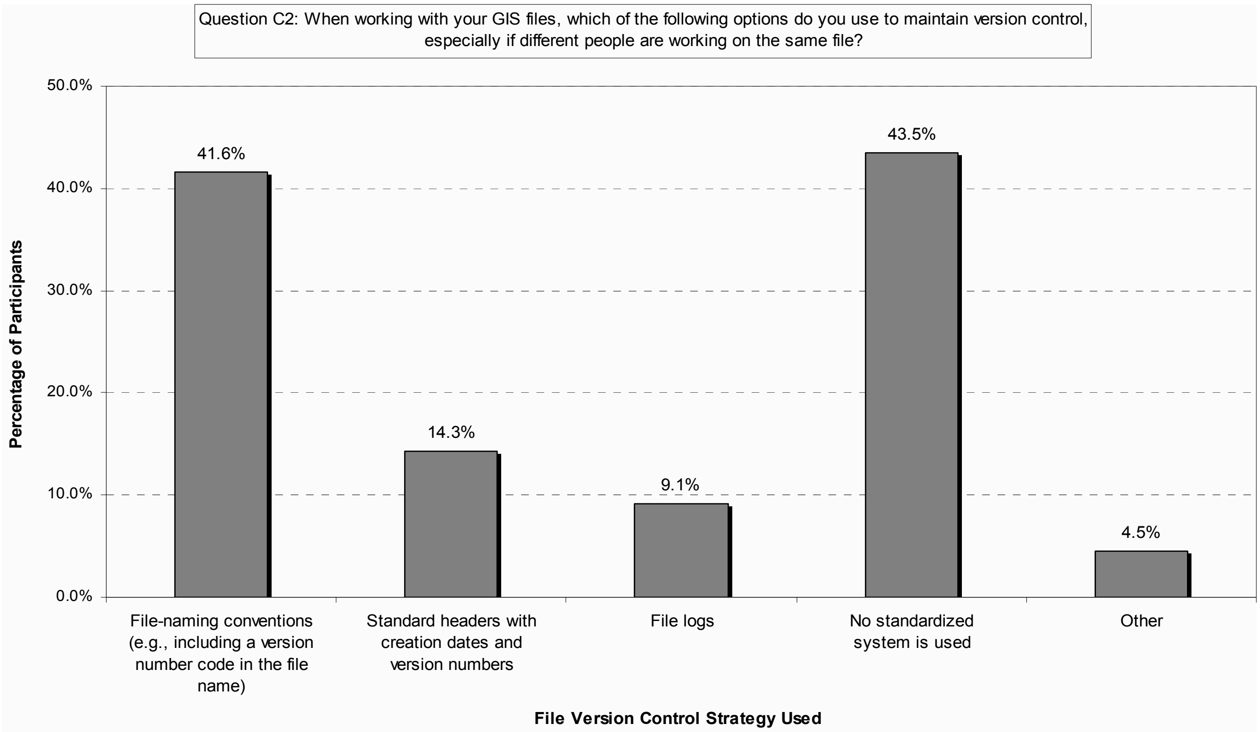


Figure 23. Summary of strategies used by survey participants to manage version control of their GIS files (n=154).

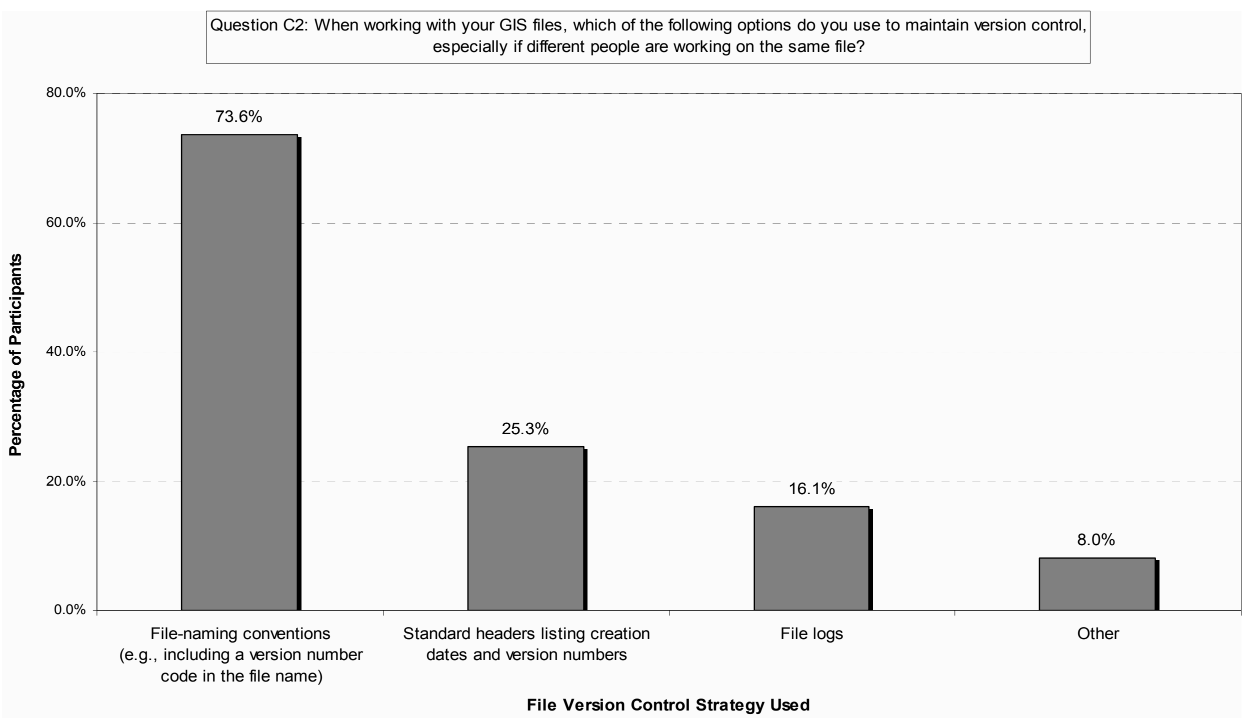


Figure 24. Summary of strategies used by survey participants to manage version control of their GIS files, excluding participants who use no standardized system (n=87).

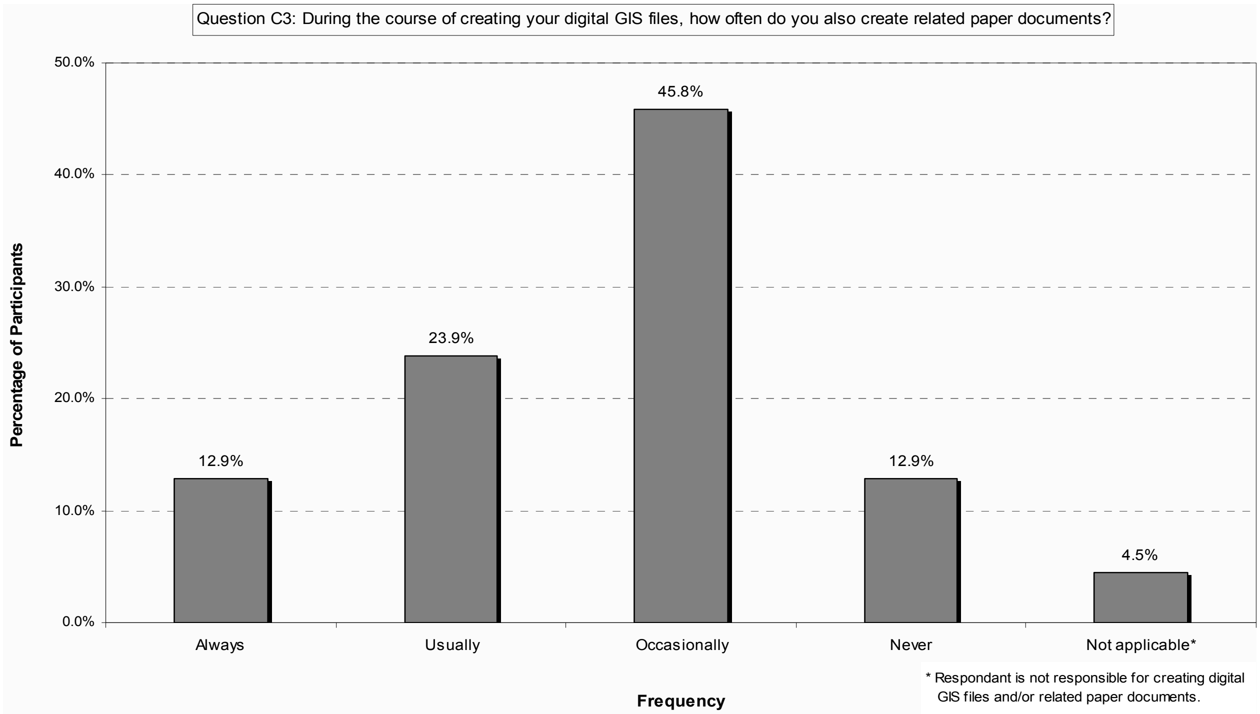


Figure 25. Frequency with which survey participants create paper records related to their digital GIS records (n=155).

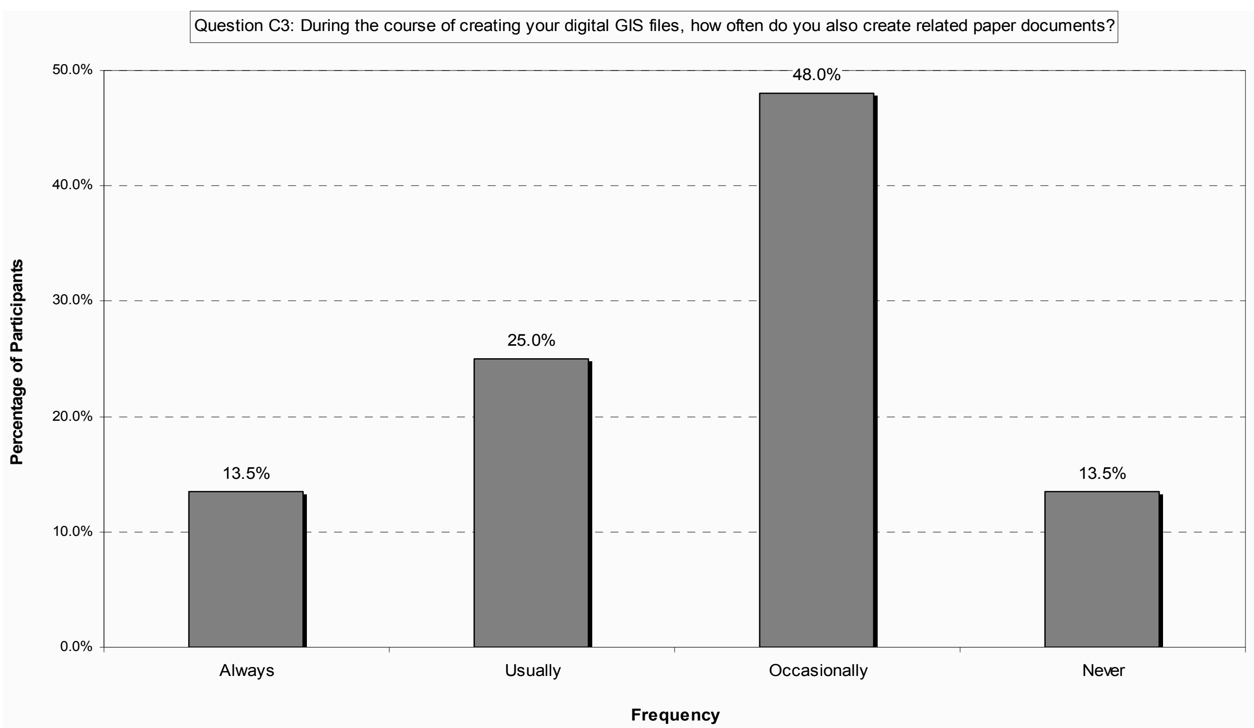


Figure 26. Frequency with which survey participants create paper records related to their digital GIS records ("not applicable" responses excluded) (n=148).

they often produce digital word-processing documents and/or PDF files in place of paper documents, especially those intended for dissemination

Although only a very small percentage of participants (3.2%) said they are not responsible for documenting the links between the digital and related paper files they create (Figure 27), nearly half (44.2%) of the remaining participants (i.e., those who ostensibly *are* responsible in some way for such documentation) claim that they only occasionally, if ever, document such links (Figure 28). As one participant noted, "there is a tendency to overlook this link." Moreover, 16.4% of the participants responsible for this type of documentation, indicated that the final decision to do so varies depending on the nature and importance of the relationship between the paper documents and the digital files. For example, for these participants, the relationships between draft or otherwise ephemeral paper documents (such as reports to supervisors, conference presentations, etc.) and their digital counterparts are rarely, if ever, documented. Judging from the comments to Question C4, one of the critical factors affecting the extent to which paper document-digital file relationships *are* documented is the degree to which such documentation can be *automated* (e.g., via plotter stamps, automatically generated file lists, etc.). Only two participants noted having processes in place for automatically generating such links (including, for example, an archive management system which "employs the same Key_ID structure [for the paper documents] as the digital records"). Two participants also indicated that they currently are working on implementing standardized protocols for documenting their paper document-digital file links.

5.3.4 Overall Degree of Documentation

Most of the participants (83.3%) produce at least some form of documentation detailing factors related explicitly to the creation and/or subsequent manipulation/modification of their individual GIS files and/or the project as a whole, while the remaining participants (16.7%) indicated that they typically do not create any such documentation about their GIS projects. For those who do document their projects, the two most commonly documented aspects are information about the individual datasets comprising the GIS and key project details (project name, number, etc.), while identification of the actual records affected by subsequent system changes or modifications, as well as information about the computer system(s) used to create/host the GIS, are the two least commonly documented aspects out of the eight specific choices listed in the survey (Figure 29). As noted by several of the participants, other documented aspects include: (1) the methodology for the creation of GIS projects and their datasets, (2) database schema, (3) the nature of any glossaries, word-lists or look-up tables incorporated into the GIS, (4) projection information and (5) contact information.

5.3.5 Specific Documentation Strategies

As shown in Figure 30, a majority of the actual documentation approaches used by the participants fall into two distinct clusters. The first cluster is characterized by the following three distinct documentation strategies: (1) documentation is entered informally into various paper documents that may or may not later be consolidated into a single binder or folder (30.8%); (2) documentation is entered into a formal digital log file(s), such as spreadsheet or word processing

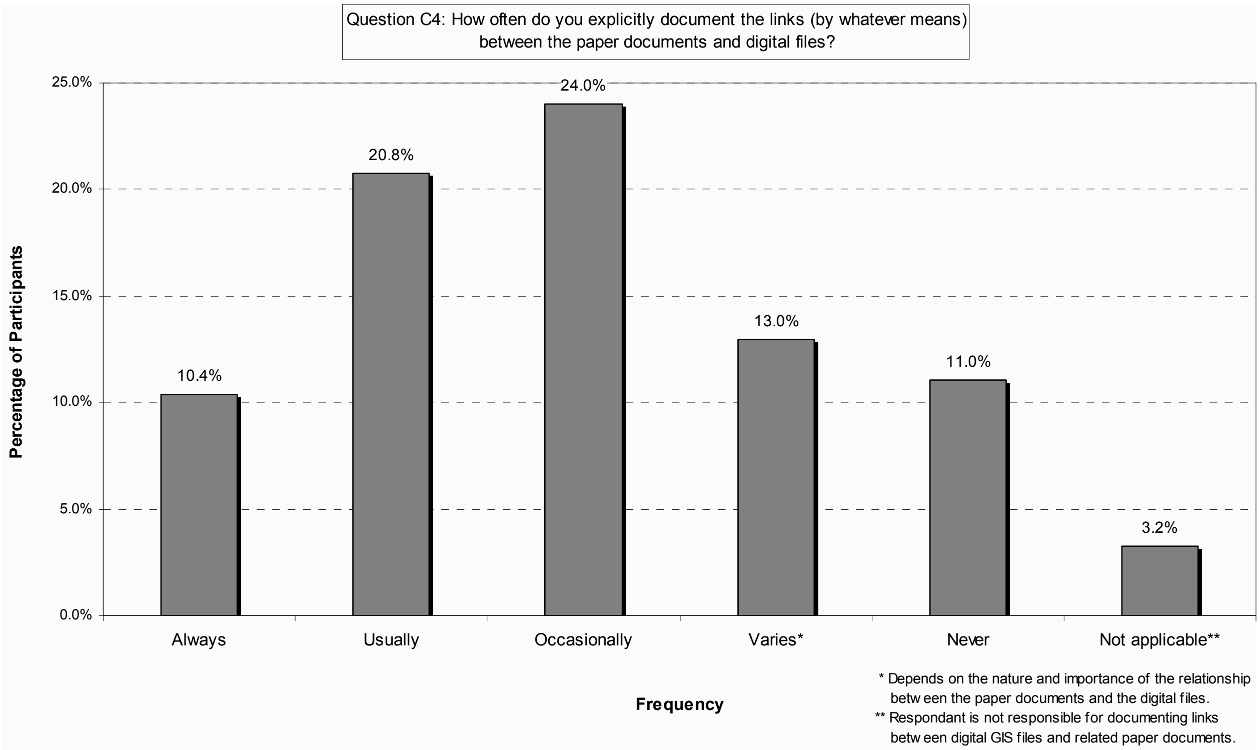


Figure 27. Frequency with which survey participants explicitly document the links between their paper and digital GIS project files (n=154).

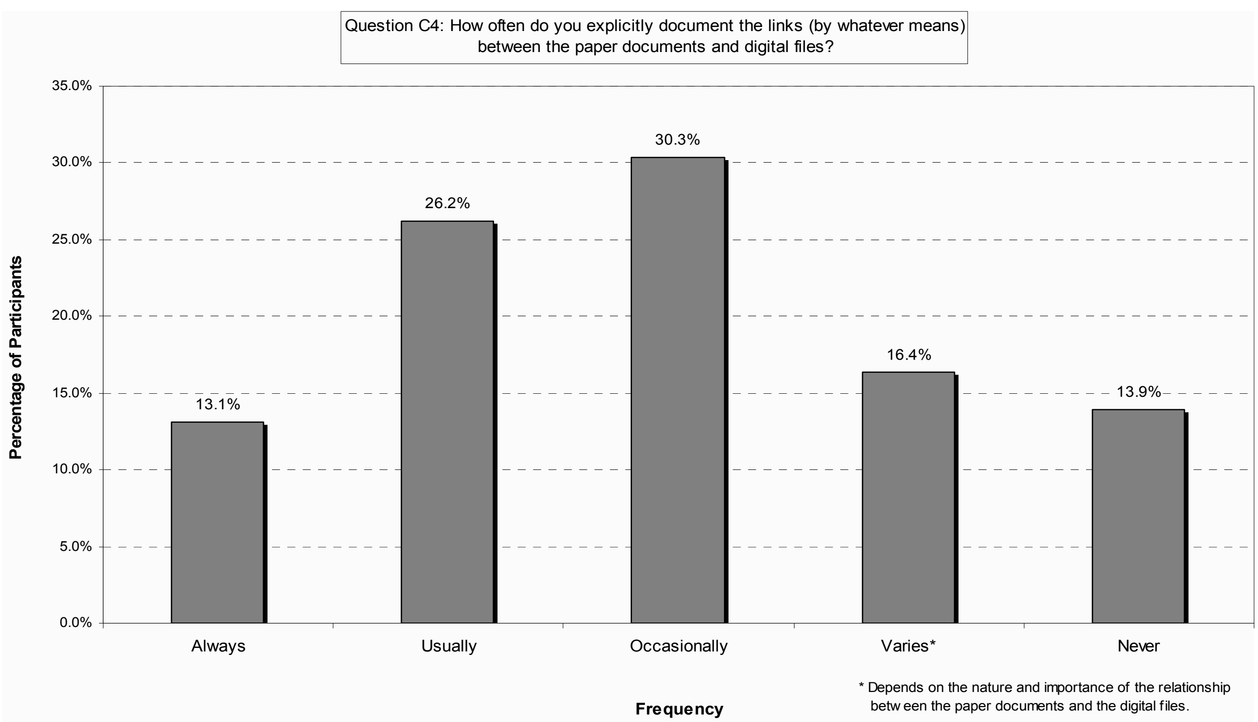


Figure 28. Frequency with which survey participants explicitly document the links between their paper and digital GIS project files (“not applicable” responses excluded) (n=122).

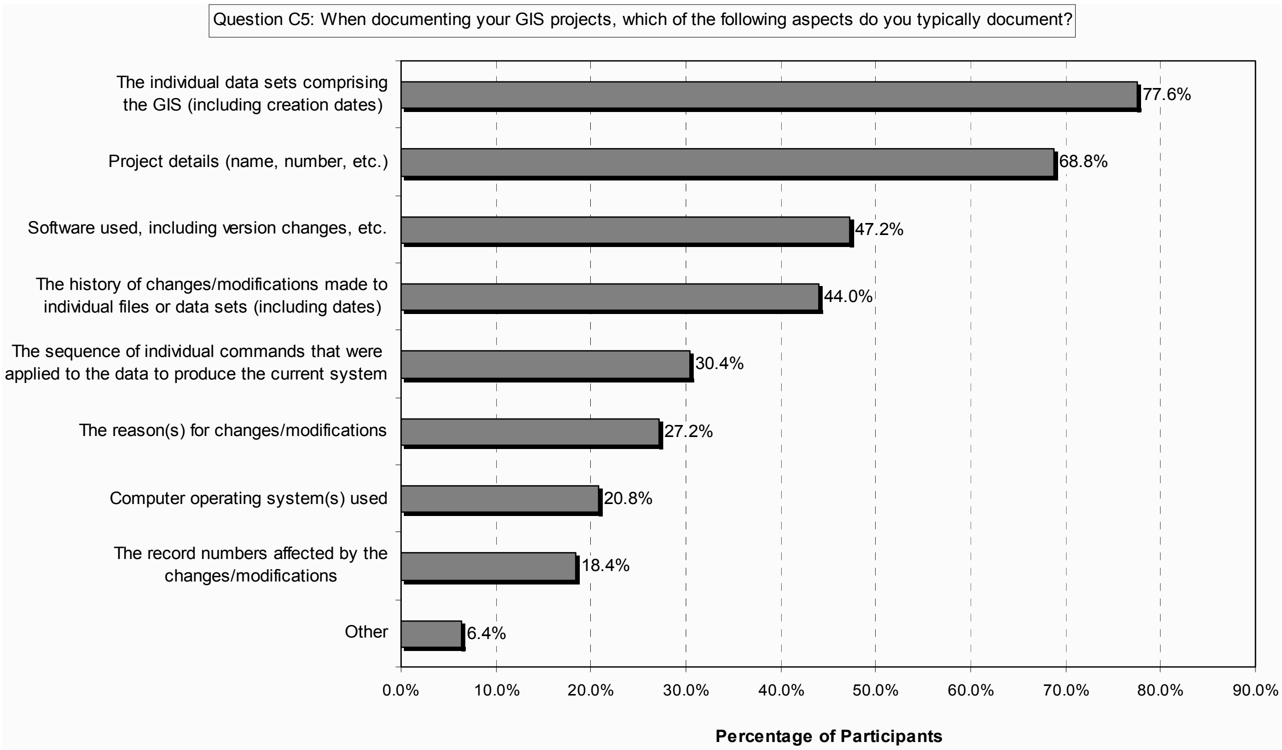


Figure 29. Documented aspects of GIS projects ("not applicable" responses excluded) (n=125).

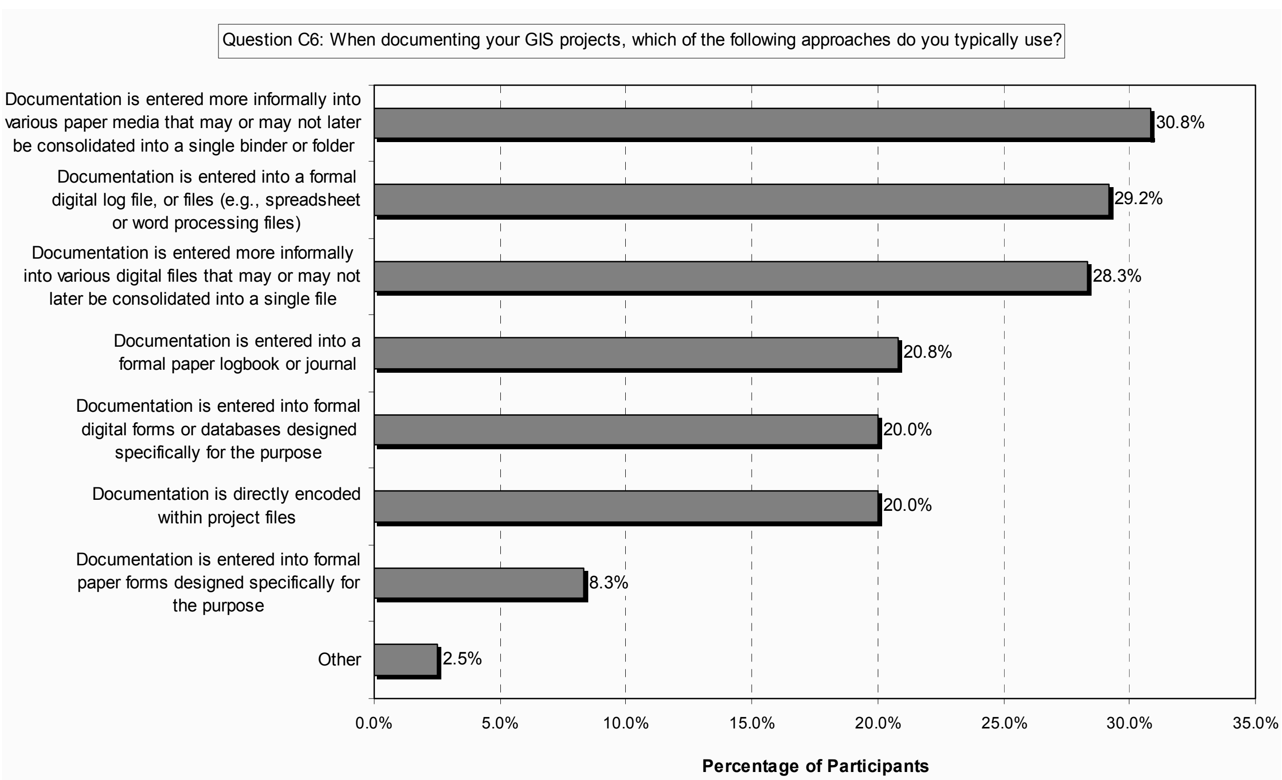


Figure 30. Approaches typically used by survey participants to document their GIS projects ("not applicable" responses excluded) (n=120).

files (29.2%); and (3) documentation is entered more informally into various digital files that may or may not later be consolidated into a single file (28.3%). The second cluster is characterized by the following three approaches: (1) documentation is entered into a formal paper logbook or journal (20.8%); (2) documentation is directly encoded within the project files (20.2%); and (3) documentation is entered into formal digital forms or databases designed specifically for the purpose (20.0%). Of those participants who do document their GIS projects, 35.2% typically use two or more distinct documentation strategies.

5.3.6 Documentation of Data/File Modifications

Of those participants who indicated that they provide some level of documentation for their GIS projects, 67.5% base their decision about when to document modifications to their data or GIS files on the nature of the changes made (Figure 31). For example, changes related to a spelling error are less likely to be documented than changes related to a systematic data recording error. The actual *number* or volume of changes made appears to be considerably less important a factor (29.8%) than the *nature* of the changes when deciding whether those changes warrant documentation, with the *type* of data or files changed (e.g., attribute text vs. numeric data vs. graphics) only slightly more important a factor in this decision (39.5%). Just over one quarter (27.2%) of these participants indicated that time and/or financial constraints influence their documentation decisions.

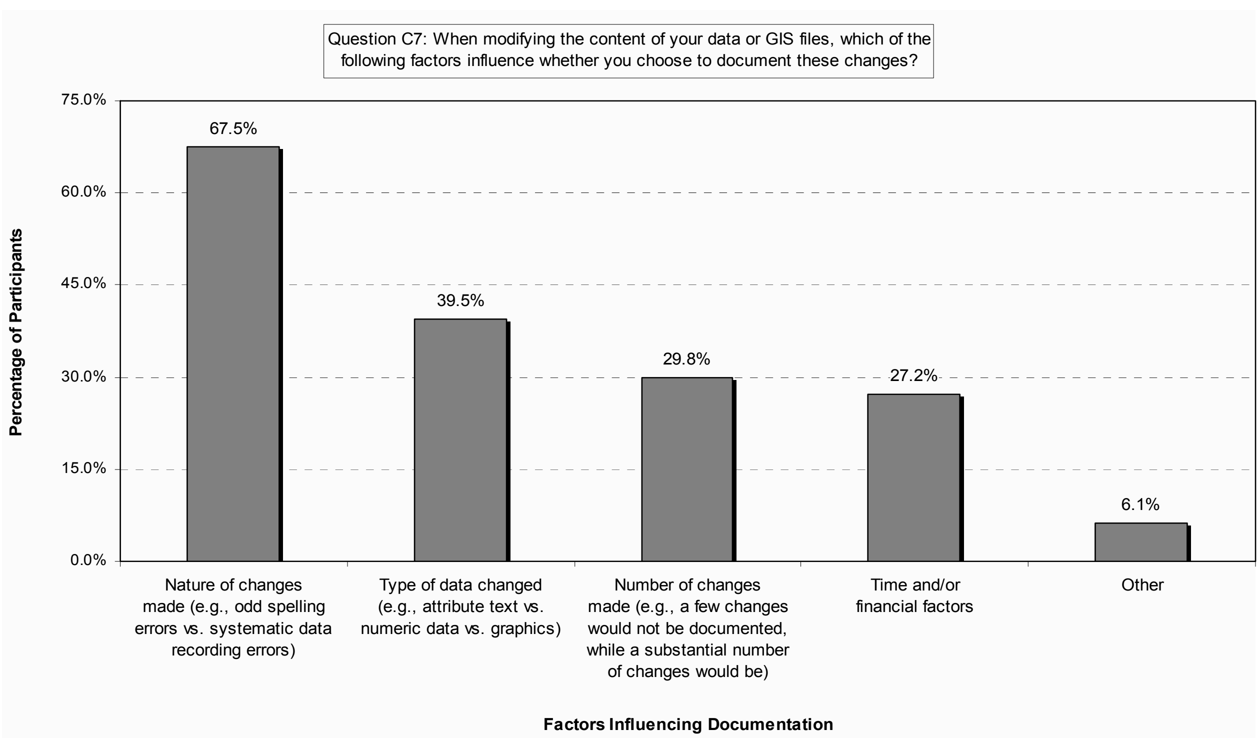


Figure 31. Key factors that influence survey participants' decisions about when to document GIS project activities ("not applicable" responses excluded) (n=114).

Two other factors identified by the participants that influence their decisions about when to document changes made to their data and GIS files include: (1) the time between changes (As one participant noted, "Active projects change daily and it's impractical to record [all these changes]…but, if a change is made to completed projects sometime later, due to a design change, etc., [such] changes will be noted in the metadata or project files."); and (2) the effect of the changes on the project's outputs or results (As one participant commented, "Changes that do not affect the end result or do not substantially contribute to analysis are not documented."). In fact, this latter consideration is not really a distinct factor per se, but rather an evaluative assessment of the overall effect of any (or all) of the first three considerations (i.e., the nature, numbers and/or types of changes made).

It is important to note that, based on the comments of several of the participants, there appears to have been some misinterpretation of Question C7. In fact, at least three participants interpreted the question (which asked, "When modifying the content of your data or GIS files, which of the following factors influence whether you choose to document these changes?") as though it was referring only to their original or primary data files, which, as these participants noted, they never modify. Several other of the participants' comments also are worth noting because of their relevance to the issue of documentation. One participant admits to being "more conscientious about documenting changes if others besides myself are accessing the data files." Another notes that, as the result of a conscious design decision, his/her GIS automatically "creates 'information history' (in multiple versions) whenever a change is made to an object." Finally, one participant was adamant that "all changes MUST be tracked so as to satisfy an absolute rollback capability."

5.3.7 Overall Documentation Processes and Procedures

Nearly three quarters (74.8%) of the participants who document their GIS projects do so more or less continuously throughout the duration of the projects, while just over one quarter (25.2%) conduct most or all of their project documentation toward the end of their projects or after the projects are completed. Of those who document continuously, 38.3% indicated that the documentation often does *not* occur at the same time as the GIS files are created or modified, whereas 36.5% usually do produce simultaneous documentation (Figure 32).

Figure 33 summarizes the degree of overall consistency of the documentation procedures used by the participants from one project to the next. Regardless of the particulars of the documentation procedures used, of those participants who do document their projects, a majority (61.7%) indicated that the procedures usually or always are consistent from one project to the next, while 36.5% only occasionally or never follow consistent procedures (Figure 34). As several of the participants emphasized in their comments, consistency, although an ideal toward which most GIS practitioners strive, is nevertheless difficult to achieve due to the sometimes highly variable nature of different projects in terms of the specific data types used, as well as the size and objectives of the projects. In addition, there occasionally are uncontrollable external factors that may influence how a project is conducted. For example, one participant noted that "additional documentation is sometimes needed by the client, so additional work is sometimes done for that purpose." Give these factors, it is not surprising that only 8.8% of the participants are always able to maintain consistency in their documentation procedures from one project to the next. For most, the reality is that their documentation processes and procedures are continuously

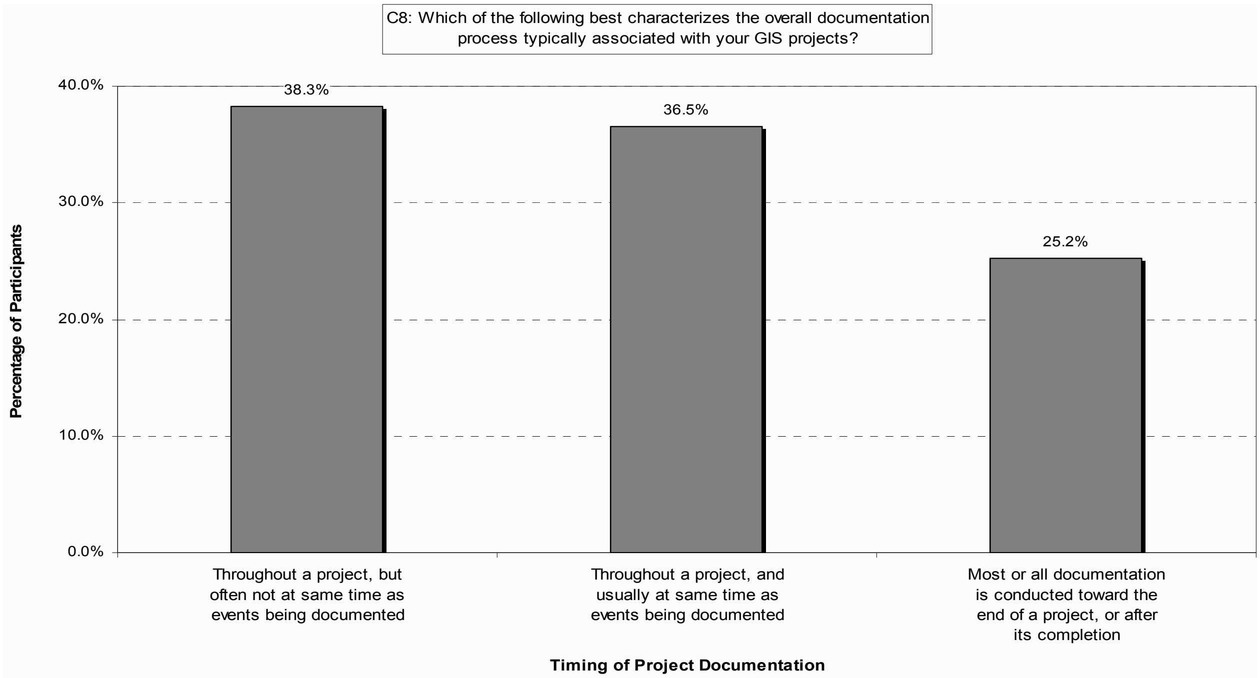


Figure 32. Characterisation of the process typically used by survey participants to document their GIS projects in relation to the timing of a project event and its subsequent documentation (n=115).

evolving and improving from one project to the next. In the words of one participant, "[my projects] start off trying to be consistent, but it seems almost impossible to keep [them] that way…the best I can do seems to be [to] build on the previous projects which allows some backwards compatibility, but each new project offers more detail."

5.3.8 Documentation Rationale

Of the five rationales for documentation listed in survey, approximately half of the participants who document their GIS projects considered each of the first four to be important factors that influence their documentation decisions (Figure 35). The single most compelling reason (57.5%) for adequately documenting the sequence of modifications was to enable "rollback" of modified files or data to a pre-modified state, followed closely (54.2%) by the ability of such documentation to provide an overall picture of the evolution of a project. Interestingly, particularly from an archivist's perspective, the ability of adequate documentation to serve as supporting evidence should one's project research be challenged was chosen as the rationale with the *least* influence on the documentation decision-making process. Other influential factors identified by the participants included: (1) to maintain data quality and results continuity in the event of project personnel changes; (2) to enable future integration of the data with a duplicate system; (3) to serve as finding and/or 'memory' aids for previous datasets; and (4) to increase awareness of inherent data errors, limitations and assumptions. As one participant emphasized, "It is important to catalogue data error and assumptions because all GIS projects have differing margins of error, dependent on whether data was collected through API, GPS, and so on."

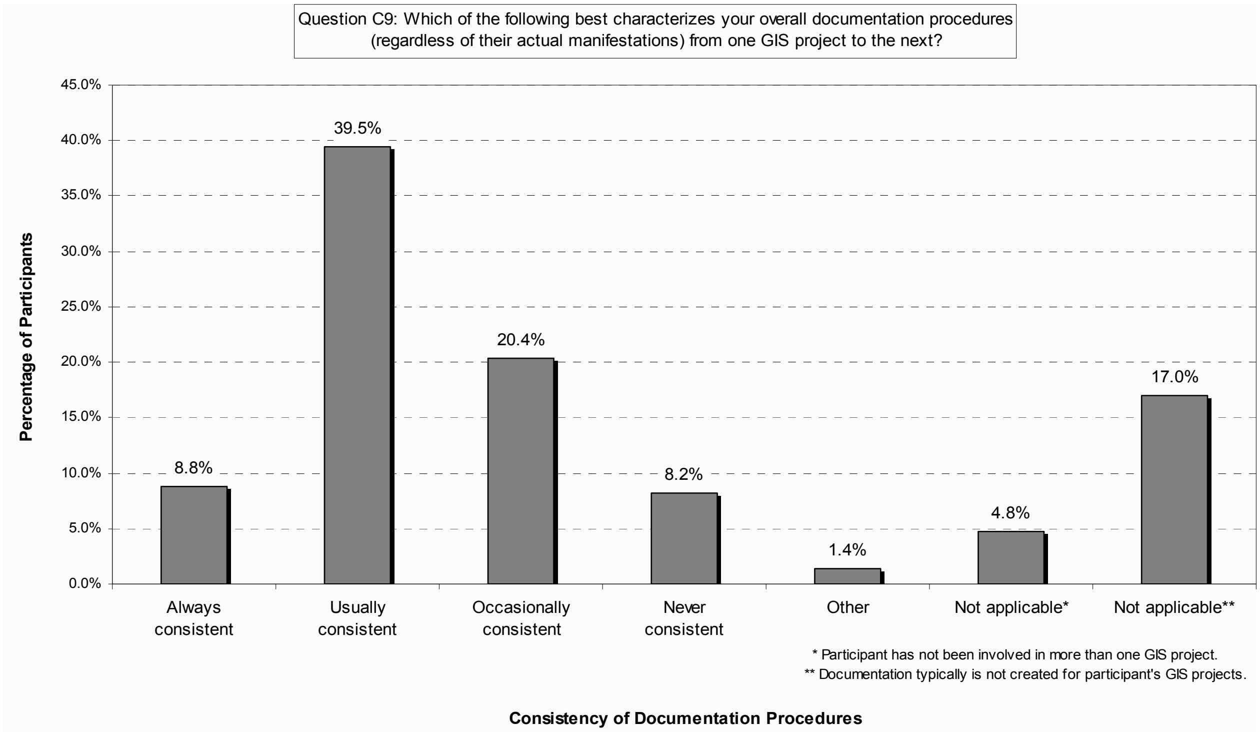


Figure 33. Consistency of documentation procedures of survey participants from one GIS project to the next (n=147).

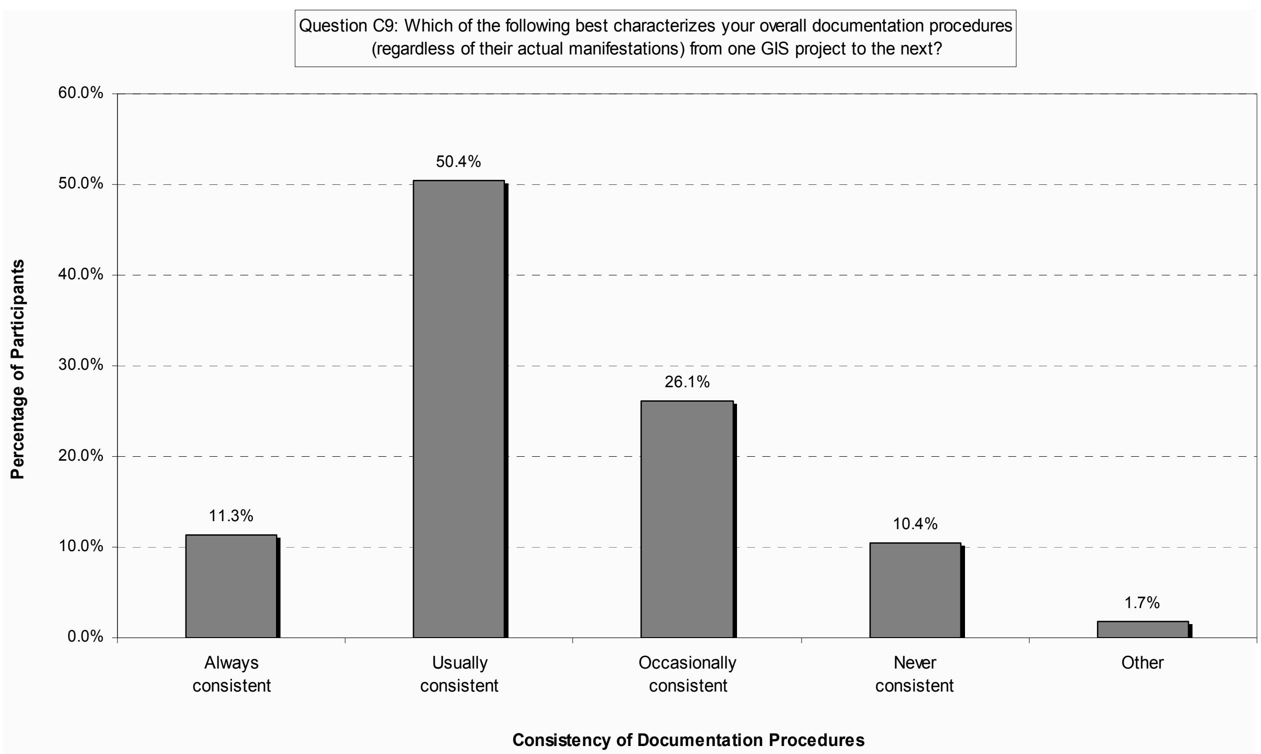


Figure 34. Consistency of documentation procedures of survey participants from one GIS project to the next ("not applicable" responses excluded) (n=115).

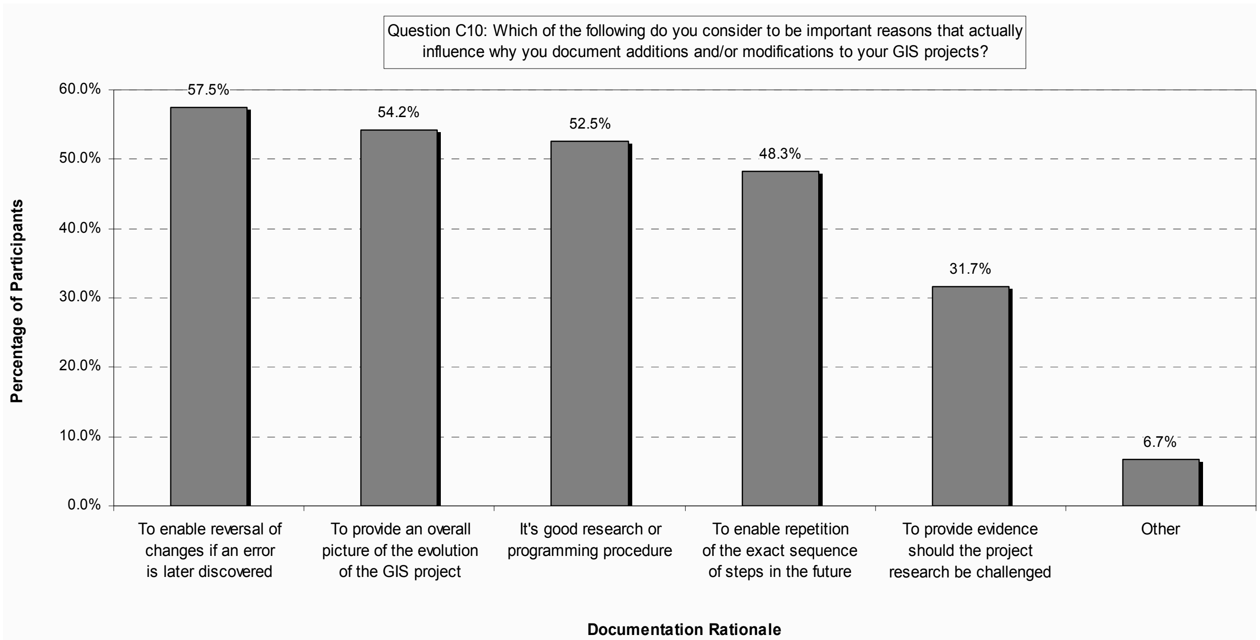


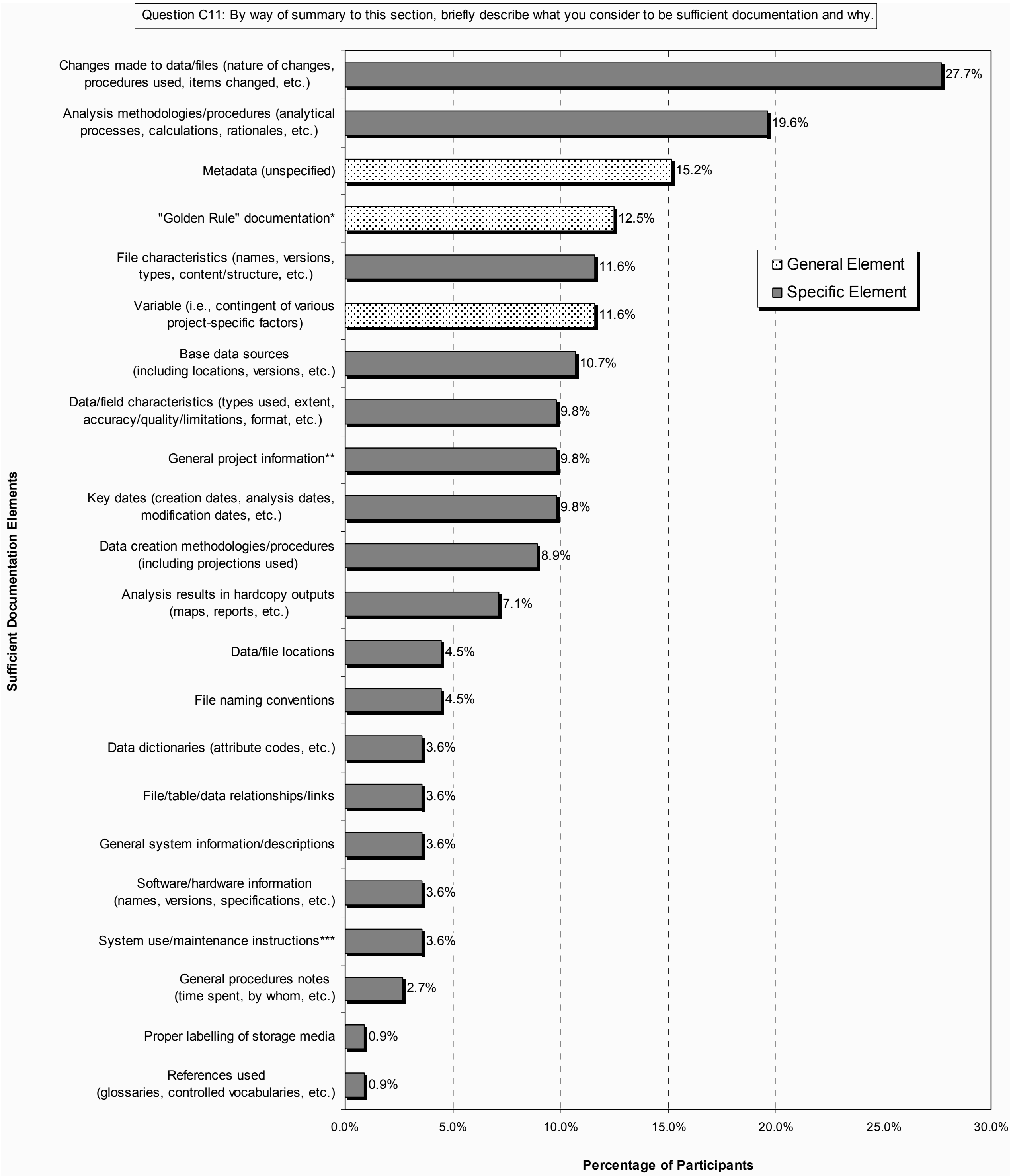
Figure 35. Key reasons identified by survey participants for documenting additions and modifications to their GIS projects ("not applicable" responses excluded) (n=120).

5.3.9 Sufficient Documentation

As a summary to the questions in this section, participants were asked to briefly describe what they considered to be sufficient documentation of their GIS projects and why. What is most apparent from the responses received to this question is that there are rather diverse opinions about what constitutes "sufficient" documentation. As shown in Figure 36, the three specific project activities, elements or functions cited most frequently included: (1) changes made to the project's data or files, including documentation of the various aspects of these changes, such as the nature of the changes, the procedures used to affect the changes and/or the actual items changed[43] (27.7%); (2) analysis methodologies and procedures used, including documentation of the analytical processes and calculations used and the underlying rationale for them (19.6%); and (3) individual file characteristics, including documentation of a file's type, name, version, content and/or structure (11.6%). Many of the comments received were too generalized to relegate to one specific category or another and so instead were lumped into three broad categories, (1) metadata, (2) "Golden Rule" documentation and (3) variable factors.

The term "Golden Rule" is borrowed from a recordkeeping concept championed by the early 20th Century British archivist, Sir Hilary Jenkinson, which he identified as the "Golden Rule of Archive Making." In the words of Jenkinson, the "Golden Rule" refers to the concept of ensuring that an organization's records are "always in such a state of completeness and order that, supposing [the organization] and [its] staff to be by some accident obliterated, a successor totally ignorant of the work of the office would be able to take it up and carry it on with the least possible

[43] A number of the participants stressed that it was only necessary to document "significant" changes, which one participant defined as, "those that would not be obvious to another GIS researcher looking at the project and final write-up."



* Sufficient documentation of all critical project procedures and components -- what was done, how it was done and why -- to enable someone with no previous knowledge of the project to step in and take over, if necessary.
** e.g., project name, creator's name, overall description of design, methodology and justifications, current status, results summary, etc.
*** Especially for customized interfaces, custom functions, etc.

