DHQ: Digital Humanities Quarterly
2009
Volume 3 Number 4

e-Science for Medievalists: Options, Challenges, Solutions and Opportunities

Peter Ainsworth <p_dot_f_dot_ainsworth_at_sheffield_dot_ac_dot_uk>, Dept of French and Humanities Research Institute, University of Sheffield
Michael Meredith <m_dot_meredith_at_sheffield_dot_ac_dot_uk>, Humanities Research Institute, University of Sheffield

Abstract

Medievalists typically resort to parchment for primary research and when editing their sources. Not always accurately catalogued, manuscripts copied onto animal skins may have started life in the same workshop but over the centuries have become dispersed, coming to rest in libraries all over the world; bringing these together entails travel, microfilm purchases and reassembly and collation of the data within reach of a microfilm reader. These unwieldy machines afford only moderate scope for exploring single manuscripts at close quarters. High-resolution digitisation yields not just better surrogates in full colour; it allows for the development of additional research tools using image compression and manipulation, and new modes of representation, e.g. juxtaposed display of several related witnesses. This paper outlines research questions underpinning the development of an electronic tool for viewing, transcribing and manipulating manuscripts; it moves on to show how the viewer can be adapted for access from remote sites, to compare and annotation one or more witnesses (interactively and in real time), and for use as an integral part of an online edition. Finally, it explores how it can be deployed for use on projects taking knowledge outside the academy: in museums, galleries and other public spaces.

Medievalists typically resort to parchment and vellum for their primary research. Housed in libraries across the world, the rolls, books and scraps bearing their precious written record are usually (though not always) reliably catalogued. Accessing manuscripts for study presents some fascinating problems. Whilst vellum is a durable support for the preservation of writing, the pigments used by the artists responsible for the miniatures and border decoration are susceptible to cumulative damage from abrasion or exposure to daylight. Some libraries direct scholars away from the originals and towards black-and-white microfilm stock of variable quality. Manuscript originals can sometimes be withdrawn for months at a time and at short notice for unavoidable conservation purposes. Yet another challenge familiar to scholars of medieval texts arises when one's target "collection" or family of manuscripts is scattered across the world in libraries each of which has its own access, "rights of reproduction" and copyright policies. Printed facsimiles exist, of course, but tend to be restricted to luxury manuscripts. Moreover they are discrete objects, whereas medievalists typically want to look (comparatively) at multiple witnesses of a text, at fragments and at ostensibly unappealing exemplars devoid of so much as a single miniature or illuminated letter. Microfilms are handy, but microfilm readers are not; the machines typically occupy a large footprint in any library and limit the scholar to work done on site. 1

Outlined below is a research trajectory starting with the scholar's need for high-quality facsimiles, for transcription and editing; we move on to consider how these might best be developed "beyond microfilm," or even CD or DVD, through recourse to the combined resources of high-resolution digital photography, image compression and other techniques derived from computer science. The outcome — a new electronic tool for viewing manuscripts — is described in some detail, with particular reference to a forthcoming electronic edition; we then explore how the same tool can be used by scholars from different institutions collaborating at a distance from one other over a shared Data Grid. The essay concludes with an exploration of how this and other e-Science tools can be deployed on projects which aspire to take knowledge outside the academy, to museums, galleries and other public spaces. 2

Microfilm and Beyond

3 Black-and-white microfilms produced in the 1940s are still considered to have achieved a peak of perfection in terms of image quality and sharpness of definition; even so the benefits they have afforded to generations of scholars have to a large extent been eclipsed — however unjustly — by the advent of high-resolution color photography. Offering considerable benefits in terms of precision, consistency and sheer quality, digital photography is still most frequently undertaken by research libraries not so much to satisfy the needs of the scholarly community as to protect original materials from repeated abrasions and exposure to light. Though medievalists will always seek access to original manuscripts, the better to evaluate their codicology and sheer physical presence, high-quality digital surrogates offer an increasingly viable alternative (for part of the work to be done, at least). One of the strangest paradoxes in the digital age is the absence to date of any new, custom-built or standardised electronic tools to take over the role of the microfilm reader. Digital copies have a potential value far beyond that of serving as a microfilm substitute, however. Digitisation surely comes into its own when harnessed to technologies able to exploit the full potential of file compression, and to set the surrogate free to perform tasks hitherto not asked of it.

4 Perhaps because medievalists are by their very training (and often by inclination) interdisciplinary, it is scarcely surprising that they have often been in the vanguard when novel partnerships have been formed with scholars from disciplines such as computer sciences. A prime example would be the pioneering electronic approach to the study of the British Library's comprehensively incinerated *Beowulf* manuscript, undertaken by Kevin Kiernan and his collaborators a full decade ago; this was to lead in turn to the creation of a range of e-Tools (known as *Archway*) for scholars working with medieval texts and manuscripts [Kiernan 1997].[1] One might also refer to large-scale lexicological and lexicographical projects such as the ongoing *ATILF Dictionnaire du Moyen Français* or indeed the print and now online versions of the *Anglo-Norman Dictionary* under development at Aberystwyth and Swansea. Medievalists continue to ask pertinent questions of computer scientists prepared to listen. Attention has shifted most recently towards high-resolution digitisation and e-Science. Eclectic yet significant collaborations are bringing together scholars, digital photographers, imaging specialists and computer programmers (but also, as we shall see below, librarians and museum curators). Scholars and programmers are beginning to ask one another how some of the most vexing research questions that preoccupy medievalists might benefit from digitisation and applications derived from it via e-Science and its particular agendas.

5 Novel technological solutions are thus being devised by teams coming together for new kinds of partnership, sometimes on an ad hoc basis, but also for more sustained programmes of collaboration fostered by grand challenge calls going out from the research councils. The forging of such partnerships can act as a catalyst for the emergence of radically rethought tools for scholarly praxis. Whilst the needs of the scholar remain paramount, the advent of e-Science is fostering kinds of dialogue, we would argue, in which researcher and programmer can jointly develop new ways of thinking around and through their respective areas of practice.[2]

6 Three instances of e-Science are showcased in this paper: starting with a tool for viewing, transcribing and manipulating (virtual) manuscripts, we move on to show how this resource is being integrated within a major online editorial programme; we then explore how it can be adapted for access by scholars at geographically remote sites for side-by-side comparison and discussion (interactively, collaboratively, and in real time). The paper concludes with a discussion of how the tool can be deployed for use on projects beyond the academy, in museums, galleries and other public spaces.

