

THE VALUE AND IMPACT OF THE ARCHAEOLOGY DATA SERVICE

A study and methods for enhancing sustainability

Final Report, September 2013

Neil Beagrie, Charles Beagrie Ltd and
John Houghton, Centre for Strategic Economic Studies (CSES), Victoria University

Copyright: Charles Beagrie Ltd and Victoria University 2013



This work is licensed under a [Creative Commons Attribution-Non Commercial 3.0 Unported Licence](https://creativecommons.org/licenses/by-nc/3.0/)

Charles Beagrie Ltd.
2 Helena Terrace
College Street
Salisbury
SP13AN



CONTENTS

| | |
|---|----|
| Contents..... | 2 |
| Acknowledgements | 4 |
| Foreword – The Impact of Research Data Centres Report Series | 5 |
| Executive Summary | 6 |
| 1 Introduction | 11 |
| 1.1 Background | 11 |
| 1.2 Aims and Objectives of the Study..... | 11 |
| 1.3 Brief Description of ADS..... | 12 |
| 1.4 Report Contents | 13 |
| 2 Approach and methods used to collect data | 14 |
| 2.1 Desk Research | 14 |
| 2.2 Interviews..... | 14 |
| 2.3 Online Surveys..... | 14 |
| 2.4 Stakeholder Focus Group | 15 |
| 3 Previous Impact and Economic Studies of Archaeology and Cultural Heritage | 16 |
| 3.1 Impact of the Heritage Industry | 16 |
| 3.2 The Archaeology Profession..... | 17 |
| 3.3 Impact of the Archaeology Data Service | 21 |
| 4 Approaches and Methods to Analyse the Data and Measure Economic Impact | 22 |
| 4.1 A Brief Review of Approaches to Valuing Research and Information Services | 22 |
| 4.1.1 Library and information services and public sector information | 22 |
| 4.1.2 Research publications..... | 24 |
| 4.1.3 Research data | 26 |
| 4.2 Approaches Used in This Study | 28 |
| 4.2.1 Investment and use value | 28 |
| 4.2.2 Contingent value | 28 |
| 4.2.3 Consumer welfare and net economic value | 30 |
| 4.2.4 Efficiency impacts and returns on investment..... | 30 |
| 4.2.5 Keeping Research Data Safe (KRDS) Benefits Framework | 30 |
| 5 Quantitative Analysis of The Value and Impact of ADS | 32 |

| | | |
|-------|--|----|
| 5.1 | Weighting the Survey Data | 32 |
| 5.2 | The Value and Impacts of ADS Data and Services | 34 |
| 5.2.1 | Investment and use value | 34 |
| 5.2.2 | Contingent value | 36 |
| 5.2.3 | Consumer surplus and net economic value | 38 |
| 5.2.4 | Efficiency impacts | 38 |
| 5.2.5 | Return on investment in Archaeology Data and Services..... | 41 |
| 5.2.6 | Longer term collection and preservation and the accumulation of value..... | 43 |
| 5.2.7 | Summary of the quantitative economic analysis | 46 |
| 6 | Qualitative Approaches and Analysis | 48 |
| 6.1 | Surveys | 48 |
| 6.1.1 | Depositors of data with ADS | 48 |
| 6.1.2 | Users of ADS data and services | 50 |
| 6.2 | Interviews..... | 57 |
| 6.3 | KRDS Stakeholder Benefit Map for ADS..... | 59 |
| 7 | Stakeholders' Perceptions of Value | 62 |
| 7.1 | Initial Stakeholder Feedback | 62 |
| 7.2 | Final Stakeholder Feedback | 63 |
| 8 | Conclusions and Recommendations..... | 65 |
| 8.1 | Recommendations..... | 65 |
| | References | 67 |
| | Appendix 1: Summary of Survey Responses..... | 73 |

Acknowledgements

We would like to thank: Daphne Charles for her input as a researcher to the study; Catherine Hardman, Stuart Jeffrey, Julian Richards and colleagues at the Archaeology Data Service for their collaboration and support; Neil Grindley and Rachel Bruce at Jisc who provided feedback and guidance throughout; and finally, all the interviewees and survey respondents, and the heritage professionals that attended and contributed to the stakeholder focus group, who all gave valuable time and input to the study.

Neil Beagrie (Charles Beagrie Ltd) and John Houghton (CSES, Victoria University)

September 2013

Foreword – The Impact of Research Data Centres Report Series

This report forms part of a series of independent studies produced by the authors on the value and impact of three UK research data centres. These reports cover the Economic and Social Research Data Service (ESDS), the British Atmospheric Data Centre (BADC) and the Archaeology Data Service (ADS). Each report was commissioned independently over a period of two years, and there are, therefore, differences in these studies arising from varying call requirements and levels of funding. However, to enable ready comparison we have deliberately structured the reports from each study in a similar way. As independent studies each report is intended to be free-standing. Consequently, there is some commonality in sections of text across the reports, particularly when methods and approaches are discussed. To summarise and facilitate dissemination of key findings, a separate synthesis *Impact of Research Data Centres Report* is being prepared from all three studies for publication by Jisc.

ADS IMPACT STUDY EXECUTIVE SUMMARY

Background, Aims and Approaches

Jisc and other funders, together with Higher Education Institutions, are investing substantial resources in projects and services for the curation and long-term preservation research data. It is a high priority area and there is strong interest in establishing the value and sustainability of these investments. The critical concept that determines how much or how little attention is paid to the long-term sustainability of digital content is how much value that content is perceived to have. This value is not solely economic, but in a tougher financial environment the economic arguments on value are increasingly important.

The proposal for this study of the Archaeology Data Service (ADS) was put forward in response to the Enhancing Sustainability of Digital Collections strand of the Jisc 16/11 Programme Call, which was designed to allow institutions to investigate and measure how effectively action can be taken to increase the prospects of sustainability for specified digital resources.

The aim of this study has been to explore and attempt to measure the value and impact of the Archaeology Data Service (ADS) and communicate the findings to key stakeholders. The project has addressed the high-level objectives in the Jisc Call by:

- Surveying and analysing perceptions of the value of digital collections held by ADS;
- Extending testing and development of economic and survey collection methods to ADS;
- Analysing their potential contribution to sustainability for the ADS and others;
- Communicating findings on the economic impact of ADS to its key stakeholders; and
- Identifying and disseminating lessons learnt to the wider JISC and research data communities in the UK and beyond.

The study shows the benefits of integrating qualitative approaches exploring user perceptions and non-economic dimensions of value with quantitative economic approaches to measuring the value and impacts of research data services. Such a mix of methods is important in capturing and presenting the full range and dimensions of value. The approaches are complementary and mutually reinforcing, with stakeholder perceptions matching the economic findings. For example, both our qualitative and quantitative analysis highlights the important contribution of ADS data and services to research efficiency.

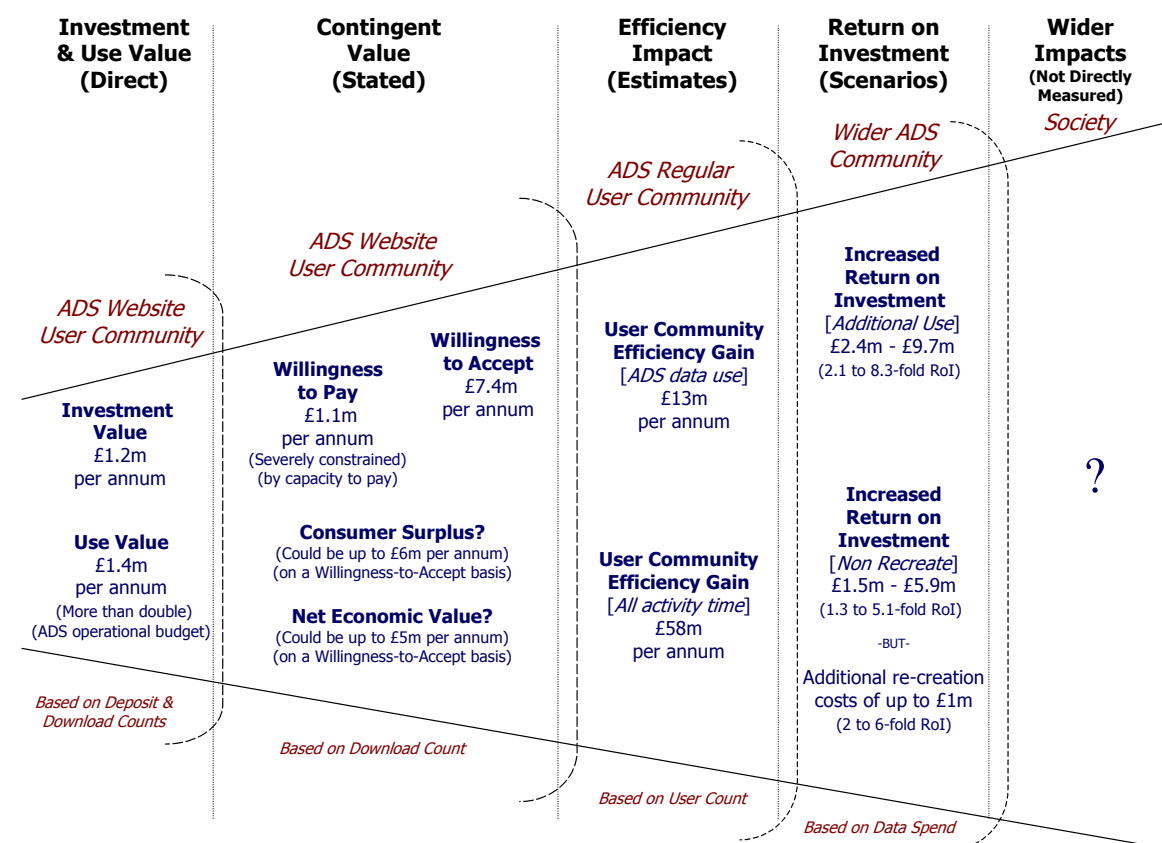
The study has changed stakeholder perceptions, increasing recognition of the value of the ADS and digital archiving and data sharing generally. Most stakeholders already valued ADS highly, but felt the study had extended their understanding of the scope of that value, and the degree of its value to other stakeholders. They were positive about seeing value expressed in economic terms, as this was something they had not previously considered or seen presented, but they also felt it was important not to dwell exclusively on economic measures of value.

These messages and our key findings on the impact of the ADS have been disseminated widely to ADS stakeholders over the last 6 months via an information leaflet from the study (in print and PDF), and in conference presentations and posters.

The Economic Impact and Value of ADS

Our *quantitative analysis* suggests that the economic benefits of ADS substantially exceed the operational costs. A very significant increase in research efficiency was reported by users as a result of using the ADS, which we calculate to be worth at least £13 million per annum – 5 times the costs of operation, data deposit and use. We also identified a potential increase in return on investment in data creation/collection resulting from the additional use that was facilitated by ADS that may be worth between £2.4 million and £9.7 million over thirty years in net present value from one-year's investment – a 2-fold to 8-fold return on investment. Due to the conservative treatment of use and user statistics, the value estimates presented are likely to be conservative. We have not directly measured wider impacts on society, although the returns on investment provide some window onto those impacts. Nevertheless, there are many forms of wider economic and social values that are not being captured. As such, these estimates should be seen as minimum values.

Figure 1: The economic value and impacts of the ADS research data service



Source: Authors' analysis

Overall, we found that:

- The investment value of ADS is around £1.2 million per annum. The direct investment in ADS by its funders is c. £600,000 per annum and we found a similar amount is invested indirectly by depositors (in terms of the time and effort they invest in depositing collections). This is

the first time the value of depositors indirect investment has been calculated, and it emphasises their importance as stakeholders in the service.

- The direct use value of ADS to its user community is around £1.4 million per annum at current prices and levels of activity – more than double the operating budget.
- Willingness to pay is an expression of value by the users, who reveal that they value their access at around £1.1 million per annum. However, it was very clear that the capacity of many users to pay is severely constrained, and the free-to-access model of ADS is important in achieving the impact it does.
- When capacity to pay is limited, the amount that users would be willing to accept in return for giving up their access to ADS for a year can be a better indicator of the value they place on it. Looked at this way, ADS data and services are worth around £7.4 million per annum to its users.
- The contribution of ADS to its user community can also be seen in terms of its impact on their research, teaching and studying efficiency, and we found that the efficiency impacts of ADS might be at least £13 million and could be as much as £58 million per annum.
- Exploring scenarios relating to the impacts of ADS data and services on returns to investment in the data held, we found that ADS facilitates additional use which may realise additional returns to the research and data creation/collection activities underpinning it that could be worth between £2.4 million and £9.7 million over thirty years in net present value from one-year's investment – effectively, a 2-fold to 8-fold return on investment (Figure 1).

While our study is a snapshot, there is evidence to suggest that the value of ADS data and services has increased over time, with the growth of collections. We believe that the value of ADS should continue to increase in the future.

The Qualitative Impact and Stakeholder Perceptions of ADS

Our *qualitative analysis* shows that the interviews and survey comments reveal strong support for the ADS, with many aware of the value of the services for them personally and for the wider archaeology community. They are also aware of the increasing importance of digital data curation in an increasingly digital environment, and the considerable benefits that can accrue from digitisation and curation.

In Archaeology, perhaps more than many other disciplines, there is only one accredited place to deposit which ensures long term curation and open access. ADS sets the standard for good practice in these areas and in this way has been instrumental in changing the professional landscape. It is pivotal in facilitating the shift from the necessarily narrow focus of print publication, to more comprehensive digital dissemination including specialist reports. Depositors have already so far committed themselves to this route that they would find it practically impossible to turn the clock back. For funders and depositors, the ADS is important for dissemination, impact, reaching the widest possible audience, and ensuring a long term legacy for their work.

While ADS has wide strategic importance for heritage management and the planning framework, the government and the Higher Education sector, its business model of paying once for in-perpetuity access is one that also appeals to the commercial sector.

Even as they struggled to fully quantify the benefits they received from ADS, users recognised that as an integrated and comprehensive source it saves them time and money, and thus improves the quality and quantity of the research they can undertake with the resource available to them.

We summarised Stakeholder benefits using approaches and formats developed by the Keeping Research Data Safe (KRDS) projects, particularly aspects of the KRDS Benefits Framework. The benefits identified in Table 6.2 of this study were drawn from the desk research, interviews and surveys, and arranged by key stakeholders identified at the stakeholder focus group.

RECOMMENDATIONS

1. The approaches used in the ADS Impact Study have now been applied to three UK data centres spanning very different disciplinary domains. The experience suggests that the approaches are transferable, but they require significant customisation to fit disciplinary and service differences. **There would be benefits from further research developing, refining and further exploring applications of the methods used in this study, as making the “business” or funding case for data centres and services plays an increasing role in ensuring their sustainability.**
2. To date these approaches have only been applied to national subject data centres. However, they should be equally applicable to other international or local institutional repositories holding research data. **Hence, we should consider also applying these methods of valuation at international and/or local levels.**
3. It is also clear that different data centres and services collect financial and operational data, such as user statistics, data deposit, data access and download statistics, to varying levels of detail. **There would be considerable advantage to providing guidance regarding the collection of such data as it is essential in making the funding case.** This should ensure a greater degree of standardisation of statistical records across centres, as well as providing the basis for more comprehensive and reliable data for economic analysis.
4. The study has looked at the aggregate value of the ADS. There is also significant scope for more granular studies that focus on the value of specific collections, such as the grey literature library, or the economic value of ADS services to specific groups. **Hence, we should consider applying these methods of valuation at more granular levels than the overall collections or all stakeholders.** There may also be some practical advantages to a narrower focus in simplifying some of the statistics and the analysis of different usage patterns across collections and user groups.
5. Value and perceptions of value change over time. The ADS user community was surveyed in 2010 for the RIN study and both ADS users and depositors surveyed for this study during June 2012. **ADS and funders should consider opportunities to repeat the ADS surveys and extend the available time series of comparative studies in future years. Ideally another survey should be considered within the next five years.**

6. During presentations given by the ADS on this impact study, there was a general audience reaction that the heritage sector outside of Higher Education has been slow to adopt measures of economic value and that the approaches used in this study have a lot to offer to other heritage organisations. **The wider heritage sector and its funders may wish to consider this study and any implications and applicability for museums, libraries and archives and their digital collections.**

1 INTRODUCTION

1.1 Background

Jisc and other funders together with Higher Education Institutions (HEIs) are investing substantial resources in projects and services for the curation and long-term preservation research data. It is a high priority area for Jisc and HEIs and there is strong interest in establishing the value and sustainability of this investment. The critical concept that determines how much or how little attention is paid to the long-term sustainability of digital content is how much value that content is perceived to have. This value is not solely economic, but in a tougher financial environment the economic arguments on value are increasingly important.

The proposal for this study was put forward in response to the Enhancing Sustainability of Digital Collections strand of the Jisc 16/11 Programme Call, which was designed to allow institutions to investigate and measure how effectively action can be taken to increase the prospects of sustainability for specified digital resources. The study was funded by Jisc and undertaken by Neil Beagrie of Charles Beagrie Ltd and Professor John Houghton of the Centre for Strategic Economic Studies (CSES), Victoria University.

The study has analysed and surveyed perceptions of the value of digital collections held by the Archaeology Data Service (ADS). Importantly, we have also assessed and quantified the economic value of those collections with the ultimate objective of improving their prospects for sustainability. A range of economic approaches drew on data gathered through online surveys, user and depositor statistics to supplement and extend other non-economic perceptions of value. Together they provide a powerful insight into the value of ADS.

In the final stages of the project, we have focussed resources on one practical enhancement measure for adding value: communicating our findings on the value and economic impact of ADS to its key stakeholders.

The issues and approaches in this study are likely to have wider relevance beyond archaeology. Through this report and other presentations we aim to disseminate to the wider Jisc and research data communities our findings and recommendations on measuring impact, and how best to communicate that effectively to others.

1.2 Aims and Objectives of the Study

The aim of this project has been to explore and attempt to measure the value and impact of the ADS and communicate the findings to key stakeholders.

The project has addressed the high-level objectives in the Jisc Call by:

- Surveying and analysing perceptions of the value of digital collections held by ADS;
- Extending testing and development of economic and survey collection methods to ADS;
- Analysing their potential contribution to sustainability for the ADS and others;
- Communicating findings on the economic impact of ADS to its key stakeholders; and

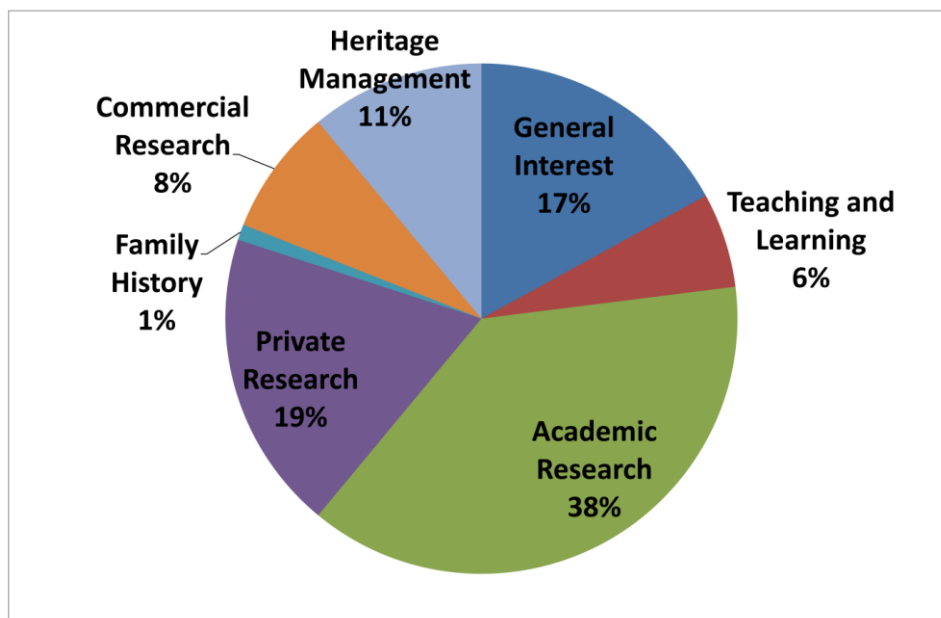
- Identifying and disseminating lessons learnt to the wider JISC and research data communities in the UK and beyond.

1.3 Brief Description of ADS

The ADS, based at Kings Manor, York, was established in September 1996. It supports research, learning, teaching and practice with freely available, high quality and dependable digital resources. It does this by preserving digital data in the long term, and by promoting and disseminating a broad range of archaeology data. The ADS promotes good practice in the use of digital data in archaeology, it provides technical advice to the research and other user communities, and supports the deployment of digital technologies.

It has a broad range of users including commercial archaeological contractors, as well as academic archaeologists in universities. Compared to many data services it also has a very high proportion of private individuals as users.

Figure 1.1: Primary re-use of ADS data



Source: ADS

ADS has been particularly influential as a pioneer in the field of digital preservation, promoting good practice in the use of digital data in archaeology, providing technical advice to the research and other user communities and taking the lead in a wide range of projects.

The ADS is also in the vanguard of work designed to transform the flow of data, finding new ways of linking archives around the world, a genuinely global network for communication that cuts across traditional research boundaries. That is why the ADS is important to the discipline and why it is highly regarded. For example, the ADS 'Making the Leap' project won the British Archaeological Award for Best Archaeological Innovation in 2008. And, in December 2012, the ADS received the prestigious Digital Preservation Coalition Decennial Award for the most outstanding contribution to digital preservation in the last decade.

1.4 Report Contents

This report presents a comprehensive account of the research undertaken for this study, as it both explores a range of qualitative and quantitative methods that can be used in assessing the value and impacts of a research data services such as ADS, and presents the results of the analysis.

Beginning with a brief introduction to the approaches and methods used to collect the data necessary for analysis (Section 2), this report then presents a brief review of previous studies of the value and impact of archaeology and cultural heritage (Section 3). While such studies are much broader in compass than ours, they provide some context to it.

This is followed by a brief review of previous studies focusing more narrowly on a range of related information and data services, which highlights the quantitative approaches used (Section 4.1). This provide the platform for an outline of the economic approaches used in this study (Section 4.2).

Section 5 presents a detailed account of the quantitative economic analysis of the value and impacts of ADS undertaken for this study. Examining the practicality and merits of the various economic approaches employed, it concludes by suggesting that the methods used can be successfully deployed and that they demonstrate the “business” or funding case for ADS.

Section 6 presents an account of the complementary qualitative approaches and analysis, and shows not only how the qualitative survey responses, interviews and stakeholder interactions are in concordance with the quantitative economic analysis, but also how they provide insights into the nature and dimensions of value and impact.

Section 7 presents a brief review of how the conduct of this project has changed stakeholders’ perceptions, primarily though demonstrating the value of ADS to others in their immediate community and to wider communities of research and practice. This is followed by a brief summary of findings, conclusions and recommendation (Section 8).

Appendix 1 presents a summary of the depositor and user survey results.

2 APPROACH AND METHODS USED TO COLLECT DATA

In selecting conceptual approaches for this study, we have taken account of the practical limitations of collecting the necessary data through survey and interview techniques, and sought to maximize economy in data collection through commonality (i.e. the same data can be used to inform more than one of the approaches). We combined:

- Desk-based analysis of existing evaluation literature and reports, looking at both methods and findings;
- Existing data from KRDS and other studies of the costs and benefits of research data infrastructure and services;
- Existing management and internal data collected by ADS, such as user registration and access statistics, deposit records, internal operational and financial reports, and
- Original data collection in the form of an online surveys of ADS users and depositors, and semi-structured interviews.

After a brief introduction to the data collection methods used, we present the economic analysis which is based primarily on the online surveys – the results of which are described in Appendix 1.

2.1 Desk Research

Desk research included: analysis of existing evaluation literature; existing data from KRDS and other studies of the costs and benefits of research data infrastructure and services; and analysis of existing management and internal data collected by ADS, such as user registration and access statistics, deposit records, internal operational and financial reports.

2.2 Interviews

Fifteen interviewees were selected to provide a cross-section of ADS stakeholders. A semi-structured interview was conducted using a pre-defined questionnaire.

2.3 Online Surveys

Two online surveys were conducted, aimed at depositors of data with ADS and ADS users, respectively. The survey questionnaires were developed iteratively by the project team with external review and input from ADS staff and others. Given the nature of some of the economic valuation approaches being explored, and the range of affiliations, roles and seniority of the survey populations, substantial effort was needed to design questionnaires suitable for an online survey.

Significant effort, therefore, was spent on trying to reduce the likely burden on recipients in terms of time to complete the questionnaires, wording of the invitations to participate, advance notice of the surveys via the ADS website, and offering five £50 Amazon vouchers in a draw for participants. As a result, the surveys enjoyed high response rates and good completion rates given the topics and number of non-mandatory questions.

The questionnaires used a range of standard survey approaches, including the use of “critical instances”, such as the last data accessed/downloaded (for users) or last data deposited/updated (for depositors). A number of questions sought specific information on: the costs of creating and collecting the data; the time and cost involved in preparing it for deposit; the time and cost of access for users; the benefits and efficiency impacts of access; and contingent valuation (i.e. willingness to pay or accept) using stated preference techniques. Answers to these questions were interpreted carefully, in the context of open-ended text comments in the surveys and other findings from the interviews and desk research, to ensure that protest and outlier answers are excluded from the economic analysis, or included with suitable caveats. These questions were supplemented by qualitative questions asking for views on the importance and impact of ADS for both depositors and users, to ensure that the quantitative and qualitative findings were in accord.

2.4 Stakeholder Focus Group

As part of the study, a focus group workshop was held in York on 21st November 2012, to present and discuss emerging results from the study with a range of ADS stakeholders. The aims of the workshop were to seek stakeholder feedback on the emerging results, establish any change of perception of the ADS amongst participants as a result of the study, and seek their views on how the study results might be presented to the archaeological community and its funders.

3 PREVIOUS IMPACT AND ECONOMIC STUDIES OF ARCHAEOLOGY AND CULTURAL HERITAGE

Our desk research reviewed recent literature relating to the economic impact of both the heritage sector in general, and of research data centres, including the ADS, and of the archaeology profession.

3.1 Impact of the Heritage Industry

In a recent report (HLF and VisitBritain 2010) the Heritage Lottery Fund and VisitBritain argued that the sector's important economic contribution has not been sufficiently understood. They suggested that it provides highly skilled employment and training opportunities, bringing in new visitors and supporting businesses. Heritage boosts tourism, the UK's fifth-largest industry. The heritage-tourism sector is in excess of £12.4 billion a year and supports an estimated 195,000 full-time jobs – more than the advertising, car or film industries. Once economic 'multiplier' effects are included, the numbers increase to a GDP contribution of £20.6 billion a year, supporting a total of 466,000 jobs. Heritage is the most important factor behind the 10 million holiday trips made by overseas visitors to the UK each year; four in 10 leisure visitors cite heritage as the primary motivation for their trip to the UK – more than any other single factor.

In the same report, Deloitte Consulting and Oxford Economics estimated that the tourism economy will grow by 2.6% a year between 2009 and 2018 – much higher than the 0.8% forecast for manufacturing, and similar to the growth rate of the retailing and construction sectors. In employment terms, these growth figures are even more significant, since the labour intensity of the tourism sector means that more jobs are created per cent of sector growth than in other industries.

English Heritage (EH 2010) has argued that investing in the historic environment brings real economic benefits to local places. On average, £1 of investment in the historic environment generates an additional £1.60 in the local economy over a ten year period, and half of all jobs created by heritage tourism are in the wider economy that supports and supplies heritage attractions.

Ecorys and Fitzpatrick (2011) demonstrate that in Ireland too, the historic environment is a highly significant contributor to the national economy, directly supporting almost 25,000 FTE jobs. Including indirect and induced effects, it is estimated that the historic environment sector supports approaching 40,000 FTE employment positions in Ireland. In terms of contribution to national income, Ireland's historic environment is estimated to account for some €1.5 billion annually of the nation's Gross Value Added (GVA), and the historic environment contribution to the national economy is equivalent to 1% of total Irish GVA, and some 2% of overall employment. Their analysis highlighted the central importance of public funding in its role as a critical enabler or catalyst to wider investment in the historic environment, and as a mechanism for inducing private sector resources and further unlocking the scale of economic benefits embodied in built heritage assets.

Table 3.1: Total Economic Impact – Historic Environment Sector in Ireland

| | <i>Direct</i> | <i>Indirect/Induced</i> | <i>Total</i> |
|------------------|---------------|-------------------------|--------------|
| Output (€m) | 1,555 | 1,110 | 2665 |
| Employment (FTE) | 23,971 | 12,976 | 36,947 |
| GVA (€m) | 855 | 605 | 1,460 |

Source: Ecorys analysis (2011), incorporating Irish Input-Output tables.

The picture is similar on a wider European canvas (European Commission 2006). The Commission, defining the cultural sector as non-industrial sectors producing non-reproducible goods and services including heritage (museums, heritage sites, archaeological sites, libraries and archives), concluded that despite shortcomings in data availability leading to under-estimation of the real weight of the sector, the figures show that the cultural and creative sector plays an important economic and social role in Europe. The sector is performing well, increasing its trend share of economic activity and turning over more than €654 billion in 2003. In comparison, the turnover of the car manufacturing industry was €271 billion in 2001 and that of ICT manufacturers was €541 billion in 2003. The cultural sector contributed to 2.6% of EU GDP in 2003.

3.2 The Archaeology Profession

Three reports look at the state of the archaeology profession from different perspectives in 2001, 2008 and 2011. Although figures cannot be compared directly, together they offer a view of trends in research publication and/or in salaries in the archaeological profession over a ten year period.

The Council for British Archaeology (CBA 2001) looked at publication in archaeology, which was at that point in a state of transition from wholly print. As this reports is ten years old, its conclusions may have less relevance today, but it does provide a useful snapshot of life before the ADS had fully developed and had addressed many of the concerns raised.

The CBA survey highlighted widespread dissatisfaction with the structure of reports, diversity of opinion about the purposes of writing them, and problems with burgeoning “grey literature” and inadequate synthesis. When the survey started, the discipline's familiarity with electronic media was partial, but by 2001, two years later, electronic communication had arrived. ‘Publication’ and ‘dissemination’, once regarded as virtually synonymous, were no longer necessarily the same thing. A majority now considered the primary purpose of publication to be the provision of information to facilitate research, whereas the dissemination of knowledge was for public benefit.

