What Does "Born Digital" Mean?

ELIAS MUHANNA
Department of Comparative Literature, Brown University, Providence, R.I.;
e-mail: elias_muhanna@brown.edu
doi:10.1017/S0020743817000976

When asked for a definition of the digital humanities, I often fall back on a crisp formula. Digital humanists use computation to generate, organize, publish, or interpret humanistic data. This covers most of the bases, but it's also a bit abstract. "What sort of computation?" a colleague sometimes asks. "I use a computer at work; why aren't I a digital humanist?" To this, I agree; there isn't much daylight between digital and analogue scholarship. Less interesting than what separates the two domains is the question of what unites them. In my view, the most significant shared ingredient is technology.

It is not often appreciated that the basic tools of nearly all humanistic scholarship are forms of technology. I have in mind here the encyclopedia, the dictionary, the map, the gazetteer, the concordance, the timeline, the library, the diagram, the card catalogue, the index, and so on. Each of these tools "generates, organizes, publishes, or interprets" humanistic data, and sometimes does all of these at once. In thinking about the relationship between old and new types of scholarship, it is helpful to recognize that most digital humanities projects are, structurally speaking, digital renderings of analogue technologies.

This relationship is most apparent in initiatives that model themselves explicitly on familiar templates. The ARTFL Project, a collaboration between the University of Chicago and the French government, is known for its digital editions of historic encyclopedias, such as Diderot and d'Alembert's *Encyclopédie*, Ephraim Chambers' *Cyclopaedia*, and Pierre Bayle's *Dictionnaire historique et critique*. The Pleiades Gazetteer is the world's largest resource on ancient places. The Perseus Project is a digital library of classical texts. ORBIS is a geospatial network model of the Roman world, allowing users to calculate the time and expense of transportation between cities around the empire. Although all of these projects deploy new technologies in sophisticated ways, each is based on one or more analogue models (i.e., the encyclopedia, the gazetteer, the map, the library). The contribution of the digital realm is not, in many cases, to introduce a new vessel for humanistic data, but rather to offer a set of new mechanisms to obtain, visualize, and manipulate that data. In so doing, digital technologies contribute dimensionality, efficiency, accessibility, and interactivity to structural models that have been around for millennia.

In the past few years, however, there has been a growing sense that more can and should be done with digital technologies. As the digital humanities have become a magnetizing force in university life, academic institutions have grappled with the criteria for accrediting different sorts of projects. What should count for tenure and promotion? Which projects have the potential to break new ground in a field, and which merely serve as useful resources? Should digital projects that pattern themselves on analogue models—and particularly the sorts of models that aren't as highly prized as monographs and journal articles—be treated differently by funding agencies and promotion committees, just because they are digital? It's in the context of these discussions that one encounters an increasingly prevalent term: the project that is "born digital."

What does born digital mean? In some cases, it may refer to records or archives that have no physical footprint: things that began life as digital entities rather than analogue artifacts converted into digital form.[1] However, this is not the meaning I have in mind. When university and foundation administrators refer to initiatives that are born digital, they often seem to be imagining projects that sever the link with the analogue world. The born-digital project does not merely reproduce the organizing principles of analogue technologies such as the dictionary or the museum. It refigures and subverts them. It gives us something new. If it is difficult to put our finger on what that new something is, so the argument goes, that is because we are beholden to analogue ways of thinking about our research objects. This is the sense of born digital that interests me, which is as much an epistemological problem as it is an administrative one.

I am not convinced that projects need to be born digital to make groundbreaking interventions in humanistic scholarship. That is partly because the aforementioned contributions of the digital realm—dimensionality, accessibility, interactivity, and efficiency—have transformative potential. Digital renderings of analogue technologies put different sets of data in dialogue with each other, instantaneously. An online database of epigraphic inscriptions may be nothing more than a digital filing cabinet, but its ability to organize and cross reference data based on textual, chronological, and geographic variables, and to visualize it in various ways—all in a few moments—creates new ways of approaching old questions.

The value of such analogue-digital hybrid projects is particularly true for scholars of the Middle East. Despite the advances in optical character recognition that have been made in recent years, there is comparatively little material that has been digitized in this field; we are still at a capacity-building stage. When I speak to graduate students in Middle East or Islamic studies who express an interest in getting involved in the digital humanities, I try to impress upon them the substantive contribution that even a simple digitization and markup project would make, as opposed to a complex study involving the latest text-mining algorithms and network analyses. Before we can think past the analogue, we need a basic digital infrastructure in place.

Let me offer an example from my own research. A few years ago, I began work on a digital project about classical Arabic poetry with the help of an undergraduate research assistant at Brown University.[2] At the outset our aims were ambitious: we hoped to build a corpus of several thousand Arabic poems and to develop a variety of text-analysis tools that would help us detect different sorts of syntactic and semantic patterns. With that information in hand, we would cross reference it against other genres of historical data, which promised to tell us a great deal about the transformations of an art form often seen to be static. Naturally, the final product would be completely accessible to the nonspecialist, and interactive as well! Users would be able to upload their own samples of poetry and the program would compare them to the archetypes it had distilled from the corpus.

Needless to say, the project did not deliver as planned. We assembled a corpus and developed several text-analysis algorithms. Some of them worked; others flopped. Some confirmed my hypotheses; others generated surprising results. The public dimension of the project was neglected entirely as we struggled to produce a stable prototype. Interestingly, the problems that were consistently the most difficult to solve were not high-tech born-digital issues; they were the sorts of things that vex the most hidebound

analogue processes: establishing stable texts, accounting for variations in those texts, and dealing with problems of authorship. All along, I found myself appreciating the labor that goes into publishing a print edition of a classical poetry *dīwān*. If we had committed to doing nothing more than reproducing that particular analogue technology in digital form and making it available to other researchers, the outcome would have been far more valuable than what our abortive experiment rendered. At the very least, it would have established a platform upon which other scholars could build.

Rather than seeking to avoid replicating traditional approaches to our subjects, I feel strongly that aspiring digital humanists in Middle East studies should be building the infrastructure that will facilitate born-digital projects down the road. By the same token, instead of seeking out the latest digital methodologies and tools in the hope that they will unlock the secrets of our archives, scholars would be better served by asking the same questions they would in an analogue project, and only then investigating how a certain digital tool might facilitate their search for answers. In a funding environment where scholars feel pressured to produce cutting-edge projects in order for them to "count," it is understandable that many digital initiatives begin with conversations about methodologies rather than research objectives. But methodologies change quickly when technology is involved. In order to make lasting contributions, humanists have to think beyond the circumstances in which their projects are "born."

As the stuff of humanities research—poems, paintings, historical archives—slowly becomes accessible in digital form, it may be that the very term "digital humanities" will cease to mean anything at all within the next generation. In the meantime, there is much more work to do.

NOTES

[1]See Matthew Kirschenbaum, "The .txtual Condition: Digital Humanities, Born-Digital Archives, and the Future Literary," *Digital Humanities Quarterly* 7 (2013), accessed 13 September 2017, http://www.digitalhumanities.org/dhq/vol/7/1/000151/000151.html.

[2]See Elias Muhanna, "Hacking the Humanities," *New Yorker*, 7 July 2015, accessed 13 September 2017, https://www.newyorker.com/culture/culture-desk/hacking-the-humanities.