7 Digital research, the legacy of form and structure and the ResearchSpace system

Dominic Oldman

Introduction

Data is present in ever expanding digital repositories, fuelled by new digital technologies. However, these infrastructures have contributed little to advancing adequate forms of knowledge representation let alone solving problems of knowledge fragmentation, and from many perspectives have made matters worse. To many technologists this sounds counter-intuitive but the underlying forms and structures of data that technology creates to serve its needs are highly problematic. While this data can be technically linked it inherently lacks the necessary expressiveness required for progressive, diverse and inclusive knowledge processes. It relies on reduction and simplification to operate, and fails to express real world relations producing connections with limited value, inequality and built-in obsolescence. This makes the decision of subject experts to invest in it also problematic (Agar 2006). However, while textual narratives provide intellectual weight they are also inherently difficult to digitally interrelate effectively because of their semantic complexity, style and variety. Textual narrative provides qualitative richness but is difficult to relate to the wider social systems that they exist within. These narratives find it difficult to describe and relate the overarching structures and complex relationships that impact on and situate lived experience.

Research is a dialectic process in that there is a constant dynamic between sources and materials, their ongoing historical relations, analysis and explanation, and their appropriate representation, but textual narratives limit this dynamic to within their static bindings. They create stand-alone publications from which the tracing and connected growth of knowledge over time, within and across communities, is increasingly impossible, and this cries out for some type of technological framework. Equally data models and structures, and the associated mindsets of artificial categorisation that they promote and protect, are fixed and static and fail to provide an effective and sustainable framework for the evolution of knowledge crucial to challenging established histories.

The modern database, developed after the Second World War, was created to meet the needs of efficient business and commercial information processing. It quickly became the workhorse of general computing. Most organisations use and maintain different institutional databases and the same database systems are also used on personal desktop computers. We interact with highly visual websites and web pages, but behind them sit database information systems that orchestrate content through inert structures and provide access to various competing datasets for which we, and the companies who provide them, have little long term intellectual investment in the content. Despite the imperative behind the original design vision for databases, they have also taken up a significant role in humanities and social science research, and their terms of use have been accepted with little resistance. There has been little investment in research to find appropriate structures and forms of data capable of supporting the full dynamic of historical research to unite both experience and thought. While relatively large amounts of time and money are ploughed into new data systems, they contribute little to the overall understanding of human history evidenced by their temporality. Repurposing the database from storing customer records, supply chain information and financial transactions, without rethinking how historical patterns and real-world abstractions should be represented, means they simply repackage old data and old thinking and propagate inequality and a lack of diversity in terms of sources of valid knowledge.

This chapter explains and contextualises the Andrew W. Mellon Foundation supported ResearchSpace project at the British Museum. ResearchSpace is an open source system designed to reflect the complexities of relational and processual approaches to knowledge generation, integration, dissemination and preservation. It enables researchers, cultural heritage professionals and other community groups to empirically investigate and represent history from different contextual vantage points providing objectivity through diversity. It can construct propositions and arguments so that different explanations can be transparently synthesised, compared, rejected, reconciled and adopted. It allows transparency across research programmes providing a means of identifying progressive branches of research while delivering a knowledge base with meaningful provenance. It creates internally related data narratives at different levels of generality. However, the history behind the traditional data systems that have and are adopted by the humanities is important in understanding the ResearchSpace project and how different professional and non-professional communities can re-imagine their relationship with data to produce a diverse, contextualised and historically driven contribution to wider information infrastructures (<https://www.researchspace.org>).

This chapter provides a background section linking the issue of neoliberal influence in the Digital Humanities to the structure and form of data and technological approach, creating an incompatibility with progressive research methods. The second section on related research sets out a short

history of databases and their effect on data-oriented research projects, again relating this to a conflation of commercial and scholarly objectives or structures that undermine the challenge of representing meaningful historical knowledge in data to achieve intellectual sustainability. The last section provides a brief case study of the ResearchSpace system itself, concentrating on describing its methodological approach rather than individual technical functions which have preoccupied data driven Digital Humanities projects.

Background: Data to represent knowledge

One significant element of the Digital Humanities – structured data – is burdened by a commercial history which may explain concerns about neo-liberal influence (Allington, Brouillette, and Golumbia 2016). A burden because while many humanities scholars have ventured into digital methods asking new questions, they are forced to adopt structures and forms derived from commercial computing imperatives that were not designed to support historical research. This problem continues because of a lack of knowledge required to engage with an essential component of data orientated systems which is dominated by a different disciplinary tradition, computer science. Computer science, among other things, studies the structure and form of data including the meta-models of different systems of data representation. Humanists often examine the structure and form of textual narrative which plays an important part in conveying meaning, thinking and plausibility. However, when it comes to “data” the humanists’ concerns with structure and form evaporates. There is little attempt to critically analyse a given data structure which is not simply accepted, but ignored. The assumption is that data is data, and its reductive raw form is fixed, neutral and immutable.

This reduction, usually based on physical characteristics and identity, is incompatible with representing the complex outputs of research methods, except as a relatively narrow quantitative index. While these indexes have utility they ultimately create fragmentation and build obsolescence. Databases constrain the use of data such that it can only operate as a poorly contextualised reference, statistical dataset or finding aid, and in doing so creates inequality by what it omits, resisting progressive forms of knowledge generation and tending towards the preservation of old knowledge, rather than extending knowledge to new sources and new knowledge categories. The Digital Humanities has focused on function above the data layer, for example, the preoccupation with scholarly primitives, and as a result experiences the effects of technology commodification (Unsworth 2000).

There are two main issues with the structure and form of data in conventional database approaches. It restricts what can be represented and therefore ensures that data cannot support abstractions from the “real concrete” world. Its artificiality contributes towards its own static nature, narrowness

of representation and lack of synthesis, providing a simple referential pointer to other sources and never seeking to represent wider relations. It ignores the researcher's need to continually grow knowledge and change the focus of research as it evolves. This makes data and databases disposable resources because the resulting stunted surrogate is always inadequate beyond a narrow design brief and restricted by our understanding of what structured data is and what it can provide.

The potential of data orientated computing in research is that it could, with the correct structures and forms, remove the fragmentation of scholarship; to represent and interconnect, *inventio* (*investigation of materials and sources*), *dispositio* (*explanation and argument*) and *elocutio* (*appropriate representation*) which are themselves not standalone linear processes but dynamically related (Ricoeur 1995). In other words, while text can perform all three tasks – but with significant limitations – the database, by design and convention, cannot. The second issue is its conformance to instrumentation built into computer databases and associated conventions of documentation that impede the creation of “patterns and linkages of facts” and resulting perspectives (Evans 2001, 252). Despite 20th-century modernist visions of a totality of world knowledge through technology, modern data systems, both through top-down human control but also technological instrumentation, filter out essential political, social and epistemology aspects (Day 2014, 566).