Many identified a discrepancy between the greater level of information which was published about site structures, deposits and features, and the lack of space allowed for specialist reports, which they felt was driven by financial expediency (i.e. the limitations of print publication were already causing concern). There was still a preference for print as the primary medium for archaeological publication, but among those who had consulted electronic publications, half felt that their search

facilities were better than conventional print indexes, and that non-linear narrative was preferable to conventional narrative.

A report by Aitcheson and Edwards (2008) for the Institute of Field Archaeologists (IFA) in 2008 collected information on 2,733 archaeologists and support staff working in 808 jobs with 519 different post titles. This represented one post title for every 5.3 individuals. In 2002-03 there was one post title for every 5.5 individuals. The authors interpret this as a symptom of greater fragmentation in the profession. In 2008, on average, full-time archaeologists earned £23,310 per annum. The median archaeological salary was £20,792 per annum. The average salary for those employed in the private sector, which employed 51% of the archaeological workforce, was £20,916 per annum. By comparison, the average for all UK full-time workers was £29,999 per annum – so, overall, the average archaeologist earned just 78% of the UK average. Over the five years from 2002-03, the average earnings of archaeologists increased by 22%, on a par with the national average increase of 23% over the same period. Twelve per cent of archaeologists held a Doctorate or post-doctoral qualification, 40% held a Masters degree or higher, and 90% of archaeologists held a Bachelors degree or higher. Effectively, all those now entering the profession are graduates.

Southport's 2011 report *Realising the Benefits of Planning-Led Investigation in the Historic Environment* (Southport Group 2011), included an economic analysis of the market for archaeological services in the planning process by the LSE. The report argues that there needs to be greater emphasis on digital methods of publication, dissemination and research, and sees ADS as central to this.

Planning Policy Guidance Notes (PPGs) 15 and 16 led to an expansion of planning-led investigation in England; it was estimated in 2008 that developers were spending around £125 million per annum on archaeological work. The annual investment by developers in archaeology has been estimated at between £120 million and £170 million on projects of all sizes. Because developers generally perceive little direct benefit from archaeological investigations, contracts are often won on price alone. This has given rise to concerns within the profession about quality control. The sector continues to be largely self-monitoring and self-regulating; there is no requirement for an individual to be registered or chartered to operate as an archaeologist. The report sees the future market for services investigating the historic environment placing greater emphasis on quality than it has done to date. Planning Policy Statement 5 (PPS5) replaced PPGs 15 and 16 in 2010, and replaces *recording* with *understanding* as the prime objective of conditioned investigation. Projects need to be explicit about their intended research outcomes, and be managed according to them.

The report argues there is an apparent lack of awareness in higher education of the wealth of information generated by planning-led research and how to access it. Researchers need to be aware that results are published in a variety of formats and media, not just books and journals: more direction is needed on the use of digital resources including GIS-based HERs, ADS' Grey Literature Library collection and English Heritage's Heritage Gateway.

It also argues the Total Economic Value of a heritage asset to an individual is the sum of use value (the value they place on using or observing the asset), option value (the value they place on preserving the asset for themselves/future generations to use later) and existence value (the value they place on the asset's existence, even though they never expect to use or see it). As the asset is available to all, the value to society is the sum of all individual values. The report suggested that the

method most suitable for valuing heritage assets, such as archaeological knowledge and artefacts, is the contingent valuation method.

The report also suggests the outcomes of the current system have high existence value, as assets are generally investigated and the information they contain extracted and analysed. However use value is relatively low, as the public is not normally involved in investigations, the results are published in specialist journals and the artefacts are stored in museum warehouses. The market currently produces a least-cost means of meeting regulatory requirement. The objective should, however, be to maximize the net value to society, including use, existence and option values.

It states that the post-PPG16 boom in planning-led archaeology saw annual expenditure on archaeological services grow from an average of £8 million in the late 1980s (all from the government) to up to £100 million in the early 1990s (almost all from developers). By the late 1990s, there was a three-fold increase in the number of contracted services, such as desk-based assessments and field evaluations, and some of the larger contractors were reporting £50 million worth of annual business each. Today, developer spending is by far the largest source of funding for archaeological activity, having risen from £68 million in 2000 to £144 million in 2004, compared to the stable figure of £19 million that was spent over the same four-year period by central government and the EU together. Its relevance is also evident when looking at the most recent labour-market figures for the archaeological profession. The private sector, or contractors that cater to developers as their clients, employs 51% of the total workforce. A survey of the archaeological market carried out in 2009 found that more than 50% of all new business came from competitive tenders. One of the most recent studies of the professional sector estimated that 3,189 out of a total of 5,827 people employed as archaeologists worked in applied commercial archaeology.

From the first time Labour Market Intelligence was gathered in 1996, the total numbers in the profession increased by 55% to 6,865 in 2008. Of those, 10% worked for national government agencies, 17% in local government, 15% for universities, 51% in the private sector, and 8% for other types of organizations. The most recent figures show a decline in the number of archaeologists since 2008 to around 5,827.

Much of the literature about the archaeological profession identifies salaries as an issue, particularly at junior levels as diggers. Seventy-three per cent of contracts in archaeology are temporary and short term. An average full-time archaeologist earns just £23,310 per annum, while the median archaeological salary is £20,792 per annum. Often the labourers on a big building site earn more than the archaeologists. While most major players in commercial archaeology comply with IFA salary guidelines of pay minima, job insecurity remains high.

Interviewees in the Southport report suggested that an average local authority gets about 15,000 planning applications per year. Archaeology officers will advise on about 500, and approximately 200 ultimately require excavation – or some other form of intervention. It concludes that governments regulate markets or provide goods or services themselves in order to correct what economists call market failures. This is a technical economic term describing situations where the market prices of goods or services do not capture their full value to society. Under such conditions the operation of an unfettered free market might produce an optimal result for individual buyers and sellers, but a suboptimal result for society as a whole. At a societal level, the outcomes would be inefficient,

inequitable, or both. Development planning regulations address market failure, as do the specific provisions relating to the protection of the historic environment.

Table 3.2: Typical prices for archaeological services in connection with development

| <i>Service</i> | <i>Typical range of costs</i> | <i>Comments</i> |
|---|---|--|
| Pre-purchase desk-based advice and pre-planning desk-based assessments ('Heritage Statement') | Up to £3,500 | In current economic conditions competition has become more severe, driving prices for these services down. Interviewees say quality has suffered. |
| On site pre and post-planning investigation and excavation | From £2,000 up to £5m | A small project would be one costing £2,000 to £5,000; anything costing over £500,000 would be considered a large project. |
| Total price for projects requiring excavation | £30,000 min For infrastructure projects: up to, but not beyond 1% of total construction costs | A project costing £30,000 to £50,000 would be considered small; up to £200,000 medium; over £200,000 large. The biggest projects can cost many millions – excavations in connection with Heathrow Terminal 5 cost £8 million – i.e. about 0.2% of total development costs. |

Source: Southport (2011) p. 64

The incentive for the developer is to fulfil the planning authority's requirements for archaeological investigation with certainty and at the lowest cost in terms of time and money. This can be regarded as a cost of securing planning permission, which the developer can be expected to want to minimise. Developers generally perceive little direct benefit from archaeological work while it is happening, except to the extent that it enhances public relations and adds cachet to a development.

Archaeologists face two competing sets of incentives. As businesses the incentive is to maximise profits; as individuals archaeologists have a professional and personal interest in investigation and in the dissemination of results. Publication in peer-reviewed journals is important to many, who contribute to, use and peer-review such journals. They do not own or benefit directly from 'finds'.

Looked at through the lens of Total Economic Value, the Report argues that there appears to be an over-reliance on publication in 'large dusty academic journals', with a lot of technical detail but very limited public readership. Regulation could be shifted to emphasise the value of other methods of public engagement as an end in itself, as well as more traditional forms of print publication. The more technical information could be transmitted through other means such as websites.

3.3 Impact of the Archaeology Data Service

The most significant previous study of the impact of the ADS itself is the RIN study on *Data centres; their use, value and impact* (Technopolis, 2011), which included ADS as one of eight research data centres examined.

Technopolis note that the most widely-agreed benefit of data centres is research efficiency – which we also find and try to quantify in this study (see Section 5.2). Data centres make research quicker, easier and cheaper, and ensure that work is not repeated unnecessarily. Research quality is another important benefit, although not rated quite as highly as efficiency. Data centres provide services and support which are highly valued by researchers, including: user support; access to otherwise-unavailable datasets via reciprocal sharing arrangements; and curation, preservation and long-term access for datasets, both for their own research and for datasets created by others. It proved more difficult to identify areas where research based upon data centre resources had gone on to have significant social, economic or environmental impacts, but a few cases did illustrate how research based on data centre resources has had a positive impact upon wider society and the economy through the development of new tools and methodologies, new policies and regulatory controls, and new products or services.

Table 3.3: ADS Research Benefits (extrapolated from Technopolis (2011))

| <i>Data Centre Research Benefits</i> | <i>ADS n=</i> | <i>ADS %</i> |
|--|---------------|--------------|
| Research efficiency | | |
| It has reduced the time required for data acquisition / processing | 67 | 79% |
| It has improved the efficiency of research | 67 | 79% |
| It has reduced the financial cost of data acquisition / processing | 66 | 65% |
| It has reduced duplication of effort (i.e. unnecessary recreation of data) | 65 | 57% |
| It has enabled me to undertake a greater quantity of research | 63 | 52% |
| Research quality | | |
| It has increased the use of data in my research | 63 | 48% |
| It has improved the quality of the data I use within my research | 62 | 55% |
| It has improved the evidence base of my research | 65 | 58% |
| It has helped to improve the quality of my research outputs | 66 | 56% |
| Research novelty | | |
| It has created new intellectual opportunities (e.g. merging of several data sets to answer new questions) | 63 | 51% |
| It has enabled research to go ahead that otherwise might not have done | 65 | 62% |
| It has permitted more novel research questions to be answered / tackled | 63 | 46% |
| It has enabled new types of research to be carried out | 63 | 56% |

Source: Data extrapolated from Technopolis (2011) Figures 23, 24, 25 pp 34-39

Technopolis surveyed users of five of the data centres and quantified the research benefits under three main categories. The figures for ADS have been extrapolated from three separate figures in their report and presented in table 3.3.

4 APPROACHES AND METHODS TO ANALYSE THE DATA AND MEASURE ECONOMIC IMPACT

Moving the focus more directly to research and data services like the ADS, this section presents a brief review of past studies that seek to estimate the economic value of research-related and publicly-funded information and services, and then outlines the economic and quantitative approaches that we use in this study.

While there are many forms of value and many ways to estimate the economic costs, benefits and impacts of an activity, we focus on the more direct value of the ADS data service to its users. We seek to examine and where possible quantify the value of the data service to its users, rather than examining what those users subsequently do and the wider impact of their work. Consequently, there are wider economic and social benefits and impacts of value that are not covered directly or explicitly in the economic analysis, but which we seek to show in more general qualitative terms in Section 6 (See also Box 4.1).

4.1 A Brief Review of Approaches to Valuing Research and Information Services

Much has been written about the costs and, to a lesser extent, the benefits of more open access to research publications. To date, somewhat less attention has been paid to the value of the open curation and sharing of research data, although such studies are emerging. There are also related literatures on the value of the arts, museums, library and information services, and public sector information (PSI). Our purpose here is to briefly explore and draw ideas from these literatures, and to assess which approaches might be most usefully applied and adapted for this study. A key criterion in selection of approaches is the practicality of data collection using survey and interview techniques.

4.1.1 Library and information services and public sector information

Library and information services have been the focus of many studies estimating their value in quantitative and qualitative terms, with the former including numerous studies based on investment and use value and estimates of consumer surplus, and a number based on contingent valuation, leading to cost-benefit analyses (Noonan 2003; Missingham 2005; Imholz and Arns 2007; Svanhild 2009; Oakleaf 2010). Most focus on the more direct economic impacts or value and calculate a Return on Investment (RoI), but some explore wider impacts (e.g. the contribution of library spending to employment and gross domestic product) and outcomes (e.g. the relationship between library spending and successful grant applications (Tenopir et al. 2010)). One example, measuring the economic impact of the British Library, combined contingent valuation in the forms of willingness to pay and willingness to accept with investment value and estimates of the cost of alternatives (British Library 2004). Such an approach to library valuations has become common.

Indeed, the use of contingent valuation has become increasingly common in a range of areas where there are intrinsic values associated with what is being valued, including the environment, the arts and museum services, as well as library and information services (Bakhshi et al. 2009). Based on public choice, contingent valuation in the form of willingness to pay and/or willingness to accept

captures the preferences of people as they choose to spend their money and time on A rather than B, thereby foregoing the latter. Contingent valuation is also increasingly widely used in marketing, and in that context is often the basis for conjoint analysis, which involves respondents being asked to make explicit, rather than implicit, trade-offs (Ramirez 2009).

Public sector information (PSI) has been the focus of a number of studies seeking to estimate its value and the benefits to be derived from making it freely available. Some forms of PSI are similar to social, scientific and observational data (e.g. national statistics, fundamental geospatial data, and meteorological data).

PIRA (2000) combined measures of the investment cost (i.e. the amount spent on the collection/generation of the information) and expenditure on PSI by users and re-users, then for final users, estimated the value as expenditure on PSI or, where the PSI was freely available, as the investment cost of its collection/generation. They estimated the investment value of PSI (i.e. what governments invest in the acquisition of PSI) and the economic value of PSI (i.e. the national income attributable to activities built on the exploitation of PSI). In the European Union, they put the former at around EUR 9.5 billion per annum in 1999, and the latter at around EUR 68 billion (equivalent to approximately 1.4% of EU GDP).

The PIRA report was popularised by Weiss (2001) in the influential report *Borders in Cyberspace*. Drawing on the PIRA report, Weiss highlighted the comparison between the US and Europe, noting that the US invested twice as much as Europe in PSI, but earned 40 times more from it. Weiss suggested that this was because the US had an open access model for PSI, whereas the EU countries used a cost recovery approach. However, a number of subsequent analysts have questioned both the numbers and the conclusions. Aside from the many difficulties in estimation and attribution, a potential problem with the PIRA approach is that it may overestimate the value of PSI because it does not account for the possible use of alternative information.

In the MEPSIR study of Dekkers et al. (2006), demand and economic performance were measured in an extensive survey by directly asking both PSI holders and re-users for key economic data, such as total turnover against turnover related to PSI, total number of staff against the number of staff dedicated to handling PSI, and estimates of the domestic market size for particular types of PSI. The market value was then estimated from the average revenues multiplied by the average number of re-users per PSI domain, minus the cost of PSI collection/generation. This produced a much lower number than suggested by the PIRA study, despite it being market size rather than value added, and coming five years later. Making some adjustments to the MEPSIR estimates with the benefit of hindsight, te Velde (2009) suggested that the value might drop further – to between one-fifteenth to one-twentieth those reported in the original PIRA study.

In their report to the UK Office of Fair Trading, DotEcon (2006) and collaborator Pollock (2009) adopted a bottom-up approach to estimating the economic value of PSI products and services in the UK in an effort to overcome some of the limitations of the PIRA approach. DotEcon adopted a contingent valuation approach and estimated the net economic value of PSI from willingness to pay for PSI minus the cost of supplying it. Using a survey and published sources, the value of PSI was estimated from the net consumer surplus from PSI (i.e. the amount that customers might be prepared to pay over and above what they do pay), and the total producer surplus that arises from the provision of PSI (i.e. the extent to which revenues exceed the costs of supplying the product or

service). Adding these two estimates gave the net economic value of PSI in the UK - around £590 million per annum in 2005.

The DotEcon report also provided estimates of the value lost from not making PSI freely available, by looking at the consumer detriment resulting from: unduly high pricing; restriction of downstream competition, such as refusing to supply or discrimination; and failure to exploit PSI. This goes to the heart of important counterfactual issues, by starting to look at the value of things that do not happen unless data is made freely available. They suggested that the net value of PSI in the UK could have been approximately doubled by resolving the problems identified. While much less subject to over-estimating the value of PSI, a potential weakness of this approach lay in estimating price elasticities of demand, especially where the PSI was supplied free of charge – throwing the necessary assumption of linearity into question.

These studies suggest the considerable potential of approaches based on investment and use value, and estimating net economic value from a survey-based approach to willingness to pay and/or willingness to accept (i.e. contingent valuation using stated preferences). They also suggest that the challenge of better integrating the economic value with other qualitative forms of value largely remains.

4.1.2 Research publications

Research publications have been the focus of quite extensive economic analysis, although most focus on the costs of creation, access and preservation, and few studies look at the value and benefits of access. Those addressing the value and benefits have done so from the supply-side, in the form of macro-economic analysis based on estimates of the impacts of changes in accessibility and efficiency on returns to R&D expenditure; and from the demand-side, in the form of estimates of the impact of research on innovation and the value of that innovation to firms (Mansfield 1991, 1998; Beise and Stahl 1998; Houghton et al. 2011). The latter approach is not readily applicable to many research data services as commercial users often make up a relatively small share of total users. While the former approach is relatively data intensive, some aspects of it could be applied to research data services.

Houghton et al. (2006) and Houghton et al. (2009a) were among the first studies to explore the costs *and* benefits of open access to research publications. Houghton et al. (2009a) outlined a detailed activity cost model, based on the IDEF0 modelling standard that is often used in business process reengineering, and used it as the basis for their analysis of the potential costs and cost savings throughout the scholarly communication lifecycle arising from alternative publishing models. The focus of this part of the work was activity-based costing, looking at the activity cost differences between the alternative models and efficiency gains in terms of research information search, discovery and access, research library negotiation, acquisition and handling, publisher handling, and use efficiencies.

Exploring the wider impacts of more open access to research findings Houghton and Sheehan (2009) and Houghton et al. (2009a) developed a modified Solow-Swan model to estimate the impacts of changes in the accessibility of the information and efficiency in its access and use on returns to R&D expenditure. The standard Solow-Swan approach makes some key simplifying assumptions, including that:

- All R&D generates knowledge that is useful in economic or social terms (efficiency of R&D); and
- All knowledge is equally accessible to all entities that could make productive use of it (accessibility of knowledge).

Obviously, these assumptions are not realistic. In the real world, there are limits to efficiency and barriers to access. Addressing these real world limitations, they introduced *accessibility* and *efficiency* into the standard Solow-Swan model as negative or friction variables, then explored the impact on returns to R&D of reducing the friction by increasing accessibility and efficiency.

Houghton et al. (2009a) then put the costs, cost savings and returns to R&D together into a cost-benefit analysis. Because there is a lag between research expenditure and the realisation of economic and social returns to that research, the impact on returns to R&D was lagged and the value of those returns discounted accordingly. The cost-benefit comparisons were made over a 20 year transitional period, and suggested that the cost savings and additional returns to R&D resulting from enhanced accessibility and efficiency would exceed the costs of open access publishing models.

To date, this is one of the few methods to have explored both the direct efficiency impacts of more open access to information, in the form of activity costs and cost savings, *and* the wider economic and efficiency impacts, in the form of increased returns to R&D arising from increasing the accessibility of the information. It has been applied in a number of subsequent studies by the original authors and colleagues and by others (Houghton et al. 2009b; Houghton 2009; CEPA/RIN 2011).

There is an increasing number of research funding agencies mandating that the findings from the research they fund be made openly and freely available. In the United States, the National Institutes of Health (NIH) spends USD 28 billion on research annually, resulting in around 65,000 peer-reviewed articles. The NIH Public Access Policy “requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.”¹ Through the PubMed repository, these manuscripts and materials are made available to the public. *The Federal Research Public Access Act* (FRPAA), currently under consideration in the United States, seeks to extend the NIH policy by requiring that US government agencies with annual extramural research expenditures of more than USD 100 million make manuscripts of journal articles stemming from research funded by that agency publicly available via the Internet.²

Houghton et al. (2010) outlined one possible approach to measuring the impacts of the proposed US *Federal Research Public Access Act* (FRPAA) on returns to public investment in R&D. The project involved further development and refinement of the modified Solow-Swan model (discussed above), particularly in relation to the most appropriate lag and distribution over time of returns to R&D, the most appropriate depreciation rate for the underlying stock of R&D knowledge arising from federally funded R&D, and metrics to measure potential changes in accessibility and efficiency. Key data required for the modelling included: the implied archiving costs, the volume of federally funded

¹<http://publicaccess.nih.gov/>

²<http://www.taxpayeraccess.org/issues/frpaa/index.shtml> The FRPAA has now been superseded by the Fair Access to Science and Technology Research Act (FASTR).

research outputs (e.g. journal articles), and the levels of federal research funding and expenditure trends. The preliminary analysis used publicly available sources and published estimates. Data relating to federal research funding, activities and outputs were taken from the National Science Board (2010), and the analysis explored three sources for archiving costs: the LIFE² Project lifecycle costs (Ayrís et al. 2008), and submission equivalent costings from arXiv (2010) and NIH (2008).

Preliminary modelling by Houghton et al. (2010) suggested that over a transitional period of 30 years, the potential incremental benefits (i.e. over and above the existing NIH mandate) of the proposed FRPAA archiving mandate for all US federally funded R&D might be worth around :

- Four times the estimated cost using the higher end LIFE² lifecycle costing;
- Eight times the cost using the NIH costing, which it was suggested would probably be the best estimate; and
- Twenty-four times the cost using the historical arXiv costing.

Perhaps two-thirds of these benefits would accrue within the US, with the remainder spilling over to other countries. Hence, the US national benefits might be of the order of five times the costs, with the benefits from increased accessibility worth more than USD 1 billion over 30 years (Net Present Value).

These studies demonstrate that more open access to publicly funded research publications can bring cost savings and efficiency improvements for both the producers and users of the information, as well as increasing the return on investment in the research by making it more accessible. Crucially, they address the former through activity costing and the latter through a return on investment approach.

4.1.3 Research data

Research data repositories have been the topic of a number of studies over recent years. Efforts to understand the costs and benefits involved in research data curation and sharing typically mix quantitative and qualitative methods, and rely primarily on case studies and extrapolation. Some have provided templates for assessing costs and benefits, but few have tried to look at the value or benefits of open access to a wide range of research data types.

In a series of projects under the title Keeping Research Data Safe (KRDS) Beagrie *et al.* (2008; 2010) explored the costs and benefits of research data curation and sharing in the UK and elsewhere in Europe. The initial KRDS study investigated the medium to long-term costs of the preservation of research data to Higher Education Institutions (HEIs), and provided a brief overview of the potential benefits to HEIs from the preservation and sharing of research data. It developed a framework and guidance for determining costs consisting of: a list of key cost variables and potential units of record; an activity model divided into pre-archive, archive, and support services, and by duration of activity; and a resources template, including major cost categories.

A series of case studies from Cambridge University, King's College London, Southampton University, and the Archaeology Data Service at York University illustrated different aspects of costs for research data within HEIs. Selective illustrations of cost-benefits and costs over time were also provided. Importantly, the study noted that the costs of a central data repository are an order of magnitude greater than that suggested for a typical institutional repository focused on e-publications alone – although likely less than the user and producer costs that would result from simply opening data

without appropriate curation (e.g. the creation of related metadata, sourcing information and guides).

Fry et al. (2008) sought to identify the value of, and benefits arising from, the curation and open sharing of research data. They suggested that potential benefits include:

- Maximised return on investment in data collection;
- Broader access, where costs would be prohibitive for individual researchers/institutions;
- Potential for new discoveries from existing data, especially where data are aggregated and integrated;
- Reduced duplication of data collection costs and increased transparency of the scientific record;
- Increased research impact and reduced time-lag in realising those impacts; and
- New collaborations and new knowledge-based industries.

The Fry et al. (2008) study used a mixed-method approach, including a literature review and qualitative case studies, to inform the development of a model on which to build a business case for data sharing in UK Higher Education. This was based on extensions to the research data preservation cost model proposed by Beagrie et al. (2008, 2010), to allow estimation of costs and benefits to users depositing or accessing data. Based on the work of co-authors Houghton and Rasmussen, the report presented a simple example of cost-benefit analysis applicable to an individual dataset or repository, based on costs and potential cost savings. It described the data requirements and walked the reader through the process step-by-step. The approach was then extended to explore the more diffuse benefits of data curation and sharing at the institutional and disciplinary levels. Importantly, the report included an outline questionnaire and template to facilitate cost-benefit analysis.

Recognising that no single approach has dominated across the studies of research publications, research data and PSI that attempt to measure the value and economic impacts, Beagrie et al. (2012) drew on a number of approaches to explore the economic value and benefits of the UK Economic and Social Data Service (ESDS). They began with approaches that can be seen as estimating minimum values, and moved progressively toward approaches that can be seen as measuring some of the wider values. These included: investment and use value, contingent valuation using stated preference techniques, and economic welfare in the form of consumer surplus and net economic value. Wider benefits and impacts were explored by looking at the efficiency gains enjoyed by users and assigning an economic value to them (e.g. activity cost savings), and by estimating the impacts of increased data use on returns to investment in the data collection/creation and the related data infrastructure services necessary for hosting and sharing the data. Beagrie et al. (2012) based their analysis on extensive user and depositor surveys. In addition to the economic analysis, Beagrie et al. (2012) drew on approaches, such as the KRDS Benefits Framework, and impact case studies to illustrate qualitative benefits defined in the surveys, and agency and user interviews that could not be reflected fully by economic analysis alone.

These studies explore the direct costs and benefits and wider economic impacts, and demonstrate the possibility of using a number of approaches to estimating the value and impacts of research data services.

4.2 Approaches Used in This Study

What this brief of previous studies of the value of information services, research publications and data suggests is that the field is relatively new and no single approach has dominated across the different but related fields. Consequently, building on our study of ESDS, we propose to draw on a number of approaches to explore the economic value and benefits of ADS data and services, beginning with approaches that can be seen as estimating minimum values and moving progressively toward approaches that can be seen as measuring some of the wider value. These include:

- Investment and use value;
- Contingent valuation, using stated preference techniques;
- Welfare approaches to estimating consumer surplus and net economic value;
- An activity-cost approach to exploring the efficiency impacts of ADS data and services; and
- An approach that seeks to explore the impacts of increased use on returns to investment in data creation/collection.

In selecting these approaches, we have taken account of the practical limitations of collecting the necessary data through survey techniques, and sought to maximize economy in data collection through commonality (i.e. the same data can be used to inform more than one of the approaches). Each of these approaches is described below. The Keeping Research Data Safe (KRDS) Benefits Framework is then used to present a summary of benefits and value to different stakeholders.

4.2.1 Investment and use value

The most direct indicators of value are investment value (i.e. the amount of resources spent on the production of the good or service) and use value (i.e. the amount of resources spent by users in obtaining the good or service). Measures of the investment in access suggest the *minimum* amount that the good or service is worth to the consumers.

Both investment and use value can be established from user and depositor surveys through questions about the time and costs involved in the creation of the data, preparation and deposit of the data, and its discovery, access and use, together with usage statistics and financial information from the ADS.

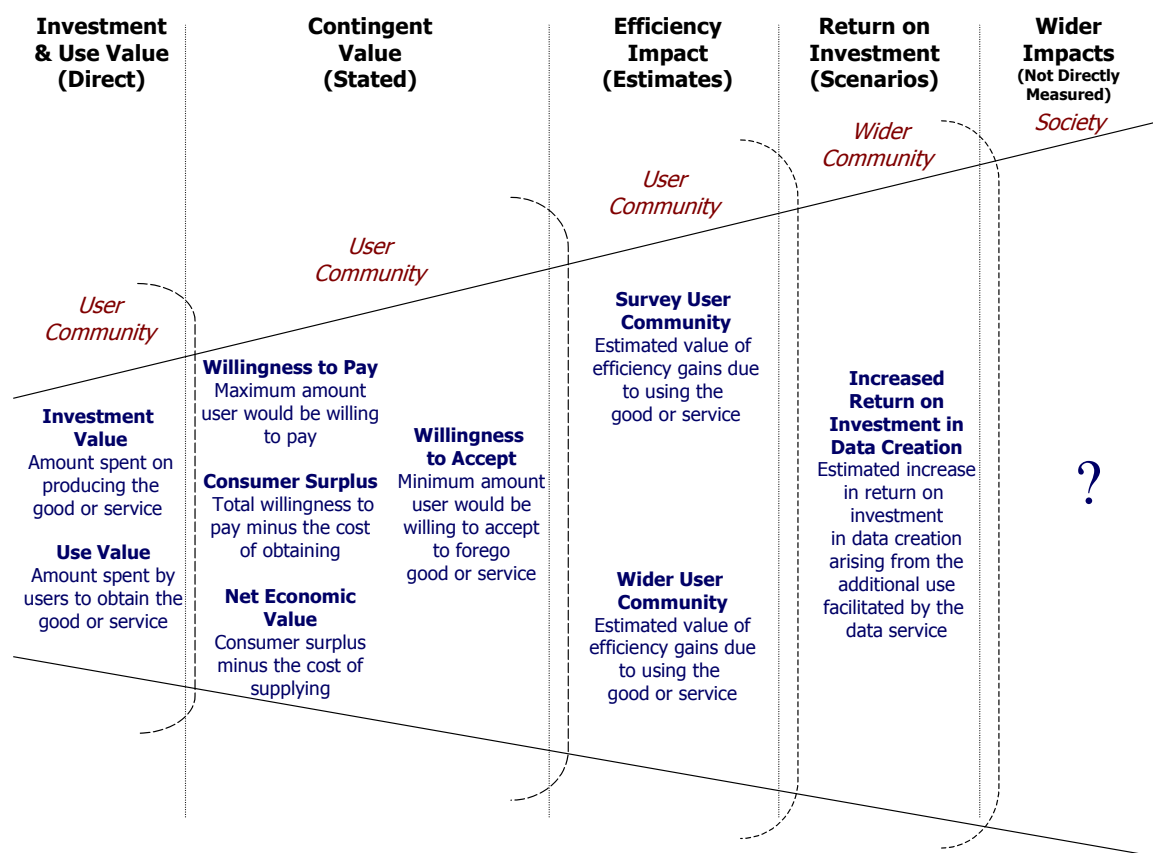
4.2.2 Contingent value

Contingent valuation involves the assignment of money values to non-market goods and services based on preferences (DTLR 2002). If a good or service contributes to human welfare, it has economic value, and whether something contributes to an individual's welfare is determined by whether or not it satisfies that individual's preferences. An individual's welfare is higher in situation *A* than situation *B*, if the individual prefers *A* to *B*. Preferences are revealed by what an individual is willing to pay for a good or service and/or by the amount of time and other resources spent obtaining the preferred good or service. Where preferences are not revealed in the market, individuals can be asked what they would be willing to pay and/or willing to accept in return for the good or service in a hypothetical market situation (i.e. stated preferences). For a public good the value is the sum of "willingnesses", as consumption is non-rivalrous (i.e. the same information can be consumed many times).

