Evaluation in the Arts and Humanities Data Service

Martin Wynne

The Author:

Martin Wynne is Head of the Oxford Text Archive and AHDS Literature, Languages and Linguistics. As well as working in the field of digital archiving and preservation, he is active in research in linguistics, and has worked as lecturer and researcher at universities in the UK, Poland and Germany.

Abstract

Purpose

There are a large number of free electronic resources available on the web, from various sources, including the outputs of funded research projects. There are not yet well-established procedures for the evaluation of the quality and usefulness of these electronic resources. This article examines two preliminary approaches to the problem.

Approach

An *Investigation into Free E-books* was carried out for the JISC to examine the potential for the use of free e-books in teaching and learning in Further and Higher Education in the UK. Questionnaires, a dissemination workshop and focus groups were used to obtain the research results.

Three possible approaches to evaluation are discussed, involving (i) the collection of high-quality resources in a trusted repository, (ii) large scale evaluation of available resources and (iii) evaluation by end-users. The

practical problems of each approach are discussed. The prospect of future large-scale digitisation projects and their possible impact in this field is also considered. It is concluded that it is necessary for end-users of electronic resources to get involved now in the planning stage of such initiatives to attempt to ensure that the outputs are in suitable formats and of the necessary quality.

In a separate section, the question of evaluation of electronic resources is briefly considered in the context of collections development in the Arts and Humanities Data Service (AHDS).

Findings

The conclusion is drawn that it is more cost effective to intervene early in the life-cycle of a digital resource and to help ensure good practice in the creation of the resource, than to attempt to evaluate and enhance a resource at the end of the creation process.

Originality

This paper draws on the unique experience of the AHDS as a central, national service for researchers in the Arts and Humanities creating and using electronic resources.

Introduction

The Oxford Text Archive (OTA) hosts *AHDS Literature, Languages and Linguistics*, part of the Arts and Humanities Data Service (AHDS). The AHDS is a distributed service offering support services to various arts and humanities subject areas through five subject centres. The services offered include the archiving, distribution and preservation of electronic resources; technical advice to creators of electronic resources; and advice to funding bodies, such as the Arts and Humanities Research Board (AHRB) and the British Academy. The AHDS is funded by the AHRB and the JISC. The AHDS is currently involved in several activities relating to the evaluation of electronic resources, and this article will focus on two such areas in which the OTA is currently active.

Evaluating free e-books

The OTA recently completed a report for JISC on the potential for the use of free e-books in teaching and learning in HE and FE in the UK. The report is now freely available at http://www.ahds.ac.uk/litlangling/ebooks/. Among the main findings of the report was the conclusion that there are some important barriers to the uptake of free e-books in these sectors. These barriers included the following:

- Lack of availability of a complete range of titles for any given course
- Doubts about quality assurance
- Lack of confidence in the persistence of availability of resources
- Costs involved in the cataloguing, archiving, management of resources
- Costs involved in computing support for users
- Poor design of free e-books and poor ergonomics of reading on screen

However, the large potential of this resources was clear and some important opportunities to promote the uptake of free e-books were identified. The project was recommended to the JISC that it offer support in various areas to promote the uptake of free e-books, including three of particular relevance to this article:

- Institute a system of quality assurance (of text integrity and metadata)
- Ensure the permanence of collections
- Support the professional, standardised cataloguing of electronic resources

One of the key findings of the report was that the lack of quality assurance associated with currently freely available electronic textual resources is a serious barrier to the use of these resources. The findings include a survey of currently available free resources and an analysis of the problems of finding and evaluating them for use in teaching and learning.

> It is evident from this brief survey of web sites that there are a vast number of free e-books available to the arts and humanities scholar. The problem this community faces is not one of quantity, but of quality. It is debatable what percentage of existing free e-books are of any practical use to the academic at present, but we would estimate that

> the numbers are low. In terms of quality and supporting bibliographic information, e-books created in academic departments or by research projects provide the best option to the academic user, while the multitude of 'enthusiast' sites provide little that can be used or trusted. (Berglund, Morrison, Wilson and Wynne, *An Investigation into Free E-Books*, http://www.ahds.ac.uk/litlangling/ebooks/)

In the focus groups there was a strong feeling that quality assurance was a problem. A handful of representative quotations from the focus group discussions are given below. There was a general concern about the quality of available resources:

> If you look at texts of Shakespeare on the internet the quality is appallingly low. It's lower than that of any printed text of Shakespeare since probably 1670. Humans have spent centuries building up ways of ensuring textural transmission as reliable, that problem texts are documented and ways of ensuring that the words are text or not, and a lot of that has gone by the board with e-text.