7 As is well known, much of the most intellectually challenging scholarship in the Arts and Humanities is delivered by lone scholars working in research libraries on original primary and secondary source material. This continues to be the mainstay of research in the UK and in most other developed countries and should never be decried; but research councils and other funding bodies have sought in recent years to encourage interdisciplinary and collaborative work wherever it looks set to open up opportunities that might not otherwise be feasible or thought conceivable. Beyond "complementarity of expertise, synergy, cross-disciplinary insights" and other fashionable buzzwords, the Arts and Humanities Research Council and the Engineering and Physical Science Research Council in particular have taken the lead in seeking to stimulate fruitful cross-disciplinary research. Two particularly exciting initiatives come to mind. The

most recent is their joint Science and Heritage programme which in 2007 funded a raft of interdisciplinary studentships, and which in 2008 launched a new programme to encourage research clusters expected to bring together curators, scientists and humanities scholars around collections of material culture and their real or potential audiences. The other joint initiative was a programme pursued in 2006, this time with support from the AHRC's own ICT Methods Network and the HEFCE-funded Joint Information Systems Committee (JISC); it funded just three e-Science "demonstrator" projects, one of which — described in this essay — focused on digitisation and data grids.

Virtual Vellum: An Online Viewing Environment

8 The first objective of this six-month project was to take a prototype manuscript viewer originally developed in Flash by digital photography consultant Colin Dunn of Scriptura Ltd, and to convert it into a more flexible, open access tool for use by the wider scientific and scholarly community. The resulting tool was named *Virtual Vellum* to reflect our desire to use high-quality digitisation to take people as close as possible to the materiality and detail of the medieval artefacts under consideration. The project aimed to explore the potential of high-resolution digital surrogate "manuscripts" from the standpoint not so much of the curating partner libraries involved (though their active support and involvement were key to success) as of the scholar-researcher and his or her particular needs.

9 Research questions to be addressed included the critical issue of how best to prepare, configure and mount high-resolution image datasets within an electronic viewing framework calculated to respond flexibly to the precise requirements of medievalists, rather than to those of multi-million pound software houses with generic products to sell at high volume. Design options and desiderata were in our case to be arrived at via conversations with practising medievalists, editors and art historians. Display and processing features would be evolved to respond to scholarly need or ambition rather than to the lowest common (or commercial) denominator. Whilst recognizing that such first-rate proprietory products as Microsoft's PowerPoint were more than fit for the very broad range of purposes to which they could be put and were designed for, the *Virtual Vellum* team sought above all to concern itself with ways in which a medievalist or art historian might wish to access, display, process, manipulate and marshal images on their PC or laptop screen. Use of the plural "images" is in itself significant; art historians delivering public or faculty lectures typically use a pair of 35mm slide projectors (with all the attendant advantages, risks and drawbacks).

10 Medievalists similarly wish to view multiple images, for a whole range of scientific reasons. These include the requirement to compare in fine detail two or more sections of text, or two or more folios from the same codex or from several related codices, the better to establish the objective differences between, say, the realisation of a given bookhand by Scribe A, Scribe C and Scribe J. Scholars have always been able to do this more or less intuitively by comparing the respective *ductus* of a given letter as executed by this or that copyist. The facility to move electronic "fragments" right up against one another and to zoom in on them independently of each other, was a further, entailed desideratum. Other functionalities to be incorporated arose from the scholar's desire to move quickly through a manuscript, to locate a particular folio in seconds, to measure the respective sizes of illuminated letters across a virtual "collection" of related digital manuscripts whose originals are now scattered across the globe, or (in the case of our art historian collaborators from Urbana-Champaign, Illinois) to measure the width of a king's mouth or the dimensions of his crown. More recently, users have asked us to incorporate a "last stage reached" button, permitting a user to resume work left off at a previous session and to set the screen to the precise display parameters last used.

11 The viewing pane was also to be susceptible of replication: scholars should be able to access two, three or even four image panes within the same screen space, each of which would have its own autonomous panning and zooming tools. It should be possible for users to add still more panes if so desired, and to tile them vertically or horizontally on their PC or laptop screen.

12 "Zooming and panning" evokes one of the most controversial issues encountered during the project. Librarians have to date been notoriously reluctant to allow high-resolution images of their manuscripts to feature on web pages, essentially and understandably for fear of unauthorised downloads leading in turn to copyright violation. They might, on the other hand, be prepared to sanction use of images as high as 500dpi on a commercially available DVD. The usual approved maximum, however, is 300dpi. From the programmer's and scholar's point of view this is a sorry state of affairs: high-

specification digitising camera scans can capture images at higher rates than this, and the results are visually stunning as well as optically more faithful to the original. However, a drawback of using higher resolution 500dpi images had to do, as we commenced the project, with the potential limitations, this time, of users' PCs and laptops. Not all of our users could be guaranteed to have machines with large amounts of RAM, super-performing graphics cards or gargantuan hard drives on which to store the rather hefty image files.

13 Just one scan of an A3-sized manuscript folio at 500dpi typically results in image sizes around 150MB. The raw files are captured as TIFFs and archived as such. To the uninitiated the raw TIFFs captured during a scanning operation can appear disappointing: dull, lifeless, even dark, they are hardly ever seen in this pre-processed state, but have to be catalogued using industry-standard metadata descriptors. For this and similar projects, the processed files are saved as JPEGs, "tweaked" to bring the colours to life once again. Even so, processed JPEGs can still be high-volume files. A second, major research question for the team to address was therefore how to "treat" or render the processed JPEGs in such a way as to reduce their volume and the time required to download them. In other words, to identify steps that would need to be taken to enhance the storage, retrieval, display and manipulation in real time of very high-resolution images (typically in excess of 8k x 6k pixels).

14 From Colin Dunn's original prototype using Flash, the project researcher moved to a solution involving what was then the new JPEG2000 standard. This was in part a response to issues arising from potential use of the viewer as an internet tool delivering image files over the web or via a Data Grid (a topic addressed in more detail below) and entailing the need for the efficient and prompt retrieval over such networks of what were, after all, still quite large images. But JPEG2000 also obviated the need for "tiling" or "file smash-ups". Instead, one file generated both a thumbnail and the principal image (the former to be displayed at the top of the screen with all the others in a kind of strapline format, the latter occupying most of the screen-space "theatre" immediately below). File compression enabled us to reduce the image file size to around 10MB, a very considerable economy.

15 A key element in taking the project forward proved to be the cordial relationships we were careful to establish with each of our partner libraries. Unlike many other institutions they were prepared not only to allow us to use our own photographer (which in turn meant that the project could maintain common and consistent standards from one institution to the next); they were ready also to let us use and reproduce the captured images at 500dpi, and to experiment with modes of delivery both locally on a PC or laptop, and over an internet connection. Once again, speed of delivery and manipulation were an issue: our potential users might not have access to sophisticated home broadband; even if they did, there was a likelihood that the images would still take an unconscionably long time to download.