The key difference between willingness to pay and willingness to accept is that the former is constrained by the person's ability to pay (typically by disposable income), whereas the latter is not. Hence, willingness to pay directly measures the demand curve with a budgetary constraint, while willingness to accept measures the demand curve without a budgetary constraint (British Library 2004). In the case of some research data services, where many users expect institutional support and where there is a relatively large number of student users and users who are private/unaffiliated individuals, retired persons and volunteers, willingness to pay will be highly constrained and willingness to accept can be the better indicator of the value users ascribe to the service.

Where there is a bundle of different goods and services these can be treated in the aggregate or disaggregated and re-aggregated in a way that reflects the bundling and/or use (e.g. multiplying the average willingness to pay expressed by users of specific types of research data by uses of those data types), thus weighting individual survey respondents' expressed preferences by the structure of the bundle and its use. This can be particularly important where most users of a data service use just some part of the service and not all, and so the value that they express (i.e. would be willing to accept or pay) relates to just some parts of the service. In such cases, weighting the user survey responses by the totality of uses is crucial.

Figure 4.1: Methods for exploring the economic value and impacts of ADS data and services



Source: ADS Survey, Authors' analysis

4.2.3 Consumer welfare and net economic value

The benefit or welfare impact of a good or service for a consumer is measured by the consumer surplus. In a market situation, willingness to pay is made up of what is actually paid and any excess willingness to pay over and above the price paid (i.e. consumer surplus). Hence, consumer surplus is the net gain derived by the consumer from the purchase of a marketed good or service. In a non-market context, all the willingness to pay is consumer surplus because there is no market price. In practice, however, some expenditure, be it in the form of time or money, will be incurred in obtaining the non-market good or service (e.g. time spent accessing the data service). In this case, the consumer surplus will be the net gain (i.e. willingness to pay minus the cost of obtaining). Net economic value is the consumer surplus minus the cost of supply.

Box 4.1: What value is and is not being captured?

Think of the example of pharmaceuticals. Imagine that a pharmaceutical company conducts research into and develops a new drug. It then sell the drug around the world for 10 - 20 years. If one did a return on investment calculation, one would look at the expenditure on R&D and the revenue from sales.

The wider value and benefit of the new drug is in the lives saved by the better drug, or the efficiency gains in hospitals through using a better drug, with shorter hospital stays, etc. A return on investment calculation does not directly measure these things, but it not true to say that they are not captured, to some extent, because the revenue from sales is an expression of the value of the drug. Doctors prescribe the new drug because it saves lives, governments, patients and doctors pay what they do for the new drug because it has the effects it does (e.g. saving lives, improving hospital efficiency, etc).

So the methods for economic valuation that we are using in this study can, to a limited extent and by proxy, capture the wider value and impacts, even though we are not directly measuring them.

4.2.4 Efficiency impacts and returns on investment

Wider benefits and impacts can be explored by looking at the efficiency gains enjoyed by users and assigning an economic value to them, and by estimating the impacts of the increased data use facilitated by the data service on returns to investment in the data collection/creation and the related data infrastructure services necessary for hosting and sharing the data. As these latter impacts are lagged and recurring during the useful life of the data it is necessary to use a simple Perpetual Inventory Method to estimate the overall value of the impacts over time.

4.2.5 Keeping Research Data Safe (KRDS) Benefits Framework

Measuring benefits is often quite challenging, especially when these benefits do not easily lend themselves to expression in quantitative terms. Often a mixture of approaches will be required to analyse both qualitative and quantitative benefits and value, and present the differences made.

The conceptual framework used for the assessment of wider qualitative benefits and value in this study is the Keeping Research Data Safe (KRDS) Benefits Framework (Beagrie et al 2010). The KRDS Benefits Framework is a tool for identifying, assessing, and communicating the benefits from investing resources in the curation/long-term preservation of research data.

To assist institutions in applying the Framework, the Keeping Research Data Safe project created a KRDS Benefits Analysis Toolkit (KRDS, 2011). This aims to help institutions identify the full scope of benefits from management and preservation of research data and to present them in a succinct way to a range of different stakeholders (e.g. when developing business cases or advocacy).

The KRDS Benefits Framework uses three dimensions to illuminate the benefits investments potentially generate. These dimensions serve as a high-level framework within which thinking about benefits can be organised and then sharpened into more focused value propositions. The Framework can be customised and extended as needed to visualise and present benefits in different ways. For example, Who Benefits (Internal/External) in the Framework can be further sub-divided by more specific groups of stakeholders. The University of Bath recently published a detailed example of this form of KRDS stakeholder benefit analysis (Beagrie and Pink, 2012). This has formed a model for application in section 6.3.

5 QUANTITATIVE ANALYSIS OF THE VALUE AND IMPACT OF ADS

The focus of the economic analysis is to try to shed light on the value of ADS data and services to its users (i.e. both the users of ADS data and services and depositors of data with ADS) and to the wider archaeology community. Self-evidently, our survey respondents are a self-selected sub-group of users and depositors (e.g. because they responded to the survey). The response rates to both user and depositor surveys were good and a brief analysis of responses by the major possible "cuts" of the data by affiliation, role, and data type used/deposited suggests few major differences between respondents and the known ADS user community. Nevertheless, on average, one might expect that those taking the time to respond to a survey are likely to use and value ADS data and services more than those who did not respond, and respondents' reported frequency of use does suggest that they are among the more frequent users. It is also important to note that few users use all of ADS, but rather experience just part of it, and they can only express costs and value relating to the part they use. Consequently, it is necessary to consider how best to weight the survey responses to better reflect the wider depositor and user communities, and the deposits with, and uses of, ADS data and services.

5.1 Weighting the Survey Data

As the focus of the study is on the value to users and in use, user survey weighting is done on the basis of data use (i.e. ADS user access statistics). Similarly, the depositor survey is weighted by data collections deposited. This involves a two-step process of dis-aggregation of responses by data type accessed for users and deposited for depositors, and re-aggregation by data types delivered/acquired, followed by the calculation of weighted means from survey responses about costs, time, values, etc.

User survey respondents reported their frequency of use by data types used (Appendix 1, Q9) and the data type they last used – a critical incident (Appendix 1, Q10). For each of the seven data types in Q9 we create a cut, so that every respondent in that cut uses that data type. However, they also use other data types, so the same respondent can appear in multiple cuts. The mean values for willingness to pay or accept relate to overall use (Q9), so each of the seven data type cuts is then distributed according to the percentage share of data type responses within that cut, and reassembled by data type by adding the distributed values for each of the contingent valuation variables from each of the seven data type cuts. These data type mean values are then weighted by the percentage share of ADS data accesses (website page views) by data type, with the most recent nine months data used to generate annual estimates for 2012.³ Access time and cost data relate to the last data type used (Q10), and are weighted more simply by overall accesses.

³ This is necessary because of a change to a new web data analysis system at ADS.

Table 5.1: Weighting by ADS data deposited and accessed by data type

| <i>Data Type</i> | <i>Deposited (events)</i> | <i>Deposit Weighting</i> | <i>Accessed (page views)</i> | <i>Use Weighting</i> |
|-----------------------|-------------------------------|------------------------------|----------------------------------|--------------------------|
| Project Archives | 265 | 17.1% | 184,464 | 17.1% |
| Journals and Series | 48 | 3.1% | 165,965 | 15.4% |
| Archsearch (HER/NMR) | 7 | 0.5% | 333,813 | 31.0% |
| Teaching and Learning | 2 | 0.1% | 3,212 | 0.3% |
| Bibliographies | 6 | 0.4% | 2,763 | 0.3% |
| Reference Resources | 17 | 1.1% | 293,826 | 27.3% |
| Grey Lit. | 1,208 | 77.8% | 91,626 | 8.5% |
| Total | 1,553 | | 1,075,669 | |

Notes: Deposit events over 3 1/2 years, annualised, and website views between April and December 2012, annualised. See Box 5.1 for details.

Source: ADS. Authors' analysis.

For the **depositor survey**, we weight by data deposits by data type, reflecting the need to explore depositor costs. Again, the data types deposited are redistributed to data type by responses to Q4 last data type deposited (Appendix 1). The process is the same as the use weighting described above for the cost related (but not contingent value related) variables, with mean values by reported last data type deposited dis-aggregated, then weighted by data deposits by data type during 2009-12, annualised (i.e. allowing a three year period matching Q3 (Appendix 1) and accounting for the irregularity of deposits of some data types). Cuts are produced for all responses, initial deposits and updates. A further key parameter for depositor survey analysis is the number of data items acquired (i.e. should it be new data additions or should it include both updates and new additions?). On balance, we felt that data creation costs should be for new data collections only (approx. 399 per annum), while preparation and deposit costs should include updates and new editions (approx. 444 per annum in total).

Box 5.1: ADS Usage and Access Statistics

The usage statistics used in this report come from two sources, 'Analog' processed raw apache log files and a local instance of PiWik web analytics software. The PiWik web analytics are much more detailed than the information supplied by Analog, giving visits as opposed to simply page views and including useful detail on the number of return visits from specific addresses. User profile information, including the 'constituency/role' and 'primary use of data' is gathered during the optional process of registration with the ADS website. The ADS currently has 3,032 registered users.

PiWik has been running on the ADS website since the 5th of April 2012 and the data were consulted for the last time for reporting on the 12th of December 2012 (252 days). The figures used in this report for the full year are extrapolated from the figures for these 252 days. The figures have also been adjusted to remove 'bounces' (i.e. visits that last for less than 30 seconds or arrive and leave on the same page and are therefore considered to be either accidental or not a genuine 'use' of the web site). Return visits are assumed not to be bounces, as return visits are quite likely to arrive and leave on the same page, but still represent a genuine use of the website.

An important point to note is that a large number of downloads from the ADS do not traverse the website at all, but come directly to the object for download via search engines. This means that the

figures of 170,757 visits and 173,818 downloads refer only to uses of the ADS website. Comparison with Analog statistics, which monitor all downloads, show that 18.5 times more files are actually downloaded than are logged by PiWik. This gives an annual figure of 3.2 million downloads, which is borne out by the figures for data transfer.

There are three reasons why this large volume of download traffic has not been included in the figures used to calculate the economic impact of the ADS. First, because there is no way of knowing whether these files have been downloaded with meaningful use in mind or are the equivalent of a 'bounce'. Second, a proportion of these downloads will have been instigated by web crawlers that are configured to ignore exclusion files or pre-loaded by browsers. Third, Analog's file download figures do not discriminate between website specific files (e.g. a JPG embedded in a web page) and what would be considered an archived data file (e.g. a JPG illustrating an archaeological excavation).

The decision to exclude these figures from the usage figures for the ADS was based primarily on their unreliability in representing actual website use, which is the subject of this report. However, it should be noted that it has the effect of making the usage statistics extremely conservative, particularly with regard to text documents, such as the Grey Literature Library. While it is undoubtedly true that a proportion of these additional downloads do represent meaningful use of the ADS website, it is simply impossible to be definitive on what that proportion is.

Similarly, a decision has had to be taken regarding what represents 'regular' use of the website. Registration is optional (and in no way required to access the site) and it has always been assumed that the number of regular users is more than the number of registered users. PiWik has allowed us to track returning visits to the website, and shows that 44% of visits to the site are return visits, and in excess of 11,000 visitors have visited more than 51 times in a year. This can be thought of as representing weekly usage and has been adopted as the definition of 'regular' use in this report. Again this could be considered very conservative because if 'regular' use was considered as bi-monthly use then the figure jumps to more than 18,500 users, and if more than one annual use were considered 'regular' the users figure would exceed 75,000.

Due to the conservative treatment of both use and user statistics (See Box 5.1), the value estimates presented herein are likely to be conservative.

5.2 The Value and Impacts of ADS Data and Services

This section explores the value of ADS data and services circa 2012 based on then current levels of activity and use. The following section explores the longer term value as it has evolved through the operational life of ADS. In view of the limitations in available access and user data, the following should be treated as no more than estimates.

5.2.1 Investment and use value

The most direct indicators of value are investment value (i.e. the amount of resources spent on the production of the good or service) and use value (i.e. the amount of resources spent by users in obtaining the good or service). Measures of the investment in access suggest the minimum amount that the good or service is worth to the consumers.

Investment value includes annual ADS operational funding, and the costs that depositors face in preparing data for deposit and in making that deposit. For simplicity, each can be annualised (i.e. expressed as an annual cost in current prices and at current levels of activity). Activity times have been converted to costs using the annual average salaries for private/unaffiliated individuals, staff and students and the Green Book method of costing (See Box 5.2 for details).

Based on the mean cost of the last deposit (Depositor Qs 4 through 7, Appendix 1) and weighted by types of data acquired by ADS during 2009-12, depositor preparation and deposit costs amount to an estimated £465,000 per annum. The operating budget for ADS is around £698,000 per annum. Hence, treating data collection costs as sunk costs (i.e. assuming that the data would have been collected whether or not ADS existed), *investment value* amounts to almost £1.2 million per annum.

Table 5.2: Investment and Use Value (per annum)

| | | | | | |
|------------------|---|---|---------------|---|-------------|
| <hr/> | | | | | |
| Investment Value | = | ADS Operational Budget | | | |
| | | + | | | |
| | | Data Preparation & Deposit Costs | 698,000 | | |
| | | <i>Weighted Mean Cost of Last Deposit</i> | + | | |
| | | x | (1,048 x 444) | = | £ 1,162,892 |
| | | <i>Total Number of Deposits</i> | | | |
| <hr/> | | | | | |
| Use Value | = | ADS User Access Costs | | | |
| | | <i>Weighted Mean Cost of Last Access</i> | 8.43 | | |
| | | x | x | | |
| | | <i>Total Number of Accesses (Visits)</i> | 170,757 | = | £ 1,439,091 |
| | | | | | |
| <hr/> | | | | | |

Source: Authors' analysis.

Use value includes ADS user access costs. Again costing is done on the basis of average annual salaries (Box 5.2). Multiplying the use-weighted mean of user access costs (User Qs 10 through 14, Appendix 1) by the estimated number of accesses from ADS website access statistics (i.e. approximately 171,000 visits per annum, ignoring bounces) suggests annual user access costs or *use value* of around £1.4 million.⁴ The estimated cost of the time spent using the data accessed would be around £30 million per annum, but this represents the use value of the data rather than of the data service per se.

⁴In some circumstances those depositing data might also be considered to be users, implying the addition of depositor preparation and deposit costs of around £465,000 per annum, and a total use value of around £2 million per annum. However, ADS does charge some depositors, implying that there may be some double counting involved in including depositor costs in this case. Hence, these costs are not included.

Box 5.2: Costing Activity Times

Activity times were converted to costs by assigning each respondent to a salary group based on (Aitchison and Edwards 2008), adjusted for inflation (using UK CPI), then scaling to include non-wage labour costs using a 30% uplift, based on the Green Book method (Green Book 2011). For students, we used the lower end graduate salaries reported in the Times and Guardian Higher Education pages, for undergraduates (£15,000 pa) and postgraduates (£20,000 pa), as we felt this best reflected the opportunity cost of earnings forgone. For the small number of retired and voluntary worker respondents we set the salary to zero and applied the average Green Book cost mark up of the other users to reflect the non-salary costs involved. Across the respondents, this resulted in an average costing of £18 per hour - £21 per hour for those saying that their duties included teaching, £20 per hour for those researching, and £13 per hour for those studying (these categories are not exclusive).

Source: Authors' analysis.

5.2.2 Contingent value

The contingent value of a non-market good or service is the amount users are "willing to pay" for it and/or "willing to accept" in return for it. For a public good the value is the sum of "willingnesses", as consumption is non-rivalrous (i.e. the same information can be consumed many times). The key difference between willingness to pay and willingness to accept is that the former is constrained by the person's ability to pay (e.g. by disposable income), whereas the latter is not. It should be noted that Archaeology, being closely linked to construction, is one of the sectors that has suffered most heavily in the recent UK recession, with a 15% decline in curatorial and commercial archaeological staff posts between August 2007 and October 2011 (Institute of Field Archaeologists 2008-2011). This may be reflected in some of the survey responses to contingent valuation questions. For example, expressed willingness to pay may be severely impacted by limited capacity to pay.

Willingness to pay (WTP)

In the user survey, respondents were asked to express their *willingness to pay* in terms of: (i) an annual (subscription) fee; and (ii) on a pay-per-access (PPV) basis (User Q35, Appendix 1). Removing 11 protest answers that are typical of this technique, we had 181 responses. The use-weighted mean of the individual willingnesses to pay is £227 per annum. The use-weighted mean on a pay-per-access basis is £10. Thus, multiplying these use-weighted means by the estimated number of accesses from ADS website access statistics suggests a willingness to pay for access to ADS data and services of around £1.13 million.⁵

Willingness to accept (WTA)

Willingness to accept is not constrained by capacity to pay, which may have been an important limitation to willingness to pay given the large number of student, private/unaffiliated and volunteer

⁵ While consistency demands the use of access data to estimate total willingness to pay/accept, somewhat higher estimates would be produced were estimated user numbers employed instead. For example, on a user number basis, willingness to pay might be as high as £4.4 million per annum, and willingness to accept almost £16 million per annum.

users. However, some respondents expressed a willingness to accept nothing in return for giving up their access to ADS because they believe that the data should be free, rather than it being of no value to them.⁶ Including the 18 non-protest zero responses, the use-weighted mean of the individual willingness to accept was £8,603 per annum. However, there was one regular user who suggested he/she would be willing to accept a minimum of £1 million to give up all access to ADS for a year. This answer was 15 times higher than the second highest answer, although there was nothing else to suggest that it was a protest answer. If this response is excluded as an outlier, the use-weighted mean of the individual willingness to accept falls to £1,441 per annum or £43 per access (User Q34, Appendix 1), and willingness to accept amounts to £7.4 million per annum.

Table 5.3: Contingent Value based on Stated Preferences (per annum)

| | | | | |
|--|--|--|--|--|
| <div> <div>Weighted Mean Willingness to Pay Per Annum (per use)</div> <div>227 / 69</div> <div>x</div> <div>Total Number of Accesses</div> <div>170,757</div> <div>+</div> <div>Willingness to Pay = (Individual Willingness to Pay Per Access</div> <div>10</div> <div>x</div> <div>Frequency of access)</div> <div>170,757</div> <div>x</div> <div>Total Number of Accesses</div> <div>÷</div> <div>2</div> <div>÷</div> <div>2</div> </div> | | | | |
| <div> <div>43</div> <div>x</div> <div>Total Number of Accesses</div> <div>170,757</div> <div>=</div> <div>£ 1,127,283</div> </div> | | | | |
| <div> <div>Weighted Mean Individual Willingness to Accept Per Access</div> <div>43</div> <div>x</div> <div>Total Number of Accesses</div> <div>170,757</div> <div>=</div> <div>£ 7,396,094</div> </div> | | | | |

Source: Authors' analysis.

When the individual pay-per-access answers are multiplied by the individual frequency of access answers it amounts to a mean of £449 per annum (i.e. 2.3 times the mean annual willingness to pay reported by respondents). This is not surprising, as one would expect there to be a premium for the pay-per-access model, because there is no longer-term commitment to pay involved.

⁶ Some also expressed a zero willingness to pay, for the same reason.

5.2.3 Consumer surplus and net economic value

The welfare impact or *consumer surplus* for a non-market good or service with public good characteristics is indicated by the total willingness to pay minus the cost of obtaining or accessing the good or service. From the data above, this suggests a net consumer surplus that is negative (-£312,000). However, as noted, capacity to pay among the archaeology user community is limited. If willingness to accept were taken as an alternative expression of value, the implied consumer surplus would be around £6 million per annum.

Table 5.4: Consumer Surplus and Net Economic Value (per annum)

| | | | | | |
|--------------------|---|--|----------------|---|---------------------|
| Consumer Surplus | = | Willingness to Pay | 1,127,283 | = | -£ 311,807 |
| | | - Cost of Obtaining ADS User Access Costs | - 1,439,091 | | |
| Net Economic Value | = | Consumer Surplus | -£311,807 | = | -£ 1,009,807 |
| | | - Cost of Provision ADS Operational Budget | - 698,000 | | |

Source: Authors' analysis.

The *net economic value* of a good or service is indicated by the consumer surplus (i.e. the willingness to pay minus the users' cost of obtaining access) minus the cost of supplying the good or service (i.e. ADS's annual operating budget). Due to the user community's limited capacity to pay, the implied net economic value of ADS data and services is negative (-£1 million). Again, however, if willingness to accept were taken as an alternative expression of value, the implied net economic value of ADS would be around £5.3 million per annum.

In light of the limited capacity of the archaeology user community to pay and the consequently low expressed willingness to pay, the contingent valuation (willingness to pay) method has not worked well. Nevertheless, if willingness to accept is taken as an alternative expression of value, the net economic value of ADS data and services to its users would be around 8 times the annual operating budget of ADS.

5.2.4 Efficiency impacts

User survey respondents were streamed into those studying, teaching and researching, and asked a series of questions about their activities, including:

- How many hours per week, on average over the last twelve months, they had spent studying/teaching/researching;
- What share of their studying/teaching/researching time they spent with data from ADS and data from all other sources; and
- What impact ADS data and services had on their studying/teaching/researching efficiency.

Following feedback during pilot testing of the questionnaire, the questions asking users to estimate the share of their studying/teaching/researching time spent with data from ADS and all other sources offered percentage scales, which were in 10 percentage point increments from 0% to >90%. While not ideal (e.g. it involves interpreting categorical data as continuous), these scales provide an approximate guide to respondents and encouraged more responses than might a simple open-ended question.

The questions asking users to estimate what impact ADS data and services have on their studying/teaching/researching efficiency also offered percentage scales, for the same user testing feedback reasons noted above. Beginning from "negative change" and "no change", respondents were offered 5%, 10% and then 10 percentage point increments to 90%, and then >90%. Again this is not ideal, but did provide some guide as to impacts on efficiency. It is worth noting that no one chose the "negative impact" response.

Student respondents spent between 10% and up to 40% of their time with data from ADS, with an approximate mean of 25%. They reported spending a further mean of 48% of their study time using data from other sources (User Q20, Appendix 1). Combined with their study time (User Q19, Appendix 1), this suggests that student respondents spent an average of around 5 hours a week with ADS data over the last 12 months (median 3 hours), and a total average of 15 hours a week with data from all sources (median 11 hours). Asked to what extent their use of ADS data and services had changed their study/learning efficiency (User Q22, Appendix 1), the reported mean was a 44% efficiency gain, which translates to an efficiency gain equivalent to almost 9 hours a week at current activity times.

Research respondents spent between 10% and up to 30% of their time with data from ADS, with an approximate mean of 21%. They reported spending a further mean of 56% of their research time using data from other sources (User Q25, Appendix 1). Combined with their research time (User Q24, Appendix 1), this suggests that research respondents spent an average of around 4 hours a week with ADS data over the last 12 months (median 2 hours), and a total average of 14 hours a week with data from all sources (median 10 hours). Asked to what extent their use of ADS data and services had changed their research efficiency (User Q27, Appendix 1), the reported mean was a 44% efficiency gain, which translates to an efficiency gain equivalent to around 7 hours a week at current activity times.

Teaching respondents spent between 10% and up to 30% of their time with data from ADS, with an approximate mean of 16%. They reported spending a further mean of 42% of their study time using data from other sources (User Q30, Appendix 1). Combined with their teaching and preparation time (User Q29, Appendix 1), this suggests that they spent an average of around 2 hours a week with ADS data over the last 12 months (median 1 hour), and a total average of 8 hours a week with data from all sources (median 4 hours). Asked to what extent their use of ADS data and services had changed their efficiency (User Q32, Appendix 1), the reported mean was a 32% efficiency gain, which translates to an efficiency gain equivalent to around 3 hours a week at current activity times.

Each of the respondents was allocated to a salary scale according to their role and affiliation and an activity cost was calculated using a 30% uplift based on the Green Book method (see Box 5.2). This enabled us to estimate the approximate value of the reported efficiency impacts of ADS data and

services. With the caveats noted above in mind, we found that the efficiency impact of ADS data and services might be worth around £9 million per annum among regular teaching users, £4 million per annum among student users, and as much as £45 million per annum among regular research users. Hence the total estimated efficiency impacts of ADS data and services among its regular user community might be as much as £58 million per annum at current activity times - £54 million per annum if students are excluded.

Table 5.5: Efficiency Impacts (per annum)

| | | | | |
|---|---|----------------------------|--------------|----------------|
| <hr/> | | | | |
| Efficiency Gains from ADS Use (All Activity Time) | | Number of Regular Users | 11,020 | |
| | | x | x | |
| | | Cost Per Hour | (13 to 21) | |
| | = | x | x | = £ 58,623,783 |
| | | Activity Time | (9 to 21) | (per annum) |
| | | x | x | |
| | | Per Cent Efficiency Impact | (32% to 44%) | |
| <hr/> | | | | |
| Efficiency Gains from ADS Use (ADS Data Time) | | Number of Regular Users | 11,020 | |
| | | x | x | |
| | | Cost Per Hour | (13 to 21) | |
| | = | x | x | = £ 13,126,830 |
| | | Activity Time | (2 to 5) | (per annum) |
| | | x | x | |
| | | Per Cent Efficiency Impact | (32% to 44%) | |
| <hr/> | | | | |

Source: Authors' analysis.

However, these impacts appear high and it seems that some respondents may have interpreted the question as relating to the efficiency impact on their studying/teaching/researching time spent with ADS data and/or data from all sources, rather than their total studying/teaching/researching time – which had been intended.⁷ If this is taken into account, then the implied efficiency time saving impacts would be lower, and the value of the efficiency impacts of ADS data and services among its regular users community would be around £13 million per annum at current activity times - £12 million per annum if students are excluded. This is still a very substantial impact – being some 5 times the sum of operational, depositor and user access time costs.

These estimates are likely to be upper bound values. First, because it might be expected that those responding to the survey might be more intensive ADS users than non-responding users – although, most notably, responses about their impression of data use by others in their field does not reflect this (User Qs 21, 26 and 31, Appendix 1). Second, because it expresses the impact in time and money

⁷ There may also have been a tendency to pick a central answer from the efficiency percentage range offered.

equivalents, rather than quality impacts, such as completeness, appropriateness, etc., thereby focusing on one dimension of efficiency. Conversely, estimates are based on regular users and current activity times, rather than on all users during the year and the implied activity times before the efficiency impact. In the end, the efficiency impacts are likely to be more about how much gets done, and how well, in a given time, rather than the amount of time spent. As such, we are not talking about real hours or real pounds, but rather using hours and pounds as proxy measures of the value of the efficiency impacts.

5.2.5 Return on investment in Archaeology Data and Services

It is possible to explore some scenarios relating to the potential impacts of ADS data and services on returns on investment in the data. There are a number of data elements required for such an analysis, including: annual investment in the research data and services; average returns to that investment; and the level of and increase in access and use resulting from ADS activities. The user and depositor surveys, together with ADS operational data and reports provide a foundation for estimates:

- Multiplying the acquisition-weighted mean of the reported annual data creation costs by the average annual number of new data deposits made, suggests annual data creation costs of more than £13 million. Acquisition-weighted mean depositor preparation and deposit costs amount to around £465,000 per annum, and ADS operation costs to some £698,000 million per annum, suggesting total data and service investment approaching £15 million per annum.
- There is an extensive literature on returns to R&D, which, while varied, suggests that returns are high - typically in the region of 20% to 60% per annum (Bernstein and Nadiri 1991; Griliches 1995; Industry Commission 1995; Salter and Martin 2001; Scott et al. 2002; Dowrick 2003; Shanks and Zheng 2006; Martin and Tang 2007; Sveikauskas 2007; Hall et al. 2009). Much of this literature relates to the natural sciences and one might imagine that average returns to archaeology may be lower. Hence, to be conservative, we explore a range of returns at and below the lower bound average returns to R&D expenditure reported in the literature (i.e. 20% and 5%).
- The other issue is what impact ADS data and services have on access and use of the data hosted and delivered. Some 44% of respondents to the user survey indicated that they could not have obtained the data in any other way if ADS had not existed (User Q15, Appendix 1). All of these represent additional use. However, responses to Q16 (Appendix 1) suggested that, of these, 61% thought the data was beyond the scope for them to collect themselves. So, for preliminary estimation, we have assumed that the remainder could have collected the data themselves in some form and that this is typical of ADS users. This suggests that at least 44% of ADS use may be additional use, of which 27% ($44\% \times 67\%$) could not have collected the data themselves and the remaining 17% ($44\% \times 33\%$) saved data re-collection costs (e.g. would otherwise have re-collected all the data they thought they were able to).

We proceed on the basis of these data, estimating the increase in annual return on investment due to additional use. As these returns are recurring during the useful life of the data we use a simple Perpetual Inventory Method to estimate the overall value of the impacts. Drawing on preliminary

work on the UK R&D Satellite Account (Evans et al. 2008) we depreciate publicly-funded research data at 5% per annum, and following the lead of the US R&D Satellite Account (Sveikauskas 2007) we set the useful life of the data/knowledge created each year at 30 years. For preliminary estimation we distribute the returns normally over year 1 through year 9. Applying a 3.5% discount rate to estimate net present value (NPV) (Green Book 2011), we then model the recurring returns as follows.

Additional use

At the lower-bound average 20% return on data investment, if 44% of ADS use is additional use (i.e. the share of respondent users who could not have got the data elsewhere) the implied increase in returns on one-year's data and services investment would be £9.7 million over 30 years (NPV), and at the lower 5% average return the implied increase in returns would be £2.4 million (NPV) [given non-sunk data services costs of around £1.2 million]. This suggests that there may be a 2-fold to 8-fold return on investment arising from the additional use of the data facilitated by ADS services.

Table 5.6: Return on Investment to Additional Use Facilitated by ADS

| | | | |
|--|--|--|---------------------------------|
| | | | 14,603,270 |
| | | | ÷ |
| | | Weighted Mean of Total Costs <i>Creation + Deposit + Operation</i> | (170,757 x 44%) |
| | | | x |
| | | | (5% to 20%) |
| | | | = |
| | | ÷ | |
| | | Number of Additional Uses | (£2,422,459 to £9,689,836) |
| | | x | (return from 1 years investment |
| | | Average Return | over 30 Years NPV) |
| | | | |
| | | ÷ | ÷ |
| | | Non-Sunk Data Services Costs <i>ADS Operational + Deposit Costs</i> | 1,162,892 |
| | | | = |
| | | | 2.1 to 8.3 |

Source: Authors' analysis.

