



JISC Project Plan

Overview of Project

1. Background

- 1.1 Over the course of the last 8 years the Centre for Computing in the Humanities (CCH) at King's College London has been directly involved in four projects focussed on Anglo-Saxon studies. The four projects represent a broad range of scholarly disciplines and perspectives:
 - The Prosopography of Anglo Saxon England (PASE) aims to provide a searchable database of all recorded people living in early medieval England;
 - The Language of Landscape ('LangScape') project is a comprehensive, searchable corpus of Anglo-Saxon boundary clauses - surveys drawn up by those who lived in, owned or worked the land, and who described it in their own words;
 - The Electronic Sawyer (eSawyer) is a revised online edition of Peter Sawyer's seminal 1968 catalogue of all extant Anglo-Saxon charters;
 - The Anglo-Saxon Charters (ASChart) project was a 6 month British Academy pilot project to develop a rich XML encoding model to represent the diplomatic discourse within Anglo-Saxon Charters.
- 1.2 At present the projects are independently managed and funded; the different academic goals of each are reflected in a variety of underlying technical approaches (PASE is database-driven whilst eSawyer is a dynamic XML publication, for example).
- 1.3 Apart from the obvious correlation of subject domain, all of the projects are implicitly related by a common focus on the Anglo-Saxon Charters as a collection of core primary sources. Significantly a universal system of identifiers already exists for uniquely identifying the 1875 known charters. These are known as Sawyer numbers, and were conceived by Peter Sawyer in his *Annotated List and Bibliography* of 1968, of which the eSawyer project is a new version. This common point of reference provides in effect a linking mechanism that is already well understood by the scholarly community. Our aim in this project is to build upon this and other points of potential interconnection to develop a new, highly interconnected hybrid resource – the Anglo-Saxon Cluster (ASC) – to unlock and bring together the repositories of knowledge embodied in each of the four component resources.
- 1.4 One of the key premises of the Anglo-Saxon Cluster will be that it should enrich and increase the visibility and utility of the component resources, without requiring them to be fundamentally altered in any way.

2. Aims and Objectives

2.1 Aims

- to develop a new web-based digital resource articulated around the Anglo-Saxon charters as core material, through which the data and the corresponding metadata embodied in each of the component projects will be available together in a thematic cluster.

- to build up an unprecedented picture of the relationships and associations which *implicitly* exist between the data in each component resource, but which are masked by the fact that each project has a distinct repository and interface.

2.2 Objectives

The project will achieve its aims by:

- Assessing the available and developing technological approaches to data aggregation
- Scoping the potential for charter-based integration of the four base projects
- Assessing user requirements, including researchers and the wider public
- Developing at least two models and prototypes for aggregation/integration
- Updating all charter-related encoding to TEI P5 standard
- Testing and disseminating the project website and tools.

3. Overall Approach

3.1 Strategy/project structure

The project will take as a starting point the body of Charter texts and the encoding model assembled for the ASChart project. The texts currently exist in XML encoded according to version P4 of the Text Encoding Initiative (TEI) guidelines; these will be brought up to the P5 standard.

As a starting point the Charter encoding will be modelled so as to represent the diplomatic discourse (i.e. mainly recording of formulas such as proem and invocation) and to refer to relevant external authority data (e.g. markup of occurrences of names and roles of persons mentioned in the charter), but, by using the TEI P5 standard, it will be conceived so as to allow for extension potentially useful for encoding additional structural and semantic distinctions.

The development of an appropriate encoding model will also take into consideration the preparatory work undertaken by the Charter Encoding Initiative (CEI), an international community interested in developing encoding models to capture the complexities of medieval and early modern charters.

Within the encoding model, particular analysis will be dedicated to those components that constitute potential interconnections with data and metadata that are external to the text. Names of individuals and locations, for instance, will be encoded appropriately, so as to allow for thorough reference systems to the PASE project. Any element that contains metadata information (e.g. content of the `teiHeader`) will be also a suitable potential candidate for creating connections.

The use of TEI P5 and the ODD documentation language will provide a robust framework that, once documented and published, could be extended and adopted by others interested in the encoding of charters within and outside the TEI community.

A functional specification will be developed during the first phase of the project based on an analysis of the data available in each underlying resource. A high priority throughout the development process will be ensuring a highly iterative process of prototype interface development and feedback from the extended project team and core resource project partners (following the Agile software development methodology). This process will result in a final functional specification and wireframe diagrams, which together will form the reference documents for the remainder of the web application and interface development.