16 The requirement somehow to square this circle, allowing us — potentially — to use the highest resolution images we possessed without slowing down processing speed to the point at which use of *Virtual Vellum* (online or offline) would become self-defeatingly fastidious, presented us with one of the most formidable challenges of the entire project. It was eventually resolved through implementation of a series of mathematical algorithms to optimise memory usage and processing power, combined with exploitation of *threads*.[3] By the conclusion of the project we had produced an online viewing tool specifically designed for day-to-day use by researchers whose projects typically involved close scrutiny and manipulation of high-resolution image files: not just medievalists but art historians and scholars from other disciplines working with large corpora of digital photographs, all of whom might wish to use the tool for research purposes, but also for more public deployment, during a lecture or workshop perhaps.

17 *Virtual Vellum* emerged from the processes described above as a fully-fledged and versatile viewing tool for the manipulation of high-resolution images. The AHRC-JISC-EPSRC e-Science project allowed the team at Sheffield to devise a highly flexible viewing environment equipped with a range of associated tools, all compatible with Windows, Mac and Linux. *Virtual Vellum* allows scholars to present research papers with confidence whilst manipulating their high-resolution image files at speed, efficiently and flexibly. It has the additional benefit of encouraging scholars to use datasets that might otherwise remain dormant due to their prohibitively large file-size volume. An open access and open source product, it is henceforward available to scholars entirely free of charge. Downloadable from the project website complete with user manual in plain English; all that it needs to show its paces is a set of high-resolution images. From the very outset the tool was designed to appeal to any researcher involved in the regular use of such datasets. Fields

envisaged included, to mention just a few of the more obvious, manuscript studies, art history, iconography, theatre and film studies, museum and gallery studies. We were less prepared for the interest shown in *Virtual Vellum* by researchers working in the natural, physical and social sciences, described below.

The core features of *Virtual Vellum* are illustrated below by a short series of figures using sample images from our datasets. Whilst these obviously showcase manuscript material, it is worth mentioning at this juncture the interest shown in our viewing tool, used online via a data connection, by programmers at the Humboldt University in Berlin busy developing software for entomologists working in the field for Germany's Natural History Museum. A scholar coming across a new species of beetle in the Amazonian rain forest might well find it useful to be able to take a digital photograph and port it back to Berlin, or compare it to all similar genus categories on their own laptop-driven database connected via satellite, so we are told. 18

Figure 1 below illustrates the quality of the digitised (virtual) manuscripts used on the Sheffield project, the image on the left providing a preview of the complete image, whilst the image on the right is shown magnified to 100%. Figure 2 illustrates how frontispiece pages from four different but cognate manuscripts can be viewed side by side; the originals of the four manuscripts are located in three physically different locations (eastern France, Lancashire and Brussels); prior to the project they had never before been shown "together." 19

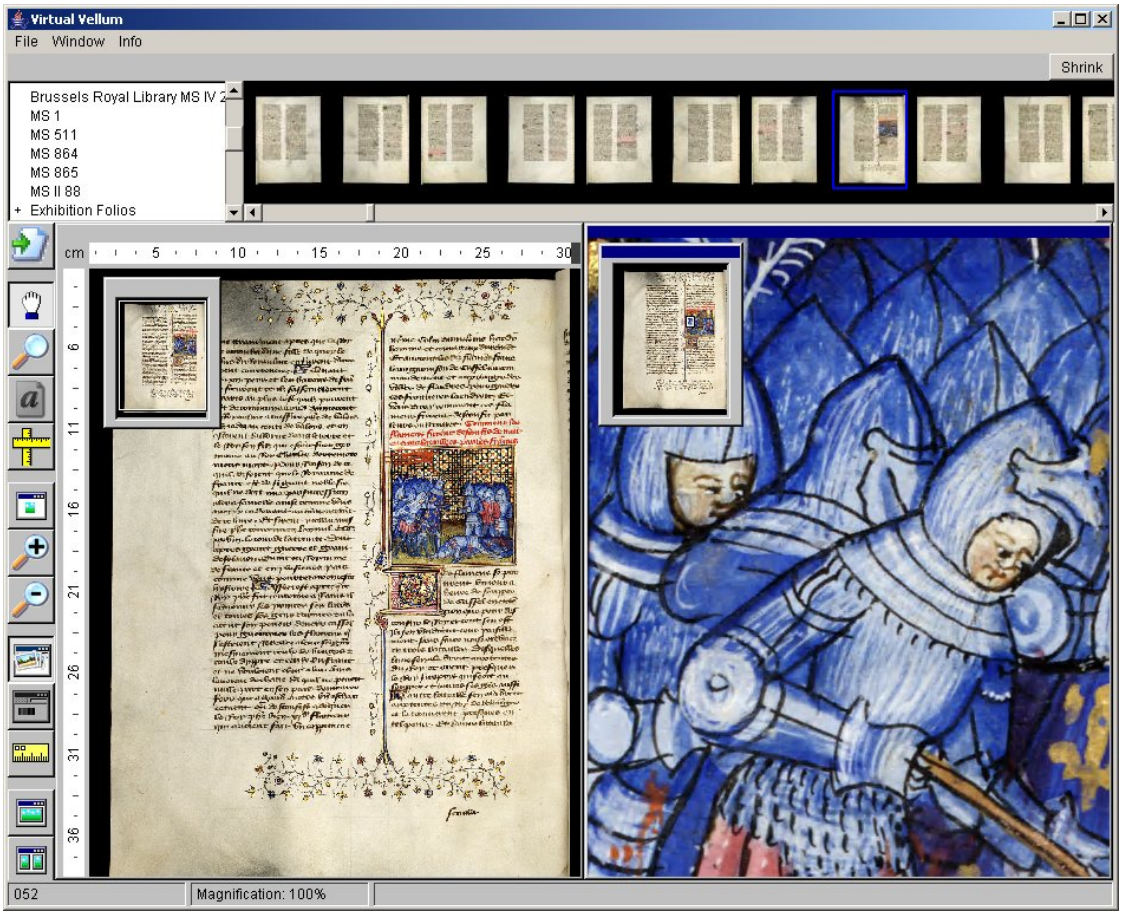

Figure 1. Illustrates image quality; each image can be viewed at any level of magnification, and in real-time. The left-hand image shows the folio scaled so that it can be viewed in its entirety, while the image on the right is displayed with a 100% magnification ratio. Images are © Bibliothèque Royale Albert 1er, Brussels, and Scriptura Ltd.

