



Prescott, A. (2008) The Imaging of Historical Documents. In: Greengrass, M. and Hughes, L. (eds.) The Virtual Representation of the Past. Series: Digital research in the arts and humanities. Ashgate: Aldershot, pp. 7-22. ISBN 9780754672883

Copyright © 2008 The Editors

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

Content must not be changed in any way or reproduced in any format or medium without the formal permission of the copyright holder(s)

When referring to this work, full bibliographic details must be given

<http://eprints.gla.ac.uk/44847>

Deposited on: 06 February 2015

Enlighten – Research publications by members of the University of Glasgow_
<http://eprints.gla.ac.uk>

The Imaging of Historical Documents

Andrew Prescott

If we look back at what ‘history and computing’ has accomplished, the results are slightly disappointing. They are not disappointing because ‘computing’ failed to do what it intended to do, which was to provide ‘history’ with computerised tools and methods historians could use to expand the possibilities and to improve the quality of their research, but because ‘historians’ failed to acknowledge many of the tools “computing” had come up with.

Onna Boonstra, Leon Breure and Peter Doorn¹

Charles Darwin’s field notebooks, and particularly those from his epoch-making voyage as naturalist on *H.M.S. Beagle*, provide the fundamental evidence for understanding and tracing the development of his scientific thought. Darwin was insistent on the importance of keeping detailed records of his observations, declaring ‘Let the collector’s motto be, “Trust nothing to the memory;” for the memory becomes a fickle guardian when one interesting object is succeeded by another still more interesting’.² The importance attached by Darwin to his notebooks is evident from the meticulous way in which he kept them, using a sharp pencil to prevent water damage to the text and buying the finest quality notebooks. One of Darwin’s notebooks was described by its first editor, Nora Barlow, as follows:

The paper is excellent; and on the inside of the cover is a beautifully engraved little plate, surmounted by an engraved lion and unicorn – ‘Velvet Paper Memorandum Book, so prepared as effectually to secure the writing from erasure; with a Metallic Pencil the point of which is not liable to break. The point of the pencil should be kept

smoothly scraped flat and in writing it should be held in the manner of a common Pen'.³

Apart from their significance as artefacts, the physical makeup of Darwin's notebooks is an important clue to the chronology of Darwin's thought. The use of both pencil and ink, and the different types of ink, are vital evidence for the chronology of the entries in the Red Notebook kept by Darwin after he returned from his voyage on the *Beagle*.⁴

In 2006, with an enormous press fanfare, the complete works of Darwin, including new transcriptions of his notebooks, were made freely available online by the University of Cambridge.⁵ The BBC report on the new resource declared that 'The project run by Cambridge University has digitised some 50,000 pages of text and 40,000 images of original publications - all of it searchable. Surfers can even access downloadable audio files to use on MP3 players'.⁶ Darwin's great-great-great grandson was quoted as saying that '"It is astonishing to see the notebook that Darwin had in his pocket as he walked around the Galapagos - the scribbled notes that he took as he clambered over the lava'.⁷ However, users of Darwin Online wishing to share in this excitement will perhaps be disappointed to find that images of only one of the Galapagos notebooks are available. Moreover, these images are poor quality greyscale scans from microfilm, which convey little of the physical character of the original notebook – it is impossible, for example, to tell from the web image whether the notebook was written in pencil or ink. The directors of Darwin Online explain that the cost of providing high-quality colour scans of the manuscripts was prohibitive. However, this is a disingenuous statement - the cost of re-editing the notebooks would have been much greater than that of scanning them, but editorial costs were nevertheless found bearable and were considered by the directors of the project a

greater priority. In other words, the editors of Darwin Online felt that the provision of an edited text was more important than good quality images of the original notebooks.

A similar outlook is evident from other large collections of the writings of major scientists recently made available on-line. For example, a project based in Imperial College London is currently producing an on-line edition of the works of Sir Isaac Newton.⁸ Like the Darwin project, the Newton Archive seeks to assemble all the known writings of Newton, so as 'to grasp the organic unity of Newton's writing by garnering all his astonishingly diverse productions into a single, freely accessible electronic edition'.⁹ Again, like Darwin Online, the Newton Archive has given priority to the preparation of new transcripts. While the Newton project aspires eventually also to make available digital images of the manuscripts, at present only a very limited selection of images are available.¹⁰ By contrast, the Boyle Project based at Birkbeck College London offers a complete set of high quality colour images of the most important papers relating to Robert Boyle in the archives of the Royal Society.¹¹ Moreover, the related Boyle Work Diaries project at the Centre for Editing Lives and Letters at Queen Mary University of London has produced a complete edition of British Library, Additional MS. 4293 with a facing colour digital facsimile of the manuscript.¹² However, in both these cases, the web version of the images has unfortunately been scaled down in such a way that, while the text can be read in the manuscript, any closer examination of complex or difficult readings in the manuscript is impossible, even when the image is saved and viewed separately. This suggests that, for the editors of the Boyle Project, the images of the manuscripts are seen primarily as didactic tools. Their underlying assumption appears to be that researchers will chiefly be interested in the searchable edited text. This is apparently confirmed by

the way in which the search facility for the Boyle work diary links not to a view of the text and image but rather to the edited text only.