Two technical approaches will be assessed and used as the basis of designing models for aggregator systems. The first model is based on the development and integration of a Web Services client into each of the four underlying projects. This would be implemented using

the Apache Axis2 web services framework (WSF). The second model involves a more direct ingest process, with four instances of Apache Lucene (one per project) interfacing directly with the individual project data repositories. Other models may suggest themselves during the research and analysis phase of the project, but these will be explored in depth.

Following the assessment of the models, one will be selected as the basis for developing a prototype aggregator. Alongside the prototype, one or more user interfaces will be developed to enable complex queries to be tested and the results analysed.

3.2 Issues

The development of the charter encoding model will take into account the work to date of the international Charter Encoding Initiative (CEI), whose interests cover all types of charter from all cultures and periods, and will contribute to the further development of that work. The web interface for the project will be expressed in XHTML 1.0, and CSS level 2.1, and will be developed to comply with priority AA of WAI WCAG 1.0. Charter documents will be stored in a repository of XML encoded according to the TEI P5 standard. Web services will be described using Web Services Description Language version 2.0 and published in SOAP 1.2 encoded XML.

3.3 Scope

In scope: The scope of the project is indicated by the foregoing sections and by the deliverables listed below (4.1).

Out of scope: The focus in the project is on researching key issues and approaches, and in developing at least two prototypes. It will not be possible to build more than prototype systems nor to develop a wider range of prototypes to represent all the different models for data aggregation that are emerging.

3.4 Critical Success Factors

- CCH supports and if possible extends the project prototypes beyond the life of the project;
- The project work forms the basis for future research and development proposals;
- A useful set of use cases is developed for wider reference;
- A valuable contribution is made to the wider community interest in data aggregation methods;
- A valuable contribution is made to wider research community engaged in the encoding of charters and similar documents, e.g. the Charter Encoding Initiative.

4. Project Outputs

4.1 Tangible Deliverables

Project Planning

- Project web page
- Project plan
- Progress reports (x2)
- Final report
- Completion report

Server Infrastructure

- All four base resources plus ASC prototypes on new servers

Scoping and technical research

- User scenarios (use cases)
- Analysis and report on system design for the four base resources
- Report on aggregation/integration methods and implications
- Bibliography/resource listing of relevant research and projects
- Functional specification of prototypes

Prototypes

- Draft TEI P5 schema for encoding Anglo-Saxon charters
- Two prototype data aggregator/integrator models
- Development of prototype system based on selected model
- Development of corresponding user interface(s)
- User testing of prototype and interface(s)
- Report on user testing
- Specialist seminar on prototypes and report

Digital Publications (on project website)

- Prototype interrogation system
- Draft TEI P5 schema for encoding Anglo-Saxon charters
- User scenarios (use cases)
- Project reports

4.2 Knowledge and experience

The project will build on and develop the team's knowledge and experience of:

- The four base Anglo-Saxon projects and the information and technical structures on which they are based
- Metadata systems for complex digital publications, including TEI XML encoding
- Research methods and needs in medieval studies
- Rapid prototyping of digital publication interfaces.

5. Project Outcomes

The project will provide:

- practical prototype systems with guidance for making complex integrated searches across four resources holding related materials
- pointers for how to extend the scope of the systems to include other Anglo-Saxon resources, and other medieval resources
- indications of how fully developed systems of this kind will benefit both research and teaching
- indications of how an aggregated resource may be used by the wider public
- an evidence and experience base for further research
- a rich encoding model for the TEI and medieval charters communities

6. Stakeholder Analysis

The major stakeholders are represented on the Advisory Board, Project Team and/or User Group:

Stakeholder	Interest / stake	Importance
Project partners in Anglo-Saxon projects (Nelson, Keynes, Baxter, Jenkyns)	Content creators and users	High
Wider community of Anglo-Saxon scholars	Users	High
Centre for Computing in the Humanities	Technical developments	High
Those involved in other Anglo-Saxon projects	Similar/cognate requirements	High

Those involved in other medieval charters projects	Similar/cognate requirements	High
--	------------------------------	------

7. Risk Analysis

No	Risk	Probability (1-5)	Severity (1-5)	Score (PxS)	Action to Prevent/Manage Risk
1	Failure to complete to time/budget	2	5	10	Realistic project plan. Regular monitoring. Effective escalation procedures.
2	IT solution not fit for purpose	1	5	5	Thorough specification produced. Comprehensive user testing completed in a timely fashion.
3	Website inaccessible for some users	1	5	10	Specification to include reference to WAI protocols and JISC-endorse standards. Testing to include performance over remote connections via a variety of servers.
4	Failure to recruit personnel	2	5	10	A number of key personnel already in place. Outstanding recruitment in progress; work packages adjusted accordingly.
5	Loss of key personnel	2	5	10	Good staff contracts and working conditions. Retention of some contingency in contract length to account for early departure.

