**Eggert — Harpur**

<no figures>

**The Harpur Critical Archive**

Eggert, P. R. and Schmidt, D. A.

The Harpur Critical Archive (HCA) is both a digital archive and a scholarly edition of the poetic works of Charles Harpur. It addresses complex problems of availability that are a direct result of Harpur’s life story.

Born in 1813 in colonial NSW only 25 years after European settlement, Harpur was Australia’s first important native-born poet. Although of ex-convict stock he acquired a literary education. His poetry would be deeply in tune with the contemporary taste for wild nature and the sublime. Although he would never become a radical original in his poetry, he was easily able to master 18th-century verse forms and was equally at home with the blank verse of the Romantics, with its flexible imitation of the tones of the speaking voice and the inner rhythms of thought.

His mission would be to give poetic expression to the shared colonial experience of a natural antipodean world. He began contributing poems to the Sydney newspapers in 1833. The young man soon proved adept not only at love poems but at satires and squibs usually directed at his natural enemies, the Exclusivists, who separated themselves from those like Harpur with the convict taint. An Australian patriot, he was soon preaching republican, as opposed to monarchist, politics in the vain attempt to secure independence from Britain for the still-young colony. He signed one of his manuscript volumes resonantly: ‘Charles Harpur / An Australian’. He added long prose notes to the newspaper appearances of his poems to prosecute his various agendas.

Unluckily for Harpur, the local book-publishing industry would not get onto a royalty-paying footing for literature until the 1890s, more than two decades after his death in 1868. Never wealthy, Harpur could not pay to be published in book form. Newspapers were his only option. From the 1840s, he kept cuttings of his poems as they appeared. In his later years, yearning for a London book publication that would never happen, he copied out, revised, and sequenced his newspaper and manuscript poems into manuscript books. At his death there were 24 such compilations, now in the Mitchell Library in Sydney, as well as nearly 800 newspaper appearances. There were also some pamphlets of his poems; a number were appended to his play *The Bushrangers* published in 1853, and a posthumous volume appeared in 1883, a substantial selection organised by his widow but edited by a literary gentleman who adapted and abridged the poems to meet the new tastes of the day.

In total, then, for the period to 1900 we have over 2,800 versions of some 700-odd works, some of which are very long. Capturing the documentary evidence, presenting edited reading texts, annotating them, and collating the variant versions has been recognised as a scholarly necessity since the late 1940s but has never been realised in book or any other form. The HCA aims to make good by seeking simple rather than complicated digital solutions.

The conceptualisation behind the technical design has been crucial. The project has two, linked expressions: archival and editorial. The archival expression aimed to gather and display the basic data: digital images of each manuscript page and newspaper or other printed appearance, and transcriptions of each one. This took about four years. The 5,000 manuscript pages were transcribed in a simplified subset of TEI-XML by master’s students at Jadavpur University in India, specially employed for the job. The transcriptions were painstakingly checked in Melbourne by a specialist research assistant, who also transcribed the newspaper poems. Along with copies of the tiff and jpeg images, the transcriptions are now archived by the Sydney University Library in its D-Space database. It can store files for preservation purposes but never delete or modify them. Done under a signed agreement, this arrangement ensures long-term storage by an institution well set up to provide it and with an interest in doing so.

This arrangement is a fail-safe; a Google search will find these files, which by themselves are not user-friendly; but metadata directs the user to our site, charles-harpur.org. That in turn flicks the user to the government-funded NeCTAR cloud-server that actually hosts the HCA. The project has a back end where the files are stored and managed, and where the editor and any registered collaborators work. There is also a public-facing front end, where readers will find a variety of ways of accessing and understanding Harpur’s oeuvre.

For the editorial expression of the HCA project, the facsimile images of the manuscripts and ephemeral printed objects are the central resource. Apart from the 5,000 manuscript pages professionally scanned in colour from the originals in the Mitchell Library, we have nearly 800 images from newspapers, using TROVE wherever possible, and images of another 500 pages from book, pamphlet, and broadside printed materials. Harpur’s correspondence is also represented. Readers are thus now in a position to understand the actual material basis of the editorial endeavour that has produced the reading texts that they encounter.

We envisage various categories of reader. Schoolteachers not knowing the titles or first lines in advance but wanting a few poems on rivers, or about the bush or Aborigines, will be able to find them via the subject index. It was semi-automatically generated from the subject terms given to Harpur’s poems indexed for the AustLit database. Local historians from, say, the NSW south coast or the Hunter Valley or the Hawkesbury—all of which Harpur celebrated—can similarly find what they are looking for. Social historians wanting to gather evidence of Harpur’s interventions in colonial debates, or literary critics wishing to follow the gradual unfolding of his poetic career, will have, via a timeline of biographical events and Harpur’s compositions, reliable datings and reading texts for the actual versions written at any particular time—rather than, as previously, being restricted to texts of works of indeterminate date. That traditional aesthetic perspective is now being productively complicated by the historical one at version rather than work level offered by the HCA.

The back-end system is based on AustESE, a NeCTAR-funded project for writing tools for scholarly editing, based at the University of Queensland during 2012–2013. Its ontology provides a model of the primary associations between people, material objects, and the concepts commonly used in scholarly editing. The primary classes of the ontology are Artefact, Version, Work, Agent, and Event. The AustESE WorkBench acts as a repository for transcription files and image files. Each such digital resource then serves as a target for analysis, specifically for annotation and textual commentary. As the user-creator adds metadata, relationships or associations among the digital surrogates are automatically asserted. A text-to-image linking tool allows the user to create links between individual words on the page of a manuscript or printed text and its transcription, mostly automatically. A minimal markup editor allows editing to be performed in an environment that presents together the page image and its transcription, together with a live preview of how the edited text will appear in the final reading text. These three components are scrolled synchronously, so that the editor never loses track of which part of the page-image corresponds to the section of text being edited. There is also an events editor that can deal with fuzzy dates and an editor for creating ordinary HTML documents intended for the scholarly editor’s use in creating annotations and introductory materials.

A Multi Version Documents (MVD) service allows the assembly and textual collation of versions of the same work, thus providing the basic data for the editor’s commentary and text-editing. The TEI-XML files are first separated into text and markup. If those versions contain deletions and additions, they are automatically separated out into layers of alteration, which may also be manually adjusted to show chronologically linked revisions. No traditional textual apparatus display is necessary since the archival recording and listing of editorial emendations that it provided are now taken care of by other means. Simplicity has been our touchstone, together with our belief that the archive does not replace the edition, especially not its provision of reading texts, textual commentary, and other annotation.

Although the users of the system are, in the first instance, the project creator and any assigned collaborators, AustESE allows later users to be granted permissions to add, for example, new annotations, transcriptions, or images. Thus the project need never be closed.

One obvious moment to publish is at the point when the archival capture appears to be complete, and the more interpretative editorial stage is about to begin. As the editorial phase progresses, further publication in tranches may be conveniently done. It is likely that Sydney University Press will be the publisher of the HCA, probably in EPUB3 format. These products will likely be sold commercially. Each one will be a subset of the project aimed at a particular audience and responding to a perceived need, even as the website of the whole HCA remains open and free. The press appears content with this arrangement of combining two of the dominant digital media of today in a single scholarly publishing enterprise.