Additional use by those who could not (re)collect the data

At the lower-bound average 20% return on data investment, if 27% of ADS use is additional use (i.e. the share of user respondents who could not have got the data elsewhere or re-collected it themselves) the implied increase in returns on one-year's data and services investment would be £6 million over 30 years (NPV); and at the lower 5% average return the implied overall increase in returns would be £1.5 million (NPV). Given non-sunk data services costs of around £1.2 million per annum, this suggests a 1.3-fold to 5.1-fold return on investment.

Of course, if the other element of additional use, namely that by users who could not have got the data elsewhere but could have (re)created it themselves, is taken into account, then the implied cost

saving of not (re)creating it should be added to the implied additional returns to investment. That would suggest the addition of up to around £1 million per annum in saved (re)creation costs (Box 5.3).

Box 5.3: Proxy indicators of value to users: re-creation costs

Indicators of value to users include the implied cost savings resulting from reuse (i.e. the costs of data collection/creation that are saved multiplied by the number of additional uses). Of course, this is an upper bound estimate of what the users may be saving as they may alternatively forego use. Nevertheless, just to explore a possible reuse related cost savings: 44% of user respondents said they could not have obtained the data they downloaded in any other way, of which 39% said they could not have re-collected the data, so perhaps 17% of respondents could have collected the data.

If this were characteristic, and 17% of all users were able to re-collect the data and had done so instead of accessing ADS during the last year, at the average of depositor reported data creation costs, less ADS operation, user access and deposit costs, it might have cost them around £1 million.

Of course, some of the accesses are repeat accesses to the same data, it is more likely that they could have re-collected the smaller and cheaper-to-create data collections than the larger and more expensive-to-create data collections, and some would have foregone the data if faced with re-creation costs, so this is likely to be very much an upper bound cost saving.

Source: Authors' analysis.

Hence, ADS data and services facilitate additional use which realises additional returns to the research and data creation/collection activities underpinning the data it hosts, that, to an approximation, are worth around £2.4 million to £9.7 million from one year's investment. As such, ADS exhibits a 2 to 8 fold return on investment.

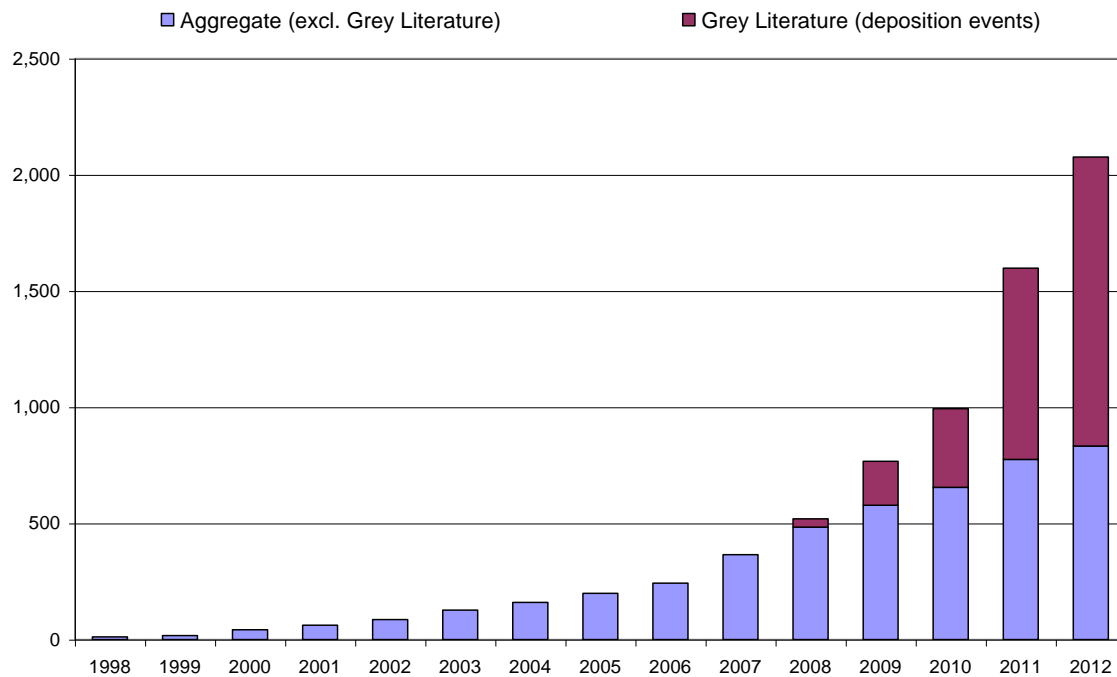
5.2.6 Longer term collection and preservation and the accumulation of value

For simplicity, we have presented the economic analysis above as annual costs, benefits and impacts at current levels of activity circa 2012 and current prices. However, there is also the issue of the potentially increasing value of the accumulating resource (i.e. at whatever current use levels, the value of ADS is likely to grow as the collection grows, just as a bigger library might be thought to be worth more than a smaller library even if they both have the same number of users and uses). As our study is a snap-shot, based on a single survey, we can say little about this dynamic and how it has evolved over time. However, ADS have provided historical data for their data deposits and uses and related operational costs back to 1997, which enables us to explore some indicators of cumulative value.

Cumulative investment and use value can be estimated from historical data, although estimates depend on the use of current average per deposit and per use costs. An advantage of this is that it automatically expresses the costs and values in current prices, but the disadvantage is that we cannot know how deposit and use costs have changed over time (e.g. as the ADS site is improved and becomes easier to use and both depositors and users gain more experience with it). Nevertheless, such estimates are indicative.

Figure 5.1 shows how data deposit events at ADS have grown over time, with deposit events increasing from 11 in 1998 to more than 2,000 in 2012. The addition of grey literature to the collection from 2008 onwards has made a substantial difference, with more than 14,000 items added in an estimated 1,244 individual deposition events.

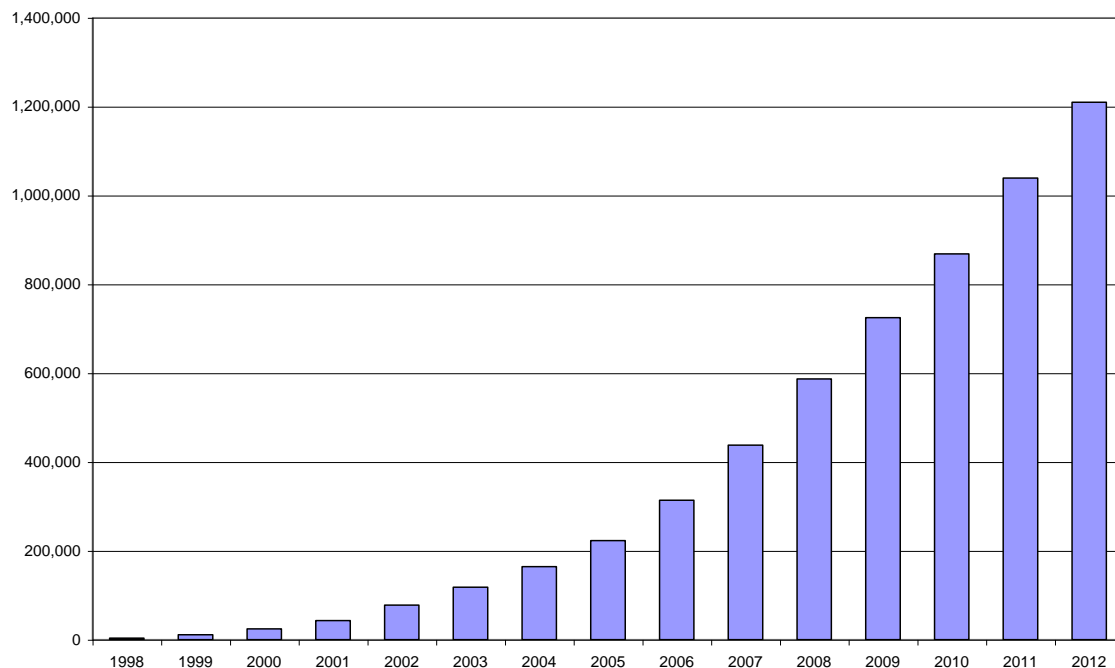
Figure 5.1: Cumulative ADS deposits (deposition events), 1998-2012



Source: ADS, Authors' analysis.

Figure 5.2 shows use trends over the same period. Adjusting historical page view data to accesses and excluding bounces (See Box 5.1, above), shows uses increasing from 3,400 in 1998 to around 171,000 in 2012, and cumulatively exceeding 1.2 million.

Figure 5.2: Cumulative website accesses, 1998-2012



Source: ADS, Authors' analysis.

At current time/costs, the *cumulative investment value* of ADS data and services (i.e. the sum of operational and deposit costs) is around £7.8 million, with depositors investing around £2.2 million. *Cumulative use value* (i.e. user access costs) amounts to around £10 million.

Another way of thinking about the value of a collection is *replacement value*, which is often used to value library collections (e.g. for insurance purposes). Replacement value can be estimated as the cost of creation plus the cost of deposit/ingest multiplied by the number of items held. Treating historical ADS operational costs as sunk costs, replacement value would be around £72 million. However, elsewhere we have treated data collection/creation costs as sunk costs (i.e. assuming that the data would have been collected/created whether ADS existed or not). To do so here would suggest a replacement value of around £2.2 million, although this is to assume that there would be no recreation costs involved in obtaining the historical data.

It is also possible to explore the cumulative return on investment arising as a result of the additional use facilitated by ADS data and services. Using the same basic parameters as the return on investment estimates described above and assuming that the current level of additional use as a share of total use has prevailed throughout the operational life of ADS, we estimate the cumulative returns from additional use at between £1.5 million and £6.1 million.

At current prices, we estimate cumulative costs and benefits over 30 years from 1997 through to 2026, and find that:

- At the lower bound 20% average return, cumulative benefits from additional use would amount to around £37 million and with cumulative costs of around £7.8 million the benefits would be 4.7 times the costs; and
- At a 5% average return, cumulative benefits from additional use would amount to £9.7 million and the benefits would be 1.2 times the costs.

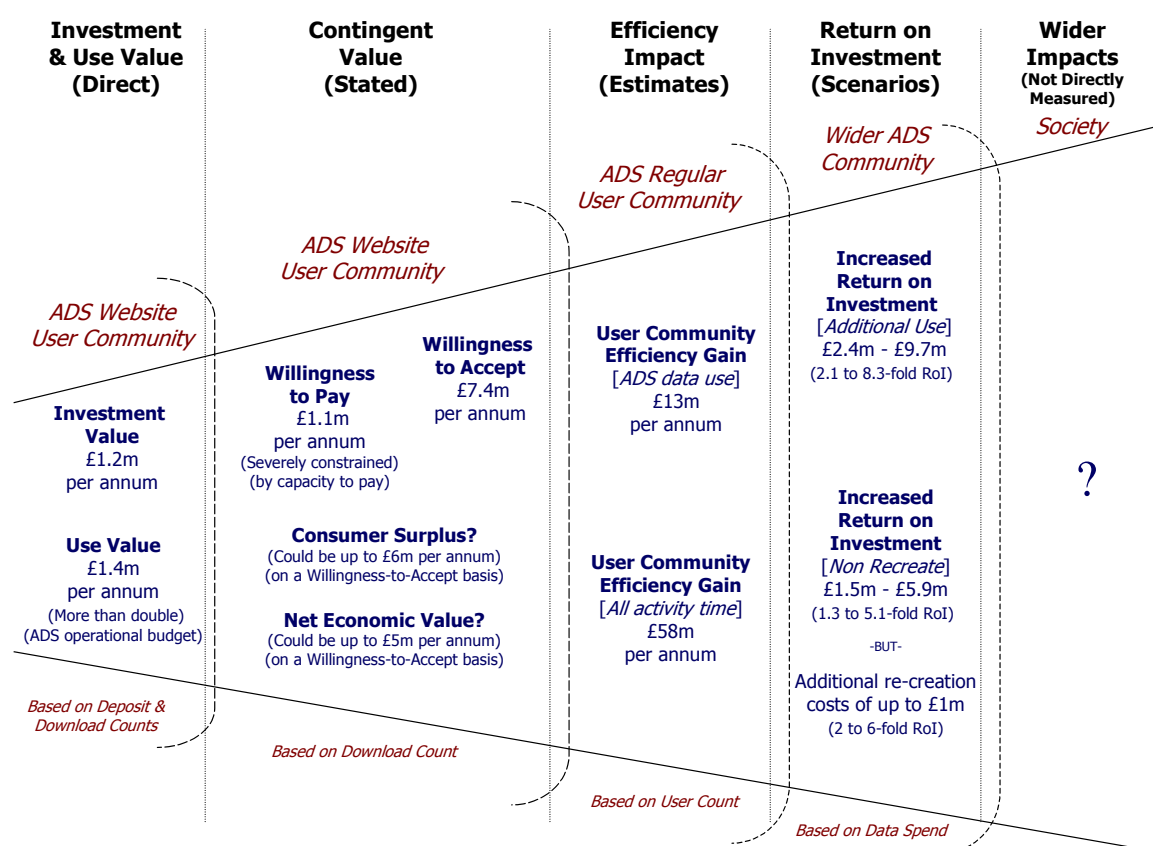
While no more than indicative these longer term cumulative values suggest an increasing value as the ADS data collection grows and the service develops. For example: the current use value of £1.4 million per annum is 2.3 times the average annual use value over the last 15 years (in today's prices); and the current investment value of £1.2 million per annum is 2.4 times the average annual investment value over the last 15 years. Similarly, the estimated return on investment from the additional data use facilitated by ADS services during the last year at 2.1 to 8.3 times the non-sunk operational and deposit costs, is around double the estimated cumulative return on investment (i.e. 1.2 to 4.7). Hence, it does appear to be the case that the value of ADS data and services has been increasing as the collection has grown and the service developed.

5.2.7 Summary of the quantitative economic analysis

While there are many limitations in the data, our survey results and ADS operation data provide a basis for estimation. Figure 5.3 summaries these results. It shows a direct investment and use value to the ADS user community of £1.2 million to £1.4 million per annum at current prices and levels of activity. Willingness to pay is an expression of value by the users, who reveal that they value their access at around £1.1 million per annum, despite resource constraints and limited capacity to pay.

When capacity to pay is limited the amount that users would be willing to accept in return for giving up their access to ADS for a year can be a better indicator of the value they place on it, as it is not constrained by their capacity to pay. Looked at this way, ADS data and services are worth around £7.4 million per annum. If one adds the cost of access to what users are willing to pay it indicates what they do and are willing to pay, which is more than £2.5 million per annum. At £7.4 million per annum willingness to accept is higher, and the difference relates to their financial constraints and incapacity to pay as much as they think it is worth (e.g. they will/can only pay one-third what they think it is worth), and, perhaps, to other forms of non-use value, such as 'existence value' (e.g. people value having ADS even if they do not use it much and so would not be willing to pay as much to use it as they think it is worth – just as people might value having a national park without ever going there).

Figure 5.3: The value and impacts of the ADS research data service



Source: Authors' analysis

The contribution of ADS data and services to its user community can be seen in terms of its impact on their research, teaching and studying efficiency. We found that the total estimated efficiency impacts of ADS data and services among its user community might be as much as £58 million per annum at current activity times. However, these impacts seem high and it seems that some respondents interpreted the question as relating to the efficiency impact on their studying/teaching/researching time spent with ADS data and/or data from all sources, rather than their total studying/teaching/researching time – which had been intended. Taking this into account, the lower bound implied efficiency time saving impacts would be around £13 million per annum. This is still a very substantial impact – some 5 times the sum of operational, depositor and user access time costs.⁸

Exploring some scenarios relating to the potential impacts of ADS data and services on returns to investment in the data, we found that ADS data and services facilitate additional use which realises potential additional returns to the research and data creation/collection activities underpinning the data it hosts, that, to an approximation, are worth around £2.4 million to £9.7 million from one-year's investment. As such, ADS exhibits a 2 to 8 fold return on investment.

⁸ In fact, the implied lower-bound ADS data use time efficiency impact is equivalent to around £1,201 per regular (i.e. weekly or more regular) user per annum. This is close to the (un-weighted) mean willingness to accept of £1,282 per annum, suggesting some convergence in the various perceptions and forms of value.

6 QUALITATIVE APPROACHES AND ANALYSIS

In addition to the quantitative economic analysis, we undertook qualitative analysis based on user and depositor surveys and interviews. This provides the opportunity to check that the quantitative and qualitative findings are in accord and sheds light on the dimensions and nature of the value of ADS data and services.

6.1 Surveys

Both the user and depositor surveys included a number of questions seeking respondents' qualitative perceptions of value.

6.1.1 Depositors of data with ADS

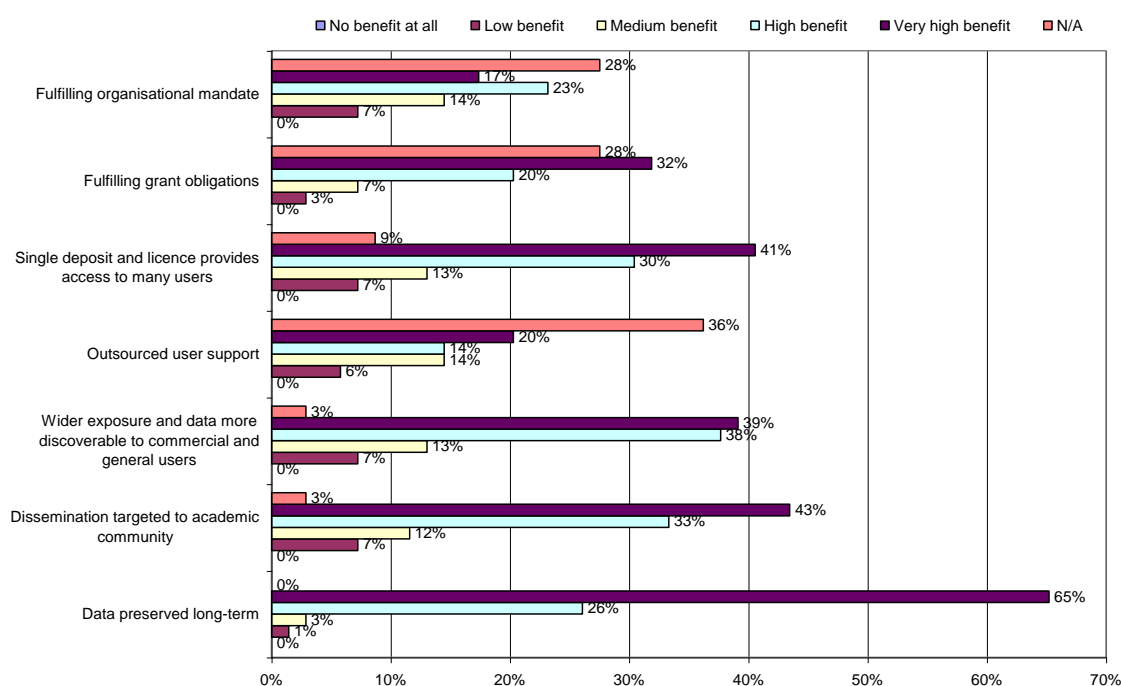
The main affiliations of depositor respondents were: higher education (33%), local government (22%), commercial depositors (17%), private unaffiliated individuals (10%), and non-government organisations (8%). Ninety-five per cent were based in the UK, with the remainder based in the United States, Canada and Sweden.

Grey Literature (i.e. unpublished reports) was by far the most common data type deposited by survey respondents during the last three years, with a total of 419 deposits/updates per annum, followed by Project Archives with 70, Local/National Historic Environment Records with 25, and Journals and Series (CBA Research Reports, PSAS, etc.) with 11. Overall, depositors reported making 179 deposits/updates per annum – a mean of 3 per depositor per annum (median 0.67). Similarly, 41% of depositors reported that Project Archives were the most recent type of data deposited, and 38% Grey Literature (i.e. unpublished reports). Few had most recently deposited other data types. For 79% of respondents their most recent deposit was an initial or one-off deposit, while for the remaining 21% it was an update.

The main benefits from providing data to ADS reported by depositors included:

- Data preserved long-term, which was rated as a high or very high benefit by 91% (rating average 4.42);
- Dissemination targeted to academic community, rated high or very high by 77% (rating average 4.00);
- Wider exposure and data more discoverable to commercial and general users, was also rated high or very high by 77% (rating average 4.00);
- Single deposit and licence provides access to many users, rated high or very high by 71% (rating average 3.78);
- Fulfilling grant obligations, rated high or very high by 52% (rating average 2.68); and
- Fulfilling organisational mandate, rated high or very high by 41% (rating average 2.38) (Figure 6.1).

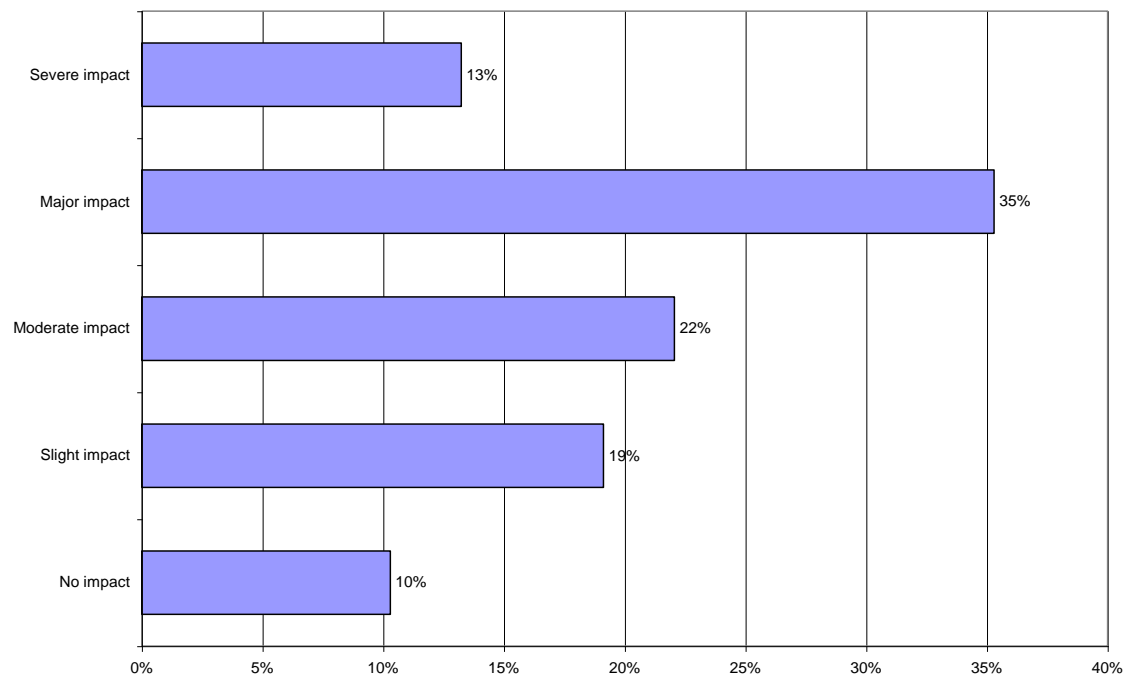
Figure 6.1: Depositors' perceived benefits from providing data to ADS (N=69)



Source: ADS Survey, Authors' analysis

Forty-eight per cent of depositors reported that not being able to provide data to ADS would have a severe or major impact on their work, and a further 41% said it would have a moderate or slight impact. Just 10% said it would have no impact (rating average 3.22) (Figure 6.2). Comparing answers by affiliation and roles, it seems that those saying that it would have no impact were often required to deposit, and may have been assuming that the requirement would be dropped if ADS no longer existed.

Figure 6.2: Impact on depositors' work of not being able to provide data to ADS (N=68)

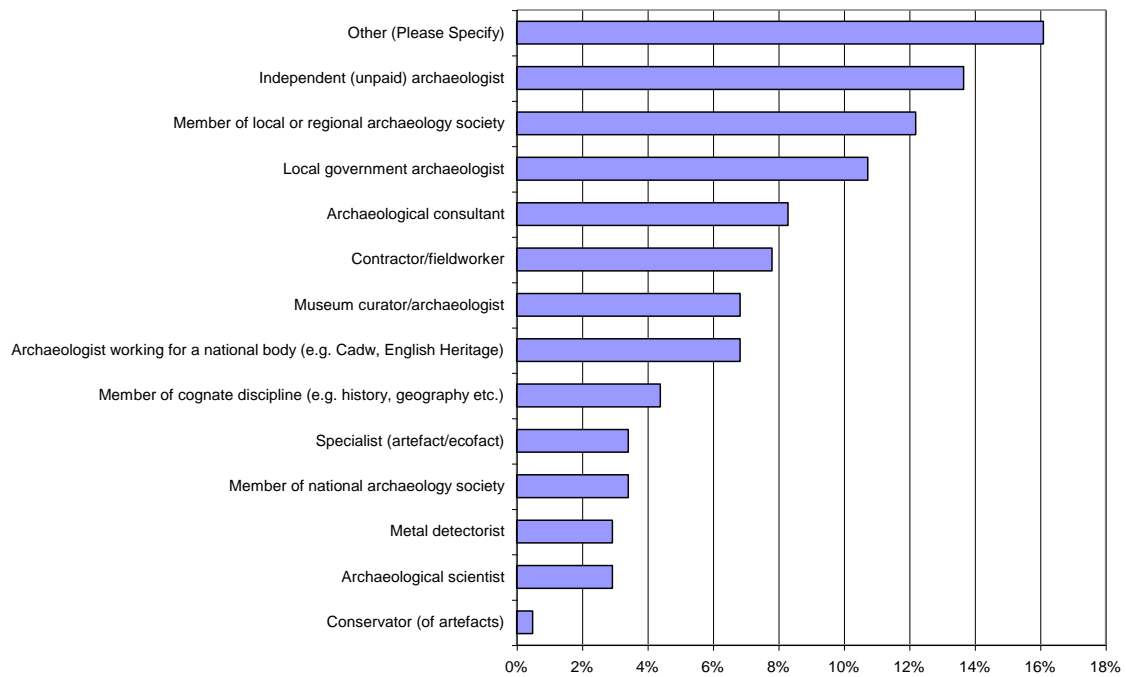


Source: ADS Survey, Authors' analysis

6.1.2 Users of ADS data and services

The main affiliations of ADS user respondents included higher education (29%), and private/unaffiliated individuals (27%). Respondents were then streamed into separate questions asking their role within their affiliation, according to whether they were in higher and further education or other affiliations. Despite being offered a long list, 16% of non-higher and further education respondents chose "other" and cited a wide range of roles. A wide range of private individual, public and private sector roles were also selected (Figure 6.3).

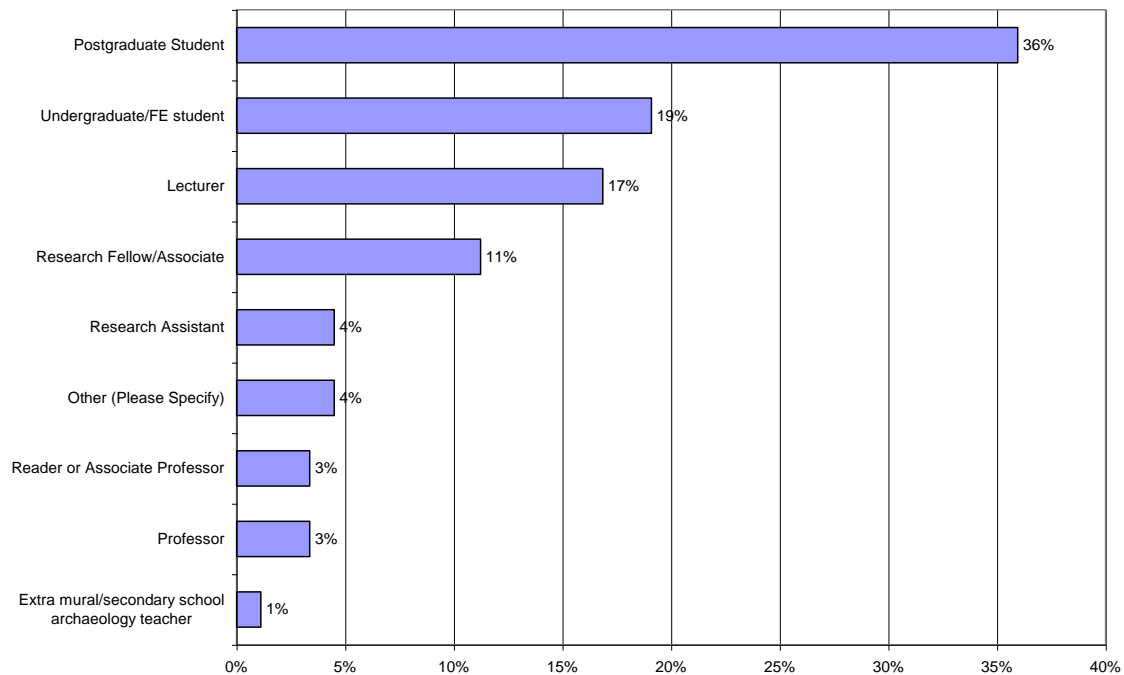
Figure 6.3: Main role within affiliation for non-education sector ADS user respondents (N=205)



Source: ADS Survey, Authors' analysis

Within the education sector, 36% of respondents were postgraduate students and a further 19% undergraduate or further education students. The shares of responding research and teaching staff reflect their workplace shares, with more lecturers (17%) and research fellows (11%) than readers (3%) and professors (3%) (Figure 6.4).