In short, the most common treatment of images of the manuscripts within electronic editions of historical papers has been as an ‘added extra’ rather than as an integral component of the research resource provided by the edition. Thus, while the Thomas Jefferson Digital Archive at the University of Virginia contains transcripts of over 1,500 letters by Jefferson, digital images are provided only for a small selection.¹³ The enormous electronic edition of the Proceedings of the Old Bailey London from 1674 to 1834 includes images of each of the contemporary printed reports of these trials, but again these are low quality black and white scans from microfilm in pdf format, and are very much subsidiary to the main edited text.¹⁴ Even archives providing commercial copies of images of historical documents tend to assume that users will be satisfied with greyscale scans from microfilm – as for example in the Documents Online service of National Archives, where the images of probate copies of wills are in greyscale pdf format apparently taken from microfilm.¹⁵ Collections of high quality colour images of historical records tend mostly to be provided not as part of on-line scholarly editions but rather as part of packages provided by libraries and archives chiefly for educational and general public use, such as the British Library’s Collect Britain website¹⁶ or the Staffordshire Past Track website of the Staffordshire Archives Service.¹⁷ Valuable though these sites are, material is rarely presented in such a form as to be useful for research purposes.

By contrast with historians, literary scholars preparing electronic editions have placed great emphasis on the provision of high quality colour digital images of key manuscripts. For example, Kevin Kiernan’s *Electronic Beowulf* includes large colour images of the early eleventh-century manuscript of *Beowulf*, together with an

enormous number of scans under special lighting conditions and images of early transcripts and collations made before the restoration of the original manuscript (figure one).¹⁸ An edition of *Beowulf* by Kiernan is included, but this is just one element of a complex digital archive and the formal edition is by no means the most prominent component. Likewise, the CD of the Hengwrt Manuscript of Chaucer's *Canterbury Tales* edited by Estelle Stubbs includes colour images of the manuscript which, when magnified, reveal details which cannot easily be seen with the naked eye.¹⁹ The various CDs produced by the Piers Plowman Electronic Archive also provide full-colour digital facsimiles of key manuscripts of Langland's poem.²⁰ The need for access to high quality digital facsimiles of literary manuscripts is not restricted to medieval texts. Emily Dickinson never formally published her poems, and for this reason the way in which she laid out her poetry in her notebooks is of great significance. The Emily Dickinson Archive therefore includes digital images of one of her notebooks and aspires to create a complete image archive of Dickinson manuscripts.²¹ Similarly, the Wilfred Owen archive includes many very high quality colour images of the poet's manuscripts, which vividly illustrate the process of composition (figure two).²²

In these early stages of the development of digitally-aided scholarship, this disparity of practice between historians and literary scholars is striking. It apparently reflects a profoundly different view as to how texts should be used and explored. The historians seem to be chiefly interested in making available large quantities of text and in quickly recovering particular pieces of information. Literary scholars are more concerned with the detailed exploration of the genesis of texts and in assembling all the available evidence for this. However, these approaches need not be mutually exclusive. Manfred Thaller, who for nearly thirty years has been the outstanding

visionary in historical information science and who has made some of the most penetrating analyses of the ways in which historians might use digital images, has pointed out that:

A change in the color of the ink a given person uses in an official correspondence of the 19th century could be an indication of the original supply of ink having dried up; or of a considerable rise of the author within the bureaucratic ranks. Let us just emphasize for non-historians, that the second example is all but artificial: indeed the different colors of comments to drafts for diplomatic documents are in the 19th century quite often the only identifying mark of which diplomatic agent added which opinion.²³

Likewise, J. Burt and T. B. James have described how the appearance and alteration of entries in census enumerator's book can provide important historical information:

The refusal of a young woman to reply to the question on occupation in the 1881 census for Winchester, coupled with the comments of the enumerator, whose definition of her occupation as 'on the town' (implying prostitution) provides an important glimpse behind the curtain of the surviving sources – the enumerators' book, and towards an understanding of the process through which the original census schedules (which have not survived) were transformed into the documents we have today. Conversely, a two-line entry in the census which reads 'Assistant Classical Master /BA Trinity College Dublin' which is reduced by the editorial pen to 'Prof' helps the researcher to grasp some of the smoothing out process of categorisation which went to contribute to census statistics overall.²⁴

Despite these salutary examples, historians in constructing electronic editions remain reluctant to give priority to the provision of images of the original documents.

The recent CD ROM of the medieval Parliament Rolls, for example, surprisingly only includes sample images of the records, despite the fact that scribal practice in compiling the original document was considered sufficiently important for the details to be carefully noted in the edition itself.²⁵ A contrasting approach is that of the Calendar of Fine Rolls project at Kings College London, where digital images of the original record are provided and the only searchable component is a summary English calendar.²⁶ The difference between these two approaches suggests that there is still considerable uncertainty as to how digital images can most effectively be used in editions of historical documents.

