

History and computing



essay: History and computing , essay19601970? 1993: Historians have made use of computers in their research and teaching almost as long as computers have been in existence. The 1960s saw the first revolution in history and computing as historians harnessed the potential of the computer to analyse far more information than had previously been possible, provide greater precision to their findings, open up new avenues of research and enable verification and comparison of their research data. It was these advantages that led to computers becoming the handmaiden of the new, quantitative, social scientific and cliometric history that sprung up in the post-war years. Thus computing became indelibly, and often pejoratively, associated with quantitative history, an association that it has struggled to shake off ever since.

By the 1970s, the perhaps predictable backlash had begun, even among some of those who had been converts to the cliometric cause. The new history was attacked for turning historians into statisticians, slaves to quantitative analysis, narrowly focused and divorced from the people, places and events they purported to research. The debate was epitomised by the controversy over Fogel and Engermans 1974 two-volume quantitative analysis *Time on the Cross: the Economics of American Negro Slavery*.

(1) The authors were unfairly accused of endorsing slavery by suggesting that on economic grounds it was not an inefficient or uneconomic institution and that slaves had a better standard of living than many northern industrial workers. More substantive criticism, particularly by Herbert Gutman, could be made of the assumptions, sampling and mathematics used.(2) Even though later analysis supported Fogels and Engermans conclusions, *Time on*

the Cross provided all the ammunition needed for those seeking to discredit the new history as a computational tail wagging the historical dog. Like many revolutionary movements, the first wave may have inspired, challenged and threatened the mainstream majority in equal measure, but ultimately failed to persuade them. But the flame was not extinguished entirely. The early converts retreated into the newly formed departments of economic and social history, safe in the embrace of social science faculties rather than the hostile or indifferent environment of traditional historians in the faculty of arts. It should come as no surprise to discover then that the subsequent development of history and computing can appear as complex, contradictory and confusing as the past events that historians seek to explain. Nevertheless, the history of history and computing, its current status and future potential is revealing not only of historians various struggles with information technology but of their wider approach to the discipline.

In their 1993 publication *Computing for Historians: an Introductory Guide*, Evan Mawdsley and Thomas Munck entitle their first chapter *History and computing: the second revolution*.⁽³⁾ This second coming for history and computing in the mid 1990s was based on a number of technological advances: the advent of affordable microcomputers from the mid 1980s, improvements in networking and storage capacity and the development of generic, and relatively easy, programs for common applications. Historians no longer had to grapple, sometimes literally, with paper tape, punch cards or reels of magnetic tape. They did not need to become accomplished programmers or beholden to those who were. Historical sources did not need

to be reduced to a series of numeric codes, nor quantified to the nth degree for meaningful results to emerge.

There were no claims that computer methods would turn history upside down; in the UK at least, it sought to persuade rather than exhort, and emphasised the computer as a tool that could help the historian do what he or she has always done, but more efficiently and effectively. The computing historians weapon of choice was, and remains, the database, particularly in relational form. Long text fields enabled historians to retain something of the integrity of original sources and rapid advances were made on issues such as classification and record linkage. By this time historians could also include sophisticated text analysis and retrieval software, adopted from literary and linguistic studies, to analyse unstructured or semi-structured sources, qualitative analysis tools and sophisticated statistical packages in their armoury. Furthermore historians benefited as much as anyone from improved word-processing software, email, digitisation, CD-ROM content, Usenet groups, TELNET, networked library catalogues, bibliographic tools and the general improvement of computer speed, memory and storage.

At last history and computing seemed poised to break into the mainstream. During this period undergraduate options that included computing techniques emerged, as did Masters courses dedicated to history and computing, and instruction on database techniques could be found in doctoral training programmes. The international Association for History and Computing (AHC), founded in the mid 1980s, and its journal History and Computing went from strength to strength and a multitude of national AHC organisations sprang up under its aegis. A plethora of acronyms, CTICH, <https://assignbuster.com/history-and-computing/>

TLTP, ALT, AHDS, HCC [J1] (4), reflected a concerted effort by UK research councils to establish a national network of expertise to support computer use in research, teaching and data preservation.

A new sense of optimism and momentum emerged as historians adopted and adapted generic applications and tools from cognate disciplines to historical research and teaching. The mid 1990s were, however, the high water mark of the second revolution. Even at the time, proponents such as Speck were lamenting: And yet, and yet, while there is much to celebrate about the last decade, the fact remains that the profession is still divided between the small minority of historians who use computers as tools for analysing historical data and the vast majority who, while they might use a PC for word processing, remain unconvinced of the case that it can become a methodological asset.⁽⁵⁾ Indeed, Speck argued that aside from historical demography and psephology, computing had made little contribution to British history and there had been a retreat to narrative and high politics. The remaining years of the 20th century were to see little change in this position. By the early 2000s history and computing had made few inroads into traditional historical methods.