Figure 36. GIS project elements identified by survey participants that are required to provide sufficient documentation of their projects (n=112).

inconvenience and delay simply on the strength of a study of the Office Files."[44] This is similar to the concept, referred to by modern-day current records managers, as "vital records management," where "vital records" are defined as those records that are essential to the continued functioning of an organization and which, if lost, deleted or destroyed, would be both prohibitively costly and time-consuming to recreate—if they could be recreated at all. Thus, the broad category of "Golden Rule" documentation, as used in this report, encompasses the general notion, expressed in one fashion or another by 12.5% of the participants, that to be considered "sufficient," the documentation of one's project must enable someone else who is technically knowledgeable, but who has no prior affiliation with the project, to step in and, on the basis of the available documentation, be able to develop, within a reasonable time frame, a clear understanding and working knowledge of the GIS project in terms of its purpose, structure, evolution, data content, current status, etc. As well, "sufficient" documentation, according to many of these participants, should enable such a person to trace and/or recreate the original steps or analytical procedures used to bring the project to its current state.

With respect to the "variable factors" category, a number of participants emphasized that the type and amount of documentation required to "sufficiently" document a project will vary depending on several factors, including: (1) the nature, purpose, size and scope of the project, (2) the type(s) of data involved, (3) the task(s) being performed and (4) the nature of the creating entity. Even here, however, opinions vary. For example, with regard to project size, one participant suggests that, "some small projects where the data is in house for a short period do not require as much documentation as multi-year projects,"[45] while another suggests it is instead the smaller projects for which documentation often is most critical, given the risk of having all the vital project information residing "only in one man's head." Two participants consider the distinction between private and public domain projects to be an important factor, where considerably more documentation—sufficient to ensure "replicability" of the project—is critical for the former, while a more "black box" approach to documentation is acceptable, within reason, for the latter, "provided researchers can 'show their work' if challenged."

One final observation regarding the architecture of the survey; given the location of this question in the survey, situated as it was at the end of a series of question about the participants' project documentation procedures and strategies, raises the question of whether, or to what degree, the comments received were influenced by the information presented in the preceding questions. In retrospect, it may have been more informative to have *opened* the section with this question.

5.4 Section D: Digital Preservation Practices

5.4.1 Influence of Preservation Concerns on GIS Projects

Question D1 asked whether concern for the long-term preservation of their GIS projects in an archival setting influenced the planning, design and/or implementation phases of their projects. As shown in Figure 37, the question did not apply to a small percentage (2.9%) of the

[44] H. Jenkinson (1922), *A Manual of Archive Administration* (London: Percy Lund, Humphries & Co.), p. 153.
[45] However, as this participant also acknowledged, "I am learning the hard way that you can not always predict which projects will be short term."

participants because they are only involved in the analysis or preservation phases of their projects. Of those to whom the question did apply, nearly three quarters (73.3%) admit to being concerned about the long-term preservation of their GIS projects (Figure 38). However, only 28.1% these individuals actively incorporate those concerns into the planning, design and implementation of their projects, while the majority (45.2%) acknowledge that preservation concerns generally have no direct influence on the execution of their projects. Even more illuminating is the fact that one quarter (25.2%) of the participants admit to having no concerns about the long-term preservation of their projects, at least not during the planning, design and implementation phases. These results, when considered together with the results from Question C8, in which 25.2% of the participants indicated that they tend to be concerned with documenting their projects only after they are completed (or are nearing completion), suggest there is a fairly high level of complacency among many GIS archaeologists with respect to documentation and long-term preservation issues.

Although the exact causes of this complacency are impossible to pinpoint, it is clear from many of the participants' comments that one important factor is the lack of suitable archival repositories for digital archaeological records. For example, as one participant noted, "I'm not aware of any archival projects for GIS data/projects in my area [Midwest and Great Lakes region of the United States]. If I had been, I would have attempted to make [my GIS project] transferable." Another lamented that, "there is no equivalent in Argentina, and no general consensus on data archives." Another participant from the United States notes that, "there currently is no established long-term curational facility for the projects that I work on." And then there was this telling comment from yet another United States participant (in Utah), "Paper data is the archival medium. Curational facilities and regulatory agencies will not accept electronic data in place of paper data, but some will accept it as supplemental. As such, long term data planning is a low priority and is only for internal convenience rather than legal compliance—which also means that it is a low priority for funds." Indeed, funding is another key factor identified by several participants who noted that more often than not the funding needed for incorporating adequate documentation and preservation procedures into project designs simply is not available, especially for those projects done on a contract basis. As well, the availability of adequate funding often is influenced by a project's nature and scope, with many small-scale, and what one participant calls "exploration," projects unable to secure the necessary funding for long-term preservation.

Still another factor seems to be what is perhaps best characterized as a "middle man mindset," in which the participant (especially those working as consultants) considers long-term preservation to be solely their client's problem. As one participant states, the only concern is "that the client can incorporate the data into their system or distribute it." Similarly, another participant matter-of-factly notes that, "we provide the data to our clients and what they do with it is generally up to them." Finally, the following comment likely is representative of the way many view this issue: "…my data is always (nearly daily) changing so there is no need for archival preservation of the dataset at any one time (other than immediate backups)." Obviously, this type of reasoning, in which the exclusive focus is on a project's very immediate day-to-day uses and requirements, fails to acknowledge that all projects are finite. Given this, it is understandable why there sometimes is so little cognizance of, and concern for, the ultimate disposition of these projects.

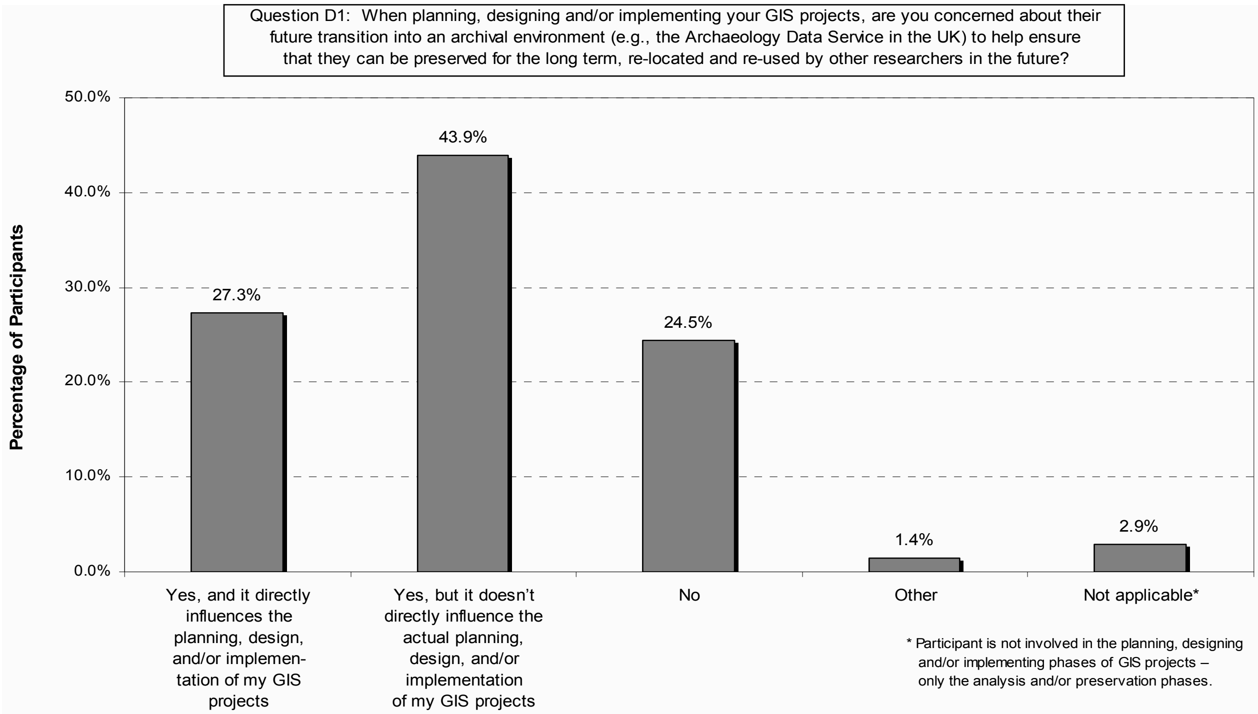


Figure 37. Characterisation of survey participants' concern for the long-term preservation of their GIS projects in an archival setting and the impact of this concern on overall project planning, design or implementation (n=139).

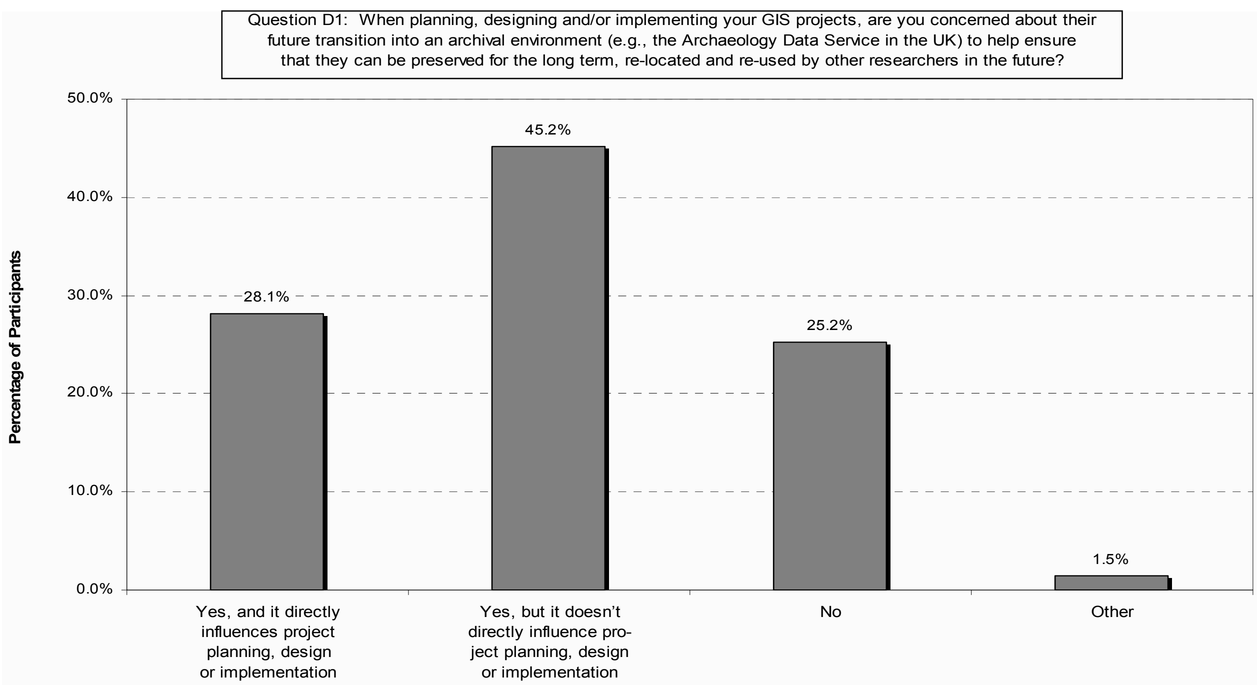


Figure 38. Characterisation of survey participants' concern for the long-term preservation of their GIS projects in an archival setting and the impact of this concern on overall project planning, design or implementation ("not applicable" responses excluded) (n=135).

On the other hand, many of the participants are very cognizant of the gravity of the issue, as the following comment demonstrates: “This is a nightmare. We had planned to archive with the CSA.[46] We could not use ADS[47] because we could not finance it. At the moment we back everything up, but what will happen in the long term is very unclear and a very major concern.” In a similar vein, another participant asks, “How does one cope with the changing costs of production and the backup legacy of the data? Does the client in CRM cut and run? Usually. I’m not sure there is an answer for this in the US CRM that I have worked on.”

5.4.2 Long-term Preservation: Use of Designated Repositories vs. In-house Preservation

When asked about where their GIS projects are preserved for the long-term, 20.4% and 29.0% of all participants who answered Questions D2 and D3, respectively, indicated that their GIS projects still are in active use and so have not yet had to face the decision about where to store their projects for the long-term (Figure 39). Of the remaining participants, more than half (67.9%) said they only occasionally if ever transfer their completed projects to a designated repository (such as the Archaeology Data Service, or a state, museum or university archives) for long-term preservation, while just one-quarter (25.7%) said they always or usually do (Figure 40). For roughly one third of participants (29%), the concern of where to preserve their inactive GIS projects—’in-house’ and/or in a designated repository—has yet to become a pressing issue, since all of the projects in which these individuals are involved are still active (Figure 41). However, among those participants for whom the need to preserve inactive projects is a pressing issue, a clear majority (65.3%) said that they currently always or usually rely on so-called ‘in-house’ preservation strategies (Figure 42), while less than one quarter (23.4%) said that they only occasionally, if ever, preserve their projects for the long-term ‘in-house.’

Of particular interest are the 10.8% of participants who neither preserve their projects in-house nor transfer them to a designated repository. Although the ultimate fate of these projects is not known, one possible explanation, as inferred from the comment of one participant who noted that “projects are usually in the hands of their creator,” is that these represent consulting projects in which the finished products are transferred to the clients, with no copies being retained by the consultants. Closer inspection of the data does not appear to support this hypothesis, however. In fact, a majority (63.6%) of those participants who answered “never” to both questions listed their professional affiliation as “College/University,” while another 18.2% indicated they are involved in a combination of “College/University” and some form of public/private consulting. In all, only 9.1% of these participants listed consulting as their sole affiliation.

[46] Center for the Study of Architecture. The participant is here referring to the CSA’s now terminated, Archaeological Data Archive Project (ADAP). As originally conceived back in the fall of 2001, this pilot project had as one of its goals the creation of the Archaeological Data Archive, an online repository for digital data from archaeological research. Unfortunately, however, participation in the project proved inadequate to justify its continuation and the project was terminated in August of 2002. More information about the project and the reasons for its demise is available online at http://www.csanet.org/archive/adap/adaplond.html and http://csanet.org/newsletter/fall02/nlf0201.html.

[47] Archaeology Data Service. Based in the United Kingdom, ADS provides long-term preservation of digital archaeological data, with preference given to data from the United Kingdom or from researchers based in, or funded by, the United Kingdom. More information about ADS is available online at http://ads.ahds.ac.uk/.

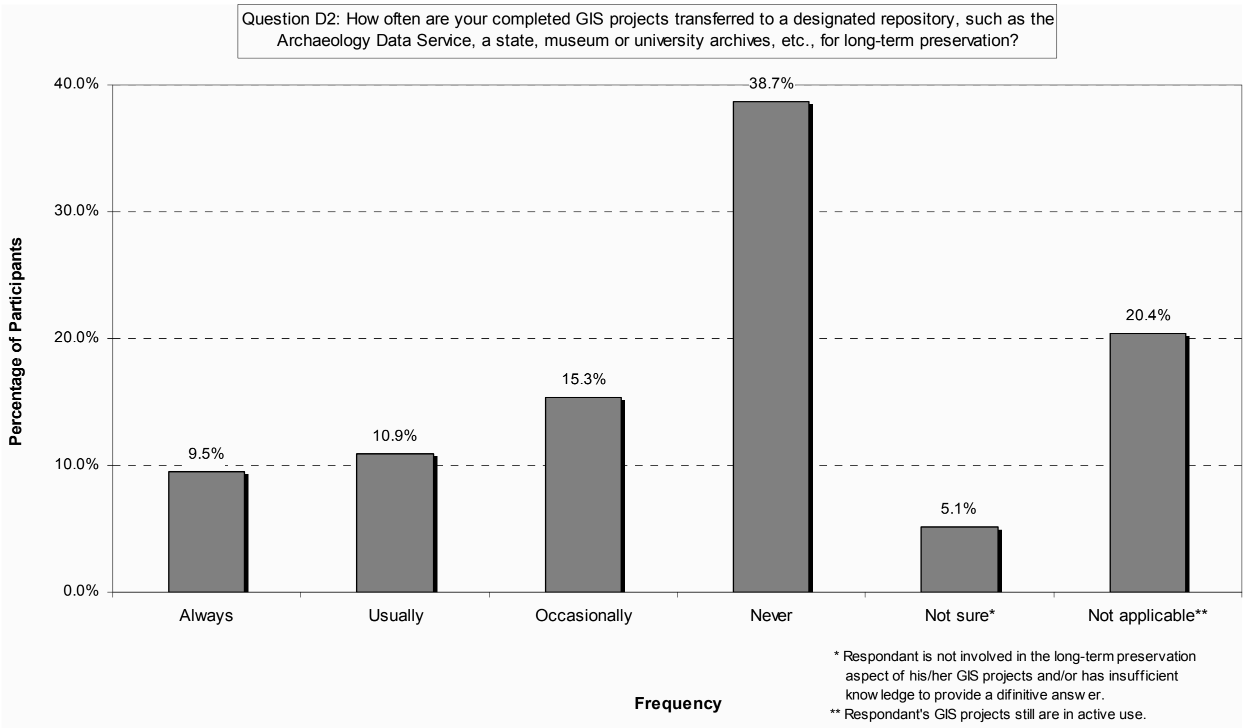


Figure 39. Frequency with which survey participants transfer their inactive GIS projects to a designated repository for long-term preservation (n=137).

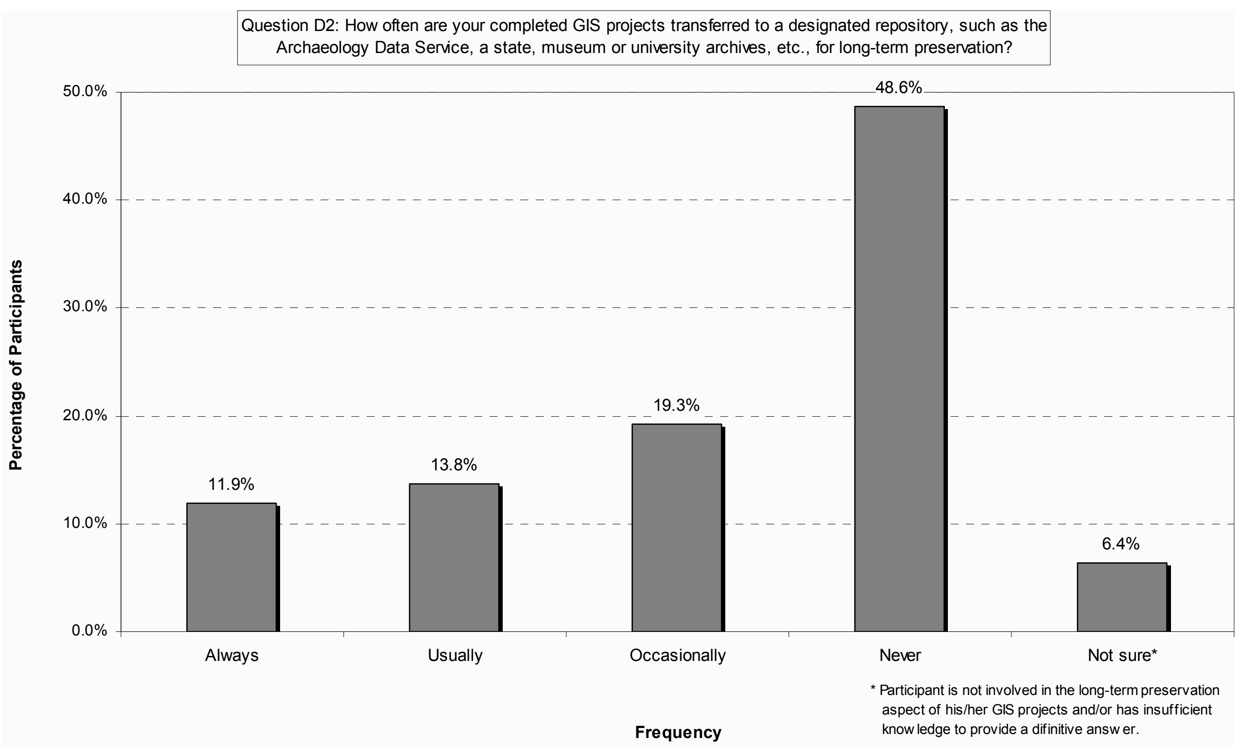


Figure 40. Frequency with which survey participants transfer their inactive GIS projects to a designated repository for long-term preservation (“not applicable” responses excluded (n=109).

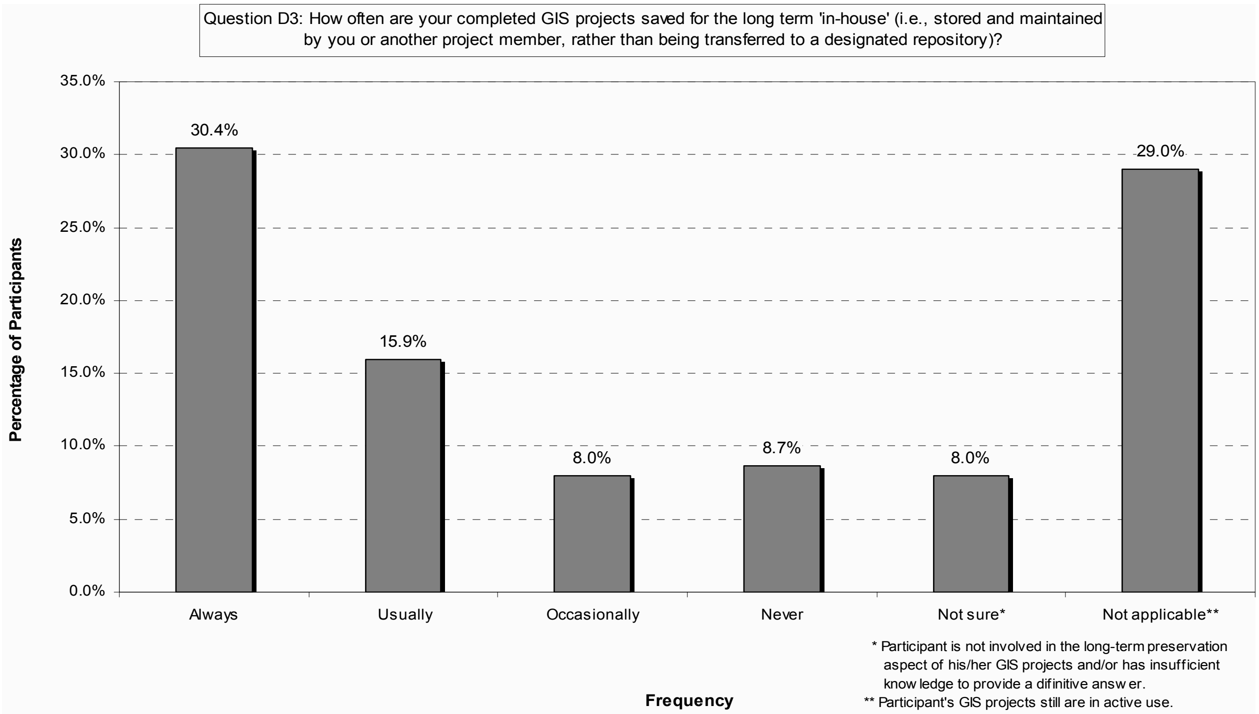


Figure 41. Frequency with which survey participants rely on "in-house" long-term preservation strategies for their GIS projects, rather than on preservation in a designated repository (n=138).

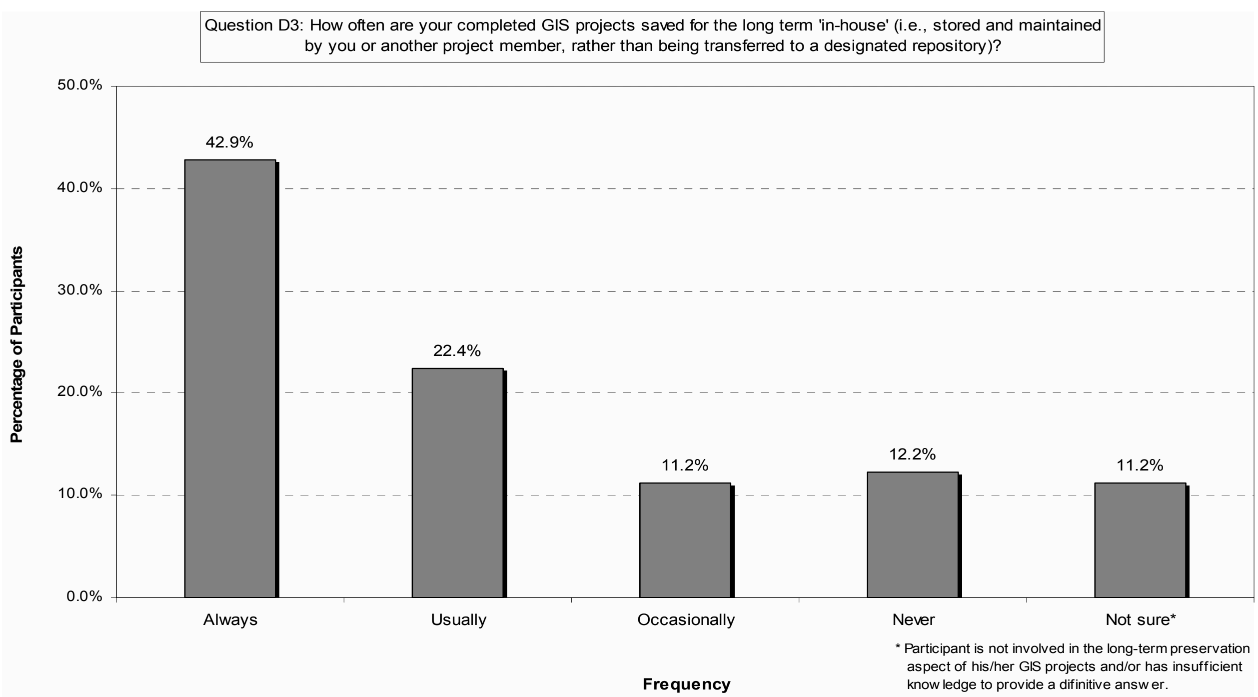


Figure 42. Frequency with which survey participants rely on "in-house" long-term preservation strategies for their GIS projects, rather than on preservation in a designated repository ("not applicable" responses excluded) (n=98).

As the following sampling of comments from Question D2 suggest, the low use of repositories for the long-term preservation of GIS projects appears to be more a consequence of the lack of suitable repositories than a reflection of any conscious reluctance on the part of the participants to transfer their completed projects into archival custody:

> Virginia is so backward, there's no means of doing so [i.e., transferring completed GIS projects to a designated repository]. I hope I live long enough to see a repository formed. I have been advocating such a system since 1982.
>
> --
>
> …the cultural agency in Spain is not prepared to receive this type of information and/or projects…
>
> --
>
> No such requirements or repositories [exist] for Kentucky or the US.
>
> --
>
> There is no such repository in existence in this state [state not specified].
>
> --
>
> In my country [Brazil] this kind of repository does not exist.

Indeed, in the wake of the termination of the Archaeological Data Archive in the United States, the ADS (see footnote 47) appears to be one of the only non-government repositories in the world at this time of any sizable scale that is willing and able to provide long-term preservation of digital archaeological data. However, as several of the participants pointed out, ADS's repository fees preclude it as a viable option for many. Another alternative mentioned by one of the survey participants is the Electronic Cultural Atlas Initiative (ECAI).[48] Better defined as a worldwide consortium of cultural researchers than as a repository, per se, the ECAI nevertheless may, within the parameters of its stated goals and objectives, provide some researchers with a viable alternative for the long-term preservation of their GIS datasets and mapspaces.

5.4.3 In-house Preservation Strategies

Data refreshment, a process that involves copying files from one medium to another as the original medium nears the end of its reliable shelf life and/or approaches technological obsolescence, is the most common long-term, in-house preservation strategy used by the survey participants (47.4%) (Figure 43). Significantly, nearly as many participants (42.3%) actually employ no particular long-term, in-house preservation strategies, despite indicating that they do "preserve" their GIS projects. In fact, what many of these participants said they do to "preserve" their projects is best summed up by the participant who wrote, "[I] put them on CD and hope I can access them in the future." In other words, "preservation" for close to half of the participants who save their projects in-house merely consists of backing up to CD/DVD. Data migration, which involves converting files into a format that can be read by current software, is used by one third (33.3%) of the participants, while some type of data documentation (e.g., documentation to track and explain data migrations, abbreviations used, file naming conventions, etc.) is used by

[48] Begun in the late 1990s, the ECAI acts as an electronic clearinghouse for historical and cultural data by maintaining a controlled catalogue of on-line data held by other agencies or individuals. Participants are first required to document their projects using ECAI metadata (which is based on Dublin Core) and then mount the projects on a web server. More information about the ECAI is available at http://www.ecai.org/.

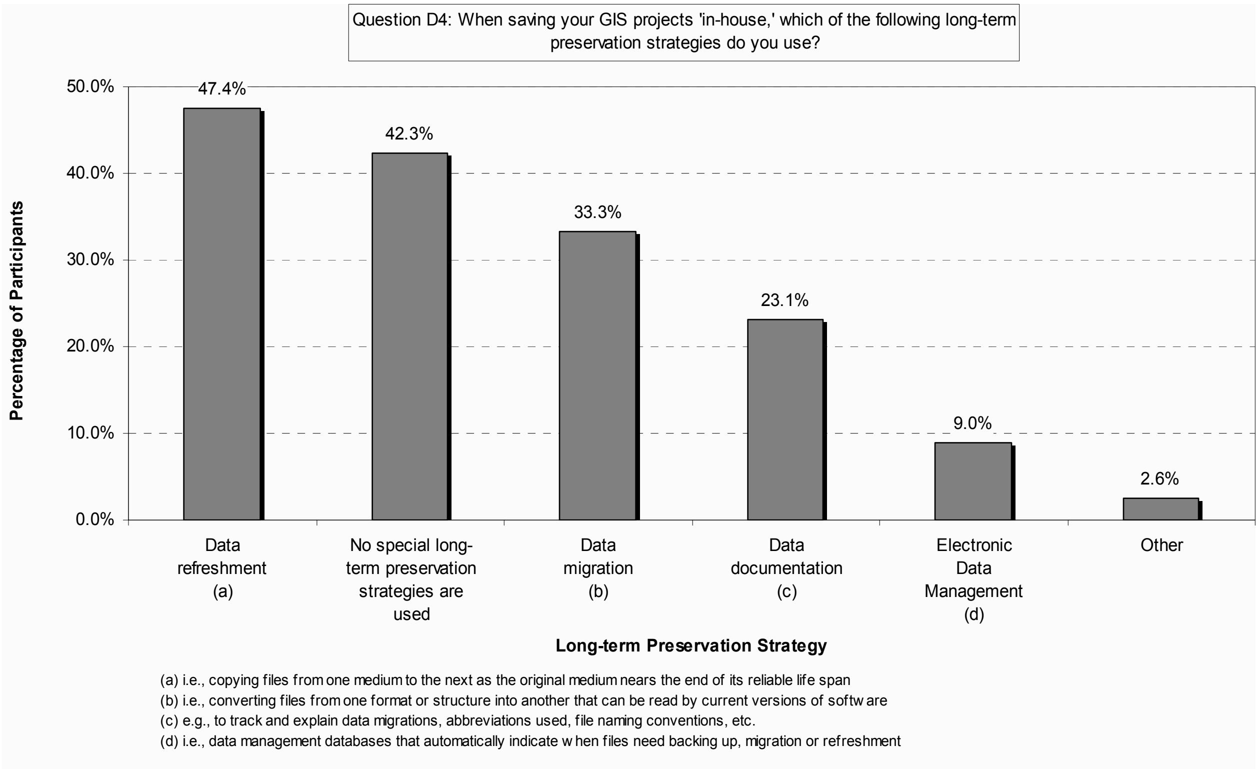


Figure 43. In-house, long-term preservation strategies
("not applicable" responses excluded) (n=78).

nearly one quarter (23.1%) of the participants. Only 9.0% of the participants take advantage of more sophisticated preservation aids, such as Electronic Data Management (EDM) systems.[49]

Given the relatively high percentage of participants who employ no particular long-term preservation strategies, coupled with the low percentage of participants who augment their preservation regimen with EDMs, it should come as little surprise that regardless of the actual preservation strategies used, more than half of the participants (52.6%) said that they implement those strategies in an irregular or subjective manner that follows no established, standardized or scheduled procedures, and without regard to special considerations for different files types or digital media (Figure 44). Another 23.7% said that although they tend to implement their preservation strategies systematically, they do so only for certain file types and digital media and not for others. In fact, just over 1 in 5 of the participants (21.1%) claim to implement their long-term preservation strategies systematically across all file and digital media types.

5.4.4 Structure of Preserved Projects

A large majority (79.6%) of the participants who save their projects for the long-term, whether in-house or in a repository, always or usually do so while preserving the organizational structure of the project's files and folders that was in place when the project still was in active use (Figure 45). Only 5.7% said that they never retain the project's original organizational structure when

[49] EDMs essentially are databases designed to automatically keep track of when files need backing up, migration or refreshment.

transferring it to a repository or preserving it for the long-term in-house. One of the participants noted that the decision of whether to retain the project's original organizational structure depends on whether the GIS project is Web-based. Another participant said that it depends on where the projects end up, with repository data being amalgamated into one central database and "local projects" being retained in their original file structure.

5.4.5 Factors Influencing Preservation Decisions

Question D7 sought to discern which of several key considerations were responsible for influencing which, or how many, of the participants' GIS projects were preserved for the long term. These considerations included: A. Insufficient funding, B. Insufficient time, C. Insufficient personnel, D. Lack of an available/applicable repository and E. Not seen as important/ necessary.[50] In addition to specifying which considerations influenced their preservation decisions, the participants were asked to rank their choices according to what they considered to be the most and the least important, or influential, considerations.

Although 'insufficient time' emerged as the most commonly cited consideration (57.1%), it nevertheless ranked last in terms of relative importance (Figure 46). In fact, only 20.0% of the 57.1% of participants who cited 'insufficient time' as a consideration that influences their preservation decisions judged it to be the "most"[51] influential consideration. At the other extreme is consideration 'E. Not seen as important/necessary,' in which a clear majority (71.4%) of the 40.0% of participants who cited this as a consideration that influences their preservation decisions deemed it to be the most influential consideration, while just 10.7% felt that is was the least influential consideration. Equally intriguing is the fact that only 38.6% of the participants chose 'D. Lack of an available / applicable repository' as a factor that influences their long-term preservation decisions, particularly considering the number of comments received for Question D2 lamenting the lack of available repositories. This result appears to be heavily influenced by the fact that, as summarized above, considerably more participants avail themselves of in-house preservation than preservation in a repository. Thus, while many participants consider the lack of available repositories to be a regrettable situation, most simply compensate by preserving their projects in-house, rather than allowing the lack of available repositories to significantly influence their preservation decisions. Had the question asked the participants to differentiate between those factors that influence their decisions to preserve their projects in-house verses those that influence their decisions to use a repository, it is likely that the results would have been quite different for consideration "D."

Finally, a couple of the comments received for this question are worth quoting, as they seem to speak to the broader societal mindset with respect to the issue of the long-term preservation of digital records:

[50] Obviously, these are not all mutually exclusive considerations. In fact, consideration A (funding) undoubtedly has both direct and indirect influences on the available levels of considerations B (time) and, especially, C (personnel). The reader is therefore advised to bear this in mind when drawing any inferences from the results.

[51] It should be noted that a number of participants listed multiple considerations as being the most and/or least important. Thus, while one may speak of the "most" or "least" influential consideration for the participants as a whole, it is acknowledged that doing so obscures the situation somewhat with respect to certain participants.

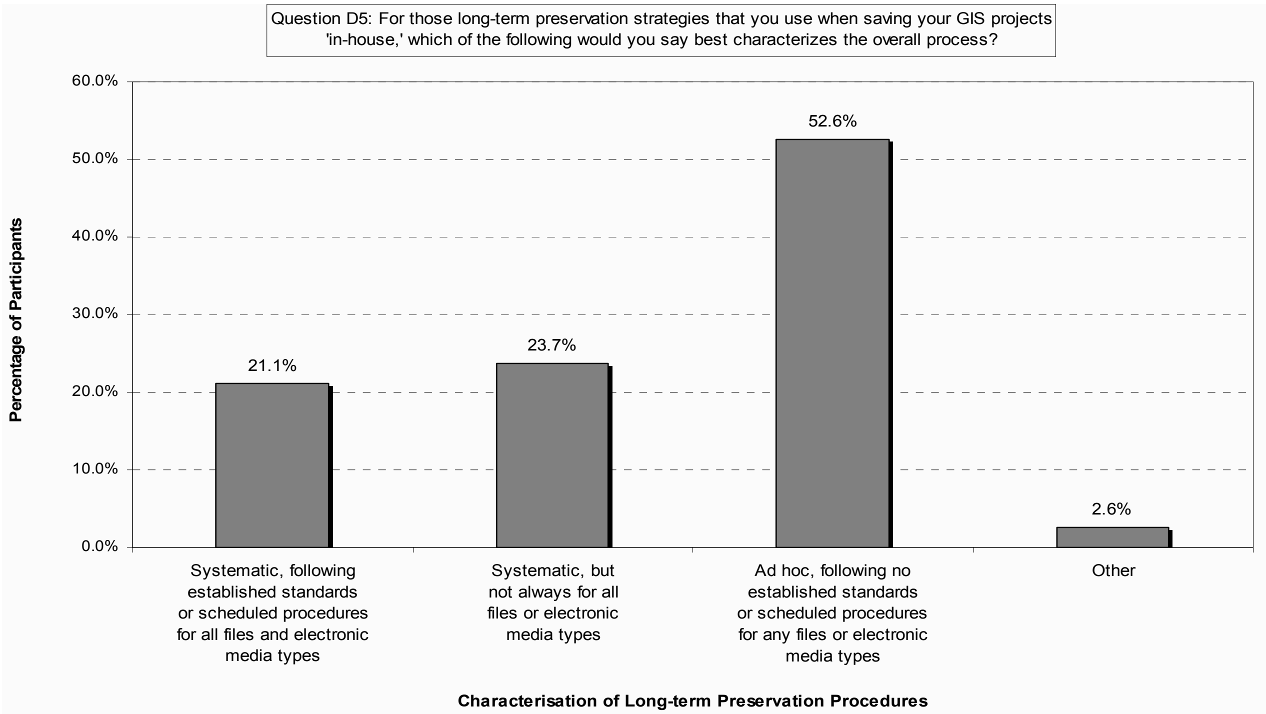


Figure 44. Overall characterisation of in-house, long-term preservation procedures used by survey participants (“not applicable” responses excluded) (n=76).

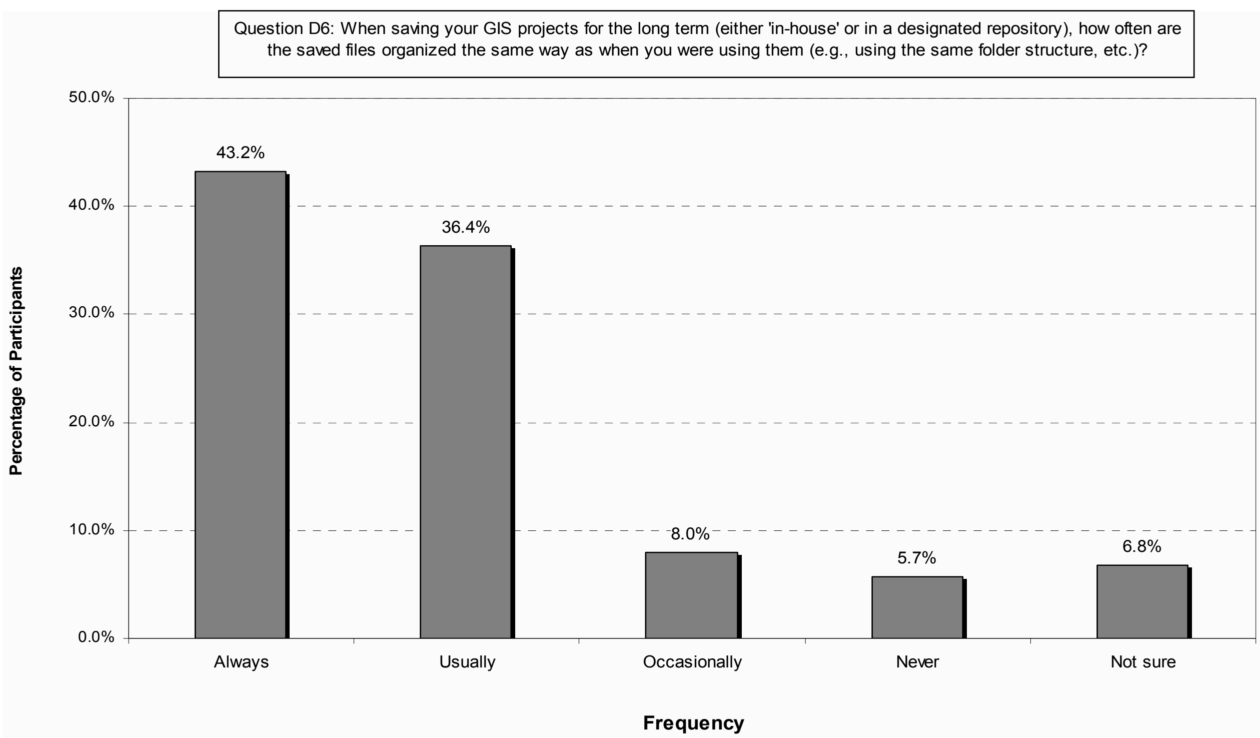


Figure 45. Frequency with which survey participants retain the original file organization of their active GIS projects when preserving them for the long-term once they are inactive (“not applicable” responses excluded) (n=88).

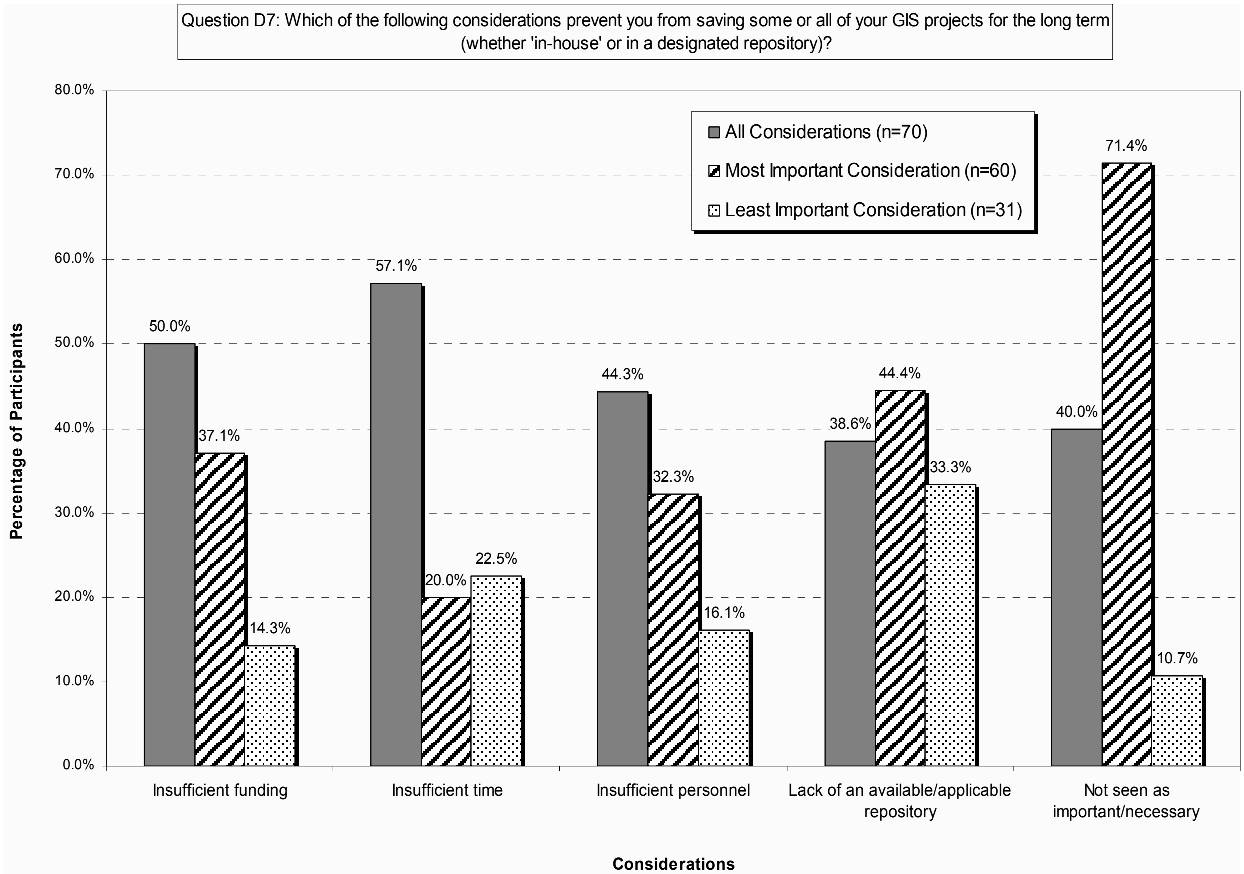


Figure 46. Key factors identified by survey participants that prevent the long-term preservation of their GIS projects ("not applicable" responses excluded) (n=70).

> California doesn't see it as necessary as a whole. It's only important on a personal level, but no policy mandates proper storage.
>
> --
>
> Long-term storage of GIS data is still viewed in some European institutions as unnecessary and sometimes unwanted.

5.4.6 Metadata: Information Recorded

Of those participants who answered Question D8, 18.6% indicated that they do not record metadata for their GIS projects, while another 3.9% were not sure if they did. Among those who do, more than half said that they routinely include metadata about the following information: (1) file format (80.0%), (2) date of file creation (77.0%), (3) file creator's name (72.0%), (4) name of the software used to create the file (66.0%), (5) date when the file was last updated (63.0%) and (6) version of the software used to create the file (56.0%) (Figure 47). Fewer than half routinely include metadata about: (1) the procedures used to create the file (i.e., how the different components of the GIS system work together to maintain functionality) (44.0%), (2) the computer operating system used to create the file (28.0%) and (3) the hardware used to create the files (23.0%). Other metadata elements routinely recorded by 8.0% of the participants included one or more of the following: (1) analysis methodology (e.g., type of interpolation used), (2) mapping/digitization scale, (3) GIS type (i.e., point, polygon or line), (4) projections, (5) precision of measurements, (6) types of data used and (7) copyright (including information about the source or supplier of the original data and any special conditions under which those data are supplied).

5.4.7 Metadata: Procedures Used

Of those participants who record metadata for their projects, fewer than one quarter (23.0%) said that they always or usually follow an established metadata schema or standard, such as Dublin Core (Figure 48). In contrast, more than half (53.8%) said that they never follow an established standard, this despite the fact that 35.9% of these participants are aware that metadata standards are available. Participants in Gourad's survey were asked to comment on whether the metadata generated for their projects met USGS metadata standards and what effect this knowledge had on their projects. Gourad found that 43% of his participants were not familiar with the standard, while 29% were, but nevertheless had not considered how it impacted their projects. Another 21% said that they had considered its impact on their projects, while just 6% said that they had actually modified their projects to account for it.[52] Although the nature of the questions asked precludes any direct comparison between the two surveys with respect to the issue of metadata, some general observations are worth discussing. Most obvious is the overall shift toward increased awareness of the issue of metadata and metadata standards, with a much lower percentage (17.9% vs. 43%) of participants in the current survey claiming to be unaware of the existence of metadata standards, and a much higher percentage (39.7% vs. 6%) claiming to have actively put that awareness to use in their projects.[53]

The comments recorded for Question D9 indicate that even among those who use an established standard, application of the standard rarely, if ever, is wholesale. Complexity of the available standards appears to be one impediment to their more widespread and wholesale application. In the words of one participant, "…we tend to make up our own metadata standards, which is more simplistic but suited to our own requirements. No-one in this organization [a public CRM organization in the United Kingdom] knows how to deal with the more complicated Dublin Core standards—would be nice to have proper training on metadata management though." Application of metadata standards also depends, in part, on the nature of the project. One participant only uses Dublin Core standards when doing academic-based GIS projects, for example. In a similar sentiment that speaks to the influence that both of these factors (i.e., metadata complexity and project type) have on the overall use of metadata, one participant, in his/her comments to Question 11, emphasizes that, "for routine work on short-term projects…extensive metadata documentation is not cost effective."

[52] Gourad, *supra* note 2, pp. 51–52.

[53] One must bear in mind, however, that Gourad's survey dealt specifically with participant awareness of a particular metadata standard, whereas the current survey made no such distinction. In fact, one generally would expect to see the overall level of awareness diminish in relation to increasing or narrowed specificity. Nevertheless, the magnitude of the difference in the results between the two surveys does suggest a real increase in overall awareness and implementation.

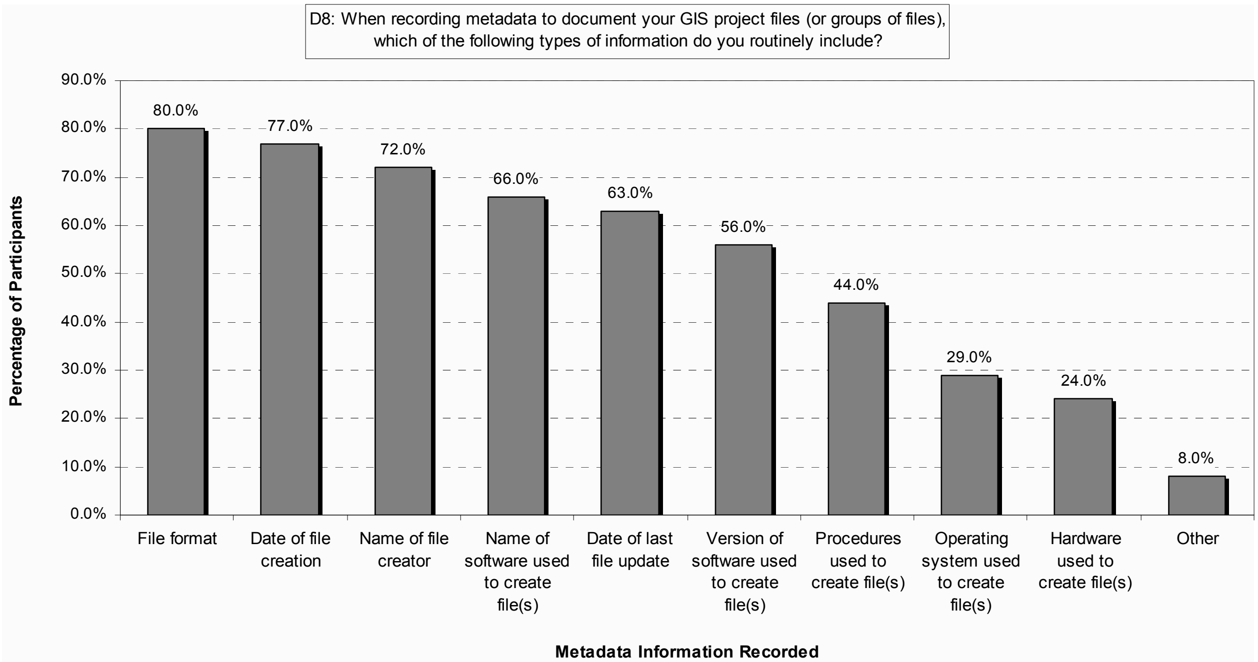


Figure 47. Types of metadata used by survey participants to document their GIS projects ("not applicable" and "not sure" responses excluded) (n=100).