The cost of making available the large quantities of information demanded by historians has always been an important factor in determining the form of publication of historical records. In his guide to *Editing Records for Publication* published in 1977,²⁷ before electronic techniques were widely available, Roy Hunnisett, one of the most distinguished British records scholars, argued that in view of the great cost of printing full transcripts of modern historical documents, most post-1200 records should be only be published in calendar form unless there were exceptional reasons otherwise. Hunnisett's advice reflected the limitations of print technology and is arguably no longer valid in a digital environment. However, the cost of providing digital images of full runs of administrative records is also likewise often cited as a reason for preferring edited texts. Manfred Thaller has argued that this is a false position and in his Duderstadt project demonstrated how, by very precise control of production techniques, it was possible to digitise the entire archive of a city before 1600 at a very low unit cost.²⁸ The Duderstadt electronic archive comprises over 80,000 images (figure three). These are mostly greyscale, but Thaller has gone on to show how similar methods can be used cost effectively to produce colour images of

decorated medieval manuscripts in the *Codices Electronici Ecclesiae Colonensis*, which already contains complete digital facsimiles of over 300 medieval manuscripts and when finished will present the entire manuscript library of a medieval cathedral.²⁹

In 2001, Thaller commented on the issues of cost that:

The creation of digital collections does not *have* to be particularly expensive anymore. One of the more spectacular technical developments in recent years has been the drop in the pricing of digital cameras, where the resolution achievable by a \$1,000 camera has been climbing sharply. At the other end, cameras like the 4096 x 4096 pixel camera offered by Kodak, with an observed workflow of ca. 5-10 seconds per exposure, are today still in the six-digit price range. With an emerging mass market of digital hobby photographers, it seems to be a safe bet that high speed digital cameras at a professional resolution will become achievable for routine projects in less than ten, presumably within the next five years. This means that with 2,000 exposures per campaign day - being a serious barrier for quite some time - 1,000,000 page digitization projects will be possible with a limited budget and over a two-year or 500-day time frame.³⁰

Thaller's predictions as to the cost of digital cameras proved very accurate, but the large-scale digitisation projects he anticipated are still not very commonplace. This suggests that the barriers to the more widespread use of high quality digital images in historical editions are not simply those of cost, copyright or the issues associated with access to and handling of original materials. The difficulties are more profound, and are perhaps connected with the way in which historians make use of documentary materials. This raises some of the issues which were recently identified as problematic in 'historical information science' by Onna Boonstra, Leon Breure and Peter Doorn.³¹

Boonstra, Breure and Doorn describe a situation in which digital technologies have had a limited impact on historical practice, because historians see their primary function as identifying, analysing and debating changes in society over periods of time. In this context, the exploration of archival materials is seen to a large extent as an activity which is ancillary and preparatory to the true business of the historian, namely debating the nature of past societies. Since much work on the use of digital technologies in history is concerned with representing and improving access to source materials, historical computing is consequently also seen as an ancillary skill. Charles Harvey has argued that a form of historical computing which focuses on how textual sources were represented in electronic form would always be marginal to the main historical profession.³² In Harvey's view, true historical computing should address wider questions in the methodology of the reconstruction of societies in time, a vision which is only just beginning to become a realistic prospect. Harvey's underlying concern, that historical computing as it has been most commonly defined in the past thirty years risked becoming marginalised from mainstream historical research, seems to have been largely realised.

The way in which electronic resources for the study of history have been developed and used suggests that historians are primarily interested in rapidly assembling large quantities of factual (and increasingly statistical) information to test the validity of preconceptions of the nature of past societies. The way in which these historical texts operate and were put together is of concern only insofar as it challenges these preconceptions. This concern rapidly to assemble information doubtless explains the preoccupation of historians with databases and their more limited interest in automated text handling, although, as other essays in this volume describe, this is shifting as the use of XML becomes more widespread. Too often,

perhaps, historians view their textual sources in a positivistic fashion, as quarries of raw factual material.

Such a view of historical research can be seen as representing a paradigm particularly appropriate to the study of more modern periods of history. It overlooks the fact that there is a long and distinguished tradition of historical research in Britain which has been less concerned with the reconstruction of past societies and more interested in how we know about the past and in exploring the limitations, failures and deceptions of the documentary sources on which historians rely. This is a tradition closely associated with medieval studies, and is represented by such distinguished records scholars as V. H. Galbraith, Charles Johnson and Roy Hunnisett himself. The distinction can perhaps be seen by comparing the work of Barbara Hanawalt and Hunnisett on medieval coroners' records. Hanawalt uses information in coroners' inquests to undertake direct reconstructions of many aspects of everyday life in the medieval period.³³ By contrast, Hunnisett exposes the complex processes by which the information in returns by coroners' juries was assembled and emphasises that much of the detail in it is fictitious. Hunnisett engages profoundly with the textuality of the archive and ultimately leaves one uncertain how far archival materials can be used with any confidence to reconstruct the past.³⁴ For a scholar such as Hunnisett, the interest is more in how the archive operates and less in pursuing the chimera of past societies.

The way in which many historical records are presented in electronic form, with the assumption that geographical and statistical presentations can readily be generated from the electronic resource, suggest that many historians have still not broken away from the complacent assumptions of their Victorian forebears that administrative records are somehow more inherently trustworthy than other texts, no

matter how many warnings we have recently been given by ‘dodgy dossiers’ and fabricated police evidence. Yet for a historian in the tradition of Galbraith or Hunnisett, the way in which a ‘dodgy dossier’ might be compiled and the influence it might have are fascinating issues, connecting to themes which reach back through the archives to Domesday Book and beyond. As Galbraith, whose work epitomises this tradition puts it, ‘administrative records, no less than the literary sources, are generally compiled from other documents, often unknown to us, and they rarely tell the whole truth’.³⁵ In Galbraith’s view, administrative documents are as much artificial constructs as literary texts and are as prone to the same inherent deceptions of all textual communication. Galbraith’s vision of a history in which our concern is not to pursue the reconstruction of the past but rather to engage in an open-ended dialogue with the textuality of the past is inherently a relativistic one:

The past itself is dead, and the books we write tombs of learning, except insofar as they live in the consciousness of their readers. So conceived, we travel pleasantly, but by the nature of things we never arrive.³⁶

