Fluid Layering: Reimagining Digital Literary Archives Through Dynamic, User-generated Content

Jon Saklofske
Acadia University

INKE Research Group

Abstract
This article promotes a theoretical evolution in the conceptualisation and operation of digital literary archives via *NewRadial*, a prototype archive application that models the following distinction: Whereas a *digital edition* continues to function as a primary source, the root of a secondary discourse field much like its print-based predecessor, the *digital archive* should be reconceived as a broader, active, dynamic public record, an information commons that substantiates a foundational collection of primary texts with a continuous aggregation of critical contexts and conversations that grow from that foundation.

The INKE Research Group comprises over 35 researchers (and their research assistants and postdoctoral fellows) at more than 20 universities in Canada, England, the United States, and Ireland, and across 20 partners in the public and private sectors. INKE is a large-scale, long-term, interdisciplinary project to study the future of books and reading, supported by the Social Sciences and Humanities Research Council of Canada as well as contributions from participating universities and partners, and bringing together activities associated with book history and textual scholarship; user experience studies; interface design; and prototyping of digital reading environments.

Jon Saklofske is Associate Professor at Acadia University, Wolfville, Nova Scotia, Canada. His interest in the ways that William Blake's composite art employs words and images has inspired his NewRadial digital environment. He co-leads INKE's Modelling and Prototyping group, and experiments with digital gamespaces in university-level research and learning. Email: jon.saklofske@acadiau.ca.

Over the last 15 years, digital archiving efforts have resulted in a significant number of digital editions of printed literary history. These electronic artefacts are broadly distributable singularities: that is, single-source editions that transcend the individualized circulation history associated with each material copy of their printed counterparts. While one might mourn the loss of ability to mark up, annotate, and trace the marginal encrustation associated with individual print copies, digital editions enable an as yet untapped potential for the centralization of critical response, dialogue, and data related to the history of that edition's use. At a digital textuality seminar during the 2011 North American Society for the Study of Romanticism (NASSR) conference, Neil Fraistat and Andrew Stauffer (2011) offered the possibility that while many print and digital editions exist as isolated data silos, our changing textual/digital conditions are generating emerging properties for the digital scholarly edition. Specifically, reconceiving the environment surrounding an edition as a dynamic field of intelligently structured, addressable possibilities introduces ideas such as interoperability, modularity, multimodality, dynamism, and scalability into traditionally static ideas of archives and editions.

As a follow-up and response to this discussion, this article promotes a theoretical evolution in the conceptualization and operation of digital literary archives via *NewRadial*, a prototype archive application that models the following distinction: Whereas a digital edition continues to function as a primary source, the root of a secondary discourse field much like its print-based predecessor, the digital archive should be reconceived as a broader, active, dynamic public record, an information commons that substantiates a foundational collection of primary texts with a continuous aggregation of critical contexts and conversations that grow from that foundation.

This announcement arrived in my inbox from Oxford University Press:

> March 2012 sees the launch of a major new publishing initiative from Oxford University Press – Oxford Scholarly Editions Online (OSEO) – the first phase publishing online the complete text of more than 150 scholarly editions of material written between 1485 and 1660. Oxford Scholarly Editions Online will provide an interlinked collection of authoritative Oxford editions of major works from the humanities. This content constitutes the cornerstone of research in the fields of English Literature, as well as Philosophy, History, and Religion.
>
> Each title within the collection presents the full text of the work, as established by an authoritative editor, accompanied by the editor's record of important variations in that text, and interpretative and explanatory notes. Most also have introductions placing the work and the author in a historical context, and explaining the editorial principles and the history of the text. Online publication of these essential scholarly resources facilitates navigation within and between editions, whilst retaining the traditional elements familiar to users of the printed editions. The more flexible online presentation opens up

new possibilities for search and comparison. (Oxford University Press, Email, October 5, 2011, paras. 1–2)

While this might be new in relation to Oxford UP's approach to digital publication, such editions are nothing new in the realm of digital humanities work. Oxford UP's institutional subscription and purchase model for access to such a database is understandable, given their corporate, authoritative motivations, scholarly reputation, and existing investment in their printed editions, but is antithetical to the "open access" practices associated with many digital editions created by university scholars and the opportunities for federation and interoperability that such an open approach creates. As a result, this "new" Oxford idea of an electronic collection of scholarly editions is actually a step backwards: it offers an already outmoded (given its focus on simple navigation, search and comparison features that have been in place for nearly two decades) and relatively limited and limiting opportunity for scholarship due to its for-profit accessibility and exclusive use of an Oxford-defined canon of primary texts. This is clearly a twenty-first century twist on the emperor's new clothes, in which the emperor-publisher has also taken on the role of the swindler, and ignorantly parades in front of a modern crowd that can easily see through the aristocratic and anachronistic pomp and circumstance.

