

Dean Irvine

Dalhousie University

Scholarly and Research

Communication

VOLUME 6 / ISSUE 2 / 2015

Abstract

This article positions historical and contemporary formations of the digital humanities in relation to different economic and business models. It examines the prototypical business partnership and economic relations between Father Roberto Busa and Thomas J. Watson as well as the collaboration between Busa and Paul Tasman at IBM. It also proposes a new business and economic prototype modelled on the principles of agile development and networks.

Dean Irvine is Associate Professor in the Department of English at Dalhousie University. He is Director of Editing Modernism in Canada, Founder and Director of Agile Humanities Agency, and Director of the Modernist Commons. Email: dean.irvine@dal.ca

Keywords

IBM; Digital humanities; Father Robert Bosa; Thomas J. Watson; Humanities computing

CCSP Press

Scholarly and Research Communication

Volume 6, Issue 2, Article ID 0201208, 8 pages

Journal URL: www.src-online.ca

Received May 29, 2015, Accepted July 13, 2015, Published September 3, 2015

Irvine, Dean. (2015). From Angel to Agile: The Business of the Digital Humanities, 6(1): 0201208, 8 pp.

© 2015 Dean Irvine. This Open Access article is distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc-nd/2.5/ca>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

[D]o not change IBM into International Busa Machines

— Thomas J. Watson, in conversation with
Father Roberto Busa, Busa, 1980, p. 20

To account for the business of the digital humanities is to compile genealogies of its economic, corporate, disciplinary, and institutional formations. These genealogical conditions are premised in the first instance on the historical conjuncture of computational linguistics and record-keeping technologies designed for census statistics and commerce, specifically International Business Machines (IBM) punch card machines. This is not only to gesture toward an archaeology of mid-twentieth century business machines and their application to humanities research but also to recognize that the conjuncture of commercial technologies and humanities computing has continued over the past half-century or so. By taking a longer historical view of the digital humanities and its formative partnerships with commercial enterprise, I want to advocate for the ongoing development of innovative business models as a means of achieving the sustainability for digital infrastructure required by research projects whose longevity is projected beyond the end of grant-based funding.

Contrary to anachronistic origin stories of the digital humanities, Father Roberto Busa's earliest experiments in humanities computing were conducted using analogue technologies and mechanical instruments. In the preface to his first machine-generated concordance, the *Varia Specimina* of 1951, Busa foregrounds the analogue mechanics of its computation and production: "The concordance which I am presenting as an example is precisely an off-set reproduction of tabulated sheets turned out by the accounting machine" (p. 28). Already a specialized type of counting, his concordances enlisted and evolved into instruments of accounting.

Busa's partnership with IBM, which extended over six decades, inaugurated a prototypical business model for humanities computing. "I could recompense IBM in any way except financially," Busa (1980) recalled telling then IBM chairman and CEO Thomas J. Watson prior to their now-legendary meetings in 1949. There were two meetings: the first, at which Busa made his pitch and Watson requested a formal proposal to distribute to his engineers, and the second, at which Watson was prepared to reject the proposal based on a report from his technical team, but changed his mind and decided to support it for a trial period. Busa's initial partnership was premised, from the onset, on IBM's ownership of Herman Hollerith's patents for the cardpunch machine, as well as card tabulating and sorting machines. Once he entered into this partnership, there was no economically feasible means of porting his proprietary data to another company's processing machines. Like IBM's business clients, Busa effectively entered into a licensing agreement, which guaranteed IBM recurrent – if unpredictable – returns on its investment. To end the agreement would not only have brought an end to their partnership; it also would have brought an end to his career-long investment in IBM's proprietary systems. Exemplary of a sustainable business model, Busa's partnership with IBM underwent massive technological and institutional changes across the course of more than half a century – a period from 1949 to 2010 that saw

their collaborations transition from punch cards to magnetic tape, card readers to mainframes, RAM to CD-ROM, and multivolume bound concordances to Web-based databases.