In a world where knowledge is becoming increasingly google-ised, there is a serious risk that information simply becomes a commodity which is consumed and that many of the critical skills in handling information which underpin much of humanities scholarship will vanish. It would be ironical if historians, by constructing electronic resources in which textual sources are treated as relatively simple in structure, actually hastened this process of google-isation. One way in which the complexity of texts can be effectively conveyed in electronic editions is through the inclusion wherever possible of digital images of the records, scanned in colour at a sufficiently high resolution to allow issues connected with the genesis of the text to be

explored by the user. Galbraith dreamed of a new type of history, which would be an archivist's history:

To him [the archivist], the past presents itself as a vast collection of 'original documents' ... To name a century to him is to call up a mental picture of the relevant records, the progress of history appearing to him as a slow pageant of slowly changing records, marked from time to time by the occasional disappearance of one class and the gradual emergence of another.³⁷

This is a vision of a history which is less concerned with the incestuous and self-serving study of historiography, is far more focused on text and closer in its intellectual approach to the study of literature. To quote Galbraith once more:

History is once more consciously, almost self-consciously, allying itself with literature ... What matters in the long run is not so much what we write about history now, or what others have written, as the original sources themselves. They are an inexhaustible and an invaluable inheritance.³⁸

Digital editions which incorporate digital images of historical records have the potential to create something akin to Galbraith's vision of an archivist's history. For example, images of documents can convey a sense of the process by which the information in the record was assembled in a way that it is impossible to do in a transcribed edition. As Galbraith emphasised, historical documents of any type are complex montages of different types of information, with many different layers:

Domesday Book, for example, is no original document, but a condensed summary of a complicated series of semi-judicial proceedings, and, moreover, full of errors. Most records – from Acts of Parliament to balance sheets of public companies – have some sort of bias of their own, and seek to conceal the truth or part of it.³⁹

Disentangling these levels of information within the record is a difficult process and one in which visual clues are often very important. For this reason, digital images are an indispensable tool in investigating further the genesis and structure of the record. I would like to illustrate this fundamental point further by giving some extended examples from the records of the judicial proceedings arising from the Peasants' Revolt in England of 1381.

Following the slaying of the rebel leader Wat Tyler at Smithfield and the collapse of the rebellion which had overwhelmed much of southern England in June 1381, a series of commissions were issued to 'chastise and punish' the insurgents.⁴⁰ These form part of a series of legal records relating to the revolt which record the names of thousands of insurgents and provide details of local disturbances throughout the affected areas. On the face of it, the Peasants' Revolt is one of the best documented events of its kind in the middle ages. But the legal proceedings against the rebels were extensively affected by false litigation, as claimants took advantage of the disturbed situation and the extra-judicial proceedings against the rebels to settle a variety of old scores. Men described in indictments as leaders of the revolt submitted petitions to the crown protesting their innocence and insisting that the allegations against them were brought by their enemies, claims which were accepted by the crown.⁴¹ In such a context, it is vital to know how the accusations against individuals in the judicial proceedings were pieced together.

Much of our information about local incidents in Kent during the revolt comes from the records of a commission comprising the Earl of Kent, John Middleton and Thomas Trevet which heard cases against participants in the rising at sessions in the county during July 1381. Some of the accusations collected by this commission are preserved in two files in the National Archives, one containing indictments relating to

East Kent⁴² and the other covering West Kent.⁴³ Digital images of the West Kent file reveal aspects of the commission's procedure which are not apparent from the printed edition of this record published by Powell and Trevelyan in 1899. The way in which this material is presented in Powell and Trevelyan's edition might lead the unwary user to suppose that the accusations were fairly straightforward reports of information presented by local juries. The records of the commission's work were kept by at least two clerks, one of whom apparently had a supervisory role. These clerks shaped and manipulated the information presented in the indictments, apparently undertaking, presumably with the commissioners, detailed investigations into what had gone on in Kent during the revolt.

The active character of the inquiries undertaken by the commission is apparent from odd working notes preserved in the records, such as a small scrap of vellum on which one of the clerks has noted that 'Robertus Man cognovit quod cepit episcopum per totam Cantuariam apud la Tour in Schep' (figure four).⁴⁴ This is apparently an *aide memoire* that Robert Man, who confessed to taking part in the seizure of the Archbishop of Canterbury at the Tower of London, was in Sheppey. The words 'totam Cantuariam' are obscured by an accidental inkblot but, as Powell pointed out, they were apparently struck through by the clerk, presumably because this was not the correct version of the Archbishop's title. The interlineation of the word 'la' also suggests that the clerk was struggling to get a correct verbal form for the note of the offence, presumably with a view to eventually working up the final indictment. The informal character of this note is confirmed by the dorse on which is a reminder to investigate John Gylot of Dartford and another incomplete note (figure five).⁴⁵ This scrap of vellum was apparently one of many such rough notes made in the course of the work of the commission as it sought out rebels in Kent.