Contrast this with something like the *William Blake Archive* (http://www.blakearchive. org), an unparalleled and constantly growing digital collection of the various editions of William Blake's 200 year-old art, engraving, and writing. Blake's in-house production process resulted in individualized and unique copies of his illuminated books — in essence, each copy of his work (a combination of relief-etching, engraving, and hands-on customization processes) is a unique art object and its own edition, despite his employment of mechanical reproduction techniques. Prior to the existence of the *Blake Archive*'s digital initiative, it was extremely difficult and expensive for a researcher to see all existing versions and variations of Blake's work, given their dispersion among private and public collections. The *William Blake Archive* is an open access effort to centralize this work in a virtual electronic space, making the large variety of Blake's creative output available to all who choose to visit the website. This approach has also allowed the *Blake Archive* to become a part of NINES (http:// www.nines.org), which is an acronym for Networked Infrastructure for Nineteenth-Century Electronic Scholarship. NINES is an umbrella initiative dedicated to promoting a networked, interoperable federation of similar open access databases, which feature nineteenth-century literary material. The *Blake Archive*'s identity as a collection justifies the use of the term "archive," but its focus on published work and its exclusive collection of primary texts causes it to function more like a digital library — an institutionally established collection offered through a publically accessible environment for the purposes of broader circulation, research, and study. Alternatively, the *Blake Archive* can be more properly understood as an attempt at a comprehensive digital edition of Blake's work, albeit one with built-in, but still limited research tools.

While the misleadingly-named *Blake Archive* is nearly two decades ahead of the not-yet-released Oxford online edition, and light years beyond Oxford's exclusive, for-profit approach, it is still an emulation and a safe extension of print-based practices and institutions into digital space. Like a print edition, both the Blake Archive

and the Oxford initiative are closed systems: as final products of diligent editorial processes, they do not make room to host user-generated data, discourses or tools, or incorporate any such dynamic content as part of their overall collection. As well, for practical reasons, neither aspires to be any more inclusive beyond the limited definition of its primary source material, and a select few contextual references. While these kinds of choices are familiar (dare I say, habitual), digital editions and collections have a singular, significant advantage over their printed counterparts. Small alterations, corrections, and additions to both print and digital editions and collections are relatively easy to make, though it is arguably cheaper to do this with digital versions. More importantly, changes to the content of an online digital edition, even if they become substantial, do not require the "publication" of a new edition that must be re-purchased and that outdates previous versions. "Versioning" becomes less important when there is no material object or local installation to distribute or replace. Versioning history remains important to digital historians but scholars, researchers, and students are more interested in accessing a facsimile repository of primary texts hosted and delivered by these archival initiatives.

However, larger conceptual changes required after the publication of a print edition or the launch of a digital edition can be equally prohibitive, unless the architecture of the digital edition is designed with modularity, flexibility, expansion, and future modification in mind. If this kind of foresight occurs, then subsequent changes will require less of an overhaul or reinvention. Such thinking can and should be extended to the kinds of work that become possible around the content of a digital edition. Looking beyond the limitations of closed-access electronic conversions of pre-digital print material, we are at a point where the familiar and comfortable (if still tedious) practices of converting and encoding pre-digital literary history need to be extended by curative and editorial frames that move away from the tightly-controlled environments of their print predecessors. Instead of a static, closed repository that enables simple browsing, searching, and correlative operations, what are now only digital editions should be augmented by or become a part of larger, more dynamic and co-developed digital archives that continuously collect and centralize critical and cultural reception and responses within a virtual environment that includes that edition itself.

