Notes

1. See, for example, Swanson and Smalheiser, "Interactive System" and "Assessing."

2. The exception is the null value, which has its own problems, as discussed above.

3. Discussing narrative, Bruner also emphasizes the importance of causality, identifying crucial components as agency, sequential order, sensitivity to the canonical (or context), and narrative perspective (77).

Works Cited

Bal, Mieke. *Narratology: Introduction to the Theory of Narrative*. 2nd ed. Toronto: U of Toronto P, 1998.

Bruner, Jerome. *Acts of Meaning: Four Lectures on Mind and Culture*. Cambridge: Harvard UP, 2002.

Connolly, Thomas, and Carolyn Begg. *Database Systems*. New York: Harlow; Essex: Pearson Educ., 2002.

Genette, Gérard. *Narrative Discourse: An Essay in Method*. Ithaca: Cornell UP, 1983.

Kroenke, David M., and David J. Auer. *Database Concepts*. 3rd ed. New York: Prentice, 2007.

Liu, Alan. "Transcendental Data: Toward a Cultural History and Aesthetics of the New Encoded Discourse." *Critical Inquiry* 31 (2004): 49–84.

Manovich, Lev. *The Language of New Media*. Cambridge: MIT P, 2001.

Mathews, Harry. *The Journalist: A Novel*. Boston: Godine, 1994.

Richardson, Brian. *Unlikely Stories: Causality and the Nature of Modern Narrative*. Newark: U of Delaware P, 1997.

Ricoeur, Paul. *Time and Narrative*. Vol. 1. Chicago: U of Chicago P, 1990.

Swanson, Don R., and N. R. Smalheiser. "Assessing a Gap in the Biomedical Literature: Magnesium Deficiency and Neurologic Disease." *Neuroscience Research Communications* 14 (1994): 1–9.

———. "An Interactive System for Finding Complementary Literatures: A Stimulus to Scientific Discovery." *Artificial Intelligence* 91 (1997): 183–203.

Tukey, John W. *Exploratory Data Analysis*. Reading: Addison, 1977.

Turner, Mark. *The Literary Mind: The Origins of Thought and Language*. New York: Oxford UP, 1998.

Reply

ED FOLSOM

Ah, the power of metaphors indeed! To describe the relation between narrative and database, N. Katherine Hayles offers an astute alternative to Lev Manovich's "natural enemies" metaphor: she suggests "natural symbionts," a metaphor I plan to appropriate and use from now on. Her claim that "database catalyzes and indeed demands narrative's reappearance as soon as meaning and interpretation are required" incisively articulates what she calls the "dance" of narrative and database. I've thought of the relation as an endless battle (once narrative begins to win, database rallies, and vice versa), but Hayles's metaphor more efficaciously captures what she rightly characterizes as "the complex ecology" of these two modes of organizing and accessing the represented world.

And, as Hayles makes clear, the metaphors are essential. The term *database* itself is a metaphor, a base onto which we put things that are given (data). The word is less than fifty years old and has mutated in meaning over the decades. Few of us (certainly not I) can approach a database without an array of metaphoric terms that make it seem something it is not. Years ago, when I used to hit a key on my old typewriter, I could follow and even explain the mechanical process that struck an inked ribbon with a typebar to impress a letter on a page. Now, when I hit a key on my computer keyboard, my knowledge of the process that makes a letter appear on my screen is hazy, to say the least, not to mention the process that transfers it to paper. How this sentence I'm writing gets preserved on my USB stick and in what form is a mystery to me. Without the metaphoric apparatus that

https://doi.org/10.1632/S0030812900168592 Published online by Cambridge University Press

allows us to save, open, cut, paste, and create files that can be read by other computers, this world of data entry and retrieval would be inaccessible to most of us. It's no accident that the term *user-friendly* followed *database* by a decade or so and that we all now depend on user interfaces, where many of our most useful metaphors reside.

So when Jerome McGann complains that my referring to *The Walt Whitman Archive* as a database is "seriously misleading—more accurately, it is metaphoric," I accept his second (more accurate) characterization. But when he says the archive "is not—in any sense that a person meaning to be precise would use—a database at all," I have to disagree. Of course it's a database. It is, in fact, several databases—the thousands of bibliographic entries are stored in one, the photographic images in another, and so on. A database, as defined in *The Oxford English Dictionary*, is "a structured collection of data held in computer storage; *esp.* one that incorporates software to make it accessible in a variety of ways." McGann's insistence that "[n]o database can function without a user interface" that "embeds . . . many kinds of hierarchical and narrativized organizations" is certainly true, because, for most of us, that's what a database is: a vast vault of unseen data that are retrieved and organized by our metaphoric commands, which, as Hayles explains, prompt a database-management system to employ "set-theoretic notation to query the database and manipulate the response through SQL and related languages. . . ." My interest in database as an emerging genre, however, has more to do with the wild and unpredictable intersections of the data that the interface allows us to generate, what Wai Chee Dimock in her introduction to this issue calls "[t]he links and pathways that open up [and] suggest that knowledge is generative rather than singular, with many outlets, ripples, and cascades, randomized by cross-references rather than locked into any one-to-one correspondence."