The AHC withered on the vine, many of its national associations became defunct as did its journal, after some prolonged death throes. While there was no backlash as experienced in the first revolution the vast majority of historical research remained untouched by computer methods. This stasis masked an underlying split in proponents of history and computing. On the one hand there were advocates of history and computing being recognised as a distinct discipline in its own right. Just as the life sciences have bio-

informatics and geography has geomatics, then history should have historical information science (or historical information processing or computational history).

Proponents of this view argued that historical sources are simply too ambiguous, relationships too complex and research questions needed to be formalised and generalised in a way that makes generic IT applications and traditional methodology unsuitable. In this view history and computing is not about the application of computer technology to solve specific historical problems but a way of modelling or representing the past that is simply made operational by information technology. On the other hand pragmatists argued that generic database, spreadsheet and statistical software could do 90 per cent of what historians demanded of them.

And other software, such as text retrieval and analysis packages, could be readily adopted from literary and linguistic computing. They countered that the real challenge lay in persuading the vast majority of historians of the benefit of even relatively simple information technology, not in developing specialist historical tools and methods that would only ever be of relevance to a minority of historians. The historical information science approach risked alienating mainstream historians even further and turning history and computing into a ghetto. The truth of course lies somewhere in between, or rather with aspects of both approaches. Historians in continental Europe, particularly The Netherlands, Germany and Russia, with their greater emphasis on theory and methodology, gravitated towards the historical information sciences approach. British historians, rooted in the empiricist tradition with their attachment to primary sources, and generally less

interested in theoretical models, methodology and process than outcomes, tended towards the pragmatic approach. Meanwhile, historians in the USA, stung by the fallout from Time on the Cross in the first revolution, tended to shun quantification in any shape or form, save for a handful of historically orientated economists.

These approaches, however, are not diametrically opposed to each other or mutually exclusive. They should be viewed as two points on a methodological continuum ranging from the computationally and methodologically light use of word processing, email and web searches, through the use of generic database and text analysis software that requires a modicum of computer experience and understanding of computer assisted research to the computationally intensive, technically demanding and methodologically innovative terrain of historical information processing. If the story thus far seems one of perpetually unfulfilled potential, history and computing, in whatever manifestation, could still point to some notable achievements. The sterile debate over quantitative versus qualitative has hopefully been put to rest. Historians of all persuasions make quantitative statements all the time, the distinction being that some express these statements in words and others in numbers. Whether analysing change over time or the relationship between cause and effect it is impossible to avoid talking about extent, range, scope, degree, duration, proportion or magnitude, whether one is using adverbs and adjectives or decimal points and chi-squares. Historians may still write in the narrative tradition but it is analytical narrative not descriptive narrative that has come to predominate. History and computing can not claim all of the credit for this development,

but debates about the role of quantitative analysis undoubtedly encouraged greater precision in history writing.

The historian today can also draw on decades of experience in accurately structuring relationships in databases, creating reliable record linkages and classification schemes that allow meaningful aggregation without blurring subtle distinctions. Nor does the employment of computer assisted research inevitably lead the historian to structured sources or statistical analysis.

Qualitative research tools are an equally adept, if under utilised, compatriot of relational databases, as are the text analysis tools of literary and linguistic computing that can provide revealing insights into historical texts with no, or minimal, encoding to corrupt the integrity of the primary source. In Britain we can also lay claim to over a decade of leadership in teaching history and computing, creating teaching resources, disseminating good practice and in the field of historical data creation, preservation and access.

Subjects as diverse as historical demography, psephology, prosopography, elite structures, entrepreneurial history, family history, urban history, political history, social history, economic history, transport history, medical history and education history have all benefited from the speed, precision, capacity and verification that the application of computer assistance brings. Furthermore, historians have successfully tackled many technical and methodological challenges; these include data modelling, record linkage, family reconstruction, multiple regression analysis, event history analysis, simulation, fuzzy data, content analysis and Geographic Information Systems (GIS). What then of the future “ can we envisage a third revolution in history and computing? Although crystal ball gazing is a hazardous occupation and <https://assignbuster.com/history-and-computing/>

the track record of history and computing should caution against making any kind of prediction, it might well be that one is already underway. There are two aspects to this third revolution. The first is the broader context of research funding and information providers in which historians conduct their research. The second is developments within the field of humanities computing.