Digital archiving, reconceived of as a dynamic, continuous process of collecting and correlating secondary textual responses to primary digital editions, includes interoperable, layered databases of user-generated dialogue, response, and argument relating to its primary material. Centralizing the post-publication life of a digital text involves continually collecting an organic corpus of contextual, digital marginalia and critical work in a virtual space that includes the primary edition. While the edition data remains unchanged, then, the archive will operate like a kernel, managing, co-ordinating, integrating, and updating multiple user-generated, interoperable datasets. See Figure 1 below:

Figure 1: This diagram illustrates the various layers of dynamic content that can exist around a static edition. The middle layer of secondary scholarship constitutes an archival environment that builds up around the original edition.
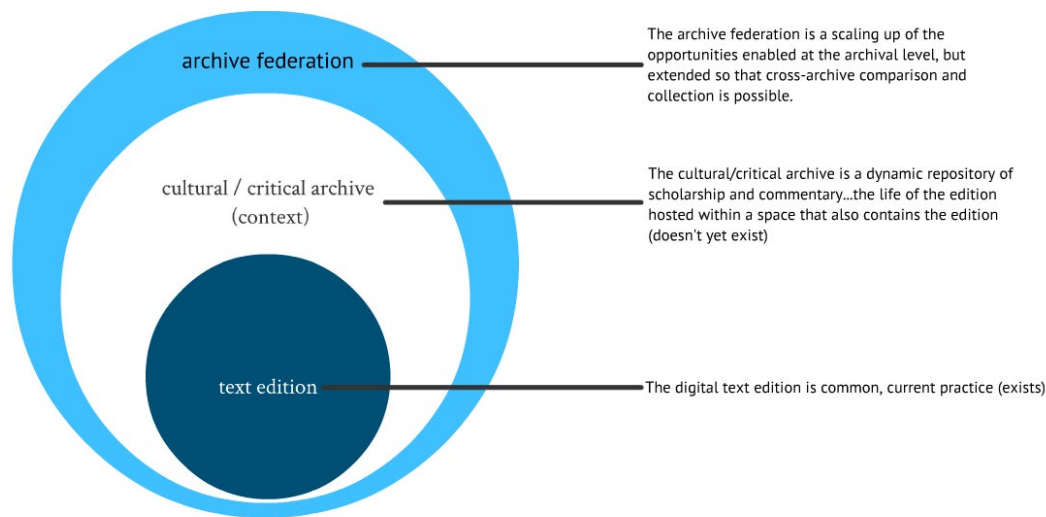
Naturally, the question of "Why?" needs to be addressed. More specifically, given that the translation of pre-digital texts into digital editions is usually the primary goal of these initiatives, why should that focus be extended to include a continual collection of secondary scholarship relating to that edition? To answer simply: the conditions of digital distribution generate an opportunity to centralize critical commentary about a specific edition within the same virtual space as that edition. Further, non-commercial editions create a critical space that remains open to a range of scholarly activity, and collecting this activity within a larger archive structure promotes open access models of scholarly exchange and communication. This opportunity does not generate new critical methods. Rather, it more efficiently and less restrictively does what we do already. We write, present, and publish critical ideas and arguments related to primary literary texts, but these efforts are distributed widely throughout books, journals, listservs, blogs, email exchanges, and conference talks. The archival aggregation of these critical activities layered on the digital edition of work that these criticisms focus on does not replace this broad range of scholarly dialogue, but encourages a centralization of such activity and makes it easier to follow and participate in the history of critical conversations that a work generates. As well, user-generated data contributed within the space of an archival collection that includes and surrounds a particular digital edition facilitates the generation and preservation of a rich dialogue field that can be mined and used comparatively via the larger archive federation sites (such as NINES). The scholarship resulting from such research is then fed back into the federation and edition databases to further reinforce the centralization of user-generated content within digital archive spaces.

Using the *Blake Archive* material as its foundation, the open source visualization application I have designed with the help of student programmers and using the *Prefuse* visualization toolkit, aims to re-present Blake's books as visual "playspaces" that encourage innovative critical approaches and reflections. The visualization

Saklofske, Jon. (2012). Fluid Layering: Reimagining Digital Literary Archives Through Dynamic, User-generated Content. *Scholarly and Research Communication, 3*(4): 040155, 11 pp.

5

application, called *NewRadial* (http://sourceforge.net/projects/newradial), thus serves as an alternative means of visualizing the *William Blake Archive*, one that furthers the unbound nature of the digital image by presenting the pages as iconic nodes in a relational field. While I have published two papers on ideas relating to the initial development of NewRadial, (Saklofske, 2010; Saklofske 2011), this article uses NewRadial to model the larger potential of digital humanities archives through user-generated data. Although the original archive is transformed through the filter of this application, it is not directly affected by such operations, and neither are Blake's original page designs. The work done in *NewRadial* is stored in a new database layer (xml) without altering any aspect of the Blake Archive's data. This is not a space in which users can creatively remix Blake's original designs. Although there might be a critical and scholarly reason to initiate such further experimentation, the two main aims of the current initiative are:

1.   To challenge book-based critical paradigms by re-presenting digitized copies of each page of Blake's (1794) *Songs of Innocence and of Experience* within a kinetic, mutating, experiential field.