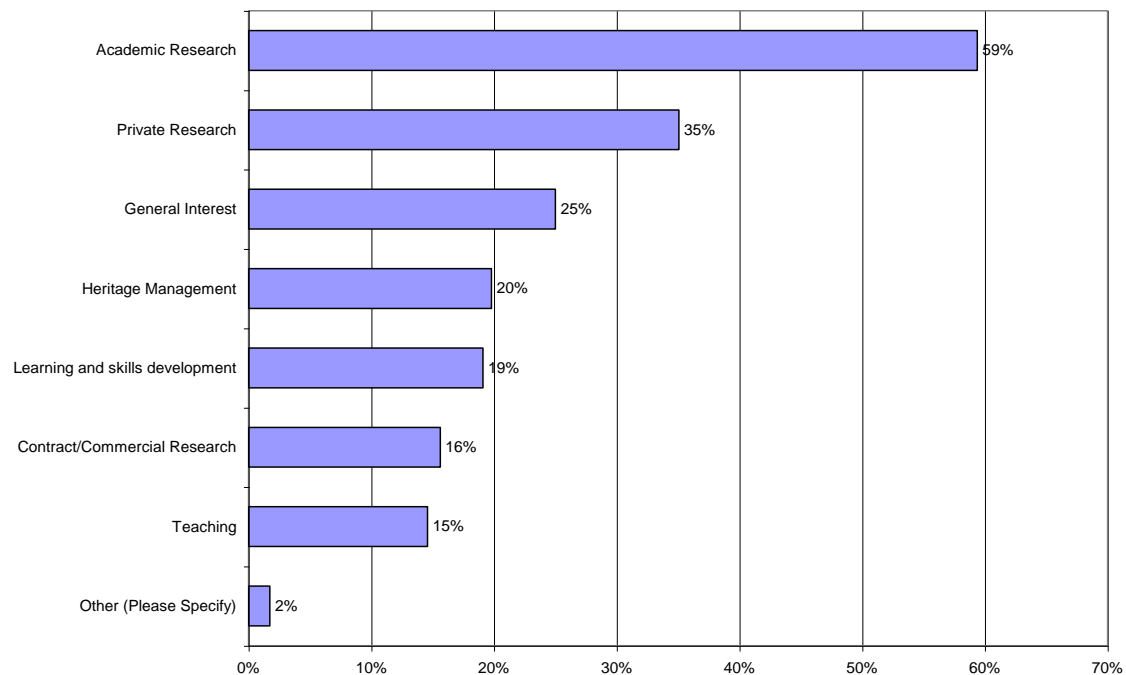
Figure 6.4: Main role within affiliation for education sector ADS user respondents (N=89)



Source: ADS Survey, Authors' analysis

Research is the main purpose of use of ADS among respondents, with 59% reporting their main purpose as academic research, 35% private research, and 16% contract or commercial research. Learning and skills development (19%) and teaching (15%) are also widely cited as a main purpose, with general interest (25%) and heritage management (20%) also providing motivations for use (Figure 6.5). Many respondents ticked multiple options, such as combining academic research, contract research and teaching (there were a total of 553 responses).

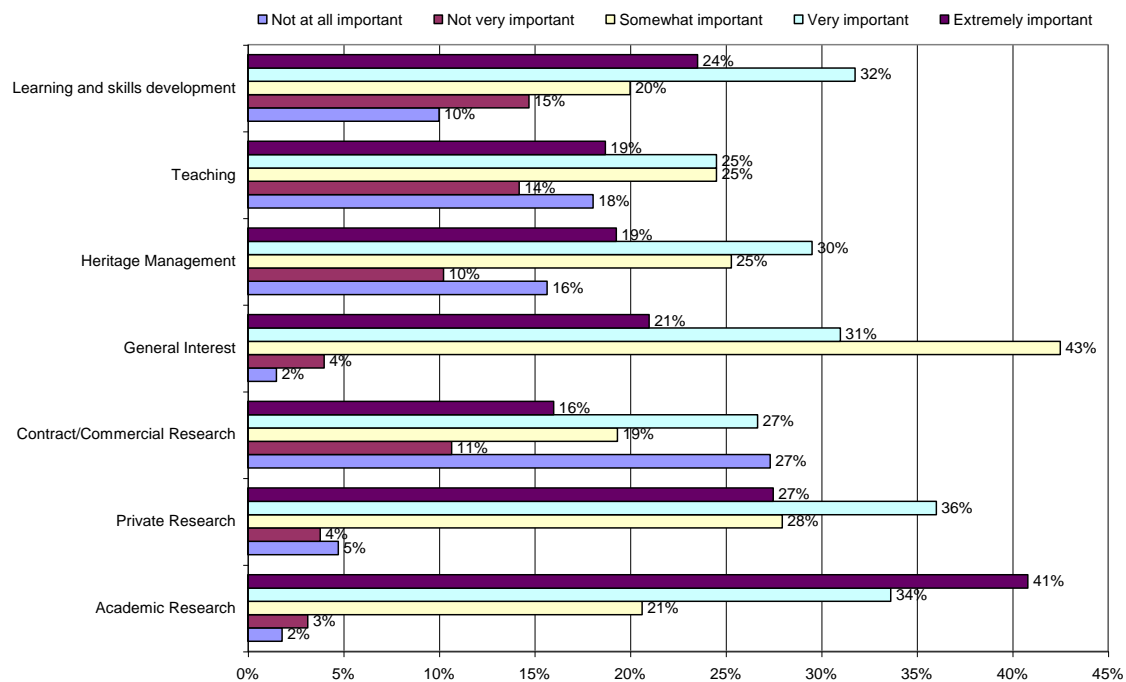
Figure 6.5: The main purposes of respondents use of ADS (N=288)



Source: ADS Survey, Authors' analysis

Seventy-four per cent said that ADS was very or extremely important for their academic research, 64% said it was important for their private research, and 55% said it was important for their learning and skills development – reflecting the shares of academic, private and student users (Figure 6.6). Rating averages reflect the importance of academic and private research (4.09 and 3.78, respectively), and general interest (3.66).

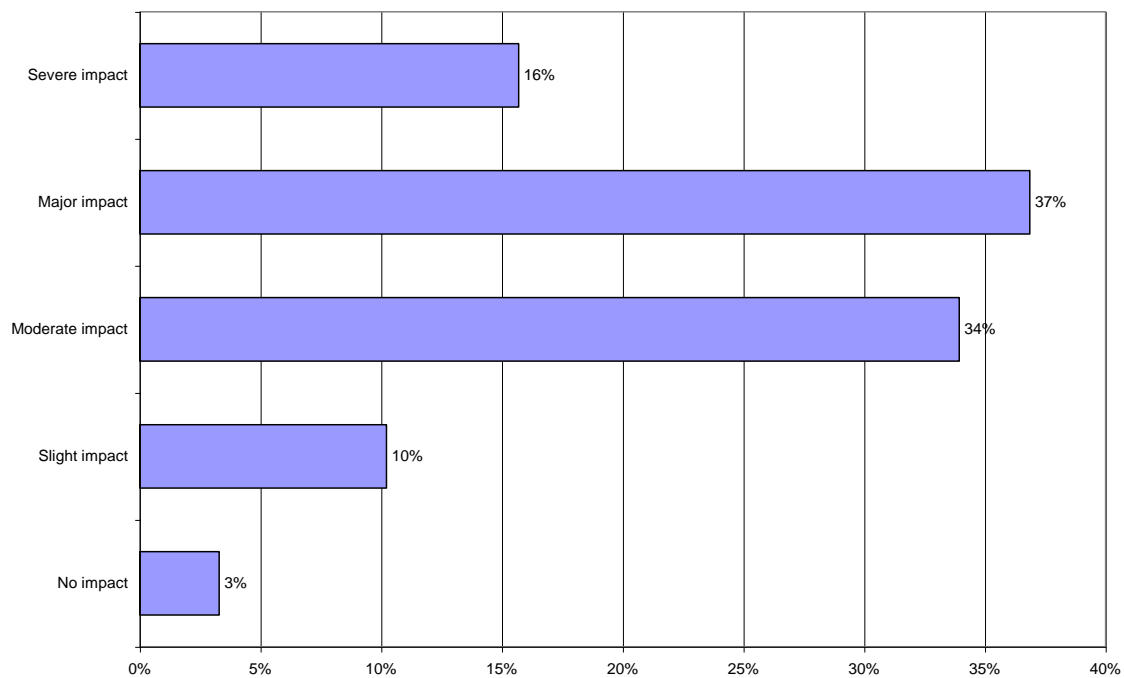
Figure 6.6: Importance to users of data and services from ADS for different areas of their work (N=276)



Source: ADS Survey, Authors' analysis

Asked the importance question in reverse, 53% said that not being able to access ADS data and services would have a severe or major impact on their work, and a further 44% said it would have a moderate or slight impact (Figure 6.7). Just 3% said that not being able to access ADS would have no impact on their work. Those saying 'no impact' were very infrequent users (median 6 accesses per annum). It is also worth noting that the question asked about impact on respondents' work, and 4 of the 9 saying 'no impact' were private/unaffiliated individuals, 2 described themselves as 'independent/unpaid archaeologists', and one was a research assistant in higher education. Hence, some may not have thought there could be any impact on their work.

Figure 6.7: Impact on respondents' work of not being able to access data and services from ADS (N=274)



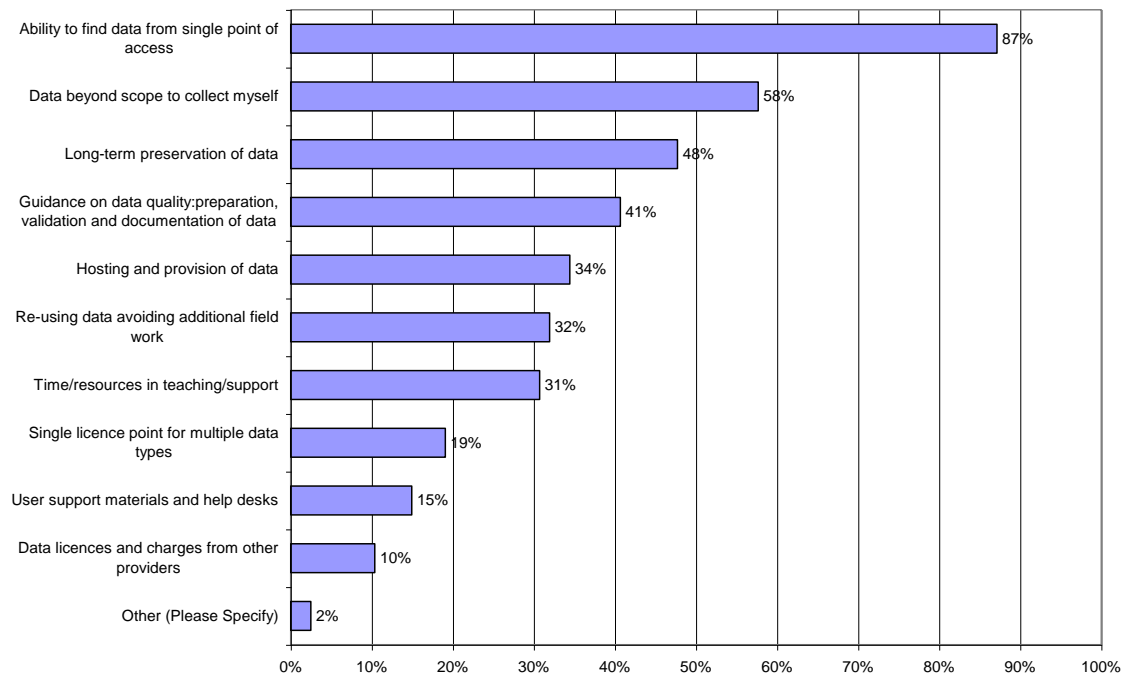
Source: ADS Survey, Authors' analysis

The question also allowed comments, of which there were 46. The major themes included:

- Alarm – their work would be impossible without access to ADS (N=20);
- Concern – they could use alternative sources, but would incur extra time and financial costs (N=10); and
- Dismay – relating to the strategic implications for the profession as a whole (N=9).

By far the most widely cited factor contributing to savings was the ability to find data from single point of access (87%) (Figure 6.8). That the data were beyond their scope to collect themselves (58%), long-term preservation of data (48%), and guidance on data quality through preparation, validation and documentation of data (41%) were also widely cited areas of benefit. These responses reveal significant efficiency impacts in terms of time and cost savings.

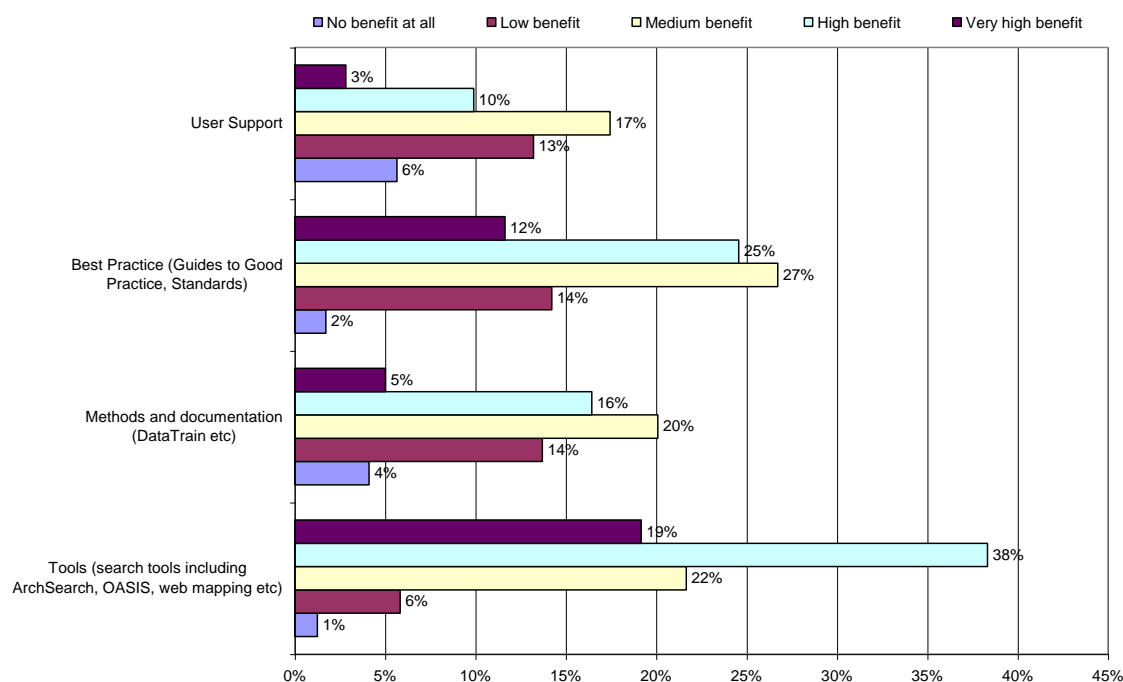
Figure 6.8: Areas of savings for respondents resulting from accessing ADS (N=241)



Source: ADS Survey, Authors' analysis

Fifty-eight per cent said that they derived a high or very high benefit from tools (e.g. search tools including ArchSearch, OASIS, web mapping, etc.), 36% said they derived high or very high benefit from guides to best practice and standards, and 21% said they derived high or very high benefit from methods and documentation (e.g. DataTrain, etc.). Very few respondents reported deriving no benefit from any of the items or services they had used (Figure 6.9).

Figure 6.9: Extent of other benefits for respondent ADS users (N=249)



Source: ADS Survey, Authors' analysis

6.2 Interviews

Fifteen interviewees were selected to provide a cross-section of ADS stakeholders. A semi-structured interview was conducted using a pre-defined questionnaire. Table 6.1 presents a synthesis and provides an anonymised summary of responses to each of the questions highlighting key messages.

Table 6.1: Key messages from the interviews

| | Key Message | Detail |
|---|--|--|
| 1 | ADS is the only place to deposit which ensures curation in perpetuity, and open access, and is setting the standard for good practice in these areas. | While some museums and universities accept digital data, there is doubt about whether they have the required understanding of long term preservation. Many in the profession are trying to persuade Local Authorities to delegate this responsibility to the ADS on a national basis. |
| 2 | ADS has been instrumental in changing the professional landscape. | Funders (EH, AHRC etc.) are mandating deposit with ADS as a condition of grant, and this is helping to educate and acclimatise commercial contractors. Contractors are able to benchmark their own work by viewing other project archives and this drives up standards overall. The Aggregates Levy Sustainability Fund (ALSF) was a major digital project that ran from 2002-2010. About £25-28 million from this “green tax” went on the historic environment. Soon after the start, EH realised they needed to ensure long term accessibility so began stipulating that all funded projects should deposit a digital copy with ADS. Other grants programmes followed suit. It is important to funders that ADS migrates data. |

| | | |
|----|--|---|
| 3 | <i>Deposit with ADS is seen as the future by much of the profession.</i> | Information is being digitised and electronic data is being generated in increasing quantities from field activity. Electronic publication facilitates reuse and reinterpretation of the raw data where traditional print publication is by nature selective, so it would be a retrograde step to have to return to this model. |
| 4 | <i>Funders and depositors see ADS as important for dissemination, impact, reaching the widest possible audience, and ensuring a long term legacy for their work.</i> | Their use of ADS is not solely for convenience and efficiency – it enhances their own reputation to be seen to be following best practice. |
| 5 | <i>The business model of paying once for in perpetuity access is one that appeals to contractors.</i> | Clients are happy to pay one-off but not ongoing costs, so the ADS business model is well aligned to the commercial world. |
| 6 | <i>The ADS is everywhere held in very high esteem.</i> | It is seen as a sector-wide asset, an exemplar of good practice both nationally and internationally, and is widely praised for its professionalism, efficiency, openness and collaborative approach. It has also been an effective mediator between academic and commercial archaeology, and government and curatorial sectors. Due to its spatial nature, archaeological data is particularly varied and complex and ADS staff are valued for their technical expertise in handling this. Their role in shaping archaeological computing and standards is appreciated, as is their visibility at conferences, where they have done much to raise awareness of issues around digital archiving. |
| 7 | <i>Users appreciate the ADS as “an integrated and comprehensive source”.</i> | The website search interface and navigational structure is highly thought of by the more technical users, but perhaps less liked by some others who prefer to get HER data from the Heritage Gateway. |
| 8 | <i>The Grey Literature library is unique, unparalleled and seen as critical to the future of UK Archaeology.</i> | As it gathers critical mass this encourages more use and more deposit, and the cumulative effect will continue, magnifying the capacity for future impact. Individual files have more value and visibility as part of the ADS Grey Literature Library than they could if held in separate repositories. |
| 9 | <i>ADS produces savings in time for researchers.</i> | Costs to do the same level of research without ADS would be incalculable, so many would probably not do this in practice. Savings are therefore in terms of quality and quantity of research outputs from the resource available. |
| 10 | <i>There are both time and financial savings arising from not having to go to different sources for the information as a result of using ADS.</i> | Depositors also save time and money on data management and infrastructure and have benefited from help with data restructuring and metadata creation. |
| 11 | <i>“If we didn’t have the ADS we would have to invent it”.</i> | It would be impossible to search nationally; users would have to seek information from a wide range of diverse sources, creators and individual HERS, often relying on personal contacts and goodwill. They might not even be aware of what resources existed. Depositors would |

| | | |
|----|---|--|
| | | have to find alternative repositories or set up their own; they would view this prospect with alarm and trepidation due to prohibitive costs and their own lack of specialist staff expertise and availability. They would therefore have concerns about long term security of digital archives, fragmentation of the resources and eventual diversification of standards, ultimately leading to loss of access. |
| 12 | <i>ADS has strategic importance.</i> | ADS has strategic importance for heritage management and the planning framework; the government (House of Lords Science and Technology report no. 5); the HE sector (The QAA benchmark statement for Archaeology degrees identifies transferable skills from data analysis), and the profession as a whole. By changing the way research is conducted ADS is also helping to shape the research agenda. |
| 13 | <i>Interviewees usually struggled to put a value on cost savings or wider benefits although they recognised they are likely to be significant.</i> | <p>Some interviewees provided figures e.g.</p> <ul style="list-style-type: none"> • C. £120 million is spent annually on investigation, and the number of events/investigations per year is in the thousands. So much data is being generated electronically from excavation, fieldwork and Local Authorities, ADS is absolutely essential. • It was estimated that for one major digital project which ran from 2002-2010, ADS received £220,000 to secure about £20 million worth of data (as measured in project cost). • ADS is significantly cheaper than Local Authorities – an extra £250 on an £800 deposit cost is levied by some Local Authorities, so ADS can cost 25% less. The fact of ADS's existence saves depositors having an archive of their own – setup costs would be in £10,000s, as would running costs even for a single organisation. • If all HERS had put their Grey Literature on ADS this would have saved thousands in travel and subsistence. For example, costs for one project could have been substantially cut; the travel part of the grant for one specific project was about £250,000, and this could have been saved. |

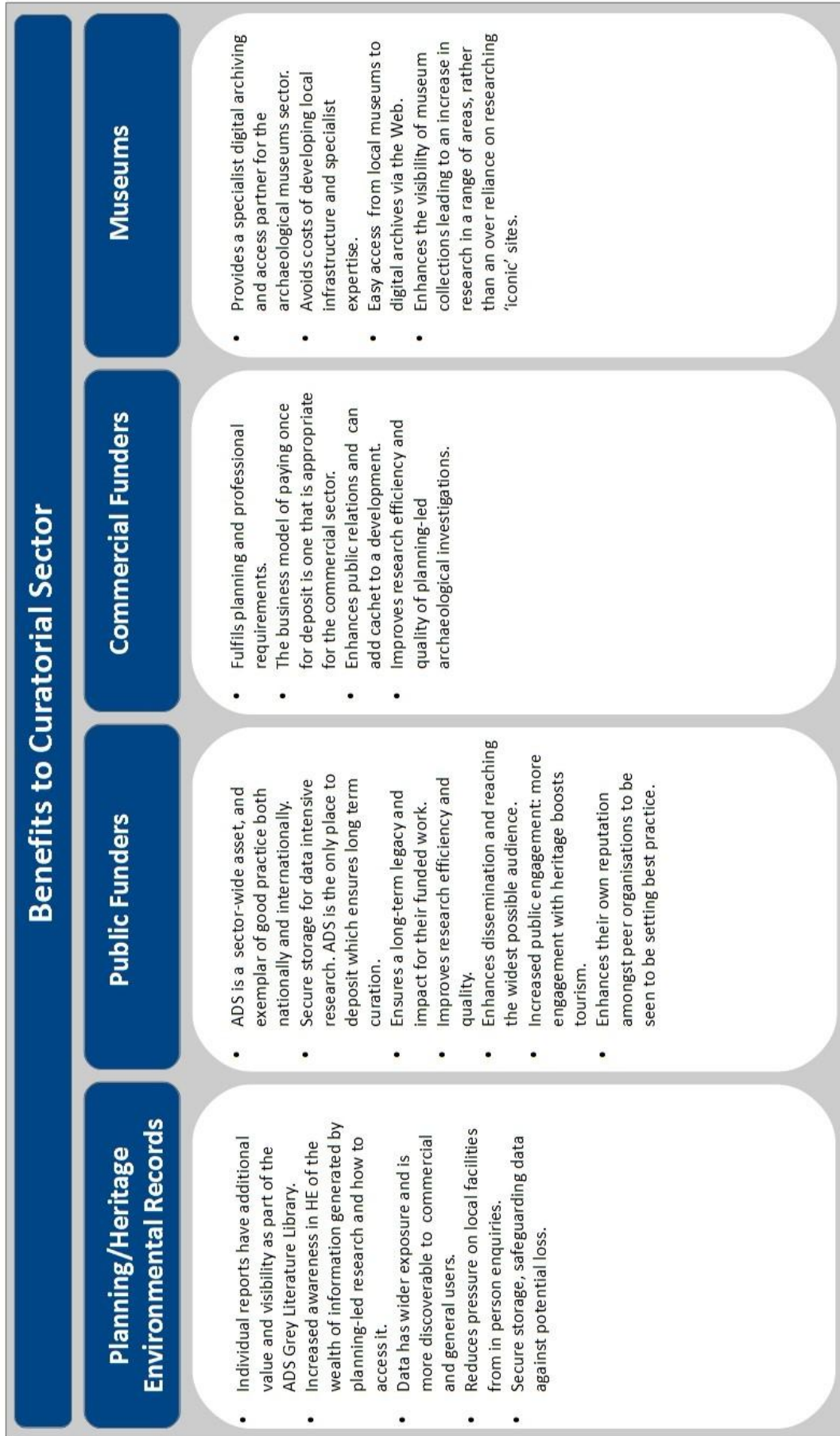
Source: Authors' analysis

6.3 KRDS Stakeholder Benefit Map for ADS

We have based the summary list of Stakeholder benefits on approaches and formats developed by Keeping Research Data Safe (KRDS), particularly aspects of the KRDS Benefits Framework (see section 4.2.5). The benefits identified are drawn from the desk research, interviews and surveys, and arranged by key stakeholders identified at the stakeholder focus group (see section 7). Benefits are listed for specific partners: first the curatorial sector, sub-divided by its key stakeholder groups (planning/heritage environment records, public funders, commercial funders, and museums respectively); and secondly the research sector, sub-divided by its key stakeholder groups (universities, archaeological contractors, professional and learned societies, and private individuals).

Table 6.2: Stakeholder benefits for the curatorial and research sectors

Source: KRDS and Authors' analysis





7 STAKEHOLDERS' PERCEPTIONS OF VALUE

As part of the ADS Impact study, a focus group workshop was held in York on 21st November 2012 to present and discuss emerging results from the study with a range of ADS stakeholders. The aims of the workshop were to seek stakeholder feedback on the emerging results, establish any change of perception of the ADS amongst participants as a result of the study, and seek their views on how the study results might be presented to the archaeological community and its funders.

Invitations were sent to a range of sector representatives, but not all could attend. Eleven delegates attended the workshop, of whom four were from the Local Authority sector, three from National Authorities, one from Universities, one from the Commercial sector, one shared university/commercial sectors, and one from Publishing.

The agenda for the workshop was as follows:

| | |
|---------------|--|
| 13.00 – 13.20 | Welcome (Julian Richards) and Introductions (Neil Beagrie + participants) |
| 13.20 – 13.45 | Workshops Aims (Neil Beagrie) and Initial Feedback (participants) |
| 13.45 – 14.45 | Study methods and Initial Results (Neil Beagrie) and Questions |
| 14.45 – 15.00 | Tea/coffee break |
| 15.00 – 15.30 | Discussion – Value and Impact of ADS (participants – facilitator Neil Beagrie) |
| 15.30 – 15.45 | Final Feedback (from participants via feedback forms) |

Facilitated discussion at the event and a structured feedback form were used to assess post-dissemination perceptions of value of the ADS. This is synthesised below. The workshop was held using “Chatham House Rules”: all feedback has therefore been anonymised.

7.1 Initial Stakeholder Feedback

The initial feedback session was structured as a discussion focussing on identification of key stakeholders of the ADS that should be made aware of the study findings. Participants suggested the following key stakeholder groups:

- Curatorial sector involved in preservation by record (planning/HER);
- Research funders (HE and others, such as HLF) both “willing” funders and developers required by planning to fund – it was noted those forced to fund are most likely to complain at any costs and most in need to be aware of likely benefits;
- Users including university academics and students, contract archaeologists, private individuals, learned societies, national heritage bodies;
- Other archives and repositories, particularly museums with archaeological collections and institutional repositories in universities – it was noted there is a wide range of un-deposited digital archives and a missed opportunity cost that could be quantified;
- European and other international peer repositories and organisations;
- Publishers and learned society journal editorial boards for linked data and articles;

- Other digital archives in the humanities and beyond with a relevant professional interest in the study. It was felt important not just to sell the ADS, but also the benefits of digital archiving generally.

7.2 Final Stakeholder Feedback

The final feedback was gathered via a form, with four questions and sections for associated comments, completed by all participants. The number of participants is insufficient to detect any overall patterns or differentiation by sector, but the number of overall responses to questions is collated in brackets in the summary text and tables below, together with individual associated comments. The final feedback can be summarised as follows:

All eleven attendees felt the study results had changed their perceptions of ADS value. For most (6/11) this was a moderate change, because they felt that they already valued ADS highly (6) but the study had extended their understanding of either the scope of that value (2), particularly economic (3), and/or the degree of its value to other stakeholders (2). Several workshop delegates commented positively about seeing value expressed in economic terms, as this was something they had not previously considered or seen presented (4).

Everybody agreed to a greater or lesser degree (with most people agreeing strongly) with all three statements provided on the value of a stakeholder benefits map, quotations from users and depositors, and the economic analysis of the benefits. In the comments, the necessity of targeting messages to the specific audience was emphasised (3) and it was also felt to be important not to dwell exclusively on economic measures of value (2). Clear messages (1) and good graphical presentation (2) were felt to be key.

Participants made a range of suggestions on how to disseminate results of the ADS Impact study to maximise their value for ADS including the following targets (Table 7.1).

Table 7.1: Stakeholders' suggested dissemination targets

| Organisations/ Institutions/Roles: | Publications | Email Discussion Lists | Conferences |
|---------------------------------------|--|---------------------------|-------------|
| FAME (2) | Current Archaeology (2) | ALGAO (1) | IFA (1) |
| IFA (1) | British Archaeology (2) | HERForum (1) | CAA (1) |
| ALGAO (2) | Public Archaeology (1) | Britarch (1) | AAA (1) |
| English Heritage (1) | The Archaeologist (1) | FISH (1) | |
| Jisc (1) | Antiquity (1) | FAME (1) | |
| AHRC (1) | Heritage Counts (2) | CBA (1) | |
| HLF (1) | Planning, construction and architecture publications (2) | | |
| SCUPHA (1) | | | |
| University Depts (1) | New media | | |
| Heritage Champions (1) | Twitter (1) | | |
| Budget-holding Directors (1) | Facebook (1) | | |

Source: Authors' analysis

Finally participants were asked for suggestions on what additional resources or services the ADS could implement that would have a further impact on its value (economic or otherwise) to the archaeological community in the UK. The most popular suggestions were:

- Improve or expand OASIS, e.g. to include all grey literature (3); and
- Publication/archiving of more archaeological or local journals (2).

8 CONCLUSIONS AND RECOMMENDATIONS

Jisc and other funders, together with Higher Education Institutions, are investing substantial resources in projects and services for the curation and long-term preservation research data. It is a high priority area and there is strong interest in establishing the value and sustainability of these investments. The critical concept that determines how much or how little attention is paid to the long-term sustainability of digital content is how much value that content is perceived to have. This value is not solely economic, but in a tougher financial environment the economic arguments on value are increasingly important.

The study has analysed and surveyed perceptions of the value of digital collections held by the Archaeology Data Service (ADS). Importantly, we have also assessed and quantified the economic value of those collections with the ultimate objective of improving their prospects for sustainability. A range of economic approaches drew on data gathered through online surveys, user and depositor statistics to supplement and extend other non-economic perceptions of value.

The study has changed stakeholder perceptions, increasing recognition of the value of the ADS and digital archiving and data sharing generally. Most stakeholders already valued ADS highly, but felt the study had extended their understanding of the scope of that value, and the degree of its value to other stakeholders. They were positive about seeing value expressed in economic terms, as this was something they had not previously considered or seen presented but they also felt it was important not to dwell exclusively on economic measures of value.