Other evidence of the process by which the information in this file was pieced together is a list headed ‘Tenentes abbatissae de Mallyng’ (figure six).⁴⁶ The list comprises more than thirty names. The majority have been struck through. The men whose names were not deleted were afterwards indicted by a presenting jury from the Hundred of Larkfield for coercing the Abbess into giving them a charter releasing them from labour services. Moreover, the last five names in the list were added separately by the supervising clerk. This suggests that the tenants of the Abbess of Malling were systematically interrogated by the commissioners or their clerks as to which of them had been involved in the revolt. The addition of the last five names indicates that there may have been more than one such examination. The preparation of this list of names was evidently only a first stage in assembling information about the insurgents. The indictment from Larkfield Hundred bears clear traces that information was still being gathered while the indictment was being compiled (figure seven).⁴⁷ For example, details of goods seized by insurgents were inserted as an interlineation, presumably as additional information came to hand. The names of some of the tenants accused of attacking the Abbess of Malling were also inserted later, presumably as a result of the second interrogation of the tenants. The final indictment against John Leg of Birling also seems only to have been squeezed in at a late stage.

The information in these indictments was compiled as the result of a complex iterative process in which the commissioners and their clerks took an active role. Much of the evidence for this process, and thus the status of the information in the indictments, depends on visual clues which are difficult adequately to convey through a conventional edited transcript. Further illustrations of this process can be found in another set of proceedings against the insurgents, this time from East Anglia.

The discovery and publication by André Reville and Edgar Powell of some records of proceedings against rebels in Norfolk and Suffolk by the commission of William Ufford, Earl of Suffolk and others helped make historians more aware of the great extent and violence of the disturbances in East Anglia in 1381.⁴⁸ However, the records of Ufford's commission are complex in structure. The earl died in 1382 and after his death a writ dated 13 May 1382 was sent to his executors summoning to chancery all records in their possession relating to the commission against the rebels.⁴⁹ The reason why the records were summoned to Westminster is not clear. Possibly it was related to process arising from pardons being granted to the rebels. It may also have reflected concern that the powerful commissions issued after the revolt were getting out of control. Whatever the reason, it is only thanks to the issue of this writ that this roll and file, the most detailed source of information about the rising in Norfolk and Suffolk, survives. Without it, the records would have vanished completely, as happened with the records of other commissions relating to the revolt in, say, the north of England. The commission to Ufford and the others was not recorded on the patent rolls, but is described in the writ requesting the records as a commission to 'chastise and punish' the rebels. We know from other surviving commissions that those authorised to chastise and punish the rebels had very sweeping powers. They could take proceedings against the rebels both according to the law of England and following their discretion. This could, and did in places such as Essex, include military action against the insurgents. Arguably, Ufford and his fellows were not obliged by the terms of their commission to keep records, but they evidently did keep very careful ones.

On the back of the writ requesting the commission records is a return made by a clerk on behalf of Sir Roger de Boys saying that he was sending all the documents

relating to the commission, but that those relating to the liberty of St Edmund were held by Thomas Morreaux, another member of the commission. However, the clerk helpfully sent to chancery a list of those named in the indictments held by Morreaux.⁵⁰ The question arises how the clerk had access to this information. Did he check through the documents held by Morreaux, or did he have his own set of notes? In this list, the clerk pays particular attention to those who participated in one of the most notorious incidents in the revolt, the killing of John Cavendish, the Chief Justice, giving brief summaries of the allegations against each of these individuals (figure eight). Those involved in Cavendish's death were excluded from the general pardon, which is presumably why this extra information was provided. Again the question arises of where the clerk compiling this list obtained this information. The corrections in the list suggest that the clerk was indeed working from some other form of documentation. For example, he noted that Katherine Gamen released a boat so that Cavendish could not escape death. He subsequently added the words 'from the land' after boat, as if he was uncertain whether his summary was clear and wished to make it more emphatic. The list of those named in the indictments held by Morreaux emphasises again the point that records are often compilations of other records, now lost. Moreover, this list is mysterious because it could not have formed the basis for formal proceedings in a court like King's Bench.

As in West Kent, the indictments heard by Ufford's commission bear clear traces of their origin in an iterative process of interrogation of the presenting jurors. This can be seen, for example, in an indictment by the jurors of Hartismere Hundred in Suffolk (figure nine).⁵¹ The first item in this indictment states that 'The jurors of Hartismere say that James of Bedingfield was chief leader of a company of the commons which robbed Edmund of Lakingheath of his goods and chattels worth 40s.

at Stoke, which goods were afterwards restored to Edmund.' As with the Kentish indictment, at first sight this seems to be highly circumstantial historical evidence, made by local men within a month of the rising, but its words are a long way from any ever uttered by any jurors. The forms used in an indictment followed strict legal formulae. For example, the value of the goods stolen was critical. Thefts were only punishable by death if the goods stolen were worth more than twelve pence. In the case of this indictment, an image shows clearly that the value of the goods stolen was only established at a late stage in the interrogation of the jury, and the value (40s) has been inserted by the clerk in a gap specially left for the purpose.

Nor were the juries the 'twelve good men and true' of popular imagination. Attached to each indictment is a *panellum* listing potential jurors. On the panel, one of the clerks associated with the commission has carefully selected who would be a suitable juror, marking those who were sworn and striking through those who were apparently not suitable. In an empanelment of a jury from Mutford hundred in Suffolk, one juror has been sworn and then his name struck through by the clerk who swore in the jurors (figure ten).⁵² Another name has been added late to the list of jurors, but nevertheless sworn, while a third was sworn, then removed from the jury and finally sworn again. In the case of Ufford's commission, it seems that particular care was taken to ensure that the jurors contained the best representatives of local society, by making them swear allegiance and even give a recognisance of forty pounds each for the faithful performance of their duties.⁵³