2.   To promote ways in which user-generated content can be centralized and preserved in the space surrounding a digital edition as part of a larger archive of cultural commentary and critical scholarship.

Blake's pages appear through *NewRadial* as individual, but associable nodes. Users can browse this re-visualization of Blake's work (see Figure 2); they can reposition or group pages categorically or individually select and isolate certain nodes away from the main categories (See Figure 3); or they can map connections by drawing a line between two nodes or grouping a selection of nodes together and associating commentary with such connections (see Figures 4 and 5). In the current version, that commentary can become part of a locally installed version of the application, but, more significantly, in the final version of *NewRadial* still in development, it will be saved to a communal, web-based (HTML5) version that accumulates and maps collective critical engagement. Like a hybrid between data visualization tools and NINES' Collex tool (see Nowviskie, 2007), this visualization supports user-based connections between nodes, which are then mapped into the visual field of the pages of Blake's *Songs* for all users to see and to potentially respond to. As more connections are generated between individual nodes by users, the connecting line between them becomes more prominent (thicker and darker). At a glance, users can then see which associations are more frequent and common and which relationships have not been explored at all. Annotated groupings or constellations of multiple nodes are also supported. Users can explore (and will eventually be able to respond to) the critical commentary left by others or follow links provided by users to additional web-based material that exists outside of this application's database. A search function also allows users to quickly isolate groups, nodes, and edges, or foreground a specific user's contributions and constellations. Through this alternative means of accessing digital versions of Blake's work, the author/reader/visitor paradigm is replaced with a community of collaborative players, whose marginal constellations, connections, and annotations become part of a critical reef that builds up around the iconic nuclei of Blake's pages. In other words, user

contributions become a new, relational database that is entirely dependent on the original archive but not limited by the original interface.

Figure 2: In NewRadial, pages are re-presented as individual, associable nodes.
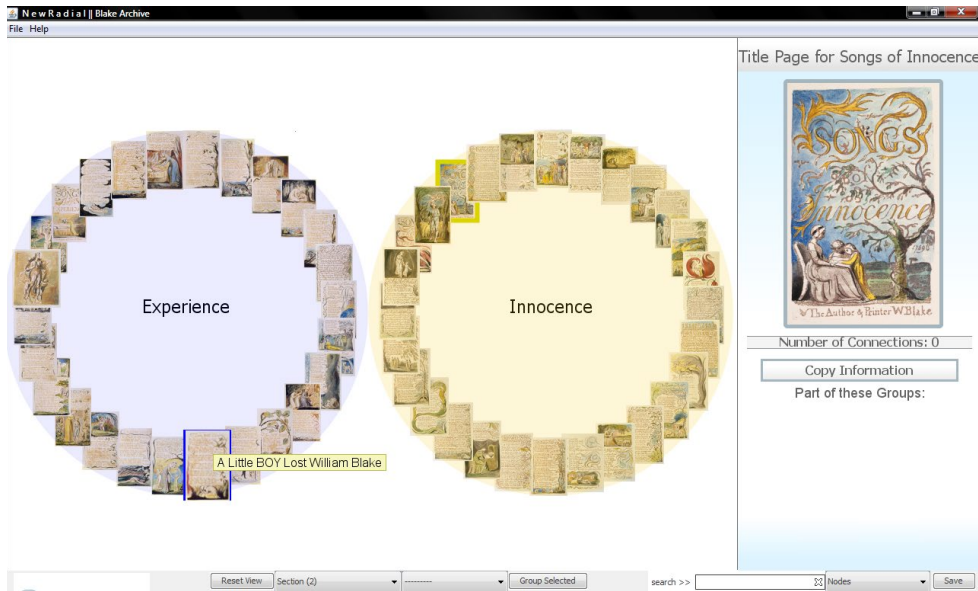