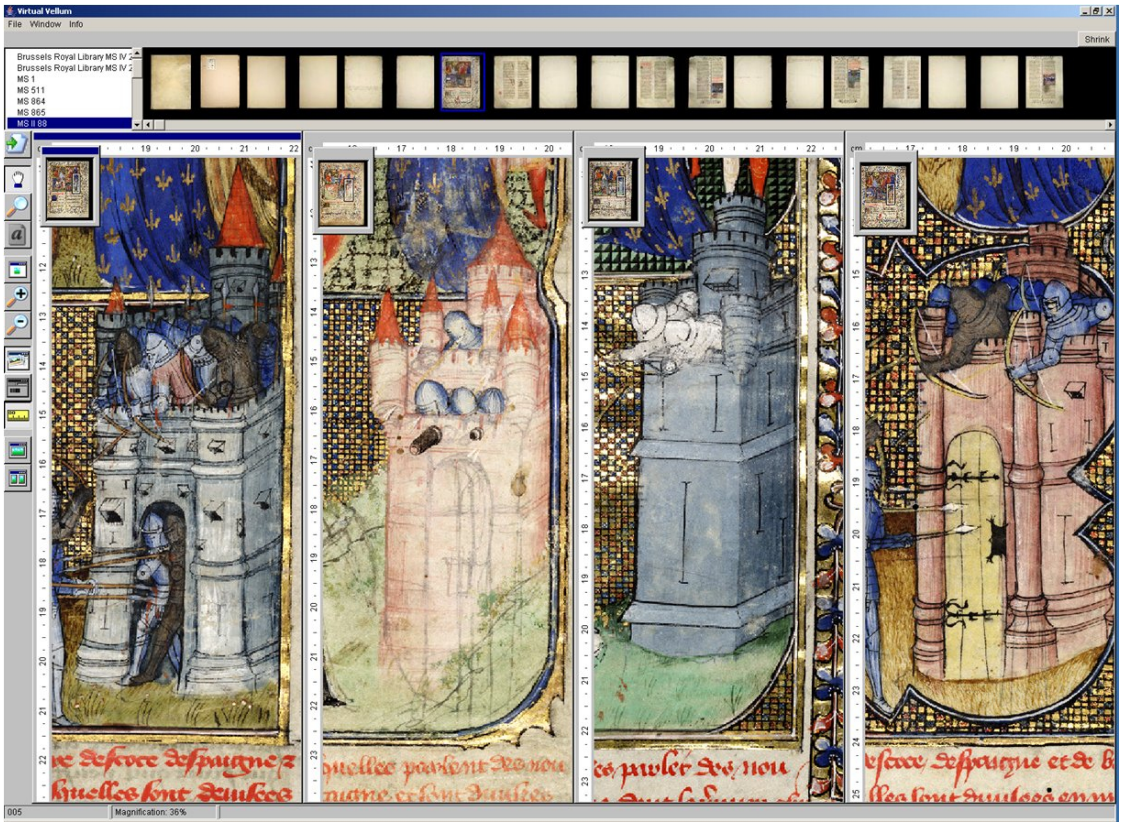

Figure 2. Compares the frontispieces from four different manuscripts located at three different physical locations (magnification is at only 30% of the originals). Images are © Bibliothèque Royale Albert 1er, Brussels (left- and extreme right-hand images), Stonyhurst College, Lancashire (second from left), Bibliothèque d'Etude et de Conservation, Besançon (second from right), and Scriptura Ltd. (digitiser of all the manuscript images shown).

e-Science and Data Grids

Scholars presenting papers live at conferences or delivering seminars online and wishing to feature visual data, typically in the guise of side-by-side comparisons, have hitherto (as noted earlier) been dependent on PowerPoint or twin 35mm colour-slide projectors. Good as these are, they were never designed by, or for use by, scholars working in these various fields, as *Virtual Vellum* was. Envisaged in part for stand-alone or seminar use, for personal or for collaborative research, *Virtual Vellum* was also conceived for deployment over a robust and secure internet connection, or better still over a supercomputing grid network. This topic is addressed in the next section of this essay. 20

An invitation to present *Virtual Vellum* at the UK e-Science All Hands Meeting in September 2006 brought the project team into direct contact with a fresh wave of potential collaborators. These included the team at the University of San Diego[4] responsible for development of SRB (Storage Resource Broker) middleware. Using this file management system over a Data Grid would in theory allow us to set up a cluster of grid nodes or networked sites affording shared access on the part of researchers separated by oceans or continents but linked via a robust Grid, to the whole of our high-resolution image datasets, totalling at that point almost a terabyte of data. The straightforward premise behind this conviction was and remains that scholars working on international collaborative projects involving image collections held on local or distributed databases may wish to consult one another, the better to explore (together) questions of mutual interest such as aspects of iconography, art history, image content, or comparison of similar or related images (or specimens!). The Access and Data Grids afforded the ideal framework and necessary computing power for the rapid and efficient deployment of securely stored, IP-protected collections of this kind, plus high processing speed and efficient file handling. 21

Ideally, a viewing environment involving scholars talking live to each other about their common "quarry" would offer scope for real-time, close-up scrutiny of single or multiple (juxtaposed) images, with independent zooming controls and 22

other functions including hotspotting (zoning and highlighting), a facility for collaborative annotation, and a sophisticated blogging tool.

The “testbed” dataset for *Virtual Vellum* was provided by a corpus of six complete (virtual) surrogate manuscripts specially photographed thanks to funding provided by Yorkshire Universities Gift Aid, the University of Sheffield and the Leverhulme Trust. They were all captured on site by Colin Dunn, at each of four partner libraries, over a period of three years: Besançon and Toulouse Municipal Libraries, Stonyhurst College Library, and the Bibliothèque Royale de Belgique in Brussels. Approximately one terabyte’s worth of raw archive image TIFFs were securely archived on a server at the University of Sheffield. The processed files, converted to the *JPEG2000* standard (each no more than 10MB in volume) are kept on a server at Sheffield’s Humanities Research Institute, federated to the White Rose and WUN Data Grids, ready for sharing and distribution. SRB is the middleware which does all the gophering, fetching and carrying. 23

As mentioned above, one of the prime uses envisaged for *Virtual Vellum* involved live online seminars involving scholars logged on from several remote but affiliated sites, connected as a Grid and therefore supported by networked supercomputers. The idea was that the viewing environment should typically allow scholars to “bring together” and display within a single environment for purposes of comparative research, several manuscript witnesses (i.e. successive or alternative scribal copies or versions) of the “same” text, whose originals were housed, however, in libraries scattered across the globe. This objective has been achieved; scholars using *Virtual Vellum* can retrieve the data files they wish to look at either from a local hard drive or over the internet, or better still via a robust Data Grid (which affords greater security and confidentiality). 24

