"Expansión, elasticidad y reelaboración de un archivo como base de datos: Entrevista a Kenneth Price del Archivo Walt Whitman" by Mariana Garzón Rogé. *RedHistoria* 4 (2013).

**Available at **

REDHISTORIA. Recursos y experiencias online para historiadores

Kenneth Price from The Walt Whitman Archive (whitmanarchive.org)

1. Which was your goal when you began the project of Walt Whitman Archive and how did it changed through the years?

We wanted to edit Walt Whitman's writings online, to make them freely available to a world audience. We wished to make a comprehensive collection of his writings in all its various forms—manuscripts, notebooks, corrected page proofs, and printings in periodicals and books. We couldn't reproduce the material forms of his writing, but we could provide accurate transcriptions along with facsimile images. We also realized that we could include additional related content that would aid research on Whitman—for example, a collection of all known contemporary reviews of his writings.

Our sense of the expansiveness and elasticity of the digital archive has only increased since we began work in 1995. The project has developed in ways we didn't foresee (and in some cases couldn't have foreseen). For example, I recently identified approximately 3,000 documents written by Whitman when he was a clerk working in the U.S. Attorney General's office. We have also become much more interested in Whitman's world-wide reception and have done far more with translations of Whitman's work than was imagined at the start of the project. In the coming years, I expect that we will do more to open up the project to contributions from users. This type of functionality was not anticipated in 1995.

We have also been conscientious about documenting the project itself so that people can understand our technical decisions, the nature of our collaborations, and the history of labor on the project. We have become increasingly self-conscious about digital publishing and its opportunities and perils, and we are preparing as best we can for transitions within the leadership of the *Archive* and for long-term preservation of our data.

Over time we have become increasingly committed to open access. Initially, we thought we could only achieve our goals through a partnership with a major publisher. We hoped such a partnership could provide a revenue stream to support the development of the *Whitman Archive*. In the late 1990s it seemed too optimistic to believe we could get steady federal grant support for our project, but, somewhat surprisingly, we have been able to do so. In addition we have established a $2 million permanent endowment to support our ongoing work, with most of the endowment support coming from private

donors. A combination of hard work and good fortune has put us in a favorable financial position relative to most other projects in the humanities. We make all of our content freely available, distributing it under a Creative Commons license. Fortunately, the emergence of our project happened to coincide with the open access movement. We recognize the importance of free distribution and have found ways to make that a successful publishing model.

From the beginning, graduate students have had a central role in the project and have contributed groundbreaking ideas. It has been rewarding to employ graduate students not as low-level workers but as partners in the project who are crucial to its development. As a result, our sense of graduate education has changed over the course of the project: we are no longer training only literary scholars but instead individuals capable of contributing to a variety of fields.

2. Is there something in Walt Whitman's work that makes it special to be explored as a database? Could we construct databases from other pieces of the past such as a historical period, for example?

The fluidity and enormity of the textual record left by Whitman make it better suited to electronic rather than print-based treatment. Whitman revised his poems frequently: in addition to the many versions of *Leaves of Grass*, he also left manuscript drafts, notebook versions, and corrected page proofs. The economics of print publishing make it impossible to represent all versions of his works. In contast, with digital publishing we can treat his writings more completely and in ways that are more convenient and more economical for users. Unique items no longer need to be locked away behind the doors of rare book rooms and special collections libraries. Whenever possible, we make high quality digital images of these items available to users everywhere.

At the *Walt Whitman Archive* we make our material available through numerous databases including a bibliography dedicated to commentary on Whitman and another database containing images of Whitman and associated metadata and commentary. We also have many XML files including implementations of TEI, EAD, and XHTML. Some of these are open to the public and others are internal databases—for example, our tracking database which allows us to manage record keeping at the level of individual documents. Our goal is comprehensiveness, though we can also highlight key items, if we choose. Importantly, the database allows for different views of the same material. The content can be reconfigured and sorted in various ways: for example, alphabetically by title or in chronological order.

No database could adequately represent a historical period. Along with other colleagues at the University of Nebraska-Lincoln, I work on another project, *Civil War Washington* (civilwardc.org), attempting to illuminate certain key aspects of the city's transformation during four years of war. We have had some success, but we also recognize how partial and selective are our foci. We have concentrated on the development of forts, the growth of hospitals, and the changing nature of the population as recorded in Washington, DC,

census records. We have also concentrated on collecting and transcribing documents produced in response to the Compensated Emancipation Act that led to the freedom of 3,300 slaves living in Washington, DC. Thus far our project has devoted much time to the study of race, slavery, and emancipation. We have not studied other potentially illuminating aspects of the city's history—crime statistics, foreign delegations, weather reports, business records, ward-level politics, environmental degradation, and so on. If we gain further support, we may be able to study some of these matters in the future. Ultimately, what could be studied in connection with *Civil War Washington* is nearly limitless.