And it is not just amateur enthusiasts who are held responsible for this:

> Even academics are creating e-text and they suddenly throw away all their training and stick them up there with no indication of what edition they are following, or what editorial principles they are using. Not all of them but some.

The potential for re-editing and repurposing electronic texts can mean that there is a lack of accepted standards for electronic editions:

> I've got a concrete example of [poor text quality], which is an edition of Catullus, when someone chose to renumber the lines in all the poems with numbers which seemed good to them.

There is an issue of who is to offer quality assurance, and how:

> Quality is a big issue for us. We are very wary in the library of putting resources on our web pages for students to use, for students assume there is a quality endorsement there. We haven't got time to look at all of those e-books, have we? It's difficult. Plus we are not subject-specialist enough to know necessarily whether that's the best

> *translation, or best edition, there's not a lot of data around on the computer to tell you exactly where that text is coming from.*

As a result, quality assurance tends to be left to the judgement of the end user:

> *[as a librarian] you are not guaranteeing any level of quality – it's a bit like using websites from Google, isn't it? Users have got to make their own minds up, evaluate the material themselves. You haven't got the guarantee you get to a certain extent with print.*

There was strong anecdotal evidence of a lack of text integrity in free e-book products:

> *Strangely enough I just used an e-book yesterday and it falls right into this category of E-books that are created badly. Someone created an E-book from existing text and left out all footnotes, all bibliography, everything. At the end of the day I would say that rendered it useless -- you couldn't get any of the references, anything. I had to go to the library and take out a hard copy.*

> *I would like to mention that the first e-book I read was on a hand-held, was some sort of detective story and the last 5 pages or so were not there.*

This experience can be backed up by the informal surveys of available free web resources. Not surprisingly, pirate editions seem to be particularly prone to particularly poor quality editions.

In view of the importance given to quality assurance by potential users, and in order to start to address this problem, at the OTA we have considered it useful to develop some *ad hoc* procedures for evaluation of free e-books. A checklist has been drawn up for use at the OTA in evaluating e-books, and is given in full below.

Checklist for the Evaluation of Free e-Books

Is it what it says it is?

1. *Is the text really available and free to the user?*

2. *Check for existence of metadata;*

3. *If the metadata claims to conform to an external standard (e.g. Dublin Core) check its grammaticality, completeness and relevance;*

4. *Is the relevant information about the particular edition of the text present and accurate;*

5. *Are intellectual property issues covered in the metadata or text? Is the treatment accurate? Does the resource provider have the right to distribute the resource, and are the creators of the resource credited in the documentation or metadata or text (as appropriate)?*

6. *Check the accuracy and completeness of the metadata for individual texts, where the resource is a collection of texts or samples;*

7. *Where there is no metadata covering these issues, check in particular whether the following particulars are as expected: text, language, file type, text encoding format.*

8. *Where there is more than one file, check that all relevant resource files are present in the correct file structure (i.e. as documented), and that file naming conventions are suitable;*

9. *Assess the file format: is it as documented, is it valid according to the normal standards for that format and is this a suitable format for interchange, storage, use and preservation?*

Fitness for purpose

10. *Assess the appropriateness of the format for the intended purpose (e.g. quality of design, representativeness, sampling etc);*

11. *Duplication: is the text available elsewhere in a usable way, or even in a more useful form?*

12. *Is this text likely to be of use in HE and/or FE? If so, where and how?*

Text Integrity

13. *Check integrity of textual material (Are bits missing? Have some elements been silently omitted?);*

14. *Check for erroneously repeated textual material;*

15. *Are footnotes, endnotes, other editorial interventions encoded, and if so are they done correctly;*

16. *Are front matter such as foreword, preface, introduction etc. correctly encoded;*