The study shows the benefits of integrating a range of quantitative economic approaches to measuring the value and impacts of research data archiving and sharing services, with qualitative approaches exploring user perceptions and wider dimensions of value. Such a mix of qualitative and quantitative methods is important to capture and present the full range and dimensions of value of a data service such as ADS.

These messages and our key findings on the impact of the ADS have been disseminated widely to ADS stakeholders over the last 6 months via an information leaflet from the study (in print and PDF), and in conference presentations and posters.

8.1 Recommendations

1. The approaches used in the ADS Impact Study have now been applied to three UK data centres spanning very different disciplinary domains. The experience suggests that the approaches are transferable, but they require significant customisation to fit disciplinary and service differences. **There would be benefits from further research developing, refining and further exploring applications of the methods used in this study, as making the “business” or funding case for data centres and services plays an increasing role in ensuring their sustainability.**
2. To date these approaches have only been applied to national subject data centres. However they should be equally applicable to other international or local institutional repositories holding research data. **Hence, we should consider also applying these methods of valuation at international or local level.**

3. It is also clear that different data centres and services collect financial and operational data, such as user statistics, data deposit, data access and download statistics, to varying levels of detail. **There would be considerable advantage to providing guidance regarding the collection of such data as it is essential in making the “business” or funding case.** This should ensure a greater degree of standardisation of statistical records across centres, as well as providing the basis for more comprehensive and reliable data for economic analysis.
4. The study has looked at the aggregate value of the ADS. There is also significant scope for more granular studies that focus on the value of specific collections, such as the grey literature library, or the economic value of ADS services to specific groups. **Hence, we should consider applying these methods of valuation at more granular levels than the overall collections or all stakeholders.** There may also be some practical advantages to a narrower focus in simplifying some of the statistics and the analysis of different usage patterns across collections and user groups.
5. Value and perceptions of value change over time. The ADS user community was surveyed in 2010 for the RIN study and both ADS users and depositors surveyed for this study during June 2012. **ADS and funders should consider opportunities to repeat the ADS surveys and extend the available time series of comparative studies in future years. Ideally another survey should be considered within the next five years.**
6. During presentations given by the ADS on this impact study, there was a general audience reaction that the heritage sector outside of Higher Education has been slow to adopt measures of economic value and that the approaches used in this study have a lot to offer to other heritage organisations. **The wider heritage sector and its funders may wish to consider this study and any implications and applicability for museums, libraries and archives and their digital collections.**

REFERENCES

- Aitchison, K. and Edwards, R. (2008) *Archaeology Labour Market Intelligence: Profiling the Profession 2007-08*. Reading, Institute of Field Archaeologists. Available <http://www.archaeologists.net/profession/profiling>
- Archaeology Data Service (2011) *ADS Annual Report 2010-2011*. Available <http://archaeologydataservice.ac.uk/attach/annualReport2010/Annual%20Report%202010-2011.pdf>
- arXiv.org (2010) arXiv Business Model White Paper, Cornell University. Available <http://arxiv.org/help/support/whitepaper>
- Ayris, P. et al. (2008) *The LIFE2 Final Report*, London & Bristol: The Joint Information System Committee (Jisc). Available <http://www.life.ac.uk/2/documentation.shtml>
- Bakhshi, H., Freeman, A. and Hitchen, G. (2009) *Measuring intrinsic value*, MPRA Paper No 14902, University of Munich. Available <http://mpra.ub.uni-muenchen.de/14902/>
- Beagrie, N. (2009) *Draft Guide to Cost/Benefit Analysis for Research Data Services*, Charles Beagrie, Salisbury. Available http://www.beagrie.com/DMIcost&benefit_programmeguidev1.pdf
- Beagrie, N., Chruszcz, J. and Lavoie, B. (2008) *Keeping Research Data Safe*, Jisc, London and Bristol. Available <http://www.jisc.ac.uk/publications/publications/keepingresearchdatasafe.aspx>.
- Beagrie, N., Lavoie, B., & Woollard, M. (2010) *Keeping Research Data Safe 2 Final Report*. London, Jisc. Available <http://www.jisc.ac.uk/publications/reports/2010/keepingresearchdatasafe2.aspx#downloads>
- Beagrie, N., and Pink, C. (2012) *Benefits from Research Data Management in Universities for Industry and Not-for-Profit Research Partners*. University of Bath. Available <http://opus.bath.ac.uk/32509>
- Beagrie, N., Houghton, J.W., Palaiologk, A. and Williams, P. (2012) *Economic Evaluation of Research Data Service Infrastructure: A Study for The ESRC*, Economic And Social Research Council. Available http://www.esrc.ac.uk/images/ESDS_Economic_Impact_Evaluation_tcm8-22229.pdf
- Beise, M., and Stahl, H. (1998) *Public Research and Industrial Innovations in Germany*, ZEW Discussion Paper No. 98-37. Available <http://ideas.repec.org/p/zbw/zewdip/5209.html>
- Bernstein, J.I. and Nadiri, M.I. (1991) 'Product demand, cost of production, spillovers and the social rate of return to R&D,' NBER Working paper 3526.

- British Library.(2004) *Measuring the Economic Impact of the British Library*, Power point presentation on unpublished report. Available <http://www.bl.uk/aboutus/stratpolprog/increasingvalue/publicvalue/confpres/pungelwesmar ks.pdf>
- CEPA and Ware, M. (2010) *Transition dynamics of improving access to research*, RIN, London. Available <http://www.cepa.co.uk/publications.php>.
- CEPA/RIN (2008) *Activities, costs and funding flows in the scholarly communications system*, RIN, London. Available <http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/activities-costs-and-funding-flows-scholarly-commu>.
- CEPA/RIN (2011) *Heading for the open road: costs and benefits of transitions in scholarly communications*, RIN London. Available <http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/heading-open-road-costs-and-benefits-transitions-s>.
- Council for British Archaeology (2001)*From the Ground Up. The Publication of Archaeological Projects, a user needs survey*. Available: <http://www.archaeologyuk.org/publications/puns/>
- Dekkers, M., Polman, F., teVelde, R., & de Vries, M. (2006) MEPSIR: Measuring European Public Sector Information Resources, Brussels: European Commission. Available http://www.epsipius.net/psi_library/reports/mepsir_measuring_european_public_sector_resources_report
- DotEcon (2006) *The commercial use of public information (CUPI)*, Report oft861, London: Office of Fair Trading. Available <http://www.oft.gov.uk/OFTwork/publications/publication-categories/reports/consumer-protection/oft861>
- Dowrick, S. (2003) *A Review of the Evidence on Science, R&D and Productivity*. Canberra: Department of Education, Science and Training.
- DTLR (2002) *Economic Valuation with Stated Preference Techniques*, London: Department of Transport, Local Government and the Regions, for a summary guide. Available <http://www.communities.gov.uk/documents/corporate/pdf/146871.pdf>
- Ecorys and Fitzpatrick Associates (2011)*Economic Value of Ireland's Historic Environment*.The Heritage Council. Availablehttp://www.heritagecouncil.ie/fileadmin/user_upload/Publications/Corporate/Economic_Evaluation_of_the_Historic_Environment_Ireland.pdf
- English Heritage (2010)*Heritage Counts 2010*. Historic Environment Forum. Available <http://hc.english-heritage.org.uk/content/pub/HC-Eng-2010>
- English Heritage (2011)*Heritage Counts 2011*.Historic Environment Forum. Available <http://hc.english-heritage.org.uk/content/pub/2011/hc-2011-england.pdf>

- European Commission (2006) *The economy of culture in Europe*. Available http://ec.europa.eu/culture/key-documents/economy-of-culture-in-europe_en.htm
- Evans, P., Hatcher, M., & Whittard, D. (2008) The preliminary satellite account for the UK: a sensitivity analysis, *Economic & Labour Market Review*, 2(9), September 2008, 37-43.
- Fry, J., Lockyer, S., Oppenheim, C., Houghton, J.W. and Rasmussen, B. (2008) *Identifying benefits arising from the curation and open sharing of research data produced by UK Higher Education and research institutes*, Jisc, London and Bristol. Available <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/jiscdataproposal-public.pdf>.
- Green Book (2011) *Appraisal and Evaluation in Central Government*, HM Treasury, London. Available <http://greenbook.treasury.gov.uk/index.htm>. Available http://www.hm-treasury.gov.uk/d/green_book_complete.pdf
- Griliches, Z. (1995) R&D and productivity: Econometric Results and Measurement Issues, In Stoneman, P. (Ed.) *Handbook of The Economics of Innovation and Technological Change*. Oxford: Blackwell, 52–89.
- Hall, B.H., Mairesse, J., & Mohnen, P. (2009) Measuring the returns to R&D, NBER Working Paper 15622, NBER, Cambridge MA.
- Heritage Lottery Fund and VisitBritain (2010) *Investing in Success. Heritage and the UK tourism economy*. Available http://www.hlf.org.uk/aboutus/howwework/Documents/HLF_Tourism_Impact_single.pdf
- Houghton, J.W. (2009) *Costs and Benefits of Alternative Publishing Models: Denmark*, DEFF, Copenhagen (September 2009). Available http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fdownloads%2fDK_Costs_and_benefits_of_alternative_publishing_models.pdf.
- Houghton, J.W. (2010) 'Costs and benefits of alternative scholarly publishing models: Lessons and developments,' in *Publishing in the networked world: Transforming the Nature of Communication*, Eds. Turid Hedlund and Yasar Tonta. ISBN 978-952-232-086-5, pp.385-403. Available <http://hdl.handle.net/10227/599>.
- Houghton, J.W., & Sheehan, P. (2009) Estimating the potential impacts of open access to research findings, *Economic Analysis and Policy* 39(1). Available http://www.eap-journal.com/vol_39_iss_1.php.
- Houghton, J.W., Rasmussen, B., Sheehan, P.J., Oppenheim, C., Morris, A., Creaser, C., et al. (2009a) *Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits*, Report to The Joint Information Systems Committee (Jisc) by Victoria University & Loughborough University. Available <http://ie-repository.jisc.ac.uk/278/>. Addendum <http://www.cfes.com/EI-ASPM/Jisc%20EI-ASPM%20Report%20%28Addendum%20April%2009%29.pdf>

- Houghton, J.W., de Jonge, J. and van Oploo, M. (2009b) *Costs and Benefits of Research Communication: The Dutch Situation*, SURFfoundation, Utrecht (June 2009). Available <http://www.surffoundation.nl/wiki/display/economicstudyOA/Home>.
- Houghton, J.W., Rasmussen, B. and Sheehan, P. (2010) *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*, The Scholarly Publishing and Academic Resources Coalition (SPARC), Washington DC. Available <http://www.arl.org/sparc/publications/papers/vuFRPAA/index.shtml>
- Houghton, J.W., Steele, C. and Sheehan, P. (2006) *Research Communication Costs in Australia: Emerging Opportunities and Benefits*, Report to the Department of Education and Training, Canberra. Available <http://hdl.handle.net/1885/44485>.
- Houghton, J.W., Swan, A., & Brown, S. (2011). *Access to Research and Technical Information in Denmark*, Forsknings og Innovationsstyrelsen and Denmark's Electronic Research Library (April 2011). Available <http://www.fi.dk/publikationer/2011/adgang-til-forskningsresultater-og-teknisk-information-i-danmark>
- Imholz, S., and Arns, J. (2007) *Worth Their Weight: An Assessment of the Evolving Field of Library Valuation*, New York: Americans for Libraries Council. Available <http://www.ala.org/ala/research/librarystats/worththeirweight.pdf>
- Industry Commission (1995) *Research and Development*, Report No 44, Industry Commission, Canberra.
- Institute of Field Archaeologists (IFA) (2008-2011) *Job Losses in Archaeology Reports*. Available <http://www.archaeologists.net/profession/recession>
- KRDS 2011, *KRDS/I2S2 Digital Preservation Benefits Analysis Toolkit*. Charles Beagrie Ltd, Salisbury. Available <http://beagrie.com/krds-i2s2.php>
- Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy* 20(1), 1991, 1-12.
- Mansfield, E. (1998). Academic research and industrial innovation: an update of empirical findings. *Research Policy* 26(7/8), 1998, 773-776.
- Martin, B.R., & Tang, P. (2007) *The benefits from publicly funded research*. SWEPS Paper No. 161, Science Policy Research Unit, Brighton: University of Sussex. Available <http://www.erawatch-network.eu/reports/sewp161.pdf>
- Missingham, R. (2005) *Libraries and economic value*, National Library of Australia Staff Paper, Canberra. Available <https://www.nla.gov.au/openpublish/index.php/nlasp/article/viewArticle/1213/1498>
- National Institutes of Health (2008) *Analysis of Comments and Implementation of the NIH public access policy*, Washington DC.: NIH. Available http://publicaccess.nih.gov/analysis_of_comments_nih_public_access_policy.pdf.

- National Science Board (2010) *Science and Engineering Indicators 2010*, National Science Foundation (NSB 10-01). Available <http://www.nsf.gov/statistics/>
- Noonan, D.S. (2003) 'Contingent valuation and cultural resources: a meta-analytic review of the literature', *Journal of Cultural Economics*, Vol 27, pp159-176.
- Oakleaf, M. (2010) *Value of Academic Libraries: A Comprehensive Research Review and Report*, Chicago: Association of College and Research Libraries (ACRL). Available http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_bibliography.pdf
- PIRA (2000) *Commercial exploitation of Europe's public sector information*, Brussels: European Commission. Available http://www.epsipus.net/psi_library/reports/commercial_exploitation_of_europe_s_public_sector_information_pira_study.
- Pollack, R. (2009) *The Economics of Public Sector Information*, CWPE 0920, Cambridge: University of Cambridge. Available <http://www.econ.cam.ac.uk/dae/repec/cam/pdf/cwpe0920.pdf>.
- Quality Assurance Agency for Higher Education (QAA) (2007) *Subject benchmark statements. Archaeology*. Available <https://www.qaa.ac.uk/academicinfrastructure/benchmark/statements/Archaeology.asp>
- Ramirez, J.M. (2009) "Measuring: from Conjoint Analysis to Integrated Conjoint Experiments". *Journal of Quantitative Methods for Economics and Business Administration* 9: pp28-43. See http://en.wikipedia.org/wiki/Conjoint_analysis
- Royal Society (2012) *Science as an open enterprise*. Available <http://royalsociety.org/policy/projects/science-public-enterprise/report/>
- Salter, A.J. & Martin, B.R. (2001) The economic benefits of publicly funded basic research: a critical review, *Research Policy* 30(3), 509-532.
- Scott, A., Steyn, G., Geuna, A., Brusoni, S., & Steinmueller, E. (2002) *The Economic Returns to Basic Research and the Benefits of University-Industry Relationships*. Report to the Office of Science and Technology, London.
- Shanks, S. & Zheng, S. (2006) *Econometric modeling of R&D and Australia's productivity*, Staff Working Paper, Canberra: Productivity Commission
- Southport Group (2011) *Realising the Benefits of Planning-Led Investigation in the Historic Environment: A Framework for Delivery*. Available <http://www2.lse.ac.uk/geographyAndEnvironment/research/london/pdf/southport.pdf>
- Svanhild, A. (2009) 'Libraires and return on investment (ROI): a meta-analysis', *New Library World*, 110(7), pp311-324.

Sveikauskas, L. (2007) R&D and Productivity Growth: A Review of the Literature, Washington DC.: US Bureau of Labor Statistics Working Paper 408.

Tanner, S. (2012) *Cultural Value and Impact*. Blog post: http://simon-tanner.blogspot.co.uk/2012_07_01_archive.html

TeVelde, R. (2009) 'Public Sector Information: Why Bother?,' in Paul Uhlir (ed.) *The socioeconomic effects of public sector information on digital networks: Towards a better understanding of different access and reuse policies*, National Academies Press, Washington DC. Available http://www.nap.edu/openbook.php?record_id=12687&page=40.

Technopolis (2011) *Data centres: their use, value and impact*. Research Information Network. Available <http://rinarchive.jisc-collections.ac.uk/our-work/data-management-and-curation/benefits-research-data-centres>

Tenopir, C., Love, A., Park, J., Wu, L., Baer, A., & Mays, R. (2010) *University Investment in the Library, Phase II: An International Study of the Library's Value to the Grants Process*, Elsevier Connect White Paper. Available <http://libraryconnect.elsevier.com/whitepapers/roi2/lcwp021001.html>

Weiss, P. (2002) *Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impacts*, National Oceanic and Atmospheric Administration. Available <http://epsiplatform.eu/reports/borders-in-cyberspace>

Appendix 1: Summary of Survey Responses

As described in section 2.3, two online surveys were conducted, aimed at depositors of data with ADS and ADS users, respectively. The results are set out in full in this Appendix. The economic analysis based upon the quantitative cost, time and value questions is described in section 5.

The most difficult part of the user survey for respondents is a set of questions asking for responses in terms of percentage changes (rather than a more qualitative 5-point scale with values such as “low” or “high”). The original approach of asking for these to be stated by the user with positive or negative values underwent several modifications during development in light of test user feedback requesting pre-defined scales. Hence, there are some questions that give respondents a choice between percentage point categories that are subsequently used in estimations which, ideally, would have been based on continuous data. Given the preliminary nature of the economic estimations and the necessities of designing survey questionnaires that generate strong response and completion rates, these compromises were deemed necessary.

User and depositor surveys were linked from: Facebook (363 followers); Twitter (approximately 5,000 followers); ADS RSS (unknown number of contacts); and other mailing lists (e.g. HER Forum for local authority archaeologists (338 contacts); Antiquist archaeology and digital data list (345 contacts); and Britarch a CBA list open to all with an interest in archaeology (1,563 contacts). There was also a link to the surveys on the ADS website.

The Depositors Survey

A total of 293 email invitations to participate were sent to ADS depositors, of which 45 bounced or were known to be undelivered, leaving a sample of 248. We received 86 responses, a 34% response rate.

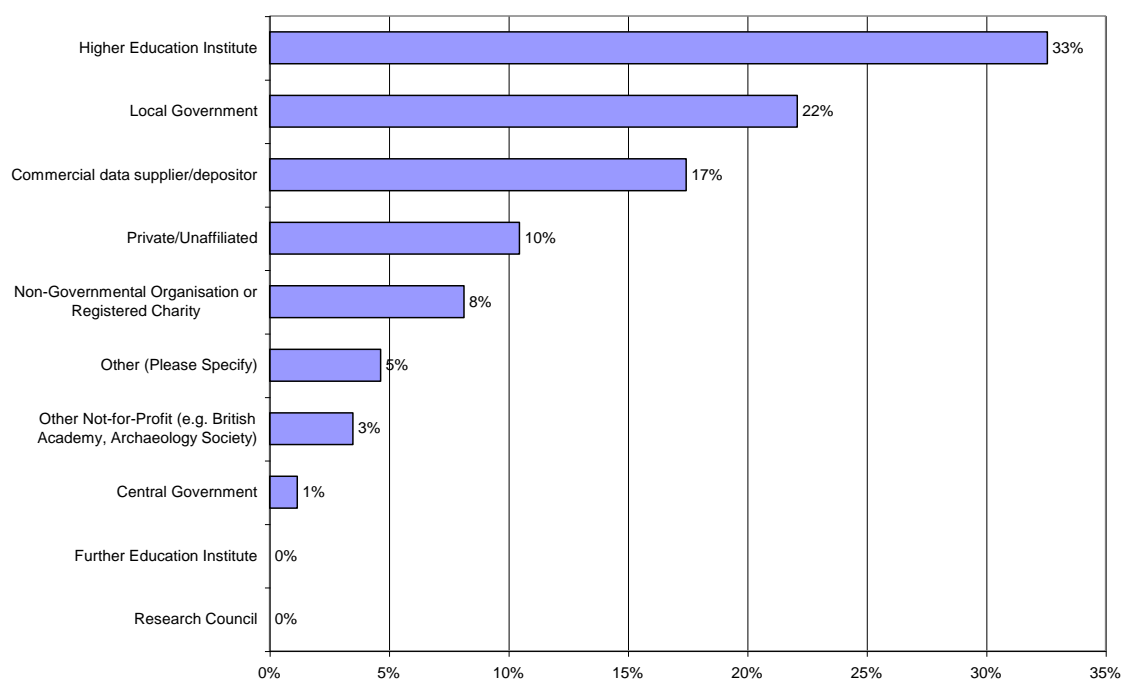
Affiliation and location

The first two questions established the nature and affiliation of the respondents, their location and local currency.

Q1: Main affiliation

The main affiliations of depositor respondents were higher education (33%), local government (22%), commercial depositors (17%), private unaffiliated individuals (10%), and non-government organisations (8%).

Figure A1: Main affiliation of ADS depositors (N=86)



Source: ADS Survey, Authors' analysis

Q2: Your currency?

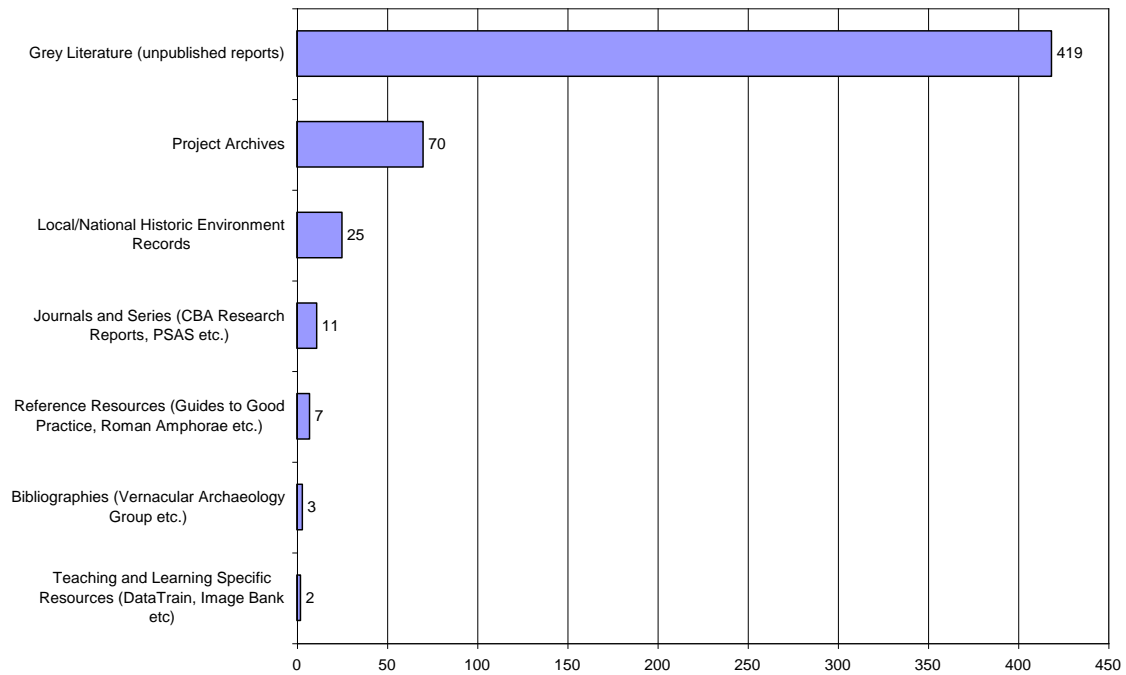
Ninety-five per cent were based in the UK, with the remainder based in the United States, Canada and Sweden.

Frequency and type of data deposits

Q3: Approximately how many times did you deposit/update the following data types to the ADS over the last three years?

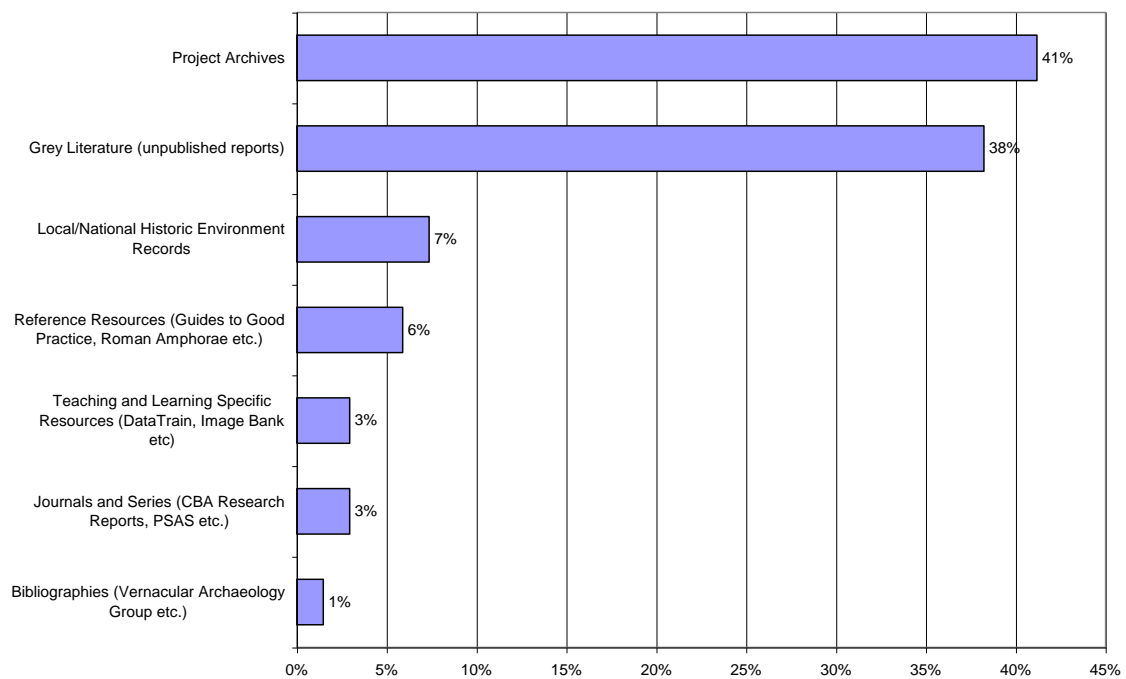
Grey Literature (i.e. unpublished reports) were by far the most common data type deposited by survey respondents during the last three years with a total of 419 deposits/updates per annum, followed by Project Archives with 70, Local/National Historic Environment Records with 25, and Journals and Series (CBA Research Reports, PSAS, etc.) with 11. Overall, depositors reported making 179 deposits/updates per annum – a mean of 3 per depositor per annum (median 0.67).

Figure A2: Frequency of ADS deposit/update by data type during the last three years, per annum (N=59)



Source: ADS Survey, Authors' analysis

Figure A3: Last data type deposited with ADS (N=68)



Source: ADS Survey, Authors' analysis

Q4: Considering the last data you deposited, what type was it?

Similarly, 41% of depositors reported that Project Archives were the most recent type of data deposited, and 38% Grey Literature (i.e. unpublished reports). Few had most recently deposited other data types.

Q5: Was this an initial deposit or an update to a previous deposit?

For 79% of respondents their most recent deposit was an initial deposit, while for the remaining 21% it was an update.

Data preparation, deposit and creation costs

Q6: Considering the last data you deposited, approximately how long did it take you to prepare them specifically for deposit and then to submit them to the ADS?

Depositors reported a wide range of preparation and deposit times, from as little as 15 minutes to more than 1,000 hours. The mean reported preparation and deposit time was 82 hours (median 15 hours) [N=65].

One might expect initial deposits of the data to take longer than updates, but that was not the case for our respondents. Initial deposits took an average of 79 hours (median 15) and updates an average 95 hours (median 15). This is likely to reflect the types of data involved in the updates and initial deposits and the low cell frequency when assigning initial deposit and update to seven data types. For example, all the deposits of journals and series (two deposits) and of bibliographies (one deposit) were updates, while all the deposits of teaching and learning (two deposits) and reference resources (four deposits) were initial deposits.

The time taken in preparation and deposit was converted to costs by assigning each respondent to a salary group based on their affiliation (Aitchison and Edwards 2008). These were scaled to include non-wage labour costs using a 30% uplift (Green Book 2011). Across the sample, the mean reported per deposit cost was £1,925 - equivalent to £805 per annum. With the above caveat in mind, the mean reported costs for initial deposits was £1,644 and for updates £2,947.

Q7: Considering the last data you deposited, can you estimate the financial cost of CREATING them?

A similarly wide range of data creation costs was reported, with two depositors reporting zero and one depositor £1 million. Indeed, there were a number of curiously low responses, suggesting that some respondents may not have understood the question and/or reported the components of the total creation cost that related specifically to themselves. No fewer than 11 reported data creation costs of £50 or less. Hence, responses may somewhat understate true data creation costs. However, this will have little impact on the economic value analysis as data creation costs are treated as 'sunk costs' (i.e. it is assumed that the data would have been created whether ADS existed or not).

The mean of reported data creation costs was £60,110 (median £3,250) [N=60]. Forty seven per cent said that the reported costs were average, with just 6% saying they were above average and 7%

saying they were below average. Nevertheless, the few reported variations from average were substantial – albeit in both directions.

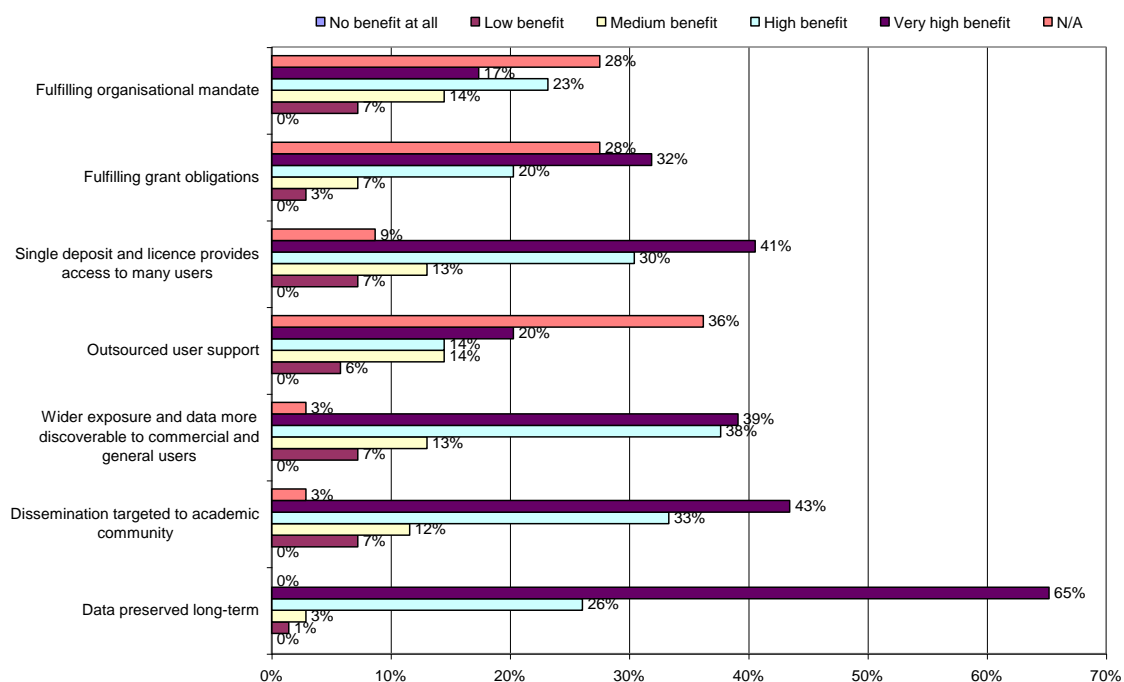
Benefits derived from deposit with ADS and impacts of access to ADS

Q10: Please select the level of perceived benefit to you as a result of your depositing/providing data to the ADS, for each of the following possible outcomes.