This picture is not true of all computer scientists. Some have also expressed dissatisfaction with forms of data representation that serve the technology more than the user, but these have been mostly confined to the 20th century. In the current era, criticism has centred on the quality of implementation. For example, in the use of alternative structures like Linked Data, which theoretically exerts, through computer ontology, improved semantics. Here, some computer scientists have asserted the need to apply classic scientific principles. They define a research object as, “a container for a principled aggregation of resources, produced and consumed by common services and shareable within and across organisational boundaries” (Bechhofer et al. 2013, Introduction). However, this is not enough for humanities scholars because it fails to consider data beyond the mechanical empiricism of reproduction and repeatability, or confront the issue of the artificiality of representation (Bechhofer et al. 2013). The humanities researcher must spend significant capacity conforming to predetermined data systems which force them to disconnect their abstractions from the real world that they investigate (Feenberg 2005). This means that the opportunity for researchers to use computers to meaningfully synthesise and reconcile different empirically-based histories – different vantage points or perspectives – is resisted by the fallacy of a single scientific reproducible neutrality, which persists in the ideology of data systems, and “vulgar” empiricism. This is not just an issue for researchers but also organisations that hold and disseminate historical materials and sources.

Related research

Always historicize! (Jameson 2002, 9)

One of the main impediments to the mass commercial uptake of the computer after the second world war was the availability of an integrated information management system. Instead of creating all the different components themselves, something not feasible for most companies, a software system was needed that integrated operational and management tools into one package making it feasible for companies to adopt (Haigh 2016, 26). The first information or database management system was created by a commercial systems engineer, Charles Bachman (https://amturing.acm.org/award_winners/bachman_1896680.cfm). Bachman created the Integrated Data Store (IDS) at a time when there was no relevant academic community, “as a practical tool, not an academic research project”. It had a straightforward objective, to provide, “efficient and flexible handling of large collections of structured data [which] was the central challenge for what we would now call corporate information systems departments, and was then called business data processing” (Haigh 2016, 26). Bachman’s Integrated Data Store was the forerunner of all the subsequent database systems created by companies like IBM, and the database system that many desktop users would have used in the 1980s and 1990s, namely dBASE, as well as the subsequent relational database management system which still dominates today (Dbase, based on the xbase programming language, was produced by Aston Tate and later bought by Borland, specialists in programming tools, and it competed with other systems like FoxBase, purchased by Microsoft and used with the Windows operating system). These systems were aimed at commercial supply chains and customer record keeping (the prime example given in many of the database textbooks of the time), but the need to categorise and organise information was, and is, universal.

Personal computing made database systems like dBase accessible to everyone, and some humanities researchers made use of them, but mostly as statistical tools, or digital versions of their card file systems. The database system paradigm is not just one that demands technical conformity but engenders a particular mental approach to structured categorisation contrary to natural thinking which is based on sensory experience. Eleanor Rosch, the cognitive psychologist, showed that people naturally use experience or “prototypes” to best organise things rather than abstract definitions. In contrast, “[t]he classical theory...defines categories only in terms of shared properties of the members and not in terms of the peculiarities of human understanding” (Lakoff 1990, 8). Databases, founded on this classical view of categorisation, were not designed to support processes of human abstraction but to process transactions and keep track of commodities. Their design architecture, based on the Bachman tradition, was not based on social history, but margins and costs and efficient inventory.

Constructive challenges to established structures of data (as opposed to general rejection) have been rare (Stone 1979). Manfred Thaller, a historian often cited by digital humanists, concluded that relational (SQL) databases could not provide answers to historical questions. In his detailed analysis, they lacked support for the variability or “fuzziness” of historical knowledge (Thaller 2017). The process of “data normalisation” that removes redundancy from data schemas to obtain technical efficiency in quantity, as a side effect also removed, and continues to remove, the elements necessary to address the types of questions historians needed to ask, for which Thaller gave empirical examples (Thaller 2017, 196–199). While computer scientists in the 1960s, many of which had a psychological grounding, developed new ways of looking at computers from the perspective of augmenting human skills, these have not been the focus of mainstream computing (Thaller 2017, 199). These pioneering computer scientists were interdisciplinary and trans-disciplinary, actively investigating the interaction and relationship between humans and machines in complex processes. This approach has been lost to mainstream Computer Science and is largely absent from data-oriented Digital Humanities. The computer industry has provided researchers with new and innovative functions, and different ways of visualising data, but underneath the bonnet little has changed in terms of what these tools ultimately operate on, bringing many humanities researchers to the conclusion that while computers may be useful for some quantitative tasks, they are mostly ancillary and disposable. Researchers and technologists seem to agree that data is not for asking complex “why” questions, but provides scalable storage, statistical processing and, through networking and aggregation, greater access to existing sources.

Practical examples

In practical terms, these limitations, both technological and mindset can be seen in a programme of digital projects from the 1990s that have focussed on existing cultural heritage sources used by humanities researchers. These include museums, galleries, archives and libraries, which in many countries are maintained with public funding, or are the traditional custodians of sources within universities. The EU, for example, has maintained a programme of projects designed to integrate the accumulated “knowledge” of cultural heritage institutions from their previously private collection databases. The first significant EU project was Remote Access to Museum Archives (RAMA) a project initiated in 1992 involving museums such as the Musée d’Orsay in Paris, the Uffizi Museum in Florence, the Ashmolean Museum in Oxford and the Goulandris Museum of Cycladic Art in Athens (European Commission (Cordis) 1994). Its Principal Investigator, Dominique Delouis was the Director of a French telecommunications company. His objective was to provide a single interface for retrieving data and images from remote museums. It would generate a commercial service, aimed at

schools, researchers and other audiences, as well as commercial enterprises. A source of knowledge greater than the sum of its parts, and at the same time, a one-stop shop to facilitate income generation based on a single point of access to cultural heritage assets.

Delouis' reflections at the end of the project indicate why his aims were unachievable and why databases have done little in terms of progressive knowledge representation. His pronouncement was, "can one hope for the 'virtual museum'? Yes, if one holds a universalist view of the world where different contents could be moulded into identical forms. No, if one thinks that each system of representation should keep its characteristics regarding form as well as contents" (Delouis 1993, 127). This approach to data goes back to the very first data management systems which serviced commercial processes, and which informed the design of the modern database.