17. *Are appendices such as afterword, endnotes, bibliography present and correctly encoded?*

Text format and encoding

18. *Assess the character sets which are used: is the character set as per the documentation, if this exists? Is it suitable? Are there any invalid characters or entities?*

19. *Assess the choice of textual markup scheme: is it suitable for interchange, use and migration?*

20. *Validate the textual markup and evaluate the semantic accuracy and appropriateness (e.g. are chapter or paragraph tags correctly used?);*

21. *Validate the design, markup and annotation against external criteria; check that it actually works with software for the processing of the format, e.g. check XML is valid and parses, check Acrobat Reader can read PDF files;*

Factors external to the text

22. *Search for and follow up documented bug reports, comments and reviews which may be available at the repository or elsewhere;*

23. *Contact, and maintain ongoing dialogue with, the resource provider (where this is possible) to ensure the accuracy and completeness of metadata, and to manage enhancement of the resource where necessary.*

However, it should be pointed out that this was developed more a checklist for the ingest of e-texts to the archive, than as a procedure for evaluating external resources for use in teaching and learning. Creators of the resource come to the archive to deposit it, and it is usually possible to enter into a dialogue with them to obtain information about the resource and the process of its creation.

One of the problems with resources found on the web via search engines is to work out their provenance, and it is not normally possible to find someone who can be asked questions about the resource..

Who evaluates, and how?

The problem then is that there are a huge number of potentially useful electronic texts freely available on the web, but it is difficult to evaluate their quality and usefulness .The problem may be partly due to inflated expectations for the web, or a misunderstanding of what it is. Users want to use it as a library, but it is more like a city, or a virtual world, where there are many sources of texts, equivalent to booksellers, evangelists, people selling stolen goods, practical jokers, rubbish dumps, as well as the university and the library. So, while web gives us the opportunity to quickly find and download an electronic text, it is hard to know whether it has come from the virtual equivalent of the national library or the rubbish dump. And even the places which we know can be trusted as repositories of printed books, such as the university library, when they operate in the virtual world they are not on such firm ground. The technical infrastructure and the knowledge and experience necessary for dealing with electronic resources may be lacking. There are several possible approaches to evaluating such resources. These can be characterised as follows:

1. collect quality resources in a trusted repository;
2. evaluate resources held in disparate sites;
3. offer tools for users to evaluate resources.

The first approach is the one taken by the Arts and Humanities Data Service, and by some libraries, which is to ingest high-quality materials into their collections and to act as a trusted repository, offering their imprimatur. This is part of a suite of services offered by the AHDS, and also including advising resource creators, assessing research grant proposals, archiving and long-term preservation. The AHDS' evaluation and distribution services cannot however solve the more general problem of evaluating the quality of available

resources. It is not feasible, or even sensible, for the AHDS to attempt to secure the deposit of all useful free e-books. The plethora of such resources on the web is both enormous and transient, and the administrative tasks involved in obtaining permissions would be enormous, even before approaching the evaluation tasks. In preference to this, the AHDS has a collections development policy which prioritises ingest of outputs of funded research in Higher Education in the UK. The focus here is on the preservation of resources, as they may be at risk after project funding has expired, and these resources can, because of their provenance, be reasonably be expected to be of a high quality. Evaluation and quality assurance procedures are however being developed for these resources, and are discussed below in the following section.

So, while such an approach can give some offer of quality assurance for the resources in the collections, it cannot offer comprehensive collections of resources to the user, and there will always be resources discovered elsewhere which the user needs to evaluate.

The second approach is to offer an evaluation of the resources held in various places on the web. This basically what the Resource Discovery Network does in the UK, although the granularity of their system of categorisation is not suitable for evaluating individual texts. The RDN hubs catalogue, and to a certain extent evaluate, websites. So an RDN catalogue entry for a website which offers lots of free e-books will contain a general description and some subjective evaluation, but not a detailed evaluation of each e-book hosted on the site. This is perhaps an area that it would be useful for the RDN to expand into if the resources were made available.