Scholarly and Research

Communication

VOLUME 6 / ISSUE 2 / 2015

While we await Steven Jones' forthcoming book about the meetings of Busa and Watson (Smyth, 2014), the details of their agreement and the financing of the projects that came out of the partnership between the Jesuit priest and the CEO of IBM can be based only on published accounts. To take Busa's recollection of the meetings at face value, as it were, the agreement resembles what we would now call an "angel investment" – that is, an investment by affluent individuals, companies, or trusts that demonstrates a willingness "to assume bigger risks and accept lower rewards when they are attracted by the nonfinancial characteristics of an entrepreneur's proposal" (Cetindamar, 2003, p. 42). Notably, the concept of angel investment has its origins in the history of early Broadway, where the arts and business meet directly, when so-called "angels" would finance theatrical productions (Cetindamar, 2003, p. 40). Unlike some angel investors, who sometimes engage more directly in their investments, Watson assigned IBM executive Paul Tasman to oversee the company's partnership with Busa. In any event, it is undeniably serendipitous that one of the period's most powerful capitalists should have played the role of "angel" to the priest. All the better that Watson and IBM backed Busa's work on machine- and computer-generated concordances and not a Broadway show about St. Thomas.

As the "proof of concept" (Winter, 1999, p. 8) for Busa's (1974-1980) magnum opus, *Index Thomisticus*, the *Varia Specimina* (Busa, 1951) is a signal example of production models adopted by mid-century engineers and, eventually, programmers. This stage is a typical business-model requirement and a step toward securing additional investment. As it happens, the *Varia Specimina* could also be classified as a prototype, since the proof of concept had already been worked out by means of "trials which were carried out on one of Dante's *Cantos*" (Busa, 1951, p. 24). The forensic detail that Busa provides in documenting his procedures in the *Varia Specimina* is akin to a laboratory report – or, given its potential audience at IBM, a marketing report. Perhaps Busa's (1951) most prominent finding is the predictable, yet telling, discovery that the "greatest hindrance" in conducting trials with punch card technologies is "transposing the system from the commercial and statistical uses to the sorting of words from a literary text" (p. 26). For IBM, the capitalization on these trials would require the realization of the obverse: to convert the literary text into a data system for commercial and statistical use. This, as it happens, was the advent of natural language processing and machine translation.

Working to create the successor to the analogue *Varia Specimina*, Busa collaborated with Tasman on digital projects in the mid-1950s, which included programming a machine-readable index to the Dead Sea Scrolls on magnetic tape read by an IBM 705. After working for several years out of the IBM offices in New York and Milan, Busa raised enough capital by 1956 to found the Centro Automazione Analisi Linguistica (CAAL), his "laboratory" (Busa, 2004), "training school for keypunch operators" (Busa, 1980, p. 85), and "Literary Data Processing Center" (Tasman, 1957, p. 256) at Gallarate, Italy. Reporting on their experiments at Gallarate in the July 1957 issue of IBM's in-house research and development journal, Tasman (1957) offered a prescient account

and made explicit some of the ways in which IBM expected to derive value from their collaboration: “The indexing and coding techniques developed by this method offer a comparatively fast method of literature searching, and it appears that the machine searching application may initiate a new era of language engineering” (p. 256). Although Tasman’s predictions were relatively modest, suggesting that the algorithmic processes that they had developed could “lead to improved and more sophisticated techniques for use in libraries, chemical documents, and abstract preparation, as well as in literary analysis” (p. 256), the history of information technology that has since transpired would support far more ambitious outcomes. If Busa’s legendary 1949 meetings with Watson initiated a business model for humanities computing, the returns on that investment would prove far more substantial than either man could reasonably have anticipated. Tasman’s promise of a “new era of language engineering” is the one in which we live now: it is the era of IBM as big data corporation. Watson’s meetings with Busa may well have launched the priest on a course to seek investors in his own research laboratory and data-driven empire, one that effectively translated “IBM into International Busa Machines” (Watson quoted in Busa, 1980, p. 20), but the greater empire would be built on the investment in Busa: this is the transmutation of Watson the angel investor into Watson the linguistically intelligent supercomputer, and thus the computational transvaluation of linguistic data into capital.