Looking at both these aspects reveals tentative signs that developments are engaging historians in ways previous revolutions have failed to do. The shift in research funding towards strategic rather than responsive modes, and towards collaborative and interdisciplinary research, may create an environment that is more conducive to the use of computer assisted research. It is by no means certain that this will be the case, but addressing broader strategic issues through interdisciplinary collaboration may well expose historians to computer resources and applications they would not otherwise encounter. Of more immediate significance is the vast increase in digitisation activity by cultural heritage organisations that is putting a hitherto unimagined range of primary and secondary sources on historians desktops. It remains to be seen if historians make the leap from using a computer to find and retrieve this information to using it to analyse it. Nevertheless, the availability of such sources, for some historians at least, removes one of the biggest barriers to the uptake of history and computing “ that of creating machine readable material. One of the biggest factors in the cost/benefit analysis is the time it takes to transcribe source material in the first place, particularly for those who are uncertain about the benefits it would bring.

Moreover, these developments are engaging historians with librarians, archivists and curators in issues of resource creation and discovery that did not tend to take place in the print culture and providing a useful testbed of the application of e-science and Grid technology to the humanities.(6) If one turns to the broader field of humanities computing there are further signs of innovative work of relevance to the historical community. History and computing has always had a symbiotic relationship with humanities computing but there is evidence that the nature of this relationship is changing. No longer is it simply about adopting techniques from related disciplines, but as Boonstra, Breure and Doorn have suggested in *Past, Present and Future of Historical Information Science*,(7) a more collaborative relationship that includes those in the information and computer sciences. For example, the Armadillo project is a collaboration between the Department of History and the Department of Computing Science at the University of Sheffield. It has evaluated the benefits of semantic web technology for distributed historical materials involving natural language processing, data mining, knowledge management and ontologies, all aspects that have not hitherto been given much consideration by the historical community, although they have exercised those in humanities computing for some time. There is also evidence of renewed interest in the methods and tools being developed by linguists, particularly corpus linguists, in areas such as text mining, corpus annotation, automated tagging and historical thesauri. Despite having different aims for and outcomes of their research, significant common ground exists between linguists and historians in their concern for the provenance, reliability and authenticity of texts, the

challenges of variant spelling and meaning, and the opportunity that text mining tools provide to serve both communities.

In this field the work of Paul Rayson (University of Lancaster) and Dawn Archer (University of Central Lancashire) in corpus annotation and retrieval, Clare Llewellyn and Rob Sanderson at The National Centre for Text Mining (NaCTeM) in text mining and Christian Kay (University of Glasgow) in historical thesauri are worthy of particular mention. Aside from the well established text analysis tools such as TACT and Wordsmith, developments such as Wmatrix, The Historical Thesaurus of English, nora project and Monk project all offer exciting possibilities. The field of GIS has also excited historians and alliances of this with techniques such as visualisation and data mining present intriguing opportunities. # Computational Polichart Cartography for Visualization of Historical GIS Patterns and Processes#, by Prof. Claudio Cioffi-Revilla (Center for Social Complexity, George Mason University) and #Integrating GIS and Data Warehousing in a Web Environment: a Case Study of the US 1880 Census#, by Richard Healy (Dept. of Geography, University of Portsmouth) are just two papers at recent AHC-UK conferences (8) that demonstrate innovative and productive applications of new techniques to historical data. The work of Gidon Cohen (University of Durham) with multiple recapture methods also demonstrates historians continuing refinement of methods more commonly found in the natural and social sciences to the challenges posed by fuzzy historical data.

Meanwhile, John Bonnett (Brock University) provides a thought-provoking scenario for the future of history and computing in #Abductive Reasoning, A-Life, and the Historians Craft: One Scenario for the Future of History and Computing#. (9) Such developments all provide for a more positive outlook

<https://assignbuster.com/history-and-computing/>

than a decade ago, but if the third revolution in history and computing is to realise its potential several challenges remain. Research funding, infrastructure, education and training are all vital components, but above all historians need to acknowledge that their sources and methods are not so unique that they can not be abstracted and generalised to a certain level to make the application of computer technology worthwhile. Equally, those in the information and computer sciences need to recognise that the one size fits all mentality of generic applications and information systems is ill-suited to academic research and even less so within the heterogeneous field of humanities computing.,;? essay essay -X