A Discourse upon method, historical knowledge and the information technology
OR
Why all historians should know and understand what some of their fellows do with the computer including those who for reasons of age and aptitude never intend to use one.

Professor R J Morris
Department of Economic and Social History
William Robertson Building
George Square
Edinburgh EH8 9JY
Scotland
Rjmorriss@ed.ac.uk

c. rjmorris. This paper is for discussion only. Not to be quoted without permission. All comments welcome. The paper is not footnoted but a bibliography is provided at the end.

Statement: no historian can ignore impact of computing upon practice of history, even if no direct use or intention of use. Why?

1. In the last 30 years computing using historians have had a deeply subversive effect on many national and other histories.
   • US myth of a nation made by the railway [many movies]; careful analysis of economic data indicate that always made very little difference to economic output
   • US myth as land of opportunity; careful study of urban census data show that social mobility was very much the same as in Europe with the exception of unskilled which was better
   • UK the concept of an ‘industrial revolution’ as a sharp break in continuity destroyed
   • Family historians long ago destroyed the mythic past of large extended families with their secure marriages and granny by the fire. There are many more examples.

2. Why did computers have this effect?
they can handle large amounts of information
- they can analyze this information in a highly systematic manner
- at their best these machines are powerful pattern seeking tools.

This is a strength and weakness; computer users have had a major impact on the study of 18th and 19th century politics, especially on the importance of party BUT if look at historiography of say 1832 Reform of Parliament, then computer has moved attention away from speeches, diaries, letters and newspaper reports towards the voting records, especially the parliamentary poll books. Why because they have that regularity of structure which the machine loves. BUT this is a move that directs attention away from the intricacies of personality, alliance, interest and ideology and towards the behaviourist evidence of ‘the vote’ and associated party label.

Two techniques have dominated computer use by historians in last 30 years

**List Processing** e.g. census, poll book, rate books, poor law, hospital and asylum registers; the regular structure plus the fact that returns can only be gained by the systematic processing of large amounts of information makes them ideal for the machine.

**Record Linkage.** Usually this involves linking records referring to named individuals in order to create new forms of information Equally the records may refer to a building, an object, a place, an area, but at its most powerful the technique reconstructs individual histories in large numbers.

*Example.* Work done by Stana Nenadic on business in the 19th century [a golden age of enterprise for family business we have often been told]; linkage shows that they were predominantly small, fragile and short lived and highly dependent on the contingencies of family life for survival.

How do you begin to criticise the authority of the computer and the huge quantity of information behind these claims.

First. The machine is pattern seeking and has a hunger for specific types of document – hence privileging the information in those documents.

Second. The processing of such information can only proceed by adopting a series of rules. Some were applied directly by the machine controlled by appropriate algorithms. In other instances the historian sits patiently in front of the screen annotating a data base line by line.

Third. There are aspects of research design which can have hidden effects on results.

*Example.* Nominal record linkage itself can only proceed through a complex set of rules which often depend upon probabilistic judgements.
It is usually impossible to quantify such judgements. How do you know that Bob Morris is the same as Professor Morris or even R J Morris? There are three Professor R Morris in Edinburgh University and two Professor Bob Morris. Whilst my personal bibliographic record used to be attributed to three people R J Morris/Bob Morris/Professor Robert J Morris. This has since been reduced to two due to the sharp eyes of a Library of Congress MARC record cataloguer. The researcher might look at department and subject area to solve the problem and be happy BUT if the conclusions about Professors indicate that they tend to stay in the same department whilst at Edinburgh, how should such conclusions be evaluated?

Record linkage has a bias in favour of stability. This has been intensified by the tendency of research design to be based upon area sampling, – all middle class people from Leeds, the population of Philadelphia, etc. NOW we know that geographical mobility and social mobility tend to go together, so is that the results embodied in the USA studies may well be a result of using area based studies.

French Annales style studies showed enormous stability, but the recent linkage of the TRA sample the demographic records of the whole of France showed that the movers had much higher rates of mobility. The peasant who stayed became a peasant, whilst the one who moved became the craftsman, trader or labourer.

Record Linkage has a bias in favour of individuals about whom we have more and better quality information and in general this means higher status individuals – at a very simple level, higher status people have more complicated names.