Figure 3: Nodes can be quickly reorganized by categorical groupings defined by the set's metadata, and individual nodes can be collected and isolated from these main groups by the user.
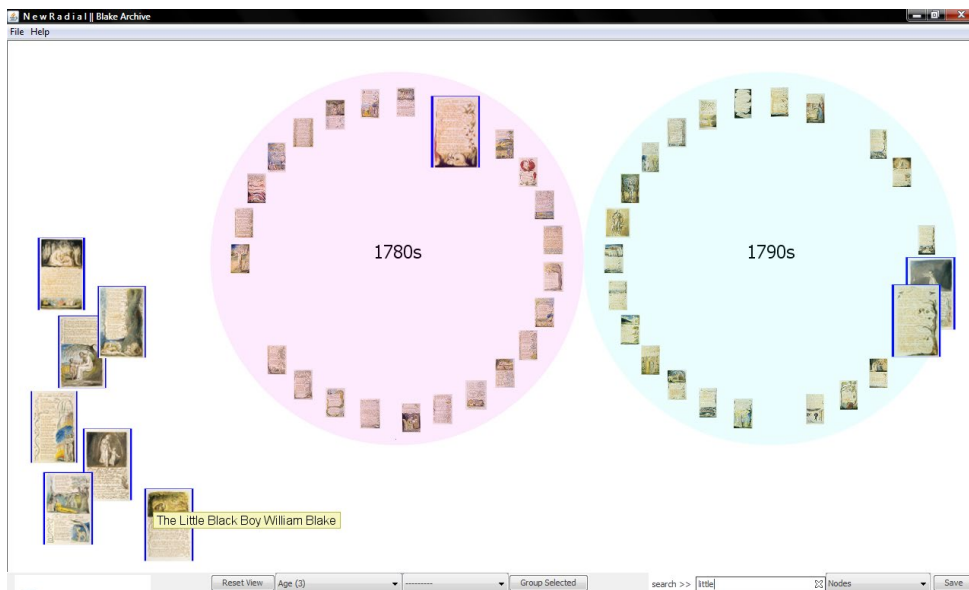

Saklofske, Jon. (2012). Fluid Layering: Reimagining Digital Literary Archives Through Dynamic, User-generated Content. *Scholarly and Research Communication, 3*(4): 040155, 11 pp.

7

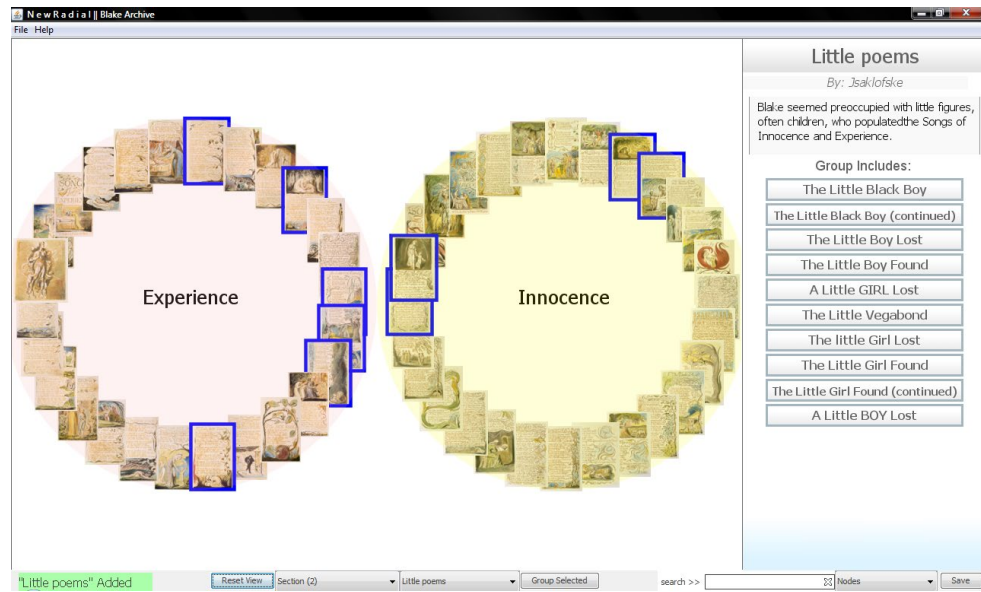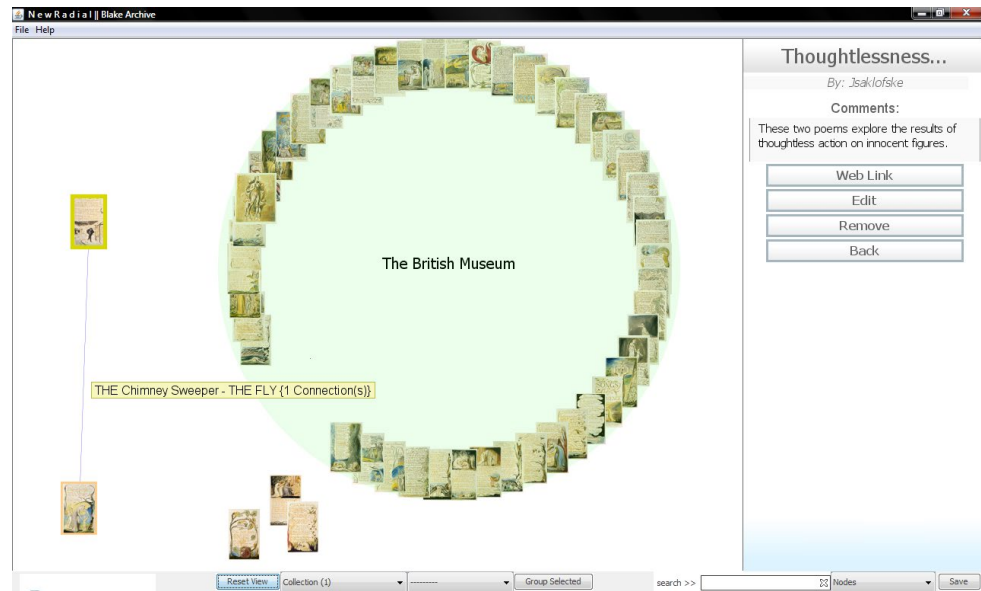**Figure 4: Users can also define unique groups and add their own annotations to such collections.**