Written entirely in Java, *Virtual Vellum* can be run either as an applet or as a desktop application. Java version 1.2 was chosen because most web browsers (minimally) have support for this version of the Java run-time environment; people who have not yet got access to the later versions of Java are thus not excluded from use of *Virtual Vellum*. The software package offered to users includes a *JPEG2000* Image Decoder and a “multiple-pane” user interface designed to work from within a memory-limited system (less than 65MB, typically the maximum amount of memory used for the deployment of Java applets). It also takes full advantage of modern hardware developments such as multiple processor cores. Open source access is provided because the product has been designed for total computer platform independence. Support is included for single, double or multiple views, for comparison of two, three or more different images. *Virtual Vellum* has already been used successfully during several live AHRC-funded collaborative Access Grid workshops, in the context of stand-alone conference presentations, and at events such as UK All_Hands and Digital Resources for the Humanities and Arts. An associated tool called *Collator* has since been developed to allow rapid and easy conversion of small or large numbers of image files (including TIFFs, BMPs and JPEGs) to the *JPEG2000* file standard. 25

Virtual Vellum has been adopted for use by Stonyhurst College in the UK and by the Ceccano Library in Avignon. At the time of writing it is being evaluated by the Department of Manuscripts at the Bibliothèque nationale de France (Paris), who in July 2008 confirmed their invitation to the project PI to undertake a sustained programme of digitisation around the BnF’s extensive collection of Froissart manuscripts. A collaborative endeavour of this kind, and on such a scale, is not possible without a very significant degree of mutual trust, respect and recognition of (equally mutual) benefit, involving the scholar, photographer and of course the directors and curators of each of the research libraries with whom the *Virtual Vellum* team has always sought to nurture good working relationships. 26

Virtual Vellum and the *Online Froissart*

Digital surrogates derived from Besançon Public Library mss 864-865 are currently seeing service in the context of an AHRC Resource Enhancement award made to the author in association with Godfried Croenen of the University of Liverpool for an online edition of Froissart’s *Chroniques*. The *Online Froissart* is scheduled to complete by March 2010. With an interface and search engine designed by the University of Sheffield’s Humanities Research Institute, and incorporating the *Virtual Vellum* viewer, the online resource will deliver complete texts for Books I, II and III of Froissart’s Chronicles. The base texts will be collated against transcriptions of other witnesses prepared by the project team, 27

managed with the aid of Peter Robinson's *Collate* software. Historical, political and cultural annotation will be supplied together with new translations into modern English. A concordance, glossary and index will also be provided. The *Online Froissart* is part of an international consortium which includes the University of Edinburgh's *Christine de Pizan Queen's Manuscript* project (Professor James Laidlaw and team) and the University of Nancy's ATILF (*Dictionnaire du Moyen Français*) project; the consortium enjoys joint support from the British Academy and Centre National de la Recherche Scientifique. Users of the *Online Froissart* will have access to complete full-colour reproductions of the base manuscripts and of several cognate witnesses collated against these. Our colleagues on the Edinburgh project have contributed to ongoing development of *Virtual Vellum* and have adopted it for their own purposes.

Perhaps the most obvious contribution brought to the development of *Virtual Vellum* by the *Online Froissart* is a feature whereby the lines of each column of text in the surrogate's electronic folios have been associated electronically with their equivalent lines in the edited transcription (this is achieved automatically by analysing each image for lines of text and matching these with the transcription markup files). Panning and resizing of the image entails linked, proportionate positioning and resizing of the transcription, and vice-versa, as illustrated below in Figure 3.

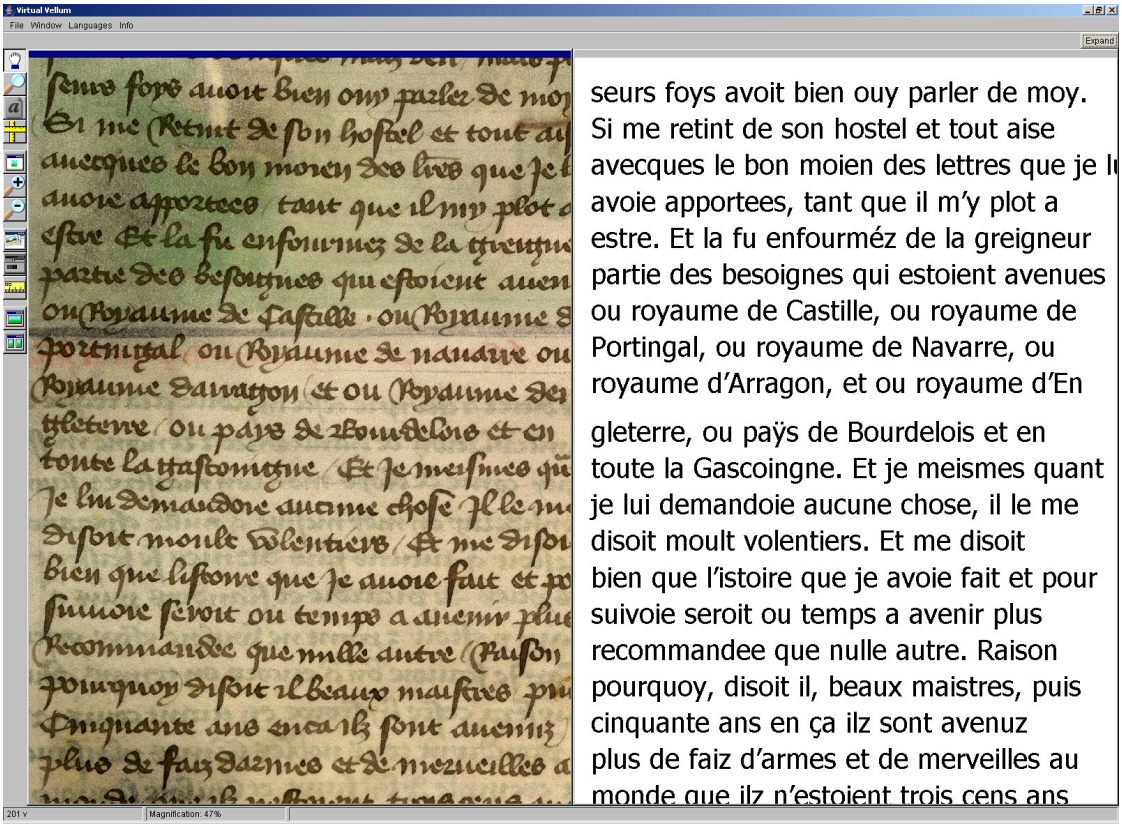

Figure 3.