3. What of the Future for historians? There are major gains to be made in text analysis. The application of authorship analysis techniques to the historical record will produce interesting results. The major barrier now is the historian’s imagination and the need for very extensive electronic databases of text. The historians use of text tends to be extensive rather than intensive as in most literary studies. How would political historians react if it was demonstrated that all the speeches in Hansard by Grey, Wellington, Brougham, Melbourne and the rest were written by the same person [not impossible given primitive state of shorthand in the 1820s and 30s].

Immediate Future. Historians need to respond to a world in which an increasing number of documents will be available in electronic form.

The problems posed include.
opportunities of analysis and access
problems of quality control
opportunities of criticism and verification.

Finally, the historian of the recent past is entering a world in which data was created in and preserved by way of electronic means – often to be accessed only by a dated technology. When I read Gladstone’s letters I touched the same bit of paper as Gladstone, got to know his handwriting and used the same technology as Gladstone. When Tony Wedgwood Benn deposited his diaries in the British Library, he handed them a set of disks, readable only by a now dated technology. For other historians, the relationship between document and data has been broken. Plan a 19<sup>th</sup> century railway journey and you look for the solid bulk of a Bradshaw much as did the clerks and passengers of the 19<sup>th</sup> century. Plan 21<sup>st</sup> century air travel and your document exists for fleeting moment on the travel agents screen, created from a data base of continually changing information.

Add to this the fact that the quantity of information created and preservable in vast electronic warehouses is vast. No historian will be able to read all those dispatches, e mails, drafts, etc. New strategies will be required to survive in such a world. One of these perhaps is that of ‘noise’ – something derived from intelligence analysis – it is impossible to organize anything violent and import like a war or a coup without an increase in telephone calls, e mails etc [a rule Bin Laden et al got around by using old fashioned foot soldiers].

Some of these strategies usable by any historians.

When did the Catholic question create most interest in 19<sup>th</sup> century Britain. Knowing that the RsLib project has just catalogued a wide range of pamphlets I logged onto the combined on line catalogue, asked for all titles which included the word ‘catholic’, sorted them by date and came up with the following, - lots of problems and need for crit of this result but it’s a start ...
Examples of the impact of computing on historical understanding are everywhere. For a recent survey see Terry Coppock (ed.), *Information Technology and Scholarship. Applications in the Humanities and Social Sciences*, British Academy 1999. See especially Jean Phillipe Genet on cultural history and R J Morris on the impact of computing on the historiography of the 1832 Parliamentary Reform Act.

Two examples

Stana Nenadic, Record Linkage and the small family firm, *Bulletin of the John Rylands Library*, vol.74, autumn 1992 [this issue included a wide variety of articles on computer based historical enquiry]

John A Phillips and Charles Wetherell, Parliamentary parties and Municipal Politics: 1835 and the party system, *Parliamentary History*, 1994 [again one of several articles based on computer methodologies]
A different sort of problem is touched upon by Edward Higgs (ed.), *History and Electronic Artefacts*, Oxford 1998. This looks at questions raised by the fact that a wide range of potential sources for historians are being and have been created in electronic form. The articles by Zeig [a diplomatic historian] and Higgs [ex PRO] are worth especial attention.

The journal *History and Computing* published by Edinburgh University Press covers a wide range of the issues debated here.


Coles, Tim, Alexander, Andrew and Shaw, Gareth (1997) Following the script: Optical Character Recognition Technology and the British Town and Trade Directory, *History and Computing*, vol.9, 1-16


Relevant web sites include

Institute of Historical Research, London
http://ihr.sas.ac.uk/ihr/bbs.ihr.html

The Association for History and Computing
http://odur.let.rug.nl/ahc/

Statistical Accounts of Scotland (both Old and New)
http://edina.ac.uk/StatAcc/

The Census of Norway from the year 1801, Jan Oldervoll, the
Department of History, University of Bergen.
http://www.uib.no/hi/1801page.html

Two major photographic archives for Scotland
George Washington Wilson, Aberdeen
http://www.visual-evidence.ac.uk/aberdeen/controller.

William Valentine of Dundee
http://specialcollections.st-and.ac.uk/

The Drawn Evidence. Scotland’s Development through its architectural
archives from industrialization to the millennium, 1780-2000
http://www.drawn-evidence.dundee.ac.uk/dundee_dr/index.jsp
[you will be asked to register but this can be done on line and should
only take a short time]