The way in which the clerks attached to the commission were the key figures in determining the wording and final form of the indictments is particularly apparent from the way in which the issue of treason was treated. Curiously, it was not clear from existing legislation that popular rebellion of this kind was indeed treason, and in

the indictments various textual strategies were used to try and establish the treasonable status of the rebels.⁵⁴ This might include the highlighting of certain actions (such as the use of banners) or the use of phrases such as 'in a warlike manner and against the dignity of the crown of the lord king'. Particular clerks can be seen preferring particular phrases. In another example, uncertainty as to the treasonable status of the offence seems to have continued until the very last moment, and the word 'produciose' (treasonably) has been inserted by the clerk in a gap (figure eleven).⁵⁵ Evidence of the way in which the indictments were the result of an iterative process, in which judge or clerks interrogated the presenting jury can be found throughout this set of documents. In another example, which describes the murder of a justice of the peace near Ipswich during the revolt the date in the indictment has been altered by another scribe.⁵⁶ When was this done and why? The answer is not immediately obvious, but clearly it has considerable importance in appraising the evidential value of this indictment. Moreover, the answer to this question will affect our view of the entire group of documents, since the correction is made to an indictment in the hand of the scribe responsible for most of the Suffolk indictments. All these features of these documents are not immediately evident from an edited transcript but are given immediate prominence by images.

One of the most widely discussed indictments relating to the revolt in East Anglia is this one which alleges that Thomas Sampson of Ipswich led 'the great society of Suffolk' (figure twelve).⁵⁷ There has been extensive discussion as to whether this can be taken as an indication that there was a central organizing body controlling the insurgents. In all this discussion there has not been as far as I am aware any mention of the fact that this indictment again contains later amendments by another clerk. It looks as if this was done because the indictment was felt to be

deficient in form – it lacked for example a verb and another clerk has duly added ‘rose’ at an appropriate point. The form of date given is also deficient. This tends to suggest that the clerk responsible for the original indictment, whose hand does not appear very much in the commission records, was struggling with the correct form of the indictment, which may perhaps explain why he plumped for the unusual reference to ‘the great society’. If images of the indictment against Sampson had been available to historians, perhaps their lengthy discussion of whether or not ‘the great society’ actually existed would have been rendered otiose.

It is by no means surprising to find that the information about the revolt was gradually pieced together in this way; more surprising is that historians using these records have been so willing to accept this information at face value and have shown little interest in the sources of the information on which they rely. The gradual assembly by administrators of the complex layers of information in historical documents is a feature of many records from the time of Domesday Book onwards. To cite another example from the 1380s, Jan Gerchow has shown in a brilliant study how the various return of the regulations of local guilds made in 1389 are a mixture of oral declarations made in Chancery at Westminster, returns written by officials of the guilds themselves and returns compiled by local writing offices, often using the same template for different guilds.⁵⁸ In attempting to differentiate these various records, images are essential for comparing the scribal practice in the hundreds of surviving returns.

Close examination of records such as these brings home forcefully how we still have very little information about how the documents on which so much of our history depends were put together. In better understanding the evidential character and textuality of our ‘primary sources’, digital images can make a fundamental

contribution and should be regarded as an indispensable component of future online editions. This implies the development of an approach to electronic editing which puts the digital image more at the heart of the edition and gives less priority to the database or transcribed text. Here much can be learnt from literary scholars. A fascinating vision of an image-based edition is provided by the Edition Production and Presentation Technology which is being developed by Kevin Kiernan and Emil Iacob.⁵⁹ This suite of generic editing tools was developed from concepts first worked out in editing very badly damaged Old English manuscripts but its use with any documentary materials is currently being demonstrated in trials with many independent projects.⁶⁰ EPPT comprises a range of innovative editing tools, but here I will just briefly review features which seem to me relevant to the particular issue I have discussed in relation to the Peasants' Revolt records. but its potential for dealing with a wide range of other materials is currently being demonstrated in a range of other projects. EPPT comprises a range of innovative editing tools, but here I will just briefly review features which seem to me relevant to the particular issue I have discussed in relation to the Peasants' Revolt records.

First, the relationship between the clerks who produced the various records in 1381 is mysterious, and we can only start to understand it if we can compare and manipulate images of the large number of membranes which constitute an individual record. In particular, we need to be able to link the transcribed texts, duly tagged to note changes in hand, directly to images of those hands. This is a concept at the heart of EPPT, in which the editor can map the position of particular portions of text on the image and automatically encode the coordinates in the XML underlying the text. In principle, this will enable all instances of the appearance of a particular hand to be drawn together and relevant patterns identified (figure thirteen). Were different

scribes used for particular sessions of the commission or in different parts of the country? Was there an overall supervising scribe? How far did say towns provide their own scribes? These are all questions fundamental to our understanding of the evidential value of these documents, and ones which we can start to answer through an image edition of the sort that EPPT facilitates.