In reality, aggregating museum inventory systems could never create a meaningful virtual museum and didn't represent their accumulated knowledge which remains fragmented and inaccessible. It only produced a secondary reduction to the ones already created by individual institutions adopting the database paradigm for their administrative purposes. A cynical reading might suggest that the data was simply a finding aid for licensing images, an explicit focus of many subsequent projects. This rationalisation of content to promote efficient data processing and the conflation of commercial objectives with academic purposes, works against knowledge development and sustainability, meaningful data preservation and collaborative knowledge building, and even efficiency – for the reasons provided above – but is an approach adopted by many funded projects investing considerable funds but rarely providing significant results. RAMA and its successors illustrated that these new computer data systems did not reveal new insight but demonstrated that even minor differences between museum catalogues constantly trip up database software developers. Rather than researching and understanding how to accommodate the complexity of humanities knowledge in data, the technological transformation from above simply repackaged old data and forms. Useful in the sense of improved technical access, but bypassing the deficiencies of the content itself.

We now acknowledge that the data systems that Delouis sought to integrate and transform were themselves preservations of old Empire ordering and acquisition record-keeping, activities with little interest in wider global perspectives. Many modern open data strategies host the same anachronistic and homogenised data suggesting that new technology in the modern period is about, "giving you today another version of what you had yesterday and never a different tomorrow" (Curtis 2021). Similarly, Alan Lui points to a lack of any "sense of history" in digital networks and identifies mash-ups of shallow fragmented information masquerading as sophisticated networks of interconnections (Liu 2011, 22). Databases and networked data avoid the meaningful relations and context required by serious scholarship

and community campaigns for greater diversity, limited by a particular “*structure and form*” statically optimised to preserve a status quo, a perfect system to document Fukuyama’s end of history (Fukuyama 1989).

In this context, the open data movement’s call for data reuse is bound up with this repackaging of an old commodity for use in new digital networks bypassing current social issues voiced through other mediums. The discourse of social justice is not matched by changes to back-office infrastructures. From the front office, cultural heritage organisations now engage in discussions of inequality whether race, culture or gender, but in the back-office data is divorced from this debate and is part of a perpetuated static infrastructure. Only in a computer database project could you deliver the statement made by Delouis (still replicated today) claiming at the same time to be promoting educational and scholarly aims, and not be challenged. Colonially derived record-keeping practices still have a cloak of digital invisibility created through data’s reductive, hidden and deceptively innocuous state.

The trajectory of Delouis’ project was clear. It collaborated with Computer Interchange of Museum Information (CIMI), a Museum Computer Network initiative, to replicate what had been produced in Libraries with Machine-Readable Cataloging (MARC) – a single standard for interchange that would lower costs and make bibliographic transactions more efficient – similar to a clearing bank system. Library systems show a particular disregard for wider knowledge and context precisely because of their proximity to it (Delouis 1995). The “stack itself offers seemingly limitless opportunities for the prepared mind to find conjunctions and synchronicities or wander productively. And yet those possible conjunctions among the books, however vast in number, are limited - not only by the size of the library... but by the catalog, by whatever ordering principle determines which books stand next to another on the shelves”. Digital systems potentially provide, through a network of books, an interconnected map of world history and society, but instead we get a “tattered ruin” which keeps those networks of knowledge at arm’s length (Schnapp and Battles 2014, 91). The largely unsuccessful aim was to transfer the library clearing house system of homogeneous data interchange and apply it to other aspects of cultural heritage, but perhaps because of the independent nature of other heritage organisations and the lack of resources compared to library infrastructures, this has never been enforceable, but is a constant short-sighted objective.

The data machine – H2

Data is an increasingly valuable form because it can be processed in vast quantities by global networks of computers that inject it, employing opaque algorithms, into large numbers of digital services. We see everyday evidence of old data or old forms of data, transported through new networked technology, for example, through our smart televisions and digital assistants,

mobile devices and particularly through social networks algorithms, underlying how data is controlled and distributed with no way of understanding its provenance. The “subsumptive process of the modern documentary tradition is, from a certain historical perspective, the dominant *Zeitgeist* of our age but one that leaves remainders, which are increasing - in not only epistemological and technical senses, but even more importantly, in political and social senses, left out of the literal counting, processing, and indexing of knowledge, technology, politics, and professional and everyday social and cultural being” (Day 2014, 566). It is the omission of these remainders in this subsumptive process that speaks to the question of a neoliberal undercurrent in the Digital Humanities, and the participation or acceptance of it. Regardless of good intentions the construction of these systems inevitably comes into contact with the legacy of commercially driven data structures – and lack of facility is not a defence.

Usually defended on the basis that researchers come to digital systems with progressive questions, many of the tools and associated data structures used to investigate these questions were not designed for the purpose. The inherent design parameters of structured data systems, and their influence on research, have not been fully appreciated (Allington, Brouillette, and Golumbia 2016). Researchers, as they might for any other method, need to pay attention to the underlying structures and forms on which their questions are pursued, rather than commission tools at a high functional level and leave these underlying data design questions to those not invested in the subject matter itself. It is this fundamental area that the Digital Humanities has consistently failed to address. Delouis set the scene for a series of global projects, including many in Europe and the United States, that repeated the same pattern, using different technologies and providing new functionality, but always glossing over the nature of data itself. This focus on function rather than subject and content has parallels in the Digital Humanities with its equally high-level concern for function through what is known as scholarly primitives (high-level academic functions), without any study of how these functions are affected by the form of content they are applied to (Oldman, Doerr, and Gradmann 2016, 266).

In software development “information hiding” refers to both the segregation of components to protect against the effects of design changes, but also to prevent the user from deviating from the prescribed parameters that might lead to greater development overheads. The same mindset that is applied to the data is extended to the user interface which becomes a tool of constraint as well as access. This is not simply about efficiency, but about control and costs. In database systems designed to fulfil specific tasks the greater the flexibility given to the user, the more involved the development of the user interface and the greater the problem of maintenance of software distributed at scale.

Just as Charles Babbage aimed to construct mechanisms with his knowledge locked in to create pre-determined processes (mechanical information

hiding) so that workers in a production line would not have to confront variables or rely on their own experience and knowledge, so his digital successors adhere to this Taylorist tradition, despite its inappropriateness for the design of a digital knowledge-oriented environment. Computer database systems are designed to scale technically, but not intellectually; however knowledge processes require the latter more than the former, which in any event has become less problematic with modern computing technology. Typically, a knowledge worker or researcher would naturally want to change the underlying model of their systems to adapt to new circumstances, processes and relationships. The greater the dependency between the data and software the less able it is in dealing with real-world abstractions and the more inappropriate it becomes as a knowledge orientated tool. A digital version of Taylorism. We are all aware from our daily work that dealing with the real world means that knowledge workers must actively find workarounds to “fixed model” systems, like the use of unstructured notes, the overloading of fields, the relaxation of validation for data types and most clearly, the use of other disconnected systems, which generate a much larger but hidden inefficiency compared to the business problem they were aimed at. In reality, most of our valuable knowledge is still represented in personal word processor documents, spreadsheets and notebooks.