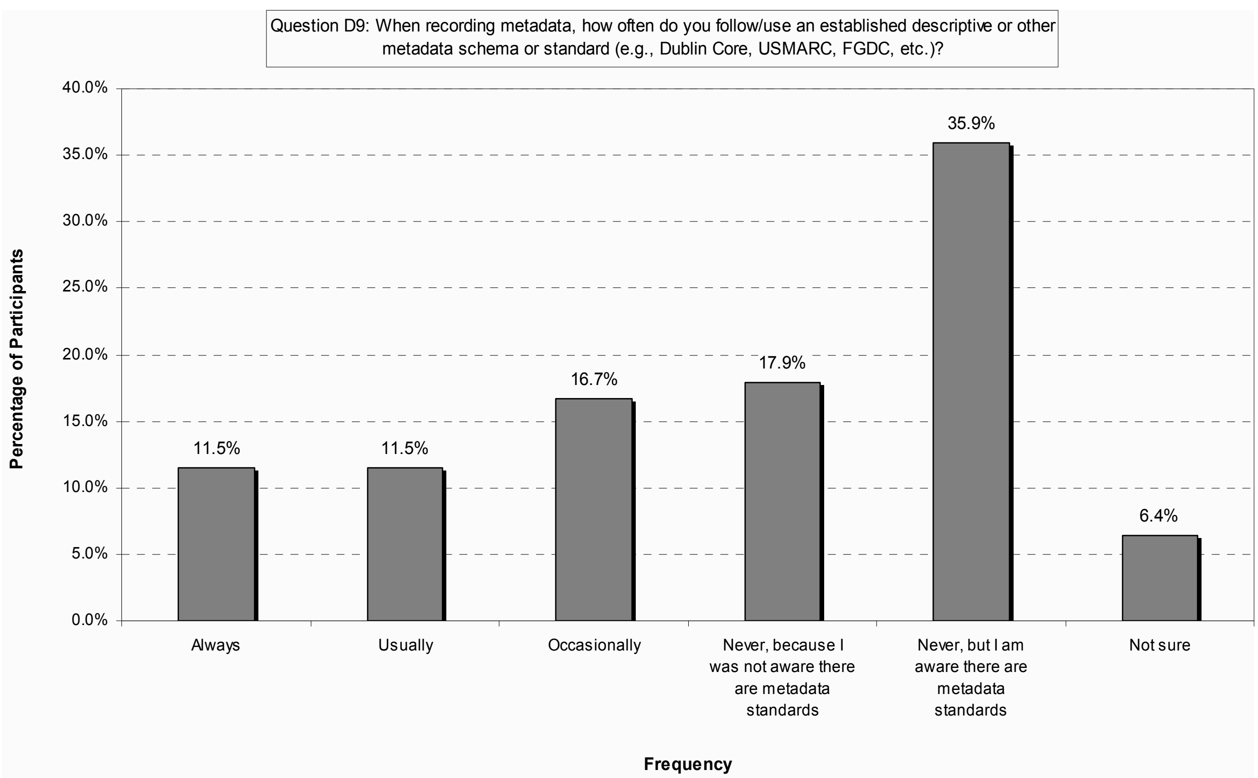
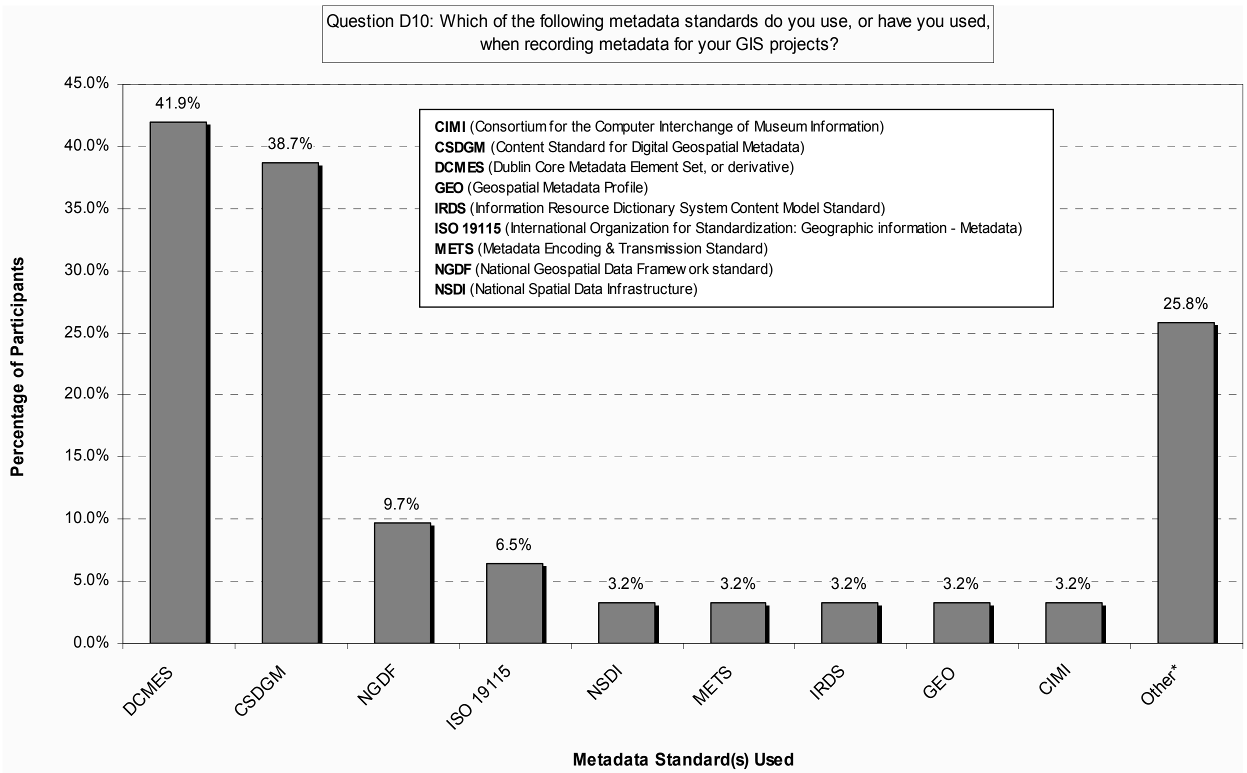


Figure 48. Frequency with which survey participants follow an established standard when recording metadata (n=78).

5.4.8 Metadata: Standards Used

As shown in Figure 49, the two most commonly utilized metadata standards (or derivatives thereof) are the Dublin Core Metadata Element Set (DCMES) and the Federal Geographic Data Committee's (FGDC) Content Standards for Digital Geospatial Metadata (CSDGM), cited respectively by 41.9% and 38.7% of those participants who record metadata for their GIS projects. The National Geospatial Data Framework (NGDF) standard and ISO standard 19115 (Geographic Information—Metadata) follow in a distant third (9.7%) and fourth (6.5%) place, respectively. Nearly one third (32.3%) of the participants use, or have used, two or more standards.



* Includes: MIRIS (metadata files for Michigan), "In-house" metadata, Landlýsing (Icelandic metadata standard), Museum of London archive guidelines, as well as one respondant who was not sure which standard was used, and two respondants who did not specify what the "other" standards they used were. (n=8)

Figure 49. Metadata standards used by survey participants (n=31).

5.4.9 Metadata: Methods of Association

At present, three options exist for associating metadata with the resource it describes: (1) linking, (2) embedding and (3) a combination of linking and embedding. Linking involves creating pointers within the resource being described to one or more referenced items containing the actual metadata descriptions for that resource. These pointers, or links, can be simple descriptions, such as text citations, to the referenced items, or they can be computational objects, such as hypertext links. Conversely, embedding involves including the actual metadata descriptions within the resource being described. A combination approach uses both linking and embedding, although not necessarily simultaneously within a single resource, to associate metadata descriptions with the resources being described. As used here, the combination approach refers to situations in which both linking and embedding simultaneously are used within a single resource, as well as situations in which both techniques are used, but not simultaneously within a single resource.

When asked which method typically is used to associate the metadata with the resources being described, 37.5% of those participants who record metadata said they typically use a combination approach in which their decision to use linking or embedding varies depending on the specific resource type (e.g., textual vs. graphic) being described (Figure 50). Among participants who typically use only one technique, embedding is employed more often than linking (28.1% vs. 18.8%). Curiously, 15.6% were not sure which of these three techniques they typically used.

5.4.10 Metadata: Use of Controlled Vocabularies

In their most basic form, controlled vocabularies consist of alphabetized pick lists of authorized and/or standardized terms, each of which is used to name or describe a specific concept, entity, etc.[54] More comprehensive controlled vocabularies include synonym rings of non-preferred, equivalent and related terms arranged into various taxonomies and hierarchies that often are presented in the form of a thesaurus. Among the many advantages of using controlled vocabularies are that they help improve the comprehensiveness and precision of search results both within and, when using the same vocabulary, across projects.

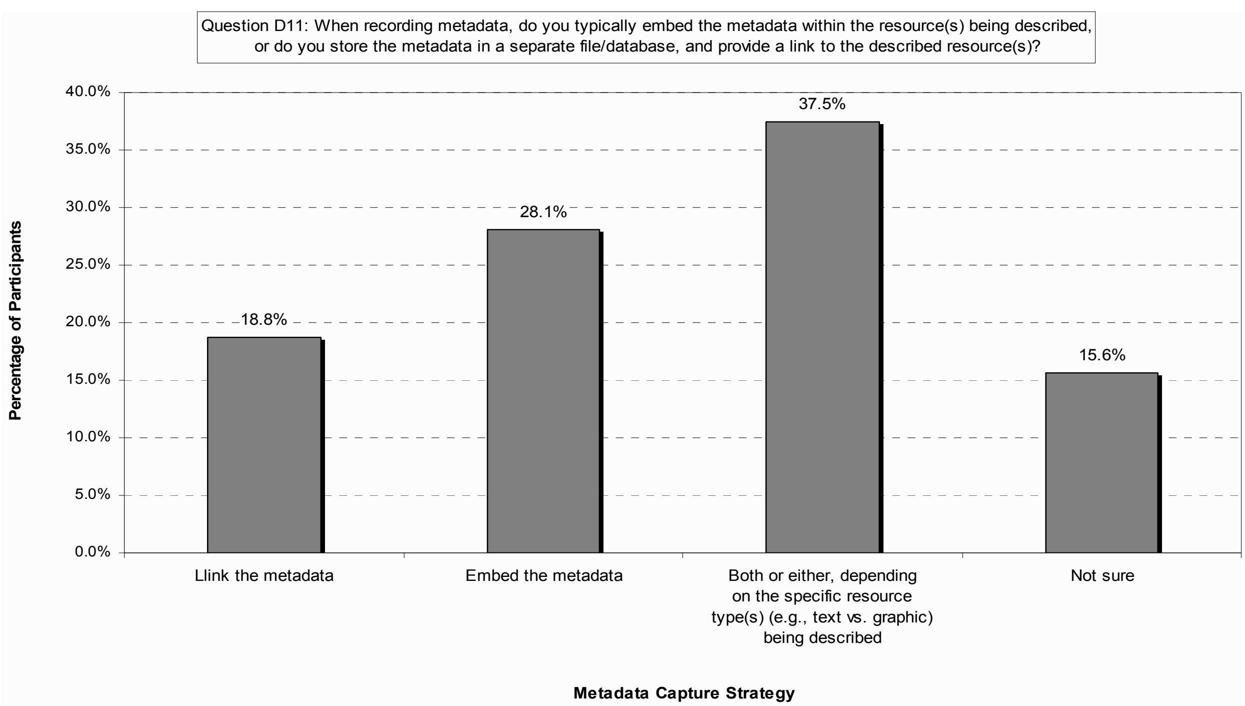


Figure 50. Relative reliance by survey participants on linking vs. embedding strategies in relation to metadata management (n=32).

[54] As used here, controlled vocabularies are synonymous with "name authority files."

Among those participants who record metadata for their GIS projects, more than half (54.9%) always or usually derive their metadata terms from some form of controlled vocabulary (Figure 51). Significantly, 16.1% said controlled vocabularies were never used, while another 9.7% were unsure. One participant noted that his/her use of a controlled vocabulary was project-dependent.

5.4.11 Most Important Elements of GIS Projects for Long-term Preservation

The final question in this section was a free-text question that asked participants to briefly describe which elements and/or outputs of their GIS projects they thought should be preserved for future use or reference and why. The primary intentions of this question were two-fold: (1) to ascertain the degree to which the participants were cognisant of, and concerned about, the issue of long-term preservation with respect to their projects and (2) to better understand which aspects of their GIS projects the participant's considered to be especially important to preserve.

Unfortunately, several limitations inherent in the structure of this question preclude all but relatively broad generalizations and tentative conclusions to be drawn from the results. First, as with question C11, the location of this question in the survey, situated at the end of a series of question about the participants' digital preservation practices, raises the question of whether, or to what degree, the comments received were influenced by the information presented in the preceding questions. Secondly, the wording of the question itself appears to have had a strong influence on the responses received, with quite a number of the participants referring specifically to the GIS element and output examples cited in the survey question (i.e., original reports, final reports, maps and tabular data). Finally, the free-text structure of the question resulted in highly variable responses. In fact, it was impossible in many cases to satisfactorily separate the responses (or portions of responses) into specific element categories, necessitating the creation of several broad element categories, such as "unspecified maps" and "unspecified data." In retrospect, this question would have been more informative and effective had it been moved to the beginning of the section and been presented in multiple choice, rather than free-text, format.

Despite these limitations, however, several general observations are worth noting. As detailed in Figure 52, and more generally summarized in Figure 53, the participants identified a number of key elements that can be divided into five broad categories, including: (1) *documentation elements*, such as original, interim and final reports, project documentation and metadata, and archaeological site records and field notes, (2) *data elements*, such as base, derived and interpreted data, (3) *graphic elements*, including original base and final derived maps, aerial photos and other images, (4) *functional elements*, such as working copies of GIS deployment and reader software and (5) *"general" elements*, which simply represents a catch-all category for those responses which stated either that "everything" should be saved or that the decision about what to preserve varied from project to project. It merits noting that 6.6% of the participants *explicitly* emphasized the importance of preserving the data, documentation and graphic elements in both analogue (i.e., hardcopy printout) and digital forms, thus highlighting a certain level of awareness among participants regarding the issue of long-term digital preservation.

Among the top three specific elements cited by participants for long-term preservation were two documentation elements, including final reports and general documentation/metadata about the projects, along with base, or original, spatial and tabular data (see Figure 52). There appears to be a clear consensus on the importance of the *original base* spatial and tabular data, relative to the *interim* and/or final *derived* data, with 31.1% of participants specifically citing the original base data as an element worthy of long-term preservation, in contrast to just 3.8% and 10.4% of participants affording the same importance to interim and final derived data, respectively. Moreover, the relative importance of base data is further emphasized by the fact that "base maps and images" (cited by 8.5% of participants) and "archaeological site records/field notes" (cited by 3.8% of participants) are themselves types of base data. As was emphasized by several of the participants, the primacy afforded base data elements reflects the fact that they are, in many cases, the most difficult element to recreate. In fact, given the inherently destructive nature of archaeological excavation, base site data are utterly impossible to recreate once lost. Underscoring the relative emphasis on preserving base data is the comment of one participant who noted that, "[derived] maps and reports can be recreated from the data and the documentation providing the steps of analysis." Derived data and graphic elements, on the other hand, are the easiest elements to recreate, provided of course that the underlying data and, preferably, the documentation of the steps involved in manipulating those data, are preserved.

Not too surprisingly, data elements (of any kind) were identified as the most important element to preserve long-term, being overtly cited as such by nearly three quarters (73.6%) of all participants (see Figure 53). Perhaps more significant, at least from a recordkeeping perspective, is the fact that nearly as many participants (67.0%) identified some type of documentation as an

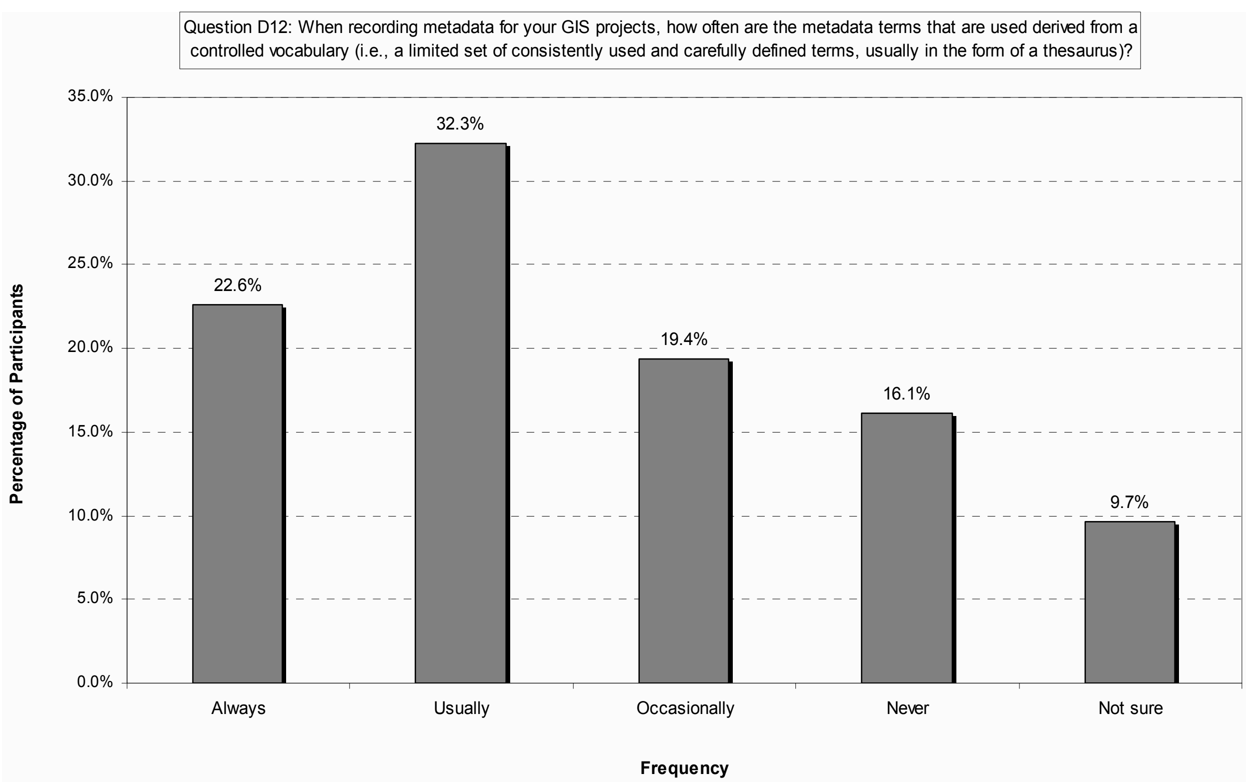


Figure 51. Frequency with which survey participants use controlled vocabularies when recording metadata for their GIS projects (n=31).

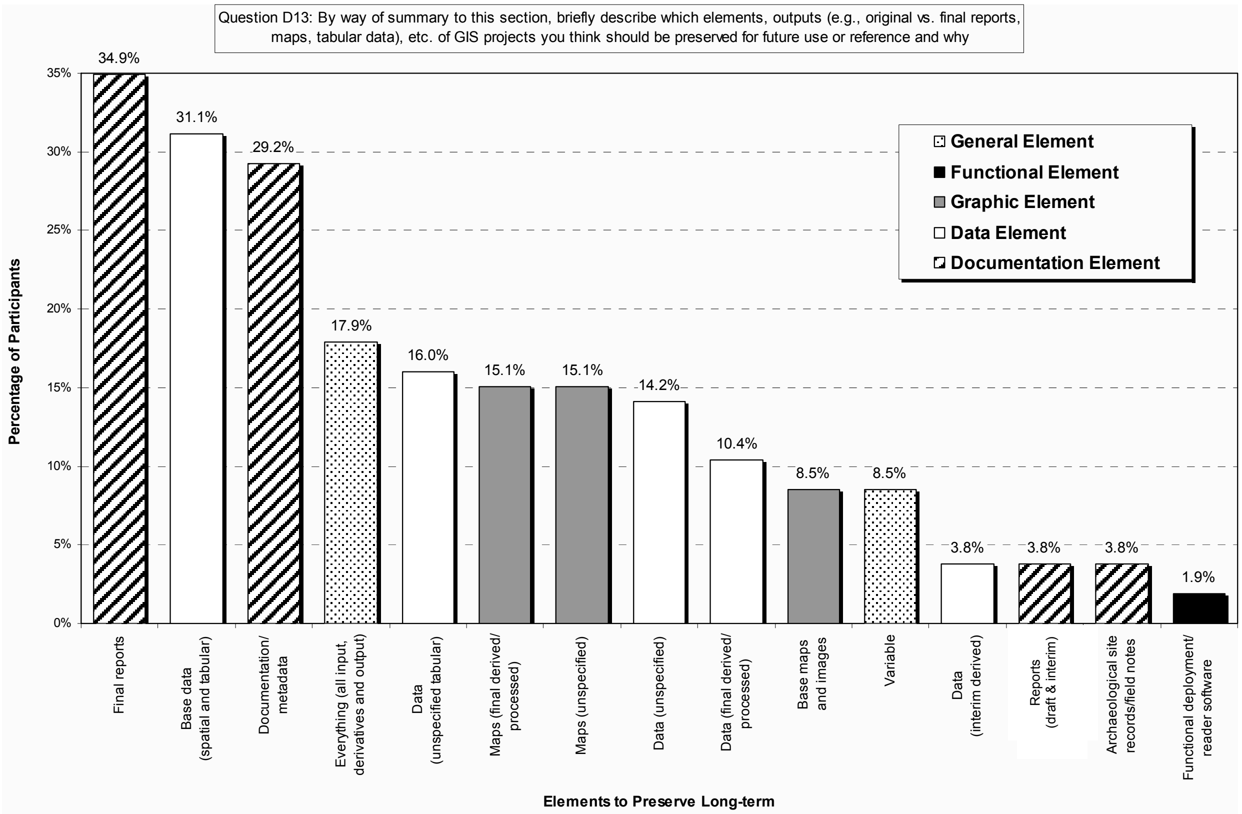


Figure 52. Detail of GIS project elements deemed desirable by survey participants to preserve long-term (n=106).

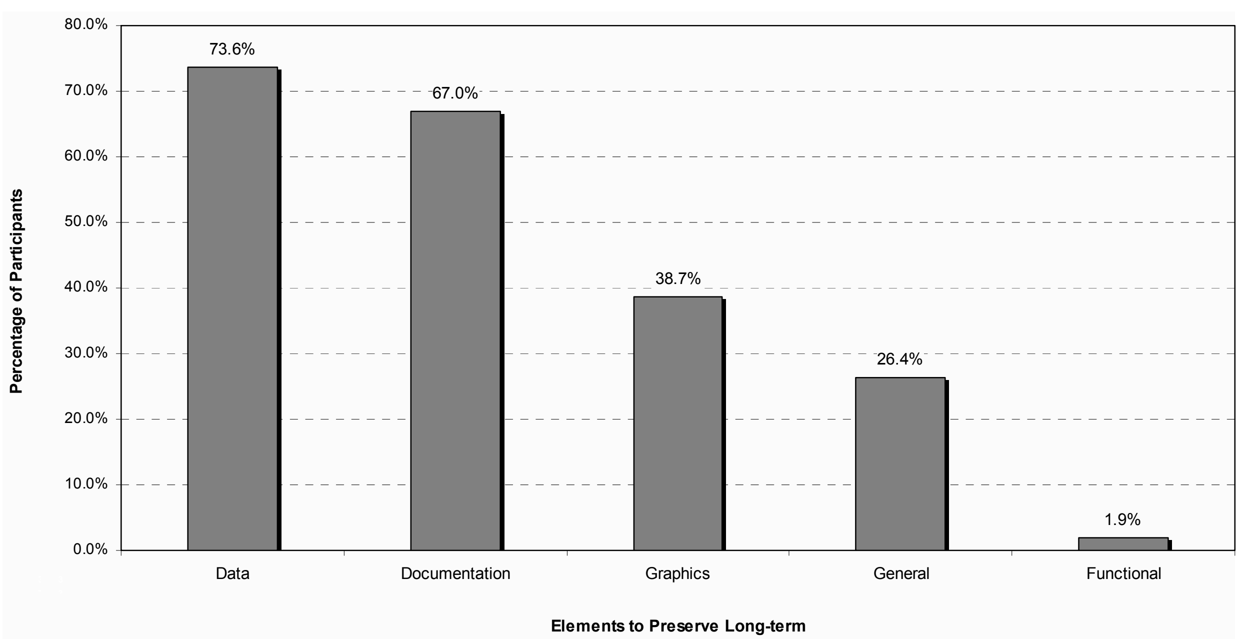


Figure 53. Summary of GIS project elements deemed desirable by survey participants to preserve long-term (n=106).

important element worthy of long-term preservation. Upon closer inspection, however, the significance of this situation is attenuated somewhat by the fact that the most commonly cited documentation element (cited by 34.9% of participants) involved some type of final project report, with more general project documentation and metadata trailing in second place at 29.2% (see Figure 52). From an archivist's perspective, it is the latter types of records that would be the most crucial and useful records to preserve, since these are the records that, if generated and maintained properly, would provide the most comprehensive and least biased account of the entire evolution of a project. In contrast, final reports, due to their 'polished' summative nature, typically provide a far more restricted and biased perspective on the overall evolution of projects. For reasons that will become more evident in a moment, the significance of this distinction is a critical one for archaeologists to grasp, as it has important ramifications for the efficacy of their long-term preservation efforts.

So what can be said about *why* the participants selected the elements they did as worthy of long-term preservation? A careful reading of the participants' responses to Question D13 shows that the factors influencing selections can be grouped into five broad categories, including: (1) the potential for any *re-use* of the GIS project or its data by other researchers, (2) the desire to enable *verification* of the project's data, analyses, methodology, documentation and/or system functionality by other researchers, (3) the desire to enable actual *replication* of the project by other researchers, (4) the need to provide a measure of *accountability* should the project's data or results be challenged and (5) the physical, logistical and financial *practicalities* of long-term preservation. Nearly one half (45.3%) of the participants identified re-use as a factor that influences their decisions regarding the long-term preservation of their GIS projects, while concerns about project verification were cited by just over one quarter (26.4%) of the participants (Figure 54).[55] By far the least influential factors were accountability and practicality, both of which were cited by just 5.7% of the participants.

5.5 Section E: Data Input/Output Practices

The primary purpose of the questions in this section was to gain a better understanding of the typical procedural habits and controls of the participants with respect to the creation of the various digital components and outputs of their GIS projects. A thorough understanding of the procedural habits and controls associated with the creation of records (digital or otherwise) is necessary for adequately assessing the *reliability* of those records. As used in this context, *reliability* refers to the trustworthiness of a record as a statement of fact and exists when a record can stand for the fact it is about. Reliability is established by examining the completeness of the record's form and the amount of control exercised on the process of its creation.[56] As such, the concept of reliability is exclusively and inextricably linked to the process of record making. It is important to bear in mind, however, that reliability is a *relative* concept in which a record may be judged to be more or less reliable according to the degree to which any established rules or pro-

[55] As noted in Figure 54, these percentages are based on a sample size of 53. This smaller sample size is due to the fact that only half of the 106 participants who answered question D13 included any discussion of their rationale for selecting the elements of their GIS projects that they identified as important to preserve for the long-term.

[56] This definition of *reliability* is paraphrased from the InterPARES 2 Project *Terminology Database Glossary*, *supra* note 1.

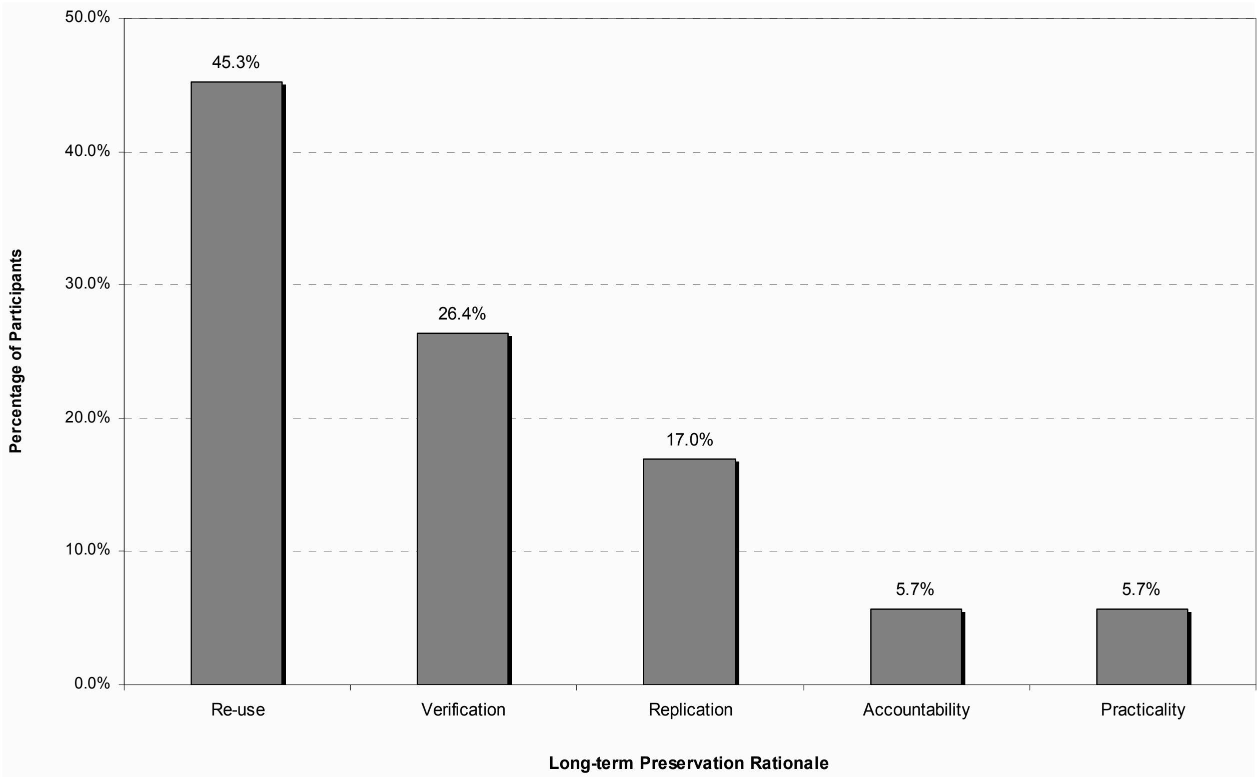


Figure 54. Summary of reasons cited by survey participants for the long-term preservation of identified GIS project elements (n=53).

cedures for making it have been respected.[57]

In the context of digital records, the degree of reliability is perhaps the most crucial element of a record; more important even than a record's authenticity. In fact, as Duranti emphasizes, "Authentic, unreliable record are of no use to present and future users…Users need to know that the record was made under controlled circumstances as part of the regular workflow…and that it was generated by somebody who was competent to make that specific record, with either the duty or the direct interest to make it accurate."[58] When dealing with digital records, the importance of establishing their reliability cannot be overemphasized. Unfortunately, however, most current digital records creation practices do little to ensure the reliability of records. Duranti unequivocally characterizes the essence of this situation when she writes,

> …the easiness of electronic records creation and the level of autonomy that it has provided to records creators, coupled with the exhilarating sense of freedom from the chains of bureaucratic structures, procedures, and forms, have produced the sloppiest records creation ever in the history of record making. Too many persons

[57] L. Duranti, T. Eastwood and H. McNeil (1997), "Template 7: How is a Record Created Reliable in the Electronic Environment?" in *The Preservation of the Integrity of Electronic Records* (Vancouver: The University of British Columbia, School of Library, Archival and Information Studies). Available at http://www.interpares.org/UBCProject/tem7.htm.

[58] L. Duranti, "Reliability and Authenticity: The Concepts and Their Implications," *Archivaria* 39 (1995): 8–9. PDF version available at http://journals.sfu.ca/archivar/index.php/archivaria/article/download/12063/13035.

> and too many records forms generated in too many different contexts participate in the same transaction; too much information is recorded; too many duplicates are preserved; and too many different technologies are used. In other words; electronic records, as presently generated, might be authentic, but they are certainly not reliable.[59]

5.5.1 Systematisation of Procedures

Just over half (51.6%) of the participants who answered Question E1 said that, overall, they always or usually follow a routine or systematic sequence of procedures when creating the various files and outputs for their GIS projects, while nearly one out of every ten (9.8%) said that they never do (Figure 55).[60] A little more than one third (38.5%) said they occasionally follow a routine or systematic sequence of procedures, or else that it varies depending on the type of file or output in question. According to many of the comments received for Question E1, the creation of certain elements, such as maps and TINs (triangulated irregular networks), and the execution of certain data creation and modification activities, such as those associated with the digitization of map data or the implementation of GPS and geophysics routines, generally are more "rule bound," and thus typically involve more standardized or structured procedures.[61] Several participants noted that they often create templates and flow charts, or set up batch processes, to help them automate many of these activities and outputs. On the other hand, certain data acquisition and post-processing activities, such as scanning section drawings or rectifying digital photographs, as well as certain activities associated with the actual analysis of the data, such as producing views, typically are far less rule bound and therefore subject to greater procedural variation.

5.5.2 Reliance on GIS Procedures Manuals

Overall, just 31.7% of those participants who answered Question E2 said that the procedures they follow (whether systematic or not) when generating the various component and outputs of their GIS projects always or usually are based on, derived from, or guided by, procedures from a GIS procedures manual (Figure 56). Nearly as many (30.0%) said that they never derive their procedures from a GIS procedures manual (no doubt due, in part, to the fact that a majority of these individuals do not have access to such a manual). The remaining 38.2% said that they only occasionally consult a procedures manual, or that their use of a manual varies from one project (or project component) to the next. Not too surprisingly, comparative analysis of the results from questions E1 and E2 reveals that those participants who always or usually use a systematic set of procedures are nearly three times as likely (30.7% vs. 11.4%) to also use a procedures manual as are those who only occasionally or never follow any systematic set of procedures.

---

[59] Ibid, p. 9.

[60] For some of these individuals, the absence of systematic procedures may be due, at least in part, to the fact that 50% reported having no access to a GIS procedures manual.

[61] The inclusion of the qualifiers "generally" and "typically" are intentional, as not all participants would agree with these characterizations. As one participant noted, for example, "an overview map can be created in several ways (sequence of data layers added, when [to] add [the] north arrow, etc.)."

5.6 Section F: Record Quality, Reliability & Authenticity Issues

5.6.1 Concepts of Accuracy

The first question in this section was a free-text question that asked participants to briefly describe or define what "accuracy" meant to them with respect to their GIS projects and/or their individual data files. The primary purpose for asking this question was to better understand how the participants view the concept of accuracy and how the concept is applied within the context of their archaeological GIS projects. The use of the term accuracy has a particular meaning in an archival context that does not always coincide with its use in an archaeological GIS context. For example, the InterPARES 2 Project Glossary defines *accuracy* as: "the degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter."[62] This is a very broad, all-encompassing definition that conflates a number of concepts that in many scientific contexts are considered to be distinct, including, for example, precision, accuracy and pertinence. To avoid confusion, therefore, it is important that archivists and archaeologists understand the nuanced ways in which each uses these terms.

5.6.1.1 Background on Accuracy and Related Concepts

Accuracy and precision often are, mistakenly, used interchangeably. This may be due, in part, to the fact that they are indeed related; a relationship that was succinctly summarized by Richards and Ryan when they wrote, "Precision implies that the degree of measurement of an attribute is refined; accuracy that the measurement taken is correct within the degree of precision indicated."[63] Although subtle, the distinction between these two concepts is an important one to understand, especially within the context of GIS research, which commonly integrates datasets of widely varying accuracy and precision.

As used in empirical scientific research, *precision* refers to how *exact and reproducible* a measurement or estimate is, irrespective of its accuracy, while *accuracy* refers to how *close* a measurement or estimate is to the correct value. For example, while a scientific instrument may be able to produce very precise measurements (i.e., repeated measurements that are always, or nearly always, the same), this does not guarantee that those measurements are accurate. The classic illustration of this distinction is a dartboard, as shown in Figure 57. The first board provides an example of a series of randomly scattered darts (or repeated measurements) that are neither precise nor accurate. The series of darts in the second board are tightly clustered within the bulls-eye and therefore are both precise and accurate. Although tightly clustered in the third board, the cluster of darts is far from the bulls-eye and therefore is precise, but inaccurate. The final board shows an imprecise but accurate pattern of darts loosely clustered around the bulls-eye. Precise but inaccurate data typically are the result of systematic error introduced during data collection. Systematic error is introduced, for example, whenever a rodman has a tendency to consistently hold the stadia rod at a slight angle, or whenever a theodolite, GPS or other measuring instrument consistently is calibrated incorrectly.

[62] InterPARES 2 Project, *Terminology Database Glossary*, *supra* note 1.
[63] J. D. Richards and N. S. Ryan (1985), *Data Processing in Archaeology* (Cambridge: Cambridge University Press), p. 20.

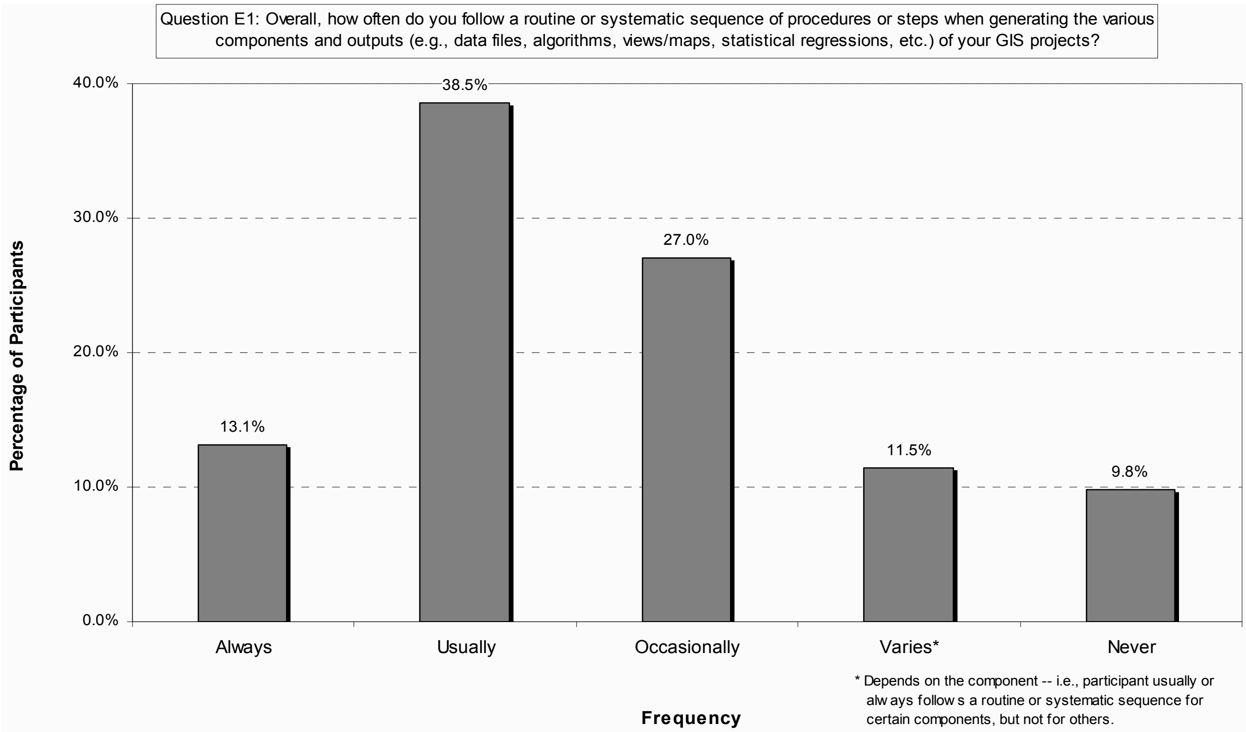


Figure 55. Frequency of use of routine procedures by survey participants when generating various GIS components and outputs (n=122).

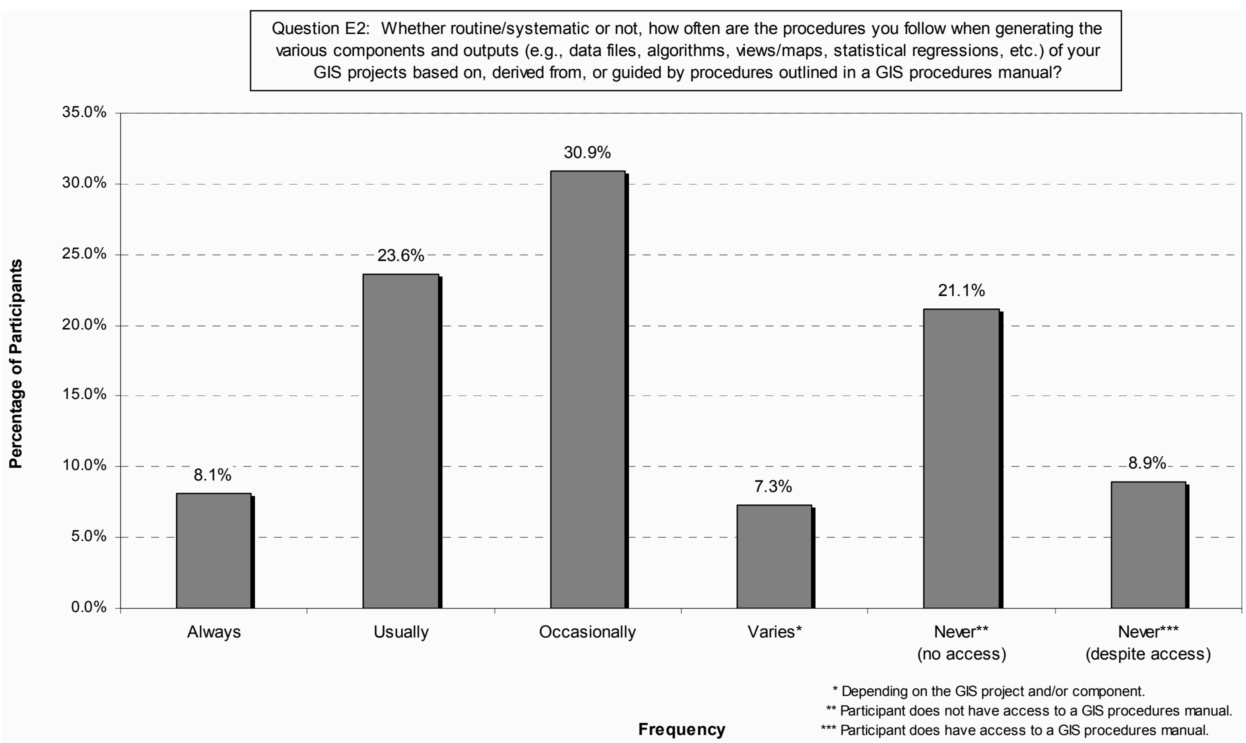


Figure 56. Frequency of use of a GIS procedures manual by survey participants when generating various GIS components and outputs (n=123).

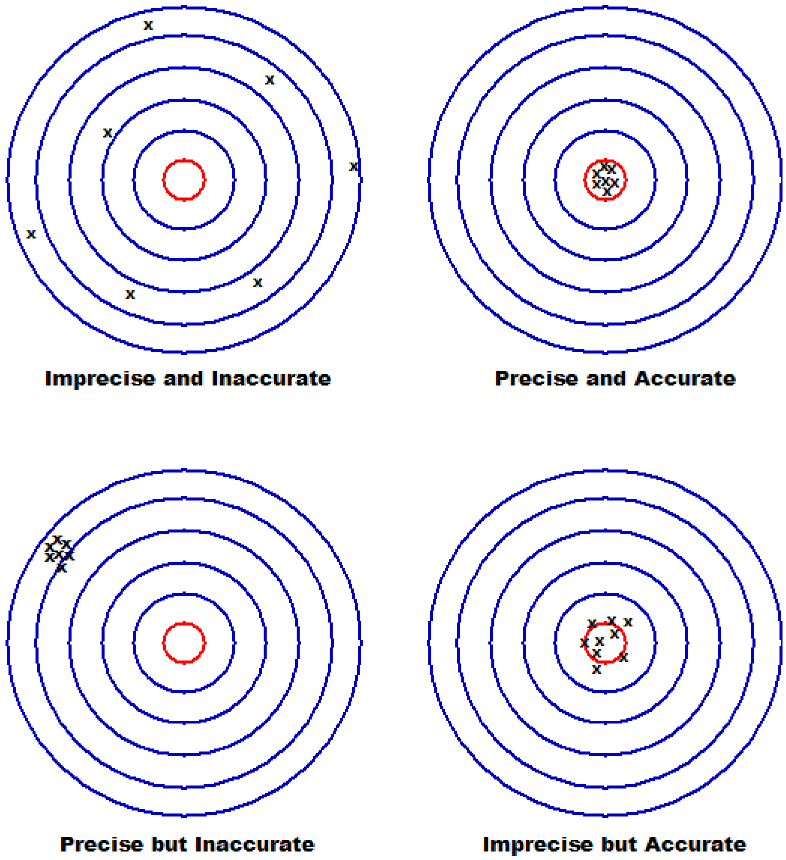


Figure 57. Precision versus accuracy in scientific measurement.

Pertinence, refers, in general terms, to the quality or state of something that has "a clear and decisive relevance to the matter in hand."[64] More specifically, within the realm of scientific research, pertinent information or knowledge is that which both fits within "the general pattern of a larger area [of research]," and, "is a recognizable and recognized part of the consensus among experts in that area."[65]

Considerations of data accuracy, precision and pertinence constitute some of the most important technical issues confronting GIS practitioners. The old data adage about "garbage in, garbage out" is particularly apt when discussing a GIS. Ironically enough, the reason for this stems from one a GIS's greatest strengths; namely, the ability to integrate large and disparate geospatial datasets. Data inaccuracies and errors are inherent features of all geospatial datasets and these problems are inherited by a GIS each time a new data set is imported. Left unchecked, these imported errors may combine with those already in the GIS to form propagation and/or cascading errors,[66] resulting, ultimately, in suspect analyses.

In the context of a GIS, it is common practice to differentiate between several types of accuracy, including *spatial* (both vertical and horizontal), *attribute*, *conceptual* and *logical*. In simple terms, *spatial accuracy* (a.k.a., *positional accuracy*) is a measure of the degree to which the

[64] *Merriam-Webster's New Collegiate Dictionary* (1981), s.v. "Pertinent."
[65] Foskett, D. J. (1970). "Classification and Indexing in the Social Sciences," *ASLIB Proceedings* 22: 91. It is noted that the definition as here cited actually is ascribed by Foskett to the concept of *relevance*, which, within the context of information retrieval, Foskett feels is important to distinguish from pertinence. For Foskett, *relevant* information refers to that which is *objective* and publicly agreed upon, whereas *pertinent* information has a more *subjective* quality that is linked to, and contingent upon, a researcher's private knowledge.
[66] Propagation error occurs when one error is responsible for, or otherwise leads to other errors. A common example of this type of error is where a mis-digitized map registration point from one coverage is used to register subsequent coverages. Cascading errors are those that are allowed to propagate and mutate unchecked from one layer to the next during the selective (re)combination of erroneous, imprecise or inaccurate data into new layers or coverages. Cascading errors are particularly pernicious because of their unpredictable additive or multiplicative variability, depending on how different GIS datasets are combined. Source: K. E. Foote and D. J. Huebner (2000), "Error, Accuracy, and Precision," *The Geographer's Craft Project*, Department of Geography, The University of Colorado at Boulder. Available at http://www.colorado.edu/geography/gcraft/notes/error/error_f.html.

locational data for a feature in the GIS correspond to the actual location of that feature on the ground. Spatial accuracy may be expressed in either *absolute* or *relative* terms, depending on whether reference is made with respect to a coordinate system or to the position of other features in the GIS, respectively. *Attribute accuracy* refers to how closely the descriptions of the features in the GIS match the reality of those features on the ground. *Conceptual accuracy* relates to an assessment of real-world phenomena relative to both the degree to which they are abstracted and the nature of their classification when incorporated into a GIS. It is the GIS practitioner who determines what level or amount of information about a phenomenon to incorporate and how best to classify it into appropriate categories. The misclassification of information or the use of inappropriate categories results in reduced conceptual accuracy. Finally, *logical accuracy* refers to the degree to which GIS data are used appropriately, relative to the goals of the research and the questions being addressed.[67] Data that are used or compared inappropriately will not yield useful, or logical, results. Together, conceptual and logical accuracy speak to the issue of *pertinence*.

In addition, GIS practitioners often assess data in relation to its *logical consistency*, *completeness* and *lineage*. *Logical consistency* concerns the internal consistency or "faithfulness" of the GIS data structure, especially its topological constructs and its spatial ↔ attribute links. Typical examples of logical inconsistencies include incorrect line intersections, line gaps (resulting in open polygons) and duplicate lines, boundaries or features. Collectively, such inconsistencies often are referred to as spatial or topological errors. *Completeness* refers to the degree to which all possible and relevant data objects are included in the database such that the GIS satisfactorily describes the area(s) of interest at the specified time(s). Incomplete datasets contain unclassified areas or "holes," resulting from inadequate data collection and/or compilation and processing procedures that caused the inadvertent elimination of data. *Lineage* refers to all the historical and compilation aspects of a dataset, including, for example, its content, its geographic coverage, its source and the manner in which it was created, captured, measured and/or processed.[68] In as much as such information is documented and linked to the dataset it can be said to be synonymous with metadata. Collectively, positional accuracy, attribute accuracy, logical consistency, completeness and lineage comprise what is referred to in ANSI standard, NCITS 320-1998—*Information Technology-Spatial Data Transfer,* as *data quality*. In short, *data quality* may be defined as the "fitness" of a specific dataset for use in a particular GIS application. As a consequence of the exact nature of its data quality components, a dataset deemed fit for one application may be entirely inappropriate for another. The ease with which disparate GIS datasets can be imported, used, manipulated and recombined at any scale, underscores the importance of generating explicit documentation of the quality of each dataset, via a *data quality report*. This is especially crucial if other researchers will use the dataset(s) in the future. A data quality report, which forms part of the metadata associated with each dataset, documents the key information that is necessary for making informed decisions about the fitness of each dataset with respect to a particular GIS application. Such reports typically will include a description of the datasets and their original source(s), how they were created, collected and processed, and how, or if, they were checked for their accuracy, logical consistency and completeness.

[67] Information in this paragraph was drawn largely from the following two sources: D. J. Buckley (1999), *The GIS Primer: An Introduction to Geographic Information Systems*, (Emeryville, CA: Pacific Meridian Resources). Available at http://www.innovativegis.com/basis/primer/primer.html; and Foote and Huebner, *supra* note 66.

[68] Foote and Huebner, *supra* note 66.

5.6.1.2 Survey Results

As the variability of responses to Question F1 indicate, accuracy, as used by archaeologists in the context of GIS research, is a multi-faceted concept that is distinguished by a number of distinct elements or characteristics (Figure 58), each of which is, in turn, influenced by one or more data creation and manipulation factors or procedures over which the end user may or may not have any direct control (Figure 59). As shown in Figure 58, of the seven accuracy characteristics identified by the survey participants, "closeness to ground truth" (or some derivative thereof) was by far the most common, cited by nearly one half (44.9%) of the participants. The essence of this characteristic was most succinctly represented by one participant who described it as "how closely the GIS data represents the real-world." It is interesting to note that of the seven characteristics identified, this is the one that is the most congruous with the "how close a measurement or estimate is to the correct value" sense in which accuracy often is defined within the broader scientific community.

Because the functional distinction between many of the seven characteristics is subtle, at best, it is possible to conflate them into the three general concepts of accuracy, precision and pertinence, as shown in Figure 60. In this figure, those characteristics that most closely align with the scientific definitions of accuracy, precision and pertinence given earlier have been lumped together under their respective labels. Thus, the labels "accuracy" and "precision" each include three of the original seven characteristics, with accuracy consisting of *closeness to ground truth*, *degree of error* and *completeness*, and precision consisting of *precision, reliability* and *consistency*. What is immediately apparent when looking at these conflated data is that nearly one half (48.0%) of the participants define the term *accuracy* exclusively in a manner that is consistent with the way it is defined in the empirical sciences, while roughly one-quarter (24.0%) define accuracy exclusively in a manner that is more consistent with the term *precision*. Far fewer (6.7%) exclusively equate accuracy with *pertinence/relevance*. Finally, it is noted that approximately one of every six participants (17.3%) defines accuracy using a combination of two of more of these three concepts (i.e., accuracy, precision and/or pertinence).

What these results demonstrate is that, while there is some underlying uniformity in way the participants approach the issue of accuracy (as indicated, for example, by the fact that 48.0% define accuracy from roughly the same analytical perspective), there nevertheless is considerable overall variability stemming, in part, from the incorrect and inconsistent use of terms such as accuracy and precision.

