The Value and Impact of Research Data Infrastructure

Neil Beagrie
(Charles Beagrie Ltd)

PASIG, Karlsruhe, September 2014
Value + Impact Analysis of Three UK Data Preservation Services

John Houghton (Victoria University) + Charles Beagrie Ltd

Methods applied to:

- Economic & Social Data Service
- Archaeology Data Service
- British Atmospheric Data Centre

*studies funded by ESRC, Jisc, Jisc/NERC respectively*
Presentation Overview

- Introduction - Measuring Value and Impact of Research Infrastructure
- UK Value and Impact Studies
  - ADS results
- Concluding Remarks
- UK Focus but the research and lessons are international and for digital preservation more generally
UK Research and Innovation

• Why Research matters
  – £3.5 billion a year currently spent on publicly funded (RCs) research generates an additional annual output of £45 billion in UK companies (Haskel and Willis 2010)
  – Wider non-economic value to policy, heritage, and knowledge

• Where disciplinary data centres and services exist they represent approx 1.4-1.5% of total research expenditure (OSI PC WG)

• Why does preservation of digital data matter for research?
Astronomy

• Data archives
  – “Central to astronomy today”
  – HST, 2MASS, and SDSS archival research is major contributor to scientific productivity

![Graph of HST Publications]

**Figure 1:** Number of annual publications using Hubble Telescope data. The publications have been divided into non-archival papers written by the original investigators (blue), totally archival publications not involving none of the original proposers (yellow), and papers that include data from multiple proposals with some being archival and some not (red). The number of archival papers has exceeded the number of PI-led papers since 2006.

![Graph of Most-cited HST papers 1997-2000]

**Figure 2:** Highly cited HST publications between 1997 and 2000. All 71 papers with more than 150 citations (as of March 2009) are included in the sample. Note the y-axis is logarithmic. As in Figure 1, the publications have been divided into non-archival (blue), totally archival (yellow), and partially archival (red) depending on whether the original proposers were authors on the paper. Totally archival papers make up 37% of the highly cited sample, which is slightly above the rate expected based on their frequency of publication during this period.

c/o R. White (STScI) and pp. 5-11, 5-12 of NWNHAA
Research Infrastructures and Impact

Data
Intellectual Capital

Training & Skills
Human Capital

Buildings / Equipment

Grid / ICT Networks

Staff

Technical / Organisational Environment
Organisational Capital

Professional Networks
Relationship Capital
Previous Work

Big Science and Innovation Study for BIS July 2013

• Desk review of c. 100 studies internationally;
• 3 studies highlighted to BIS as being particularly good examples of ‘good practice’ in the measurement of economic impacts:
  – Berkeley Lab 2010
  – Human Genome Project 2011
  – Economic and Social Data Service 2012 (authors)

[ Two further studies of BADC and ADS published later – we think/hope even better!]
Best Practice from ESDS study

- Applies range of methods;
- Includes counter-factual;
- Data collection tailored to different stakeholders: depositors, users, research, teaching;
- Data weighting - survey value responses weighted to reflect the overall pattern of use from weblogs;
- Case studies/ KRDS benefits illustrate benefits and impact pathways;
- Research- partnership with Service: adaptation to community and knowledge transfer.
Approaches and Methods

• We combined quantitative and qualitative approaches, to quantify the value and impacts of the data services and explore other, non-economic impacts.
• The combination is important in capturing and presenting the full range and dimensions of value.
• All three studies combined:
  • Desk-based analysis of existing literature and reports, looking at both methods and findings;
  • Existing management and internal data collected by the data services; and
  • Original data collection in the form of online surveys of users and depositors, together with semi-structured interviews.
The Value and Impact of the Archaeology Data Service
Economic
Key Findings - Economic

Series of different economic approaches used:

• The investment value of ADS is around £1.2 million per annum;

• ADS impact on research, teaching and studying efficiency: worth at least £13 million per annum;

• Contingent Valuation - What ADS data and services are worth to its users:
  – Willingness to pay around £1.1 million per annum.
  – Willing to accept around £7.4 million per annum.