This resulting inefficiency is only justifiable in the sense that it satisfies the requirements of detached management information or regulatory requirements on which financial and organisational decisions are based. Conversely, the need for information systems that support knowledge-based processes relies on the ability to incorporate new knowledge and new categories of knowledge regularly, and often spontaneously. The gap between the information system and real changing processes grows over time draining resources, constraining innovation and giving misleading organisational indicators. When this gap becomes too great either a change is made to the system, which is generally avoided if possible, or the organisation bears a greater resource overhead and manages around it with reduced capacity. This gap is directly experienced by the everyday users of the system and is invisible to those that deploy them. However, even changes to information systems are likely to focus more on function and updating technology rather than on improving the value of the underlying data. The additional function may allow knowledge workers to find more workarounds but stuck with a similarly fixed data model. In a corporate environment, legacy databases can be sustained for a long time, for example, banking systems, and similarly in cultural heritage organisations where collection systems have only just come under scrutiny. While these problems are also relevant to the modern business that increasingly rely on knowledge workers, they represent a significant barrier for progressing research methods in digital environments in which they can only provide a nominal role, but at great expense – financially and intellectually.

Research and general knowledge systems need to express semantics and logic and produce a “data narrative” that is responsive and effective. They need to provide a general ontological commitment, but not mechanical empiricism, and include social perspective and the ability to represent epistemology. Research is a process of abstraction from the real world, “the real concrete...the world in which we live, in all its complexity”, and breaking it down into manageable parts, the “thought concrete”, bearing in mind that it all belongs to one universal reality (Ollman 1990, 27). True interdisciplinary and transdisciplinary digital research is impossible if digital abstractions are not representations of the real world.

Since the real world cannot be unpacked all at once by any one person or body, different vantage points or historical perspectives are produced from different contextual starting points (Ginzburg 2014, 15). A museum may represent objects using an empirical base but they select from a set of facts and from a particular vantage point which increasingly requires wider enrichment. A museum visitor may get a different narrative from an independent tour compared to the official one, and while not incompatible, are nevertheless separated and disconnected, rather than mutually enhanced and reconciled. Cultural heritage information systems tell you more about the objects’ place in the institution rather than their place in the wider world. From one perspective data projects that publish institutional data, like the one managed by Delouis, make the situation worse by transferring legacy data issues into wider networks, but from another perspective, they have made transparent the problems of the internal processes that maintain and perpetuate them. The question is now whether this realisation is confronted by reaching for the next generation of off-the-shelf technology and tweaking existing data vocabularies and standards, or whether we now aim the significant skills and experience of humanities scholars and community groups towards a critical analysis of the data structures that wash over and dilute richer representations, and radically redesign them to support the complexity of social history.

Contextualised representations that embrace variability and intellectual rigour solve issues of sustainability, value, longevity and inequality – but this is completely counter-intuitive to managers and technologists trained in efficient and scalable transactions but who dominate technical infrastructures defending the bedrock of the Babbage legacy. The tradition of professional archivists, curators and librarians, alongside their Academy counterparts is not one of simple physical characteristics, but about contextualisation – the dynamic interconnection of the physical and intellectual (Eastwood 2000). Even in commercial circles, business consultants like Peter Drucker advised against companies hanging on to modern forms of Taylorism which prevented innovation in a knowledge society – a phrase he coined. The computer systems which were developed after the two world wars to address commercial requirements never considered this intellectual element and raced to provide universal computing infrastructures targeting

managerial administration and quantitative transaction processing – not to encourage human creativity. Drucker himself separated the computer from the processes of knowledge calling computer systems, “morons”, but as a result they required more intelligent operators. He never envisaged that computers would ever take up an active role in decision-making (Drucker 2012, chap. 10). However, from a business sector perspective he saw that in ignoring the shift from the physical production line to the new information and knowledge-driven economy, a great inefficiency and cost was being created in organisations (Wartzman 2014). In humanities research, this mind-set also equates to a poor investment of research funds and intellectual fragmentation.

Babbage, considered by computer scientists as the founder of computing, was fascinated with Adam Smith’s concept of division of labour and was concerned with reducing cost and gaining control over previously craft processes (Tinel 2013, 266). He laid the foundations for Taylor’s, “scientific management”, (Babbage 1971) and was interested in how technology could remove humans from sight (Schaffer 1994). He aimed to use his privileged knowledge to relegate people to the role of servicing technology rather than using technology to enhance and augment human skill. The psychologically oriented computer scientists of the 1960s challenged Babbage’s approach to technology and researched on the basis that computers should enhance human intellect. Douglas Englebart’s report to the Director of Information Sciences, Air Force Office of Scientific Research, Washington DC, was called, “Augmenting Human Intellect: *A Conceptual Framework*”, setting out the potential of computers to help humans tackle ever greater complexity (Englebart 1962). However, much of modern data computing continues the tradition of the Babbage and Taylor vision, now in the form of machine learning that threatens a new stage of its development, never allowing true knowledge systems the space to dislodge the computers’ traditional hard-coded role.

What is the complexity that modern humanities disciplines might need computers for to help them with? One answer came from historians, a discipline well known for its technological caution. In 1970, the historian Eric Hobsbawm outlined the major challenge for historians which he thought could only be answered by technology. In a conference on Historical Studies, he stated that “[t]he most serious difficulty may well be the one which leads us directly toward the history of society as a whole. It arises from the fact that class defines not a group of people in isolation, but a system of relationships, both vertical and horizontal. Thus it is a relationship of difference (or similarity) and distance, but also a qualitatively different relationship of social function, of exploitation, of dominance/subjection” (Hobsbawm 1998, 115). This system of relations at different layers marks a slow but fundamental change in progressive humanities methodologies but it is exactly this type of representation that computer systems prevent. More than 20 years later, sociologists also declared that they needed to

make a fundamental decision of “whether to conceive of the social world as consisting primarily in substances or in processes, in static ‘things’ or in dynamic, unfolding relations...increasingly, researchers are searching for viable analytic alternatives, approaches that...depict social reality... in dynamic, continuous, and processual terms” (Emirbayer 1997, 281). There are other examples of this new trajectory in other disciplines which need to address complexity, but to date, the provision of computer data systems with appropriate structures and forms that might address these fundamental requirements are not even on the radar of most information and computer scientists. While technologists see scholars as technologically conservative, it is the technologies offered to humanities scholars that are unsophisticated and regressive.