Another significant finding is that only one third of the participants (33.3%) mentioned the integral role that the documentation of datasets plays in the assessment of their accuracy (see Figure 59). Virtually all GIS databases integrate disparate datasets, often of different scales and quality. Moreover, as Gourad emphasizes, "Accuracy issues are present in virtually every GIS data source."[69] These two facts alone should provide ample warning of the ease with which serious propagation and cascading errors can be generated, especially in the absence of adequate data quality documentation. Without adequate documentation about the quality of the data used in a GIS, it is difficult, costly and, at times, impossible, for others to verify or assess the accuracy of the data and, by extension, the results of any analyses based on those data. This highlights what

[69] Gourad, *supra* note 22, p. 20.

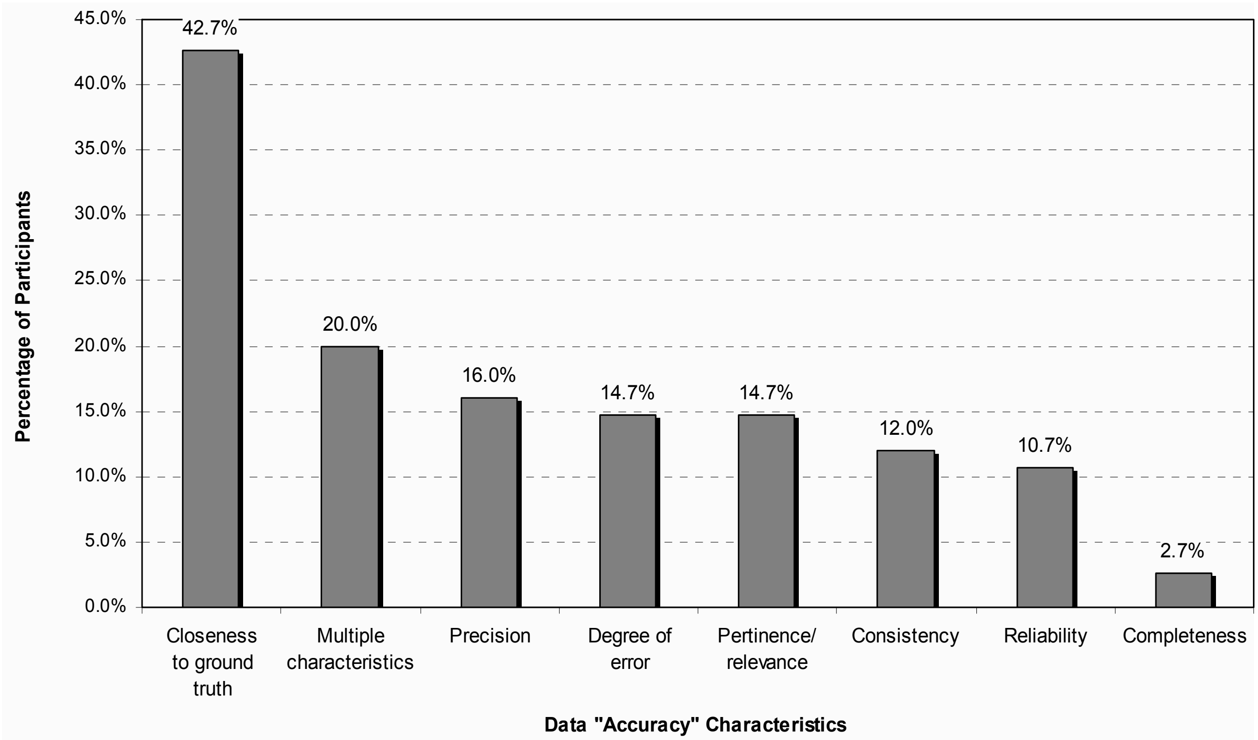


Figure 58. Characteristics of GIS data that distinguish or define their accuracy, as identified by survey participants (n=75).

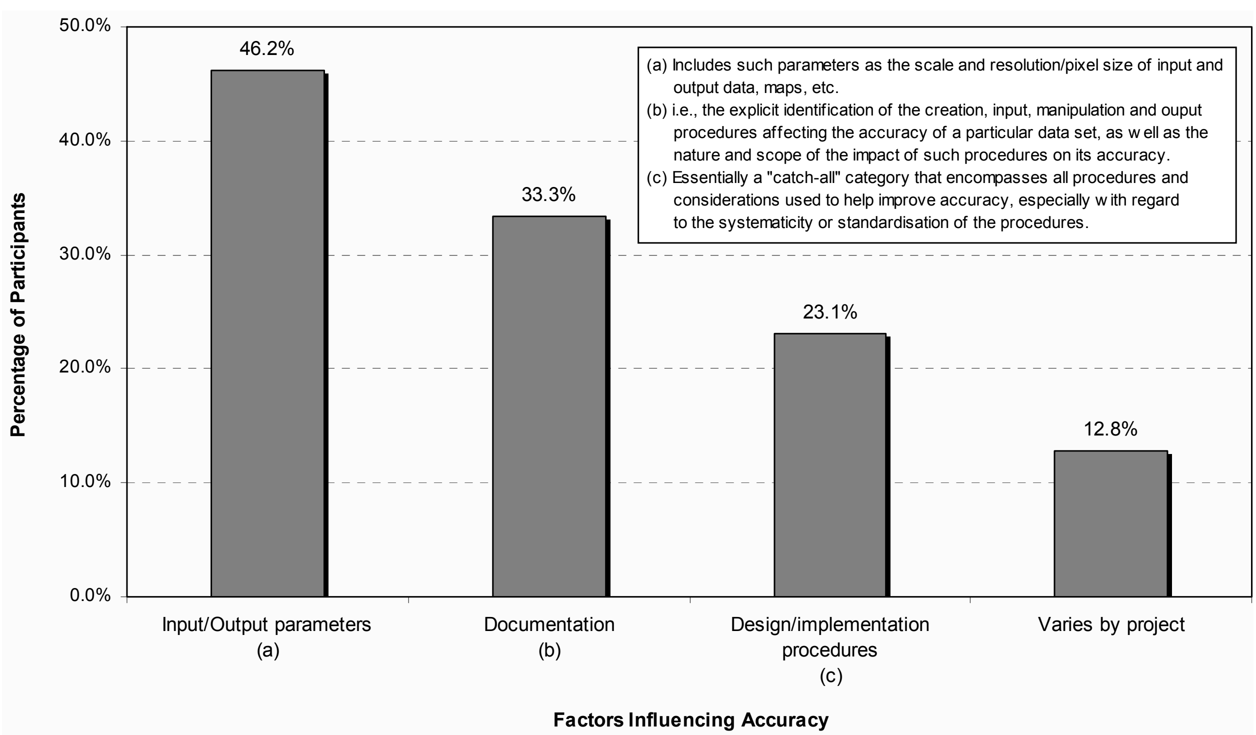


Figure 59. General factors and procedures that influence the accuracy of GIS data, as identified by survey participants (n=39).

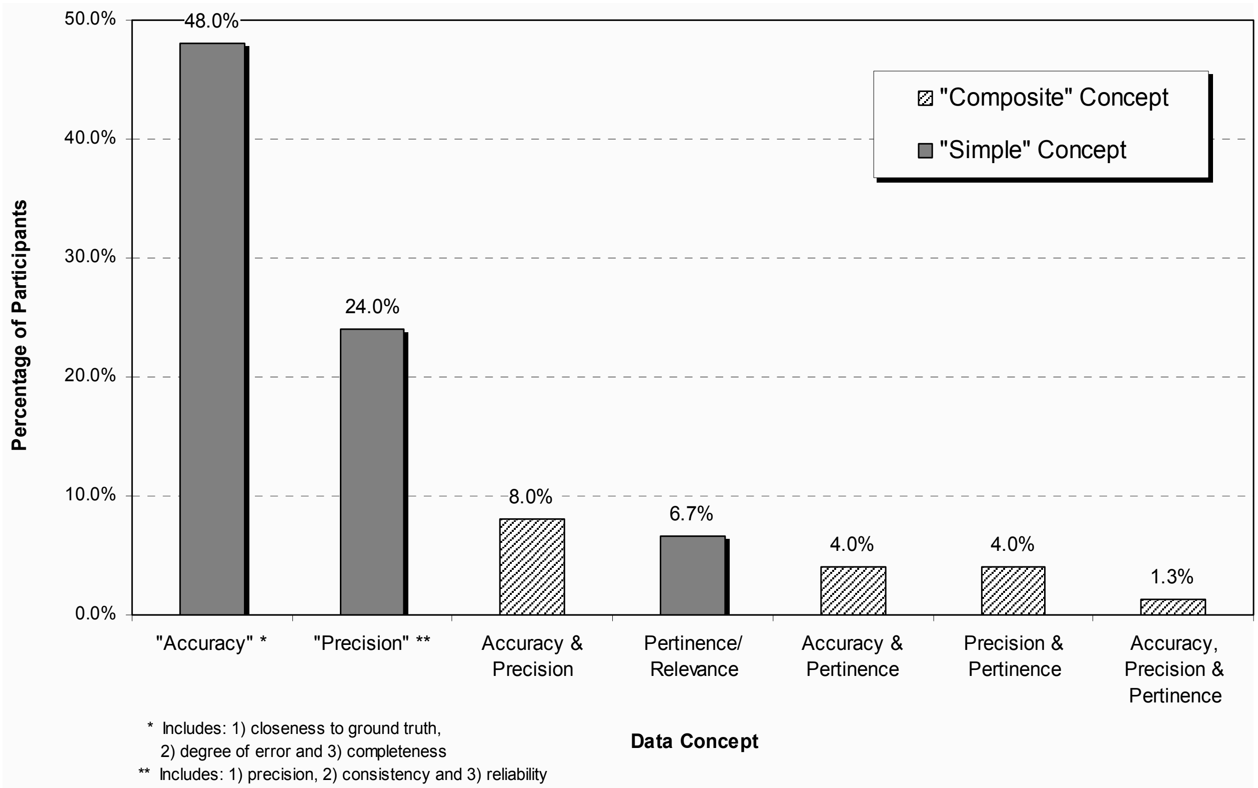


Figure 60. Comparison of the degree to which characteristics related to scientific concepts of accuracy, precision and pertinence were cited by survey participants in their definitions of "accuracy" (n=75).

Buckley identifies as one of the major problems currently confronting GIS researchers; namely, the "aura of accuracy" associated with digital geospatial data.[70] As Buckley points out, although many analogue map sources include a map reliability or confidence rating in the map legend, such information rarely is encoded in the digital data during digitization. Thus, in the absence of data quality documentation, the original data quality information will be unavailable to subsequent users of the digital data. Once in digital form, the data can be, and typically are, represented with a misleadingly high precision by the GIS software;[71] a precision well beyond the true precision of the original analogue data. The concern here is that high precision will, in the minds of unwary GIS practitioners and end-users, erroneously be equated with high accuracy.

Although generally better informed about the potential limitations of GIS data, GIS practitioners (i.e., those involved in GIS planning, development and/or maintenance, as opposed to end-users) are not immune to such data pitfalls.[72] In fact, Gourad found that 14% of his survey participants either were not even familiar with the issue of data accuracy and its impact on GIS research, or

[70] Buckley, *supra* note 67, citation 1.
[71] According to Gourad, for example, "Most GIS systems use at least 8 decimal digits to allocate coordinates. [Eight] decimal digits applied to the entire planet could map spatial entities up to the nearest 10-cm, which is much more precise than any available data set" (Gourad, *supra* note 22, p. 23).
[72] For a concise summary of many of the data, analytical, theoretical and interpretive "pitfalls" to which any GIS research may be exposed, the reader is directed to section 2.5, GIS Pitfalls, in Gourad, *supra* note 22, pp. 18–28.

else were aware of the issue but nevertheless had never considered its impact on their research.[73] Even more telling is the fact that nearly one half (48%) of Gourad's survey participants either were not familiar with the procedures available for testing the accuracy of spatial data (e.g., logical consistency checks, sensitivity analyses, statistical modeling, etc.), or else were aware of the procedures but had never bothered to apply them to the GIS data used in their research.[74] Gourad's findings seem to suggest the existence of a fairly high level of ambivalence toward, and in some cases total unawareness of, certain data accuracy issues among archaeological GIS practitioners. In the six years since Gourad's survey, however, there has been a marked increase in awareness among GIS practitioners, including archaeologists, about the importance of addressing data quality issues and the detrimental impact that ignoring them can have on one's research. Indeed, in the words of Foote and Huebner, "It is now generally recognized that error, inaccuracy, and imprecision can 'make or break' many types of GIS project[s]."[75]

5.6.2 Accuracy and Auditing of Data Files

Approximately the same percentage of participants said that their GIS data files (whether created in-house or imported) always or usually are formally or systematically audited for accuracy, as said that they never are (33.6% vs. 35.2%, respectively) (Figure 61). Nearly one quarter said that the files occasionally are audited, while the remaining 6.6% were not sure how often, or if, their data files were subject to audits. Available time and funding, as well as the nature of the project/task, were three factors cited by the participants as influencing whether, how often, or how thoroughly their data files are audited for accuracy. For example, one participant noted that, "For all 'real world' major projects, ALL data undergo QA/QC by a separate individual at each step of the process. This is all documented on paper and the logs are kept. For small academic or student projects this is sometimes not done." In the words of another, "It depends on the task. If say, only one area is required to work on but the general map shows a whole region, then that region (whole) is not checked for errors." One participant emphasized that their approach to data quality control was more preventative than diagnostic, in which more effort was placed on the use of sample audits during user training than on more full-scale, post-training data audits.

In his 1998 survey, Gourad approached the issue of data accuracy from a slightly different angle. Instead of asking how often his participants' data were audited, Gourad asked his participants how familiar they were with the issue of data accuracy and what impact this knowledge had on their projects. In particular, Gourad asked a series of ten questions related in one way or another to the issue of data accuracy, the results of which he summarizes in section 4.5, Awareness of Pitfalls. In particular, Gourad asked his survey participants to indicate their familiarity with the effects the following issues had on their GIS analyses:

1. How the original data was collected
2. Whether the metadata meets USGS standards
3. Applying accuracy testing methods on data
4. Issues of generalizations in map making

[73] Ibid, pp. 50–51.
[74] Ibid, pp. 52–53.
[75] Foote and Huebner, *supra* note 66.

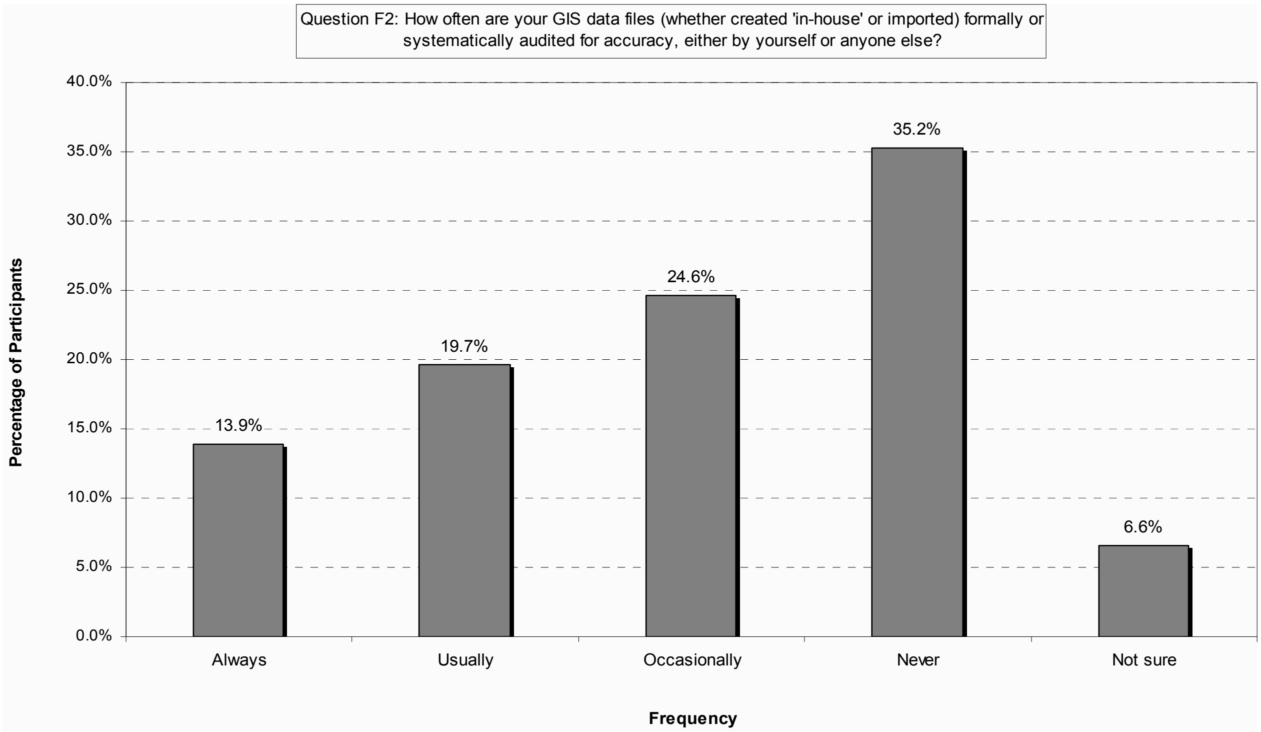


Figure 61. Frequency with which survey participants systematically audit their GIS data files for accuracy (n=122).

5. Computer errors
6. How distances are measured in vector and raster environments
7. Overlays in vector and raster environments
8. Conversions between raster and vector formats
9. Different GIS packages use different algorithms to interpolate
10. Issues of map scales

Although, again, direct comparison between the two surveys is not possible because of differences in the nature of the questions asked, there does appear to be a general correspondence between the two surveys with respect to the participants' attitudes toward the issue of data accuracy, with perhaps a slight shift toward increased implementation of accuracy auditing among the participants in the current survey. With respect to the issue of the participants' attitudes toward the issue of data accuracy in general, the results from these ten questions, when considered collectively, compare fairly closely with the results of the current survey. For example, while 96% of Gourad's participants indicated they were at least aware of the problems that the issue of data accuracy could present, fewer than half (49%) said that they have in fact taken steps to account for accuracy issues.[76] This compares favourably to the 58.2% of participants in the current survey who said that they formally audit their project data at least occasionally, if not more often (see Figure 61). There appears to be even greater congruence between the attitudes and actions of participants in the two surveys over the specific issue of

[76] Gourad, *supra* note 22, p. 50–51.

testing one's GIS data for accuracy (issue no. 3 in the list above), with 52% of Gourad's survey participants stating that they have either considered the importance of accuracy tests, or else modified their research to account for them.[77] Likewise, while 48% of Gourad's participants either were not familiar with the data testing issue, or else chose to ignore its impact on their projects, a slightly lower percentage (35.2%) of participants in the current survey said that they never formally audit their project data.

5.6.3 Identification of GIS Project Creator

When asked how often measures are taken to ensure that, when sharing their GIS projects with other researchers or the general public, the participants are identified as the creators of their GIS projects (through the use of logos, researcher/institution names, etc.), 17.9% said that the question does not apply to them since they never share their projects. Of those who do, a full two thirds (66.4%) said that they always or usually take measures to ensure their identity as the project's creator, while just over one quarter (26.7%) said that they only occasionally, if ever, do (Figure 62). Another 6.9% were not sure if, or how often, such measures were taken.

Concerns about the protection of copyright were cited by a number of the participants in their comments to Question F3. One participant noted that their projects contain a designated copyright

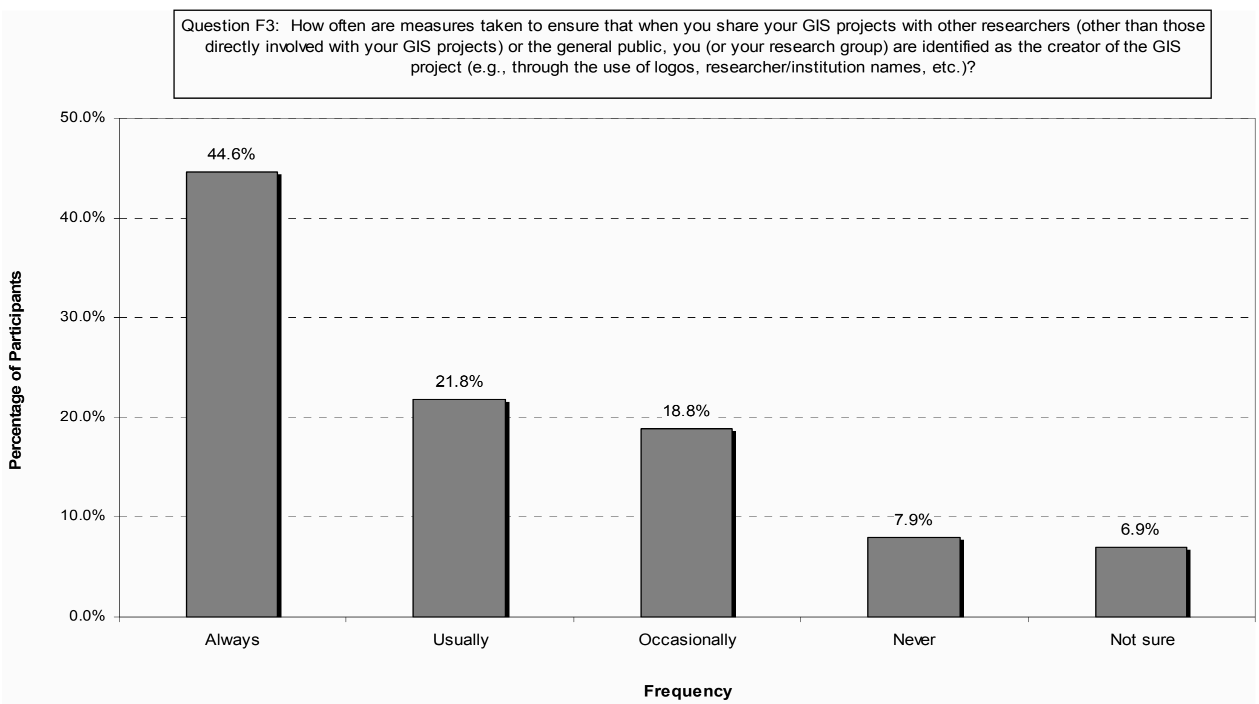


Figure 62. Frequency with which survey participants take measures to ensure that the creator of the GIS project is identified when sharing the project with others ("not applicable" responses excluded) (n=101).

[77] Ibid., p. 52–53.

field. Another noted that, while contact details are always included as part of the metadata supplied to external researchers, the metadata are included in a separate "readme.txt" file and thus easily prone to separation from the actual datasets to which they pertain. In contrast, another participant felt the issue of creator identification was not important, but noted that such information, was, nevertheless, embedded in the project's dataset metadata. Finally, a participant working for a commercial unit within a government agency emphasized that, as such, he/she is required to "acknowledge the source of [the] data/GIS project when…provid[ing] it to others or members of the public to view or use." This likely is the norm for most, if not all, GIS projects created under similar circumstances.

5.6.4 Project Access and Security

When asked which measures are, or have been, used to restrict access to, or otherwise protect their GIS projects (whether still active or not) and their projects' underlying data from unauthorized access or modification, 13.1% said that no particular measures are implemented because they do not restrict access to their projects (Figure 63). Another 2.8% said that they never allow access to their projects (e.g., project files on their home computer to which only they have access) and so see no reason to implement access security measures. Of those who do implement security measures to restrict and control access, some form of privileged system or file access (e.g., via authentication systems using passwords, access control lists, etc.) is the most commonly employed measure (82.2%), followed by privileged facilities access or physical site security measures (75.5%), file ownership and Digital Rights Management measures (e.g., protected vs. unprotected PDF files, read-only vs. read-write access files, etc.) (62.9%), user and record logging software (e.g., automated tracking fields) (40.4%) and, finally, file/data encryption measures (20.4%).

One participant noted that the relatively limited availability of the GIS software used for his/her projects served as a de facto form of security that helped control access to the data. Another emphasized that the nature and degree of the security measures employed varied by project, with "virtually all 'real' projects that are done under contract [having] high security."

5.6.4.1 Privileged System/File Access

Of those participants who control access to their GIS projects, 59.8% said that they always or usually use some type of software-controlled measures to restrict system-wide or file-level access to privileged users (Figure 64). Such control measures may include, for example, authentication systems using passwords, user access control lists, etc. Another 36.5% said that they only occasionally, if ever, use such access control measures, while the remaining 3.3% said that such measures were used, but that they were not sure how often.

5.6.4.2 Privileged Facilities Access / Physical Site Security

The use of some type of physical facilities access measure(s) (e.g., restricted building or computer access) closely mirrored the use of the aforementioned software-controlled security measures, with 59.4% of participants indicating that they always or usually use such measures and 36.8% indicating that they do so only occasionally, if ever (see Figure 64). Again, 3.3% said that such measures were used, but that they were not sure how often.

5.6.4.3 User and Record Logging Software

Substantially fewer participants (23.1%) said that they always or usually employ some type of software-controlled measure to track or log user access and record modification activities, while most (69.2%) said that they only occasionally, if ever, implement such controls (see Figure 64). The remaining 6.7% said that such measures were used, but that they were not sure how often.

5.6.4.4 File Ownership and Digital Rights Management

Just over one third (35.2%) of the participants indicated that they always or usually make use of file ownership and Digital Rights Management measures to protect their GIS project files from unauthorized access or modification (see Figure 64). Such measures may include using protected PDF files, making files read-only, etc. A little more than half (56.1%) said that they only occasionally, if ever, use such measures, while the remaining 7.5% said that such measures were used, but that they were not sure how often.

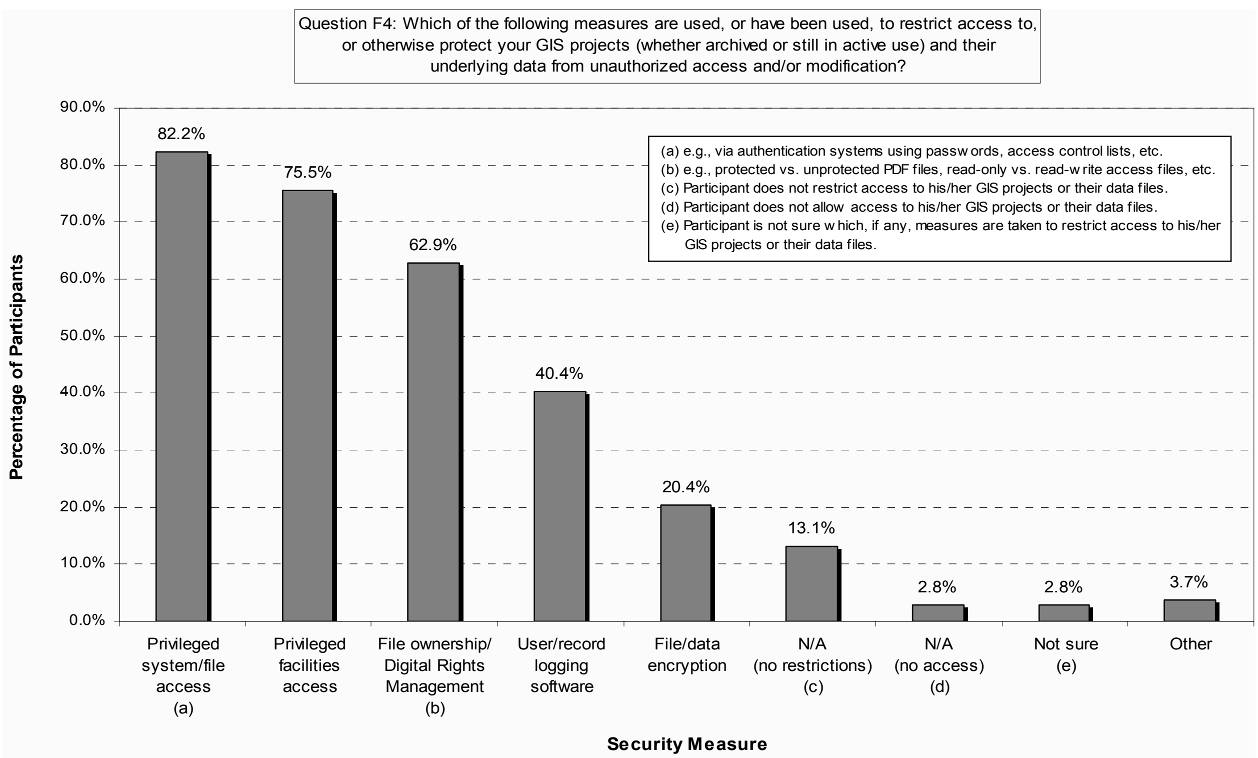


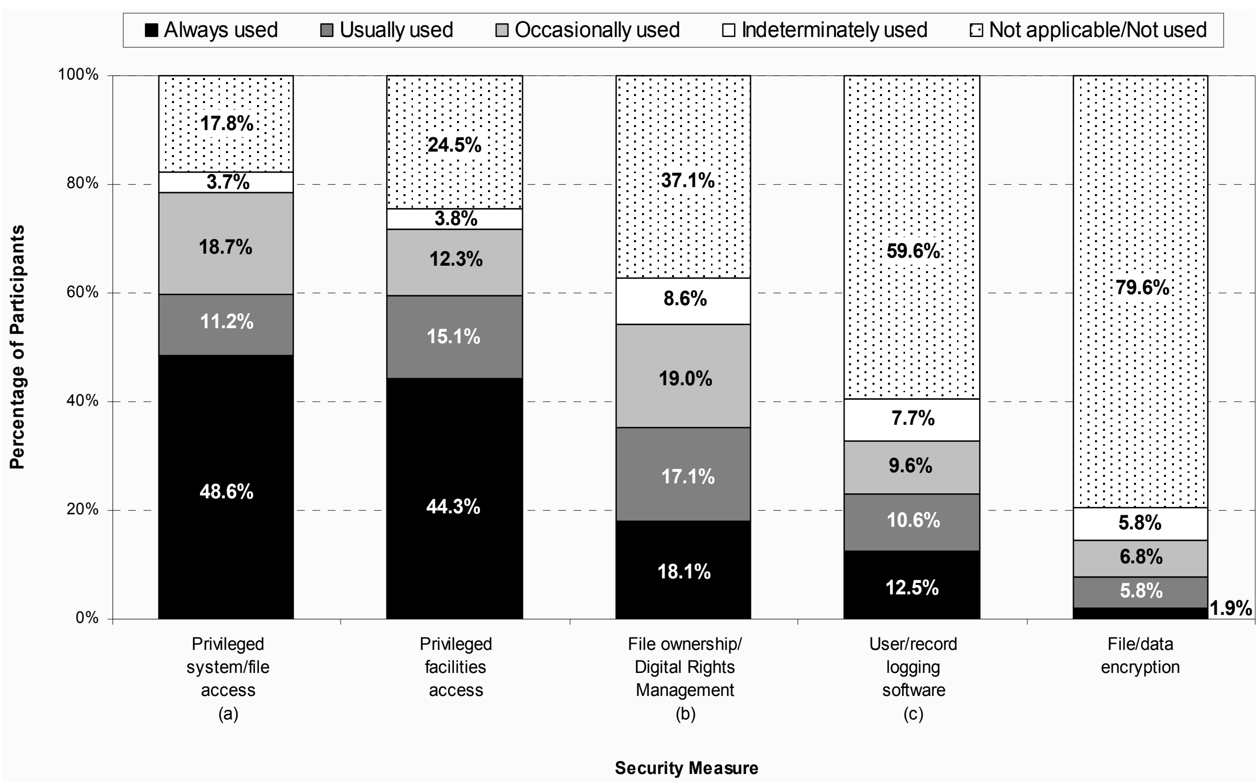
Figure 63. Security measures used by survey participants to restrict unauthorized access to, or modification of, their GIS project files (n=107).

5.6.4.5 File/Data Encryption

The use of file or data encryption was, by far, the least common security approach used by the participants, with only 7.7% indicating that encryption was something that they always or usually used to protect and control access to their GIS data and files (see Figure 64). An overwhelming majority (86.4%) said that they only occasionally, if ever, use encryption, while the remaining 5.1% said that encryption was used, but that they were not sure how often.

5.6.5 Project Data Integrity

As shown in Figure 65, overall confidence in the continuing integrity of their GIS data generally is high, with 54.1% of the participants expressing strong or absolute confidence that their data (associated with both active and inactive projects) have never been tampered with or corrupted over time in such a way that would reduce the value of those for future use or make it impossible to reproduce or recreate the participant's results. Another 27.9% said that they were fairly confident about the continuing integrity of their GIS data. In fact, only 11.5% expressed little or no confidence. The remaining 6.6% were not able to say one way or the other.



(a) e.g., via authentication systems using passwords, access control lists, etc.
(b) Digital Rights Management (e.g., protected vs. unprotected PDF files, read-only vs. read-write access files)
(c) e.g., automated tracking fields

Figure 64. Frequency with which survey participants use various file access/modification security strategies.

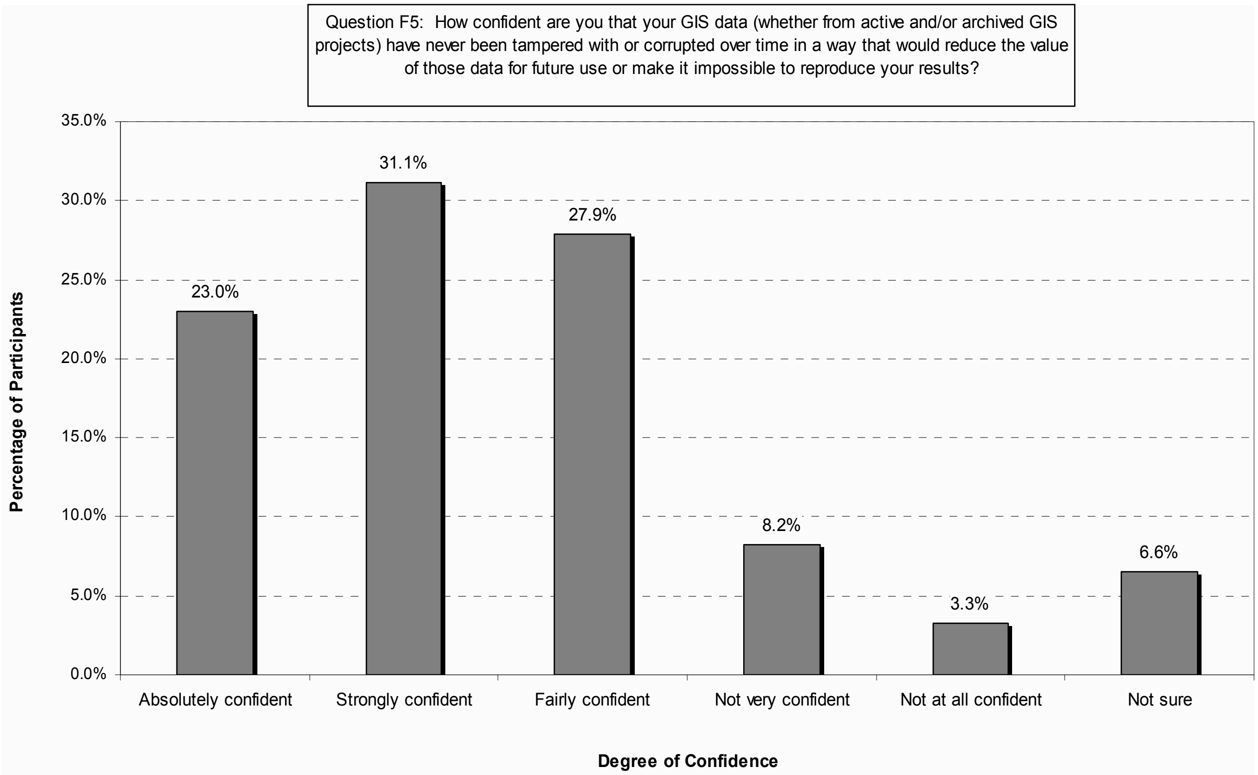


Figure 65. Degree of confidence of survey participants in the long-term integrity of their project data (n=122).

Several participants addressed the distinction between deliberate tampering and incidental corruption, suggesting that it was easier to characterize the potential level of impact of the former than it was the latter, due in no small part to the relative newness of the GIS profession. In the words of one participant, "I am confident that the data has not been deliberately tampered with, but have no indication of whether it has been corrupted accidentally over time, partly because none of our GIS projects are very old." Another participant drew attention to a related data integrity issue: "un-noticed modification." Neither an instance of deliberate tampering, nor a type of incidental corruption, this phenomenon instead occurs in cases where coverages that were originally used as the basis for an analysis are subsequently modified by another member of the GIS group, but without notifying those responsible for the earlier analysis. In fact, the person modifying the coverage data may not even be aware that an analysis based on the pre-modified data even exists. In any event, this type of data integrity issue can result in the inadvertent invalidation of an earlier analysis.

5.7 Section G: General Comments

The final section in the survey consisted of an optional, free-text that asked participants to comment on anything else about their GIS recordkeeping activities or experiences that they thought might be useful for InterPARES researchers to know. Of the 40 participants who provided a response, most chose to elaborate on their past and present GIS experiences in relation to what they see as some of the most pressing long-term preservation challenges with which they, and the organizations for whom they work, are faced. Many even went a step further and offered suggestions and solutions for meeting these challenges. A number of participants provided feedback on the survey itself, with all but one commenting favourably on its content and structure. The one unfavourable comment came from a participant who felt that the "survey [was] too detailed and ask[ed] the user to think too hard and long." This participant went on to say that, "While I realize that you are trying to gather very important information, if you have problems with survey reliability you'll know why. This is especially true for those of us who aren't in the records business, but who are gathering a lot of GIS data."

The length and complexity of the survey were indeed issues of concern, in part because, as noted earlier, participants were required to complete the survey during a single online session. These concerns likely influenced some or all of the 38% of survey log-ins that resulted in partially-completed surveys. Nevertheless, it is worth emphasizing that, despite these concerns, 62% of all survey log-ins resulted in fully-completed surveys, while another 19% resulted in sufficiently-completed surveys.

The concern raised by the participant cited above about the reliability of survey responses is a valid one. However, as discussed earlier in section 4.5.1 Reliability, statistical correlation tests conducted on data from participants who dropped out of the survey and those who completed the survey suggest that there are no significant differences in the answers between these two groups for the six questions examined. These results provide compelling, albeit somewhat anecdotal, evidence to support a conclusion that participant dropout did not significantly bias the reliability of the survey results.

6.0 CONCLUSIONS AND COMMENTARY

6.1 Overall Assessment

The foregoing analysis offers a number of encouraging and not-so-encouraging insights into the recordkeeping practices of GIS archaeologists. To the extent that the results of this survey can be considered representative of the habits of GIS archaeologists in general, it appears that, on the whole, there is a considerable, and growing, level of awareness among GIS archaeologists of the many technical, administrative, professional and societal issues surrounding the long-term preservation of their archaeological GIS data and research records, particularly when compared to the results of Gourad's 1998 survey.[78] On the other hand, the survey also reveals that many participants currently engage in questionable file creation, management, preservation and/or documentation practices that have the potential to seriously compromise the accuracy, reliability, authenticity and accessibility of the files, especially over the long-term.

6.1.1 File Creation Practices

Although just over half of all participants (52%, see Figure 55) claim to always or usually use systematic procedures when creating their various project files, this figure is somewhat deceiving. In fact, upon closer examination, it is clear that, overall, the file creation practices of many participants are, at best, characterized as *idiosyncratic*. This conclusion is supported by the fact that nearly two-thirds of participants (61%, see Figure 56) admit that their file creation practices are only occasionally, if ever, based on, derived from, or guided by a procedures manual or other similar procedures documentation. Given that the reliability of a record is a function of its completeness and the amount of control exercised on the process of its creation, the reluctance (or inability) of many participants to support their file creation practices through the use of documented file creation procedures has the potential to seriously undermine both the accuracy and the reliability of the records generated.

6.1.2 File Management Practices

Like their file creation practices, the file management practices of many participants are, at best, characterized as idiosyncratic and, in many cases, entirely *ad hoc*. This is particularly evident with respect to file version control and file audit practices, where nearly half of all participants (44%, see Figure 23) admit to using no standardized file version control strategy, while most (60%, see Figure 61) only occasionally, if ever, systematically audit their data files for accuracy.

No doubt, some of these more haphazard approaches to GIS file creation and management are due, in part, to the fact that GIS archaeologists often have little or no formal GIS training. Perhaps nowhere are the consequences of these 'self-taught,' idiosyncratic approaches more apparent and more detrimental to long-term preservation efforts than in the general reluctance (or inability) noted by many of the participants to implement and adhere to formal, systematic and documented file creation and management procedures.

[78] Gourad, *supra* note 22.

6.1.3 Preservation Practices

With respect to preservation practices, one of the more distressing findings, especially from an archivist's perspective, is the fact that fewer than one-third of participants (29%, see Figure 38) are concerned enough about the long-term preservation of their projects to factor such concerns into their project planning, design or implementation. Instead, a significant percentage of participants (40%, see Figure 46) actually cite 'lack of importance' as one of the key factors that prevent them from attempting to save some or all of their GIS projects for the long term. Even more telling is the fact that nearly three-quarters (71%, see Figure 46) of these same participants explicitly identify 'lack of importance' as the *most important* reason why they do not attempt to preserve some or all of their projects (either in-house or in a designated repository).

This ambivalence toward long-term preservation appears to stem, for many participants, from a general lack of exposure to, and hence awareness of, the broader preservation and recordkeeping issues. More troubling, however, is the fact that, for many others, the perspective that there is little or no point in preserving their GIS projects appears to be the result of a more *purpose-driven* ambivalence that has its roots in one or more of the following beliefs. First, there is a widespread assumption within archaeology, and indeed within the broader scientific community, that publication alone constitutes sufficient long-term preservation of one's research. Among other things, this assumption fails to acknowledge the critical role that related, unpublished research records play in establishing, supporting and ensuring the long-term integrity and authenticity of the research as a whole, including any of its published components. Second, there is a fairly common belief that, due to the speed with which GIS tools and research techniques are evolving, one's own research likely will be obsolete within a relatively short period and is, therefore, not worth the effort required to facilitate its long-term preservation. Finally, there is a perception, especially among those participants working as consultants, that concern for long-term preservation is entirely the client's problem. As one participant matter-of-factly noted, "we provide the data to our clients and what they do with it is generally up to them." This perspective, of course, fails to recognize the critical importance of actively addressing and integrating concerns for long-term preservation *throughout all phases* of a project, from initial conception through to completion and preservation. In fact, as is now clearly evident, digital records require, in all but trivial cases, much more concerted and sustained preservation efforts than do traditional analogue records; efforts that, to be successful, must be integrated and co-ordinated throughout all phases of a record's life cycle.

On the other hand, the survey also revealed some encouraging signs with regard to the participants' concerns about long-term preservation. For example, as was clearly conveyed in many of the participants' comments, there is a growing sense of frustration (even desperation, in some cases) over the current lack of suitable long-term preservation repositories available to archaeologists, as well as over the continuing absence of any concerted, profession-wide response to these particular issues and concerns (the efforts of the Archaeology Data Service in the UK, notwithstanding). This sense of frustration is perhaps best exemplified in the words of one participant who lamented that, "This is a nightmare. We had planned to archive with the CSA. [Center for the Study of Architecture and its now abandoned Archaeological Data Archive Project, conceived in fall 2001 and terminated by August 2002 due to "inadequate participation"]. We could not use ADS (Archaeological Data Service) because we could not finance it. At the moment we back everything up, but what will happen in the long term is very unclear and a very major concern." In one sense, this participant's comments and the overriding sense of concern and frustration it conveys is an encouraging sign, as it signals an increasing sensitivity to, and awareness of, long-term preservation issues and, more importantly, the information preservation disaster that awaits should archaeologists continue to allow ambivalence to 'guide' their long-term preservation decisions.

6.1.4 Documentation Practices

Although nearly two-thirds of the participants (62%, see Figure 34) claim that their documentation practices are always or usually consistent from one project to the next, the effectiveness of these practices may, in many cases, be compromised by the fact that just as many participants admit that project documentation often does not occur at or near the time of the event being documented and, in fact, for a full one quarter of participants, project documentation typically only occurs either near the end or after the end of the project (see Figure 32). This practice violates one of the keystones of good, effective recordkeeping. Not only does it significantly increase the chances of compromising the accuracy of the documentation that is eventually generated, it also seriously compromises the evidentiary value of the documentation, particularly its admissibility and weight, in the eyes of the courts; a fact that will likely be of particular concern to those GIS archaeologists working as consultants.

The significance of these findings is particularly illuminating when contrasted against the information summarized earlier (question D13) about which project elements participants identified as most worthy of long-term preservation, together with the above-noted (misguided) belief among many that publication alone constitutes sufficient long-term preservation of one's research. Specifically, by focusing on the preservation of final reports (and related types of documentation, such as published articles), at the expense of more comprehensive on-going project documentation, archaeologists run the risk of not preserving the types of records required to ensure that the key preservation goals identified by survey participants—i.e., project/data re-use, verification, replication and accountability—can in fact be met.

The following quotes, drawn from responses to question C11, offer some valuable insight into the relative level of awareness among participants of the importance of good documentation and its relation to long-term preservation requirements and concerns. When asked to briefly describe what they considered to be sufficient documentation and why, here is what several participants wrote:

> Short journal of changes made, otherwise too much is spent on documentation with no time left for analysis. Most people are interested in results not the tedious change to datasets. GIS is not an excavation, thus it has not to be treated with the same record keeping mania
>
> --
>
> I am increasingly concerned with the issue of documentation, and I realize that I have not been aware enough about it for a long time. I believe that data should be documented up to the point in which anybody could be able to access and understand (de-construct) the process of design, development and storage of any GIS project, either he/she had been somewhat involved in it or not.

> Documentation should provide the full contextual information that make data meaningful.
>
> --
>
> A file management/documentation system is sufficient if and only if others not directly involved in the project can examine the system and get a clear understanding of how the GIS project has evolved and its current status.
>
> --
>
> Sufficient documentation is recording the whole procedure from the initial step of planning a G.I.S. project to the more detailed practical alterations until the information on the last steps of presenting the data. Metadata is crucial to preservation and reuse of archaeological record and G.I.S. projects.
>
> --
>
> Sufficient documentation is more documentation than I'm currently employing…
>
> --
>
> What I would consider sufficient documentation and what I actually do for documentation are two different things…
>
> --
>
> I document all aspects of the project that I need to recreate the project. The documentation for this is very detailed consisting of steps taken as well as the reasoning behind the steps and what I might want to do different next time and why….
>
> --
>
> It should be enough to allow people in the future to be able to understand the processes that went into creating the data, without any input from the original authors, i.e., the history of the data creation process should be transparent (in an ideal world).
>
> --
>
> You've actually made me realize I don't document as well as I should- One should create a regular file name- document its existence and relation to other files and track any changes made to the file and as a result the project.
>
> --
>
> … To be able to distinguish the succeeding versions of the files and how the data was transformed in the evolution of the project in order to understand how results were achieved….
>
> --
>
> Documentation is a necessary evil. It can be extremely time consuming, and thus, most of us procrastinate. However, we know better, and we should be training our students to keep records as the project evolves in a systematic fashion.

As these quotes clearly suggest, there already exists substantial awareness within the profession of the importance of good documentation for the effective long-term preservation of digital records. Especially encouraging is an acknowledgement among many of the participants of the importance of preserving metadata and other forms of *ongoing* project documentation, either in addition to, or in place of, the more summative, end-of-project type of documentation encapsulated in final reports and publications. On the other hand, the benefits of this awareness are tempered somewhat by both a sense of complacency among many of the participants about actually following through with what they acknowledge to be adequate documentation practices, and by a palpable tension over the need to balance project results against documentation practices that, in the absence of standardized policies and procedures, have the potential to morph into what at least one participant characterizes as "record keeping mania." This suggests that the real challenge for archivists (and other preservers) is not so much the need to convince archaeologists of the importance of providing (and preserving) an adequate level of project documentation to achieve their long-term preservation goals, as it is the need to highlight the key preservation issues with respect to documentation practices, and to offer archaeologists more practical guidance as to what types of documentation would be most effective to generate and preserve and at what level(s) of detail.

Ultimately, however, successful long-term preservation of accurate, reliable, authentic and accessible digital archaeological data and research records depends on the willingness of archaeologists to develop and implement more effective documentation practices. To some degree, this may well prove to be a difficult challenge since it will require a fundamental change in the way some in the profession view the relative importance of published research and the underlying documentation that is required to support or substantiate that research. On the other hand, since, for many aspects of archaeological research, meticulous and thorough documentation is already standard practice, both the skills and the mindset needed to make this happen are already in place, so that all that is needed is the will and the necessary resources to support it.

6.2 Recent Developments

In addition to the ongoing efforts of the Archaeological Data Service in the United Kingdom, encouraging developments have recently emerged within the archaeological community in the United States with respect to concerns surrounding the long-term preservation of archaeological data and research records. In fact, in December 2004, a 31-member team composed of 21 archaeologists, one physical anthropologist and nine individuals from the fields of computer science and ecology with expertise in data integration and informatics, held a workshop, funded by a National Science Foundation grant and hosted by the National Center for Ecological Analysis and Synthesis (NCEAS), titled, "Enabling the Study of Long-Term Human and Social Dynamics: A Cyberinfrastructure for Archaeology." The goal of the workshop was to assess "the informatics needs of archaeology and the potential of a cyberinfrastructure for archaeology to benefit the discipline and sciences more generally." The team released its final report in May 2005.[79]

The report opens with the admonition that for archaeology to meet its fundamental disciplinary requirement of providing "long term, scientific understandings of human history, there is a pressing need for an archaeological information infrastructure that will allow us to archive, access, integrate, and mine disparate datasets." Of particular note is the acknowledgement that

[79] Departments of Anthropology and Computer Science and Engineering, Arizona State University, *Final Report of the Workshop, The Promise and Challenge of Archaeological Data Integration, December 5 and 6, 2004, the Upham Hotel, Santa Barbara, California,* organized by Keith W. Kintigh, Arizona State University, Under National Science Foundation Grant SES 0433959, May 10, 2005. Available on Cybertools for Archaeological Data Integration (CIDI) Web site at http://cadi.asu.edu/CyberInfrastructureWorkshopFinal.pdf.

"much important archaeological information also resides in unstructured documents.…[and that] the discipline also needs ways to gain better access to these resources and to extract knowledge from them." This actual relevance of this "unstructured documents" statement to the corpus of *unpublished* research records that comprise the bulk of virtually all research projects is made more explicit in a recent letter of endorsement for the cyberinfrastructure project from Society of American Archaeology (SAA) president, Kenneth Ames, in which he states that:

> Over the last century, archaeologists have amassed large amounts of irreplaceable unpublished or gray literature reports, databases, maps, and images that are critical for understanding human history and that can contribute unique scientific data to many other important scientific topics, such as long-term socioecological stability and change.[80]

The cyberinfrastructure report goes on to stress that, "[r]esearchers have a pressing need for an information infrastructure that will allow them to extract a sensibly integrated and appropriately scaled database of analytically comparable observations from multiple datasets employing different recording protocols." Considering that the goal is to provide a profession-wide information infrastructure that, first and foremost, facilitates inter-institutional/researcher sharing of datasets (and, presumably, all key research records supporting those datasets) regardless of the types of data they contain or the methods by which they were generated, one of the most crucial elements on which the success or failure of such an ambitious undertaking will turn will be the quality and consistency of the *recordkeeping documentation* (aka., recordkeeping metadata) accompanying those datasets.

Fortunately, the authors of the report seem to be well aware of this fact, as is suggested by their insistence, for example, that "[t]he scientific utility of digital data is *absolutely* contingent on the availability of adequate metadata that document the datasets…[and that it] is only through these metadata that we can assess the comparability of observations in different datasets and determine the kinds of operations that can be meaningfully performed on them (e.g., the means of aggregating data), [81] as well as the statement that "[r]ealizing this vision will entail both the development of innovative software tools that permit cross-project integration of data and *a sustained effort to document existing and newly created datasets.*"[82] Despite statements such as these that highlight the important role that recordkeeping documentation will play in the ultimate success (or failure) of the project, the report nevertheless seems to equivocate on just how such documentation is likely to be generated. In fact, one often gets the sense when reading the report, that the authors view recordkeeping documentation as largely, if not wholly, a *technological* issue that can satisfactorily be addressed via the development of some sort of clever (perhaps even fully automated?) metadata capture technology. This appears to the implication, for example, in the report's executive summary where the authors proclaim that, "[n]ew

[80] K. M. Ames (SAA President). Letter to Dr. John Yellen, Director for Archaeology, Archaeometry, and Systematic Collections, National Science Foundation, 16 May 2006. Available at http://cadi.asu.edu/SAALetter.pdf.