The third approach would be to offer a checklist, such as the one listed above, for end-users, or for LIS professionals working closely with the end-users, to apply to resources which they find. This has the advantage of being a more practical proposition than trying to 'evaluate the web', but risks duplication of effort because many popular resources will be evaluated separately in similar ways for similar purposes, and this experience will not shared. A method for sharing peer reviews would address this problem. There would however still be the problem that such a peer review forum would be referring to a resource

held elsewhere, and ensuring the viability of the link or the integrity and persistence of the resource could not easily be guaranteed.

It should also be pointed out that while the above is a survey of current challenges, it is possible that the landscape will change in fundamental ways in the near future. A future development which is already on the horizon is the mass digitisation by professionals of the holdings of research libraries. Project Ocean in the USA already promises this, in the shape of the digitisation of all out-of-copyright holdings of Stanford University Library. This threatens to make redundant any attempts to come to grips with the mass of data of dubious provenance on the web, as one can hope that these texts will be replaced by new high-quality versions of known provenance. Electronic deposit in copyright libraries in the UK also offers the possibility of a large trusted repository, at least of future digital publications. However, it remains to be seen whether these projects will be successfully concluded, or how long they might take. For the outputs to be of sufficient quality to be used in teaching, learning and research in higher education, it will also be necessary for the products to be well encoded in suitable file formats, with adequate descriptive metadata. It seems likely that the most effective way to ensure that good practice is followed in initiatives of this type is for the intervention of academics and information professionals in the planning stage. This would be more cost effective than attempting to evaluate the outputs.

Evaluating accessions to *Arts and Humanities Data Service* Collections

The Arts and Humanities Data Service collects, catalogues, archives and preserves electronic resources created by or useful to researchers in UK Higher Education in arts and humanities subject areas. These resources must meet minimum thresholds in terms of quality, documentation and the appropriateness of the encoding formats. The precise requirements are currently under development, and much progress has been made in specifying documentation requirements and appropriate file formats. The question of the quality of the resource is a more difficult proposition, as this

must normally be done by someone with detailed knowledge of the highly specific subject area. In most cases, the provenance of the data from a UK HE institution or source of similar standing has to take the place of detailed quality inspection.

The best guarantee of quality, at least at the technical level, is the involvement of the AHDS from the early stages of the project life cycle. This approach is encouraged by the AHRB, and applicants for funding from this body who are proposing to create an electronic resource are asked to contact the AHDS at an early stage to discuss their project plan. The technical appendix of the funding proposal is then also assessed by the AHDS as part of the competition for funding, and must demonstrate that he techniques and technologies proposed are fit for purpose and that the project planning is realistic. It is then hoped that ongoing contact between the funded project and the AHDS will help to ensure that good practice is followed and that the two parties can learn from each other's experiences. At the conclusion of the project, it is required that the completed resource be deposited with the appropriate AHDS centre. At this stage, if all has gone well, good practice has been followed and the ingest of the resource is unproblematic.

However, even in what is generally considered an ideal, or at least 'good practice', scenario, where the archivists and technical experts have input form the start, the timely delivery of a well-documented, high quality resource with appropriate technical characteristics is not guaranteed. This points to the complexity of the problem, and perhaps to the futility of attempting to ingest and use almost randomly discovered resources from the web. Most of the problems can be addressed with improved procedures and communications between the parties, and it should be stressed that improvements are expected to come mainly through better early interventions, not through improvements in the evaluation procedures.

There is also a further problem faced by the OTA in particular, which has been collecting electronic texts since 1976. The legacy data in the archive was not collected with any coherent evaluation of quality, and the text encoding conventions and formats have in some cases become outdated. The OTA is embarking on a full-scale evaluation of the legacy data held in the

archive. Over the years, many different standards and practices have been applied to the creation of these texts. We aim to make the archive more accessible for the academic community by updating and enhancing usability of the texts. This will allow users to define a set of texts according to their own specifications and obtain concordances via a simple online interface. The first stage in this process of standardising the text encoding and markup is an evaluation of each of the texts. A detailed workflow and checklist have been devised and employed in order to evaluate and then discard, replace or convert each text. So far two pilot studies have been completed, the project plan has been fully developed, and the work is ongoing.

Reference

Berglund,Y., Morrison, A., Wilson, R. and Wynne, M. (2003), "An Investigation into Free E-Books", available at:
http://www.ahds.ac.uk/litlangling/ebooks/