By the time Busa (2004) wrote the foreword to *A Companion to Digital Humanities*, who could possibly imagine writing a sequel to the story of a humanities scholar making the kind of pitch that he delivered in 1949? Writing in the aftermath of the attacks on September 11, 2001, he observed that we were living in “an unforeseen season of *lean kine*,” that he had witnessed “reductions in public funds for research,” but held out the promise that the “period will pass” and that “cutbacks in finance” could lead to “the according of priority to a definitive solution … which could facilitate the fulfillment of the globalization of economic exchange.” With these words, the priest became the CEO of a field that had recently rebranded itself as the “digital humanities.” These prognostications resounded with the belt-tightening neoliberal rhetoric of austerity economics and the triumphalist discourse of global capitalism. After six decades of working with IBM, it comes as no surprise that Busa could convincingly play the spokesman for the business empire that had backed his enterprise.

Upstarts and start-ups

Agile processes promote sustainability. The sponsors, developers, and users
should be able to maintain a constant pace indefinitely.

— “Manifesto for Agile Software Development,” Beck, Beedle,
van Bennekum, Cockburn, Cunningham, Fowler, Grenning,
Highsmith, Hunt, Jeffries, Kern, Marick, Martin, Mellor,
Schwaber, Sutherland, & Thomas, 2001

Contrary to Busa’s economic vision for the future of the digital humanities, the period of financial trouble and defunding research did not pass. Precarity became the new normal. This is the economic midden into which the digital humanities had sunk when we launched Editing Modernism in Canada (EMiC) in 2008. Precipitated by the global

financial crisis in the fall of that same year, the fallout of austerity economics descended upon university campuses. Despite EMiC's good fortune in securing public funds from the Social Sciences and Humanities Research Council of Canada (SSHRC), which guaranteed the sustainability of our work for seven years, it became increasingly apparent that longer-term support for our digital humanities research would require rethinking an economic model predicated upon leveraging our grant funds to procure additional resources from postsecondary institutions already partnered with the project at start-up. While the majority of our resources were already targeted toward training and funding research by emerging scholars, we realized early on that one key way to sustain these investments in a new generation of digital humanists would be to allocate resources toward the development of a common Web-accessible repository of digitized resources, a Web-based workbench of digital editing tools, and a customizable publication interface. This was the beginning of the Modernist Commons.

Of necessity, given the predictive model of grant applications and their corresponding budgets, it seemed logical at the outset to adopt a traditional "waterfall" methodology – that is, a sequential system of software development that proceeds stage by stage from requirements, modelling, and design to development, testing, and operations. After commissioning a white paper on available technologies, we invested resources from our grant in an open-source software services start-up, DiscoveryGarden, Inc., which had already secured capital investments from government agencies, universities, and public-institution and private sector clients. As it happens, however, the implementation of waterfall quickly proved incompatible with our strategy to adopt and adapt open source tools, many of which were still proofs of concept or prototypes. Along with the transformation of the start-up's methodologies from waterfall to agile – principally user stories, rapid feedback loops, iterative implementation, sprint-based rollouts, and test-driven development – we repositioned our project in partnership with our principal software developer. Given our desire to develop an adaptive open source system, we reached out to new institutional partners that already had alternative resources to develop tools that we wanted to adapt, incorporate, and re-release to the open source community. We subsequently parlayed our own software investments to back these allied and emergent partner projects. Meanwhile, austerity budgets and neoliberal restructuring of universities into corporate entities ran amok. By the time we looked up from our screens at the end of seven years, the sustainability model that we had devised by way of creating a distributed network of allied institutions and initiatives was still intact, albeit shaken at times by the inevitable tremors of running a longer-range project. Not only had EMiC entered into agile software development, it had shifted its sustainability model toward the formation of agile networks.