3. In a database, which is the relationship between detail and categories? Which knowledge effects' can produce the fact to see as a database what before we could only see fragmented in several archives? Which is your experience in the Walt Whitman Archive?

A database is not an undifferentiated sea of information out of which structure emerges. Argument is always there from the beginning in how those constructing a database choose to categorize information — the initial understanding of the materials governs how more fine-grained views will appear because of the way the objects of attention are shaped by divisions and subdivisions within the database. The process of database creation is not neutral, nor should it be. The design of the database's user interface also imposes restrictions on the way that the information can be categorized, searched, and accessed. Users of databases often don't consider these points.

Databases can be misleading, if people don't understand what is presented, and what biases may inhere in the selection of content. Readers need to be critical and to ask questions about the results they see and fail to see.

4. To explore is different from simply searching for results… that is something that you say when you refer to what users can make with the Walt Whitman Archive. Is there a "model of reader" for databases?

As I've mentioned, readers or users of a database need to question the assumptions on which it was built. A database offers a particular vision based on the ideas of the people who created it. Their organizing principles determine what is selected and what is excluded. We shouldn't think of data structures or interfaces as natural or inevitable. Instead, they are constructed and are expressions of the interests of their creators.

An ideal user of the *Walt Whitman Archive* would use some of our XML downloads and perhaps make use of our data in ways we had never anticipated but are nonetheless highly illuminating. Such a user might want to do something different than what we have done with the TEI-encoded texts and the visual output using XSLT or some other method of delivery. A different user might pluck out some of our content and aggregate it with non-Whitman material to explore or illustrate certain ideas. We hope to enable ingenious reworking of our material by releasing it through a Creative Commons license.

5. Your coeditor Ed Folsom has said that it is possible to build a database toward inclusiveness rather than exclusiveness. This could improve the users' chances of questioning and challenging narratives in which the data has been read before. But, how can this impact in our capacities to built alternative historical narratives?

It's certainly possible for databases to build toward inclusiveness, but databases can be misunderstood to be more inclusive than they are. Users of databases of historic newspapers, for example, should not assume that they represent a complete record of past journalism. Records are almost invariably partial, so it is important to understand the nature and extent of any collection represented in a database so that any claims made are responsible and in keeping with the evidence.

Readers of databases have to be astute. They need to be critical rather than merely accepting the content they receive. There has always been filtering and interpretation in the setting up of categories, no matter how inclusive a project tries to be. Even the most inclusive of artists (Whitman would be one) and the most comprehensive of historians are selective in important ways.

It may be possible to build alternative historical narratives merely by detecting patterns in the data that others have overlooked. Another way to build alternate historical narratives is to become so aware of the limits of what is included in a dataset that one creates a new one better suited to reveal alternative histories. Of course, sometimes the data itself is very hard to locate, however one might want to find it. Information about the emotional lives of working-class nineteenth-century women, for example, is hard to document, just as it is often hard to find much detail about many nineteenth-century African Americans.

6. Do you imagine an end for the database you created? There would be a day in which the Walt Whitman Archive will be completed?

Some parts of the *Whitman Archive* could, logically speaking, reach a state of conclusion. All known manuscripts in Whitman's hand could be located and transcribed, for example. But other parts of the site do not have a logical end point. Our bibliography of criticism is updated with each passing year, and it will need continuous updating for as long as people are interested in the life and writings of Whitman.

With the *Walt Whitman Archive*, we are creating a research environment, not just an edition. The edition is central to our ambition, but we are also creating a space that promotes and publishes scholarship that first appears there. Our translation work, like the bibliography and like general scholarship on Whitman, could be ongoing indefinitely. In important ways, the *Whitman Archive*, by its nature, defies attempts at completion.

When so much could be done or is needed to be done, we face continually the question of what work should be done first. We prioritize according to current scholarly needs. Thus, we have worked on scribal documents before, say, Whitman's temperance novel *Franklin Evans*, because the scribal documents would be completely inaccessible without our work, and good print editions of *Franklin Evans* are available.