[81] *Final Report of the Workshop, supra* note 79, p. 10 (emphasis as in the original). It is interesting to note, also, that in cautioning researcher's that their "*ability to reconstruct the necessary metadata decays rapidly with time, and often catastrophically with the death of the investigator*" (Ibid., emphasis as in the original), the report clearly emphasizes the potentially dire consequences, from a long-term preservation perspective, of allowing the contextual knowledge of one's project to reside entirely, or largely, within the heads of one or two researchers, which is a concern that was also highlighted by a number of the survey participants in their comments.

[82] Ibid., p. 3 (emphasis added).

technologies in information integration will enable archaeologists to…sustain the scientific utility of existing digital data that are critically endangered by media degradation, software obsolescence, and inadequate data documentation (metadata)."[83] A similar sentiment is, in some respects, more explicitly expressed in the SAA's letter of endorsement for the project, where SAA president Kenneth Ames states:

> It is essential that we, as a profession, find ways of assuring our texts, data, and images are preserved in the long run and that they will be available in a form that can meet current and future needs of both focused and synthetic archaeological research. To do this, we must participate in *developing the technologies that allow full documentation*, stable storage, and integrated access to these resources in ways that can maximize their availability and scientific utility.[84]

The fact that the 31 participants of the cyberinfrastructure workshop, a workshop aimed, ultimately, at laying the foundation for creating what is, effectively, a sophisticated, distributed, information and records management system, included a number of computer scientists, yet not one professional current records manager or archivist (both of who's professional training and raison d'être centre around ensuring preservation of, and continued access to, recorded information in all formats), is not an encouraging sign.

As noted earlier, purely technological factors are but one part of what is a very complex long-term preservation equation involving many equally important socio-cultural elements. Indeed, anyone who thinks the solution to the current digital records preservation dilemma is entirely, or even primarily, a technological one, is sadly mistaken. The actual effectiveness of computer technology, as with any other tool or technology, is, in fact, governed and constrained by the weakest link in the system, which invariably proves to be human interaction with the system. Thus, to have any chance of success, a long-term preservation strategy must effectively integrate both technological *and* socio-cultural components. Although many problems involving the technological components of the long-term preservation equation (vis-à-vis software/hardware obsolescence, digital media volatility, etc.) still have yet to be fully overcome, as the results of this survey suggest, satisfactory solutions to problems related to the socio-cultural components of the equation (vis-à-vis recordkeeping practices, attitudes toward the need for long-term preservation, etc.) currently lag even farther behind. Thus, unless or until the archaeological community (and, indeed, society in general) is able to reconcile its long-term digital preservation needs, its digital recordkeeping practices and the types of digital objects it generates as by-products of the various technologies, such as GIS—upon which more and more archaeological research so heavily relies—the long-term preservation of accurate, reliable, authentic and accessible archaeological data and research records is likely, in all but isolated instances, to remain an elusive goal.

[83] Ibid., p. 2 (emphasis added).

[84] Ames, *supra* note 80, paragraph 3 (emphasis added). To be fair, Ames also goes on to acknowledge that the proposed project faces both technological and *sociological* challenges (see paragraph 4).

7.0 BIBLIOGRAPHY OF REFERENCES CITED

Ames, K. M. (SAA President). Letter to Dr. John Yellen, Director for Archaeology, Archaeometry, and Systematic Collections, National Science Foundation, 16 May 2006. http://cadi.asu.edu/SAALetter.pdf

Archaeology Data Service. http://ads.ahds.ac.uk/

Berdie, D. R. (1973), "Questionnaire Length and Response Rate," *Journal of Applied Statistics* 58(2):278–280.

Bogen, K. (1996), "The Effect of Questionnaire Length on Response Rates: A Review of the Literature," in *Proceedings of the Section on Survey Research Methods, 51st Annual Conference of the American Association for Public Opinion Research, May 16-19, 1996, Salt Lake City, Utah* (Alexandria, VA: American Statistical Association), pp. 1020–1025.

Buckley, D. J. (1999), *The GIS Primer: An Introduction to Geographic Information Systems*, (Emeryville, CA: Pacific Meridian Resources). http://www.innovativegis.com/basis/primer/primer.html

Burchell, B. and C. Marsh (1992), "The Effect of Questionnaire Length on Survey Response," *Quality and Quantity* 26:233–244.

Center for the Study of Architecture, *Archaeological Data Archive Project*. http://www.csanet.org/archive/adap/adaplond.html

Clayton, R. L. and G. S. Werking (1998), "Business Surveys of the Future: The World Wide Web as A Data Collection Methodology," in M. P. Couper, R. P. Baker, J. Bethlehem, C. Z. F. Clark, J. Martin, W. L. Nicholls II and J. M. O'Reilly (eds.), *Computer Assisted Survey Information Collection* (New York, NY: John Wiley and Sons), pp. 543–562.

Couper, M. P. (2000), "Web Surveys: A Review of Issues and Approaches," *Public Opinion Quarterly* 64(4):464–494.

Departments of Anthropology and Computer Science and Engineering, Arizona State University (2005), *Final Report of the Workshop*, *The Promise and Challenge of Archaeological Data Integration, December 5 and 6, 2004, the Upham Hotel, Santa Barbara, California*, organized by Keith W. Kintigh, Arizona State University, Under National Science Foundation Grant SES 0433959, May 10, 2005. http://cadi.asu.edu/CyberInfrastructureWorkshopFinal.pdf

Dillman, D. A. (2000), *Tailored Design of Mail and Other Self-Administered Surveys* (New York, NY: Wiley-Interscience).

Dillman, D. A. and D. K. Bowker (2001), "The Web Questionnaire Challenge to Survey Methodologists," In U. D. Reips and M. Bosnjak (eds.), *Dimensions of Internet Science* (Lengerich: Pabst Science Publishers), pp. 159–178.

Dillman D. A., M. D. Sinclair and J. R. Clark (1993), "Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys," *Public Opinion Quarterly* 57:289–304.

Dillman, D. A., R. D. Tortora and D. K. Bowker (1998), "Influence of Plain vs. Fancy Design on Response Rates for Web Surveys," in *The 1998 Proceedings of Section on Survey Research Methods* (Dallas, Texas: American Statistical Association). http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf

Duranti, L. (1995), "Reliability and Authenticity: The Concepts and Their Implications," *Archivaria* 39 (Spring):5–10. http://journals.sfu.ca/archivar/index.php/archivaria/article/download/12063/13035

Duranti, L., T. Eastwood and H. McNeil (1997), "Template 7: How is a Record Created Reliable in the Electronic Environment?" in *The Preservation of the Integrity of Electronic Records* (Vancouver: The University of British Columbia, School of Library, Archival and Information Studies). http://www.interpares.org/UBCProject/tem7.htm

Electronic Cultural Atlas Initiative (ECAI). http://www.ecai.org/

Gourad, K. (1999), "Geographic Information Systems in Archaeology: A Survey" (unpublished Masters thesis, Hunter College of the City University of New York, Department of Anthropology), 75 pp. http://khalid.gourad.com/thesis/

Foote, K. E. and D. J. Huebner (2000), "Error, Accuracy, and Precision," *The Geographer's Craft Project*, Department of Geography, The University of Colorado at Boulder. http://www.colorado.edu/geography/gcraft/notes/error/error_f.html

Helgeson, J. G. and M. L. Ursic (1994), "The Role of Affective and Cognitive Decision-Making Processes during Questionnaire Completion," *Public Opinion Quarterly* 11(5):493–510.

Herzog, A. R. and J. G. Bachman (1981), "Effects of Questionnaire Length on Response Quality," *Public Opinion Quarterly* 45:549–559.

Hogg, A. and J. Miller (2003), "Watch Out for Dropouts," *Quirk's Marketing Research Review* July/August. http://www.quirks.com/articles/a2003/20030706.aspx?searchID=14722804&sort=9

InterPARES 2 Project, *Terminology Database Glossary*. http://www.interpares.org/ip2/ip2_terminology_db.cfm

Jenkinson, H. (1922), *A Manual of Archive Administration* (London: Percy Lund, Humphries & Co.).

Manfreda, K. L., Z. Batagelj and V. Vehovar (2002), "Design of Web Survey Questionnaires: Three Basic Experiments," *Journal of Computer-Mediated Communication* 7(3): no pagination. http://www.ascusc.org/jcmc/vol7/issue3/vehovar.html

Miller, J. (2003), "Online Survey Length: Can Research Findings be Impacted?" *American Marketing Association, 24th Annual Marketing Research Conference, September 14-17, 2003, Los Angeles, CA.*

Miller, P. and I. Johnson (2005), "Archaeologists using GIS," Archaeological Computing Laboratory, University of Sydney. http://www.acl.arts.usyd.edu.au/index.php?option=com_content&task=view&id=96&Itemid=120

Nunnally, J. C. (1978), *Psychometric Theory* (New York: McGraw-Hill).

Richards, J. D. and N. S. Ryan (1985), *Data Processing in Archaeology* (Cambridge: Cambridge University Press).

Spain, S. W. (2004), "Top-10 Web Survey Issues and How to Address Them," *iResearch*. http://www.iresearch.com/pages/library/top_10.cfm

APPENDIX A

Survey of Recordkeeping Practices of Archaeologists: Screen-shots of Web-based Survey

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

Informed Consent Information Letter

Survey of Record-Keeping Practices of Archaeologists
Funded by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC)

You are being invited to participate in a brief survey of archaeologists and their use of digital technology. We are requesting your participation in this survey because of your past or present first-hand experience with archaeologically-related Geographical Information System (GIS) projects, and the electronic records that such projects generate. InterPARES 2 is an international research project investigating problems surrounding the reliability, permanence, and accessibility of digital records.

The goal of this particular survey is to obtain some objective data on archaeologists' use of digital technology. The sciences are one of the areas being studied by InterPARES 2 (the others are artistic and government activities) and we are particularly interested in dynamic and experiential documents; that is, documents which may change over time or take on different forms depending on how they are used. One of the longer-term goals of the InterPARES 2 project is to ensure that the records which are produced will continue to be accessible and reliable in the face of rapid technological change. You are invited to visit the InterPARES 2 website at www.interpares.org for more detailed information.

The emailed invitation to participate was sent to a number of archaeologists whose email addresses appear in publicly available lists of organizations such as ADS-ALL List Server, Maptitude List Server, AAC-L List Server, Mapinfo-L List Server, ESRI Archaeology Discussion Conference and others.

The survey can be completed online in approximately 30 minutes. There are no known risks to participants, nor are there any rewards, other than those which may come from looking at one's professional activities in a different light. Your name will not appear on the survey itself; your responses will therefore be anonymous.

After the survey has been removed from the InterPARES 2 website, the results will be collated and a short summary produced. If you would like to receive a copy of this summary we would be happy to email it to you; simply send an email request to Randy Preston at rpreston@interchange.ubc.ca .

Any further questions about this project should be directed to:

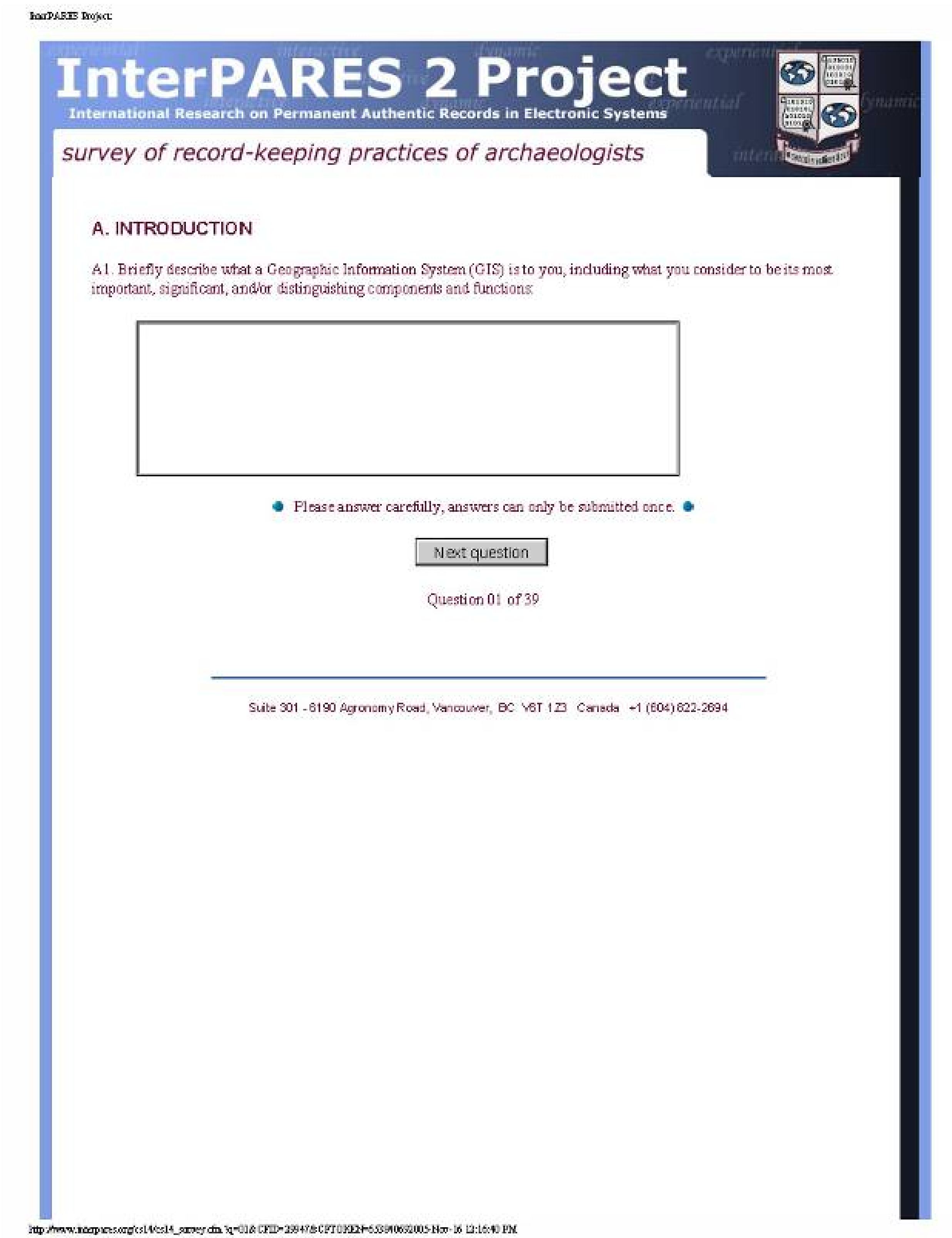
Co-Investigator: Richard Pearce-Moses
Arizona State Library and Archives
rpm@lib.az.us

Research Assistant: Randy Preston
University of British Columbia
rpreston@interchange.ubc.ca

Your participation in this study is entirely voluntary and you may refuse to participate or withdraw from the study at any time without jeopardy to any future participation with the InterPARES Project.

By clicking on the Accept button below, you are accepting the terms of the Informed Consent Information Letter. To decline for any reason, click Decline.

Accept Decline



InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

A. INTRODUCTION

A1. Briefly describe what a Geographic Information System (GIS) is to you, including what you consider to be its most important, significant, and/or distinguishing components and functions:

Please answer carefully, answers can only be submitted once.

Next question

Question 01 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=01&CFID=259478&CFTOKEN=65390069 2005-Nov-16 12:16:40 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

B. GIS EXPERIENCE/BACKGROUND

B1. How long have you been using GIS?

- ○ Less than 1 year
- ○ 1 to 2 years
- ○ 2 to 3 years
- ○ 3 to 4 years
- ○ 4 to 5 years
- ○ 5 to 6 years
- ○ 6 to 7 years
- ○ 7 to 8 years
- ○ 8 to 9 years
- ○ 9 to 10 years
- ○ More than 10 years
- ○ Not sure

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 02 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=02&CFID=259478&CFTOKEN=65390665 2005-Nov-16 12:16:56 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

B. GIS EXPERIENCE/BACKGROUND

B2. How often do you use GIS?

- ○ Daily (I usually use GIS every day)
- ○ Weekly (I usually use GIS one or more times during most weeks, but not every day)
- ○ Monthly (I usually use GIS one or more times during most months, but not every week)
- ○ Yearly (I usually use GIS one or more times during the year, but not every month)
- ○ Rarely (I rarely use GIS one or more times during any given year)

Comments - Use this space if you wish to comment further on this question:

● Please answer carefully, answers can only be submitted once. ●

Next question

Question 03 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=B3&CFID=29947&CFTOKEN=65590065 2005-Nov-16 12:17:01 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

B. GIS EXPERIENCE/BACKGROUND

B3. With which of the following phases of a GIS project(s) are you or have you been involved with? (check all that apply)

Note: Please answer this question in reference to GIS projects, not GIS systems

- ☐ A. Planning
- ☐ B. Development/Design
- ☐ C. Implementation/Data entry
- ☐ D. Data analysis
- ☐ E. Data preservation

Note: If you chose more than one of the above phases, please indicate those phases with which you have the most and the least experience by typing the corresponding letter (A-E) in the appropriate boxes below.

Most experience (A-E): [ ] Least experience (A-E): [ ]

Comments - Use this space if you wish to comment further on this question:

[ ]

Please answer carefully, answers can only be submitted once.

[Next question]

Question 04 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=04&CFID=... 2005-Nov-16 12:17:05 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

B. GIS EXPERIENCE/BACKGROUND

B4. Which of the following GIS system designs do you typically use? (check all that apply)

- ☐ File processing system (i.e., each data set and function stored as a separate file and linked together during analytical processing)
- ☐ Hybrid design system (i.e., attribute data stored in a conventional DBMS with separate software used for geographical data)
- ☐ Extended design system (i.e., both geographical and attribute data stored together)
- ☐ Not sure
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question.

Please answer carefully, answers can only be submitted once.

Next question

Question 05 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=05&CFID=269478&CFTOKEN=65390065 2005-Nov-16 12:17:09 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

B. GIS EXPERIENCE/BACKGROUND

B5a. Please indicate the geographic focus of your research at whatever scale is most appropriate (e.g., North America, United States, American Southwest, Arizona, etc.):

B5b. **Optional Question:** Please indicate your geographic location (i.e., the country in which you are employed, which may or may not be the same as the geographic focus of your research)

Please answer carefully, answers can only be submitted once.

Next question

Question 06 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

B. GIS EXPERIENCE/BACKGROUND

B6. What is your current professional archaeological affiliation? (check all that apply)

- ☐ Cultural Resource Management (public)
- ☐ Cultural Resource Management (private consulting)
- ☐ College/University
- ☐ Private consulting (other than CRM)
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 07 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C1. How often do you use a standardized and/or documented procedure for naming the files you create related to each GIS project?

- ○ Always
- ○ Usually
- ○ Occasionally
- ○ Varies
- ○ Varies from one project to the next (i.e., standardized methods are used for some projects, but not for others)
- ○ Not applicable (I'm not responsible for creating or naming digital GIS files)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 08 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=08&CFID=259478&CFTOKEN=65590069 2005-Nov-16 12:17:21 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C2. When working with your GIS files, which of the following options do you use to maintain version control, especially if different people are working on the same file?(check all that apply)

- ☐ File-naming conventions (e.g., including a version number code in the file name)
- ☐ Standard headers listing creation dates and version numbers
- ☐ File logs
- ☐ No standardized system is used to maintain version control
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 09 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C3. During the course of creating your digital GIS files, how often do you also create related paper documents?

- Always
- Usually
- Occasionally
- Never
- Not applicable (I don't create digital GIS files and/or related paper documents)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 10 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C4. How often do you explicitly document the links (by whatever means) between the paper documents and digital files?

- ○ Always
- ○ Usually
- ○ Occasionally
- ○ Varies depending on the nature and importance of the relationship between the paper documents and the digital files (please describe in the "Comments" box below)
- ○ Never
- ○ Not applicable

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 11 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C5. When documenting your GIS projects, which of the following aspects do you typically document? (check all that apply)

- ☐ The individual data sets comprising the GIS (including creation dates)
- ☐ The sequence of individual commands that were applied to the data to produce the current system
- ☐ The history of changes/modifications made to individual files or data sets (including dates)
- ☐ The reason(s) for changes/modifications
- ☐ The record numbers affected by the changes/modifications
- ☐ Software used, including version changes, etc.
- ☐ Computer operating system(s) used
- ☐ Project details (name, number, etc.)
- ☐ Not applicable (documentation typically is not created for my GIS projects)
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question.

Please answer carefully, answers can only be submitted once.

Next question

Question 12 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C6. When documenting your GIS projects, which of the following approaches do you typically use? (check all that apply)

- ☐ Documentation is entered into a formal paper log book, or journal
- ☐ Documentation is entered into formal paper forms designed specifically for the purpose
- ☐ Documentation is entered more informally into various paper media that may or may not later be consolidated into a single binder or folder
- ☐ Documentation is entered into a formal digital log file, or files (e.g., spreadsheet or word processing files)
- ☐ Documentation is entered into formal digital forms or databases designed specifically for the purpose
- ☐ Documentation is entered more informally into various digital files that may or may not later be consolidated into a single file
- ☐ Documentation is directly encoded within project files
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 13 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C7. When modifying the content of your data or GIS files, which of the following factors influence whether you choose to document these changes? (check all that apply)

- ☐ The number of changes made (e.g. a few changes would not be documented, while a substantial number of changes would be)
- ☐ The nature of the changes made (e.g., the odd spelling error vs. a systematic data recording error)
- ☐ The type of data changed (e.g., attribute text vs. numeric data vs. graphics)
- ☐ Time and/or financial factors
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question.

● Please answer carefully, answers can only be submitted once. ●

Next question

Question 14 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C8. Which of the following best characterizes the overall documentation process typically associated with your GIS projects?

- ○ Documentation is conducted throughout the projects, and usually occurs at the same time as the data and/or files are created, modified and/or manipulated
- ○ Documentation is conducted throughout the projects, but often does not occur at the same time as the data and/or files are created, modified and/or manipulated
- ○ Most or all documentation is conducted toward the end of the projects, or after they are completed
- ○ Not applicable (documentation is not typically created for the GIS projects)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 15 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=15&CFID=259478&CFTOKEN=65590692005-Nov-16 12:17:49 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C9. Which of the following best characterizes your overall documentation procedures (regardless of their actual manifestations) from one GIS project to the next?

- ○ They always are consistent from one project to the next
- ○ They usually are consistent from one project to the next
- ○ They occasionally are consistent from one project to the next
- ○ They never are consistent from one project to the next
- ○ Not applicable (I have not been involved in more than one GIS project)
- ○ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 16 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C10. Which of the following do you consider to be important reasons that actually influence why you document additions and/or modifications to your GIS projects? (check all that apply)

- ☐ It's just good research/programming procedure
- ☐ To provide evidence should my project research be challenged
- ☐ To provide myself or others with an overall picture of the evolution of the GIS project
- ☐ To enable myself or others to reverse the change(s) if necessary (e.g., an error is later discovered)
- ☐ To enable myself or others to repeat the exact sequence of documented steps in the future
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

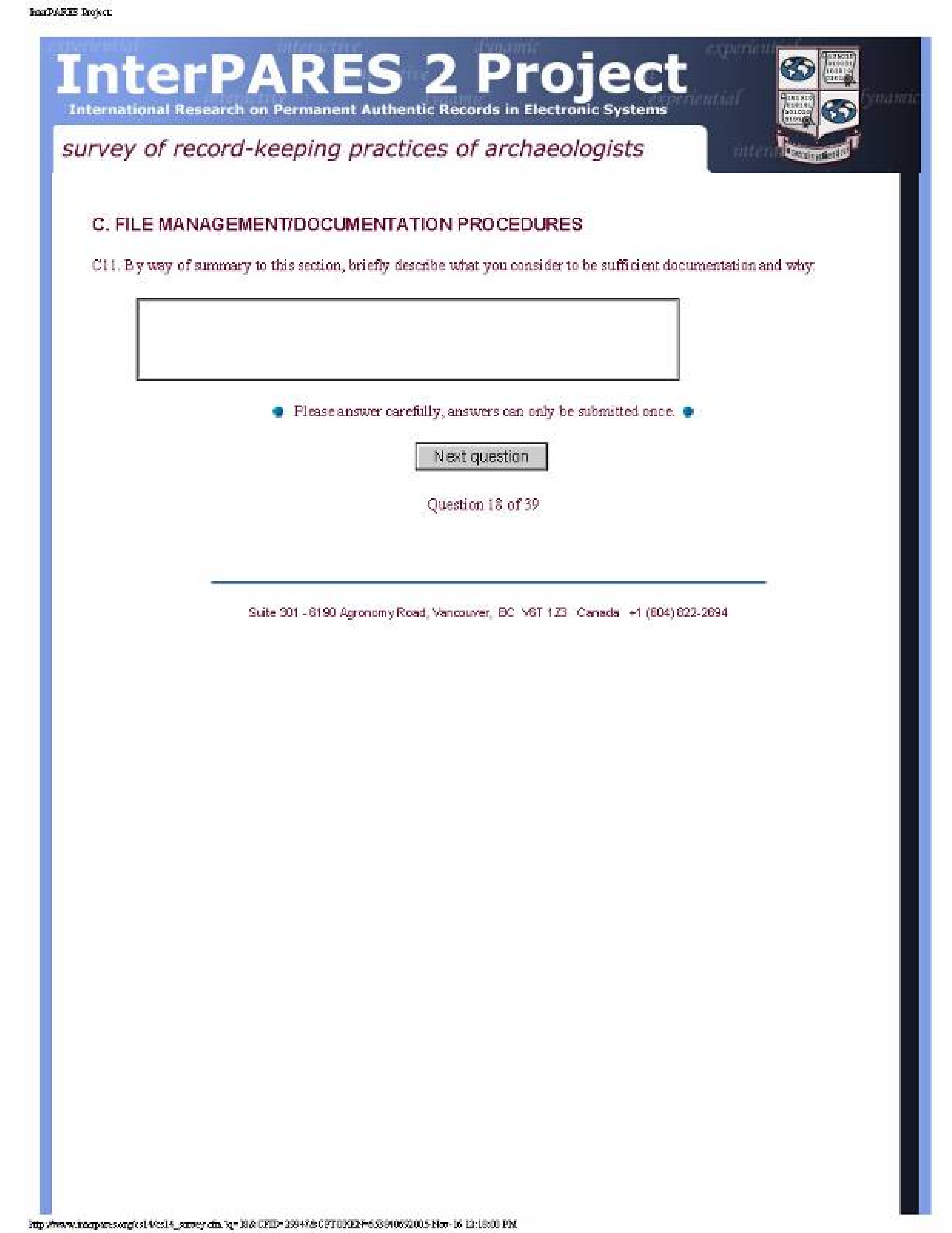
Please answer carefully, answers can only be submitted once.

Next question

Question 17 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=17&CFID=[illegible]&CFTOKEN=[illegible] 2005-Nov-16 12:17:57 PM


InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

C. FILE MANAGEMENT/DOCUMENTATION PROCEDURES

C11. By way of summary to this section, briefly describe what you consider to be sufficient documentation and why:

Please answer carefully, answers can only be submitted once.

Next question

Question 18 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=18&CFID=259478&CFTOKEN=65590692005-Nov-16 12:18:00 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D1. When planning, designing and/or implementing your GIS projects, are you concerned about their future transition into an archival environment (e.g., the Archaeology Data Service in the UK) to help ensure that they can be preserved for the long term, re-located and re-used by other researchers in the future?

- ○ Yes, and it directly influences the planning, design, and/or implementation of my GIS projects
- ○ Yes, but it doesn't directly influence the actual planning, design, and/or implementation of my GIS projects
- ○ No
- ○ Other (please describe in the "Comments" box below)
- ○ Not applicable (I am not involved in the planning, designing and/or implementation phases of GIS projects, only the analysis and/or preservation phases.)

Comments - Use this space if you wish to comment further on this question:

● Please answer carefully, answers can only be submitted once. ●

[Next question]

Question 19 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D2. How often are your completed GIS projects transferred to a designated repository, such as the Archaeology Data Service, a state, museum or university archives, etc., for long-term preservation?

- ○ Always
- ○ Usually
- ○ Occasionally
- ○ Not applicable (My GIS projects are still in active use)
- ○ Never
- ○ Not sure (I'm not involved involved in the long-term preservation aspect of my GIS projects)
- ○ Not applicable (my GIS projects are not, or have not been transferred to a designated repository.)

Comments - Use this space if you wish to comment further on this question:

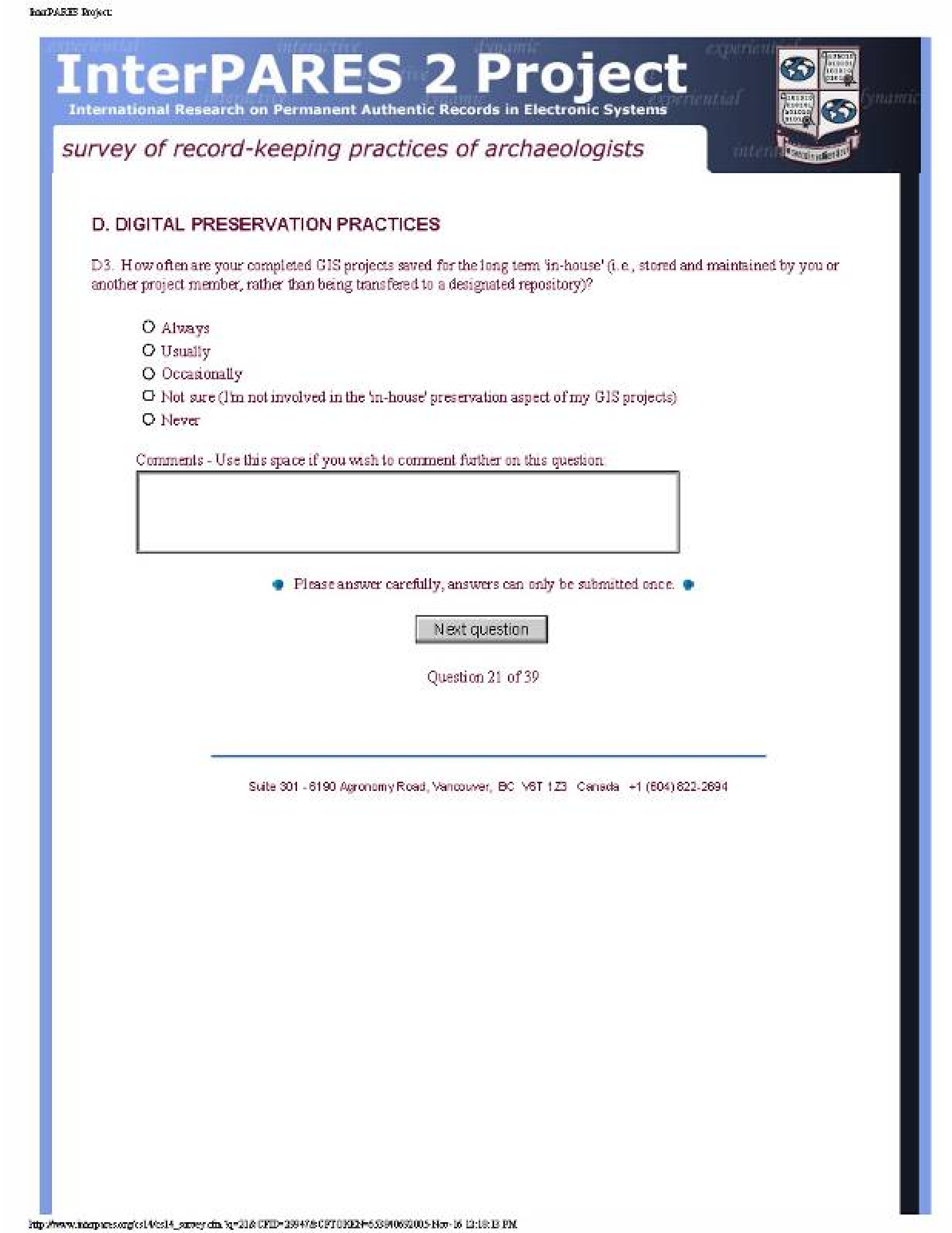
● Please answer carefully, answers can only be submitted once. ●

Next question

Question 20 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D3. How often are your completed GIS projects saved for the long term 'in-house' (i.e., stored and maintained by you or another project member, rather than being transfered to a designated repository)?

- Always
- Usually
- Occasionally
- Not sure (I'm not involved in the 'in-house' preservation aspect of my GIS projects)
- Never

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 21 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=21&CFID=259478&CFTOKEN=65590065 2005-Nov-16 12:18:13 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D4. When saving your GIS projects 'in-house,' which of the following long-term preservation strategies do you use? (check all that apply)

- ☐ Data refreshment (i.e., copying files from one medium to the next as the original medium nears the end of its reliable life span)
- ☐ Data migration (i.e., converting files from one format or structure into another that can be read by current versions of software)
- ☐ Data documentation (e.g., to track and explain data migrations, abbreviations used, file naming conventions, etc.)
- ☐ Electronic Data Management (EDM) (i.e., data management databases that automatically indicate when files need backing up, migration or refreshment)
- ☐ No special long-term preservation strategies are used
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question.

Please answer carefully, answers can only be submitted once.

Next question

Question 22 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=22&CFID=259478&CFTOKEN=65390065 2005-Nov-16 12:18:17 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D5. For those long-term preservation strategies that you use when saving your GIS projects 'in-house,' which of the following would you say best characterizes the overall process?

- ☐ The strategies are implemented systematically (i.e., following an established, standard and/or scheduled procedure) for all files, file types, and electronic media
- ☐ The strategies are implemented systematically, but not always for all files, file types or electronic media
- ☐ The strategies are implemented irregularly or subjectively (i.e., following no established, standard and/or scheduled procedure) for all files, file types and electronic media.
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 23 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=23&CFID=29947&CFTOKEN=65390692005-Nov-16 12:18:21 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D6. When saving your GIS projects (either 'in-house' or in a designated repository), how often are the saved files organized the same way as when you were using them (e.g., using the same folder structure, etc.)?

- Always
- Usually
- Occasionally
- Never
- Not sure (although the GIS projects are saved for the long-term, I'm not involved enough in the process to comment)
- Not applicable (my GIS projects are not, or have not been saved for the long-term.)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 24 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D7. Which of the following considerations prevent you from saving some or all of your GIS projects for the long term (whether 'in-house' or in a designated repository)? (check all that apply)

- ☐ A. Insufficient funding
- ☐ B. Insufficient time
- ☐ C. Insufficient personnel
- ☐ D. Lack of an available/applicable repository
- ☐ E. Not seen as important/necessary
- ☐ F. Other (please describe in the "Comments" box below)
- ☐ G. Not applicable (I save all my GIS projects for the long term.)
- ☐ H. Not applicable (I am not responsible for saving my GIS projects for the long term.)

NOTE: If you chose more than one of the above considerations A-H, please indicate the most important and least important consideration by typing the corresponding letter (A-H) in the appropriate boxes below.

Most important (A-H): ______ Least important (A-H): ______

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 25 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=25&CFID=259478&CFTOKEN=65390065 2005-Nov-16 12:18:29 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D8. When recording metadata to document your GIS project files (or groups of files), which of the following types of information do you routinely include? (check all that apply)

Note: As used here, metadata refer to data describing the context, content and structure of the GIS files and their management through time. Typically, metadata either are embedded within the file being documented/described, or are stored in a separate file/database and linked to the file being documented/described.

- ☐ File format
- ☐ Hardware used to create the file(s)
- ☐ Name of software used to create the file(s)
- ☐ Version of software used to create the file(s)
- ☐ Operating system used to create the file(s)
- ☐ Name of file creator
- ☐ Date of file creation
- ☐ Date of last file update
- ☐ The procedures used to create the file(s) (i.e., how the different components work together to maintain functionality)
- ☐ Not sure (although metadata are recorded, that is not my responsibility)
- ☐ Not applicable (I don't record metadata for my GIS project files)
- ☐ Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question.

Please answer carefully, answers can only be submitted once.

Next question

Question 26 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=26&CFID=29947&CFTOKEN=65390652005-Nov-16 12:18:36 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D9. When recording metadata, how often do you follow/use an established descriptive or other metadata schema or standard (e.g., Dublin Core, USMARC, FGDC, etc.)?

Note: As used here, metadata refer to data describing the context, content and structure of the GIS files and their management through time. Typically, metadata either are embedded within the file being documented/described, or are stored in a separate file/database and linked to the file being documented/described.

- Always
- Usually
- Occasionally
- Never, because I was not aware there were metadata standards
- Never, but I am aware there are metadata standards
- Not sure

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 27 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D10. Which of the following metadata standards do you use, or have you used, when recording metadata for your GIS projects? (check all that apply.)

Note: As used here, metadata refer to data describing the context, content and structure of the GIS files and their management through time. Typically, metadata either are embedded within the file being documented/described, or are stored in a separate file/database and linked to the file being documented/described.

- ☐ CIMI (Consortium for the Computer Interchange of Museum Information)
- ☐ CSDGM (Content Standard for Digital Geospatial Metadata)
- ☐ DCMES (Dublin Core Metadata Element Set, or derivative)
- ☐ DIF (Directory Interchange Format)
- ☐ DIG35 Specification
- ☐ GEO (Geospatial Metadata Profile)
- ☐ IRDS (Information Resource Dictionary System Content Model Standard)
- ☐ ISO 19115 (International Organization for Standardization: Geographic information - Metadata)
- ☐ METS (Metadata Encoding & Transmission Standard)
- ☐ MPEG-7 (Moving Picture Coding Experts Group: Multimedia Metadata Standard)
- ☐ NGDF (National Geospatial Data Framework standard -- e.g., FGDC)
- ☐ NSDI (National Spatial Data Infrastructure)
- ☐ RDF (Resource Description Framework)
- ☐ TEI (Text Encoding Initiative)
- ☐ USMARC (U. S. Machine Readable Cataloging)
- ☐ Federal Geographic Data Committie (FGDC) Content Standards for Digital Geospatial Metadata
- ☐ Other (please list in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 28 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=28&CFID=259478&CFTOKEN=65390652005-Nov-16 12:18:46 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D11. When recording metadata, do you typically embed the metadata within the resource(s) being described, or do you store the metadata in a separate file/database, and provide a link to the described resource(s)?

Note: As used here, metadata refer to data describing the context, content and structure of the GIS files and their management through time.

- I typically embed the metadata
- I typically link the metadata
- I typically use both or either method, depending on the specific resource type(s) (e.g., text vs. graphic) being described
- Not sure
- Other (please describe in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 29 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D12. When recording metadata for your GIS projects, how often are the metadata terms that are used derived from a controlled vocabulary (i.e., a limited set of consistently used and carefully defined terms, usually in the form of a thesaurus)?

NOTE: As used here, metadata refer to data describing the context, content and structure of the GIS files and their management through time. Typically, metadata either are embedded within the file being documented/described, or are stored in a separate file/database and linked to the file being documented/described.

- ○ Always
- ○ Usually
- ○ Occasionally
- ○ Never
- ○ Not sure

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 30 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

D. DIGITAL PRESERVATION PRACTICES

D13. By way of summary to this section, briefly describe which elements, outputs (e.g., original vs. final reports, maps, tabular data), etc. of GIS projects you think should be preserved for future use or reference and why.

Please answer carefully, answers can only be submitted once.

Next question

Question 31 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

E. DATA INPUT/OUTPUT PRACTICES

E1. Overall, how often do you follow a routine or systematic sequence of procedures or steps when generating the various components and outputs (e.g., data files, algorithms, views/maps, statistical regressions, etc.) of your GIS projects?

- Always
- Usually
- Occasionally
- Never
- Varies depending on the component -- I usually or always follow a routine or systematic sequence for certain components, but not for others (please elaborate below)

Comments - Use this space if you wish to comment further on this question.

Please answer carefully, answers can only be submitted once.

Next question

Question 32 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

E. DATA INPUT/OUTPUT PRACTICES

E2. Whether routine/systematic or not, how often are the procedures you follow when generating the various components and outputs (e.g., data files, algorithms, views/maps, statistical regressions, etc.) of your GIS projects based on, derived from, or guided by procedures outlined in a GIS procedures manual?

- Always
- Usually
- Occasionally
- Never, because I do not have access to a GIS procedures manual
- Never, although I do have access to a GIS procedures manual
- Varies depending on the GIS project and/or component (please elaborate in the "Comments" box below)

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 33 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=33&CFID=259478&CFTOKEN=65590692005-Nov-16 12:19:07 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

F. RECORD QUALITY, RELIABILITY & AUTHENTICITY ISSUES

F1. Briefly describe/define what "accuracy" means to you with respect to your GIS projects and/or their individual data files:

Please answer carefully, answers can only be submitted once.

Next question

Question 34 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

F. RECORD QUALITY, RELIABILITY & AUTHENTICITY ISSUES

F2. How often are your GIS data files (whether created 'in-house' or imported) **formally or systematically** audited for accuracy, either by yourself or anyone else?

- Always
- Usually
- Occasionally
- Never
- Not sure

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 35 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=35&CFID=259478&CFTOKEN=65390065 2005-Nov-16 12:19:30 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

F. RECORD QUALITY, RELIABILITY & AUTHENTICITY ISSUES

F3. How often are measures taken to ensure that when you share your GIS projects with other researchers (other than those directly involved with your GIS projects) or the general public, you (or your research group) are identified as the creator of the GIS project (e.g., through the use of logos, researcher/instution names, etc.)?

- Always
- Usually
- Occasionally
- Never
- Not applicable (I don't share my GIS projects with other researchers or the general public)
- Not sure

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 36 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2684

http://www.interpares.org/cs14/cs14_survey.cfm?q=36&CFID=259478&CFTOKEN=65590692005-Nov-16 12:19:28 PM

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

F. RECORD QUALITY, RELIABILITY & AUTHENTICITY ISSUES

F4. Which of the following measures are used, or have been used, to restrict access to, or otherwise protect your GIS projects (whether archived or still in active use) and their underlying data from unauthorized access and/or modification? Please select the applicable measure(s) by indicating how often each measure is used.

A. Privileged system/file access (e.g., via authentication systems using passwords, access control lists, etc.)

- Always used
- Usually used
- Occasionally used
- Used, but not sure how often
- Skip this measure

B. Privileged facilities access / physical site security measures

- Always used
- Usually used
- Occasionally used
- Used, but not sure how often
- Skip this measure

C. User and record logging software (e.g., automated tracking fields)

- Always used
- Usually used
- Occasionally used
- Used, but not sure how often
- Skip this measure

D. File ownership and Digital Rights Management measures (e.g., protected vs. unprotected PDF files, read-only vs. read-write access files, etc.)

- Always used
- Usually used
- Occasionally used
- Used, but not sure how often
- Skip this measure

E. File/data encryption measures

- Always used
- Usually used
- Occasionally used
- Used, but not sure how often

○ Skip this measure

☐ Not applicable (I don't restrict access to my GIS projects or their data files)
☐ Not sure which if any measures are, or have been, used
☐ Other (please describe in the "Comments" box below, and indicate how often the measure is used)

Comments - Use this space if you wish to comment further on this question

● Please answer carefully, answers can only be submitted once. ●

Next question

Question 37 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

F. RECORD QUALITY, RELIABILITY & AUTHENTICITY ISSUES

F5. How confident are you that your GIS data (whether from active and/or archived GIS projects) have never been tampered with or corrupted over time in a way that would reduce the value of those data for future use or make it impossible to reproduce your results?

- Absolutely confident
- Strongly confident
- Fairly confident
- Not very confident
- Not at all confident
- Not sure

Comments - Use this space if you wish to comment further on this question:

Please answer carefully, answers can only be submitted once.

Next question

Question 38 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=38&CFID=259478&CFTOKEN=65390065 2005-Nov-16 12:19:35 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

G. GENERAL COMMENTS (OPTIONAL)

G1. Please add anything else that you think might be useful for us to know about your GIS record-keeping activities or experiences:

Please answer carefully, answers can only be submitted once.

Next question

Question 39 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=39&CFID=259478&CFTOKEN=65590605 2005-Nov-16 12:19:39 PM

InterPARES Project

InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

survey of record-keeping practices of archaeologists

Thank you.

Thank you for taking the time to complete our survey. If you would like to receive a copy of the final report, please submit your email address.

[Submit]

Question 40 of 39

Suite 301 - 6190 Agronomy Road, Vancouver, BC V6T 1Z3 Canada +1 (604) 822-2694

http://www.interpares.org/cs14/cs14_survey.cfm?q=40&CFID=25947&CFTOKEN=65390065 2005-Nov-16 12:19:42 PM

APPENDIX B

Text Responses to Free-text Questions and Supplementary Text Responses to Multiple Choice Questions

Note: Text responses of participants are provided verbatim, with the following exceptions:
(1) Personal identifiers associated with participants, such as personal names, e-mail addresses, employer/institution names, project names, theses titles, Web sites, etc. have been removed and replaced with [---].
(2) English-language translations of Spanish-language responses have been added in double square brackets preceded by the word 'Translation:' immediately following the original Spanish-language text.
(3) Occasional clarifying comments (e.g., to provide the full text of an acronym used by the participant, or to indicate a participant's non-response to a required question) have been provided by the author of this report and are enclosed in single square brackets.

Question A1: Briefly describe what a Geographic Information System (GIS) is to you, including what you consider to be its most important, significant, and/or distinguishing components and functions.	
Partic. ID	**Comments**
4760	GIS: computerized mapping system composed of discrete data layers that can be used alone or combined to create new layers. Data points, lines and polygons can have associated data tables and pdf viewable files.
4761	Visual spatial research tool; database (attributes) and satellite image processing are most important for my work.
4762	I usually describe it to the uninitiated as a database that can also store and output maps and other forms of spatial data. Its ability to store, manipulate, and create spatial data are what sets it apart from a traditional relational database and are its most significant components.
4765	A GIS is software, hardware, human resources, and data. The most important component of a GIS is the human resources managing or utilizing the system. For most archaeological applications GIS falls within two utilization categories. The first is that of a management tool. CRM managers use it store information about sites and artefacts that have important spatial components. The second category of use is the analysis of information. That is to say those users who take the information stored within a GIS and use it to conduct research further adding to our understanding.
4768	Coordinate and mapping information that is able to be easily identified, checked and replicated in the field. Accuracy, and ease and reliability transmission from field to desktop.
4771	GIS enables the spatial and temporal tracking of survey activities for subsequent research. The localisation of site data enhances repeatability, data integration, analysis and visualisation.
4774	Tool for collection and mapping of original data, analysis of relationships among the data, and high-quality presentation of results
4776	- Los GIS permiten tener representaciones espaciales que invitan a formular nuevas ideas de Investigacion. -Los GIS incluyen documentos de diferenetes fuentes y con ello pueden amplra relaciones no advertidas. -Un Gis complejo puede tener caracter historico e incluir procesos. [[Translation: GIS permit one to have special representations that help to formulate new ideas and investigation. GIS include documents of different sources and with them one will exemplify unapparent relationships. A completed GIS can have historical characteristics and include processes.]]
4780	A means of visually displaying disparate data obtained from manipulation of databases for the purposes of understanding.
4781	GIS is great development of Computer Science, I think the most important function of GIS is the new spacious study method and perspective, and give us archaeologists great help to interpret ancient sites and recover them virtually.
4784	An integrated set of software tools for the acquisition, storage, manipulation, analysis and visualization of geospatial data and their attributes.
4787	GIS are tools to manage data base and to perform spatial analysis.
4788	MapInfo
4790	A GIS is a way of manipulating spatial information. Its most important components are its ability to present data from different sources, and its iterative mapping capabilities.
4791	Data analyses.
4792	Un GIS es un sistema que es capaz de relacionar bases de datos complejas para ejecutar análisis con

	ellas y obtener resultados acerca de los procesos investigativos que se desarrollen. Un aspecto muy importante es la georreferenciación de todos los datos que toman parte en el proceso, tanto los de partida como los finales, lo que puede llegar a hacer universal las fuentes de información en cuanto a su ubicación geoespacial. Las salidas cartográficas adquieren de esta manera una importancia vital en lo que a capacidad de exponer de manera fidedigna los resultados se refiere. Pero lo mejor de todo, sin dudas, es la capacidad de interrelación y generación de planteamientos de campo e hipótesis a corroborar. [[Translation: A GIS is a system that is able to relate complex databases to execute analysis with them and to obtain results about the investigative processes that are developed. A very important aspect is the geo-referenciation of all the data that take part in the process, as much the departure ones as the end, that can make the sources of information universal as far as their geospatial location. Because of this, the cartographic outputs are able to fully expose in a trustworthy way the results they talk about. But the best of everything, without doubt, is the cabability that comes from the interrelation and generation of field approaches and hypothesis testing.]]
4793	GIS is a tool of analysis and modeling of geographically registered information.
4795	I see a GIS as it pertains to archaeology as the next generation toolset for analyzing, managing, disseminating data. With technology allowing us to collect more and new types of data. The GIS gives us great flexibility to store many types of data in one central location and then access by many means. Most important about the GIS is the added component of spatial thinking.
4796	Computer programming for mapping. Useful for historic overlay studies, inter-site relationships of cultural materials, and relating site locations to environmental factors.
4808	A GIS is, for us, a powerful tool for collecting, postprocessing and analyzing all data from an archaeological mission. We use as cartographic base the orthorectified and mosaiced image of each “cut”. In the same time we take a simple DTM of the surface and all of the important “structures” in evidence. After the postprocessing and the database correlation, we can take all the thematic maps that we need for the archaeological research using “queries”. Other powerful uses of GIS in archaeology for us are the “intrasite” correlation, the “digital survey” using all the carthography in our possession and so on…
4809	GIS is a computer hard/soft-ware system that integrates database, spreadsheet & other image files, into a Multi-user geodatabase. It includes regional data libraries such as the Florida Geographic Data Library (University of Florida, Urban and regional planning department). GIS allows analysis of various geographic information, including topological (point, line & polygon; vector images) and raster (grid or pixel) data. The system creates projects or displays data. It includes various software such as ArcInfo, ArcEdit, ArcView.
4817	Initially, it was the use of evolving software and other data and meta-data as it became available in the end of the 1980’s and early 1990’s. I was involved in the archaeological survey of [---] sites in the states of New York and New Jersey. This involved infrared transit survey and hand held remote sensing data incorporated into evolving site reports as different phases of research and fieldwork progressed, in the background of preservation laws. Some of the first GIS systems I first saw were like OZ (developed in Australia) used in the demographics of voting in the Washington, D.C. area and perhaps used by the White House. AutoCad which I used starting in version 2 point something, evolved into the incorporation of GIS data, and I used some of the integrally software developed within its framework then: QuickSurg from Schreiber Instruments, Inc., MR2 close-range photogrammetry from Rollei, Inc. and Prometric Technologies, with Schneider Instruments, Inc., and other softwares and hardwares. The production of reports involved a smörgåsbord of tools and methods to create overlays and management data for preservation evaluation and documentation.
4827	For me GIS is: a. a whole system to document the dig process; b. a tool designed for evaluating the patterns of settlements; c. a tool which improves interpretation on cultural processes.
4864	GIS is a tool that allows us to better understand the landscape, modeling the world that surrounds us in such a way that we can better view an answer to a given problem/question.