One of the most adventurous digital humanities projects of the 21st century was Project Bamboo, run by the University of California, Berkeley and the University of Chicago, initiated in 2008 (Dombrowski 2014). It aimed to solve the problems of previous projects and bridge the gap between humanists and technologists providing the tools that humanists “required” to address their questions and to resolve the continuing problem of fragmentation. The project spent 18 months travelling the world visiting different institutions listening to academics but with a predetermined notion of what technological frameworks were needed; ones that relied yet again on the same data structures and forms, but through a new technology architecture. They were shot down by scholars who approached the problem from the perspective (and defence) of their own knowledge systems and methods. Each new project, like Bamboo, fails to investigate the history and issues of previous ones. Bamboo’s planning document included a phrase that sums up the technologists’ blindness to serious historical and cultural heritage data representation. After considering the technical infrastructure and tools, which themselves had already been pre-selected from existing systems, the project reserved a paragraph revealing their approach to data. It asked the question; “[w]ho will do the data shovelling?” (Broughton and Jackson 2008, 7). There is no consideration of a need to address the limitations of traditional data structures and find a basis for academic scholarly communication in a structured form. The same data, data standards and vocabularies are preserved, but in new magical re-packaging based either on the traditional database and/or its worldview, which is transferred into the next new technology, like Linked Data (Langmead et al. 2018, 158). Project Bamboo failed to provide a solution and as a parting gesture also emphasised the importance of function over structure and form, again through the lens of “scholarly primitives”, without considering the possibility and importance of combining real word abstraction, ontological commitment and human thinking. Project Bamboo rejected Linked Data as a viable technology, believing that it wouldn’t technically scale. In reality, their results would have been the same regardless, indicated by their functional-technical priorities.

Case study

ResearchSpace background

The ResearchSpace project was initiated by the Andrew W. Mellon Foundation both out of frustration with, and interest in, digital humanities projects. The Foundation had set up a funding program called “Research into Technology” which operated between 2000 and 2010 working across Higher Education and Arts and Culture.ⁱ This programme made a significant contribution to open source higher education software developed and used across the world, and provided additional skills and experience to other Mellon programmes. In terms of scholarly communication and cultural heritage programmes, the Mellon Foundation was aware of a growing and significant problem in digital humanities research projects. They had funded a large number of digital projects allowing universities and cultural heritage organisations to build up digital data repositories, but expected to see these resources actively used and built upon by their beneficiaries. The Foundation funded particular data-driven pilot projects at different institutions that produced digital monographs – ones not easily reusable by others and lacking the qualities necessary for them to challenge traditional text-based sources or established methodologies. While the Mellon Foundation focused on the need to leverage computer networks to facilitate greater collaboration and sharing of knowledge between researchers and institutions and to provide accessibility of information to wider audiences, they were confronted with structural and organisational problems.

Each project used technical infrastructures that employed different flavours of database technology which, as argued above, didn’t provide the necessary properties for the sustainable forms of scholarly communication that the Mellon Foundation sought. The projects either slotted into a traditional catalogue publishing model and were more about digitisation, or employed processes too reliant on particular software and data models which were impossible to share and build upon. While the former may have been incorporated into existing institutional catalogues and made available as static Web references, the latter was never institutionally supported beyond their external funding. Institutions were unwilling to invest in these new infrastructures. They had not bought into the Mellon’s community-based vision and didn’t understand the challenges and benefits of forging a different kind of collaborative paradigm in cultural heritage. Many institutions hadn’t even considered the potential benefits of open source models and the concept of sharing the costs of developing mutually beneficial community software. In many organisations this represented a risk rather than an investment for the future. In the main, cultural heritage organisations

i Programme Managers, Ira Fuchs and Christopher Mackie

procure basic proprietary software which has a consistent cost and support model, and in theory a minimal impact on internal resources. However, in reality, this has meant a lack of choice, flexibility and innovation, both from the market and within the institutions themselves. It has generated a considerable hidden cost in terms of internal productivity and knowledge preservation, affecting traditional social and educational aims. A “degenerative” branch.

The progressive way forward was, and is still not, clear. The lack of cross-boundary terms of reference which fully explain the sensitivities regarding technology, clearly apparent in Project Bamboo, between the subject experts (the humanities scholars and curators) and the technologists, remains undefined. The dominance of function and efficient data processing, artificial categorisations and top-down control create apprehension, caution and disinterest in academic circles. While computers provide quantity they are not able, under traditional computing paradigms, to transfer the complexity taken for granted in textual narrative and, “situate a work within the many networks from which it gains meaning and value, and then present the results within complex visual arguments—the kind that were elaborately constructed on slide tables before being reduced to side-by-side comparisons for lectures or standard print publications” (Drucker 2013, 6).

The ResearchSpace project started life investigating one thing; the structure and form of data in digital environments. In a letter dated 25 August 2009, addressed to the Directors of the London National Museums, the Andrew W. Mellon Foundation proposed a “Shared Infrastructure”. Like many digital infrastructure projects part of the rationale was to reduce costs, but not in the conventional sense. The infrastructure would aim to create intellectual and scholarly sustainability by building systems not bound to the conventions of traditional databases. It would concentrate on separating the dependency between software and information, using Linked Data as a way of representing all relevant meaning and logic. The infrastructure would allow projects with different aims to provide vantage points, authored and presented in multiple ways but which were naturally compatible, not through software or vocabularies, but true ontologies. The ontological model ensured a common framework based on reality, not artificial categorisation and escalating artificial standards. Such an approach needed a significant amount of research in its own right, funded in stages as inroads were made into understanding and solving the problem.

Design principles of ResearchSpace

Narrative is defined in the Oxford English Dictionary (OED) as, “[a]n account of a series of events, facts, etc., given in order and with the establishing of connections between them; a narration, a story, an account” (Oxford University Press 2021). Usually narrative is associated with textual, oral or visual accounts. Narrative is not usually associated with databases

that assume a referential role providing specific uncontextualised results to queries which may then be used to inform a narrative conducted elsewhere – although database content is not generally formally cited as another implication of their unsustainable form and short life. Database queries are rarely the same as the actual questions being investigated and the answer is similarly detached. Only the person making the query is aware of its potential significance concerning an investigation and a line of thinking. Understanding the context of a question is part of the systems of knowledge, but it is never divulged to the database and therefore to its community of users. It is a limited form of scholarly communication, often on purpose to protect an individualistic approach. For example, I search for a letter in a database based on a query of the sender, receiver and date, leading me to a letter that I read to only then determine its significance to the actual questions that form the basis of my research. There is no intellectual narrative in the structured data (although there may be some comments in unstructured text fields) and there is no potential for any ongoing, let alone multi-faceted, accounts. The interactions are only acknowledged at the system log level. The notion of structured data providing a narrative may seem alien, but that is the critical starting point for understanding how structured data might be re-imagined to confront the problem of providing a narrative in data. The ability to generate such a narrative and at the same time use the benefits of computers to read, support underlying processes (augment intellect), compare and traverse these patterns of data narrative, completely changes the nature of structured data environments.