Figure 5: Links (or edges) can be created between two nodes and commented on by individual users.

Since interfaces are perceptual frames, the ways that users can access and contribute to a database hinges on the tools and opportunities embedded in the interface design. A standardized approach to creating databases for digital editions would create flexible opportunities for the incorporation of various, multiple interfaces at the archival and federation levels. However, to take full advantage of the proposed interoperability between editions, archives, and federations, interface options need to be harmonized and standardized at all levels. Much work needs to be done to this end, but the first, important step is to reconsider digital archives, or dynamic environments that surround digital editions, as active sites of continuing scholarship.

Overall, then, this article argues that user-generated content needs to be given a more prominent place in the active space of a digital archive, and that such content extends but also digitally collects and centralizes the second life of a text as it is reconsidered through critical and cultural responses. The digital archive should be re-imagined as a centre for critical engagement, user-generated dialogue, argument, commentary, and response. Such allowances carry the risks and rewards of "messiness," of the energy of process rather than final product. As a result, it becomes a repository of raw material, of the cultural response to and reception of an edition and the primary texts that it contains, but also establishes the site of the digital text as a critical workspace and as a valid, respectable venue for scholarly contribution. In addition, this opening of digital archives to user-generated content encourages the same kinds of activity and annotation that would be performed on each material copy of a printed edition. This extended "life" of a printed copy is thus preserved but is not distributed among multiple, physical, remote copies. Rather, reading response and marginalia are still performed in the space that the primary text occupies, but becomes centralized in the same, singular virtual space or environment, as the edition itself. This creates a closer relationship between text and context and establishes the digital archive as a repository of the text and its living context. In these ways, the migration of critical scholarship from print to digital paradigms is shown not to be an either/or binary, but a more flexible space of preserved and extended best practices.

Think of a person's *Facebook* account, a dynamic edition of that person as represented through digital media. If that person dies in real life, status updates and profile changes are no longer contributed, but others can still post on his or her wall, tag photographs, and continue to add to the digital edition of that person in posthumous ways that extend the "life" of that person, that further the dynamic activity related to that person's digital edition. Similarly, *NewRadial* demonstrates that digital editions such as the *William Blake Archive* should be designed as starting points for active, participatory archiving rather than as inflexible destinations.

Such initiatives can be encouraged through archive federation sites, or could be developed in tandem with such federations, but can remain independent from, yet still somewhat reliant on, specific digital editions if need be. However, during the planning and development of digital editions, it is easy to imagine the broader archival motivations incorporated into the original database and interface structure. At the very least, perceptive creators of digital editions work on such editions knowing whether

Saklofske, Jon. (2012). Fluid Layering: Reimagining Digital Literary Archives Through Dynamic, User-generated Content. *Scholarly and Research Communication, 3*(4): 040155, 11 pp.

