

FORUM

University of Edinburgh
Postgraduate Journal of Culture and the Arts
Issue 12 | Spring 2011

Title	Digital Attraction: from the real to the virtual in manuscript studies
Author	Peter Ainsworth
Publication	FORUM: University of Edinburgh Postgraduate Journal of Culture and the Arts
Issue Number	12
Issue Date	Spring 2011
Publication Date	05/06/2011
Editors	Elysse Meredith & Dorothy Butchard

FORUM claims non-exclusive rights to reproduce this article electronically (in full or in part) and to publish this work in any such media current or later developed. The author retains all rights, including the right to be identified as the author wherever and whenever this article is published, and the right to use all or part of the article and abstracts, with or without revision or modification in compilations or other publications. Any latter publication shall recognise FORUM as the original publisher.

Digital Attraction : from the real to the virtual in manuscript studies

Peter Ainsworth

In a BBC Radio 4 programme broadcast on 26 April 2011 (*Tales from the Digital Archive*),¹ archaeologist Christine Finn explored some of the ways in which the digital revolution has changed writers' working methods, and the consequential impact that these have had on librarians, curators and conservators. Wendy Cope recently donated an archive of 40,000 emails to the British Library, providing scholars with invaluable material for research on drafts of her poems as they developed. Fay Weldon found herself changing quite easily from pen and paper to computer and mouse, discovering in the process that it changed how she actually wrote her novels. Altering phrases or drafts became much easier, and much cheaper too. Weldon's archive is destined for the University of Indiana. Salman Rushdie's archive went to Emory University in 2006, becoming available to researchers in 2010; it includes 15 years of electronic material from his computer.² The British Library has a Curator of Digital Manuscripts, Jeremy John, whose strongroom contains IBM and Macintosh computers, floppies, CDs and other ephemera such as post-it slips (stored in archive boxes alongside the machines they were attached to).

Flaubert scholars have for a long time pored over the consecutive drafts of *L'Education sentimentale*, and the field of textual genetics has delivered important

¹ I am grateful to the BBC and to producer Marya Burgess for permission to reproduce the material featured in my opening three paragraphs.

² Whilst one may regret so many of our literary authors committing their archives to overseas institutions, these are often the best-equipped to curate them, and digitised material can of course be disseminated much more easily than original material on print or paper.

conclusions. In the near future scholars will be asking themselves slightly different questions: How did authors use and customise their desktops? What were their characteristic habits and shortcuts? What applications were used, and how were their files organised? How was a key manuscript altered on a given day? Which emails went out that day and to whom? How did a particular phrase evolve?

Digital technology has had implications even for writers who cordially disliked computers, including Ted Hughes. Hughes favoured the typewriter; unlike Rushdie's or Weldon's archive, therefore, his contains no *digitally born* material; but the British Library's archive of the poet's Devon home will include a 360° panoramic picture of his workspace based on 50 high-resolution photographs showing the books on his shelves, his collection of stuffed birds and an ingenious personal system of wooden blocks for keeping track of his projects in progress.³

Digital curation and conservation bring their own challenges: the moment you turn a computer on, you risk changing the dates; the British Library uses a 'write-blocker' connected to a forensic workstation to access material without changing it. A screen image of an author's desktop and files as left, for example, on the day of his or her demise can be generated from their hard drive using an emulator; the hard drive is in this way not altered by date changes or software version updates. Searches can be conducted to pinpoint and delete private information such as credit card or phone numbers, though curators still have to take ultimate decisions as to what may be publishable.

The relationship between scholar and curator-conservator is becoming acutely important as the 21st century gets into its stride. It always was, of course, just as paper and print will continue to hold the attention of the majority of arts and humanities

³ Ted Hughes liked to cut and paste, but using paper and sellotape ('an archivist's nightmare').

researchers for generations to come. But the advent of high-resolution digital photography and of computing for the arts and humanities affords not only new opportunities and toolkits, but also the emergence of new kinds of research question.⁴ Scholars working in this exciting and intensively collaborative, interdisciplinary and (arguably) cost-effective field are beginning to do things they were not able to before. Space precludes comprehensive exploration of the field, so just two projects are reviewed here, one of which proceeds in part from the other. Both focus on the *Chronicles* of the 14th-century writer and poet Jean Froissart (?1337-?1404). For reasons that will become clear, we trust, we begin with the later of the two, forming part of the internationally-funded Digging into Data Challenge.⁵

In the Spring of 2011 an exhibition at the Invalides in Paris curated by Peter Ainsworth in partnership with the Royal Armouries UK and the national French Musée de l'Armée portrayed aspects of the literary and military culture of the Hundred Years' War. Items of weaponry from the period were displayed against a backdrop provided by large-scale, high-resolution photographs of miniatures from

⁴ High-resolution photography is not, of course, a *sine qua non* for the best in computer-enhanced research in the arts and humanities or social sciences, witness such eminent projects as *OldBailey Online* (*OldBaileyProceedings Online* (www.oldbaileyonline.org [accessed 04 May 2011]) or *Connected Histories* (*Connected Histories* (www.connectedhistories.org [accessed 04 May 2011]).