4865	GIS is a tool allowing the collation of a range of information in an easily accessible form.
4866	Database storage of spatial and attribute data. Visualization of such data, on-screen, printing, and other publishing (e.g., files and web). Analysis tools for such data.
4877	Tools for recording, storing and analyzing spatial data.
4878	GIS is a very useful tool.
4879	A database software package with the ability to store large quantities of attribute data, query that data and graphically display and analyse the results.
4881	A set of geographically-based data that, taken all together, make some kind of sense for any specific purpose (i.e., research). I could also agree with the idea that GIS is also the technology used to manage and analyze those data.
4882	GIS can be many things, ranging from a simple text-file describing the location of objects up to high-tech systems providing databases, development environments, analytic functionality, etc.
4884	To me, GIS is an useful, powerful tool to ‘translate’ (archaeological) (field)-data to the archaeological story of a site. I myself don’t use much excavation data, but mainly historic geographical data, modern altitude data and geological data to combine with archaeological features for archaeological heritage management. That way I can make map-like geo-referenced databases to explain features for planning and site management. NB. GIS is NOT a mapmaking program.
4891	A method for spatial analysis, display and storing.
4892	[not answered]
4896	GIS is an advance data collection system that allows your geographic data to be precise and presented in a format that can be used by many. The most important aspect of GIS is the ability to navigate and place items precisely.
4897	GIS is a software with which one can organize, analyse, and present any kind of information within its spatial context and the qualities of that spatial context. As an archaeologist, the spatial analysis, and more particularly viewsheds and cost distance are considered as human aspects of the landscape, since it is man who sees and walks the land. We consider visibility and cost distance as highly determinative for choice of site locations.
4898	[not answered]
4899	GIS is a descriptive and analytical tool used on spatially referenced data to (a) visually display it, (b) to explore the relationship between various classes of data/features/attributes, (c) delineate patterning (or the lack thereof) in b, and (d) generate new analytical queries based on the results of c. GIS can be used to model and/or simulate (“what if” questions) the past and present with an eye to outlining what could be in store for the future. For me the real power of a GIS is its ability to analyze large amounts of diverse data within one system.
4900	GIS’s provide a way to organise archaeological site data in 3 dimensions across landscapes and as a way to visualise palaeoenvironmental conditions. Its most important components are its ability to model site predictability by way of layering differing data.
4901	A GIS is a tool to organise spatial data to aid in the presentation of data and to look for patterns. Its most important components and functions are the relative ease in use and the comprehensive and ‘unbiased’ examination of the data.
4903	A GIS allows us to maintain all of our spatial data where it can be easily accessed, referenced, and edited. This includes both rasters such as hand drawn site maps, topo maps, and analytical calculations, as well as vector data such as shovel test locations, physical feature dimensions, and activity areas. Within a GIS, we are able to bring all of these elements together in order to more

	efficiently visually and statistically analyze distributions.
4909	A GIS consists of a computer component, a spatial component, a data component, and most importantly a human component. Beyond computer cartography, GIS allows us to analyze and create new data from existing data. The applications of this technology are not limited by any scale.
4910	I use GIS for research, so GIS is important for data structuring, data analysis, data manipulation, cross-data and of course spatial analysis.
4913	[not answered]
4914	A GIS system consists of several phases of systems that allow the collection, correlation, analysis, display, and presentation of spatially referenced data. Minimum components include a means of acquiring map and or image data of a spatial region; a system for acquiring tabular, spatially referenced data, ideally through a DBMS facility such as postgresql, Access, or any of a number of similar pieces of software; a means of integrating this information for analytical purposes and for conducting both standard and spatial analytical data analysis; and a means of displaying and printing the results of such an analysis.
4919	A GIS is a system for organizing, viewing, querying, analyzing, and reporting on spatial (geographic) data. A GIS should be able to import and export data, relate tables to data to vector features, georeference rasters to match the coordinates of vector features. Some GIS packages have 3D viewing and surface analysis capabilities.
4924	GIS is one of the most important tools in an archaeologist's toolkit. Everything studied in archaeology has a spatial context and the ability to display and analyse this context is crucial.
4926	GIS is another tool in the proverbial toolbox of a modern archaeologist. Just as I would call upon a geologist to interpret and extrapolate geologic data in the field, I would call upon GIS to quantify and qualify data sets concerning relevant information before, during, and after an archaeological survey, data recovery, and/or monitoring exercise. GIS's ability to capture large data sets and effectively represent them on maps is its greatest feature.
4927	The main use we put it through here is map creation and display, specifically field maps for archeological survey. I also use GIS for data analysis, predictive modeling, and designing archaeological sample surveys. The most important aspect of a GIS system, for my work, is the ease of data cataloging and display. Additionally, it greatly aids in point based spatial analysis--something I certainly do not want to do by hand anymore.
4933	A graphical link database that provides for analysis/demonstration/display of geographical elements of a geographically distributed database. It normally will allow for concatenation of elements based upon geographic data. The most important components must include a graphical element of maps, a database of findings/observations/conclusory data, and a software system for linking the elements of the database to their areas on the map. It is distinguished from a flat map in that it is dynamic: can be modified by the user based upon previous searches or information. Its most important feature is that by allowing geographical comparison of disparate data, it permits exploratory data analysis that can (and does) lead to an understanding of distribution of categories of information in relation to one or more concatenated elements for which distribution is already determinable.
4935	GIS is a database manager that uses a spatial reference as the primary key field in organizing the information. Its significance in archaeology is that space, location, is the only attribute about the archaeological record that we know a priori (except of course that the remains are the product of humans). It is therefore the single best was to archive and analyze archaeological remains.
4937	GIS = manipulation of data with a geospatial component. Most important function/use - assist in the analysis of large geospatial data sets.
4939	System that significantly assists in the visual representation of spatial data and which enables various types of analyses of these data.

4942	Ideally a hardcore spatial-analytical tool, although in practise it tends to get used mostly for illustrative mapping purposes.
4945	We use GIS at [---] to archive remote sensing data bases: balloon photography, close contour differential GPS and geophysical. We also integrate the small amount of excavation data that we have. We hope to use the analytical tools to study urban dynamics, but difficulties in creating data bases make this a slow and frustrating task. We have not used GIS for regional studies.
4949	GIS involves the use of electronic mapping and cartographic resources. Its most important and distinguishing functions is that it provides a graphical and analytical platform for the graphical and mathematical analysis and graphical display of relationships and interrelationships in the natural and cultural spatial environment.
4951	A GIS is a tool for managing spatial data in a manner that allows it to be easily and usefully collated, analysed and interpreted. Its main benefits lie in its ability to integrate and interrogate large amounts of data to reveal patterning relevant to research and management goals.
4952	G.I.S is a flexible tool for an experienced user in order to record archaeological data of any kind (architectural, ceramic, faunal remains, environmental parameters, geographical parameters) in such a way that allows him to explore both their spatial and temporal component. To my perspective, just obtained a master degree on GIS and archaeology (not experienced practitioner), their most important function is that they allow the user to manipulate data by combining its attribute and topological component at the same time. By performing queries and several types of analyses the archaeologist can obtain meaningful conclusions, only if analytical/archaeological thought overcomes his technological, “computing” aspect.
4955	I have mainly used GIS simply to present data - layers of maps of different data ‘melded’ into maps presenting relationships between these data; or specific fields embedded into background maps of various sorts. I have also used GIS programs to show relationships between features in micrographic images, treating soil thin sections as ‘maps’. Despite my limited application, I think that GIS should be used with research aims, i.e., to generate relationships that are not already clear, perhaps with many complex variables, across landscapes. Personally I have never had access to a computer with the power to do such complex analyses using GIS, nor have I had access to the training required to reach that level of skill.
4956	GIS is a state of mind, it draws together strands regarding spatial theory from a number of disciplines; GIS is certainly not simply a bit of software. A GIS can be seen as an implementation of GIS theory using some software and some data; nowadays, there are many makes of desktop GIS software application. Practically, GIS applications are distinguished by some kind of spatial interface and some kind of database management system. Personally, I find the data management and analytical functions (e.g., map calculations, visibility, etc.) of prime importance (it’s my job) but the ability to produce hi-quality maps and gain access to web-based remote resources are also highly significant.
4960	A GIS combines data base and cartographic data. Important concepts are: raster and vector data, layers (also: combination of layers by adding, subtracting, multiplying), thematic maps, buffer zones, calculations like exposition, slope, voronoi polygons.
4961	A GIS is a system for organizing and analyzing spatial and attribute data. Its most important features are the ability to handle data at multiple scales and compare attributes from different data sets. The ability to create an attractive and informative paper product is also important.
4962	MapInfo 5.x’s advantage is in combining points, lines and polygons in one table. Disadvantage - inconvenient thematic mapping, raster maps registration. ArcView 3.x - advantage in thematic mapping and layout tools, geoprocessing wizard, great plugins like Sp. and 3D analysts. Disadvantage - lack of and inconvenience of existing table editing tools.
4967	GIS is a computer package for manipulating, storing and retrieving map information. Its most important function is the ability to reprocess and manipulate map data to extract significant

	relationships between locations on the land surface.
4968	A GIS is a spatial database and mapping system. The benefit to me is that I have a quick way of screening all construction projects for known or suspected conflicts with a wide variety of cultural resource types.
4969	A mix of both computer software and remote sensing equipment. The remote sensing equipment ranges from a simple gradometer to the satellites used by NASA.
4973	A Geographic Information System (GIS) is a set of procedures and computer software for organizing, storing, analyzing, and displaying data that includes a geographic component. GIS is a mapping system and a tool for managing information according to where it's located.
4975	A GIS is a way to collect, manage and analyze data that has a spatial and database component.
4977	GIS, to me, is a spatially referenced database that, unlike other databases, can be queried on or about spatial attributes. It is a superior way to store any data type that can be linked to some sort of spatial coordinate system (i.e., UTM, latitude/longitude, grid on an archaeological site). The ability of a GIS to store data in layers or coverages, to combine those layers in many different ways in analyses and to display results in visual form seems fundamentally important. I also find the ability to do various kinds of simulations (i.e., deforestation, population growth) very useful.
4978	GIS is a means though which to organize, maintain, manipulate, and conduct analysis on data.
4980	GIS is a tool we use to display and manage data we collect in the field, particularly utilized in conjunction with GPS. From our standpoint, its most significant functions is in allowing us to track more accurately the data collected in the field and graphically display the results (both pre- and post-field) of these efforts. GIS has also allowed us to help show clients what has been done where, help our clients to manage what they need to do, and in particular, produce maps for field guidance and reports.
4982	GIS is a system of data management within a predefined space. It can include computer databases, computers themselves, global positioning systems, aerial and satellite imaging, and other graphics. Its significance is its applicability, it is used in any number of fields of study and can easily integrate those fields. Its most distinguishing component is it allows the abstract to be placed in the real world, in relation to real things, i.e., highway designs on a 3d depiction of the project area with real world coordinates. The data storage, linking, and querying functions allow the most mundane chores of archaeological data management to become greatly accurate works of art.
4983	GIS is a system of organizing and manipulating spatial information. For my purposes, the most important component of GIS (versus other types of computer mapping) is the ability to perform spatial statistical analysis. I don't use GIS for database purposes much myself.
4984	I can't believe you want us to fill this out -- you must not know anything about giving surveys...short, sweet and simple. Here's my intro lecture. GIS stands for "geographic information system":·is a special kind of "information system"; information systems are used to manipulate, summarize, query, edit, visualize - generally, to work with information stored in computer databases;·a commonly encountered application are the information systems used by airlines and travel agents to make reservations, check in passengers, etc.;·uses special information about what is where on the Earth's surface;·there are many kinds of information used in computers;·numbers: computers are used to add, multiply, divide, ...; text: computers are used as word processors ·to create, edit, send, and receive text;·pictures: computers are used as image processors ·lists, tables;·in spreadsheets;·sounds;·in music synthesizers;·maps and images of the Earth's surface ·in GIS ·why use computers to handle information? ·easy to store, retrieve, query, manipulate, send, receive, copy, display... most of these things can be done by hand, but only slowly ·paper maps are difficult to handle, store, send, receive, copy... ·GIS makes all of these operations easier.
4987	GIS is a system that I can insert data (coordinates, time, descriptions) about the excavation (trenches,

	findings) and see them afterwards altogether in a user friendly, easy to understand representation. Important: Representation of fuzziness (in time, in place, in description), 3D representation, interaction with user choices (i.e., zoom in/out, performing queries in order to receive more specific information).
4988	[not answered]
4991	A fully integrated spatial data management and analysis tool integrating archaeological data sets incorporating conventional databases, chronological matrix and thematic data-sets, 3D vector graphics at any scale, raster images at any scale and supporting image-bank seamlessly joined at the feature, site and landscape scale and incorporating facilities for output of publication quality drawings and supporting materials in a variety of formats for dissemination of results on paper and on the web. In my own case it is the principle data management tool for more than 60 acres of excavation and 10,000+acres of landscape data. It is fully integrated at the point of excavation operates in 3D and provides fully interactive access to all data sources.
4992	In short a spatial database. Its most distinct quality (vis-à-vis a regular database) is the spatial dimension with all the opportunities they provide in analyses and presentation.
4994	A combination of software and digital data that allows me to enter, store, manipulate and retrieve geographically bound data.
5003	Its most important real-world use for me is as a geographical database. I don’t do much in the way of complicated stats with it, but use it for things like assessing intra- and inter-site density, proximity of cultural resources to construction projects, etc. I also use it for visually presenting archaeological data, to both other archaeologists and non-archs. For major projects I also like to use it to centralize data in different forms, e.g., hyperlinks link to word/pdf docs, images, spreadsheets, etc.
5005	A GIS is a sophisticated database management system designed for the acquisition, manipulation, visualization, management, and display of spatially referenced (or geographical) data.
5006	A GIS is a way of managing spatial data.
5007	GIS is a tool, like a trowel, a shovel, or a notebook. Like any tool, the user must know how to use it and what the proper uses are. GIS is very important to me for its ability to display a variety of data types and forms in a easily understood manner. The data merging and number crunching that most GISs are capable of conducting also allows me to test many different hypotheses within a matter of minutes.
5010	It’s a tool that allows a set of data, both spatial and tabular, to be quickly analysed. The most important functions are those in which major number crunching is done (calculating distances, creating surfaces, etc.) that no person could do manually.
5015	GIS is a database linked to geographic space. The most important function is the ability to use queries and logical arguments to look for hidden relationships between numerous datasets, including geographic data.
5017	GIS is to me a much more sophisticated method of keeping track of the archaeological excavation, with many advantages over the more traditional methods. Advantages that I clearly perceive include: a) Greater accuracy in the surveying of the site; b) Greater accuracy in plotting finds; c) Real-time control of any possible problems; d) Real-time control of horizontal and vertical distribution of features and finds; e) Excavation records are computerized on the spot instead of months later.
5026	A GIS is a map tied to a database at its core. It also offers significant spatial analysis capabilities which are ultimately the most useful to me.
5028	Because Archaeology always dealing with data handling. GIS can use Archaeological data recording, analysis and interpretations. It is very useful tool to me.

5034	A GIS is system that allows the storage and retrieval of a wide range of data and data types. Besides the map producing quality of the GIS another important function is the ability to include a vary wide range and wide diversity of data into one location or program.
5036	Information system that contains geographical and/or time-dimension and attribute-data.
5043	An automated mapping system that allows me to compile, cross-reference, tabulate, and reorganize all of the quantitative and qualitative data presented by that map.
5048	A combined database/mapping software tool that enables locational data to be visualised on a map or on a remotely sensed image. Can be used for searching, querying and extraction of information. Most important use in terms of our business - is ability to search for archaeological point data and observe its location on the map (historical, modern or photographic). Distinguishing functions: as a visualisation tool and database/query tool.
5050	It is a spatially oriented database, viewing and manipulation system that allows the user to investigate, measure, and describe the spatial distribution of the data. Its most important function is allowing the user to visualize the spatial relationships between data with a single dataset, AND relationships to other datasets.
5055	A GIS is a tool for collection, analysis and management of spatial data. In my daily work the main task for a GIS would be data collection and map production. Not much analysis in other words. My main interest though is in making GIS a useful tool for archaeological landscape analysis. So spatial/geospatial analysis, 3D analysis are very important functions/components for me. (viewsheds, etc.)
5058	MapInfo. Significant aspect is the ease of use for basic mapping/modeling.
5060	The integration of geographic information and a database. This permits researchers to query geographic data, and non-spatial data that is associated with a geographic location, to model or discover more complex relationships than can be observed on paper maps. These relationships/model can be quantified and tested.
5061	A system for organizing data based on locational coordinates. After the locational information is supplied, various other layers of information can be overlain on it so that a point will have a number of variables associated with it. Each of these variables can then be analyzed (singly or in combination) against location, with greater facility.
5069	A system to house spatial data sets and allow analyses between these and other related data sets. Most important attribute is the visual representation of data sets to allow meaningful interpretation. Data representation is organised around real life perceptions and so facilitates non technical specialists to make use of GIS systems with increased ease.
5071	In my opinion GIS is a valid support to everyone who works in landscape management. GIS has to be easy to use, friendly in the front end and open to further technological development. In particular archeological use, GIS should answer to research problems, like population distributions or risk areas. The biggest problem in our field is that doesn't exist a specific tool to manipulate archeological data, so, at the beginning of a case study about GIS and archeology, you have to put your interest choosing the right instrument. No so easy find the right one.
5072	A GIS is a kind of computerized data base that allows for the storage and referencing of data so that it can be displayed and analyzed spatially, that is, through the use of computer generated maps. Data collected include geographic, topographic, botanical, archaeological, and others. Such GIS can be used for research, and they can be used to produce illustrations (maps) for use in reports.
5076	A spatial mapping tool.
5077	Definition: A spatial database (when used to its full potential). Key function: *The ability to integrate archaeological spatial information with proposed development plans may enable heritage protection

	and require developers to comply more fully with legislation created to protect significant archaeological/historical sites.
5083	GIS is a cultural resources management data management and research tool. Its most important feature is its ability to link maps to databases.
5096	A storage system for spatial data. Important tools: search engine, coordinate converters, select tools, graphic tools for maps and data.
5110	To me, A GIS is a powerful tool that links information about a feature to its physical location on the earth. The most important function of this in maintaining archaeological data is that it makes it available in electronic form and allows spatial analysis to be conducted in order to gain a broader understanding of site relationship. This can be taken further, in that site locations can be predicted and archaeologically sensitive zones can be defined for those areas where no cultural resource surveys have been conducted. It can also continuously be updated to allow for new information. In CRM defining archaeologically sensitive zones allows us to provide archaeological information about a specific area without giving away site locations.
5185	A GIS is a tool to store geographical data, which enables the comparison of different data with each other. It makes it possible to visualise distribution patterns and makes it possible to analyse data in a statistical and geographical way. In heritage management it enables the production of planning related layers which makes an integrated approach easier.
5188	GIS is a suite of software tools for the collection, management, articulation and analysis of spatial and attribute data. These systems allow researchers to articulate spatial and a-spatial phenomena with varying degrees of complexity and at different scales.
5190	GIS is a spatial tool and a geodatabase to link potential attributes records to spatial objects. Its greatest ability to derive meaningful spatial understandings from patterns with the ability to query subsets of data determined by its attributes; visualisation, query and reproduction of data are its most significant characteristics.
5193	A GIS is a software platform allowing the integration of spatial data and attribute data enabling data management and analysis.
5194	A indispensable tool which allows for the graphic display of complex land management information and research data. It must be relatively easy to operate and be compatible to a wide range of external data sources.
5196	A GIS system provides a digital workspace where spatial information associated with objects and databases can be interrogated and defined. The most important aspect of this software for me is the ability to view datasets, querying associated information and to have spatial relationship displayed in a digital environment, which can then be manipulated.
5198	A single location where all your geometric information is coordinated into a single coordinate system and where elements in this geometric database can be quickly & easily (preferably automatically) linked to other data sets wherever they may be stored.
5199	A GIS is first and foremost a data management system that is able to deal with geographical as well as alphanumeric data. Its ability to conduct spatial analysis is of secondary importance.
5201	A means of managing spatial data and relational information in order to manage, interpret or educate about specific aspects of the environment. In my case, this is the historic environment.
5202	This is a very big question! I'd rather not get into the definition question other than observing that a GIS has to have three elements - geography, information and systems. The significant properties of a GIS are the abilities that come with being able to manage complicated data relationships through a shared topology which mimics the real world. This allows innovative research questions to be formulated between entities whose relationships might not otherwise be modeled.

5203	A method of bringing together geographical information, and allowing that information to interact dynamically with other data sources such as databases, images etc. Most important aspect- allows information such as finds data to be displayed geographically, allowing it to be analysed in a way that is closer to how it was in the ground. Also, makes data analysis quicker and easier!
5204	A GIS is a system for managing, portraying and analysing spatial information, especially for looking at the characteristics of spatial distributions of artefacts on a landscape and their relations to natural and cultural features of that landscape.
5205	A GIS is a system for generating, manipulating, storing and displaying geospatial data. The manipulation bit is most important to me at the moment: The ability to perform spatial analysis.
5206	GIS is a tool which allows us to disseminate spatial information relating to the historic environment of our region to a variety of stakeholders, and to interrogate the information ourselves to enable an informed response on our part. The most important facility is the ability to depict areas (by polygons) rather than just points, and to allow simultaneous viewing of any number of configurable layers of data, or to allow us to relate different datasets in a spatial framework.
5212	Mainly a map production tool that allows to easily bring together a disparate array of spatial data. The analytical functions of it are mainly a tool to get there. For example sometimes certain points have to be included/excluded based on criteria, areas have to be selected based on polygon attributes, etc., but all this is for the goal of getting a good map of my data in the end. Main data are geophysical surveys.
5213	[not answered]
5214	A way to index spatial information about archaeological excavated materials along with map-based information about a specific excavated area.
5219	GIS is a computerized data system that stores, displays, outputs, and manipulates spatial data. The analytic aspect is the most important, along with the ability to integrate data from numerous sources: imagery, scanned maps, GPS, etc., etc.
5220	It's the organizational framework for geospatial database. Since archaeology is a spatial science, it is the only effective way to track information elements in time and space. Once they are tracked and managed, you can analyse them by which ever method meets your research objectives.
5223	GIS allows us to document our site types and better understand settlement patterns- we have yet to employ predictive modeling to "pre-locate" sites but will in the future. It also serves as an excellent CRM tool by allowing us to quickly review known site locations, historic photos and maps and geological information simultaneously to better understand a proposed development site.
5226	Linking of a digital map to a digital database thus enabling spatial queries and spatial statistics.
5228	Geographic Information System (GIS) is a useful tool for database storage.
5229	A GIS is an instrument that tries to give an accurate coordination for located artefacts or sites.
5231	A GIS is a database with both graphic and tabular components. The data are spatially referenced to a grid system and the database can be searched both using the tabular or spatial data. Most important - spatial analysis.
5232	GIS is the new world where for the first time (because of low-cost hi-power computing) disparate sources of seemingly discordant and diverse data is combined using common ground reference points to 'lock' each 'layer' into a presentation, the value of which to knowledge creation is far greater than the value of any item or individual layer.
5233	Tool for interpreting archaeological record using spatial archaeological and topographic information. As such it can provide modeling over space and time, however, its use is Australia has been severely limited to use as a cartographic tool and the full potential is yet to be explored. This is particularly the case for private industry, such as mining companies who all used GIS software but only use its

	cartographic functions.
5234	GIS or geographic information system is an analytical tool that provides a spatial-temporal dimension to any kind of data. GIS has three main components: data system, analytical system, and report system.
5249	GIS is an important tool to connect archaeological, environmental, cultural, ... data in order to help recognizing patterns in distribution and also to recognize sites, objects, ... which do not fit into patterns. GIS can help to find explanations; both for patterns and for "non-fitting" observations.
5256	Database, spatial analysis and visualisation tool.
5257	A means of combining spatial and non-spatial information from a variety of sources. A means of deriving new information from the comparison, iterative querying and modeling of data sets therein either individually or in myriad combinations.
5258	A GIS is a tool to combine textual- (database) and graphical- (vector/pixel) data. It helps to analyze an amount of collected data usual methods can not deal with. Most important is the visualization of data.
5262	GIS is a spatial database management system. Its functions are spatial data capture, manipulation, analysis and presentation.
5268	For me, a geographic Information system is a multi-component, computer-based technology that allows users to interactively relate digital records containing information about the spatial location and character of cartographic features in a given landscape with digital, non-cartographic information about those features resident in a database or databases. Beyond the value of GIS as a tool for storing, managing, and manipulating large amounts of information about particular landscapes, I believe the most important/significant/valuable function of GIS technology is to facilitate the analysis of cartographic and related non-cartographic data in a manner that leads to the generation of new information and insights.
5290	Geographical Information system is a computerized platform that allows the user to combine qualitative and quantitative data to a geographical location and use its analytical capabilities to produce new type of Data/information. Usually this type of data is more or less affected by the geographical location/position which is the main component of the analysis. The main components are: 1-handling of data, 2-manipulation of the data and the most important is 3-analysis of the data. Some examples of analysis: Site catchment analysis, viewshed analysis, cost surface analysis, predictive modeling, hazard (flood, fire) mapping, etc....
5292	GIS is an integrated technology/software system with the capacity to link tabular and spatial data and to manipulate these data. We use GIS for information management and research. Specifically, since 1983 we have used GIS to maintain and update the Illinois Inventory of Archaeological sites. We have also used GIS to generate predictive models of archaeological sites locations and reconstruct paleo-topography. In short, GIS revolutionized our capability and ability to manage and explore archaeological data.
5293	It is a tool by which to create spatial databases to track and analyze spatial data, including archeological sites but also other resources of interest. Most importantly it must be 'geographically' spatial and allow analysis of spatial relationships of multiple resources or feature types. It is not a CAD system, even if that CAD system can provide a georeference. CAD was designed for architecture and design, not for spatial analysis of resources.
5297	GIS is a tool and techniques to have integrated frame of analysis taking together varieties of data and the maps. By superimposition, comparison and links we reach to precise understanding and generalizations. Especially we are using GIS in the area of cultural astronomy comparing the orientations and correspondences related to the GPS and related values observed at the ancient sites of the shrines and pilgrimage centres. The GIS helped us to uncover the mystical, cosmological and mythological derivates related to pilgrimage centres in India, leading to support and develop a fresh

	area of research called "sacred geography", or "sacred geometry".
5301	A GIS is a computer program to analyse spatial data and is combination of a database, a statistical program and a graphic program. For me the combination of these functions is the most important aspect. 25 years ago you could do the same things (more or less) but you would have to switch from one program to the other.
5308	GIS is a means of compiling spatially referenced digital data and analysing these data sets in relation to a variety of other spatially referenced data sets and coverages.
5321	GIS is an important methodological approach that serves to achieve answers about landscape, social organization or any other kind of answer that can be related with spatial organization.
5323	Un SIG constituye un conjunto integrado de hardware y software que tiene las capacidades de ingresar, manipular, visualizar, analizar y consultar datos en formato tabular y gràfico. Cada una de estas capacidades tiene ventajas particulares para cada tipode trabajo, sin embargo, la capacidad analìtica es la primordial. [[Translation: A GIS constitutes an assembly of integrated hardware and software that has the capacity to enter, to manipulate, to visualize, to analyze and to consult data in tabular and graphical formats. Each one of these capacities has special advantages for each type of work; nevertheless, the analytical capacity is the fundamental one.]]
5324	I think GIS is a valuable tool that connects databases with geographic information. Because of this important characteristic its most significant function is the possibility of crossing data collected in surveys and excavations with its spatial information.
5330	Analisis sobre el desarrollo y crecimiento Urbano del asentamiento [---], Argentina. (1000-1500 d.C apróx). Se basa, con exclusividad, del analisis cronológico del material en superficie de todo el sitio para estimar la extensión y ciertas caracteristicas de los diversos momentos de ocupación. Asismismo se consideraron los Procesos de Formación (Schiffer, 1887) que incidieron e inciden en la configuración y ubicación actual de este registro mediante el análisis de mapas (layers) de topografia, suelo, cobertura vegetal, microcuencas de escurrimiento acuiferas. Con ello esperamos definir la correcta escala de resolución a la problematica tempo-espacial planteada. [[Translation: Analysis on development and urban growth of the settlement [---], Argentina. (1000-1500 BC approx.). It is based exclusively on the chronological analysis of the material surface of the entire site to estimate the extension and particular characteristics of the various periods of occupation. Also, the analysis of maps (layers) of topography, ground, vegetable coverage, water-bearing microcuencas of dripping considered the Processes of Formation (Schiffer, 1987) that affected and affect the configuration and present-day position of the intervening records themselves. We hoped to define the correct scale of resolution to the problems which its spacial-temperal characteristics presented.]]
5331	To be able to collate and manipulate data with a spatial component which otherwise would be difficult to manage and especially to visualize. It also broadens the way we may visualize data by adding the 3D aspect and alternating the emic and etic perspecives (eg., watershed views). Gathering the data with the all the indispensable variables and compiling the databases also requires a more disciplinarian approach to data-gathering.
5338	GIS for me as an archaeologist working and teaching with all digital technologies is an active tool for the archaeological research process, an active tool and not only a passive one. Me and my Italian colleagues are working from 1999 with GIS on excavation and on survey research projects with powerful results and new implementations, from spatial analyses to 3D management of excavation stratigraphic deposits. So, we are now in a very good position to say that working with GIS in our researches is to work with an important tool for analytical and study phases.
5347	A GIS system is a useful tool for storing, accessing and viewing site information in a spatial manner. Questions can be asked of the data and further investigation occur through the GIS system. The integrity of the coordinates input into the GIS system can be readily viewed for confirmation of

	accuracy much quicker than conventional methods. Searches of specific characteristics of the data are easy to perform and visible in a spatially oriented format. The most useful feature in the versions that I have access to are the ease (once the system has been set up to do what you wish it to) of accessing useful data and checking the spatial and coordinate integrity of my data.
6000	An electronic means of recording, databasing, querying, accessing archaeological field data, allowing archaeologists to develop new ways of looking at archaeology and understand archaeological processes. Significant components: possibility to integrate and share data easily across boundaries forcing archaeologists to stop and think and re-quantify and re-qualify the data we wish to capture, this should in turn force us to "talk" with each other and share methodologies.
6001	GIS is a good tool to store and analize data with a spatial component. So one can perform intra-site studies starting from locational features of archeological records (finds, monuments, structures, etc.). Main GIS feature is the topology that allow us to explore spatial patterns. Also query builder to carry out thematic maps starting from alphanumeric data is a significant component of GIS system.
6002	A geographic information system is a computer based research tool, performing an integrated analysis of descriptive and numerical data in combination with information retrieved by processing maps and drawings in digital form.
6003	[not answered]
6004	For me a GIS is a tool to calculate and visualize archaeological predictive maps. The main function is to handle with geographical data of different sources and the possibility to combine them. Important for me is also the usage off different software products, which is sometimes a problem.
6005	Computer application displaying database and spatial analysis functions. Its ability to combine both functions in one application is the greatest asset.

Question B1: How long have you been using GIS?	
4776	El uso del Gis se efectuo en relacion al proyecto de investigacion de arte rupestre en el municipio de [---] –altiplano cundiboyacenese-Colombia Suarmerica. Se ubicaron mas de 2000 puntos de sitios rupestres. Diversas capas se incluyeron relativas a la estructura historica y cartografica. [[Translation: The use of GIS can help protect rock art in the municipality of [---]—a high plateau cundiboyacenese—Colombia, South America. Over 2000 rock art site have been located. Diverse layers include relationships between historical structures and maps.]]
4780	Not as a power user, but on the periphery
4808	we started to use GIS with Dr. [---] of [---] University.
4817	After the Superfund basically became an impasse over mortgage responsible legal loopholes, much of the research was not done it seemed to me. So I personally went back to background research and digging holes for various companies.
4884	I started using GIS-like techniques during uni, when, in Holland, nobody, but a small group of students, used GIS in archaeology.
4899	While I have used various types of GISs over the time span noted above such use has not been on a full-time basis.
4910	I'm coming to GIS after a good experience of DBMS
4914	I am familiar with GIS and worked for a vendor, but have rarely been able to use such systems.
4933	The first GIS I built was a paper-model of a two county area, with more than 50 separate pen-and-ink iterations to document significant overlaps of data categories based upon elevation, drainage basin characteristics, vegetation, fauna, aspect, and others in comparison with age, material content, and

	location of a selected subset of sites. More recently I have been using the Pocket PC with GPS and mapping software to make simple GIS models to document contract surveys.
4952	G.I.S. seems very attractive to an unexperienced archaeologist but he is not able to obtain meaningful results, if he is not guided carefully, because their applications entail numerous tasks (transformation of geographical coordinates, statistical tests) that are quite unknown to an archaeologist.
4956	I began during my undergrad course at university before undertaking a postgrad qualification in GIS. I have 2.5 years working with GIS in a cultural heritage environment.
4969	I have been using it in my graduate school training
4980	From our experience, most individuals, both other archaeologists and our clients have heard of GIS but do not fully understand the tools capabilities or the time necessary to complete tasks.
4987	An excavation in Greece..
5010	Started the fall of my junior year of college (Fall 2000).
5017	My experience is limited to the use of EDM connected to hand-held computers in the field, then daily download and processing of the data with ArcView.
5034	I have used it mostly in a graduate school situation but recently have begun using it more intensively with my thesis.
5050	First trained on ODYSSEY, the first GIS system.
5055	First used ArcView 3.2 at University, then also at work. Used GRASS 4.3 and 5.x at home under Linux/FreeBSD parallel to this
5061	No formal training. Was expected to “pick it up” on the job.
5069	on and off. But also started writing GIS algorithms almost 20yrs ago
5072	Working for U.S. Bureau of Land Management in northern [---] 2001-2002 was my opportunity to work with a developing GIS system, staffed by experts with good hardware and software.
5083	If GIS had not been invented for us, archaeologists would eventually have invented it for ourselves.
5198	Started in the early 90’s using AutoCAD’s ADE to produce CAD drawings from data files, primarily downloads from survey instruments annotated entity details
5229	GIS have been used on sites I have worked on for many years, but I personally have never used one.
5232	Originally for feature film CG special effects, but the same soft and hard ware delivered far better visualisations than that as distilled from the CAD tools available at the time. Convergence now sees the best of the math and science combined with the best of visualisation.
5258	I never really used a real GIS-software, but was working with ‘homemade’ AutoCAD applications using MS-Access Data and ODBC.
5290	I began with a custom made GIS software (G-sys) designed by a British archaeologist ([---]). It had the handling and manipulation components more developed than the analytical one. It was also more vector oriented with few raster capabilities. Afterward I started with IDRISI and used it for more than 6 years and now I’m using ArcGIS for than one year.
5292	The State of Illinois invested in GIS technology in 1983. We have been using it for managing archaeological data since.
5297	Mostly we are using GIS to understand the locational mysteries of shrines and divine image

	associated to holy places in India with the collaboration of Prof. [---] (Dept. of [---], University of [---]).
5308	Started in 1982
5331	My use has been intermittent, limited mostly by slow progess in creating the databases and the elevated costs of obtaining the basic geographic/cartographic layers.
5347	Please note that my use of GIS has been basic—the systems already set up for me and for specific purposes.
6001	In Italy there aren't school to learn the use of GIS in archaeology. So I realized experience by myself.

Question B2: Which of the following most closely describes how often you use GIS?	
4780	Most of my effort is in making an abysmal system work via influencing the bureaucracy that controls it.
4790	I have months when I use GIS daily, and months when I do not use it at all. I use the data which I analyse using GIS daily, but GIS analysis is only one tool of many.
4792	El hecho de usar diariamente el GIS no significa que mi labor cotiadiana se refiera a la investigación en este campo, lo que pasa es que gran parte de la información que se genera a diario en mi empresa se ejecuta mediante GIS fabricando cartografías para nuestros proyectos que se nutren cada vez más de bases de datos propias. [[Translation: The fact that I use GIS daily does not mean that my "cotiadiana" [specialty?] work refers to the investigation in this field, what happens is that much of the information that is generated every day at my company is done by using GIS to make maps for our projects that are derived more and more from our own databases.]]
4808	We today use GIS in every archaeological mission.
4945	Members of the [---] Team use GIS daily since it is where the data is stored. I myself rarely use GIS but rely on our research students and other team members.
4951	I would use it more frequently if heritage agencies had GIS data easily available for the study areas I am interested in. Unfortunately at present in Australia much of the necessary data sets are not freely / easily available (or even exist for some parts of the country).
4956	I am a GIS specialist! The bulk of my work is either using GIS or directly related to its use.
4961	Some weeks I use GIS daily others I go for weeks without needing it.
4967	This varies depending on whether I am actively engaged in analysis, or merely writing up the results of a previous analysis.
4977	My use is sporadic; I may use it almost every day for a couple month period, then not at all for several months.
4980	Our use of GIS is very seasonal. During the field season it is used weekly and in the winter months it is used monthly at best.
4987	Only in excavation periods (summer months).
4992	I am more involved in the design and maintenance of GIS rather than the operational use of desktop-GIS.
5010	Both at work (MTThF) and in my own research (most non work days).

5017	Two months a year, during the excavation, then along the year to produce reports etc.
5034	Daily, especially considering homework assignments.
5048	I work as a Sites and Monuments Officer for regional archaeology trust, so I use a combination of database and GIS tools daily to record and update new information.
5055	Sometimes daily, most of the time on an ad hoc basis (but never less than 2-3 days a week)
5069	Frequency varies with access. I am doing a GIS-based Ph.D. part time so varies from daily to yearly.
5072	I used GIS all the time while working as an archaeologist for a federal agency, 2001-2002. Currently, I am a male version of a housewife.
5077	I work as a GIS officer.
5202	I mainly now use it to teach about GIS - and less for my own research / management.
5229	GIS are used only on sites where big money has been raised as they are expensive and the cheaper ones are fairly inaccurate.
5249	I'm working for a research project, dealing with the so called "princely sites" of the Early Iron Age in Middle Europe and their environment
5290	It depends on the assignment and needs. For some periods I use it 24h a day to finish the work to be done and then I could have other types of assignments that doesn't need GIS.
5292	My use of GIS is generally limited to ARCVIEW sessions, but other staff members use ARC/INFO on a daily basis.
5297	It depends upon the running of the project, e.g. during 1999-2002 sponsored by the [---] College, London, we have regularly used GIS. However, since last year 2003, we have not yet used GIS. By August 2004 fully developed GIS Lab will start functioning in our Dept. which will help us for regular working.
5331	Work has been intermittent and mostly due to difficulties in creating databases, obtaining cartographic layers and an abscence of peer interaction on the subject since GIS methodology is still new in our community.
5347	Use of GIS can vary from job to job and their frequency. When there is a great deal of work on various projects use of GIS increases and when there are fewer jobs available it decreases accordingly.
6001	I use much more the GIS during the excavation or when we prepare our report-excavations. So the use is "seasonal."
6002	Periods of everyday use are followed by long intervals of no use at all. "Montly" has been chosen as an average that describes best the condition of use.

Question B3: With which of the following phases of a GIS project(s) are you or have you been involved?	
4792	Me resulta difícil exponer mi mayor o menor experiencia en los aspectos arriba mencionados, pero creo que en el campo que mejorme muevo es en el del análisis e implementación de datos, aunque no podría decir que esta sea mi especialidad. Quizás el motivo de esto sea que nos encontramos en una fase de inicial de nuestros proyectos GIS. [[Translation: It is difficult for me to explain more or less my experiences in the above-mentioned areas, but I believe that in the field that my major role is in the analysis and implementation of data, although I could not say that this is my specialty. Perhaps the reason for this is that we were in an initial phase of our projects GIS.]]

4808	Obviously every mission is a team project and each time we redefine the role of people involved.
4817	For example digitizing old maps (using IDRISI for example) overlaying new maps and GIS data from the USGS, determining where historical activities took place. Or from a photo of an old map on the wall, trying to incorporate the past and the present. Also digitizing close-range photogrammetric photos on a large digitizing tablet to create 3D records of historical objects, plans and profiles of excavations. The depiction of magnetometer surveys of former active areas in a marsh, or on land to show possible locations explored later. Tree surveys for botanists, dbh measurements and species identification used in the planning process at [---] University, [---] and [---] and the [---] for example. Integrating old depth measurements into modern bathymetry around historic shorelines, etc.
4897	The matter of data preservation is a difficult one. In Greece no standard data formats have been decided for any kind of storing.
4910	I'm working only about little GIS projects but every project must to answer at one research question. so I'm necessarily involved in all phases.
4937	If given the choice I would have listed most experience as a-c and least experience as e
4945	I am not sure how to answer this. As a team we do all of these things.
4956	I have designed and built specific project-based GIS for data capture and analysis although data entry on such applications has generally been done by others. I prefer not to develop tied to a project, rather develop generic reusable components that can be used on other projects where possible.
4978	My work is such that I am just about equally involved in all phases.
4983	It seems difficult to differentiate some GIS projects (or aspects of GIS projects) based on these categories.
4991	It is impossible to rank the above in a sense as each relies on each other and it makes an assumption that one should have 'GIS Projects' as an independent activity, in a practical case all our projects utilise GIS, it is the project that comes first, the GIS is simply a day to day management tool which offers particular functionality.
4992	Design is preservation—meaning that it's an integrated part in the design process of a large-scale GIS rather than add-on or afterthought.
5055	Most of my experience with AB is from planning/developing my own project in University. I also have some experience planning/developing a (Norwegian) governmental project (member of reference group). ABCD experience is both from uni project as well as work related and private projects (involving different systems)
5072	I worked with [---], administered by the [---] State Museum.
5077	Many GIS packages are not suitably geared up to data preservation. I feel this is a future step.
5206	C is listed as most experience because it took only 3 months to plan and design a project, but 3 years to enter all the data—complicated by the fact that we had to change GIS systems part way through the project.
5220	I don't enter a lot of data, I use the data after it has been entered.
5290	I'm developing actually a project for the automation of the geographical data in my department which is responsible for the archaeological sites in all the country. So it will be a new experience in Data preservation and Development and design aspects of GIS.
5292	I directed the development of the GIS-based [---] Inventory of Archaeological Sites, and although I have used it for analytical purposes, others generally do the first-hand manipulation of the data.

5293	This is a tough question, as I've had more or less involvement with all at various times. When I first started, I did everything myself, from planning through analysis. Now that I have a team I probably do more planning and development, but I still like to do the implementation and data analysis on things I am really interested in.
6005	Adding and executing scripts is still part of my GIS involvement although this is done on a rather restricted scale (not project).

Question B4: Which of the following GIS system designs do you typically use?	
4771	My research is based upon SEG-Y data where GIS and attribute data are stored together
4808	We have two different session for an archaeological mission: the first is the collection and instant analysis of carthographic (low range photogrammetry) data, the second is the postprocess and study (with all the phase plants) of all the data collected.
4817	The complicated phasing and logistics of projects (doing too many at once) and the nature of the tasks, integrating old maps with modern data, precluded a "modern" GIS approach. However, in evaluating a number of proposed parcels in the [---], road maps and AutoCad maps were integrated using Microsoft's database Access along with evaluations so that the client had a GIS of locations and potentials for archaeological significance.
4919	I use ESRI software - ArcGIS & extensions, ArcView 3.x, ArcINFO, GRID
4945	We have spent much time and sweat converting data sets so that they are compatible, i.e. use the same co-ordinate system.
4956	Spatial data stored as Shapefiles containing geometry plus UID only; attribute data resides in Access rDBMS; ArcView3.2 provides spatial interface.
4962	Stand-alone DB application + GIS
4991	In essence we have to handle all the different GIS design structures simultaneously as we are passionate about the use of primary data from and supporting a large number of specialist areas many of whom may not be able to handle some of the more complex data structures, the key component is the multi-dimensional linking of all data regardless of its structure at origination.
4992	The system I'm currently working is an "extended design system" but with an integrated capacity for date-extraction in file formats.
5048	Use Microsoft Access linked into GIS system
5069	This is developing across all types and will change again. Mostly hybrid and will import DBMs data into GIS for certain things but not rely on it as the format is not as interchangeable and flexible.
5185	We use an extended design system in an off line GIS system (with ArcView), we also have an on-line system which uses a mySQL database-structure.
5292	We use ARC/INFO, but archaeological data are organized in a series of separate files.
5293	It really depends on the project. For our standard data management, Hybrid is the most common, as our attribute database is a corporate database, but our corporate has not spatial database requirements. So we create the GIS for our own needs, and link the attribute data. For some small scale project for a specific study, however, we may maintain all the data in one system, usually the GIS.
5331	Access with output to Dbase files; OziExplorer & Excell used as intermediaries for downloading and importing data. ArcView as GIS software.

Question B6: What is your current professional archaeological affiliation?	
4762	My employer is a joint operation between the university and the state historic preservation office
4771	Maritime archaeological research
4776	[---] es una organizacion que investiga la estetica precolombina: www.[---]. [[Translation: [---] is an organization that investigates Precolombian art.]]
4790	Part-time PhD student, Full time CRM (Public). GIS used extensively in former, slightly in latter.
4791	Museum
4792	Desarrollo mi labor cotidiana en mi propia empresa [---] que se dedica a la investigación y desarrollo del patrimonio cultural y natural, especializada en aplicaciones arqueológicas, impacto ambiental y procesos formativos. Además estoy unido a la Universidad de [---] a través del Doctorado y la colaboración con diversos profesores de esta universidad y también la Universidad de [---]. [[Translation: I've developed my daily work in my own business, [---] that is dedicated to the investigation and development of cultural and natural resources, specializing in archaeological applications, environmental impact and formation processes. In addition, I am working on my doctorate, affiliated with the University of [---], and collaborating with numerous professors of this university and also the University of [---].]]
4817	"[---]" [Editor of an online archaeology newsletter]
4881	Publica scientific research institution [---]
4919	Software documentation, interaction with archaeologists using ESRI software.
4927	Graduate student, GIS based thesis research. Also National Forest Service, Heritage Resources.
5006	Indian Tribe
5007	I am an employee of a private CRM firm who also conducts independent research
4956	I also undertake post-graduate level teaching of GIS to archaeologists on a casual basis (i.e. am not affiliated to any particular university)
4962	Institute for the [---] (i.e. Scientific Research Institute)
4991	Independent research institute
5010	2nd year Masters Graduate Student at [---] University
5034	Graduate Student also.
5055	I work in [---] County CRM as an archaeologist. Educated at [---] University, [---]. Thesis: "[---]". Still work (privately) with projects on landscape, GIS and archaeology.
5072	Currently unemployed, but my archaeology jobs have always been with the government: universities, state agencies, federal agencies.
5077	Now full-time in CRM, but I was a previous lecturer/teacher and I do ad hoc university work.
5083	Private utilities company; private CRM consulting after June 24, 2004.
5096	Research and publication
5190	My employers are a research institute that carries out government funded research as well as contracted development-led work.
5193	Contracting unit

5201	Outreach
5202	Archaeological digital archiving / research support service based in university.
5228	I am an anthropologist.
5249	Governmental research institute
5257	MIFA Member of the Institute of Field Archaeologists (a UK body).
5258	Field Director in 5 month yearly archaeological excavations.
5290	I did some GIS teaching at the University. Actually I am responsible for the archaeological map in the [---] Ministry of Culture.
5292	Although a museum curator involved in archaeological research, collection management, and cultural resource management for more than 20 years, I now serve as Director of [---] Museum, a branch of the [---] State Museum.
5331	Belong to the Departamento de [---] of the Universidad de [---].
5347	I am currently studying Honours at [---] University part time. I also do some subconsulting for Archaeologists as well as work as a contractor for a government department.
6005	[---] Museum, [---], Denmark

Question C1: How often do you use a standardized and/or documented procedure for naming the digital files you create related to each GIS project?	
4790	Never. As a researcher, I am the only user of the files, so I have no need to follow a procedure. I sometimes use names that make sense (e.g. Norway.dbf contains data relating to Norway!) and sometimes I use names that relate to presentations (e.g. York.dbf contains data for a presentation at York), and sometimes they are private jokes.
4795	I always attempt to have a file naming convention. However as I work with my colleagues training them in the procedure is a long process. But over time they see the applicability of this method.
4808	Unfortunately, today not all the archaeologists have the cognition of a GIS-based project.
4817	Standards had not been developed yet then. The USGS was just getting its meta-file definitions clear and out.
4881	We are involved in a process of technological change, that will include improvements in standardization of that type.
4882	The GIS used uses a central database with a version managed database system.
4945	Our team attempts to always use a standardised and documented procedure. This breaks down less often than it used to.
4956	Naming conventions are applied to all but temporary files, which are usually named according to the task in hand then deleted. Naming conventions may vary according to the needs of a project, but are always documented. Naming conventions are chosen to convey meaning rather than be arbitrary identifiers (e.g. SH_CV3 represents version 3 of a cumulative visibility analysis for Stonehenge; there will be associated metadata describing each version).
4978	Is really project to project basis.
4980	We are still in the process of deciding upon naming conventions for our GIS files.

4991	We use a variety of different structures which are organic as new structures are added to address different issues.
4992	We don't produce files but data entries in a spatial database. Each object of course has standardised identity.
4994	Crucial data files that are included in the end delivery to the customer will have standard file names, intermediate files used or created during processing not as a general rule.
5055	This will probably be standardized in the near future. We are working on this related to digital field equipment (PDA+GPS+digital camera).
5060	We don't yet have a consistent file naming system, largely because we have not been involved in too many (archaeological) projects. We often use USGS or military file naming conventions for basic coverages like hydrology or hypsography, but they seem inordinately complex and less intuitive than we really need. The result is we name files on a case by case basis and in a descriptive manner.
5185	We developed a standard for archiving and designing the information.
5190	Varies because of the longevity of many projects. New projects are always often carried out using standardized methods and terminologies.
5198	I always TRY and am almost always forced into non-compliance. My solution has been to build audit tools to fix anomalies after the fact.
5202	I say always - but of course I've learned to do this. I didn't always ...
5203	Trying to implement this.
5206	Files to be retained, as opposed to single-use creations, follow a standardised procedure.
5292	The site file manager names files.
5331	Since there is no established or reference protocol I am learning how to improve my system from project to project. Still feel there is a lot of room for improvement.

Question C2: When working with your GIS files, which of the following options do you use to maintain version control, especially if different people are working on the same file?	
4762	Usually only record dates not version.
4771	The Society of Exploration Geophysicists have standardised GIS headers incorporated into SEG-Y data sets.
4784	We are shifting to ISO 19115.
4790	Date of file creation.
4792	No usamos normalmente ningún sistema de standarización ni de control de versiones, pero me parece adecuado tener uno para evitar errores en la manipulación de los datos y por tanto no perder tiempo inútilmente. Estamos trabajando en ello. [[Translation: We do not normally use any standardized system nor version control, but it appears adequate to have one to avoid errors in the manipulation of data and as a result one does not spend time wisely. We are working on this.]]
4795	Still experimenting with this to find the best system for the systems I work with. I have used all of the above.
4817	Hands on Solo effort mostly.

4878	Since I'm working alone on my GIS-based research projects, the problem of different people working on different versions of files doesn't arise.
4882	The database in use has its own version management in place.
4910	But our file are on a server and we can't working together on a same file.
4956	Filenames include a version number which can be related to metadata describing each version and its release.
4961	Generally only one person works with a given project file.
4977	I'm more systematic on some projects than others.
4991	We maintain data sets at multi-levels archiving the primary data sets during data generation and at set stages during the life of a project so that most edits can be traced. This is an area where more work could be done.
4992	Our Archaeological Sites and Information System versions objects through the use of time-stamps (and also resolves update conflicts through checking the time-satmaps).
5048	Whole company uses the same version GIS.
5055	Mostly incorporating version numbers and/or dates. Will probably be changed/standardized soon.
5060	This doesn't often happen here, but I have worked in environments where all files are assigned to a particular cartographer and saved daily (at a minimum) to the network. This sometimes caused problems where there was a miscommunication, but generally worked well. Here, I am the only one doing archaeology GIS work and so all related files would go through me.
5072	[---] system is controlled by computer gurus in the administrative office.
5190	As far as is possible the same file is used and stored on our central server. These used standard names. Where copies are made the documentation associated with each file is recorded with copy date and number.
5198	Audit and renaming functions incorporated into the DBMS front end.
5203	Currently trying to come up with protocol for this.
5220	The file metadata has information on which version is current although file modification dates usually provides the answer.
5232	Where it is not convenient to use the file-name method, then unique identifiers are used and this is cross linked to a RDBS that manages file issuance, auditing and tracking. Same code as used in software development for version tracking and management. Rollback is critical.
5290	Actually we are not using a specific system. But during the elaboration of the final phase of the automation project we will use the standard versioning tool in the core of the ArcGIS software.
5292	Site manager organizes data and makes it available. Site file organization continuously updated, but file name remains constant.
5293	Generally we do not have different people working on the same file at the same time, although different people may edit a file later. The date of the file and standardized metadata documentation are used to track the currentness of a file.
5324	I am the only person responsible for GIS projects running at my institution.
5331	I usually add sequential numbers to successive versions of a same file and also rely on the date of creating of the file. Sometimes I may rename a second version of the file altogether.