In the distant past, some historians attempted to provide descriptive factual histories in the text without including an explanation, for fear of introducing subjective elements that would undermine their empirical scientific work. This type of account is commonly called a “chronicle” history. The OED defines chronicle as “[a] detailed and continuous register of events in order of time; a historical record, *esp.* one in which the facts are narrated without philosophic treatment, or any attempt at literary style” (Oxford University Press 2021). In modern scholarship these types of history, while having some historical use, generally have little benefit to ongoing scholarship and even the empirical information can be difficult to extract and understand in terms of its purpose and significance. What distinguishes historical narrative from historical chronicle is the construction of “patterns and linkages” (real processes and relations) which involve an element of judgement and imagination, but still grounded in empirical evidence (Evans 2001, 252). This process can be equated to general methodology in scientific research programmes, such as those described by Imre Lakatos, the influential philosopher of science and mathematics. According to Lakatos (summarised by Burawoy), “in a progressive program, the new belts of theory expand the empirical content of the program, not only by absorbing anomalies but by making predictions, some of which are corroborated. In a degenerating program, successive belts are only backward looking, patching up

anomalies in ad hoc fashion, by reducing the scope of the theory, or by simply barring counterexamples. In a degenerating program new theories do not anticipate new facts, and thus knowledge does not grow" (Burawoy 1990, 766).

The historian that, with others, pursues a single reality of the world and who rejects relativism, uses empirical research methods to uncover new facts. But to grow knowledge they must interpret this evidence which necessitates a selection of what they consider to be the most relevant facts and whose relations provide more than intrinsic insight (Schmidt-Glintzer, Mittag, and Rösen 2005, 156). This provides a particular view of the world, an angle or a *vantage point*. These empirically based interpretations are a catalyst for further work and the discovery of new facts and thinking. Communities of professional humanities scholars read the latest research and contribute to it both in terms of countering facts and theories as well as developing them. In other words, the process of understanding the reality of history is dependent on an ongoing, dynamic and the collaborative scientific process of interpreting empirical evidence some of which will, through the integration and synthesis of the work of others, either be progressive, generating new investigation and evidence or be degenerative with a risk of preserving static and anachronistic knowledge. These "predictions" or interpretations through the analysis of facts are a necessary part of the process of objectivity which cannot be achieved through individual detachment but through the continual dynamic activity of community development and testing. This is not a three-year funded project, but a continuous everyday process that requires the different branches of a scientific programme to be recorded and preserved in a meaningful way, both progressive branches and the degenerative dead ends. *Therefore, digital systems that don't support this long term and ongoing scientific process are themselves part of degenerative branches and have short life-spans, or are artificially sustained and are damaging to research and the dissemination of contextual and diverse knowledge.*

While historians may come up with different conclusions, the ongoing task of research is to pursue progressive branches and deprecate degenerative ones (but maintaining the provenance) considering, comparing, reconciling and reaching new conclusions all devoted to understanding historical reality. Textual narratives even over relatively short periods are extremely difficult to synthesise, not least because of the different language and style that they use. The further away the reader is from the professional community that generates the content the harder it becomes to receive a complete understanding of the direction and flow of the different branches of particular programmes, which over time become fragmented across numerous publications. Only the professional historian has the time to follow it, but even within their discipline there is fragmentation across the many specialisations. Computers provide a potential answer. They do not have the same aesthetics and storytelling qualities of a textual narrative, but they can get very close with the right "relational" structures and capture the dynamic

relationship between *inventio*, *dispositio* and *elocutio*. Data, alongside the textual narrative, can solve one of the biggest problems that historians and humanities scholars face in a world overflowing with uncontextualised information. To precisely track the progress of different interpretations and explanations, based on different, overlapping and emerging empirical facts with a transparent and traceable provenance. Existing structures and form and associated mindsets prevent this.

As Hobsbawm put it, “the wider the range of human activities which is accepted as the legitimate concern of the historian, the more clearly understood the necessity of establishing systematic connections between them, the greater the difficulty of achieving a synthesis. This is, naturally, far more than a technical problem of presentation, yet it is that also. Even those who continue to be guided in their analysis by something like the ‘three-tiered hierarchical’ model of base and superstructures, may find it an inadequate guide to presentation, though probably a less inadequate guide than straight chronological narrative” (Hobsbawm 1998, 250). Current uses of technology in the humanities are used at arm’s length to ensure that their reductive qualities do not affect the purity of research methods, but at the same time, this reluctance to engage in the details of computer information structures means that the impact and influence of humanities scholars are reduced and more importantly their scientific programmes become increasingly irrelevant.

The communities of knowledge in which scientific programmes of history are conducted, using traditional modes of authoring and publication are too narrow and immediately bring into question the validity of these research programmes. This is still, of course, a question and problem for natural sciences also. Without technology humanities researchers cannot resolve issues of fragmented research programmes let alone provide a sufficient synthesis of knowledge. In any event, history is not linear and cannot be presented as such if we value progressive research with social responsibility. Textual narratives hit a serious limitation in their ability to interconnect and communicate patterns across time and space and more particularly make the connection between past and present, or present and past, the dynamic from which we gain new insight. Computers provide a potential answer but only if they shake off their commercial constraints, provide the ability to establish “patterns and linkages” (or rather relations and processes) using new forms and structures that incorporate argument and explanation and allow subject experts to take control of how information and thinking are abstracted into structured environments.