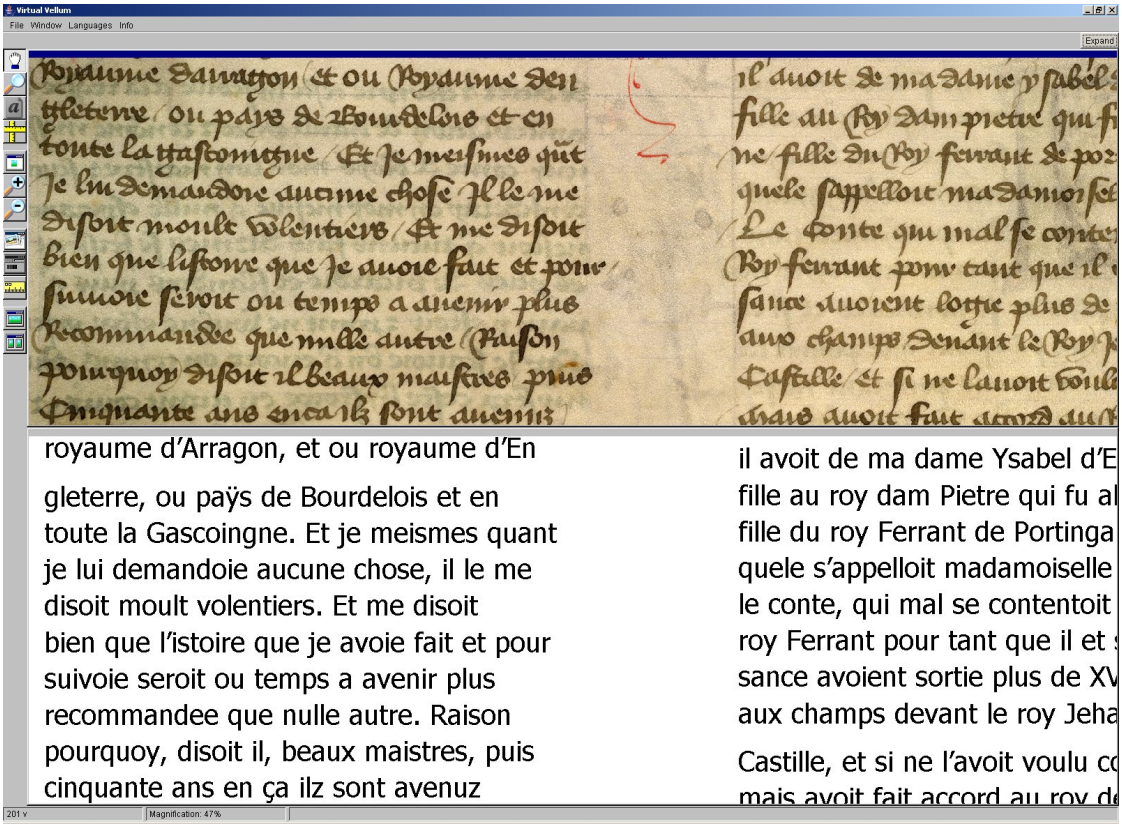

Figure 4. Screenshot juxtaposing a page of transcribed text from Besançon Public Library ms 865 and the original folio image, which are synchronised while panning and zooming.

e-Science, e-Learning

Virtual Vellum was recently adapted for yet another purpose, this time under the name *Kiosque* and for service as an exhibition enhancement tool. The aim in this case was to provide visitors to an exhibition in Leeds with interactive digitised material accessed through touch screens. The manuscript viewer therefore found itself at the core of a four-month exhibition mounted at the Royal Armouries Museum between 08 December 2007 and 06 April 2008, entitled *The Chronicles of Froissart: from conflict to cooperation*. Covering the Hundred Years' War as recounted in the Chronicles, the exhibition focused on the craftsmanship underpinning the source manuscripts and the arms and armour featured in the miniatures. On display was the (real) Stonyhurst College manuscript, shown alongside five other virtual manuscripts represented as interactive surrogates, and against the background of items specially selected from the Museum's internationally celebrated collection of contemporary arms and armour. These were displayed against the backdrop of miniatures reproduced from all of our digital surrogates showing the same items in use. 29

The purpose of the *Kiosque* software, developed jointly with e-Learning specialists Tribal (Sheffield) with support from the Department of Trade and Industry's Knowledge Transfer Partnership scheme and incorporating a version of *Virtual Vellum* adapted for touch screens was to provide a fully configurable e-Learning tool that could reach out to wider and more variegated audiences than that comprising professional medievalists and art historians. Visitors could explore image content at varying levels according to their particular interests. In addition to its presence on the touch screens, the manuscript viewer was used to support real-time transcription of sections of the Stonyhurst manuscript by Sara Mack, an experienced calligrapher. *Kiosque* is illustrated below in Figure 4. Similar though smaller-scale French-language versions of the software were on display throughout the Leeds exhibition, at three libraries in the south of France: the Alcazar at Marseilles, the Cité du Livre-Méjanes in Aix-en-Provence and Avignon's Ceccano Library. At time of writing there are plans to mount a French version of the exhibition at the Musée de l'Armée in Paris (Hôtel des Invalides), with support from the Department of Manuscripts at the Bibliothèque nationale de France. 30

Figure 5. *Kiosque* — telling the story of Jean Froissart's Chronicles using the digitised images as a learning resource.