Discussing the standard markup approaches used for encoding textual data, McGann admits that the "TEI and XML do not adequately address the problem of knowledge representation that is the core issue here—that is, how do we design and build digital simulations that meet our needs for studying works like Whitman's?" and, again, I agree. All our careful tagging and markup (further suggestive metaphors) of the texts on the Whitman archive reveal more and more features that our tagging codes cannot adequately describe. That's the wild excess, and it's one reason we have insisted on including in the archive high-quality scans of the material that we enter into the database as tagged text, so that users can test and challenge our embedded hierarchies and interpretive decisions. On every page of manuscript that we transcribe, there are features that we either name as an instance of some category or ignore. For some user sometime, what we ignored will turn out to be important; what we tagged as one thing will seem to be something else. The images linked to the tagged text (it's all data; it's all on the base at once) serve as checks. Already, as I mention in my essay, some users of the archive have been able to piece together manuscripts that had been physically separated and scattered among different archives; they have done so by examining the untagged details (glue marks, needle holes, small tears) on scans of the pages. There's a great deal in this database, in other words, that escapes the editorial markup and yet is still retrievable and valuable for users who wish to explore instead of simply searching for results.

What is true for the myriad bewildering markings on one of Whitman's manuscript pages is also the case in his printed texts. Take the first edition of *Leaves of Grass*: virtually all students of Whitman know (because they've been told so many times) that the twelve poems in that edition are untitled. But when we prepared to tag the text of the first edition, we were confronted with the

https://doi.org/10.1632/S0030812900168592 Published online by Cambridge University Press

jarring typographic fact that, while the final six poems have no titles, the first six *do*. Each of the first six poems is entitled "Leaves of Grass." Now, *Leaves of Grass* is the book's title, so most readers, editors, and critics have apparently assumed this repeated title must be some kind of running head, even though it clearly occupies the position of a title. The New York University Press's three-volume variorum edition ignores these titles, as do most reprintings of the book, like the Library of America edition. But in tagging this material to enter it into a database, we needed to describe this stubborn printed phrase. Since in later editions of *Leaves of Grass* Whitman would again use repeated titles, including "Leaves of Grass," it seemed reasonable to conclude that he had started this practice with his first edition. And since in the 1860 edition Whitman includes a cluster of twenty-four numbered poems called "Leaves of Grass," is it also reasonable to conclude that the final seven short poems in the first edition are actually his first cluster, all contained under the sixth "Leaves of Grass" title? Or, in his desire to fit everything into twelve eight-page signatures, did he begin to drop this title to save space? We editors have to make a hierarchical decision in cases like this, but the scanned pages of each edition stand in the database as visual checks on every tagging decision we make. Our decision in this case will affect title searches, but no matter what we call a particular feature, the image scans of each page will continue to portray the feature in its raw, untagged, wild state.

When McGann says, then, that "databases and all digital instruments require the most severe kinds of categorical forms" and that the "power of database—of digital instruments in general—rests in its ability to draw sharp, disambiguated distinctions," he's right (tagging requires it), but for me the real power of database rests in its equal ability to generate the materials that allow users to question each sharply drawn distinction. Jonathan Freedman, like McGann, worries that "to celebrate database as a kind of autonomous form" is "to downplay the inclusions, exclusions, choices that have gone into the making of databases and hence to occlude the possibilities for questioning those choices." But this points, once more, to the endless battle between—the symbiosis of—narrative and database. It is possible to try to build a database toward inclusiveness rather than exclusiveness, and the more we do so, the better the users' chances of questioning and challenging whatever narrative the creators have attempted to tag onto the data.

I've learned a great deal of what I know about textuality from Jerome McGann (that's truly Folsom praise), and I take to heart his cautions about how database is but one step in an endless process of mediation and remediation. I am optimistic about the possibilities of electronic editions, but, as a frequent dweller in physical archives, I am also viscerally aware of what does not get translated into the virtual archive. I've held that little notebook where Whitman first teases out the voice (and the attitude) that would generate *Leaves of Grass*, where you can see something like the DNA of his future work: there's an endless amount of information in the feel of the pages, the stubs of the cut-out leaves, in the way the book rests in the palm of the hand, not to mention in the story of how it sat in an attic for half a century after it was stolen from the Library of Congress. By examining the binding and signature construction of the first edition of *Leaves of Grass* in multiple physical archives, I've learned many things about its making that I never could have discovered on the virtual archive. But I love the challenge of trying to figure out how we can now remediate as much of that information as possible onto the Whitman archive, to try to grow the database so that the surprises of searching and juxtaposing will become richer and more frequent.