5347	I do not create the files in question. Any alterations that I request of the GIS files are conducted by another person within our section of work.
6002	Different people very rarely work on the same file.

Question C3: During the course of creating your digital GIS files, how often do you also create related paper documents?	
4795	Experimenting with this as well to find a balance.
4808	We use to print for an instant study (and surface interpretation) the low range photogrammetry of each cut.
4817	Desk top ink jet had just emerged. We started on a single plotter then rented a multiple large format pen plotter form our AutoCad upgrade distributor who designed and published architectural textbooks for Princeton University in NYC. In most cases we used a color copier which also had just emerged as office equipment. [---], Ph.D. was/is an avid public presenter of archaeological data and has also, I believe, testified about the new technology allowing an efficiency of archaeology in the public never before available to the US Senate.
4903	In the form of metadata that can be printed if necessary.
4910	But I know it must be always.
4924	The paper documents are always stored in the specific project folders, so a link between the paper and digital files is not necessary.
4956	I very often create associated documents (help files, installation instructions, tutorials) but these are rarely disseminated as hard-copy, usually as word docs or PDFs. New data is always accompanied by release notes, again electronically.
4967	Virtually never—only for attribute data relating to archaeological sites, and then only as an emergency backup.
4975	Only if I need to create a report showing what I’ve done for agencies…
4980	Paper documents are usually made for our field archaeologists and in the paper copies of our reports.
4983	I generally create flow-charts in a word processing program, and print these from time to time.
4991	We tend to maintain digital records which can be distributed on paper or PDF. At archive stage all records are exported in basic ASCII structures and as a paper record using simple fonts suitable for OCR recovery should the digital data be compromised.
5048	Usually produce a small meta-data table (very basic) to record the field types used in GIS table.
5060	It’s usually the other way around. We take the paper documentation and enter it into the database/GIS. Since I trust computers about as far as I can throw them (and sometimes I’d really like to throw them) I always keep a digital and hard copy backup of everything I do.
5069	Should do as I have no proper independent record of the state of the GIS project in terms of file versions and descriptions.
5198	As seldom as possible. PDFs are what goes out.
5257	A “GIS01.DOC” file would be filled out for all core GIS data in a project. As archaeologist work with the data they will create derived datasets and are not always as careful to name and document their contents.
5290	Paper documents are usually created during the elaboration of the geodatabase for checking purposes

	and sometimes for map printing etc…. especially if the results have to be presented to people that are not used to the digital files and eventually need to see everything printed on paper.
5292	I download data into statistical application or spreadsheet formats. I usually print a copy of the data file to review.
5293	Workers track their work in a paper log of what was completed when and what remains to be completed.
5331	I print the databases, line maps; do not usually print the colour satellite imagery, unless it is an output needed for a presentation or publication.
6005	Paper prints are created to check on different parameters such as scale, display range, color, or just for working further on drafting or analysis. Paper prints are also used in conjunction with storing attribute tables in project-related binders.

Question C4: How often do you explicitly document the links (by whatever means) between the paper documents and digital files?	
4790	When the 'paper document' is my thesis, the paper is an example or documentation of the GIS process, and is therefore documented. When the 'paper document' is 'work in progress' - e.g. report to supervisor, presentation at a conference, then the relationship is ephemeral and not explicitly documented.
4817	An evolving problem in software. I would today. Back in Intel 386/387 days backup was a problem, mostly to tape with inherent problems. If in architecture title blocks were used (rarely) I would have put the file info there I suppose. Just getting large files to print without crashing was a constant worry, however, due to what was perceived could be output and the ability of the technology to actually perform.
4899	NOT often enough ...
4903	Links are automatically generated.
4937	Draft documents intended for discard are not linked.
4945	Depends on whether paper documents are work in progress or for the archive.
4956	A readme.txt file is placed with digital files listing further resources, including web-links and paper-based resources where appropriate. Generally, associated docs are stored digitally with the digital files and hard-copy is not used.
4978	Our environmental overviews are fully documented with source data locations and granting agencies
4994	When using MapInfo, the workspace name is always printed on the paper map.
5055	Will be standardized. Currently this depends on the nature of the paper documents. Documents describing features in a map are of course referenced carefully (in the document and on the map) to avoid confusion with other maps.
5060	Basically it gets down to that I know where the back-ups are stored. There's no policy dictating how this should be handled.
5198	Insofar as this is a manual task it often doesn't get done. When it can be automated (plotter stamps, automatically generated lists, etc.) I incorporate them into the package.
5202	This is very variable because I don't always create separate paper documentation.
5203	Again, trying to implement a protocol for this.

5220	Paper is good way of documenting on how the data was made to fit into an electronic format.
5257	That is the purpose of the GIS01.DOC; lists file names and described content.
5290	Usually the paper documents produced are to be sent to a third party so I keep the original digital files. And in case the paper documents are to be preserved in our archives a reference will be printed to be able to trace the original digital files that are related.
5292	There is a tendency to overlook this link.
5293	Do you mean making a paper copy of a digital file? We make paper copies of files at project completion as a part of report production only.
5331	I sometimes put a header/footer naming the file or the route to the file; plus the date of the print. This has been easier to do with Word files. Especially if I have to edit database comparing with field notes. But I realize I have not linked them to the GIS project, especially in the cases where the files are in a different directory from the GIS project.

Question C5: When documenting your GIS projects, which of the following aspects do you typically document?	
4790	The methodology for creation of project/datasets.
4817	Not then a requirement in the EPA Region [---] office, or the [---] State Historic Preservation Office, or other agencies worked for.
4956	The database schema and the nature of any glossaries/word-lists/look-up tables in place.
4961	Info recorded as metadata.
4980	Because our GIS system is mainly used in map production, there is little documentation that takes place. If a client requests metadata to be compiled then it is done for each data set. We are also trying to implement a personal log via notes in a notebook to better document when and what changes were made to the project or data sets.
4991	The linkage between paper records and the digital data-set is implicit in the archive management of the paper record which employs the same Key_ID structure as the digital records.
5034	Projection and contact information.
5060	I've done many of these things to some extent, but we do not have a standardised system. Although I am working on it.
5069	Not good at documenting my Ph.D. research GIS project but I would use all the above for a commercial project.
5077	The system is very standard and the only changes are the client supplied data sets which are of course documented (i.e., perhaps AutoCAD base data and therefore line work issues, etc.), and projection. Projection information appears on map outputs.
5198	So far my clients NEVER learn the software so I prepare views or layouts, plot those and give the PDFs names consistent with the view.
5202	This also applies to data that we archive.
5331	At first I lost a lot of information (e.g., could not locate original files) because I did not document at all, relying on memory. Later I resolved it by putting all the images and databases in the same directory as the project, even at the cost of duplicating files in the hard disk. Faced with these questions I realize I have not been systematic at all and could greatly benefit from reference

	protocols.
5347	The GIS projects that I am involved in have been created by another person with my input for its requirements. I typically do not record information concerning the GIS system.

Question C6: When documenting your GIS projects, which of the following approaches do you typically use?	
4978	Our projects are small and individually charged.
5034	I use the ESRI metadata form.
5055	Metadata features (ArcGIS).
5083	Meta-data documentation is contained within our ArcView. ArcInfo and Trimble GPS Pathfinder Office file directories.
5198	I try to keep everything in the project folder and produce a file listing including date stamps of everything there (dir /on/s > [logfile].txt).
5202	All but the second of these have applied under different circumstances.
5290	Meta Data
5292	My activity is generally limited to project specific actions. I use files established by others.
6005	This documentation actually ends up in a binder, or binders with that project name.

Question C7: When modifying the content of your data or GIS files, which of the following factors influence whether you choose to document these changes?	
4780	Not applicable.
4897	When analysis results in completely new data these data will be saved as a separate file.
4945	We would not change the raw data from the field.
4968	I'm not so good about documenting changes…sorry!
4977	I am more conscientious about documenting changes if others besides myself are accessing the data files.
4980	We currently do not document many/any changes, but we are trying to implement a log book where changes would be documented. Once implemented, all changes will be documented.
4982	Time in between changes. Active projects change daily and it's impractical to record. But, if a change is made to completed projects sometime later, due to a design change, etc., changes will be noted in the metadata or project files.
4983	How easy I think it will be to remember what I did! (Not a particularly good criterion, but true.)
4992	The system creates "information history" (in multiple versions) whenever a change is made to an object. This was design decision.
5199	None—my data are usually 'historic' in nature and therefore are not normally modified after creation.
5202	I now only ever work on derivatives and leave the master copies untouched.
5232	All changes MUST be tracked so as to satisfy an absolute rollback capability.

5292	I would not modify the content of the primary file.
5331	I try to leave the original "as is", from then on every edition of the file (correction, addition, amendment) is assigned a different number. A sorting or query of original file sometimes gets a new name with subsequent version numbers).
6005	Changes that affect the output, that is, the result, are the ones that determine whether a file is documented or not. Changes that do not affect the end result or do not substantially contribute to analysis are not documented.

Question C8: Which of the following best characterizes the overall documentation process typically associated with your GIS projects?	
5077	GIS is not an end in itself; it is a component of CRM projects.
5198	We have an "Archival Research" module that is ongoing with components of it linked to a new project. Several "Field Work" and "Artifact" modules where documentation is more or less simultaneous. And then write the report and look up anything that's not already there.
5199	BUT only on selected aspects of the process.
5202	I teach students to keep track as they go.
5290	Usually, database design is done before to correspond with the documentation used during the project. Continuous refining of the records is made during the project.
5331	Field data documents are compiled from several paper media, later recompiled into the necessary databases when creating the GIS project.

Question C9: Which of the following best characterizes your overall documentation procedures (regardless of their actual manifestations) from one GIS project to the next?	
4790	I have only been involved in one GIS project where I have been responsible for overall documentation procedures—in other projects my involvement has usually only been at a high level (concept/funding), and not included details such as details of procedures used.
4795	Due to the nature of some resource management work. Additional documentation is sometimes needed by the client so additional work is sometimes done for that purpose.
4914	Projects may use different data types and amounts. Consistency is ideal but often not exactly practical due to variation from project to project.
4992	Lacking documentation is usually a beginner's mistake—you tend to do better in project 2 than project 1…
5055	They are tuned to the nature of the project: Big—small; map production—analysis; etc.
5198	They start off trying to be consistent, but it seems almost impossible to keep it that way. The best I can do seems to be building on the previous projects which allows some backwards compatibility, but each new project offers more detail.
5203	Hopefully once we have a recognised protocol this will change!
5220	Room for improvement.
5232	As the processes are always open to improvement, it is important to evaluate down and up stream impacts procedural change can and do bring.

5290	Each project needs a specific type of data, but the base core data usually remains the same for most projects.
5331	I strive for a workable system and improve on my previous experience and problems.

Question C10: Which of the following do you consider to be important reasons that actually influence why you document additions and/or modifications to your GIS projects?	
4780	Archival.
4899	While I believe all the above to be important I personally am less systematic about going through each of these steps. BUT filling out this questionnaire is making me seriously consider why I am being so careless in documenting.
4968	All are good reasons—and I know I should, but I do not document…
4982	To document the evolution of the highway project.
4983	I oftentimes keep copies of all (or most or many) changes to each file, so I can look at these for documentation and to "reverse" procedures. This becomes difficult with large projects though (when the files take up great amounts of disk space).
4994	To be able to find back data and still know what it's about.
5055	To keep continuity should others take over a project—important to get them up to speed quickly, as well as to ensure same quality of data/results.
5058	To enable future integration of the data with a duplicate system.
5077	It is necessary to catalogue data error and assumptions because all GIS projects have differing margins of error, dependent on whether data was collected through API, GPS, and so on. Point/line location errors are important to have an awareness of.
5083	We are few users using GIS, and projects are maintained by individuals responsible for the project. The overall database (map data) is not modified.
5110	All of these apply, and more. My overriding concern is about intellectual property rights, to show that I have not breached anyone else's.
5196	Future-proofing.
6005	Modification documentation is a basic principle not only of GIS projects, but of any other research projects in any research discipline.

Question C11: By way of summary to this section, briefly describe what you consider to be sufficient documentation and why.	
4760	[not answered]
4761	I'm more concerned about documentation that allows me to re-create the steps I've taken.
4762	Standard metadata following either the ISO or US standard in both and electronic and paper format. These standardized formats allow for unambiguous usage.
4765	For my work which involves heavy use of scripting and programming I use the scripts and pages of code to act as my documentation. The reports generated by doing describes on the datasets is also quite useful to add in the appendices. A data dictionary is also a must in this case.
4768	[not answered]

4771	Byte location of GIS data in file for later access.
4774	My documentation process is related to personal projects, but I can see the need for dramatic improvement on this if data is to be shared.
4776	[not answered]
4780	Electronic recordation of the parts and significant changes.
4781	[not answered]
4784	ISO 19115 Core Elements - because it's the international standard.
4787	[not answered]
4788	[not answered]
4790	Sufficient for someone else to follow the path of my work.
4791	[not answered]
4792	Creo que en archivo/catálogo de los datos debería comprender un nombre sistematizado, un control de la versión de os datos, un documento acerca de las carácterísticas de los datos, y un historial de los cambios efectuados en el proceso investigativo que lo relacione con otros datos. Se trataría de fabricar metadatos de nuestros datos para que el acceso por cualquier investigador fuera fácil y conociendo a que se enfrenta. [[Translation: I believe that archiving/cataloguing of the data would have to include a systematized name, version control of the data, a document about the characteristics of the data, and a record of the changes made in the investigative process that relates it to other data. This would involve creating metadata of our data so that access by any investigator was easy and they would know what they faced.]]
4793	[not answered]
4795	From a science background the documentation should provide another scientist to recreate my results. All documentation necessary to achieve this is paramount. From an archaeological perspective the idea of keeping changes as a dataset is also a somewhat unique challenge as well. There are many instances where the datasets created in the interim of change are also valid to keep these are a documentation challenge in their own right. We don't just have a raw dataset and the final dataset but you must keep some of the iterations in between.
4796	[not answered]
4808	Field data such as findings, dig annotations and so on.
4809	Labelling storage devices (discs), hardcopies of core analysis maps, notes regarding the databases or files needed (during planning of project), and documentation within the file and on disc of the project structure.
4817	I am not sure, I have not worked in this area for awhile. I would like to have seen more about the documentation of my production, however, under the rubric "proprietary" defending one's position in CRM, that was not the case usually.
4827	[not answered]
4864	The one that best records the development and changes of the GIS system
4865	[not answered]
4866	[not answered]

4877	[not answered]
4878	[not answered]
4879	Paper log of analytical processes and calculations involved in altering or producing new GIS files.
4881	I am increasingly concerned with the issue of documentation, and I realize that I have not been aware enough about it for a long time. I believe that data should be documented up to the point in which anybody could be able to access and understand (de-construct) the process of design, development and storage of any GIS project, either he/she had been somewhat involved in it or not. Documentation should provide the full contextual information that make data meaningful.
4882	Depending on the nature, size, etc. of the project documentation of the data should be done for those things which justify the business case, guarantee consistency/data integrity.
4884	Depending on the scale of GIS-operations and operators; but the main doc should at least refer to project, project leader and dates.
4891	[not answered]
4892	[not answered]
4896	Includes the documentation of the steps used to create the data, so they can be repeated if necessary.
4897	The most recent GIS project is my dissertation and sufficient documentation was sufficient for me. This means that all documentation consisted in finalised vectors, and rasters, covering the basic information layers to work with. These included DEM and residuals, geology and land use, viewsheds and cost distances, least distances, and a number of statistical analyses based on these data.
4898	[not answered]
4899	I believe ALL the steps in the previous question (#17) [i.e., C10] to be important. IF I were to keel over tomorrow ALL the work completed to date would be lost because no one else could possibly pick up where I left off. While I usually work with one other person on GIS-related projects we have a division of labor; sometimes the right hand does not know what the left hand is doing.
4900	[not answered]
4901	All of the software and its options should be documented along with the types of and names of the data used along with its qualities. And also the steps used to come to any conclusions.
4903	Documentation is maintained to FGDC standards for metadata.
4909	[not answered]
4910	Documentation can be very different for each project. The minimum documentation must be describe the procedure of data manipulation and the choice make for data analysis.
4913	Usually just the creation date of an attribute and the data source (i.e., GPS, remote sensing, etc.). Attributes (once finalized) rarely change and it is important to know if the attribute was physically mapped by a diff. corrected GPS or acquired through some other means. The unique ID number of the attribute will correspond to a much larger electronic and paper dataset that will also document administrative data.
4914	List of files, file types, file contents, and changes to data types in basic tables (e.g., text to numeric) that might affect the use of the data.
4919	It depends on the purpose of the project and the data that I’m using.
4924	For our purposes, we generally only document how to perform certain analysis steps. We do not

	typically generate or change datasets, so documentation for a scientific reason is not necessary.
4926	[not answered]
4927	Date of creation. Date of finalization. Date of changes. Attribute dictionary. The attribute dictionary is the most important documentation. I deal with very large files often covering very large areas and the only way for someone to understand the coding of attributes is through basic attribute descriptions (I wish the USGS would realize how important this is).
4933	Identification of software used, for what it was used, versions, and specific file characteristics. Links across files for a relational DBMs and the format of the linking variables. Specific sources of geographical data and its accuracy level. Procedural notes: time spent, who did the data entry/encoding, etc.
4935	I am very bad about this and the emphasis varies. Sufficient documentation would include a digital and paper backup of all changes made, sensible and consistent file/version names and numbers, and all this tied to a well thought-out plan ahead of time.
4937	We generally follow USDA Forest Service protocols
4939	I generally do not use these procedures for two reasons: firstly, the GIS projects I work on/develop are usually quite small scale and simple and therefore do not warrant such documentation. Secondly, GIS is really only the starting point in my work as more often than not I use it to generate information which I then export to alternative programs.
4942	Basic dataset evolution (commands, processes, reasons why), data dictionary.
4945	We are most concerned with raw data management. Documentation of data collection and file location is the prime concern. We keep paper records in addition to digital ones, partly for security and partly for ease of use.
4949	Written record of date, analyses performed and/or variations made, saved version (i.e., file name) & location of saved version on the hard drive. This level of documentation is adequate because I use standard data sources for each project.
4951	A file management/documentation system is sufficient if and only if others not directly involved in the project can examine the system and get a clear understanding of how the GIS project has evolved and its current status.
4952	Sufficient documentation is recording the whole procedure from the initial step of planning a G.I.S project to the more detailed practical alterations until the information on the last steps of presenting the data. Metadata is crucial to preservation and reuse of archaeological record and G.I.S. projects.
4955	Project variable.
4956	Sufficient documentation should first describe the methodology or design rationale then the datasets, their attributes and their relationships (i.e., the data structure), glossaries and controlled vocabularies in use. File naming conventions should also be described. If the interface is customized, instructions for use should be included plus descriptions of any custom functions included. This should always be put into the context of the project to provide background information. This information should be provided as it is the minimum required to work out what was done, for which project and why.
4960	[not answered]
4961	Documentation varies from project to project. Some small projects where the data is in house for a short period do not require as much documentation as multi-year projects. I am learning the hard way that you can not always predict which projects will be short term.
4962	Suff. documentation is information that may be stored on various media to: 1. track the project history; 2. teach an unprepared user; 3. explain to a developer clearly and in details the structure,

	working environment and working procedure of the project.
4967	Sufficient documentation is more documentation than I'm currently employing. It should at the very least document the major changes made to the data, as well as the basics of the GIS system used to store and make the changes.
4968	[not answered]
4969	[not answered]
4973	[not answered]
4975	Written out procedures of how the data was created, write-ups on project and analysis for replication, and metadata for files created or modified for my project.
4977	What I would consider sufficient documentation and what I actually do for documentation are two different things. I should think that some sort of log should be kept of every time (and by whom) a GIS is modified and what was modified.
4978	Depends on what is being done. Full analysis needs to have data source, manipulation and conversion, and analysis to be documented. Map generations and simple spatial analysis requires even less if any at all.
4980	Just like a references cited in a report, what how strongly can the project stand on its own if it has not been properly documented.
4982	Metadata always, including software version, projections, source of data, and project file structure.
4983	I try to document all "significant" changes, and by "significant" I mean those that would not be obvious to another GIS researcher looking at the project and final write-up.
4984	No
4987	[not answered]
4988	[not answered]
4991	Briefly this cannot be done. The objectives are simply why, what and how, the documentation is sufficient if another user can fully understand the structure, purpose and content of the data.
4992	Sufficient for what purpose? This is not a cop-out - documentation is a means to an end not an end in itself. I don't personally see the need to document with such detail that a system can be recreated 50 years from now, only that the information (including object's information history) can still be used. On this though our National Archives feel different... They want the system to be so well documented that it could be recreated 100 years from now.
4994	Document data in such a way that someone else can find his/her way in it.
5003	Included as a section in wider project design: Overall research design describing methodology and justifications - 1-2 pages +. I think there should be a clearly reasoned methodology before beginning the project. Word document: Single line descriptions of individual changes with date and time. This allows checking what has been done and when. Section in final report: Results summary of the project. Describes whether what was planned was achieved and describes the status of the project.
5005	Documentation is a necessary evil. It can be extremely time consuming, and thus, most of us procrastinate. However, we know better, and we should be training our students to keep records as the project evolves in a systematic fashion.
5006	I usually end up with a paper copy and report which could be used to remake the info if necessary.
5007	When released for public distribution, a data set needs to have lengthy documentation including who

	maintains it/created it, descriptions of the attributes, scale of the data, source of the material, etc. If the data is to be used for internal or personal use only, a small paragraph summarizing the necessary information is sufficient.
5010	The majority of my documentation is in data creation. So, it's generally the steps that were taken to create the data that are recorded. I think of this as enough to be able to backtrack or have the process repeated.
5015	List of data used. List of codes used in each data set. List of files and their locations. Brief write-up of procedures used to create new data.
5017	[not answered]
5026	Since my projects are designed and directed for myself and used for my own purposes, I usually do a write up of steps used, but not much else.
5028	Documentation is very essential part.
5034	Filling out the metadata forms and log for keeping track of what has been done and when. The metadata allows a wide range of information about the individual data files to be stored. The log will keep people from repeating tasks and will start a trail for trying to find lost of incorrect data files.
5036	[not answered]
5043	Creation date, creator's name, data type, data extent, previous version. This is the minimum I consider necessary to determine what a file is, what it represents, and where it came from.
5048	Meta-data documentation of fields used and field types used in GIS projects, ensure data entry to consistent standards. There seemed to be limited information around on how to deal with meta-data in GIS projects and what information to include.
5050	[not answered]
5055	Depends what is to be documented.
5058	Document all new additions to data and amendments to existing data. To distinguish all project data from pre-existing data and to allow most efficient extraction procedures.
5060	We typically will create formal metadata using a standard USGS format. Additional documentation is generally for internal or individual use and/or it is used to create the formal metadata.
5061	[not answered]
5069	Ideally, documentation should include descriptions of all data files, versions and structure with a record of processes applied so that the project can be reconstructed from the original source files. This documentation should exist on paper as well as on computer.
5071	[not answered]
5072	I don't understand this question. I think there needs to be hard copy files for all archaeological fieldwork and analysis, regardless of the use of computers, CAD, and GIS. Computers and associated hardware should not be considered archival documentation. Fieldwork should result in a written report, with copies sent to numerous people/offices. The report is an important part of the documentation of fieldwork.
5076	Sufficient documentation is when someone unaffiliated with the project (but an archaeologist) can figure out what is going on.
5077	Date, Projection, Base data source (GPS, drawing program, etc.). This covers all the basic possibilities of error and assumption in GIS creation. If you have indicated that the data set is created from a GPS with x range of error over an aerial photo within a certain season which is or not

	orthorectified and registered to particular spheroid another GIS person will understand the possible assumptions/errors.
5083	At present, our documentation efforts may be inadequate, especially in the face of my forthcoming retirement. The options listed in the survey, including independent written documentation and tracking the records affected by modifications, are presently being contemplated as we upgrade both our GIS and external environmental databases.
5096	I haven't found this issue to be a problem. Files are named in a non-standardized form and stored in the server along with other files associated to the project. Usually the filenames and directory are adequate enough to identify the content.
5110	I document all aspects of the project that I need to recreate the project. The documentation for this is very detailed consisting of steps taken as well as the reasoning behind the steps and what I might want to do different next time and why. The documentation provided to an agency for which a project is developed consists of information about the data, how they were developed (on a general rather than detailed basis) and directions on how to use and maintain the data.
5185	The documentation information is always stored in a database which can be linked to the GIS. By maintaining this database we can record the history of all the changes made within the GIS project. This way this log information can if necessary be linked to the GIS itself.
5188	Full technical spatial and a-spatial documentation (system, software, field definition, table relationships, etc.). Hardware and software summary (office and field equipment). Associated guidelines for users including (non technical?) summary of the above.
5190	Project name, date, creators, data sources, modifications. Also supplemented by additional description of project. Sufficient documentation should also be applied using metadata standards or similar elements. We use Dublin Core.
5193	Dublin core metadata.
5194	Printed map for file record. Not required to provide any other documentation.
5196	Keeping a record of the way the data was created, i.e., sources consulted, individuals creating both database recording and spatial data, plus an overall record for the methodology of the whole project.
5198	Everything you've gathered about a project should be: 1. in one place ; 2. is stored in non-proprietary formats; 3. have a consistent naming scheme (documented); 4. has some sort of intuitive navigation tool packaged with it. (for 3 & 4 we use a tool (carthtml) that automatically builds web pages based on the naming scheme) Also as many people as possible know where it is or have a copy.
5199	Jesus, 'briefly'? I guess in my case it would be sufficient if, when I die unexpectedly, some else who is technically knowledgeable is able to take up my projects on the basis of available documentation.
5201	[not answered]
5202	[not answered]
5203	It should be enough to allow people in the future to be able to understand the processes that went into creating the data, without any input from the original authors, i.e., the history of the data creation process should be transparent (in an ideal world).
5204	[not answered]
5205	Documentation of the creation and significant changes to the dataset; general metadata creation (Dublin Core).
5206	Completion of necessary metadata on digital form. Conforms to national standards, allows other users to view processes, and meets targets for e-government.

5212	[not answered]
5213	Changes made to files, processes used during analysis, location of attribute data.
5214	Since most of the projects I've created using GIS software are unique, I don't use standards for documentation. Documentation about how to use the project and load the files is not required. Most of the projects I've created are easy to load and use.
5219	Sufficient to come back much later and know what was done, or for a competent GIS person with no knowledge to do the same.
5220	[not answered]
5223	You've actually made me realize I don't document as well as I should- One should create a regular file name- document its existence and relation to other files and track any changes made to the file and as a result the project.
5226	Should allow replicability if the application is going to be in the public domain. If private research then more black box is ok within reason provided researchers can "show their work" if challenged.
5228	[not answered]
5229	[not answered]
5231	For example: when updating a cultural resource database, I document which resources were updated (and when) and what items were updated (boundary shape or attribute update) or added. This is entered into a digital document that records all updates.
5232	An old planning maxim is that you can never have enough of it. And the best way to record a plan is to fix it in some physical form, i.e., paper, charts, grids, etc. Each step of the way is then known and as the journey proceeds the journaling function provides the bread crumb trail back in time allowing for very quick and highly efficient change management.
5233	[not answered]
5234	To document the files used to create a GIS system as well as the changes incurring while working on it.
5249	Always be aware of your data and your data quality; be able to share this knowledge with your co-researchers; make sure, that the result of your research can be checked by others by checking the sources of your work.
5256	Short journal of changes made, otherwise too much is spent on documentation with no time left for analysis. Most people are interested in results not the tedious change to datasets. GIS is not an excavation, thus it has not to be treated with the same record keeping mania.
5257	You need to know what the content of derived GIS data files are - if that derivation has proved valuable. It is good to encourage investigation and tentative querying but when a derived data set is valuable and worthy of sharing with other team members then it needs to be properly documented.
5258	[not answered]
5262	[not answered]
5268	I am most interested in documenting the varied information sources from which cartographic and non-cartographic records were/are created.
5290	There is no thing called sufficient documentation. There will be always lacks. You just discover it when you elaborate a specific project that needs specific data and you didn't think of it before. The reason is that each type of research demands specific data and you cannot incorporate everything from the beginning. This is why the core data must be consistent and all new geodatabases can be

	created and developed according to the needs of the each specific project, making sure that it has to be related to the core data using the same standards.
5292	I manipulate data created by others. I generally organize my working files in accordance with the project at hand, and I make an effort to maintain consistent nomenclature to keep track of project development.
5293	[not answered]
5297	Independent files, but have links and access for the rest. Of course, my collaborators are dealing with all this.
5301	I make a list of files belonging to a GIS project and document the changes made to these files.
5308	Documentation should allow another researcher to follow the broad steps taken in a project and participants to go back and correct error when encountered.
5321	Files, Paper Documentation (as a backup).
5323	[not answered]
5324	To my work I think that only link the digital files created with the paper documents is enough.
5330	[not answered]
5331	Have all the data to be used in a digital file which references as to where the original data came from or was edited and elaborated (paper media, field notes, GPS, etc.). To be able to locate all the files used in the Project. To be able to distinguish the succeding versions of the files and how the data was transformed in the evolution of the project in order to understand how results were achieved. This process should really be comprehensible to other operators.
5338	For me to consider a documentation sufficient it is important that it answers to all or almost most of all questions and that the final structure of the documentation will allow the final user to answer at all his questions.
5347	I am not in a position to answer this question as I have not documented or file managed a GIS project to date.
6000	We work in a National Archaeological framework certain functions are decentralised to the various Regional Offices and others are centralised. There are no clear lines of demarcation as to what is done in the Regional offices and what is centralised. Therefore documentation does differ between HQ and the Regions and also within the regions as different individuals work with GIS in different ways. We need to develop a minimum of GIS standardisation and ensure certain documentation changes are recorded centrally and others left to the Regions or individuals involved.
6001	It depends on the number of researchers involved in the project. If there aren't many researchers we don't need to store the documentation about planning, designing and implementing the GIS System. In the other case we decide together a specific documentation. This need concerns the use and the update of the GIS-files in order to avoid mistakes during data-entry and data-exploring.
6002	Sufficient documentation is considered to be the recording of all necessary information to provide a complete overall picture of the G.I.S. project aiming at future improvements in the research procedure.
6003	Metadata automatically generated by software as datasets are created; command lists.
6004	Especially in small projects it is important to document the changes and the workflow to make it visible and changeable also for other people. It is dangerous if this is only in "one man's head".
6005	Data files, project structure, output, procedure are sufficient to document the skeleton of any project.

	Why? Because any project is like a book that reflects one's own ability to organize, relate, analyze, and disseminate research work.

Question D1: When planning, designing and/or implementing your GIS projects, are you concerned about their future transition into an archival environment (e.g., the Archaeology Data Service in the UK) to help ensure that they can be preserved for the long term, re-located and re-used by other researchers in the future?	
4780	All projects have electronic backups and are migrated as needed such that any project can be brought up and used with current software/hardware.
4795	Yes, in theory; however, in the United States the ability to do this is very regionalized and somewhat behind the times. Some regions, due to politics and money, have not thought ahead enough about the digital archiving process while other areas have.
4817	How does one cope with the changing costs of production and the backup legacy of the data? Does the client in CRM cut and run? Usually. I'm not sure if there is an answer for this in the US CRM that I have worked on. Though now they have those new CD ROM's I read about back in 1973 at [---], but barely, the last time I worked in GIS.
4896	No, my data is always (nearly daily) changing so there is no need for archival preservation of the dataset at any one time (other than immediate backups).
4899	At least not up to now. We are beginning to make serious headway into a very large dataset which has not seen systematic analyses and interpretations for a very long time. As such our data has gone through a number of migrations in terms of hardware and software. In fact some of our information remains in outdated versions of DBMS.
4913	Paper data is the archival medium. Curational facilities and regulatory agencies will not accept electronic data in place of paper data, but some will accept it as supplemental. As such, long term data planning is a low priority and is only for internal convenience rather than legal compliance-which also means that it is a low priority for funds.
4919	I am no longer involved in archaeological projects, but work on other GIS projects that are not typically archived.
5007	Currently there is no established long-term curational facility for the projects that I work on. Our firm routinely backs up our projects to CD-Rom, and produces a paper report summarizing our findings for each individual project.
4939	No, because the data I generate is confidential at the request of the communities I work with.
4942	We conform to TimeMap standards and methodology for dataset storage and distribution: http://www.timemap.net
4945	This is a nightmare. We had planned to archive with the CSA. We could not use ADS because we could not finance it. At the moment we back everything up, but what will happen in the long term is very unclear and a very major concern.
4956	The guidelines provided by e.g., ADS, are worth thinking about at the initial stages of a project; some aspects of the project may need more attention as a result and this is best identified as early as possible.
4961	Only concerned that the client can incorporate the data into their system or distribute it.
4980	Currently all of our projects are backed up by copying the project folder with all associated data to CD for storage, both on and off-site.
5010	I'm not aware of any archival projects for GIS data/projects in my area. If I had been, I would have

	attempted to make it transferable.
5058	However, due to project mismanagement the implementation of archival and data sharing procedures has so far been ineffectual and poorly planned.
5060	We provide the data to our clients and what they do with it is generally up to them. We store back-ups, but it is not generally available to the public and there is no policy regarding how long it is to be stored.
5077	This is extremely important when using just point locations for archaeological sites.
5110	I would like the information I create to be part of a larger archival environment. But for now, the data is used in-house.
5188	Archivation is essential but has to be considered against analytical utility.
5219	Usually there are no funds for this in contracting, but it has to be done.
5292	It depends on the nature of the project. The principal file design and implementation took short and long-term preservation practices into account. Smaller scale projects, especially those done as a contract, also consider preservation. Some small-scale exploration projects are not preserved.
5331	There is no equivalent in Argentina, and no general consensus on data archives.
5347	It should be noted however that this has recently been raised with the gentleman who does plan and create the GIS files I manipulate. As a govt. dept. we are required to ensure that items and work is trackable at all times and that any data can be archived, be it electronically or hard copy. Our dept. has a new system in place for electronic archives.

Question D2: How often are your completed GIS projects transferred to a designated repository, such as the Archaeology Data Service, a state, museum or university archives, etc., for long-term preservation?	
4780	Virginia is so backward there's no means of so doing. I hope I live long enough to see a repository formed. I have been advocating such a system since 1982.
4790	But my major project is still work-in-progress, with no secure home to go to.
4792	Pero la administración cultural en [---] no está preparada para recibir este tipo de información y/o proyectos, ni siquiera reclama como opcional la presentación de los resultados en estos formatos GIS. [[Translation: But the cultural agency in [---] is not prepared to receive this type of information and/or projects, not even claiming as an option the presentation of the results in the form of GIS.]]
4795	See my previous answer not all regions are ready to accept digital data. Archiving is addressed at my level to prepare for future submittal.
4881	At the moment preservation of digital archives is a purely internal concern in my institution.
4937	We are our own repository.
5007	There is no such repository in existence in this state.
4942	Datasets and mapspaces to the ECAI clearinghouse: http://www.timemap.net
4945	See previous answer.
4949	Only those projects which have a legislative archival interest are archived. Those projects which only have 'day-to-day' operational interest are stored on CD in my personal collection, but not formally sent to archives.

4978	No such requirements or repositories for Kentucky or the US.
4991	All our data is destined for the ADS, some sets cannot be deposited with ADS as they are covered by other copyright licensing, some sets may be provided with restricted access as they could be used as a source for treasure hunters.
4992	Back-ups are regularly created and stored at two separate locations. However, I feel the best way to preserve information is to let it live and be used rather than be locked up in an archive.
5055	Depends on the nature of the project. If working with known data already in the repository, only new data found during work will be transferred from project to rep.
5058	To National Monuments Record and local SMR.
5190	At present there is no system designed for the long-term storage of digital data per se. This is of course something that is being discussed at the moment. However, our data is housed on our sever and is regimed by regular back-ups and data migration/refreshment when new GIS software or versions are used.
5198	But based on the kind of questions I get, I don't think they are ever integrated into an overall GIS.
5201	It's too expensive to do this- unfortunately.
5202	Most of *my* projects are now about teaching so they fall out with the scope of the ADS collecting policy.
5206	Transfer to ADS currently under discussion, but this is part of the development of a national procedure.
5257	The Museum of [---] (for whom I work) is the archiving body for the [---] region.
5292	If only because we are the repository as well as the user.
5324	In my country this kind of repository does not exist.
5347	The projects that I work on are part of a Government Department and should therefore go towards the archival system for our department - RTA Archives, from there they should eventually go to State Records NSW.

Question D3: How often are your completed GIS projects saved for the long term 'in-house' (i.e., stored and maintained by you or another project member, rather than being transferred to a designated repository)?	
4795	This is an issue that I am currently tackling. Designing a system to keep certain types of information for in-house use and ignoring other types of data.
4899	Projects are usually in the hands of their creator.
4983	Always archived "in-house", and often elsewhere as well.
5006	We save them and use them again, but there's not much maintenance.
5007	Our GIS projects are routinely backed up onto CD-Rom at either the completion of the project, or at the end of the calendar year (dependent on how large the project is and how much space it consumes on the computer).
5198	AS WELL AS being transferred to a designated repository.
5257	Even when deposited they are retained in house for reference and comparison. The archaeology of London is like one big jigsaw puzzle.

5331	Kept on Hard disk and CD back-up by myself.

Question D4: When saving your GIS projects 'in-house,' which of the following long-term preservation strategies do you use?	
4774	Backups on CD.
4795	Here again some archaeologists in the US are not used to thinking about this. So often times it is an afterthought and they don't want to deal with the costs.
4796	Put them on CD and hope I can access them in the future.
4961	We have not been faced with this issue.
4992	I can't answer for our entire "house"—it's a big organisation. The entire database and its metadata can be exported in XML-format—that's part of the design. How these exports are stored and dealt with in the long-term I don't know (and is not my primary responsibility but our Archives' Unit).
5034	I also make several back ups on both a second removable hard drive and CD-ROM.
5055	Either kept on harddisk if it's a project that will be used again shortly/at intervals, or backed up to CD/DVD for storage. These are usually double copies, and checked/reburned after a certain amount of time.
5196	Finished data is copied onto CD and onto a shared network that has data retrieval facilities. Also paper copies are produced where possible.
5198	Migration has not been an issue because we don't store in proprietary formats. (let the next guy do the conversions).
5203	No protocol has been established yet.
5257	All server data is backed up on daily/weekly/monthly and yearly cycles and is stored off site.
5331	Kept on HD and back up on CD. No paper output kept, just the prints of the working versions of the databases and final outputs of graphics.

Question D5: For those long-term preservation strategies that you use when saving your GIS projects 'in-house,' which of the following would you say best characterizes the overall process?	
4961	All data transferred from server to CD at end of project.
4983	I plan to implement more systematic long-term archival practices, but haven't yet (and am not sure exactly how to go about it).
4984	I'm really bored by this survey—you have given me a survey that is too long and any answer from here on, including this question is bogus!
4992	Our Chief of Archives might respond differently ;)
5055	This will be standardized in the near future.
5190	This is only a recent practice phenomena, and at present no real back-log dpp have been carried out except for migration.
5198	So far new hardware has always been required long before even the most conservative estimates of data stability so each new server gets ever larger disks and the whole lot is copied over.
5331	Digital files are always preserved & backed-up, a paper version is not usually printed or kept

	together with the digital files.

Question D6: When saving your GIS projects for the long term (either 'in-house' or in a designated repository), how often are the saved files organized the same way as when you were using them (e.g., using the same folder structure, etc.)?	
4780	Files are abstracted, copied and placed into groups for ease of reference.
4795	I try and establish a consistent method but often takes a period of time as I introduce a company to a way of working that they have never thought about and there are some discussion and fine tuning. Once this is worked out then it is always. I try and take the view to make the organization by a general categories of process that can be applied to a wide variety projects so that you don't have to have different formats from project to project
4808	It depends if we create a WEB-GIS or not.
4899	Regardless as to where they are stored they are usually organized in a form that is consistent, e.g., in terms of file names, etc.
5055	This depends on where they end up. Repository data go into a central database, local projects are usually stored as files.
5198	Along with their automatically generated web pages that provide navigation.
5292	I would have to consult with the Curator of Information Technology to answer this question, and it is my hope that he is also completing this survey.
5331	This is the way used to preserve or reconstruct the process of documentation & elaboration of the project.

Question D7: Which of the following considerations prevent you from saving some or all of your GIS projects for the long term (whether 'in-house' or in a designated repository)?	
4790	I am talking about my PHD work here ... Day job work is G.
4795	In general a whole understanding and attitude needs to be addressed. Technology as a whole is often dealt with as any other tool has been dealt with in the past. They try and use it best in their already determined way of working. Often this is not the most efficient use of the tool or can also take longer for the individuals than their old way. This is why you get people often complaining about the headaches they get from technology. It truly requires that people rethink how to best do their work using the new tools of technology available to us to achieve our ultimate goals.
4919	I back up my projects to multiple computers and the corporate IT archive system.
4982	California doesn't see it as necessary as a whole. It's only important on a personal level, but no policy mandates proper storage.
5034	I would think that disk storage space, whether hard drive or CD-ROM, would be a big issue. The life span of some storage devices also is a thing that needs to be thought about.
5055	We currently lack a local repository, but have access to a national repository. This might change. Funding is perhaps the most crucial factor here as sufficient funds would help with B and C too.
5203	Majority of personnel don't consider it important (lack of knowledge rather than willingness to do it).
5223	I haven't focused carefully enough on this- I maintain paper records related to our site database, but no archives exist for the digital media.

5292	Information management staff limited and under considerable demand. Preservation of site file information is routine, but other projects are not given the same attention or priority.
5331	I am the only person in charge of the GIS project (from data entry to design & analysis) with no specific funding assigned on a regular basis. Time dedicated to the GIS project must be combined with other duties (teaching & research). Work is carried out exclusively at my home computer because of lack of proper facilities at work place. Archival procedures at my work place are now being implemented and are mainly geared to conservation of objects, not data archive at all. The absence of protocols (national or work place standards) and easy access to my own data allows me to postpone some tasks (e.g., complete printing of databbses and graphics- expensive in view of recent devaluation) and rely on more inexpensive archival methods (CD).
6005	Long-term storage of GIS data is still viewed in some European institutions as unnecessary and sometimes unwanted.

Question D8: When recording metadata to document your GIS project files (or groups of files), which of the following types of information do you routinely include?	
4792	Como he mencionado con anterioridad estamos trabajando en esto. En estos momentos no poseemos muchos metadatos, pero los aspectos indicados arriba están contemplados en nuestra idea. [[Translation: As I have mentioned before we are working on this. Sometimes we do not have much metadata, but the aspects indicated above are contemplated in our idea.]]
4817	USGS had just “invented” it (ca. 1991?). Sol Tax at the US Bureau of Land Management was sort of a public affairs officer for the data and is missed. Informally I followed the metadata discussion and once held a copy of its first public issuance.
5006	My organization is supposed to set up a common metadata, but it hasn’t been finalized yet. Currently, metadata is sporadic.
4942	Using TimeMap metadata creation tool (TMWin) or inbuilt XML metadata functionality of ArcGIS.
4956	Also methodology where appropriate (e.g. type of interpolation used).
4984	Why should i spend 30 minutes of my time answering this survey?
4991	Our data is archived in ASCII formats such as dBase III or other portable formats wherever possible to allow the data to be re-used in alternate systems with the minimum of bother.
4992	Note that in the system we use it’s data entries that are created not files. Hardware, OS and Software is irrelevant since data entered need to comply to a data standard and this data standard can be produced by a variety of software on a variety of platforms.
5048	Scale of Mapping; GIS Types (Point, Polygon, Line)
5055	Projections, information on the kind of data in the project, as well as precision of measurements.
5058	Metadata recorded as part of SMR record
5077	All basic data is recorded in a workspace file or, if I am using ArcView, this is easy enough to record. Other information is on map layout.
5202	All of these are important, but the critical Intellectual Property issue has to be documented too. That’s the source and / or supplier of derived data, and the conditions under which it is supplied.
5206	Copyright; digitisation scale.
5292	Applicable to some, but not all projects. Especially applicable to the development of GIS data for other agencies under contract.

5331	Also source of original data (is this different from “file format”?).

Question D9: When recording metadata, how often do you follow/use an established descriptive or other metadata schema or standard (e.g., Dublin Core, USMARC, FGDC, etc.)?	
4919	FGDC or ISO standards, but rarely complete.
4956	Make use of the NGDF Metadata standard, and the NGDF Metadata Explorer (MS Access97 based application) for its maintenance.
4991	Only in part.
5048	Occasionally—we tend to make up our own metadata standards, which is more simplistic but suited to our own requirements. No-one in this organisation knows how to deal with the more complicated Dublin Core standards—would be nice to have proper training on metadata management though.
5077	If doing academic work, I use Dublin Core.
5190	Again depends on the date of projects. New projects always, others only if revisited; no systematic back-log dpp has been carried out—this is something that will happen though.

Question D10: Which of the following metadata standards do you use, or have you used, when recording metadata for your GIS projects?	
4942	TimeMap & ArcGIS (although I think ArcGIS xml metadata may conform in theory to one of the above?). [Note: By default, ArcCatalog automatically records properties of the selected item in both FGDC and ISO format. TimeMap has adopted the Dublin Core standard for descriptive metadata. Participant’s original “Other” response has been changed accordingly.]
5015	MIRIS, metadata files for Michigan.
5048	Use in-house metadata.
5058	I think.
5077	For routine work on short-term projects, this extensive metadata documentation is not cost effective.
5083	I’m not sure which standards we use.
5190	Use an Icelandic standard based on a Danish one; Landlýsing http://www.lmi.is/landlysing/. As well as our own developed from DC [Dublin Core]. We are looking at CIDOC CRM as well as MIDAS map depiction standard.
5193	Museum of [---] archive guidelines.
5290	Our metadata core is designed for our specific data and use. It combines lots of the components present in the above mentioned standards.

Question D11: When recording metadata, do you typically embed the metadata within the resource(s) being described, or do you store the metadata in a separate file/database, and provide a link to the described resource(s)?	
5060	I believe metadata should always accompany the GIS data layers. You shouldn’t have to search for it.
5206	Linking, rather than embedding, is current corporate policy.

Question D12: When recording metadata for your GIS projects, how often are the metadata terms that are used derived from a controlled vocabulary (i.e., a limited set of consistently used and carefully defined terms, usually in the form of a thesaurus)?	
4991	The terms change as new terms have to be added, the thesaurus is in part an expanding table that requires flexibility.
5055	This depends on the project, but we try to use an established vocabulary to ease deposition in national repository.
5202	Where appropriate. In many cases there are no controlled vocabularies.

Question D13: By way of summary to this section, briefly describe which elements, outputs (e.g., original vs. final reports, maps, tabular data), etc. of GIS projects you think should be preserved for future use or reference and why.	
4760	[not answered]
4761	[not answered]
4762	For financial and space constraint reasons I believe only final reports, maps and data sources should be saved for future research and reference. Repositories only have limited resources to handle such data and the data created in various steps along the way are usually not as important as the final product, as long as the development procedure is well documented repetition should be possible.
4765	Everything should be saved if possible and archived onto CD. Many projects need to be revisited to update information or to correct errors.
4768	[not answered]
4771	Correlation of surface GIS data with underwater site data is somewhat problematic. Standard SEG-Y format incorporates GIS data but I am working on the linking of site visualisation with spatial data.
4774	Tabular data.
4776	[not answered]
4780	All of the text files, all of the graphics and all of the data and all of the photographic media plus whatever else went into the process.
4781	[not answered]
4784	All base maps and raw data + derivatives.
4787	Lower level original data.
4788	[not answered]
4790	It depends on the project.
4791	[not answered]
4792	El contenido de los datos, su área geográfica, la calidad espacial y espectral, la escala, el formato del archivo, la banda cronológica de los datos, la fecha de fabricación y/o actualización, y los proyectos en los que ha participado. [[Translation: The content of the data, your geographic area, the spatial and spectral quality, the scale, the storage format, the chronological span of the data, the date of creation and/or actualization, and the associated project.]]
4793	All data are valuable (maps, tabular data, digitized forms in different projection systems, hardcopy

	maps).
4795	It depends on the purpose of a GIS. If its purpose is cartographic then keeping report products would be useful. If the GIS is for management and or analytical purposes then keeping the data is most important.
4796	All data sets, maps, etc., to not only reproduce final output but to re-use data in the future.
4808	All data.
4809	final reports, maps and tabular data should be preserved, if they pertain to cultural and historical resources. The priorities of each organization or business that runs a GIS will make their individual priorities relevant as well.
4817	Depends on the project...Although I was informed that surveyors in my state (NY) or maybe elsewhere in the US are not permitted to use pencils when recording survey info as it can be erased. All errors must be struck through with a pen (or pencil?) so that any errors of recording can be seen in court. I imagine we might have all the data preserved, but officially, what is necessary? For example, the West Point Foundry, is the heritage of the US, across the Hudson River from the Academy, today very separate, once very integral to the history and training of its cadets, where has the government been? I think noticeably absent but perhaps not. Nike missile batteries were made there and the cadmium cleaned up, though the site is now in quasi-private hands, part of "Scenic Hudson, Inc." Does the data become public record of the Academy or the current owners?
4827	I think all material must be preserved.
4864	The original data layers and some of the more used maps and final reports.
4865	[not answered]
4866	[not answered]
4877	[not answered]
4878	Ideally, all digital data produced in GIS projects should be preserved and made available for future researchers, although copyright issues would probably prevent this from ever happening.
4879	Method of file creation, software used, software version used, source of primary data, purpose of file creation.
4881	[not answered]
4882	[not answered]
4884	It depends strongly on the nature of the project; some projects are set up only to come to a final (paper) publication. In that case, the final maps, complete pre-press processed should be stored, the GIS project is less important. GIS data etc. of such projects should always be stored. In continuing research the project would be saved with good descriptions of structure and data.
4891	[not answered]
4892	[not answered]
4896	Steps to creating the data and project.
4897	I think all should be preserved, basic spatial data such as digitised geology maps and land use maps as well as georeferenced rasters, site locations, for future use, and any kind of spatial analysis should also be preserved so future research can build upon and verify or even correct past analyses.
4898	[not answered]

4899	I think all input and output of projects should be preserved in paper and digital form (with backups in several different locations). Too often large projects are funded for their initial retrieval and analytical phases with no eye towards the end product and its storage. To not be concerned about the production and accessibility of the product is irresponsible.
4900	[not answered]
4901	If possible, all data used, and the steps followed to reach the end results should be preserved along with the final results and their implications. That way future researchers will be able to recreate the results.
4903	All aspects should be preserved so that reconstruction of any phase of the project can be accomplished.
4909	[not answered]
4910	[not answered]
4913	Final reports and maps for all projects; tabular data for projects with collected artifacts.
4914	[not answered]
4919	I think that original files, tabular data, maps, and final reports should be archived.
4924	Most of our GIS projects are saved for future use. We save anything that may be helpful with other projects and save time. We save everything associated with an important project: maps, files, and databases.
4926	[not answered]
4927	The final product, or most recent version should be stored in a permanent fashion. This allows future users to know the last steps used in the project. Our project boundaries change often, so keeping a running backup would overload our storage capacity very quickly.
4933	Hardcopies of all final maps, printouts of data tables, and analytical results. Digital data cannot be made archival per se—so digital migration is only a marginally acceptable option. Multiple copies of printed documentation are the only way to insure long-term (ca. 1 century) usability of the essential data.
4935	[not answered]
4937	The data sets used to create the project need to be preserved so that future workers will know what was available at the time the project was created so as to be better able to evaluate the project.
4939	I see the benefits of storing all types of data generated by specific projects, provided that the files have adequate documentation and are all stored under the same project heading (or sub-heading).
4942	ALL—for reasons of transparency and replicability—through some kind of versioning/repository system if possible (although in praxtice who really has the time/funds for this).
4945	Original data from the field. Processed GPS and geophysical data. Future studies, both in the field and in the office, will be able to make use of the raw data. The processed data would generally be what a future researcher in archaeology would use unless there was a very specific research aim (e.g., in developing new software for geophysical analysis). Further levels of processing and output for publications (graphics) are preserved in digital form, but their long-term preservation is of less concern.
4949	In addition to the final output (e.g., report, electronic or hardcopy mapping, etc.), all of the data and individual data sets used during the course of the project are stored to DAT and to CD-R. Copies of these data (usually on CD) are physically stored with the official (departmental) copy of the final

	product, and a duplicate set is held in personal storage. In addition, care is taken to ensure that at least 2 complete copies of the final output and that data used for that output are stored in 2 distinct locations (such as different buildings). Complete data sets for important projects are stored in at least 3 or 4 distinct physical locations.
4951	Original data is critical, rather than just maps, etc.
4952	All of the aforementioned elements seem important to me to be included in the metadata section of a project's report.
4955	Tabular data: the only way others can really use and assess the GIS. Maps: as above, plus the results of the research. Interim and final reports: record of procedure, interpretations, reasoning. Links (where present): enhances database.
4956	By storing raw data and functions (e.g., map definitions, query definitions, etc.), it is possible to reanalyse at a later date although this rarely happens. Certainly, all final reports and maps should be preserved digitally for future use and where possible data and functions. Interim reports which contain information pertinent to the project itself undocumented elsewhere should also be saved.
4960	[not answered]
4961	[not answered]
4962	[not answered]
4967	The maps, tabular data and final reports should definitely be preserved. Additionally, some record of the algorithms used to produce the analyses (i.e., the specific map algebraic expressions used, etc.) should be retained.
4968	[not answered]
4969	[not answered]
4973	[not answered]
4975	All files with metadata should be saved so that future researchers can either re-use the data or add it to other projects.
4977	The final reports should, of course, be preserved. Ideally the data themselves should be preserved in a format that could be accessed and used by other researchers for future analyses.
4978	Adherence to the open GIS file structuring.
4980	Any and all data possible should be preserved. We do not know what other researchers will be looking for in the future so it is best to save all of the data possible for potential use at a later time. This also allows other researchers to see the thought process or steps taken to create the final product.
4982	Final reports and raw data.
4983	Ideally, ALL steps, procedures, files, and rationale should be saved. For many reasons (listed in the last few questions), this is not always practical. I think the ideas behind the manipulations and information (what you did and why) are, in most cases, more important than the exact manipulations performed.
4984	[not answered]
4987	[not answered]
4988	[not answered]

4991	Both primary data as generated at source and the final resulting data set with a clear definition of originated and derived data.
4992	Object data, object history, metadata. It's the information that is of primary importance not the system in itself.
4994	Original input data, end product data, intermediate files depending on type of action performed on data.
5003	A working copy of the project or the ability to recreate a working copy of the GIS should be archived. This is so it can be shown that the system worked, but also so that the data can be reanalyzed in future. A summary of all the relevant documents should also be available in e.g., PDF, that do not require that the GIS be used: The final report the GIS contributed to, specific maps/views, and tables etc. that are referenced in the report.
5005	This is a hard question to answer. It depends on the nature of the archive how that archive will deal with questions of persistence and migration. Are the data to be obtained online, or in hard copy only? Obviously, we'd like to ensure migration, but some models can become very complex–who will do that? So we many be forced to maintain static copies of final results.
5006	Maps, tabular data, original data used to generate final reports
5007	Final reports, maps and tabular data should be preserved. The preserved elements should consist of all the elements necessary to re-create the GIS data or the project. I view this the same way as I view scientific experiments. Sufficient description and information needs to be preserved so that some other person at a later date can re-create the project exactly.
5010	The original digital data should definitely be preserved. The maps and reports can be recreated from the data and the documentation providing the steps of analysis.
5015	I store it all.
5017	[not answered]
5026	Digital data files and methodology write-ups, maps, etc., can be recreated.
5028	[not answered]
5034	I personally believe that a copy in each of the formats mentioned should be used including the electronic format. The electronic format should be saved so that others can go through and add information to the GIS or be able to use it for further research. The paper copies should be kept because we know the preservation quality of paper as opposed to the preservation life of CD-ROMs or other disk storage devices.
5036	[not answered]
5043	Final reports and any data used to bolster the conclusions within that report should be saved, as a minimum. If no claim is made on the data, there is no need to back up data that is not supporting any research. However, any data that are basic to archaeology should of course be preserved (e.g., a site catalog in GIS).
5048	A combination of final reports and original GIS data.
5050	[not answered]
5055	As much as possible, of course again depending on the nature of the project.
5058	All original elements should be recorded, particularly if a GIS project was not designed to be modularised.