The main benefits from providing data to ADS reported by depositors included:

- Data preserved long-term, which was rated as a high or very high benefit by 91% (rating average 4.42);
- Dissemination targeted to academic community, rated high or very high by 77% (rating average 4.00);
- Wider exposure and data more discoverable to commercial and general users, was also rated high or very high by 77% (rating average 4.00);
- Single deposit and licence provides access to many users, rated high or very high by 71% (rating average 3.78);
- Fulfilling grant obligations, rated high or very high by 52% (rating average 2.68); and
- Fulfilling organisational mandate, rated high or very high by 41% (rating average 2.38).

Figure A4: Depositors' perceived benefits from providing data to ADS (N=69)

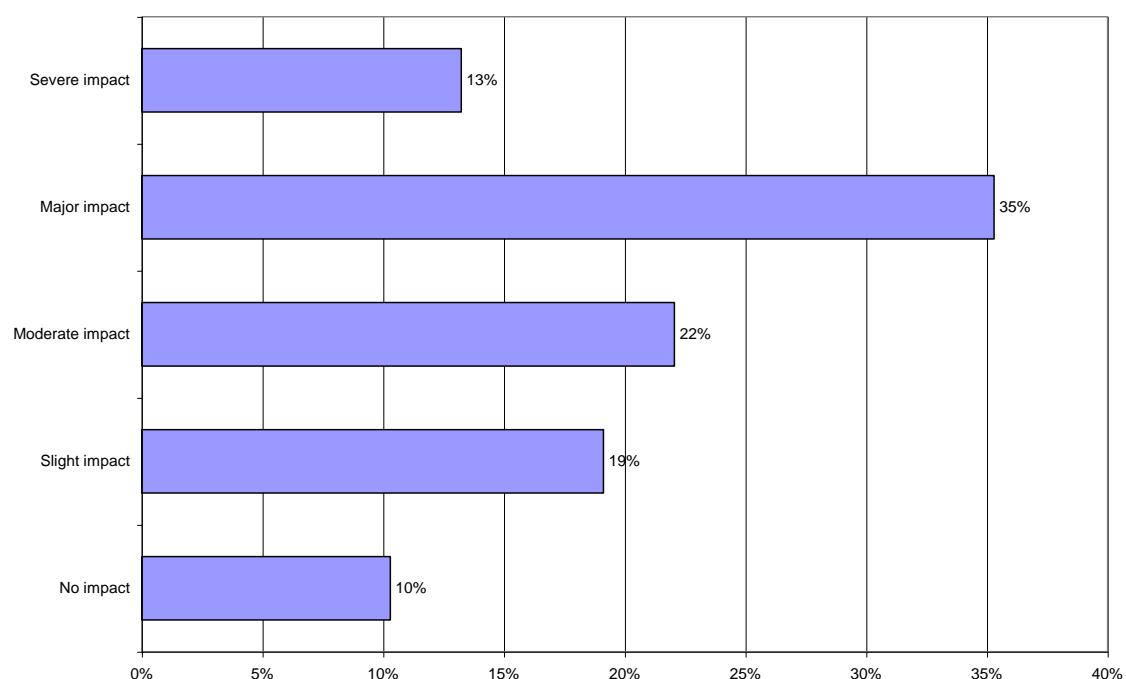


Source: ADS Survey, Authors' analysis

Q11: What impact would it have on your work if you could not deposit/provide data to the ADS?

Forty-eight per cent of depositors reported that not being able to provide data to ADS would have a severe or major impact on their work, and a further 41% said it would have a moderate or slight impact. Just 10% said it would have no impact (rating average 3.22). Comparing answers by affiliation and roles, it seems that those saying that it would have no impact were often required to deposit, and may have been assuming that the requirement would be dropped if ADS no longer existed.

Figure A5: Impact on depositors' work of not being able to provide data to ADS (N=68)



Source: ADS Survey, Authors' analysis

The Users Survey

Some 1,536 email invitations to participate were sent to ADS registered users, of which 21 bounced or were known to be undelivered, leaving a sample of 1,515. We received 299 responses in total from ADS registered and non-registered users (a 20% response rate). Of the total responses, 195 were from registered users (i.e. users who identified themselves as being registered in response to Q4), a 13% response rate.

The scale of the non-registered ADS user population is unknown, but we received 73 responses from individuals who identified themselves as non-registered users in response to Q4 (27% of respondents). It is thought that perhaps 25% or less of the regular users of ADS are registered. Given some 3,032 registered users at the time of writing this would suggest a regular user population of around 12,000. Hence, we probably reached a response rate of around 1% of non-registered users.

ADS website access data suggest that there were 11,020 users who visited the ADS site more than 51 times during 2012 (i.e. weekly use), 18,570 who visited more than 26 time during 2012 (i.e. fortnightly use), and 75,133 visits that were return visits (i.e. using the site more than once a year).

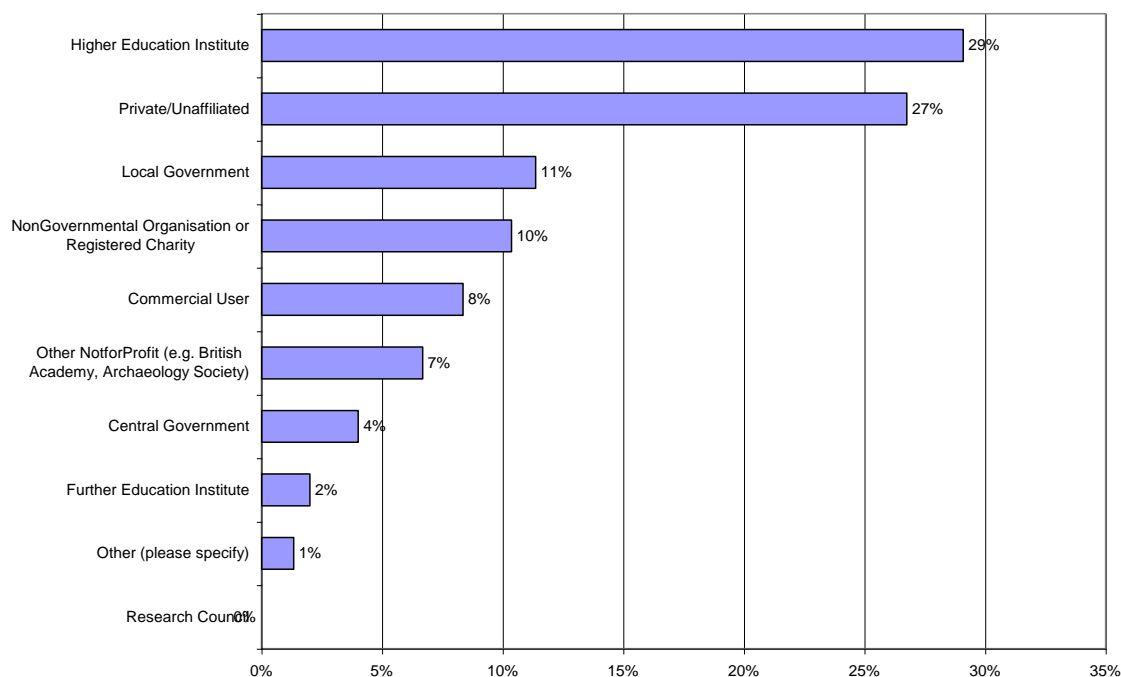
Affiliation, role and location

The first six questions established the role and affiliation of the respondents, their location and local currency.

Q1: Main affiliation

The main affiliations of ADS user respondents included higher education (29%), and private/unaffiliated individuals (27%).

Figure A6: Main affiliation of ADS user respondents (N=299)

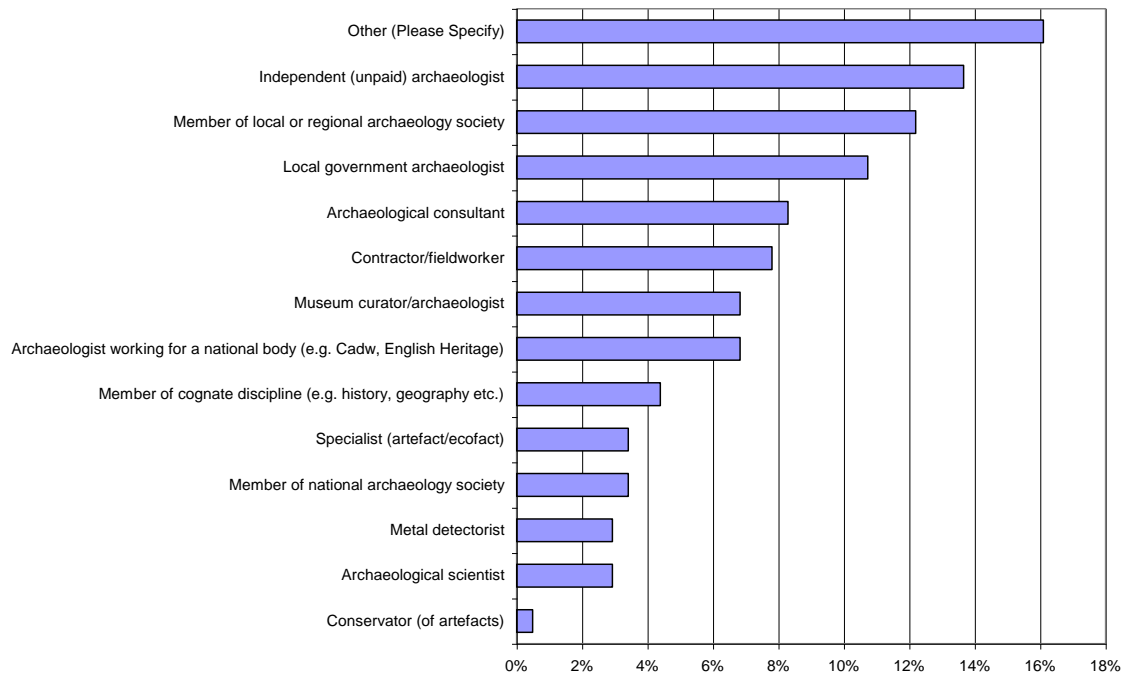


Source: ADS Survey, Authors' analysis

Q2 and Q3: What is your main role within this affiliation?

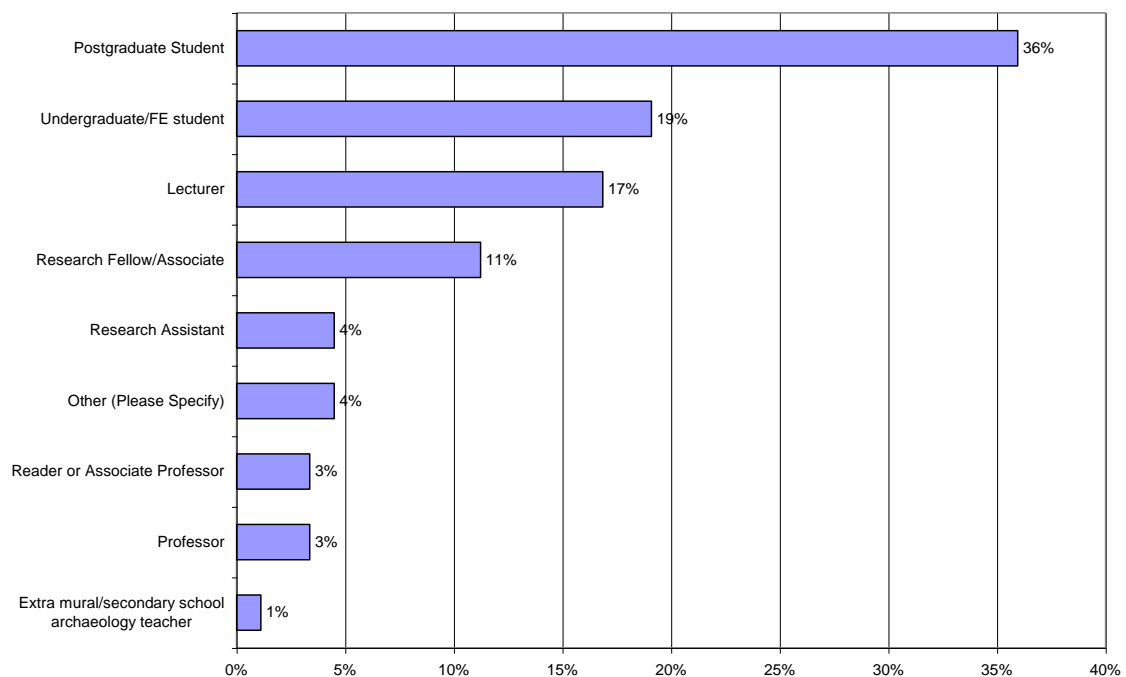
Respondents were then streamed into separate questions asking their role within their affiliation, according to whether they were in higher and further education or other affiliations. Despite being offered a long list, 16% of non-higher and further education respondents chose "other" and cited a wide range of roles. A wide range of private individual, public and private sector roles were also selected.

Figure A7: Main role within affiliation for non-education sector ADS user respondents (N=205)



Source: ADS Survey, Authors' analysis

Figure A8: Main role within affiliation for education sector ADS user respondents (N=89)



Source: ADS Survey, Authors' analysis

Within the education sector, 36% of respondents were postgraduate students and a further 19% undergraduate or further education students. The shares of responding research and teaching staff reflect their workplace shares, with more lecturers (17%) and research fellows (11%) than there are readers (3%) and professors (3%).

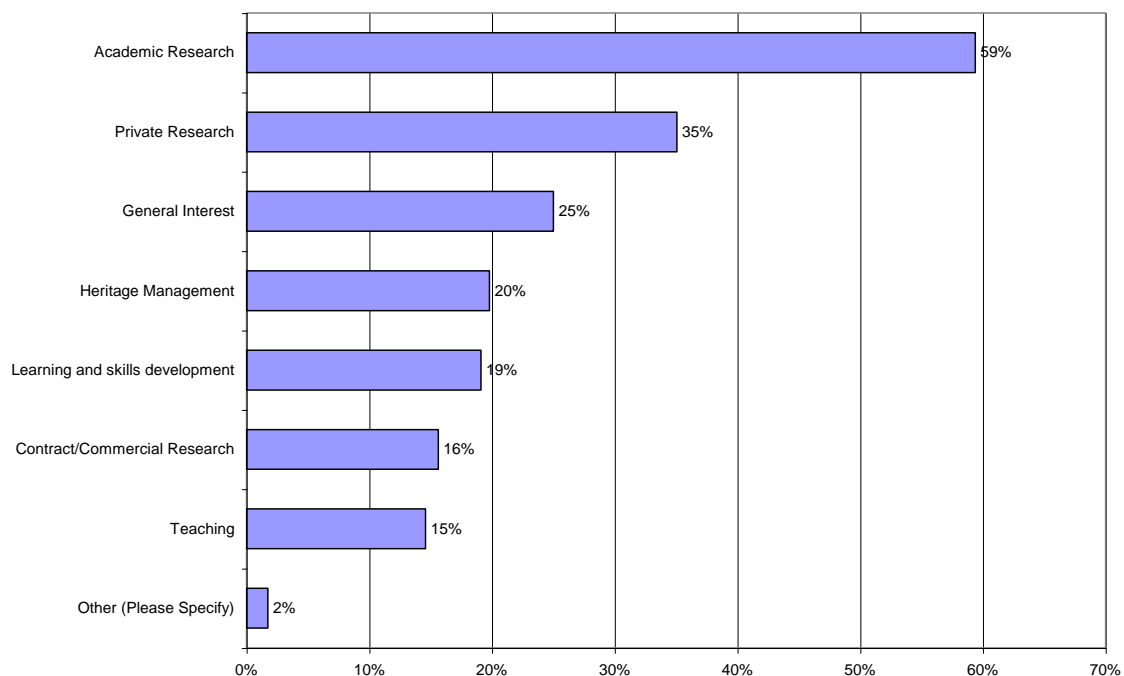
Q4: ADS allows optional user registration. Have you registered?

In order to ascertain how closely the respondents represented registered ADS users, they were asked if they had voluntarily registered with ADS. Thirty-one skipped the question, but of the 268 respondents to the question 73% (195) had registered.

Q5: What is the main purpose of your use of ADS?

Research is the main purpose of use of ADS among respondents, with 59% reporting their main purpose as academic research, 35% private research, and 16% contract or commercial research. Learning and skills development (19%) and teaching (15%) are also widely cited as a main purpose, with general interest (25%) and heritage management (20%) also providing motivations for use. Many respondents ticked multiple options, such as combining academic research, contract research and teaching (there were a total of 553 responses).

Figure A9: The main purposes of respondents use of ADS (N=288)



Source: ADS Survey, Authors' analysis

Q6: Local currency?

Respondents were asked to nominate their local currency as an input to cost calculations. This revealed that 83% of respondents were located in the UK, 8% in Euro zone countries, and a further 9% elsewhere. These latter included the US, Canada, Australia, Norway and South Africa, as well as Serbia, Albania, Croatia, and the Russian Federation.

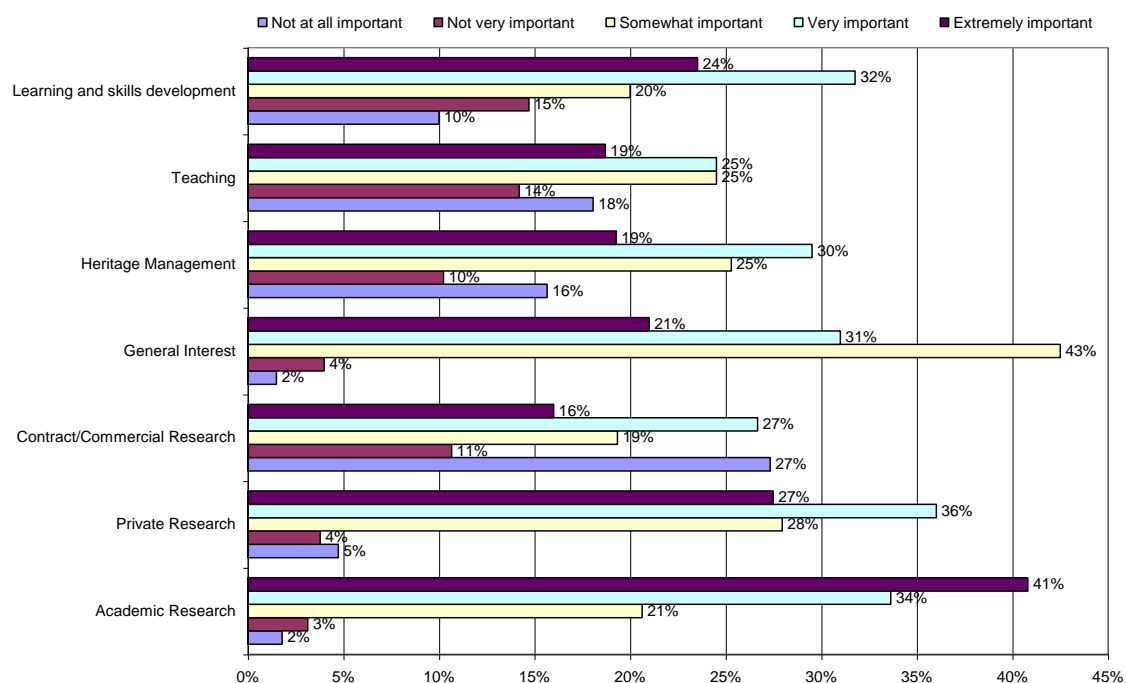
Importance of access to ADS

The following two questions explored the importance of data and services from ADS.

Q7: How important are data and services from the ADS for any of the following areas of your work?

Seventy-four per cent said that ADS was very or extremely important for their academic research, 64% said it was important for their private research, and 55% said it was important for their learning and skills development - reflecting the shares of academic, private and student users. Rating averages reflect the importance of academic and private research (4.09 and 3.78, respectively), and general interest (3.66).

Figure A10: Importance to users of data and services from ADS for different areas of their work (N=276)

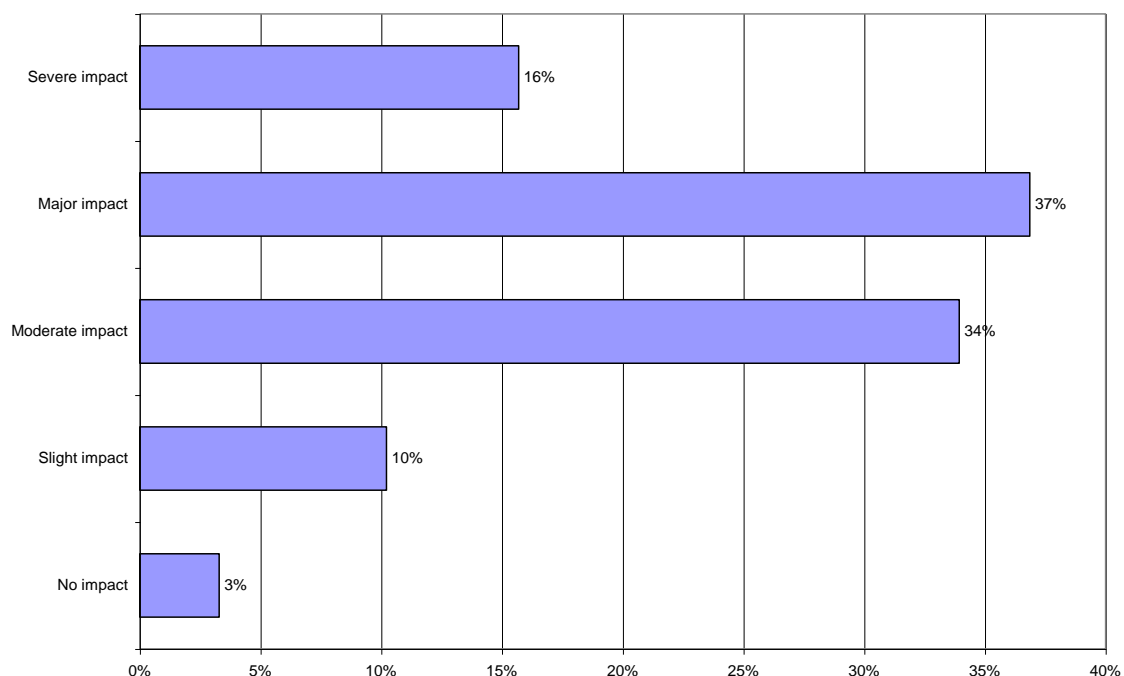


Source: ADS Survey, Authors' analysis

Q8: What impact would it have on your work if you could not access data and services from the ADS?

Asked the importance question in reverse, 53% said that not being able to access ADS data and services would have a severe or major impact on their work, and a further 44% said it would have a moderate or slight impact. Just 3% said that not being able to access ADS would have no impact on their work. Those saying 'no impact' were very infrequent users (median 6 accesses per annum). It is also worth noting that the question asked about impact on respondents' work, and 4 of the 9 saying 'no impact' were private/unaffiliated individuals, 2 described themselves as 'independent/unpaid archaeologists', and one was a research assistant in higher education. Hence, some may not have thought there could be any impact on their work.

Figure A11: Impact on respondents' work of not being able to access data and services from ADS (N=274)



Source: ADS Survey, Authors' analysis

The question also allowed comments, of which there were 46. The major themes included:

- Alarm – their work would be impossible without access to ADS (N=20);
- Concern – they could use alternative sources, but would incur extra time and financial costs (N=10); and
- Dismay – relating to the strategic implications for the profession as a whole (N=9).

Box A1: Selected comments on the impact of users not being able to access ADS data and services

"This just can't happen!"

"Access to archives is dependent on opening times which are generally during the normal working day thus limiting easy access without taking time off from work."

"I rely on the data held by the ADS for both my own research and for teaching. Not only because it gives me access to data with content directly within my subject area, but it also allows me to easily see how other researchers create and structure their data in a way which would be very difficult otherwise. It gives students access to all manner of data sets so that they can explore whatever subject area is most engaging to them within many of our areas of teaching. It also gives them a sense of the importance of publishing data for use by others, and the important role re-use of data has in archaeology."

"No other service provides access to this information so easily. In teaching students especially, access to a high quality, relevant online resource is invaluable. Maintenance of this provision is vital to archaeology as a whole to further the progress of the study."

"Though it has the potential of becoming the first port of call for acquiring datasets for research purposes at the moment it doesn't yet seem to have the critical mass to provide appropriate datasets for most specific purposes, meaning other (often more personal) arrangements around datasets and their use are made."

"Interpretation of archaeological evidence may be incorrectly biased, inconclusive or even wrong without access to otherwise unpublished grey literature and other source material"

"Given the number of pay walls in academic data the loss of access to the ADS would be catastrophic"

"ADS has been a lifeline for my research. Having access to the quantity and quality of data, along with access to grey literature has saved me months, if not years of work. ADS provides me with a starting point for research through its search engines, contact information, and links. By providing online publication of grey literature, data from 'lost' reports and little known sites are available, and provide for more complete and detailed knowledge of British archaeology."

"Key archives such as ALSF are absolutely critical for professional research and consultancy underpinning Impact Assessments and other work in support of industry and development."

"...My ADS access has been vital to widen my own background knowledge, see what else is available elsewhere in the country and to do so at no cost to myself... I cannot afford to travel or pay big sums to access documents. As secretary of a local archaeological society, much of my work is to answer telephone queries on archaeological topics - without ADS I could not reply to any width or depth."

Frequency and cost of access to ADS

The next six questions explored frequency of use, and the time and costs involved for users in accessing ADS data and services.

Q9: Approximately how frequently do you access/download data from ADS?

In Q9, respondents were given a range of time period access/download frequencies that produced non-continuous data. These were converted to annual numeric frequencies on the basis of months (12) or weeks (52) per year, with 'more than once a week' interpreted conservatively, as 53 times a year.

Grey literature, journals and series, and reference resources were the most frequently accessed/downloaded, in that order, followed by Historic Environment Records (HERs). Across all data types, ignoring the never and unable to say responses, respondents accessed/downloaded an average of 69 items a year (median 48 times).

Table A1: Frequency of use of ADS data by data type (N=250)

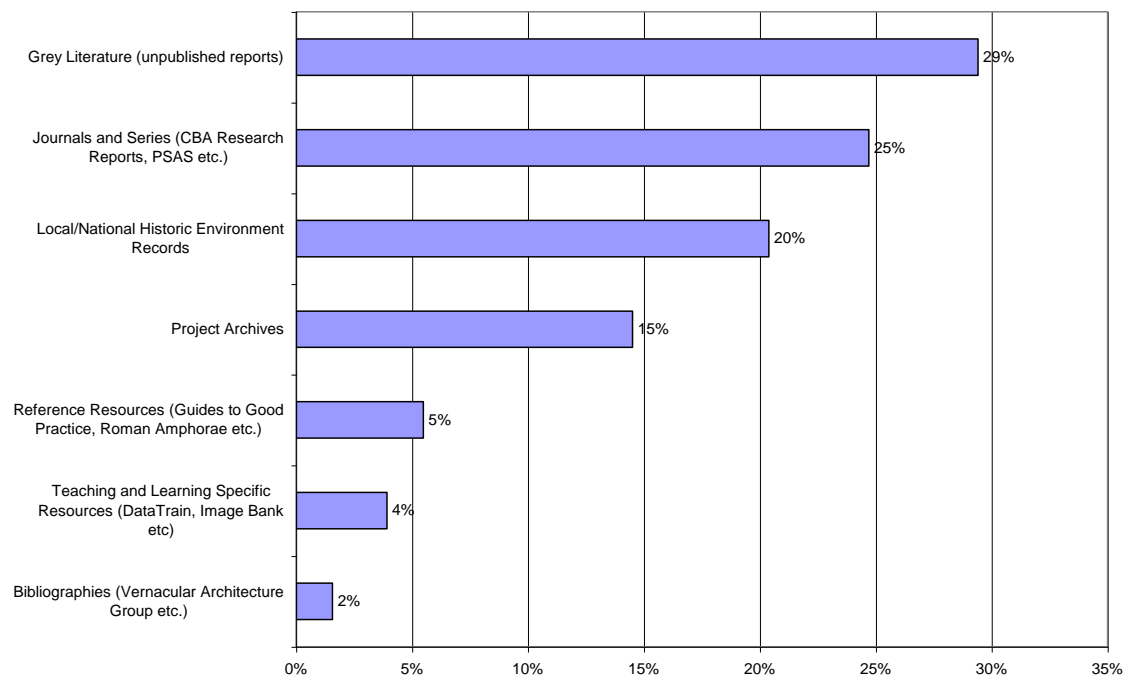
| | Local/National Historic Environment Records | Journals and Series | Grey Literature | Project Archives | Biblio-Reference graphiesResources | Teaching and Learning Specific Resources | |
|-----------------------|--|---------------------------|-----------------|---------------------|---------------------------------------|--|------------|
| Never | 12% | 5% | 5% | 9% | 29% | 14% | 39% |
| Once a year | 10% | 7% | 6% | 15% | 20% | 16% | 14% |
| Twice a year | 4% | 11% | 13% | 13% | 10% | 17% | 8% |
| Quarterly | 22% | 19% | 20% | 17% | 12% | 17% | 13% |
| Once a month | 20% | 22% | 23% | 16% | 10% | 18% | 6% |
| Twice a month | 14% | 14% | 15% | 15% | 6% | 7% | 9% |
| Once a week | 8% | 12% | 8% | 6% | 4% | 5% | 1% |
| More than once a week | 5% | 7% | 6% | 2% | 2% | 1% | 2% |
| Unable to say | 5% | 4% | 4% | 7% | 6% | 5% | 6% |

Source: ADS Survey, Authors' analysis.