Conclusion: ResearchSpace design dynamic

The ResearchSpace project was approached as a humanities research project, not an IT solution, using some aspects of a design approach described by Löwgren and Stolterman in the MIT book “Thoughtful Interaction Design: A Design Perspective on Information Technology” (Löwgren and

Stolterman 2007). The authors describe a dialectic design method that treats the process as a research project, requiring a full investigation and understanding of a given subject area, rather than simply the design medium. In addition, the products of design, “are not artefacts, but knowledge” (Löwgren and Stolterman 2007, 2). ResearchSpace is a knowledge building project investigating and understanding the processes and relations involved in aspects of humanities research and the ResearchSpace platform uses this knowledge to help researchers approach their own research dialectically (Löwgren and Stolterman 2007, 12). Löwgren and Stolterman saw that system design based on “useability and usefulness” created a narrow and limiting design position (Löwgren and Stolterman 2007, 131). The ResearchSpace project uses their concept of “thoughtful design” which is based on real research and experience, rather than rapid technological development which prevents researchers from “experimenting with and learning about all the new possibilities created by new technology and new knowledge... [and]...deal with a reality marked by complexity and change” (Löwgren and Stolterman 2007, ix). The next generation of humanities researchers need not be programmers, but they must be designers and modellers of their own data and systems just as they design the structures of textual narratives.

As such the ResearchSpace project has not been a linear process of requirements gathering, specification creation and build. It has been a dynamic process of research, communication, testing and collaborating with communities also progressing new relational and processual research methods, ones that create insurmountable challenges for textual narratives. One part of this community developed an ontology which supports the same contextual approach employed by ResearchSpace. The International Committee for Documentation of the International Council of Museums – Conceptual Reference Model (CIDOC CRM) is itself an evolving empirical research project with significant practical potential (<http://www.cidoc-crm.org/>). Its recent growth has resulted in different branches of development, some progressive while others less so. Its wider adoption means that its scope, the new specialisms of ontological commitment it addresses, places pressures to reassess its own scientific base (the sign of a true scientific research programme), particularly when addressing the issues of describing social ontology and grappling with the internal relations of social structure and agency. In contrast, many computer ontologies used in Linked Data are based on existing artificial schemas which transfer existing vocabulary and protect established models and standards, propping up degenerative programmes of research and outmoded professional standards.

The point of ResearchSpace is not simply to create tools that implement function but to provide forms of representation that are as close as possible to original research abstractions and enable data narratives. The adoption of ResearchSpace flows from the need to address complexity and wider community knowledge. As such the groups and individuals now using ResearchSpace transfer these methods into a digital environment to obtain

the benefits of computers without losing the provenance and semantics of abstractions and explanations – although in reality all systems, whether computers or pen and paper, always carry some constraints and limitations. One challenge is that many projects will be transferring data from existing platforms to ResearchSpace and therefore will be transferring elements of previous approaches and mindsets. For example, current projects like “Pharos” are bringing together art archives from institutions based in the United States and Europe to create an aggregated research resource. This initially equates to the migration of legacy data to, initially, simple ontological patterns based on the original records (Pharos Consortium 2021, <http://pharosartresearch.org>). The ongoing challenge, in contrast to projects like RAMA, is to encourage researchers to approach this data with a new mindset and assume a position of dynamic transformers of the information and its structures, changing it from a static reference to a living archive hosting new ideas, explanations and arguments. For new projects, a progressive approach can be adopted from the start. Other active ResearchSpace contributors like Villa I Tatti, The Harvard University Center for Italian Renaissance Studies, are now using ResearchSpace in various projects including, “Venice as an Archipelago” and the innovative, “Metapolis: Spatializing histories through archival sources”. These projects use temporal-spatial layering in the study of early modern Venice (Villa I Tatti 2021). They provide a detailed microcosm from which many other projects at different levels of generality (including global systems) can grow from and provide fresh insight building on the original research. Taking a microcosm and linking it to overarching historical social systems, is the type of research activity that ResearchSpace makes possible. Equally, the Linked Infrastructure for Cultural Scholarship (LINCS) consists of a network of researchers from various Canadian universities investigating socio-historical questions using and developing the ResearchSpace platform. Their particular objective is to create, “a smarter, ‘semantic’ web whose links will elucidate the diverse causes, effects, and significance of human action and expression” (LINCS 2021).

ResearchSpace is also aimed at organisational transformations, in particular long-term growth of knowledge and knowledge processes across different organisational departments and beyond the institutional walls – informing research and generating intellectual efficiencies which provide organisational and personal benefits. The inefficiencies created by the top-down deployment of fixed model technology (including non-digital) stifle knowledge-based processes (the branches of organisational research), circumvent back-office innovation, affect the long-term preservation of organisational knowledge and prevent the incorporation of community knowledge. Just as in commercial organisations, project-oriented management approaches which have no underlying knowledge infrastructure create informational and process fragmentation. However, traditional cultural heritage processes such as archiving, conservation, collection research and external engagement could be supported and interconnected through a

knowledge base structure that supports internal and external collaborative working. This type of transformation is crucial to organisation's ongoing relevance for research, education and social responsibility, effectively bringing meaningful knowledge to the surface which is accessible to wide audiences, rather than producing limited static catalogues and records.

The details of individual tools and functions can be found on the project's web site. They will evolve and change over time. The important factors are, the ability to use dynamic and dialectic data forms and structures, a progressive and collaboratively orientated community, a willingness to engage with innovation and assert new designs based on historical challenges, and a commitment to providing context and diversity in data. These aspects are increasingly important to confront the considerable issues of technology disruption and commercially led global information infrastructures.

Bibliography

- Agar, Jon. 2006. "What Difference Did Computers Make?" *Social Studies of Science* 36 (6): 869–907. <https://doi.org/10.1177/0306312706073450>
- Allington, Daniel, Sarah Brouillette, and David Golumbia. 2016. "Neoliberal Tools (and Archives): A Political History of Digital Humanities. Los Angeles Review of Books." 1 May 2016. <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>
- Babbage, Charles. 1971. *On the Economy of Machinery and Manufactures*. A. M. Kelley.
- Bechhofer, Sean et. al. 2013. "Why Linked Data Is Not Enough for Scientists." *Future Generation Computer Systems* 29 (2): 599–611.
- Bridle, James. 2018. *New Dark Age: Technology, Knowledge and the End of the Future*. London: Verso.
- Broughton, Janet, and Gregory A. Jackson. 2008. "Bamboo Planning Project: An Arts and Humanities Community Planning project to Develop Shared Technology Services for Research: A Proposal to the Andrew W. Mellon Foundation." January, 1–39.
- Burawoy, Michael. 1990. "Marxism as Science: Historical Challenges and Theoretical Growth." *American Sociological Review* 55 (6): 775. <https://doi.org/10.2307/2095745>
- Curtis, Adam. 2021. *Can't Get You Out of My Head – Series 1: 1. Part One – Bloodshed on Wolf Mountain*. <https://www.bbc.co.uk/iplayer/episode/p093wpgw/cant-get-you-out-of-my-head-series-1-1-part-one-bloodshed-on-wolf-mountain>
- Day, Ronald E. 2014. "Indexing It All: The Modern Documentary Subsuming of the Subject and Its Mediation of the Real." In *iConference 2014 Proceedings*. iSchools. <https://doi.org/10.9776/14140>
- Delouis, Dominique. 1993. "TOIOsystixnes France." *International Cultural Heritage Informatics*, 117–127. <http://www.archimuse.com/publishing/ichim93/delouis.pdf>
- Delouis, Dominique. 1995. "Standards for Cultural Heritage Information: The RAMA Perspective." *Archives & Museum Informatics (Selected Papers from the Third International Conference on Hypermedia and Interactivity in Museums (ICHIM 95/MCN 95) vol. 2: 281–282.*