5060	All original data should be saved and available in the event a discrepancy is found. This information should be saved in-house. Obviously final reports and data layers should be preserved and available to those who need them.
5061	[not answered]
5069	I sway from minimal storage, so original source data and detailed records of processes to storing intermediate data sets with final reports, etc. The latter is more useful for researchers but takes more storage but also may lead to cumulative research errors where people interpret final processed data as real maps.
5071	[not answered]
5072	Final reports, including site records and maps.
5076	[not answered]
5077	Depends what you are doing: 1) working in planning then you need maps; 2) working on stats then you need tables; 3) working with dems then you need documentation of the algorithm chosen.
5083	Metadata is described within the file directories for GIS-related projects.
5096	We lack an overall strategy for long term preservation of electronic data. Personally I try to preserve everything that I imagine will be used again, but it's not an obligation.
5110	[not answered]
5185	In the first place maps and tabular data, with a reference to the original reports.
5188	All raw data and supporting documentation for reconstruction of the project. Appropriate derivatives and syntheses if in 'grey' literature.
5190	Project name, date created, by whom, data source, update, file format, software used to generate and a short description of the resource.
5193	Primarily data that has been created 'in house' and not necessarily that obtained from national mapping bodies.
5194	Final maps as these usually have some legal or auditing significance.
5196	The original data (maps and tabular) should be preserved, so that others can work with it, rather than using data that you have created from it. Although manipulated data and resulting reports should also be preserved.
5198	First we perform an audit to ensure that file and object naming is consistent throughout the project. We reduce everything to the simplest non-proprietary file format (TXT JPG TIF DBF) supplemented with proprietary formats like PDF DWF for which there are public domain readers. We turn the whole lot into a set of HTML pages: http://[---] http://[---] This makes an archive that is likely to survive but with limited functionality. To this we add our working files (just in case the next guy has compatible software). We then put the whole thing on CDs to be distributed to whoever needs them and copy the source for these CDs to the Archive area of our server.
5199	Again this is too wide-open to answer, but certainly it is important to preserve information about the (analog) source of the data, and the method(s) that were used to digitise them.
5201	Maps and tabular data, final reports. This minimises staff time dealing with enquiries.
5202	The underlying data is critical as are the final reports which make the case and context for the data. Other outputs can be preserved as required: so long as they can be recreated.
5203	Enough data to make the creation process transparent, e.g., original data, and final reports, along

	with major interim components. Not sure it's important to keep every version of a file. Ditto map data- the components necessary for the final version, but not parts that weren't used. If tabular data was combined into a database, then only keep the database and not the original tables (i.e., no repetition/duplication of data).
5204	[not answered]
5205	Primary data. Any data (analysis results, etc.) which could be generated from the primary data does not necessarily need to be stored.
5206	[not answered]
5212	[not answered]
5213	The entire project.
5214	From my perspective, anything that can be used as archival evidence of the destruction caused by excavation at a particular site is important to keep a record of. This includes maps and any spatial information that's been generated that may be of use to researchers who come in to examine a site later or are excavated an adjacent site.
5219	It really depends, but any academic project that is peer reviewed or any contract project that could be questioned, involved in legal dispute, etc. must be fully archived. This should include all personal workbooks, data, reports, etc.
5220	[not answered]
5223	Final Reports, maps and tabular data.
5226	Ideally final reports and tabular data would be kept. No real need to keep plots or drafts except "in house".
5228	Databases and maps.
5229	[not answered]
5231	The digital data files should be saved, along with description of the data and an electronic map if possible.
5232	The hidden value of GIS is the benefit to future generations who will be seeking to compare their world to that of another time. By incorporating not only the data sets, but functional copies of the deployment or reader software and sufficient information as to the computing environment and system used originally to create the data then value-of-use is assured into the uncertain future. As the cost of the storage medium decreases along with extensions of the preservation quality of those mediums, then it would be remiss to exclude. On occasion I include JPEG photos of the locations, team and crew as part of the data backup as a way of providing a human link between the dryness of the content and the humans involved in its creation, plus it's a bit of fun for future potential users to find human perceivable data as pictures amongst the rest of the bits and bytes.
5233	[not answered]
5234	Tabular data, maps, and final report.
5249	[not answered]
5256	Last version, final report with meta-data- no earlier versions or earlier changes, just confuses matter otherwise.
5257	Raw data must be maintained from which outputs can be re-created. The report which the GIS maps will accompany should provide the bulk of methodological statement if this is out of the ordinary.

	That is if a report shows GIS layers showing an erosion monitoring programme then the report to the client must explain the interpolation algorithms used for example so that they know how the surfaces and volumes were derived.
5258	[not answered]
5262	[not answered]
5268	Ideally, it would be very useful to preserve elements and outputs from all stages of a particular project in order that others might have the potential to evaluate the conclusions reached by that project.
5290	All the elements related to a specific project should be preserved and stored together on a single data storage type/system labeled with the relevant information.
5292	The archaeological site file is routinely preserved. Information generated during contractual projects is also routinely preserved. Other, individual research activities, are not consistently preserved.
5293	[not answered]
5297	Final reports.
5301	[not answered]
5308	Original tabular data, final covers and reports.
5321	Original and final reports, raw data, maps, everything must be preserved.
5323	[not answered]
5324	I think that original reports, added with original maps and tabular data is information enough to preservation.
5330	[not answered]
5331	Certainly original vs. final reports, maps, tabular data, original location of the files & data, people who intervened in collection, entry & processing. Based on these questions a final index could be achived if a log file were kept for the analysis sessions.
5338	I think it is always important to preserve the original of all documents and maps and other type of documentation of a GIS project in order to have always at any moment the possibility to check original documentation of the project.
5347	I think both the actual outputs—like maps, etc., that have gone through the final checking and review process should be kept, but I also feel that the base data that went into the GIS system should be kept with the final output for future access and review if necessary.
6000	Tabular data, maps and final reports preserved as a reference. Since we deal here with development projects and their impact on archaeological sites.
6001	Projection system; differences between digital and analogical data storing the tools used to import data within GIS environment (scanned maps, digital maps acquired by means EDM, GPS, traditional survey, photogrammetry, etc.); scale of each map before digital conversion; alphanumeric data type (i.e., row data, analytical or synthetic data, etc.); metadata about archeological records (finds, objects, structures, monuments, etc.).
6002	Final reports, conclusions and maps, formally published, are important to be preserved and used for future reference. A geographic information system is regarded as a simple tool in the research procedure, with secondary importance after the publication of the results.
6003	All outputs and base data should be preserved (paper will last longer than data unless migration

	procedures are followed).
6004	In my opinion every element should be preserved which is important to understand the workflow. It should be possible to recreate the results and that it's possible to go on with the research with another team.
6005	This question relates to the answer to Q C7. All files, no matter what kind, that change the final results ought be preserved. Why? They represent the records of the thinking process, and of the methodological development within that particular GIS project.

Question E1: Overall, how often do you follow a routine or systematic sequence of procedures or steps when generating the various components and outputs (e.g., data files, algorithms, views/maps, statistical regressions, etc.) of your GIS projects?	
4780	Needs vary from project to project requiring flexible response, so if it works, it's added to the process.
4792	Cuando se trata de representar nuestras bases de datos para generar mapas de representación: siempre. Si se trata del proceso de investigación: ocasionalmente y dependiendo del tipo de análisis. [[Translation: When one tries to represent our database to create representational maps: always. If one deals with the process of investigation: occasionally, and it depends on the type of analysis.]]
4795	Changing the company workflow sometime prevents this from happening.
4808	We use a simple routine that may vary depending by archaeologist chief request.
4817	The nature of the past CRM was totally all over the place and parts of one were done alongside parts of another all the time.
4937	Each output depends on the analytical questions being asked—that in turn identifies the requirements for the output and how it should be created.
4945	For GPS or geophysics the routines are standardised. For scanned paper records (e.g., section drawings) and digital photographs (e.g., for rectification) they are not.
4956	Map production is very systematic and I use templates to automate this. Much of my analytical work is also highly structured/procedural and as such is scripted so that I provide the input and the computer follows the routine to provide me with output. Producing views on the other hand is rarely structured as views are created for specific purposes.
4983	Depends on the number of files and type of analysis—if there are numerous files which much undergo similar manipulations, I draw a flow chart and follow it for each one (or better yet, set up a batch process when possible).
4992	Backups are created through a standard procedure.
5055	This depends heavily on the component, as some are more "rule bound" than others. For instance: there are not too many ways to go about creating a TIN [triangulated irregular network], but an overview map can be created in several ways (sequence of data layers added, when do you add north arrow, etc.).
5077	Depends if base data is vector or raster or created from microstation/autocad/or a GPS. Depends if base data is in a surveying grid and requires reprojection/offsets. However, all projects have the same goal of overlaying cultural and natural heritage information with proposed developments.
5199	A routine procedure is followed in creating GIS data (usually at the digitising table) and in post-processing them, but other procedures tend to be much more variable and depend on the demand of the moment.

5290	It depends sometimes I create different procedures to test the best fit.
5292	One person is responsible for updating and maintaining the archaeological site file.
5331	Since I consider that I am still at the learning stage, this depends on whether it is a procedure I carry out more often. Still, I realize the need for a small informal log laying out the steps, errors and incorporating new alternatives or definite procedures. Difficult to implement at the same time and I should to it at the end of a work session.
5347	Not applicable at this time.

Question E2: Whether routine/systematic or not, how often are the procedures you follow when generating the various components and outputs (e.g., data files, algorithms, views/maps, statistical regressions, etc.) of your GIS projects based on, derived from, or guided by procedures outlined in a GIS procedures manual?	
4790	Where I have been able to find a procedures manual, I have used it.
4795	But I am in charge of creating the Procedures manual so it is based on memo currently that I put out for people.
4808	Sometimes we have to insert external data such as geochemical, palynologic, and so on.
4817	Questions of scale are involved. My projects ranged from ½ acre to 20,000 so the relevancy of procedures along with the applicability of the procedure and its defense came to mind quite often.
4896	We do not have a GIS procedures manual.
4937	Our manual does not specify how outputs are to be generated.
4983	I don’t have a formal GIS procedures manual, and most of the analysis I do is non-standard enough that there probably aren’t any (?)
4991	Derived outputs are effectively defined by the functionality of the system in use, there is no specific “do this or that” manual rather we approach it from the “if you want to do this then…” direction coupled with an awareness of “these are the things you can do.”
4992	All providers of information to the system I work with are “licensed” by the [---] Board.
5010	Not really sure what you mean by “procedures manual.” I’ve had classes on basic GIS data entry/analysis, so I probably follow from what I was taught without being totally conscious of it.
5055	Usually they are. But this again depends on the nature of the project/component.
5110	If the projects are the same or very close to the same, I use a general manual (more like the documentation) created from previous projects. The majority of my smaller projects are pretty much the same and are pretty simple so I follow the procedures I keep in my head. Maybe it’s job security!!!
5292	I must defer to the site file manager.
5297	It depends upon the running of the project and my collaboration.
5331	Still feel dependant on these and course notes. They are also usually clear, if not I add comments or notes on the margins.
5347	I do not create GIS projects directly therefore this question is not applicable to me.

Question F1: Briefly describe/define what "accuracy" means to you with respect to your GIS projects and/or their individual data files.	
4760	[not answered]
4761	[not answered]
4762	Accuracy is a very ambiguous term with a loaded meaning in geospatial analysis. I would define it as a quantifiable variable of the degree to which the data is correct or matches the real world.
4765	"Accuracy" would mean how well a given data model represents a real world object.
4768	90%
4771	Accuracy is as important as resolution at this phase of the project. Its data processing requirements require microsecond resolution.
4774	As accurate as I can get it.
4776	[not answered]
4780	The ability to recreate on the ground what I have recorded and used in the visual displays.
4781	[not answered]
4784	Two separate meanings: (1) spatial accuracy as measured by e.g., RMSE [root mean square estimate]; (2) accuracy of attribute data.
4787	Accuracy depends on data type (i.e., format for numeric data and number of class for discrete data).
4788	[not answered]
4790	A measure of the reliability of data.
4791	[not answered]
4792	Siempre tratamos de trabajar con datos dela máxima precisión y exactitud, y procuramos generar información de alta calidad, pero a veces es imposible trabajar con buenas fuentes de información lo que hace que los resultados no puedan ser muy precisos. De todas formas dedicamos una parte importante de nuestro tiempo a mejorar la calidad de los datos de partida y nos esforzamos mucho en obtener resultados finales precisos. [[Translation: We always try to work with data of the highest precision and accuracy, and we attempt to generate information of that quality, but sometimes it is impossible to work with good sources of information which leads to results that are not very precise. Of all forms, we dedicate an important part of our time to improving the quality of the departure data and we try hard to obtain precise final results.]]
4793	Accuracy depends on the type of scale and the available information.
4795	Accuracy means different things for me. Having worked in the land surveying profession the spatial accuracy of the datasets is paramount to me. The techniques of data collection and processing are critical in maintaining that accuracy. From a database standpoint attribute accuracy is just as important. Your data is only as good as your worst data. You may be spatially accurate but if it is miss-attributed than your accuracy is terrible.
4796	[not answered]
4808	High.
4809	[not answered]

4817	AutoCad's stated precision is/was unequaled, based on the core of its software routines. I have tried other CAD programs that choked, but would be good for many purposes. Close range photogrammetry requires precision of accuracy. The "accuracy" of given procedures is often not actually known without research. Often historic maps are not scalable, historic B&W photos blurry or taken at great distance, and "super accuracy" of the present a hindrance to depiction. Many stones are arranged irregularly, with irregular form, computer software replicates identical objects the best. Between the two "extremes" is the archaeologist usually.
4827	[not answered]
4864	The measure of the spatial input error (meters, cms, mms).
4865	[not answered]
4866	[not answered]
4877	[not answered]
4878	[not answered]
4879	Degree to which data can be said to relate to 'ground truth', as opposed to 'precision' which relates to level of recording.
4881	I understand accuracy as more an internal and contextual question than as an absolute one. I mean, to my view accuracy within a GIS project do not refer so much to the reliability of the representation of the real world through the digital data (the relation between data and reality) but prior to the question that every data generated and used in a project should be similarly precise and adequate to the selected scale of analysis (the relation between data themselves).
4882	[not answered]
4884	Accuracy is the way a project and its project data matches the question and need of the researcher who wanted the project. E.g., when sketches are asked for to get a feeling of the material the data and map accuracy is, and should be, very different than when printable high-res maps are asked for.
4891	[not answered]
4892	Accuracy is a measure of how close an observation/data item is to the "true" value.
4896	Accuracy is the degree to which the data adheres to the actual situation.
4897	Accuracy is the pixel size of the raster, the scale of the digitized map, the detail of knowledge on the archaeological record (intensity of survey, knowledge on chronological detail).
4898	[not answered]
4899	Scale of input/output, consistency and confidence level.
4900	[not answered]
4901	I guess accuracy would include the combined total of the data used (its individual components) and the results and how they are reported. That is to say, the data itself and the documentation that follows.
4903	How well the spatial data corresponds to the real world.
4909	[not answered]
4910	[not answered]
4913	Defined datum and a data precision of approx. 1 meter.

4914	[not answered]
4919	Source data collection spatial scale and attribute accuracy.
4924	Accuracy for my purposes, varies from map to map. I look at it as the closest I need to be to the true spatial location for the purposes of the map.
4926	[not answered]
4927	Accuracy, in my work, can best be described as reliability of location based on ground truthing. I deal with point locations more than anything else, and sites recorded before GPS are always problematic, so our measure of accuracy changes through time as more of these old sites are revisited and their coordinates updated.
4933	Accuracy means a fair reflection of the reliability of the data at the projection level in use. In most cases the accuracy can and should include positional reliability or repeatability (in terms of % error or meters of circle of confusion), and accurately (100%) transcribing names, dates, counts, masses, etc.
4935	[not answered]
4937	Accuracy means the precision with which the data were collected, entered, and manipulated - the level of accuracy of the output can not be greater than the accuracy of the input.
4939	As an archaeologist I take this term to relate to the data accuracy in terms of attribute location, description, etc. If this is what the question relates to then accuracy is essential.
4942	Acceptable levels of accuracy defined generally by the nature of the problem and the scale of the data; largely a subjective thing that needs to be carefully considered and justified.
4945	For our specific project accuracy means the ability to combine different types of map (GPS, rectified aerial photos, geophysical, topographic, trench plans, etc.) so that the same features appear in the same place.
4949	'Accuracy' is based primarily on the use of standard or standardised data sets usually developed and held by Government Departments or Agencies, which are notarised, authorised and archived to Government or departmental standards. After that, 'accuracy' is based on concise records of the particular data manipulations and/or migrations undertaken; records of the data analysis; and finally digitally archiving copies of the original data and any new data developed during the course of the project.
4951	This is a highly problematic issue, as often I don't have any control over the original collection of the data, and hence am often wary of its purported accuracy.
4952	Accuracy includes both the meaning of careful design and implementation and that of detailed description of the whole procedure.
4955	Spatial accuracy re: data (e.g., mapping to within a few metres maximum - as good as you could get doing by hand on the ground - all layers should be of comparable level); numerical accuracy re: data; accuracy re: presentation & inclusion, or access to, all relevant data (e.g., through links).
4956	Accuracy in this context means a true representation of a situation, so a land-use theme should accurately represent land-use. Furthermore, errors can be quantified in order to infer a degree of accuracy -- e.g., the land use theme is only 50% accurate as half the values are wrong.
4960	[not answered]
4961	How closely the GIS data represents the real-world.
4962	[not answered]

4967	In my case, mainly how accurately the GIS data layers mirror the 'true' situation on the ground, for example in accuracy of site position, or how closely the terrain model approximates the true ground surface.
4968	[not answered]
4969	[not answered]
4973	[not answered]
4975	Accuracy means that the data is an attempt to get as close to reality in modeling a certain spatial phenomena...
4977	[not answered]
4978	Depends on the Scale...Accuracy on a site excavation of 50 meters has a different meaning than for regional scale accuracy with sites over 50 kilometers.
4980	Accuracy for us means that we know where we were working and relies heavily upon GPS as well.
4982	Accuracy is very important! Its not worth creating a project if it's not as accurate as possible.
4983	I define accurate information as that which corresponds well with what it should (e.g., elevations in a DEM accurately reflect actual elevations).
4984	[not answered]
4987	[not answered]
4988	[not answered]
4991	3D spatial accuracy is of key importance as is the separation between fact and interpretation. Accuracy is a product of a combination of source data scale and digitising scale, so for field drawings for instance any drawn scale of 1:100 or less the data is maintained at millimetre scale for scales above data is maintained at 1 cm precision. 3D locations of all finds for instance are maintained at cm resolution, all referenced with the Ordnance Survey grid of the UK, else where we tie in to the UTM co-ordinates. For derived data sets the accuracy is usually implicit in the output file or its supporting metadata file.
4992	We speak of quality rather than accuracy - accuracy is only one aspect of quality.
4994	Correct topology and databases without errors are my main concerns.
5003	[not answered]
5005	Accuracy can mean close fidelity to spatial location. It can also mean data sets with minimal error and clear documentation as to how they arrived in their current (or final state).
5006	[not answered]
5007	Accuracy means that the data presents what it promises within specified constraints, and that it is consistent. Accuracy is not precision, and precision is not accuracy.
5010	It is the degree of confidence with which I have in the data files that they have the "correct" data.
5015	The degree in which the estimated differs from what is true.
5017	[not answered]
5026	Accuracy refers to the "truth" of the data.
5028	[not answered]

5034	Accuracy is the ability to recreate what you are representing as precisely as possible. While it is impossible to get it 100% it is possible to take precautions to cut down on the total loss in accuracy.
5036	[not answered]
5043	Quantitatively: keeping my data in a format that allows the GIS representation of archaeological items to reflect the real item to a degree necessary for a particular purpose. These should usually be kept in comparable units whenever possible. Quantitatively: keeping my data in a format that allows the GIS representation of archaeological items to reflect the real item to a degree necessary for a particular purpose. This involves utilizing a series of nested categories into which items can be placed to a fine level of detail, but also abstracted as necessary.
5048	Accuracy in GIS terms would probably refer to the accuracy of a given grid reference.
5050	[not answered]
5055	Quality in planning and implementation - resulting in accurate data/outputs.
5058	Accuracy to me means creating GIS data with as much relevant data as is appropriate to the scale and proposed application of the data. For example, ascribing an 8 figure grid reference to a poorly provenanced fieldwalking findspot is not necessarily accurate.
5060	Accuracy refers to the level to which the data reflects the real world phenomenon that it represents. The level of accuracy should be known, or estimated, and reported so that someone using the data knows what they've got.
5061	Accuracy is dependent upon the conditions under which the data was originally gathered and should reflect that level. For example, to give UTM coordinates at a 1-meter level of accuracy to a site that has been identified with boundaries determined by pacing is ludicrous, but often done.
5069	Accuracy refers to original data sets in their location data and also resolution accuracy with grid processing. Analysing different sets of data with varying resolutions and accuracy is an issue.
5071	[not answered]
5072	Accuracy is a measure of the variability in the data recorded.
5076	[not answered]
5077	A defined level of error combined with reporting/documentation.
5083	Most of our work eventually is reviewed by the [---] and [---] State Historic Preservation Offices. Our results get recorded on their 1:24000 USGS maps. In the field, we use resource-grade GPS to log datum points, photo points, and sometimes archaeological features (cairns, housepits, etc.). Therefore, accuracy for us means +/- 2 meters or less.
5096	Input accuracy (map scale, method), record-id.
5110	Accuracy is very important. In fact, for site locations I create an entry so that I can rate the reliability of the site location.
5185	Project-based, the guidelines for maintaining GIS-data are always followed in a consistent way. Standardization is very important.
5188	Accuracy is a subjective term. I interpret accuracy as an aspiration to create a consistent dataset that is aligned to both the recording system and the analytical outcomes.
5190	Accuracy is different from precision. Accuracy depends on the positional accuracy of an object; i.e., that it is a statement of the true representation of the spatial object to the real object in the real world.
5193	The extent to which data reflects the 'real world'.

5194	Accuracy is essential as the final map depictions often have a legal significance.
5196	Accuracy with regards the spatial data, the reliability of a grid reference, etc. Also the accuracy of sources used to create the data must be considered.
5198	Accuracy to me means geometric accuracy and we have to be satisfied with as good as we can get under the circumstances and an indication, whenever possible of the metric value of the discrepancies. For instance, most of our measurement are taken by survey the location of which is established by referencing benchmarks. We assume our “accuracy” to be the distances between benchmarks from different survey stations (usually sub centimeter in 3D space). To this survey data we add rectified photographs (for which we store residuals) which are inserted into the CAD drawing and we consider our “accuracy” for this step to be the distance from the point measured to the point on the image in (2D space). This is typically less than 2cm but we don’t document this separately because there are an awful lot of points and anyone viewing the image can make these assessments for themselves.
5199	This is a measure of the difference between the stated and the real. We only use it to indicate the quality of certain (point and vector) GIS layers.
5201	Consistent, double checked by separate persons for line and shape accuracy to original source or map data as well as for relational information.
5202	Accuracy is the degree to which something is correct within a declared precision.
5203	The documentation should be a truthful record of the data. There should be an indication of precision in (for example) geographical positional data, and in georeferencing.
5204	[not answered]
5205	Correct location, correct contents.
5206	Data checked for consistency; captured at appropriate scale; major development stages recorded.
5212	[not answered]
5213	Means that everything can be repeated and the same value obtained.
5214	If we are using maps generated by survey data, we try and maintain the same level of accuracy in anything that is associated with that data.
5219	Accuracy is how ‘correct’ a feature or data point is, in terms of spatial, categorical, temporal, etc. In metadata it would mean how correct are the information describing the data.
5220	[not answered]
5223	Accuracy refers to the truthfulness of the reported data, as related to its “reality” in the field.
5226	Not sure I understand the question. Usually would think of this as validity and reliability of a georeference....
5228	[not answered]
5229	[not answered]
5231	For all of our archaeological GIS projects we record the resources using high precision DGPS. We always state what the accuracy is.
5232	Reducing the variables that effect raw data increases the accuracy potential of the overall project. The more absolute the raw materials going into the system, then higher confidence can be held in the results as output.

5233	[not answered]
5234	By accuracy, I understand the lack of errors on: database entries, digitizing errors, and analysis results.
5249	[not answered]
5256	Intrinsic property of dataset which I have never time to estimate due to lack of funding.
5257	Accuracy is the measure of whether things are in the right place, while precision is the degree of control over accuracy that the various data capture and manipulation techniques allow.
5258	Those data being put into the system should be complete.
5262	[not answered]
5268	With respect to cartographic records, accuracy involves adherence to a given set of standards generally expressed in terms of RMS [root mean square] error. For non-cartographic records, accuracy for me means insuring that the data contained in those records is correct in so far as this is possible (with uncertain information indicated).
5290	I understand from your question that you're asking about the geographic spatial accuracy. Well for that the accuracy depends from the needs of the project. And sometimes we are constrained by the accuracy of the data we get even if the project demands higher accuracy. In this case we try to do our own manipulations to reach our goals.
5292	In development, we used quality assurance/quality control procedures by enlisted others to review content. When users of the data raise questions, we address them to ensure accuracy.
5293	[not answered]
5297	Accuracy means the possible best way of generalisation and use in making the policies.
5301	[not answered]
5308	Accuracy to me entails assessing the reliability of the different data sets used and matching the analytical procedures to this frame of reference, finally implementing the procedure systematically and producing a result which is based on sound procedures.
5321	[not answered]
5323	[not answered]
5324	[not answered]
5330	[not answered]
5331	How close to the original data I am, or to record the basis of any ammendments or reinterpretations of the same. That these be comprehensible to an outsider.
5338	For me "accuracy" is a very important term, in fact as an archaeologist working with GIS project I saw in several projects that from a different accuracy in data recording and structure depends a lot of final results and analyses. For example, if I consider in a GIS project to record the length of a road represented as a line or if I decided to represent it as closed polyline it changes a lot my analysis possibilities. In fact, I could ask just the length measure in the first case but also the width in the second one, and so on.
5347	For example an easy check of the co-ordinates of a particular heritage site in one of 3 different co-ordinate types can be entered into the GIS system and their plotted spatial location can be verified against where they should be identified as being placed. Accuracy is an integral part of the GIS system that I need to use for site management purposes and for site relocation purposes. While a

	system is only as good as the person inputting the data, checks can be established to confirm the accuracy or reliability of the input data and should be initiated at regular intervals.
6000	As near to exact geographic location of archaeological sites and features.
6001	Accuracy means knowing exactly scale and projection systems of each map and the tools adopted for converting spatial data in digital format. So is possible reduce the errors overlapping different map acquired in different scale. Accuracy means being able to manage the errors for avoiding misunderstandings when analyzing data. Different scale and different projection systems can increase errors and consequently isn't possible have a good data validation.
6002	Accuracy means information as brief, as precise and as comprehensive as possible. Accuracy level varies from project to project, as it depends and has to be in accordance with the aim of the research.
6003	Positional accuracy of spatial data.
6004	Accuracy depends on the scale of the GIS project and all data files should be of similar quality for all calculations.
6005	A. is the degree of quality of reflection of true information data.

Question F2: How often are your GIS data files (whether created 'in-house' or imported) formally or systematically audited for accuracy, either by yourself or anyone else?	
4762	Statewide data is distributed to contractors and governmental agencies in the region who do examine it for accuracy and inform of us of errors.
4771	As the data sets are imported, its accuracy can only be assumed.
4780	Checks are made as the project progresses. Naval Intelligence motto applies: In God we trust, everyone else we verify.
4792	No tenemos un sistema de auditoría externa, pero personalmente me encargo de comprobar la exactitud y validez de los datos antes de trabajar com ellos y en sus resultados finales. [[Translation: We do not have an external auditing system, but personally I take it upon myself to verify the accuracy and validity of the data before working with it and in the final results.]]
4817	Newbie data newbie programs newbie operator.
4913	About 2-4 times a year, specific data files are reviewed in-house.
4924	Most of our maps are for use in the field. GPS points are used to check the data.
4949	All imported data sources are formally and systematically audited before use. Data sources developed 'in-house' are archived at the completion of each project and are not reused.
4956	Whenever new datasets are rec'd/generated, they are checked against the source data to ensure accuracy relative to published sources. They can also be checked against other data sources if necessary/appropriate.
4967	They are double-checked by myself for errors, especially in the location of sites.
4983	I do most of my work alone, and don't formally or systematically check files nearly as often as I know I should. This is largely (but not entirely) a result of the limited time I have.
4992	Our primary quality control lies in the training of our users. Their actual data entries are then sampled and feedback provided to avoid mistakes made to occur again.
5010	Have created GIS projects with which I am not associated with anymore, so I'm not sure if they have been checked for accuracy.

5048	Occasionally—depending on the funding and time needed to undertake proper audits, which is very rare.
5077	Not sure where this question is going...if there is a major error it is obvious and all minor “errors” are part of the whole data management problems with differing sources/projections/etc.
5185	There is an audit per sub-project, an audit for every batch of external information when delivered, and a periodic audit for the internal data (varying from a few months to a year).
5194	In theory they should be available for auditing but this has not happened thus far.
5198	Never that I am aware of.
5219	For all ‘real world’ major projects, ALL data undergo QA/QC by a separate individual at each step of the process. This is all documented on paper and the logs are kept. For small academic or student projects this is sometimes not done.
5290	Usually when we delegate work to a third party. We always check for the accuracy at the reception of the final products.
5292	One person is responsible for entry and management. Many people use the data. The manager audits data submitted for entry, but we do not routinely audit the entered data for accuracy.
5331	At the data entry level I usually enter the information “as is”. Usually with a print out I then go over to insure that there are no typos, misspellings, or errors. I then create another file with a corrected or edited version, in the case of reinterpretation of some data.
5347	The actual data files are verified periodically—both with external auditors for quality procedures within the unit as well as internally for verification that the input data is not corrupted or erroneous.
6005	It depends on the task. If say, only one area is required to work on but the general map shows a whole region, then that region (whole) is not checked for errors.

Question F3: How often are measures taken to ensure that when you share your GIS projects with other researchers (other than those directly involved with your GIS projects) or the general public, you (or your research group) are identified as the creator of the GIS project (e.g., through the use of logos, researcher/institution names, etc.)?	
4780	Standing offers have been made.
4792	En algunas fases del proyecto debemos exponer nuestros resultados o hipótesis sin logos y casi de manera anónima. [[Translation: In some phases of the project we have to expose our results or hypotheses without logos and almost in an anonymous fashion.]]
4817	I’m not sure. I have been told a map I did of different outlines of shorelines of [---] in [---] from the 17th century to the present is quite popular among the visitors of “[---]” an historical archaeology museum at [---], administered by the [---] Museum, but I’m not sure if my name or the company I worked on it for are mentioned. I am not sure the map makers are either, though somewhere that data is recorded. It was developed to educate a public, not win a debate over the current lack of graphics permitted in Masters and Doctorates granted at the University level in [---]. Besides, Herman Melville lived at one time next door.
4945	Output is available on our web site to which copyright rules apply—at least in theory.
4956	Contact details are always supplied as part of the metadata supplied to external researchers (as readme.txt files), but these are not part of the datasets and can easily be separated. Logos are used where there is formal documentation supplied. Generally we supply data rather than share projects.

4978	Depends on who funds the project as they are the final owners of the data and have the full rights to it.
4992	The presence sites or monuments on a piece of property can have big consequences (financial or in terms of what you are allowed to do) which makes it important that information derives from a government agency.
5048	We have a copyright field
5055	Except for intermediate test results, etc. Mostly this is templated.
5188	Don't consider this important and it is embedded in the dataset metadata.
5190	Again something that is being developed, particularly with the advent of more sophisticated development for web GIS.
5198	Most of our sharing is via the automatically generated web pages where a name and email address is tacked onto each page.
5202	Though the amount of sharing as against using I do makes this largely academic.
5206	Affected by copyright issues.
5290	When our GIS project is transmitted officially we always put our signature on it.
5292	The Inventory of Illinois Archaeological Sites is maintained by the Illinois State Museum in collaboration with the Illinois Historic Preservation Agency.
5330	Not yet.
5331	Put my name or initials on the final map. Put my name on some files. That is all I have done so far.
5347	As a commercial unit within a government agency it is required that we acknowledge the source of our data/GIS project when we provide it to others or members of the public to view or use.

Question F4: Which of the following measures are used, or have been used, to restrict access to, or otherwise protect your GIS projects (whether archived or still in active use) and their underlying data from, unauthorized access and/or modification?	
4787	Don't take into account the A-E answers.
4790	Not applicable: I don't allow any access (at the moment).
4879	My data is not shared as yet.
4897	I am not sure, some of our GIS projects have been published on the net, web-based GIS, in which data is read-only. It is possible that coordinates of archaeological sites are encrypted for safety reasons against illegal excavations.
4945	Access is restricted to prevent Research Students using the systems for their own purposes (e.g., downloading music from the web), and in an attempt to get all team members to use standard methods of logging, file naming, etc. There is [no] "security" issue as such—i.e., we do not think anyone wants to steal our data.
4956	Generally, GIS data resides on a local drive with a public access version and a backup version on the network; anyone with GIS can access the public network version (read-only) from across the corporate WAN.
4967	My data and project files are stored both on a laptop computer and on the University's file server, both within password-protected accounts.

4983	I'm more than happy to let anyone look at or work with what I have!
4991	Different parts of the data sets have to be released in different ways and the level of data control maintained changes during the data collection-processing-publication stage.
4992	All answers applicable to the Archaeological Sites and Information System. Other systems might work differently.
5060	I know that some of these measures are built into our file system at the network level, but I am not sure how they work.
5077	Limited software availability prevents other end users access.
5198	During the project the data is restricted, once completed I assume it's public domain (though my partners and clients do not always agree).
5219	Again, it varies by project. Virtually all 'real' projects that are done under contract have high security.
5290	As an official Directorate working for the Ministry of Culture. Security of data is one of the most important priorities.
5292	In house security is robust. However, we provide data to other agencies with instructions that they secure the use of these data. We assume that they do so because we have a common interest. But we have not audited their procedures.
5308	When web enabled.
5331	All data & projects are on my home computer to which only I have access. Otherwise I have only added my name to files or images.

Question F5: How confident are you that your GIS data (whether from active and/or archived GIS projects) have never been tampered with or corrupted over time in a way that would reduce the value of those data for future use or make it impossible to reproduce your results?	
4762	The data itself is inherently error prone though we continue to correct data whether they be error from the original paper records or from errors to the digital data once it was entered.
4780	Murphy's Law dictates that problems will occur.
4910	But our GIS projects are design to answer at some questions in a clearly defined research project, so our databases are not necessarily useful for another project.
4945	I take this to mean deliberate tampering. We have had mistakes.
4951	High staff turnover in heritage agencies means that it is impossible to keep track of everyone who might have 'fiddled' with the data over the years - it is a worrying thought, and often leads me to unwillingly adopt a 'This is mine and no one else can touch it' attitude.
4955	Strongly confident re. tampering; not so confident re. file corruption.
4956	All datasets are made read-only and users are instructed never to edit the core GIS datasets although they can create new datasets. Once archived, the data should be stable.
4978	With digital media there is always the chance for problems.
4991	With such a large data-set there is always the chance that parts become damaged during use. By maintaining an archive audit trail we can reconstitute the data set at most stages of its creation.
5048	I have my own back-ups (i.e., the originals). Public and other staff can only access a copy of the GIS

	data (read-only), and corruption problems are easily resolved by re-instating a new copy from my original.
5083	While the data have not been tampered with or corrupted, un-noticed modification is a problem. For example, if our analyses are based on coverages that are subsequently modified by the GIS group, and if they do not advise us of the modification, our analyses can be invalidated. This problem is currently being addressed in the course of upgrading both the GIS and environmental databases.
5190	Date of use and if any modification is checked when used. Otherwise no stringent system is in place.
5199	Depends what you mean - the greatest danger for me is that I cannot reproduce results because I forgot to document all the steps I used the first time around…
5202	Because I base my work on data held with the ADS so it's their job to look after it.
5203	I am confident that the data has not been deliberately tampered with, but have no indication of whether it has been corrupted accidentally over time, partly because none of our GIS projects are very old. The only check on the health of the files that we have at present is when they are reopened as part of the same or different projects.
5206	I'm currently correcting legacy issues from switching GIS systems, since the old system could not export attribute data.
5290	Continuous check have to be done to ensure this matter. But I think concerning the media and hardware one have to rely on good quality brand names and trust the factory guaranties for their products.
5292	As far as the in-house primary record is concerned. I cannot address the accuracy of those who use copies of the data.
5331	I have not been consistent in logging the procedure and databases used.

Question G1: Please add anything else that you think might be useful for us to know about your GIS record-keeping activities or experiences.	
4780	Central repositories at the SHPO level ought to be mandated with backups at the Federal level for all projects. Data should be shared and made usable with appropriate attribution.
4792	Me gustaría participar en este proyecto porque creo que es muy importante luchar por que la calidad de los datos con los que trabajamos sea la mejor posible. Además, en nuestra labor cotidiana nos encontramos con múltiples trabas en cuanto a que en nuestro ámbito de trabajo e investigación la escasez de datos digitales y de calidad es muy alta lo que nos obliga a dedicar mucho tiempo en fabricarlos personalmente y en localizar las fuentes primarias para llevar a cabo esta tarea, lo cual ya de por si es bastante difícil, y a veces inaccesible. [[Translation: I would like to participate in this project because I believe that it is very important to fight so that the quality of the data with which we work is the best possible. In addition, in our daily work we encounter many cases, in our field of work and investigation, where there is a very high shortage of quality digital data which requires us to dedicate a lot of time for producing such data personally and locating the primary sources to carry out this task, which often is quite difficult, and at times inaccessible.]]
4795	I am working to address these issues as well as broadening the use of GIS as a multi-tool in Contract archaeology. At first many saw GIS as the next step after CAD in report production. Now we are getting people to realize it is also a tool to help managers manage a project or plan a project or create new projects. This ultimately requires a whole new paradigm in how they approach their work to best implement this new tool. A part of that implementation it is always critical to plan and consider the issues raised in this survey. I look forward to seeing the results of this survey.

4796	There needs to be an established media format for archival storage of all digital data (not just GIS). Such a format would have unchanging hardware and file formats with the anticipation of retrieval beyond 100 yrs.
4808	Good work!!!
4817	[---] College has an architecture and archaeology Institute that proposed standard layer names for archaeology that would allow comparative analysis and documentation that would at least be standardized. The convention, however would have to be adopted and I'm not sure AutoCad is universally used though the layering may be a cross-software attribute. Years ago I attended a conference "[---]" at Princeton University in New Jersey, at Forbes College. The college, (after Malcolm Forbes of "Forbes" magazine) the only facility there with a cafeteria, apparently "food clubs" are the rule, had been outfitted the basement with the latest IBM XT's networked together. I was surprised by the number of different database designs, some ("Animals") allowing the electronic input of various data with an electronic caliper, which metaphorically, is what GIS is.
4881	At the moment, GIS experience in my institution has been primarily an individual concern for a reduced number of researchers, working on an individual basis. We are currently involved in a process of shifting the use of GIS technologies as a daily tool for everyone here. Within this process we are being forced to modify a number of procedures and routines, namely those related to the integrity and preservation of data, documentation and so on. This is why we are very much interested in these kind of questions right now, so this questionnaire has come in a critical moment for us. Perhaps within a year or so some questions would be answered in a different way, but it has been very valuable for me to think about these questions again.
4897	I have been involved in just a few, but big GIS projects so far. One of these projects has a strong CRM factor involved, but I was much more participating as a data filler rather than as designer at this project. The second larger project was my PhD, which has a strong interpretational function, and consisted basically in testing hypotheses and creating hypotheses on site location.
4899	Your questions have provided much for me to consider ....
4919	As I have gained experience with GIS I have become more certain of the need for data documentation/metadata. It needs to be a part of the basic GIS curriculum, as well as the archaeological GIS curriculum. The Forest Service, at least in the western states, has a system to append project specific data files to an overall corporate file. The system is not used regularly at this moment, however this is one of the aspects of my job--streamlining the appending process and making it accessible to others. There is also the problem of different software systems. Almost everyone I work with has yet to migrate from ArcView 3.3 to ArcGIS 8.3. I use ArcGIS 8.x, however there are still things that are easier or quicker to do in ArcView. This remains a problem.
4942	Standards like those specified by the ADS in the U would be good to adhere to even for the purpose of local GIS project data storage, but again, in practise this is very difficult to achieve what with availability of funds and the large number of novice users (i.e. students) working with and producing data in our (educational) context.
4945	The major headache is keeping up with developments in hardware and software where compatibility, or lack of it, is the major issue.
4949	I have found that one of the key issues for GIS record keeping is keeping the data sources current and useable through technological and software changes. The current strategy (which might not be the best) is to keep a digital copy of the original data and software together (sometimes also archiving the essential hardware as well); and secondly to periodically update (i.e. migrate) the data to 'new' or current software environments.
4980	Much more time and funds need to be directed into record-keeping activities than are currently given. It is similar to conducting fieldwork, clients are more than willing to pay for the time necessary to complete the fieldwork, but are less than willing to pay for the time necessary to write

	the report!
4982	[not answered]
4983	Thank you for your interest in these issues -- I know they are important and I'm glad someone is trying to understand how they actually play out in the real world. I would welcome (and I know many other people would as well) any comments or suggestions you could publish along these lines!
4984	Yes = I'm just completing an on-line survey for another source. My survey has 24 questions, but takes less than 10 minutes. Your survey is too detailed and asks the user to think too hard and long. While I realize that you are trying to gather very important information if you have problems with survey reliability you'll know why. This is especially true for those of us who aren't in the records business, but who are gathering a lot of GIS data. Hopefully I'm just having a bad day.
5005	My experience in the field suggests we have much to do to educate ourselves and our colleagues in proper digital preservation. GIS is a particularly difficult arena for preservation, because much of what it produces are complex visualizations or models that are often difficult to migrate into other contexts. It's often difficult to move a model from one version of the software to another depending on the complexity of that model. We simply have to do better, and to create a conservation ethic in the field for ALL digital products.
5034	I have found that the lack of electronic copies of archaeological information and logs can throw an unforeseen wrench in a project by necessitating extra database compiling. While more and more archaeologists use some sort of electronic database, not all do.
5048	Seems to be a lack of senior management involvement in both the development of GIS projects and recognising the benefits of using GIS systems. Most staff have none or little training in the use of GIS and meta-data, and often have to train themselves. The price of training in GIS still remains expensive.
5060	The cultural resource management division is relatively new and has a less well developed file/data management procedures in place than is the case with other divisions (i.e. natural resources) in our organization. The data needs and types are different enough that we need our own system. We are working on standardizing our archaeological database and GIS procedures.
5072	Can't live with it, can't live without it.
5077	GIS are used in a variety of ways. Certain GIS packages like ArcInfo/ArcView are more geared toward the information management/metadata than for example MapInfo. MapInfo, however, is far more user friendly. A lot of base metadata is recorded within a file and separate data entry (as per ArcView) is both time consuming/not cost effective and often irrelevant. I suspect that this survey with receive a variety of responses because of the varying uses of GIS from mapping to modeling and documentation procedures differing so greatly dependent on these varying output needs. Good luck with the survey; it is an important concern.
5083	This is a very well-constructed survey. Your questions give us several ideas about what needs to be included in our upgrade efforts.
5110	I love GIS and believe it is finally bringing archaeology and analysis into the 21st century.
5198	I have come to the conclusion that standards are there to be converted, not adhered to. The pressure of field work is such that I have to be grateful for anything entered into the system and all our software is tuned to make this step as easy (and non-intimidating) as possible. If things are not quite right the audit finds most of it and if we need to integrate data from another project it will never fit perfectly so conversion tools are necessary.
5199	There were never any procedures in place wherever I started my GIS projects, so I was always left to devise my own procedures. Since GIS data are usually shared on a personal basis, I would expect (and have) to coach any new GIS users in the use of my data, until they can work independently.

5202	Have you spoken to the ADS about this? I guess they would be interested in your results / questions.
5203	Our organisation has only had a widely used GIS system for two years or so and is still developing protocols for documentation of data and future-proofing of all kinds. The concepts outlined in this questionnaire are not widely considered by more than a very small number of staff, and therefore there is no organisation-wide use of data documentation. We aim to get some protocols in place but personally I think this will take a long time to percolate through to all members of staff, who in general don't see the importance (or problems) of data storage. Procedures will need to be built into projects at the planning level in order to be taken on board.
5219	This is a very important area, and I look forward to your results. This needs to be included in all academic programs, so that people learn the right way to do things. It is, in reality, much easier to teach this stuff than to do it! It is hard and takes valuable time and energy, but for all large projects where our results could be questioned, I insist that we use a 'procedural manual' that fully documents all aspects of storage, naming conventions, security, access, QA/QC, examples of all forms, etc. etc. for the project. All staff are expected to know and use these guidelines, and the manuals are updated as needed. We probably loose a half day per week in doing all this, but we can actually go back years later and resurrect what was done, when, etc. A 'chain of evidence' approach is used. If you conduct your work such that you assume that you will be audited/sued/reviewed/challenged you pay a price in productivity and speed, but you save yourself a ton of grief if it actually happens, and you are doing what you should do anyway. Good luck on your project.
5232	Proliferation of proprietary file formats, methods and systems will hold back any significant advances the industry can or could expect. Looking to other industries with similar structures and data issues would yield a more cohesive environment.
5234	My system has allowed me to share with other people my GIS project as well as to re-create the GIS environment in other computers without problems.
5256	No more comments.
5257	The archiving issue of how to make many GIS data sets coming from a variety of archaeological contractors working in the same Geographical region (in our case [---]) available to be queried en masse. A defined region needs to offer a joined up projection of the various GIS data sets re-deposited by project that have taken place therein. Guidelines for depositors need to be clear and fairly draconian to enable this - bearing in mind if an archae offer this facility it will need to support and pay for any licence fees for the systems used to do so in perpetuity.
5258	GIS is ONE tool to analyze the excavation data. I try to avoid spending too much time simply working on software. That's not what I'm getting paid for and (more important) it's not subject of my research.
5290	I think that standardisation is needed for elaboration and long term storage of GIS projects and especially row and pre-treated data. But we have to make sure that the standards change each period of time to cope with the evolution of the software and hardware development. For that systematic migrations should be operated when necessary and according to my knowledge this has to be done whenever a change occurs in the material we work with (i.e., new software or hardware etc...).
5292	The software evolves so quickly that it requires substantial investment to be current.
5308	This survey has been useful in focusing attention on record keeping procedures. You might ask what medium was used to keep copies of GIS projects CD ROM, hard drives, etc.
5331	I am very glad to have taken the time to answer this questionnaire. It has been a learning process and will surely orient on where to concentrate my efforts for adequate data archive in the future. Nevertheless, there is no concensus at a national level of the minimal standards of record keeping and archival maintenance and certainly not in something as dynamic as digital data processing. This

	may be true for other countries in Latin or South America. Certainly more interaction with people or communities with more experience will save a lot of trial and error, as well as narrowing the “digital divide” between Latin America and the “northern hemisphere”. In this sense I see a great need for implementing more networks on a regional basis in order to: a) share experiences (avoid tiral & error); b) access information; c) training; d) initiate workshop experiences. Most researchers in LA, even within the same country, work mostly in isolation as they are initiating the experience in GIS archaeology applications and have usually received training in centers geared to other disciplines (not necessarily research).
6004	I think the kind of report keeping depends on the character of the project. We were a small team to built up archaeological predictive maps. These maps (digital output and paper) can be used now in cultural resource management and it is not neccesary to give access to the raw data to anybody else. A following project will now continue the work and maybe it is possible in the future to built up a software program for creating predictive maps. This should be used by the CRM and connected with the archaeological database.
6005	GIS has become a major tool in archaeological research, but still a lot of work is needed to tailor to the needs of the archaeological science. Besides that, more than often, GIS is synonymous with ‘data made visible’ for general ‘optical assessment’, but this ought be just the beginning not the end of GIS-supported archaeological research.