9

or not they want to tailor their data to function within federated spaces like NINES. Expanding this awareness to more consistently facilitate a broader inclusion of user-generated data within the idea of an edition-based archive does not require much more foresight.

While this initial call for a reconsideration of the content and function of a digital archive focuses on redefinitions and reconceptualizations, there are additional questions and issues that deserve further discussion. I will offer some of those challenges as an open-ended conclusion to this article in hopes of fruitfully continuing this discussion and catalyzing the further development of some of the proposals offered thus far:

1. What challenges and compromises are necessary to work towards the standardization of interfaces or database structures (or markup) to allow for easy exchanges between archive and federation datasets?

2. Are federation initiatives such as NINES (which depend on peer review to establish the scholarly legitimacy of their consortia and their members) willing to support the addition of archival environments that would support a dynamic and ever-expanding record of user-generated data?

3. What needs to be done to encourage a large-scale adoption of this initiative at both the developmental and user levels? How can existing editions be retrofitted to work within dynamic archival frames, and what kinds of guidelines will encourage in-development and future digital projects to incorporate modular flexibility and interoperability so that they will be compatible with a dynamic archival paradigm that values the collection and correlation of user-generated data?

4. Will different kinds of content be incorporated into user-generated datasets, or will these archives remain text-based?

5. Who will be responsible for editing and maintaining archival layers of user-generated data? Is such editing and maintenance necessary? Might a frequency of use or activity-based model be used to dictate ordering/editing principles of this secondary content?

The early planning efforts of the Modelling and Prototyping group within the INKE (Implementing New Knowledge Environments) initiative outline work over the next year that will attempt to address many of these questions and to develop prototypes, which will model effective ways to generate and sustain an active environment around digital editions. Overall, the idea is to preserve the validity of academically produced scholarly print editions in digital form, but to use those digital editions as a site around which scholarly activity, commentary, and dialogue can be focused and collected, then validated via community-based vetting procedures. This user-generated content, in combination with the existing primary source material, will provide a foundation upon which future digital editions can emerge. More specifically, INKE's Modelling and Prototyping team is exploring the possibility that, given appropriate toolsets and interfaces, these active digital archive environments can be used as the raw material

for user-generated editions (similar to the Collex collections that one can generate from the federated databases under the NINES umbrella) that will emerge from, but also live within the space of their source material. In other words, the archive space becomes a repository for and a network of a variety of editions and critical dialogue relating to specific pre-digital, primary material, centralizing the reception history of particular texts. Ideally, networking a number of these dynamic archives in a larger, federated workspace would encourage comparative and correlative work, expanding the context within which scholarship surrounding particular source material takes place. Accessing, collaboratively contributing to, and responding to scholarly work in progress allows for conference-type energies to exist within virtual scholarship spaces but also preserves the opportunity to sustain existing peer evaluation and validation practices.

Websites

NewRadial. http://sourceforge.net/projects/newradial
NINES. http://www.nines.org
The William Blake Archive. http://www.blakearchive.org

References

Blake, William. (1794) *The Songs of Innocence and of Experience. The William Blake Archive*. Ed. Morris Eaves, Robert N. Essick, and Joseph Viscomi. URL: http://www.blakearchive.org/exist/blake/archive/work.xq?workid=songsie [July 16, 2012].

Fraistat, Neil & Stauffer, Andrew. (2011). Digital textuality seminar. North American Society for the Study of Romanticism (NASSR) Annual Conference. 12 August 2011.

Nowviskie, Bethany. (2007). A scholar's guide to research, collaboration, and publication in NINES. *Romanticism and Victorianism on the Net*, *47*. URL: http://www.erudit.org/revue/ravon/2007/v/n47/016707ar.html [October 31, 2011].

Saklofske, Jon. (2010). NewRadial: Revisualizing the Blake Archive. *Poetess Archive Journal*, *2*(1). URL: http://unixgen.muohio.edu/~poetess/PAJournal/2_1/Saklofske.html [October 31, 2011].

Saklofske, Jon. (2011). Remediating William Blake: Unbinding the narrative architectures of Blake's songs. *European Romantic Review*, *22*(3), 381–388.

Saklofske, Jon. (2012). Fluid Layering: Reimagining Digital Literary Archives Through Dynamic, User-generated Content. *Scholarly and Research Communication, 3*(4): 040155, 11 pp.

11