Once we had identified the tools that we required, we partnered with the institutions and organizations that managed their development. Unlike the partnerships that we formed at the onset of the project – which mainly targeted resources toward training students and funding their research, and which were based on a predictive budgetary model that allocated specific amounts over the course of the grant-funded phase of the project – these new partnerships were transparently and strategically oriented toward software development and often limited in duration to the period necessary for completion (or, in some cases, abandonment). Not all of these partnerships resulted in

usable tools; others proved most productive to advance a proof of concept or prototype, but dissipated before such tools could be made ready for a production environment. This is the character of an agile network: a team is formed rapidly to respond to a specific, mutually required need and disbands afterwards (Metes, Gundry, & Bradish, 1997). These kinds of networks are designed to respond to change; they do not replace the more durable partnerships that support long-range sustainability.

Another of the working concepts for EMiC's sustainability has been the commons; that is, the digital production of social goods, so that value is generated by collaborative labour and shared co-operatively among network participants and partners. This model is manifest in the project's digital repository and editorial workbench, the Modernist Commons, which invites contributors to upload content to a co-op accessible by network participants who can, in turn, reproduce versions for use in online editions (Irvine, 2014). Even so, the transferability of the commons as an economic model to support the EMiC network as a whole seems improbable, not least because its repository and workbench were built upon the assumption that contributors need not have access to (or part with) economic resources to participate in the communal production of digital knowledge and social goods. Busa's projection of the "globalization of economic exchange" may have been conjured as a "Utopia" (Busa, 2004), but the ideological differences between his neoliberal digital utopia and the "affective economies" (Ahmed, 2004, p. 119) of a user-supported commons are incommensurable.

Rather than await the revolutionary transformation of an economic system that facilitated the emergence of the digital humanities, I developed a proof of concept so that the capitalist model with which our community of practice has been partnered for six decades might be put to service to sustain our investment in the Modernist Commons. To do so, I drew upon the agile programming methodology that we had adopted and the agile networks that we had formed. Instead of planning a grant proposal, which obviously comes more naturally to me as an academic researcher, I drafted a business plan. The result was the incorporation of the Agile Humanities Agency in 2014.

Many digital humanists pride themselves on a DIY model that readily accommodates the realization of individual research initiatives (or a modified version of the same that includes hands-on training for student researchers) funded by granting agencies, others avail themselves of software development services provided at their home institutions, and still others follow the multi-institutional and cross-sector partnership model that funds large-scale academic and industry initiatives. My start-up disrupts none of these scenarios; it regularly supplements such organizational configurations, appending agile networks to operate in collaboration with other researchers and development teams. There is nothing remarkable about a start-up implementing agile methodologies or performing open source customization, nor is there anything innovative about providing third-party software services to the digital humanities, but there is a market for projects that present challenges beyond the expertise of student programmers typically drafted into service by grant-funded researchers, that exceed the capacities of services available at home institutions, or that fall below the profit-

margin threshold of larger software-industry shops. Rather than target the start-up toward my own research priorities, I positioned it in a niche of micro-scale, limited-budget, rapid-delivery open source software customization for the digital humanities. Instead of running a shop with salaried employees, the start-up operates as an agency to form agile networks between clients and rapidly assembled teams of programmers, designers, documenters, and trainers who work under contract for the Agile Humanities Agency. This is a for-profit agency that adheres to principles of open source development by redistributing its code via GitHub, but its profits are not just reinvested to facilitate its expansion of services, they are also redirected toward the sustainability of the Modernist Commons.