8. Standards

Name of standard or specification	Version	Notes
XHTML	1.0	Strict or Transitional. All HTML will be expressed in this format.
CSS	2.1	For standards compliant browsers only. Non compliant browsers will be delivered specific CSS style sheets which may employ proprietary CSS properties as necessary to ensure wide browser compatibility.
XML	n/a	Work based on version P5 of Text Encoding Initiative guidelines. Work of Charters Encoding Initiative taken into account.
WAI WCAG	1.0	The project website will be upgraded to priority A compliance minimally; the project aims to achieve AA compliance.
Java Servlet / JSP	2.5 / 2.1	The existing PASE web application is already implemented using these technologies.
JavaScript	1.6	Deployed using the JQuery framework.
MySQL	5	The existing PASE web application is already implemented using this technology.
SOAP	1.1	Web services transfer protocol

9. Technical Development

9.1 Charter Encoding Model

The XML encoding of the ASChart charter texts will be brought up to the TEI P5 standard.

The Charter encoding will be modelled so as to represent the diplomatic discourse – e.g. proem and invocation – and to refer to relevant external authority data (e.g. person names and roles), but the TEI P5 conversion will be conceived so as to allow for extension to encode additional structural and semantic distinctions.

The development of the encoding model will take into consideration the preparatory work undertaken by the Charter Encoding Initiative (CEI).

Within the encoding model, particular analysis will be dedicated to those components that constitute potential interconnections with data and metadata that are external to the text. Names of individuals and locations, for instance, will be encoded appropriately, so as to allow for reference to the PASE project. Any element that contains metadata information (e.g. content of the `teiHeader`) will be also a suitable potential candidate for creating connections.

9.2 Prototype development

A functional specification will be developed to reflect the analysis of the data in each underlying resource. This will in turn form the basis of the initial prototype interface development, which will be refined iteratively using feedback from the extended project team and core resource project partners (following the Agile software development methodology). The initial process will result in a functional specification document including wireframe diagrams.

The data for the Anglo-Saxon Cluster will be regularly harvested from each of the underlying projects and aggregated into a offset repository which can be used for generating unified indexes for browse-type displays, and for search indexing.

Prototypes will be built to test two different technical approaches to aggregation and exposure of metadata from each project:

The first model involves the development and integration of a Web Services client into each of the four underlying projects. This will be implemented using the Apache Axis2 web services framework (WSF), primarily because it will gracefully accommodate the different web application frameworks in place for each of the core projects (which are built upon Apache Cocoon and XML, and/or Java Server Pages (JSP) and MySQL). Once the WSF interfaces are in place for each project, the data aggregation and indexing for the ASC would be handled by Apache Lucene, which will manage a unified index of all four projects. This approach will require some enhancement of the underlying resources.

The second model involves a more direct ingest process, with four instances of Apache Lucene (one per project) interfacing directly with the individual project data repositories. The task of exposing the data via WSF (again using Apache Axis2) for external use will be taken on by the ASC itself, and in theory no modification at all will be required to the underlying resources.

9.3 Web application and Interface

The ASC web application will be developed and generated using a number of standards-compliant tools developed in-house at CCH. The public facing web pages will be powered by a web publishing framework known as xMod, which runs within the Open Source Apache Cocoon application. Indexing and searching functionality would be developed and delivered using the Ereuna search framework, which is built primarily upon the Apache Lucene search engine library - a key component of the both of the possible data aggregation models we might use (see 6.4).

The interface layer will be implemented using sUPL (simple Unified Presentation Layer). sUPL is a CSS/XHTML based templating framework which allows rapid prototyping of complex web interfaces in full compliance with the relevant standards and in accordance with industry best practice for usability and accessibility.

10. Intellectual Property Rights

The Anglo-Saxon Cluster will deliver users to the underlying resources via the Cluster discovery functions. The IPR in all the content will continue to reside with the 4 underlying resources: PASE, LangScape, eSawyer and ASChart. The branding of the source data will be respected and it will always be clear to a user which underlying resource is being delivered or linked. CCH will retain IPR in the Cluster and aggregation tools.