- Dombrowski, Quinn. 2014. "What Ever Happened to Project Bamboo? The Andrew W. Mellon Foundation." 1 September 2014. <https://mellon.org/news-blog/articles/what-ever-happened-project-bamboo/>
- Drucker, Johanna. 2013. "Is There a "Digital" Art History?" *Visual Resources* 29 (1–2): 5–13. <https://doi.org/10.1080/01973762.2013.761106>
- Drucker, Peter F. 2012. "The Manager and the Moron." In *Technology, Management and Society*. London: Routledge.
- Eastwood, Terry. 2000. "Putting the Parts of the Whole Together: Systematic Arrangement of Archives." *Archivaria* 50 (1): 93–116.
- Emirbayer, Mustafa. 1997. "Manifesto for a Relational Sociology." *American Journal of Sociology* 103 (2): 281–317. <https://doi.org/10.1086/231209>
- Englebart, Doug. 1962. *Augmenting Human Intellect: A Conceptual Framework*. Menlo Park: Stanford Research Institute. <https://www.dougenlebart.org/content/view/138>
- European Commission (Cordis). 1994. "Remote Access to Museum Archives | RAMA Project | FP3 | CORDIS | European Commission." <https://cordis.europa.eu/project/id/R2043>
- Evans, Richard J. 2001. *In Defence of History*. London: Granta Books.
- Feenberg, Andrew. 2005. "Critical Theory of Technology: An Overview." *Tailoring Biotechnologies* 1 (1): 47–64.
- Fukuyama, Francis. 1989. "Have We Reached the End of History?" [http://bev.berkeley.edu/ipe/readings/Fukuyama%20\(corrected\).rtf](http://bev.berkeley.edu/ipe/readings/Fukuyama%20(corrected).rtf)
- Ginzburg, Carlo. 2014. "Distance and Perspective: Reflections on Two Metaphors." In *Historians and Social Values*, edited by Ann Rigney, and Joep Leerssen, 218–232. Amsterdam: Amsterdam University Press.
- Haigh, Thomas. 2016. "How Charles Bachman Invented the DBMS, a Foundation of Our Digital World." *Communications of the ACM* 59 (7): 25–30. <https://doi.org/10.1145/2935880>
- Hobsbawm, Eric. 1998. *On History*. London: Abacus.
- Jameson, Fredric. 2002. *The Political Unconscious: Narrative as a Socially Symbolic Act*. 2nd ed. London: Routledge.
- Lakoff, George. 1990. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: University of Chicago Press.
- Langmead, Alison, Tracey Berg-Fulton, Thomas Lombardi, David Newbury, and Christopher Nygren. 2018. "A Role-Based Model for Successful Collaboration in Digital Art History." *International Journal for Digital Art History* no. 3. <https://doi.org/10.11588/dah.2018.3.34297>
- LINCS. 2021. "LINCS Cyberinfrastructure Project | CWRC/CSEC." *Canadian Writing Research Collaboratory*. <https://cwrc.ca/news/LINCS-CFI>
- Liu, Alan. 2011. "Friending the Past: The Sense of History and Social Computing." *New Literary History* 42 (1): 1–30. <https://doi.org/10.1353/nlh.2011.0004>
- Löwgren, Jonas, and Erik Stolterman. 2007. *Thoughtful Interaction Design: A Design Perspective on Information Technology*. Cambridge MA: MIT Press.
- Oldman, Dominic, Martin Doerr, and Stephan Gradmann. 2016. "ZEN and the Art of Linked Data New Strategies for a Semantic Web of Humanist Knowledge." In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 2nd ed, 251–73. Holboken NJ: John Wiley & Sons.

- Ollman, Bertell. 1990. "Putting Dialectics to Work: The Process of Abstraction in Marx's Method." *Rethinking Marxism* 3 (1): 26–74. <https://doi.org/10.1080/08935699008657901>
- Oxford University Press. 2021. "Chronicle, n." In *OED Online*. Oxford University Press. <https://www.oed.com/view/Entry/32576>
- Oxford University Press. 2021. "Narrative, n." In *OED Online*. Oxford University Press. <https://www.oed.com/view/Entry/125146>.
- Pharos Consortium. 2021. "PHAROS: The International Consortium of Photo Archives." <http://pharosartresearch.org/>
- Ricoeur, Paul. 1995. "History and Rhetoric." In *The Social Responsibility of the Historian*, edited by François Bédarida, 7–24. Providence RI: Berghahn Books.
- Schaffer, Simon. 1994. "Babbage's Intelligence: Calculating Engines and the Factory System." *Critical Inquiry* 21 (1): 203–227.
- Schmidt-Glintzer, Helwig, Achim Mittag, and Jörn Rüsen, eds. 2005. *Historical Truth, Historical Criticism, and Ideology: Chinese Historiography and Historical Culture from a New Comparative Perspective*. Leiden Series in Comparative Historiography, vol. I. Leiden: Brill.
- Schnapp, Jeffrey T., and Matthew Battles. 2014. *The Library Beyond the Book (Metalabprojects)*. Cambridge, MA: Harvard University Press.
- Stone, Lawrence. 1979. "The Revival of Narrative: Reflections on a New Old History." *Past & Present* 85 (1): 3–24.
- Thaller, Manfred. 2017. "The Need for a Theory of Historical Computing [1991]." *Historical Social Research/Historische Sozialforschung. Supplement* 29: 193–202.
- Tinel, Bruno. 2013. "Why and How Do Capitalists Divide Labour? From Marglin and Back Again through Babbage and Marx." *Review of Political Economy* 25 (2): 254–72. <https://doi.org/10.1080/09538259.2013.775825>
- Unsworth, John. 2000. "Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?" <http://people.lis.illinois.edu/~unsworth/Kings.5-00/primitives.html>
- Villa I Tatti. 2021. "DigiTatti." <https://dh.itatti.harvard.edu/#>
- Wartzman, Rick. 2014. "What Peter Drucker Knew about 2020." *Harvard Business Review* 16 October 2014. <https://hbr.org/2014/10/what-peter-drucker-knew-about-2020>