• RoI scenarios - additional use over 30 years - a 2-fold to 8-fold return on investment.
Investment & Use Value (Direct)  
Contingent Value (Stated)  
Efficiency Impact (Estimates)  
Return on Investment (Scenarios)  
Wider Impacts (Not Directly Measured)  

---  

**Investment Value**  
£1.2m per annum  

**Use Value**  
£1.4m per annum  
(More than double)  
(ADS operational budget)  

---  

**Contingent Value**  
£1.4m per annum  
(More than double)  
(ADS operational budget)  

**Willingness to Pay**  
£1.1m per annum  
(Severely constrained)  
(by capacity to pay)  

---  

**Willingness to Accept**  
£7.4m per annum  

**Consumer Surplus?**  
(Could be up to £6m per annum)  
(on a Willingness-to-Accept basis)  

**Net Economic Value?**  
(Could be up to £5m per annum)  
(on a Willingness-to-Accept basis)  

---  

**User Community Efficiency Gain**  
[ADS data use]  
£13m per annum  

**User Community Efficiency Gain**  
[All activity time]  
£58m per annum  

---  

**Increased Return on Investment**  
[Additional Use]  
£2.4m - £9.7m  
(2.1 to 8.3-fold RoI)  

**Increased Return on Investment**  
[Non Recreate]  
£1.5m - £5.9m  
(1.3 to 5.1-fold RoI)  

**Additional re-creation costs of up to £1m**  
(2 to 6-fold RoI)  

---  

**ADS Regular User Community**  

**ADS Website User Community**  

---  

**Wider ADS Community**  

---  

**Society**  

---  

**Based on Deposit & Download Counts**  

**Based on Download Count**  

---  

**Based on User Count**  

---  

**Based on Data Spend**  

---  

**Based on Deposit & Download Counts**  

---  

**Wider Impacts (Not Directly Measured)**  

---  

**?**
Stakeholder Perceptions
Key Findings – Perceptions

• Qualitative analysis - interviews and survey comments reveal strong support for the ADS - many aware of the value of the services;

Workshop feedback on quantitative + qualitative interim findings:

• Stakeholders aware of value to them – less aware of value to others;

• Stakeholders were positive about seeing value expressed in economic terms - 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 approaches to measuring the value (and for its dissemination).
Safeguarding our heritage for the future
Preserving valuable archaeological data saves time and money

Professor Julian Richards UNIVERSITY OF YORK

The Archaeology Data Service (ADS) provides a digital archive for cataloguing and preserving data created through archaeological research. It was originally established as part of the Arts and Humanities Data Service in 1996 and became independent of this group in 2008. The resource has since gone from strength to strength, becoming invaluable to users both within and outside of academia and making a significant contribution to research efficiency, estimated to be worth £13m each year.

The ADS is a unique resource for the archaeological community; it houses a collection of digital resources created as a product of archaeological research, much of which would otherwise remain unpublished. As well as providing users with an accessible and reliable service, it preserves collections which tell the story of Britain’s past and of international historic sites.

JISC commissioned research carried out in 2012 found that the ADS has a broad user group which goes well beyond academia; whilst 38% of users are conducting academic research, 19% use ADS for private research; 17% for general interest enquiries; 16% are Heritage Management users and 8% are commercial users. 6% use it to support teaching and learning activities; and 1% use it for family history research.

The ADS is regarded as an invaluable resource saving users time and therefore money, and providing security for those who use the service to deposit their data. A significant increase in research efficiency was reported by users as a result of using the ADS, worth at least £13 million per annum – five times the costs of operation, data deposit and use. A potential increase in return on investment resulting from the additional use facilitated by ADS 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.

Feedback from users within and outside of academia is testament to its research value. The ADS can be used as a tool to support the UK’s archaeological community within and outside of academic settings. It provides both secure preservation and access to the impossible records of archaeological investigations. It not only provides valuable storage for expensive and often irreplaceable data, but also helps researchers to avoid high costs incurred through travel and the duplication of fieldwork. This economic value will increase as the service continues to grow and develop.

The ADS’s achievements were formally acknowledged in 2012 when it won two awards: the British Archaeological Award for Best Archaeological Innovation and the Digital Preservation Coalition’s Decennial Award. This highlights its value beyond the archaeological community, recognising its major contribution to the digital memory sector between 2002 and 2012.

1 Praise from ADS users, Times Higher Education Leadership & Management Awards.
2 Beagrie and Houghton, The Value and Impact of the Archaeology Data Service: A Study and Methodology for Enhancing Sustainability, September 2013.

Published by Arts and Humanities Research Council, Folens House, North Star Avenue, Swindon. Website: AHR.artsandhumanities.ac.uk. Arts and Humanities Research Council 2014. Published July 2014. Design by Kumba. Printed by JSI on paper contain high 60% post-consumer waste.
Impact Studies - conclusions

• Economic benefits exceed the operational costs
• A mix of qualitative and quantitative methods is important to capture and present the value of the Data Services
• The studies are increasing recognition of the value of the Data Services and digital preservation and data sharing generally
• Helping stakeholders see value to them and to others
• Need to extend work into other disciplines and types of research infrastructure
Further Information

Synthesis of the three studies:

http://repository.jisc.ac.uk/5568/1/iDF308_-_Digital_Infrastructure_Directions_Report%2C_Jan14_v1-04.pdf