Digital humanities and start-up communities, as Lisa Spiro (2011) has recognized, are closely aligned, often sharing common principles, such as “agile development, user-focused design, open source software, and iteration.” Although she suggests that the digital humanities brings the “spirit of entrepreneurship” to the humanities, she admits that, “DH [digital humanities] projects typically don’t form companies and don’t aim to make a profit,” even if most need to find a way to “sustain themselves.” Why does the digital humanities community embrace, as Spiro puts it, the “spirit” of entrepreneurship – “taking risks, experimenting, building something that serves a need, innovating, tolerating failure” – but not the letter? After all, if not for the entrepreneurial work of the father of the digital humanities, our origin story would have to be rewritten. This is not to insist that there is only one sequel to Busa’s meetings with Watson, and that it must pursue a trajectory toward big business and big data; it is, rather, to recompile the genealogies of the digital humanities so that we might think through alternatives to the economic models that have sustained us so far. Even if the CEOs of global IT and big data corporations are not accepting our meeting requests, it might be better to think small and develop agile start-ups as proofs of concept and prototypes for a digital humanities economy that does not always need to scale up to the pace of global capitalism and that might just as well sustain itself by scaling down.

Websites

Editing Modernism in Canada (EMiC), editingmodernism.ca

Modernist Commons, modernistcommons.ca

References

- Ahmed, Sara. (2004). Affective economies. *Social Text*, 22(2), 117-139.
- Beck, Kent, Beedle, Mike, van Bennekum, Arie, Cockburn, Alistair, Cunningham, Ward, Fowler, Martin, Grenning, James, Highsmith, Jim, Hunt, Andrew, Jeffries, Ron, Kern, Jon, Marick, Brian, Martin, Robert C., Mellor, Steve, Schwaber, Ken, Sutherland, Jeff, & Thomas, Dave. (2001). *Manifesto for agile software development*. URL: <http://agilemanifesto.org> [August 26, 2015].
- Busa, Roberto. (1951). *Sancti Thomae Aquinatis hymnorum ritualium varia specimina concordantiarum: A first example of word index automatically compiled and printed by IBM punched card machines*. Archivum Philosophicum Aloisianum, Ser. 2 no. 7. Milano, IT: Fratelli Bocca.
- Busa, Roberto. (1974-80). *Index Thomisticus: Sancti Thomas operum omnium indices et concordantiae in quibus verborum omnium et singulorum formae et lemmata cum suis frequentiis et contextibus variis modis referuntur quaeque* (Vols. 1-56). Stuttgart-Bad Cannstatt, DE: Frommann-Holzboog.

- Busa, Roberto. (1980). The annals of humanities computing: the *Index Thomisticus. Computers and the Humanities*, 14, 83-90.
- Busa, Roberto. (2004). Perspectives on the digital humanities. In Susan Shreibman, Ray Siemens, & John Unsworth (Eds.), *A companion to digital humanities* (Foreword). Oxford, UK: Blackwell.
URL: <http://www.digitalhumanities.org/companion> [August 26, 2015].
- Cetindamar, Dilek. (2003). *The growth of venture capital: A cross cultural comparison*. Westport, CT: Praetor.
- Irvine, Dean. (2014). A modernist commons in Canada] Une commune moderniste au Canada. In Marie Carrière & Patricia Demers (Eds.), *Regenerations: Canadian women's writing/Régénérations : écriture des femmes au Canada* (pp. 39-55). Edmonton, AB: University of Alberta Press.
- Metes, George, Gundry, John, & Bradish, Paul. (1997). *Agile networking: Competing through the internet and intranets*. New York, NY: Prentice Hall.
- Smyth, Patrick. (2014, December 4) Steven Jones: the Priest and the CEO. *Student research commons*.
URL: <https://arc.commons.gc.cuny.edu/2014/12/04/steven-jones-the-priest-and-the-ceo> [August 26, 2015].
- Spiro, Lisa. (2011, Dec. 6). Startups and the digital humanities. *Digital scholarship in the humanities*.
URL: <https://digitalscholarship.wordpress.com/2011/12/06/startups-and-the-digital-humanities> [August 26, 2015].
- Tasman, Paul. (1957). Literary data processing. *IBM Journal of Research and Development*, 1(3), 249-256.
- Winter, Thomas Nelson. (1999). Roberto Busa, S.J., and the invention of the machine-generated concordance. *The Classical Bulletin*, 75(1), 3-20.