



AustESE: eResearch Tools to Support the Collaborative Authoring and Management of Electronic Scholarly Editions

Anna Gerber¹, Roger Osborne¹, Jane Hunter¹

¹The University of Queensland, St Lucia, Australia, {a.gerber | r.osborne | j.hunter}@uq.edu.au

INTRODUCTION

Scholarly editions contribute to and support research in the humanities by providing accurate reading texts of works of literary, historical, theological, and philosophical significance. In addition to the reading text, a scholarly edition also includes historical and textual essays, explanatory notes, appendixes, and a scholarly apparatus that provides access to alternative readings in other versions of the work. Computer-assisted scholarly editions have been appearing for decades, but most editions continue to be published in book form, and most electronic editions do not extend beyond the traditional book model.

Since the mid-1990s, scholarly editors have experimented with Web-based electronic editions, producing exemplary models such as The William Blake Archive, The Rossetti Archive, The Mark Twain Project, Nietzsche Source, Electronic Kierkegaard and the Samuel Beckett Digital Manuscript project. But, despite such advances, these previous efforts provide one-off handcrafted solutions that address project-specific issues but don't facilitate re-use or interoperability of digital tools and data [1]. In recent years, it has been widely acknowledged that there is a critical need for a set of re-usable services to support the collaborative authoring, editing, visualization and publishing of scholarly editions [2]. To achieve this in the Australian context, the AustESE Project [3] is integrating and extending a collection of open-source scholarly editing tools within a scholar's workbench that employs a common data model and a systematic workflow.

SCHOLARLY EDITING WORKFLOW

The work required to produce a scholarly edition is usually conducted by a senior scholar who collaborates with colleagues, assistants, and various service providers (publishers, editors, libraries), located locally, nationally and internationally. In creating a new scholarly edition, scholarly editors must identify all of the relevant documents, describe those documents fully, digitize and transcribe those documents, compare alternative versions, annotate the documents and texts, edit a reading text, write historical and textual essays, create various appendixes and scholarly apparatuses, and prepare an edition for publication in an electronic and/or book form.

AUSTESE WORKBENCH

The AustESE workbench will coordinate this workflow and provide access to online tools that support scholarly editing tasks. The workflow will be further supported by assisting scholars to manage and archive digital resources for an edition within the content repository, including transcriptions, images, essays, annotations, ancillary materials and associated metadata. For the AustESE project, we have adopted a Service-Oriented Architecture, illustrated in Figure 1, with the aim of developing modular, reusable, and potentially distributed components that can be assembled and substituted according to the requirements of each scholarly edition project. To implement this architecture, we are extending existing scholarly editing tools with REST APIs to enable their integration with our content repository and workflow engine.

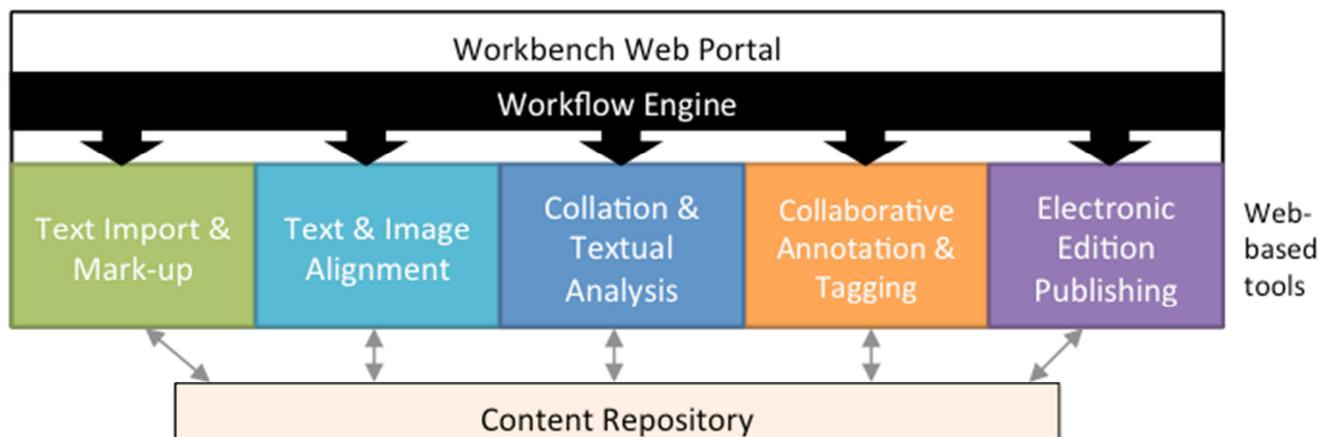


Figure 1: AustESE Technical Architecture

In Table 1 we have categorized scholarly editing tools according to their function, such as *Collation & Textual Analysis* tools, used for analyzing and comparing multiple versions of a work and to assist scholars to create an apparatus. For each category we surveyed existing Web-based tools and services and open source software projects to identify software that we can reuse and extend (in bold), and to determine gaps where we will need to develop new services to support an end-to-end workflow. Table 1 presents a summary of our survey results.