Project Resources

11. Project Partners

11.1 Partners and contributions

Project partners in the four base Anglo-Saxon projects are the key partners in this project. These include Professor Dame Janet Nelson (PASE, LangScape and ASChart), Prof Simon Keynes (PASE, Electronic Sawyer and ASChart) and Dr Stephen Baxter (PASE). They will develop an initial set of user scenarios, and will comment and advise on the functional specification of the prototypes. *Contact: Janet Nelson*

The Anglo-Saxon research community will contribute to the development of user scenarios (use cases) and the testing of the prototypes. *Contacts: Janet Nelson, Harold Short.*

CCH will be responsible for analysis, development, implementation, and management of the technical infrastructure to support the project. CCH will participate in the development of the user scenarios and evaluation of feedback from prototype testing. *Contact: Paul Spence*

12. Project Management

12.1 The project will be managed in accordance with PRINCE2 methodology and overseen by an Advisory Board.

12.2 Advisory Board

The project board, responsible for strategy, monitoring progress and for ensuring that the projects is completed to time and within budget, will comprise:

- Co-Chairs of Advisory Group: Prof Dame Janet L Nelson; Prof Harold Short
- Project Director: Paul Spence (CCH)
- Technical Research Director: Paul Vetch (CCH)
- Board members: Prof Simon Keynes; Dr Stephen Baxter; John Bradley

The Advisory Board will meet at least once every two months, or more frequently if necessary.

12.3 Project Team

The project team, responsible for delivering the prototypes and documentation, will meet every two weeks to discuss progress and to resolve ongoing project challenges. It will comprise:

- Project Manager: Paul Spence (CCH)
- Technical Research Director: Paul Vetch (CCH)
- Technical project officers (CCH): technical design and build

12.4 User Group

A user group will include members of the Anglo-Saxon research community, from senior researchers to PhD students. A subset of the group will be invited to one of the workshops, and all members will be invited to test the prototypes and provide feedback.

13. Programme Support

n/a

14. Budget

See Appendix A

Detailed Project Planning

15. Workpackages

See Appendix B

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
March 09	Achievements against aims and objectives	Is project management effective? Are the stakeholders on board? Are work packages realistic? Do we need to change the plan?	Observation; consultation with project personnel and Advisory Board	Project is on target Project staff and Advisory Board are satisfied
March 2010	Outcomes and impacts	How many people visit the website? Feedback from use of the prototypes? Has further work followed from the project – e.g. new projects or proposals? Are the use cases being extended and/or enhanced?	Web stats Feedback from on the website Observation & consultation Observation/ assessment of feedback	100 user sessions per month Feedback indicates engaged use of prototype systems At least one project or proposal. At least ten further use cases

17. Quality Plan

17.1 Principles

The person responsible for each deliverable (see Work Packages) will ensure that it is quality assured according to this plan.

Project deliverables will be reviewed or tested by members of the project team, advisory board and user group as appropriate (see below).

Once reviewed, the project director will seek formal approval by the Advisory Board.

17.2 Work Package 1: Project Planning

PROJECT DOCUMENTATION					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.5.09 – 30.12.10	Relevant templates completed: Content is accurate: Language accessible	Review by Advisory Board (AB)	AB approval	Project Director (PD) Technical Research Director (TRD)	n/a

17.3 Work Package 2: Server Infrastructure

FOUR BASE RESOURCES MOVED TO NEW SERVERS					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
28.02.09 – 30.09.09	Allows efficient continuing use of base resources	Review by CCH systems management team; load testing; user testing	CCH signoff; acceptable performance under load; user acceptance	TRD	JMeter

17.4 Work Package 3: Scoping and technical research

USER SCENARIOS (USE CASES)					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
30.11.09	Reflects project scope and objectives; scenarios are clear;	Review by Advisory Board	Advisory Board approval	PD: scope and relevance TRD: technical assurance AB: clarity & relevance	n/a

ANALYSIS AND REPORT ON FOUR BASE RESOURCES					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
30.06.09	Appropriate coverage of all four resources; highlighting of common metadata	Review by PD & TRD	PD & TRD approval	PD: relevance TRD: technical assurance and completeness	n/a

REPORT ON AGGREGATION/INTEGRATION METHODS					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.07.09	Coverage of main current work	Review by PD & TRD	PD & TRD approval	PD: relevance TRD: technical assurance and completeness	n/a

BIBLIOGRAPHY & RESOURCE LISTING					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.07.09	Coverage of main current work	Review by PD & TRD	PD & TRD approval	PD: relevance TRD: technical assurance and completeness	n/a

FUNCTIONAL SPECIFICATION					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.08.09	Meets user requirements; compliance with JISC standards; language accessible	Review by Advisory Board	Advisory Board approval	PD: accessibility and relevance TRD: technical assurance AB: accessibility and relevance	n/a