In building image-based editions, we quickly reach the limitations of our current tools for the tagging of texts. In particular, where it is necessary to indicate not only the linguistic character of the text but also to mark up information about the physical medium on which the text is preserved, complex multiple hierarchies of mark up are required. The editing tools within EPPT allow such complex XML hierarchies to be built up for individual documents using intuitive tools, all the time tied closely to images of the documents themselves. The answer to some of the questions I have raised may be very simple – but the answers can only be found if our digital edition incorporates all the relevant information about the structure and character of the manuscript and the procedures of the scribes who created them.⁶¹

In comparing different hands, it is clearly unsatisfactory to rely on the judgement of a trained (or otherwise) eye. in the absence of any illustrative evidence. The DucType tool in EPPT allows the editor to compare, define, encode and display the characteristics of individual letter forms and thus allows the scribes to be more precisely distinguished. The collection of such detailed, searchable information facilitates the identification of scribal (and thus administrative) procedures in the preparation of records such as the one we have been considering. Existing collation tools such as Collate tend to focus on individual variants, rather than identifying the scribal manipulation and transmission of larger blocks of text. Yet, in examining scribal practice in administrative documents, it is the way in which larger textual units

move which is significant. We have noted how different clerks in 1381 struggled to find a form of indictment which established that insurgents were guilty of treason. An EPPT collation tool called SaMod (Old English "together"), under development for textual analysis, automates the detailed comparison of blocks of text as well as of individual words. The tool should therefore facilitate a detailed examination of the way in which legal forms were manipulated and enhanced by different clerks in 1381.

It may be argued that these techniques are chiefly relevant to medieval records and have limited application to more modern materials, but a digital image of the examination of a leader in another uprising nearly five hundred years later, the Chartist rising at Newport in 1839, shows that the same issues that we saw in examining the records of the 1381 revolt are just as relevant for historians of Chartism (figure fourteen). This image is taken from the examination of Morgan James of Pillgwenlly held by the Newport Library and Information Service and available on the Gathering the Jewels website. As in 1381, the successive annotations reflect an iterative process of interrogation, with some key components of the examination being altered in pencil.

The examination of Morgan James perhaps could be taken as an emblem of the way in which we are increasingly beginning to appreciate that administrative records represent artificial textual constructions and are in a way literary genres. In examining the way in which these literary productions were put together, the tools developed to investigate literary texts such as EPPT are invaluable. Whether it is a chartist trial record, *Beowulf* or an indictment against an insurgent in 1381, we cannot separate the text from the medium in which it is preserved and, to secure a full understanding of that text, a high resolution digital image is indispensable. Digital images give us a new awareness of the physical character of the historical records

which should be at the heart of our historical understanding. Image-based editions of historical documents offer the potential to move towards that archivists' history of which Galbraith dreamed.

¹ Onna Boonstra, Leon Breure and Peter Doorn, *Past, Present and Future of Historical Information Science* (Amsterdam: Netherlands Institute for Scientific Information, Royal Netherlands Academy of Arts and Sciences, 2004), p. 9.

² Charles Darwin, *Narrative of the surveying voyages of His Majesty's Ships Adventure and Beagle between the years 1826 and 1836...* (London: Henry Colburn, 1839), p. 598.

³ Cited in the description of Darwin's Beagle Field Notebooks at: http://darwin-online.org.uk/EditorialIntroductions/Chancellor_fieldNotebooks.html (all cited URLs were accessed in February 2007 unless otherwise stated)

⁴ Discussed in S. Herbert, 'The Red Notebook of Charles Darwin', *Bulletin of the British Museum (Natural History) Historical Series* 7:24 (1980), pp. 5-19.

⁵ <http://darwin-online.org.uk/>

⁶ <http://news.bbc.co.uk/1/hi/sci/tech/6064364.stm>

⁷ Ibid.

⁸ <http://www.newtonproject.ic.ac.uk/>

⁹ <http://www.newtonproject.ic.ac.uk/prism.php?id=26>

¹⁰ At the time of writing (February 2007), the only images of Newton manuscripts unique to the Newton Archive are of the following manuscripts: ASC MS. N47 HER in the James White Library, Andrews University, Berrien Springs, Michigan, and Kings College Cambridge, Keynes MSS. 3, 6, 7, 9, 10 and 11 (all accessible from <http://www.newtonproject.ic.ac.uk/prism.php?id=44>). The site also holds images of Keynes's correspondence about the purchase of his Newton manuscripts: <http://www.newtonproject.ic.ac.uk/prism.php?id=19>. There are links to images of Newton manuscripts in the related *Chymistry of Issac Newton* project: <http://webapp1.dlib.indiana.edu/newton/index.jsp>. It should be noted that images on the *Chymistry of Isaac Newton* website are also greyscale scans, apparently from microfilm. Links are also provided to images of material in the Burndy Library, but at the time of writing these are temporarily unavailable. I am grateful to Dr John Young of the Newton project for providing information on the extent of manuscript images within the archive.

¹¹ <http://www.bbk.ac.uk/boyle/>

¹² <http://www.livesandletters.ac.uk/wd/>

¹³ <http://etext.virginia.edu/jefferson/>

¹⁴ <http://www.oldbaileyonline.org/>

¹⁵ <http://www.nationalarchives.gov.uk/documentsonline/> Other material provided by this service, such as the indexes of campaign medals in World Wars One and Two, is also made available in only in greyscale scan from microfilm.