Table 1: Candidate tools for integration into AustESE workbench

	Text Import & Markup	Text & Image Alignment	Collation & Text Analysis	Annotation & Tagging	Publishing
Existing	HRITServer T-PEN, Tika, CTSDH Standoff Markup editor, Tesseract OCR, Islandora editor, Scripto, XTF	TILE	nmerge (MVD) and HRITServer, Juxta, Interedition CollateX, HRIT image collate, eHinman, Voyant, CATMA	LORE & lorestore, YUMA, UVic, HRIT CaTT	epub-tools, Apache POI, SADE, xMOD
To be developed	Versioning & correction services, Service to link documents with bibliographic entities, Advanced content search, Document relationships visualization service	Improved line / word recognition service, Support for standoff properties	Advanced variant query and visualization services, Light box	Annotation of textual variants, Advanced annotation query services	One-click publication, locking and cloning of electronic editions

In order to ensure that metadata and the contents of the editions created and published through the workbench are shareable, exchangeable and interoperable across tools and platforms, the data model for our content repository builds on existing standards. Bibliographic metadata about the versions of works under study will be described using the IFLA Functional Requirements for Bibliographic Records (FRBR), and annotations will be compliant with the RDF-based Open Annotation model [4]. The publishing service will support exporting editions or selections of their content to a range of common document formats. To enhance discovery and reuse, metadata describing published editions will be harvestable and collection descriptions will be made available through ANDS Australian Research Data Commons.

IMPACT

The eResearch infrastructure that we are developing through AustESE to support the needs of the Australian scholarly editing community will enable:

- Faster authoring of scholarly editions, particularly online electronic editions;
- Greater sharing of the resources, including primary and ancillary materials, annotations and bibliographic metadata that comprise scholarly editions; and
- Wider adoption of best practices and standards for publishing and archiving electronic editions and their constituent resources.

Although there are several international projects developing scholarly editing tools, this project is unique in its workflow-driven approach to integrating and extending existing tools and services. This project will benefit the broader international scholarly editing community by providing reusable models and software to contribute to the development of the emerging global “network of digital storage, access and dissemination” [5] for electronic editions.

ACKNOWLEDGEMENTS

The AustESE project is a collaboration between The University of Queensland, University of NSW, Curtin University, University of Sydney, Queensland University of Technology, Loyola University and the University of Saskatchewan. The project is funded by NeCTAR eResearch Tools program from May 2012 to Dec 2013.

REFERENCES

- [1] P. Shillingsburg, 2009. "How Literary Works Exist: Convenient Scholarly Editions." *Digital Humanities Quarterly*, 3(3). Available: <http://digitalhumanities.org/dhq/vol/3/3/000054/000054.html>
- [2] P. Robinson, 2010. "Editing Without Walls." *Literature Compass* 7(2), p.57-61. Available: <http://onlinelibrary.wiley.com/doi/10.1111/j.1741-4113.2009.00676.x/abstract>
- [3] "AustESE: Australian Electronic Scholarly Editing." [Online]. Available: <http://itee.uq.edu.au/~eresearch/projects/austese/>.
- [4] R. Sanderson, P. Ciccarese, & H. Van De Sompel, 2012. "Open Annotation Draft Data Model." Available: <http://www.openannotation.org/spec/core/>
- [5] J. McGann, 2004. "A Note on the Current State of Humanities Scholarship. *Critical Inquiry*." 30(2), p.409-413. Available: <http://dx.doi.org/10.1086/421142>.

ABOUT THE AUTHORS

Anna Gerber is the senior software engineer for the AustESE project, developing eResearch tools to support the collaborative authoring and management of electronic scholarly editions. She is also part of the Open Annotation Collaboration, developing annotation tools and services that can be used to evaluate and demonstrate the applicability of the OAC data model in the context of annotations supporting collaborative development of scholarly editions. From July 2008 – July 2011, Anna was the senior software engineer for the Aus-e-Lit project. She had previously worked on the HarvANA: Harvesting and Aggregating Networked Annotations and AnnoCryst for PyMOL projects. Prior to joining the eResearch Group at The University of Queensland School of Information Technology and Electrical Engineering, Anna was a Research Scientist at DSTC from 2000 – 2005, focusing on Enterprise Modeling.

Dr Roger Osborne has published widely in the fields of book history, print culture and textual criticism. He completed a PhD at the UNSW in 2000 and was a Postdoctoral Fellow in the Australian Studies Centre, University of Queensland, from 2004-2007. He was Project Manager of the Aus-e-Lit Project from 2008-2011. He is co-editor of the Cambridge Edition of Joseph Conrad's *Under Western Eyes* (forthcoming 2013) and as the 2011 Nancy Keesing Fellow at the State Library of New South Wales, he is working towards an electronic edition of Joseph Furphy's *Such is Life*. Roger is also conducting preliminary research for an edition of Joseph Conrad's modernist classic, *Nostramo*.

Jane Hunter is Professor of eResearch and the Director of the eResearch Lab at the University of Queensland where she manages a team of software engineers, Post-docs and PhD students developing innovative software solutions for a broad range of disciplines. She is Deputy Chair of the Academy of Sciences National Committee for Data in Science and has published over 100 peer-reviewed papers in the fields of e-Research, Information Systems, Semantic Web and Digital Libraries.