Pegasus

In the early months of 2007 the UK's Engineering and Physical Sciences Research Council launched a programme of experimental collaboration with the US National Science Foundation under the latter's *TeraGrid* initiative. EPSRC funding was secured by the author to undertake a further demonstrator project incorporating *Virtual Vellum* and baptised *Pegasus*. The project takes the manuscript viewing tool deeper into the realm of e-Science via a new partnership established with the State University of Illinois at Urbana-Champaign, and the National Center for Supercomputing Applications based on the campus there. The project aims to establish a grid-supported infrastructure for intercontinental research based on high-volume image and text datasets. The *Online Froissart* project once again supplies the initial dataset, namely its full-colour, high-resolution digital surrogates of closely-related manuscripts of the *Chroniques* housed in libraries located in three different countries. 31

By July 2008 *Pegasus* was set to take off under the joint auspices of the EPSRC, NSF and *TeraGrid* initiatives, with the full support of the Universities of Sheffield and Illinois. Additional financial and cyber-infrastructural support was provided by a generous grant from the Worldwide Universities Network, whose Executive Director Dr David Pilsbury also ensured that the Sheffield team were able to obtain digital surrogates from both the Bibliothèque Royale Albert 1er and the Bibliothèque nationale de France. The support of WUN Grid, involving technical expertise provided by the Universities of Bergen and York, is supplemented by equally robust computing support via White Rose Grid, underpinned by the Universities of Sheffield, York and Leeds. An early event in the joint programme was the attendance of programmer and project research associate Michael Meredith at a Supercomputing 08 workshop run at Urbana-Champaign in late July 2008 by our NCSA partners. This allowed him to present and facilitate a hands-on workshop surrounding *Virtual Vellum* to a much wider audience, including the *Abraham Lincoln Papers* project team, and to scope with art historians and programmers at Urbana-Champaign a series of initiatives around pattern recognition algorithms. 32

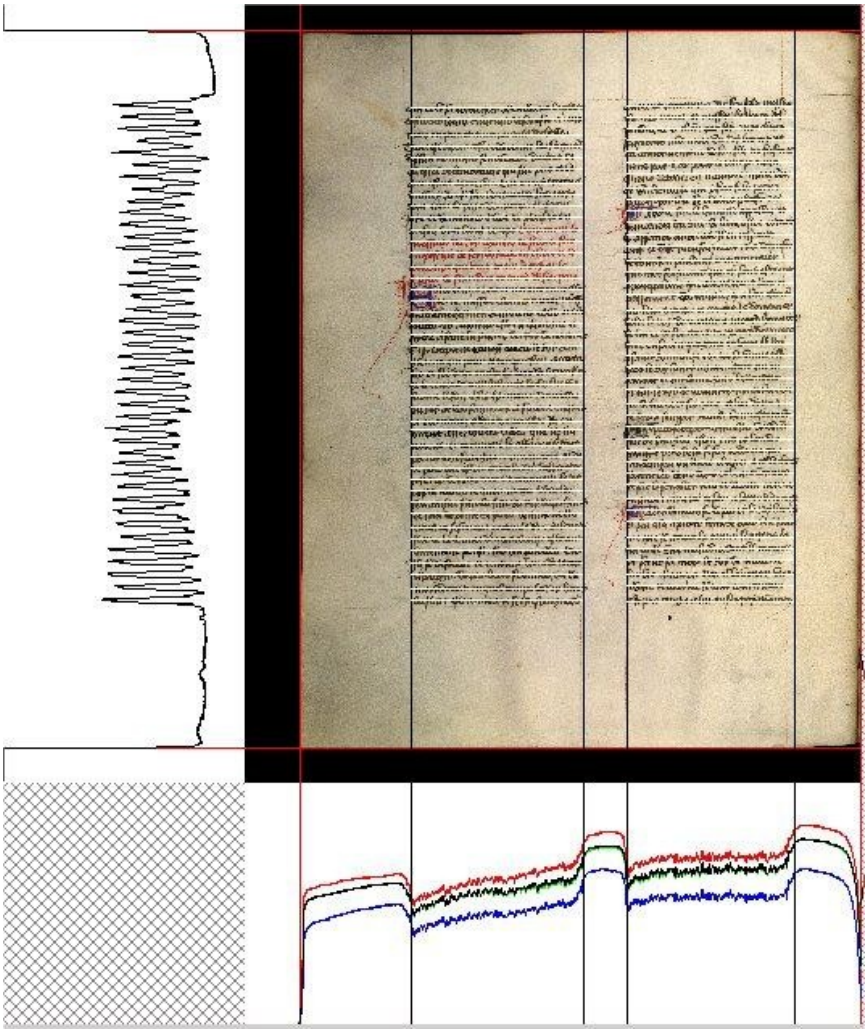

Figure 6. Using colour channel histograms and their gradient changes to automatically detect manuscript lines, folio characteristics (such as rubrics) and miniature locations.