¹⁶ <http://www.collectbritain.co.uk/>

- ¹⁷ <http://www.staffspasttrack.org.uk/>
- ¹⁸ <http://www.uky.edu/~kiernan/eBeowulf/guide.htm>.
- ¹⁹ http://www.llgc.org.uk/drych/drych_s007.htm.
- ²⁰ <http://jefferson.village.virginia.edu/seenet/piers/>
- ²¹ <http://www.emilydickinson.org/>
- ²² <http://www.hcu.ox.ac.uk/jtap/>
- ²³ Manfred Thaller, 'Text as a Data Type', paper at ALLC-ACH 1996, available at: <http://gandalf.aksis.uib.no/allc-ach96/Panels/Thaller/thaller2.html> .
- ²⁴ J. Burt and T. B. James, 'Source-Oriented Data Processing: The Triumph of the Micro over the Macro?', *History and Computing* 8:3 (1996), pp. 160-8.
- ²⁵ <http://www.sd-editions.com/PROME/>
- ²⁶ <http://www.finerollshenry3.org.uk/home.html>.
- ²⁷ R. F. Hunnisett, *Editing Records for Publication*, Archive and the User, 4 (London: British Records Association, 1977)
- ²⁸ <http://www.archive.geschichte.mpg.de/duderstadt/dud-e.htm>.
- ²⁹ <http://www.ceec.uni-koeln.de/>
- ³⁰ M. Thaller, 'From Digitized to the Digital Library', *D-Lib Magazine* 7:2 (February 2001). Available at: <http://www.dlib.org/dlib/february01/thaller/02thaller.html>.
- ³¹ Boonstra, Breure and Doorn, *op. cit.*
- ³² See, for example, Charles Harvey and John Press, *Databases in Historical Research: Theory, Methods and Application* (London: Macmillan, 1996), particularly Chapter 1.
- ³³ B. Hanawalt, 'Violent Death in Fourteenth- and Early Fifteenth-Century England', *Comparative Studies in Society and History* 18 (1976), pp. 297-320, or, more recently, 'The Voices and Audiences of Social History Records', *Social Science History* (1991), pp. 159-75.
- ³⁴ Apart from Hunnisett's indispensable monograph, *The Medieval Coroner* (Cambridge: Cambridge University Press, 1961), see also 'The Reliability of Inquisitions as Historical Evidence' in *The Study of Medieval Records: Essays in Honour of Kathleen Major*, ed. D. A. Bullough and R. L. Storey (Oxford: Oxford University Press, 1971), pp. 206-35.
- ³⁵ V. H. Galbraith, *Introduction to the Study of History* (London: C. A. Watts, 1964), p. 13.
- ³⁶ *The Historian's Workshop*, ed. L. P. Curtius (New York: Knopf, 1970), p. 7; reprinted in *Kings and Chroniclers: Essays in English Medieval History* (London: Hambledon Press, 1982)
- ³⁷ *Studies in the Public Records* (London: Thomas Nelson, 1948), p. 7.
- ³⁸ *Introduction to the Study of History*, pp. 75, 79-80.
- ³⁹ *Ibid.*, p. 13.

⁴⁰ For general background on the following, see A. J. Prescott, *The Judicial Records of the Rising of 1381*, unpublished Ph. D. thesis, University of London, 1984; 'Writing about Rebellion: Using the Records of the Peasants' Revolt of 1381', *History Workshop Journal*, 45 (1998), pp. 1-27; and "'The Hand of God'": the Suppression of the Peasants' Revolt of 1381', in *Prophecy, Apocalypse and the Day of Doom*, ed. N. Morgan, Harlaxton Medieval Studies 12 (Donington: Shaun Tyas, 2004), pp. 317-41.

⁴¹ For example, National Archives, SC 8/262/13099, a petition by John Creek al. Bettes of Wymondham, protesting that allegations of insurgency against him were false. On the accusations against Creek, see further Prescott, 'The Hand of God', n. 6.

⁴² The National Archives, JUST 1/400, printed in translation by W. Flaherty, 'The Great Rebellion in Kent of 1381 illustrated from the Public Records', *Archaeologia Cantiana*, 3 (1890), pp. 71-96.

⁴³ The National Archives, KB 9/43, printed in an abridged version in E. Powell and G. M. Trevelyan, *The Peasants' Rising and the Lollards* (Cambridge: Cambridge University Press, 1899), pp. 3-12.

⁴⁴ The National Archives, KB 9/43 m. 9.

⁴⁵ The National Archives, KB 9/43 m. 9d. This note is omitted in Powell and Trevelyan's edition.

⁴⁶ The National Archives, KB 9/43 m. 14. This membrane is noted by Powell and Trevelyan, but the list of names is not printed by them.

⁴⁷ The National Archives, KB 9/43 m. 15.

⁴⁸ The National Archives, KB 9/166/1.

⁴⁹ The National Archives, KB 9/166/1 m.1.

⁵⁰ The National Archives, KB 9/166/1 m. 2.

⁵¹ The National Archives, KB 9/166/1 m. 5.

⁵² The National Archives, KB 9/166/1 m. 40.

⁵³ For example, National Archives, KB 9/166/1 m. 25.

⁵⁴ This is discussed further in Prescott, *Judicial Records*.

⁵⁵ The National Archives, KB 9/166/1 m. 27.

⁵⁶ The National Archives, KB 9/166/1 m. 18.

⁵⁷ The National Archives, KB 9/166/1 m. 31.

⁵⁸ Jan Gerchow, 'Gilds and Fourteenth-Century Bureaucracy: the case of 1388-9', *Nottingham Medieval Studies* 40 (1996), pp. 109-48.

⁵⁹ <http://www.eppt.org>; <http://beowulf.engl.uky.edu/~kiernan/eBoethius/edit.htm#remodeling>

⁶⁰ <http://www.eppt.org/eppt-trial/EPPT-TrialProjects.htm>

⁶¹ A. Dekhtyar, I. E. Iacob, J. W. Jaromczyk, K. Kiernan, N. Moore, D. C. Porter, 'Support for XML Markup of Image-Based Electronic Editions', *International Journal on Digital Libraries* 6:1 (2006), pp. 55-69.