17.5 Work Package 4: Prototypes

DRAFT TEI P5 SCHEMA FOR A-S CHARTERS					
Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.03.09	Reflects structure and content of A-S charters; conforms to TEI P5	Review by Advisory Board; Review by TEI Council member; Review by CEI	Approval from Advisory Board, TEI and CEI	PD: general standards compliance; liaison with TEI and CEI TRD: technical compliance	n/a

AGGREGATOR/INTEGRATOR MODELS					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
30.09.09	Meet functional spec; meets usability requirements	Review by PD, TRD; Review by Advisory Board	PD, TRD approval; Advisory Board approval	PD, TRD	n/a

AGGREGATOR/INTEGRATOR PROTOTYPE DEVELOPMENT & TESTING					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.10.09 – 30.11.09	Meets functional spec; meets usability requirements; corresponds to selected model	Review by PD, TRD; Review by Advisory Board; tested by selected users	PD, TRD approval; Advisory Board approval; user feedback	PD, TRD User group	n/a

INTERFACE DEVELOPMENT & TESTING					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
31.10.09 – 30.11.09	Meets usability requirements; enables use cases to be completed	Review by PD, TRD; Review by Advisory Board; tested by selected users	PD, TRD approval; Advisory Board approval; user feedback	PD, TRD User group	n/a

USER TESTING OF PROTOTYPE & INTERFACE					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
30.11.09	Successful user testing; completion of original and new use cases	Review by User Group and Advisory Board; user group feedback	User Group and Advisory Board approval	Board Chair and PD: quality of user experience TRD: technical compliance; User Group: value of use	n/a

REPORT ON USER TESTING					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
18.12.09	Clear analysis and report of user experiences and feedback	Review by User Group and Advisory Board	User Group and Advisory Board approval	Board Chair and PD: quality of user experience TRD: technical compliance; User Group: value of use	n/a

SPECIALIST SEMINAR & REPORT					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
30.11.09	Engaged discussion of prototype and broader issues; accurate report	Review by Advisory Board and User Group	Advisory Board and User Group approval	AB Chair; PD	n/a

17.6 Work Package 5: Digital Publications

PROTOTYPE INTERROGATION SYSTEM					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
18.12.09	Technical compliance; usability	Project review by project team and Advisory Board	Project feedback	AB & PD: satisfactory performance & feedback TRD: technical compliance	n/a

DRAFT TEI P5 SCHEMA FOR A-S CHARTERS					
Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
18.12.09	Links to TEI and CEI work & websites	Review by project team	Approval from PD & TRD	PD: TRD:	n/a

Output		USER SCENARIOS			
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
18.12.09	Stable as publication	Review by project team and Advisory Board	Approval by Advisory Board	AB & PD: TRD	n/a

Output		PROJECT REPORTS			
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
18.12.09	Complete & stable	Review by project team and Advisory Board	Approval by Advisory Board	AB & PD: TRD	n/a

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
Aug 09	Existing email discussion lists, e.g. ANSAX-L.	Anglo-Saxon researchers & students	Raise awareness of project	Give us feedback
Aug 09	Create public mailing list	Anglo-Saxonists and medieval historians; wider public	Raise awareness of project	An interesting way to explore resources is in creation
Nov 09	Email drop	Academics in Anglo-Saxon studies	Encourage use of the new website	Check out this fantastic teaching resource
Dec 09	Initiate 'Charter of the Month' feature	Anglo-Saxon research community, , general public	Encourage use of the website	Try out the website to do research yourself
Nov/Dec 09	Create links on base resource websites	Existing users of base resources, general public	Encourage use of the new website	You can farther and do more

19. Exit and Sustainability Plans

Project Outputs	Action for Take-up & Embedding	Action for Exit
Prototypes	Hosted by CCH; work will be extended as resources permit – e.g. new funded project	TRD to ensure prototypes are hosted and accessible; PD to oversee preparation of proposals
Public interface	Hosted by CCH; promoted via project websites, email lists, conferences	TRD to ensure website is hosted and accessible: AB to ensure promotion via all means
TEI P5 schema for AS charters	Will be used by CCH projects; will be discussed with TEI and CEI	PD and TRD to ensure continuing contacts

Project Acronym: ASC
Version: 1.0
Contact: PS
Date: Aug 09

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Prototypes	Hosted by CCH; used by a range of users		
Public Interface	Hosted by CCH; used by a range of users		
TEI P5 schema for AS charters	Used by CCH and various projects		

Appendixes

Appendix A. Project Budget

Appendix B. Workpackages