Freedman teams me up with the "Googlizers": if *The Walt Whitman Archive* had only

https://doi.org/10.1632/S0030812900168592 Published online by Cambridge University Press

a fraction of one percent of Google's resources, we could grow our holdings quickly and make the archive more like the vast and inclusive database that I fantasize about in my essay and that Meredith McGill would understandably like to see more of now. McGill finds the archive "not a transformation but a 'remediation' of archives." Here we come back again to the metaphor of the symbionts: database cannot remediate archives without in some key sense transforming them (as McGann's comments on markup make clear), but there is no doubt that a vital part of the *The Walt Whitman Archive* is the collection of scans of books, manuscripts, and photographs, which, taken by themselves, *are* a remediation (and a combining) of archives. I'm not sure, though, why McGill believes that "[d]igitizing archives makes it harder to see the partial nature of the printed record, the limited reach of print at any moment in history, and the supersession of one edition by another." We will soon be including in the archive the results of the first complete census of extant copies of the first edition of *Leaves of Grass*, including their known original owners and the variations from copy to copy. Even now, users can for the first time put side by side on their screen the same poem as it appears in each edition of *Leaves of Grass*, creating a visual image of "supersession" of editions unlike anything possible before, short of opening actual original copies of all the editions.

McGill makes the valid point that, in its current stage of development, the archive reproduces "mass culture's reductive treatment of genre" by offering all the poetry and little of the prose. But, as she accurately notes at the end of her response, "[t]hese are still early days for the digital humanities." Yes indeed. Kenneth Price and I initially thought we'd be done with this project in five or six years; now, more than a decade later, we realize that if we can keep it supported it will continue to grow long after we're gone, because database does not handle completion well—it is voracious and thrives on revision, addition, and supplementation. McGill's exciting suggestion of how "'rhizomorphous' connections . . . might have been encouraged by providing hyperlinks to Whitman's editorials in the *Brooklyn Daily Eagle*" sounds like the continual discussion among archive staff members about how we need to include a history of translations of Whitman's work from around the world, scans of the issues of the periodicals in which he published, all the biographies of him, the letters he wrote and all known letters to him. . . . The list is endless.

And database can handle it all. What are needed are time, energy, resources, talented scholars, and the inevitable improvement in software and hardware that has made so much digital scholarship thinkable today that was unthinkable ten or even five years ago and that will make the unthinkable today doable a decade from now. Freedman notes, for example, that "[t]he creators and maintainers of *The Walt Whitman Archive* don't include much contemporary criticism (largely, one assumes, because of copyright rather than predilection) but link extensively to Whitman-era responses; the result is to institutionalize certain versions of Whitman while effacing others." That was true when he wrote his response, but it is less true now, because the University of Iowa Press generously agreed to let us put online the entire Iowa Whitman Series (currently fifteen books of criticism from 1989 to the present, three of which are already available), and we are working with authors and presses to arrange for more copyrighted material to appear. If my rhetoric is, as Freedman suggests, "utopian," my experience in working on the archive is anything but utopian. It's slow and frustrating work, but database invites big imagining, and, as more and more humanities scholarship becomes digitally based, the possibilities will grow exponentially.

Database is a genre that the next generation of humanists will take for granted. Universities that haven't yet adjusted their

https://doi.org/10.1632/S0030812900168592 Published online by Cambridge University Press

scholarship and research expectations to allow for and encourage digital scholarship will soon do so. Digital research requires collaborative enterprise of the sort that has been rare in humanities scholarship. As with any emerging genre, it's anybody's guess where it will go and what range of effects it will have. As Peter Stallybrass notes, however, already "millions of people who cannot or do not want to go to the archives are accessing them in digital form. And digital information has profoundly undermined an academic elite's control over the circulation of knowledge." Just as my work with an electronic archive has helped me discern in Whitman's work aspects of what I think of as database, so has Stallybrass found "Shakespeare consciously practic[ing] his own form of database." He goes on to point out how "some of the most powerful modern databases draw on the development of a massive range of finding aids and databases in the Middle Ages and Renaissance." Stallybrass reveals how database has fundamentally altered his pedagogical approach, since our scholarly competitors are "no longer just our colleagues; in the age of database, they are also the students whom we claim to be teaching." This overturning of "proprietary authorship" is one of many emerging realizations of the still-dawning age of database.

Like Stallybrass, I believe this age of database has a long, precomputer history, stretching back to the first epics. Like Hayles, I believe that narrative is "an essential technology for human beings," but I also believe that database is the equally essential countertechnology, the innate desire to pile up and absorb experiences and ideas and material things that don't sort themselves immediately into narrative—items we can access later as pieces of a narrative if and when they fit the story, history, or syntax of meaning we are seeking to construct. Keeping a commonplace book edges toward database; keeping a journal, toward narrative. Our greatest and most evocative narratives, including the novels we teach, paradoxically become database when we write our interpretive narratives about them, using bits of the data to construct a meaning that is always exceeded by the data that do not fit the narrative we construct. The hermeneutical enterprise finds databases everywhere—even in narratives—and accesses them to create meaning. Database, in an age of computers, provides increasingly quick access to increasingly vast realms of thought, language, facts, and works.

Work Cited

"Database." Def. 1. *The Oxford English Dictionary*. 2nd ed. 1989.

https://doi.org/10.1632/S0030812900168592 Published online by Cambridge University Press