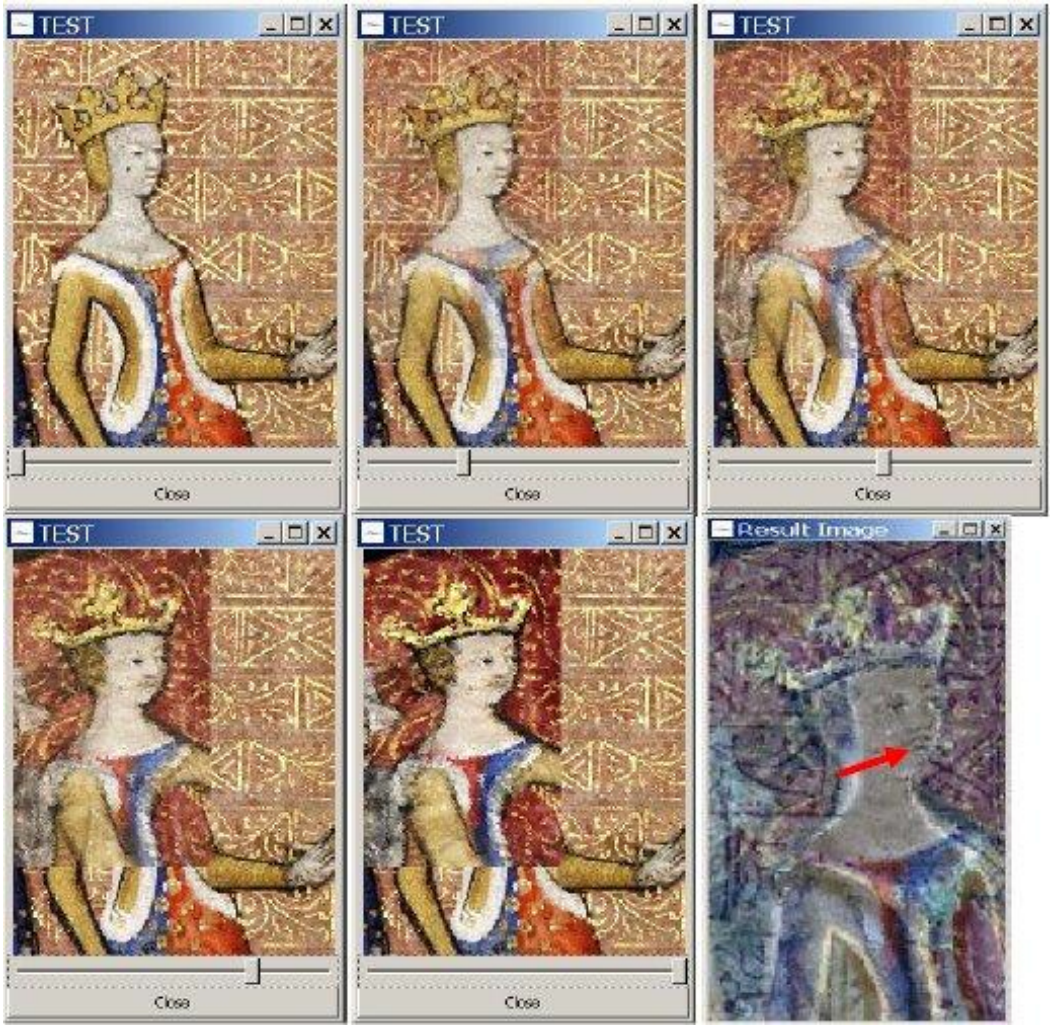

Figure 7. The result of registration of two illustration sub-areas. The result is presented as a sequence of two image overlays with variable transparency (0, 25, 50, 75 and 100% transparency of the right image on the left image) as indicated by the slider bar on the bottom of each window (top row and the bottom left and middle figures). The user driven change of transparency parameter allows not only exploring the visual differences in the crown and the colour pattern of the cloth but also quantifying the scaling offset of the mouth by each artist (bottom right figure obtained by subtracting the registered images. Image composed by Professor Peter Bajcsy of the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign.

Whilst the *Online Froissart* virtual manuscripts serve once again as the pilot dataset for *Pegasus*, partners at Illinois, Ottawa and at other WUN member institutions have initiated discussions with a view to securing the resource needed to make available for research an entirely different manuscript, the British Library miscellany (Royal 15 E.VI) more familiarly known as the Talbot Shrewsbury Book. This lavishly illuminated compilation volume encapsulates the full range of writings then suitable for a warrior prince or aristocrat, combining the literature of chivalric romances and *chansons de geste* with the historical accounts provided by chronicles, and with practical treatises and manuals of kingship, warfare and chivalry. The extraordinary range of texts and illuminations included in this manuscript means that it can only properly be understood by a team of experts, interpreting the artefact from a range of different but complementary intellectual and disciplinary perspectives. Under the auspices of the Worldwide Universities Network, an international team of scholars proposes to use a high-quality digital reproduction of the Shrewsbury Book in order to investigate numerous research questions such as how it was read and used, its influence in reshaping the cultures of chivalry and warfare in England at the end of the Hundred Years' War (1336-1453), where and by whom the manuscript was created and how it compares to other collections. A detailed codicological examination of the manuscript via the high-resolution images of the digital reproduction, via *Virtual Vellum*, is expected to shed further light on these questions, in particular by extending scholars' knowledge of the marginalia and other forms of notation on the manuscript, but also of its programme of illustration, and the different scribal hands and artists who worked on the manuscript. Initial scientific work towards achieving this objective was undertaken in Urbana in July 2008 by Michael

Meredith, Peter Bajcsy and Professors Anne D Hedeman and Karen Fresco.

Scholars provided with access to the resource would be able to explore annotations and interpretive essays on the model of the Christine de Pizan Queen's Manuscript, British Library MS Harley 4431, http://www.pizan.lib.ed.ac.uk/. Users would be able to contact team members via a blog or wiki and explore the manuscript material at their leisure using the e-Science framework and tools that have already been and continue to be developed. In sum the project bids fair to be of great interest to all those interested in chivalry, ideology of the political classes in France and England, the Hundred Years' War, patronage, canon formation, heraldry, and the iconography of the miniature. 34

In Conclusion

Using the combined resources of the Access and Data Grids, and of the *Online Froissart* project dataset, *Virtual Vellum*, *Kiosque* and *Pegasus* aspire together to establish an exemplar platform for collaborative project development based on large-volume, high-resolution image datasets, involving curators, photographers, imaging and grid technicians, programmers, and, most importantly, specialists in medieval textual, historical and art-historical scholarship. The *Shrewsbury Book* consortium offers a similar approach. It is through partnerships such as those outlined above that genuinely useful collaborative tools will be forged to allow scholars in medieval studies and other, cognate areas to benefit from the technological advances currently being made in e-Science for the Arts and Humanities. 35

Whilst there is never an adequate substitute for handling and viewing the real artefacts of our material culture, the ability to have at one's disposal high-resolution digital surrogates that one can "handle" and inspect virtually, at greater leisure than is usually possible during necessarily brief visits to research libraries, has proven to be a more than satisfactory alternative. The advantages already described are further complemented by the ability to bring together, for instant comparison and in real time, materials that are geographically remote from each other, all the while respecting local IP rights. It may perhaps be some time before such technologies are taken up by the mainstream, but it is now that they need to be developed and trialled in order to ensure their availability and usefulness to scholars when that day eventually arrives, as it surely will. 36

Acknowledgements

We are grateful to Professors Anne D Hedeman and Karen Fresco, College of Liberal Arts, State University of Illinois at Urbana-Champaign, Dr Craig Taylor, Senior Lecturer in History, Centre for Medieval Studies, University of York, and Dr Andrew Taylor of the University of Ottawa, for kind permission to use material reproduced here. 37

Notes

[1] See also [Shillingsburg 2006, 203] and [Lesk 2004, 330].

[2] Not the least appealing aspect of such work is its collaborative basis. The new opportunities for interdisciplinary collaboration described here deserve to be celebrated and promoted in so far as they can be shown to contribute significantly to the enrichment and even the transformation of disciplines which might, otherwise, risk stagnation.)

[3] This allows the software to make use of modern multi-core processors to decode regions of an image simultaneously and thus display it to the user much more rapidly.

[4] A co-member with the University of Sheffield of the Worldwide Universities Network, a key factor in the successful development of our project ambitions.

Works Cited

Ainsworth 2010 Ainsworth, Peter. "Editing, E-Science and Exhibitions". In Rebecca Dixon, *Essays in Later Medieval French Literature: The Legacy of Jane H.M. Taylor*. Manchester: Manchester University Press, 2010.

Kiernan 1997 Kiernan, Kevin S. *Beowulf and the Beowulf Manuscript*. Ann Arbor: University of Michigan Press, 1997.

Lesk 2004 Lesk, M. *Understanding Digital Libraries*, Language, Arts and Disciplines. New York: Peter Lang, 2004.

Shillingsburg 2006 Shillingsburg, Peter L. *From Gutenberg to Google: Electronic Representations of Literary Texts*. Cambridge: Cambridge University Press, 2006.