Q10: Which of the following data types did you last use?

In order to randomise the responses to questions about the time and costs involved in accessing ADS data and services and enable us to assess how closely respondents' use matches overall use of ADS, respondents were asked which data type they last used. Answers reflected those to the previous question, with 29% citing grey literature as the last data type used, 25% citing journals and series, 20% HERs, and 15% project archives.

Figure A12: ADS data type that respondents last used (N=255)



Source: ADS Survey, Authors' analysis

Q11: How long did it take you to find and access that data from the ADS?

Respondents reported that finding and accessing ADS data took them an average of 29 minutes (median 15 minutes). The maximum time reported was 6 hours (N=248).

These times were converted to costs by assigning each respondent to a salary group based on Aitchison and Edwards (2008), adjusted for inflation, then scaling to include non-wage labour costs using a 30% uplift (Green Book 2011). For students we used the lower end graduate salaries reported in the Times and Guardian Higher Education pages, for undergraduates (£15,000 pa) and postgraduates (£20,000 pa), as we felt this best reflected the opportunity cost of earnings forgone. For the small number of retired and voluntary worker respondents we set the salary to zero and applied the average cost mark up of the other users to reflect the non-salary costs involved. Across the respondents, this resulted in an average costing of £18 per hour - £21 per hour for those saying that their duties included teaching, £20 per hour for those researching, and £13 per hour for those studying.

The mean estimated cost of accessing ADS was around £8 for the last data type accessed. These costs varied by the type of data last accessed, with a mean cost of access to teaching and learning specific resources of £20, and that for journals and series £6. The costs also varied by type of user, although those with more time appeared to spend more, resulting in what might be the opposite of an intuitive result, wherein the average staff access cost £8, and the average student access £10.

Box A2: Selected comments on how long it took users to find and access data from ADS

"I know my way round the ADS reasonably well and find that it is fairly well thought out, and has clearly undergone improvement on an ongoing basis since I first started using it 5 or 6 years ago."

"It is a very cumbersome system"

"I was asked to supply locational data as a GIS layer I was able to quickly and easily navigate to the page with the archives (admittedly using Google), extract the data I required in a portable format (.txt) that allowed me to import easily into Excel and convert this into a format (csv) that would allow them to import the data into their GIS software."

"Use ArchSearch frequently to cross reference local/ national HE data and check for new entries, but can't always find all the data fields I need - so this is a hard question to answer"

"I am now very familiar with ADS but although my last search was quick I have still found it difficult to find some content which I know exists. With such a complex site site navigation is never going to be simple but I do feel it could be considerable easier."

"Would be much easier if the NGR search was re-introduced. New site looks nice, but much more difficult to search productively."

Q12, Q13 and Q14: Were there any other costs involved? If there were other costs involved how much were they? What were they for?

Asked if there were any other costs involved in accessing ADS data, and if so how much and what they were for, 99% said that there were no other costs involved (N=257). No one gave any further information.

Q15: If the ADS had not existed, would you have been able to obtain the data in any other way?

In order to explore the possible costs involved in either sourcing the data elsewhere or recreating it, respondents were asked if they could have obtained the data they last accessed/downloaded from ADS in any other way. Fifty-six per cent said they could have obtained the data in another way, and 44% said they could not (N=251).

When asked to explain how they could have otherwise obtained the data, respondents offered a wide range of qualitative and quantitative answers (146 comments). Major themes in respondents' comments included:

- Extra time (including travel), effort and delay (N=85);
- Extra financial costs (N=50); and
- There may be no other source for all or part of the resource (including the difficulty in finding other holders and the possibility that they may no longer be operational).

Box A3: Selected comments on respondents' ability to obtain the data in another way

"This would have involved a publication search, with no certainty that the search would be successful / complete. The time cost would be significantly larger; I estimate that it would have been days rather than minutes. It would turn a retrieval into a second PhD."

"It might have been possible to pull together some of the information from a range of other sources, but it would have taken 1-2 days longer, and would not have been of the same quality. Much of the ADS documentation we use has benefited significantly from their expertise. In terms of the data we link to from our journals, I don't believe that the majority of this would be publicly available at all if the ADS did not exist."

"Through reference to actual hard copies of the journals - library visit, time away from desk, parking fees/travel time. Would have taken a few hours rather than a few minutes and there would have been direct and indirect costs incurred."

"It would have been a convoluted process, involving trying to find the right person to contact, making contact, requesting a document, negotiating how to get this (hard copy, by email, etc.), all taking time. If a document was needed for a piece of research and could not be emailed, there is the delay in waiting for it to arrive, plus costs associated. There is also the factor of the time it takes to find the right person to talk to, and being able to successfully speak to that person. Moreover, ADS also allows you to search for things (e.g. search a subject) which may reveal results that you weren't previously aware of."

"Approach organisation directly. Minimal costs, phone call, email etc. Time is more complex, requests for reports can sometimes take weeks to fulfil. I have found it very difficult to get hold of literature from bodies no longer operating."

"Possibly but it would vary on a case by case basis. In the current economic climate with a number of companies having folded the experience of what is available where is lost. Facilities like ADS are vital to provide some chance of long term survival."

"I could have done it, but it would have taken months of making individual contacts and negotiating its re-use, to find data which would be held in a variety of non-archival formats, and would have taken further months of work to get into any kind of shape to actually use. I suspect it would have required a further year of work within my PhD to not only get the data I ended up using, but put it into usable shape. Because of this, my PhD would have been much less ambitious, and of a far lower quality without the ADS."

"I could have paid for hard copies of the journals and reports of inter-library loaned them but that would have been a far more expensive and time consuming option. I very much appreciate being able to download the text immediately and in .PDF form so I can search for keywords rather than having to skim read the whole text if it was in paper form. I also haven't wasted too much time or money if I find it wasn't what I was looking for. I also prefer being able to search the ADS archives for a range of papers/reports on a similar theme using keywords - this is an extremely valuable part of the ADS and saves an immense amount of time using library catalogues and internet searches."

"Visits to libraries (primarily the Society of Antiquaries, which involves a return rail fare to London at £32); correspondence with colleagues (free, but time consuming - and some colleagues are not good at answering emails)."

"A heavily qualified 'yes'. Without the ADS, I would not have known that a relevant project archive existed, though theoretically I might perhaps have found a reference elsewhere. But then to find the contact details of the digital archive holder, contact them in the hopes that they might have a coherent and readily transferable 'package' for the archive, wait for a response, wait some more, negotiate possible terms and conditions of transfer once they did respond and sort out the format(s) for the data, metadata and documentation, and then wait for the archive to arrive -- all of this could potentially take weeks if not months. Whether costs other than my time and the time of the original data creator(s) might be involved is hard to say -- it depends on whether they would seek to 'recover costs,' not just of sharing the data, but potentially of having gathered the data in the first place."

"In the case of out-of-print CBA reports, buying copies off books on amazon at ruinous cost."

"Copies of grey literature may not be available from elsewhere if the original excavator has died or gone out of business and the existence of it may not be traceable by other means"

"From an HER. Cost and time of travel to LG offices, including time lag in having booked appointment to access. LG archaeologists are also often unaware of the copyright constraints on material for public access and fair use policy."

"With difficulty, remote island location, would have had to acquire information through other university libraries. University libraries don't have the resources to subscribe to the main journals, not sure how to get the other archived data etc. It would have been difficult and very time consuming, and costly, probably meaning I would not have been able to use the data to complete dissertation"

"Writing e-mails to the concerned parties. Hoping for them to send it here."

"Would have involved expensive car or train trip to nearest university library (2hrs away)"

"My university library provides a certain number of free interlibrary loans/photocopies. The cost would have come out of my allocation for this."

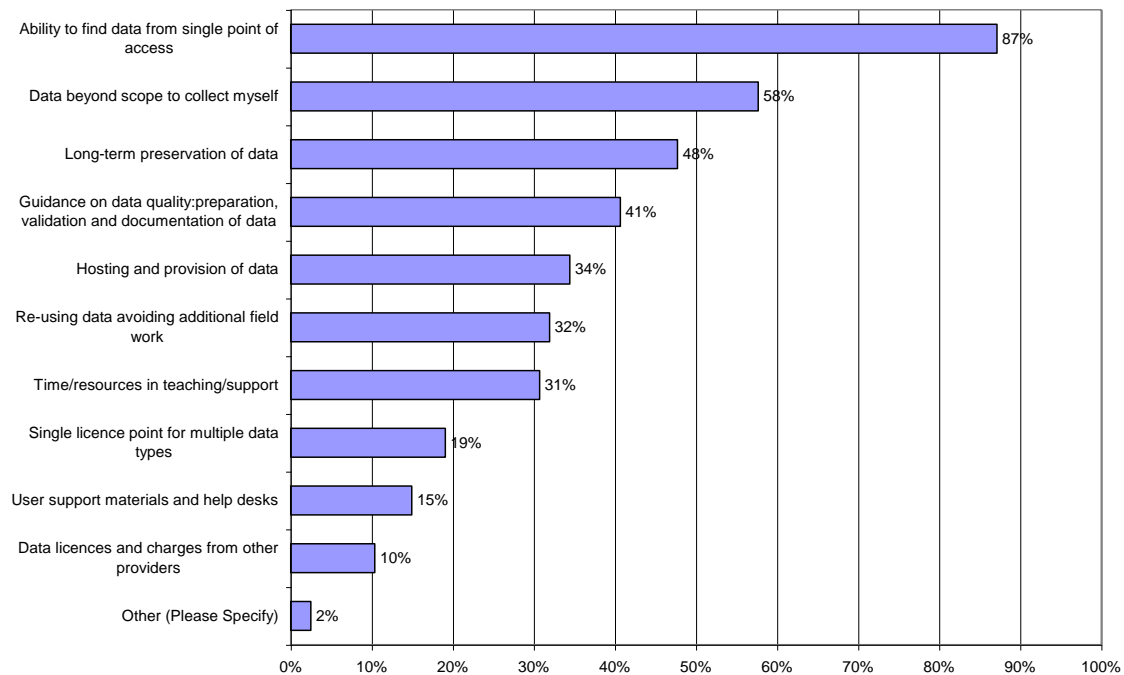
"Possibly inter-library loan if it was available to the public library service at a cost of maybe £10 plus."

Q16: Do you consider that you have saved time or money in any of the following areas as a result of using data and/or services from the ADS?

Moving from the costs involved in accessing ADS data and services, the next two questions explored the possible time and cost savings and benefits arising from accessing ADS. Respondents could tick multiple options (N=241, responses 909).

By far the most widely cited factor contributing to savings was the ability to find data from single point of access (87%). That the data were beyond their scope to collect themselves (58%), long-term preservation of data (48%), and guidance on data quality through preparation, validation and documentation of data (41%) were also widely cited areas of benefit. These responses reveal significant efficiency impacts in terms of time and cost savings.

Figure A13: Areas of savings for respondents resulting from accessing ADS (N=241)

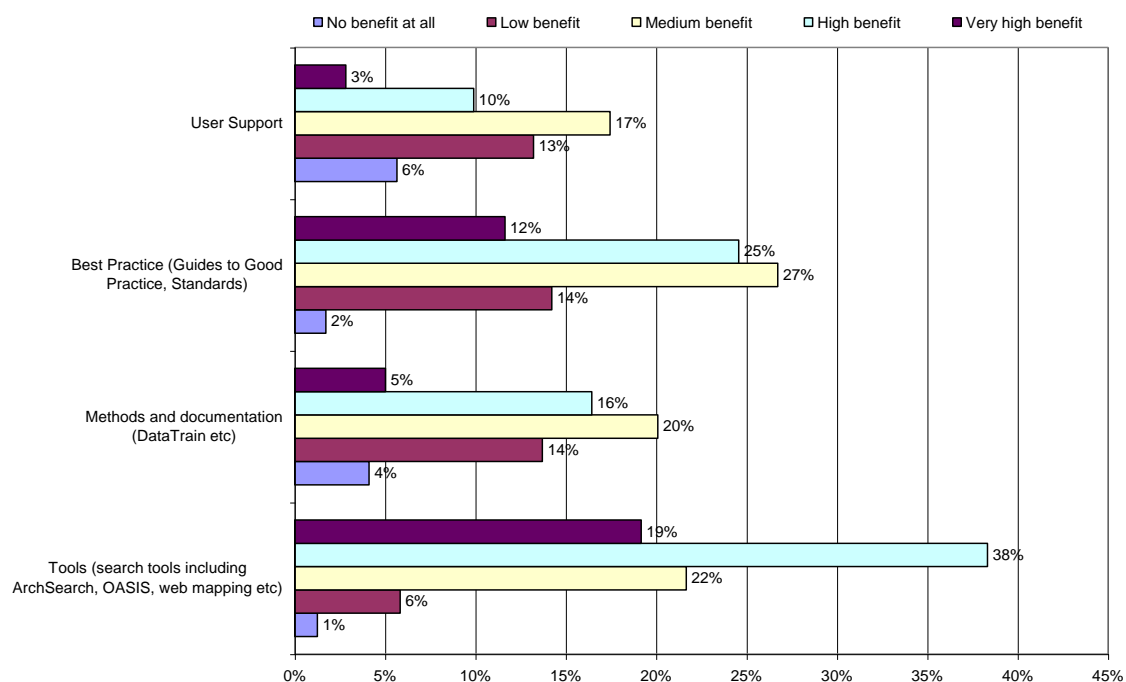


Source: ADS Survey, Authors' analysis

Q17: To what extent do you also benefit from the ADS in any of the following ways?

Fifty-eight per cent said that they derived a high or very high benefit from tools (e.g. search tools including ArchSearch, OASIS, web mapping, etc.), 36% said they derived high or very high benefit from guides to best practice and standards, and 21% said they derived high or very high benefit from methods and documentation (e.g. DataTrain, etc.). Very few respondents reported deriving no benefit from any of the items or services they had used.

Figure A14: Extent of other benefits for respondent ADS users (N=249)



Source: ADS Survey, Authors' analysis

Intensity of data use and efficiency impacts of ADS

Respondents were then streamed into sections for students, researchers and teachers, with the next five questions exploring the potential efficiency impacts of accessing ADS data and services. Some of these questions were difficult to answer, asking people to report their impression of their own or other people's activity time distributions. It was clear from responses that some had misinterpreted the questions relating to time spent with data as referring to the shares of their data time, rather than their total activity time, as had been intended. Hence, responses to these time share questions must be interpreted with caution.

Students

Q18: Are you an undergraduate or FE/school student?

Twenty-one respondents (8%) reported that they were undergraduate or further education students.

Q19: Over the last twelve months, on average how many hours a week did you spend studying and learning?

Asked how many hours a week they had spent studying and learning, on average over the last 12 months, answers ranged from 2 hours to 60 hours - the latter cited a remote island location as the reason for spending so much time studying. The average was 21 hours per week (median 20 hours

per week) (N=22). The question was open-ended, simply asking respondents to enter a number of hours.

Q20: Can you estimate the approximate share of your total study/learning time spent with data during the last twelve months (e.g. creating, manipulating and analysing data)?

In order to extrapolate from student respondents, through ADS student users to the wider student community who may not be using ADS, the next two questions explored the relative data use intensity of the student respondents and other similar students. To aid respondents they were offered a scale from 0% to 100% in 10% intervals.

Asked what share of their study time they spent with data from ADS, the majority spent between 10% and up to 40% of their time with data from ADS, with a mode of 10% and an approximate mean of 25%. They reported spending a further mean of 48% of their study time using data from other sources, although the mode was around 30%.

Analysing responses to questions 19 and 20 suggests that student respondents spent an average of around 5 hours a week with ADS data over the last 12 months (median 3 hours), and a total average of 15 hours a week with data from all sources (median 11 hours).

Q21: Do you have any impression of what might be typical for other students with the same main affiliation (sector) as you?

To enable us to extrapolate to non-users of ADS data and services, respondents were asked about the data use intensity of others in their field. Perhaps curiously, whereas the modal response for their own data use was 10% of their time with data from ADS and 30% with data from other sources, for others in their field their mode responses were 10% with data from ADS and 50% with data from other sources. However, the approximate mean data use was similar at 30% of time with data from ADS and a further 51% with data from other sources. Due to the complexity of the question and the relative lack of experience among student users, these responses should be interpreted with caution.

Q22: Do you have any sense of the extent to which your use of data and services from the ADS has changed your STUDY/LEARNING efficiency (i.e. compared to if no ADS existed)?

Asked to what extent their use of ADS data and services had changed their study/learning efficiency, respondents were offered a scale from 'negative change' to 0%, 5%, 10% then 10% intervals to >90%. Answers ranged from no change (0%) to a more than 90% gain. The reported mean was a 44% efficiency gain (median 40%). Given the number of hours spent studying reported in Q19, this translates to an efficiency gain equivalent to 9 hours a week at current activity times – allowing them to complete the same work in close to half the time and/or do almost twice as much work in the same time. These efficiency impact responses seem high, and given limited experience one could, perhaps, question the ability of students to assess the efficiency impacts. Moreover, while it was intended to refer to all study/learning time, it is possible that respondents interpreted the question as referring to the time spent with ADS data alone. If so, the reported efficiency gain would be equivalent to 2 hours a week.

Researchers

The same series of efficiency related questions were repeated for researchers and teachers.

Q23: Do your job duties include research?

One hundred and seventy-two respondents (77%) reported that their job duties included research.

Q24: Over the last twelve months, on average how many hours a week did you spend on research?

Asked how many hours a week they had spent on research, on average over the last 12 months, answers ranged from 1 hour to 70 hours. Three responses were deleted as outside the range of hours, suggesting that respondents had misinterpreted the question. The average was 16 hours per week (median 12 hours per week) (N=147).

Q25: Can you estimate the approximate share of your total research time spent with data during the last twelve months (e.g. creating, manipulating and analysing data)?

Asked what share of their research time they spent with data from ADS, the majority spent between 10% and up to 30% of their time with data from ADS, with a mode of 10% and an approximate mean of 21%. They reported spending a further average of 56% of their research time using data from other sources, and the mode was around 50%. To aid respondents they were offered a scale from 0% to 100% in 10% intervals. Four commented that <10% should have been an option.

Analysing responses to questions 24 and 25 suggests that research respondents spent an average of around 4 hours a week with ADS data over the last 12 months (median 2 hours), and a total average of 14 hours a week with data from all sources (median 10 hours).

Q26: Do you have any impression of what might be typical for other researchers in the same sector (e.g. HE/contract/commercial or private, etc.) as you?

Modal responses for their own data use were 10% of their time with data from ADS and 50% with data from other sources, and for others in their sector their modal responses were also 10% with data from ADS and 50% with data from other sources. The approximate mean data use was also similar at 24% of time with data from ADS and a further 52% with data from other sources.

Q27: Do you have any sense of the extent to which your use of data and services from the ADS has changed your research efficiency (i.e. compared to if no ADS existed)?

Asked to what extent their use of ADS data and services had changed their research efficiency, answers ranged from no change (0%) to a more than 90% gain. The reported mean was a 44% efficiency gain (median 40%). Given the time spent on research reported in Q25, this translates to an efficiency gain equivalent to 7 hours a week at current activity times.

Again these efficiency impact responses seem high, and while it was intended to refer to all research time, it is possible that respondents interpreted the question as referring to the time spent with ADS data alone. If so, the reported efficiency gain would be equivalent to less than 2 hours a week.

Box A4: Selected comments on research efficiency impacts

"It saves me time as I can research from home/ work without having to go into university/ library/ county office etc."

"ADS has not changed my research efficiency as I have always used ADS, including during my graduate course."

"Over the last 5 years I worked as a HER Officer. I saw how the availability of grey literature online greatly improved research efficiency, as it allowed researchers to check online resources before making costly trips to various HERs."

"Difficult to measure precisely, but being able to search for and access bibliographies, project data and some published works from one's desk does save time, but also provides pointers to other avenues of research which can be followed up elsewhere, and which would have involve significant time undertaking literature searches. Literature searches can be targeted."

"My archaeological research time would be trebled if I didn't have access to ADS."

"I often search first on ADS for things I think they will hold, and then search elsewhere if ADS doesn't have it. Much of the information I could possibly get in other ways, but the timescales in terms of emailing and waiting for replies, would be much longer"

"As a home worker, access to huge online archives of data/reports makes what I do achievable. I doubt I would be able to have any confidence in the research I do without access to ads data, not without considerable (days/weeks) of extra time accessing libraries and HER report collections."

"It had a huge impact on my research possibilities and 'workflow'. However it hasn't become very much efficient to do research, because of the large amounts of excavations and data generated by these."

Teachers

Q28: Do your job duties include teaching?

Seventy-eight respondents (35%) reported that their duties included teaching.

Q29: Over the last twelve months, on average how many hours a week did you spend teaching and preparing learning materials?

Asked how many hours a week they had spent teaching and preparing learning materials, on average over the last 12 months, answers ranged from 1 hour to 49 hours. The mean was 9 hours per week (median 5 hours per week) (N=60).

Q30: Can you estimate the approximate share of your total teaching/preparation working time spent with data during the last twelve months (e.g. creating, manipulating and analysing data)?

Asked what share of their teaching and preparation time they spent with data from ADS, the majority spent between 10% and up to 30% of their time with data from ADS, with a mode of 10% and an approximate mean of 16%. To aid respondents they were offered a scale from 0% to 100% in 10% intervals. They reported spending a further average of 42% of their teaching and preparation time using data from other sources, although the modal answer was 90%.

Analysing responses to questions 29 and 30 suggests that teaching respondents spent an average of around 2 hours a week with ADS data over the last 12 months (median 1 hour), and a total average of 8 hours a week with data from all sources (median 4 hours).

Q31: Do you have any impression of what might be typical for other teachers with the same main affiliation (sector) as you?

Modal responses for their own data use was less than 10% of their time spent with data from ADS and 90% with data from other sources, for others in their field their mode responses were 0% (less than 10%) with data from ADS and 90% with data from other sources. However, the approximate mean data use was similar at 16% of time with data from ADS and a further 53% with data from other sources.

Q32: Do you have any sense of the extent to which your use of data and services from the ADS has changed your teaching efficiency (i.e. compared to if no ADS existed)?

Asked to what extent their use of ADS data and services had changed their teaching efficiency, answers ranged from no change (0%) to a more than 90% gain. The reported mean was a 32% efficiency gain (median 20%). Given the time spent teaching reported in Q30, this translates to an efficiency gain equivalent to 3 hour a week at current activity times.

Again these efficiency impact responses seem high, and while it was intended to refer to all teaching and preparation time, it is possible that respondents interpreted the question as referring to the time spent with ADS data alone. If so, the reported efficiency gain would be equivalent to less than 1 hour a week at current activity times.

Contingent valuation

The final three questions were addressed to all user types and involved the use of contingent valuation techniques using stated preferences. A description of the method can be found in section 4.2.

Contingent valuation inevitably generates some protest responses, and it is important to include an open-ended question/comment opportunity for respondents to register their protest. This also makes it possible to identify and exclude protest answers. Eleven protest answers were identified in this case, leaving a total of 181 responses. ADS users were asked how much they would be willing to

accept in return for giving up their access to ADS, and how much they would be willing to pay for access to ADS.

It should not be surprising that individual responses vary widely, as different users use different data and services and use them with very different frequency. Some regular users may base their work on ADS data and services, while others may be occasional users who dip into ADS data as a matter of passing curiosity. Consequently, they are likely to value ADS data and services quite differently.

Q34: Imagine that the ADS was no longer open to new users. You have the option to either carry on using the ADS or to sell your rights as an existing user to a third party. If you sold your use of the ADS, what is the MINIMUM amount that you would be willing to accept as an ANNUAL payment in return for giving up ALL of your use of the ADS?

The minimum amount that respondents were willing to accept varied widely, ranging from £0 to £1 million. Both extremes present some difficulties for interpretation.

- Experience in this and similar surveys suggests that some people say they are willing to accept nothing in return for their access rights because they believe that access should be free, not because they do not value their access. Hence we explore the results including and excluding the £0 responses, of which there were 18.
- The upper end value of £1 million is high, 15 times higher than the second highest response. However, the respondent was a very frequent user of ADS data and services and said that their access (to four data types) was extremely important, lack of access would have a severe impact on their work, and a number of ADS services provided a high benefit. In short, a high value recorded in willingness to accept was consistent with responses to other questions. Indeed, the higher end responses were all matched by high levels of use and strong expressions of value and benefit. Nevertheless, we explore the results including and excluding this outlier response.

The mean value respondents were willing to accept was £6,807 and the median £150 (N=181). Excluding the zero responses the mean was £7,560 and the median £250 (N=163). Excluding the highest response the mean willingness to accept was £1,282 and the median £132. On balance, we are inclined to take the last (i.e. excluding the high-end response as an outlier) to be the most representative.

Q35: The ADS receives funding from a number of bodies and is committed to providing free access. For this question, however, please imagine that this funding ceased to be provided. In this hypothetical case, what is the MAXIMUM amount you would be willing to pay for access to data and services from the ADS? (If you do not value the ADS enough to pay, enter 0).

Respondents were then asked what was the maximum amount that they would be willing to pay for access to ADS data and services: (i) per year, as an annual subscription, and (ii) per use, on a pay-per-view basis. Again responses varied widely, from £4,000 per annum to £0 (N=191), and from £500 per use to £0 (N=181). The mean amount respondents would be willing to pay per annum was £196 (median £60), and the mean amount per use was £10 (median £3).

One check on the willingness to pay answers is the compare the annual amount with the pay-per-use amount multiplied by the frequency of use per annum. This generally reflects a willingness to

pay more per use, for the convenience and lack of longer term commitment to paying. However, more than a dozen respondents expressed some dislike of the 'pay-per-view' model, because subscriptions were easier to manage at the institutional level and are more familiar to them, as well as other objections. We found that:

- For the 181 respondents answering both elements of the question, the mean amount they were willing to pay per annum was £204, while the amount per use was £10. Multiplying £10 per use with the frequency of use per annum for each of these respondents individually, results in £449 (i.e. double the annual WTP).
- For the 181 respondents answering both elements of the question, the median amount they were willing to pay per annum was £60, while the amount per use was £3. Multiplying £3 per use with the frequency of use per annum for each of these respondents individually, results in £80 (i.e. on-third higher)

Hence, while there is a gap between annual and per use "willingnesses" to pay, it is in the expected direction and no greater than might be expected of the different revenue models.

As noted, willingness to accept expresses a value that is unconstrained, while willingness to pay is constrained by the respondents' capacity to pay. Given that some 55% of respondents are students, 27% are private/unaffiliated individuals, and 14% independent and unpaid, their capacity to pay is clearly constrained. For example, removing students results in a small increase in what respondents are willing (and able) to pay: to a mean of £210 per annum and a mean of £11 per use (N=158).

Whereas the students are willing to pay a mean of £161 per annum and a mean of £4 per use (N=23). Moreover, 11 respondents specifically said that the amount they or their organisation could afford was less than the value they put on the information/service.

Box A5: Selected comments on contingent valuation (excluding protest responses)

"I am a private individual, retired on a small pension. Every penny counts."

"Annual fee would be paid by company; pay-as-you-go fee is likely to never be recovered and would therefore be paid by me personally as part of general background studies/ response to queries"

"I can't afford more but if I was I would be willing to pay about double."

"Though several of the Society Council and Committee members use it, and possibly ordinary members, I doubt that would be prepared to pay for single use, although possibly for group use.

Personally it would be one more expense of many for which I do not get re-imbursed. As a voluntary worker, I would probably opt out."

"The amount I could personally afford. I also considered that much of the information I regularly use is available on the Heritage Gateway, PastScape. In all cases I would buy the HER data anyway, so I use ADS just to check I haven't missed anything and to access the reports in order to save time visiting the HER if I can."

"for personal use/skills update etc it is what I can afford ... for commercial work it is what a client might cope with. Clients are usually small developers, a few of whom resent spending anything at all on archaeology."

"If I were to have to pay to use ADS as an independent researcher gaining no monetary benefit from my research and publications it would just cost me money - a no brainer really."

"As an institution our budgets have been slashed. As this is on a user by user basis it is assumed that several accounts might be needed."

"I tried to think of the amount of data I get each visit and then how many visits I make a year. I then compared this to the amount charged by journal websites etc... I consider ADS to be more valuable as it provides access to such a range of data."

"At the moment our service relies on the ADS to maintain the digital archive If we did not have access to this service we would need to increase the capacity of the existing servers to host this information ourselves and staff time for on-going maintenance. As a charity we need to make sure that we keep subscription and pay per use costs to a minimum. This amount would be more if I was part of a non-charitable organisation."

"The annual subscription figure is the amount I would ask my funder or organisation to pay for me as a single user. I seriously doubt I would be able to afford an annual subscription, assuming that the subscription would be on the level of what many major journals now charge."

"Amount I would ask a funder or organisation to pay. Note that this is for access to the ADS as an exemplar of archiving and connectivity, rather than for specific datasets"

"Amount I can afford personally. I would have to estimate what I would be paid for my work by the institution and, in a time of tight resources, that is not always easy to do."

"Amount I would be willing with all the other calls I have for what is basically a hobby"

"My organisation would probably fund a small amount for access to the ADS, although I would have to strongly justify this. In terms of single use I would be prepared to pay a small amount to download a useful document, and again might be able to get some funding from my employer."

"If I had to pay my research could not continue. This would impact on a national archaeological research project ... - my work will be submitted to ... County Record Office for others to consult"

"It is related to my personal income; I am an unfunded postgraduate student without income (living off savings), I'm therefore more financially vulnerable than most. Nonetheless, given my status, ADS ranks highly compared to what I would pay for other services. To get a comparable amount of data without ADS would mean signing up to other expensive services, I therefore appreciate ADS being free."

"Amount I believe my employer would tolerate, given severe recent restrictions on what we are permitted to spend on fees. We would simply drop archaeological research from our services if it became too expensive, reducing the quality of our output."

"Information can just as easily be accessed via HER/NMR searches that would normally be required anyway, even if they charge. More value would be accrued if search facility was improved."

"Comparison with other large data providers for monthly subscriptions through to annual subs, e.g. genealogical websites."

Contingent valuation using stated preferences involves questions that can be difficult for people to answer. Many may simply not have thought about it before and be bewildered. Despite the careful wording of questions, some may be fearful that ADS services will be charged for, and if they say they are willing to pay a substantial amount of money it would not be long before they are asked to do

so. Consequently, there is a need for caution in interpreting these results. Nevertheless, the results obtained do seem to be reasonable and sufficient for further analysis.