⁵ The *Digging into data challenge* programme (2009-2011) funded eight 'cutting-edge' projects, arousing so much interest that a second round is open at time of writing (closes 16 June 2011). Round two is supported by no fewer than eight international funders: <http://www.jisc.ac.uk/news/stories/2011/03/digging.aspx> [accessed 04 May 2011]. See also Hannah Fearn's *Times Higher Education* article, 'Research intelligence Let's dig a little deeper', *THE* 28 April 2011, and 'Digging into Data Using New Collaborative Infrastructures Supporting Humanities-based Computer Science Research', *First Monday* (May 2011).

Froissart's *Chronicles*.⁶ The more practical aspects of manuscript culture featured in the guise of pens, brushes, pigments and other implements used by the scribes and artists responsible for the miniatures found in many manuscripts of the *Chronicles*, kindly loaned by the *Scriptorial* museum in Avranches, Normandy. The centrepiece of the exhibition was a display in two glazed cases facing one another across a darkened room, of two pairs of twin manuscripts. Besançon Municipal Library mss 864 and 865 (comprising respectively Book I and Books II-III of the *Chronicles*) thus found themselves just yards away from Paris, Bibliothèque nationale de France fonds français mss 2663 and 2664 (Book I and Book II of the *Chronicles*). This was exciting for Froissart scholars, since the four volumes had originally been copied and illustrated in central Paris around 1412-1418 under the direction of bookseller Pierre de Liffol, who seems to have glimpsed a market opportunity for the production of luxury copies of Froissart's Middle French *Chronicles* (Croenen).⁷ The four volumes were displayed open at a fixed point inside sealed cases, but their entire contents could be explored in virtual format via interactive touchscreens nearby, using the *Kiosque* software specially developed for the purpose at the University of Sheffield by the Department of French and the Humanities Research Institute (Meredith).⁸

The *Chronicles* remain one of the most important prose accounts in French of the Hundred Years' War between France and England and their respective allies; they

⁶ Images from the exhibition can be viewed at: <http://hrionline.ac.uk/onlinefroissart/> under "About the Project", iv. Related Projects.

⁷ The books were sold mainly to clients in the service of Charles VI of France, though the iconographical emphasis of the illustrations to Besançon BM, ms. 864 testifies to a client with pro-*English* sympathies: Peter Ainsworth, 'Representing Royalty: Kings, Queens and Captains in Some Early Fifteenth-Century Manuscripts of Froissart's *Chroniques*', *The Medieval Chronicle* IV, Editions Rodopi (Amsterdam/Atlanta, 2006), pp. 1-38.

⁸ *Kiosque* was developed by Dr Mike Meredith working with Peter Ainsworth and Tribal (Sheffield). The virtual versions of the Besançon and Paris manuscripts are part of a corpus comprising more than 6,000 high-resolution image files captured photographically from ten digitised manuscript volumes (2TB of data).

remain a key source for study of the conflict. Their content forms the basis of the *Online Froissart* discussed later on in this paper. The Sheffield team on the *Digging into Data to answer Authorship-Related Questions* (DID-ARQ) project engaged above all with the virtual *Chronicles*, considered not so much as flexible surrogates for the originals locked in their closed display cases or hidden away in the strongrooms of their research libraries,⁹ but rather as a supplementary source of data. However, the research began with the originals themselves (collation, codicology, palaeography, etc.). A second phase involved the careful nurturing over many months of good relationships of mutual understanding between the researchers and the conservators and librarians at the various libraries where photography was to take place. Once these were in place, the recruitment of talented specialist photographers¹⁰ led in due course to the capture, under almost identical conditions, of no fewer than ten complete facsimiles (most of which are currently viewable via the *Online Froissart*), seven of them at an unusually high resolution (500 dpi). Development of fit-for-purpose manuscript viewing and manipulation software followed, *Virtual Vellum*; final copyright clearance and permissions were then obtained to cover shared use of the images for the *Online Froissart* and by the international DID-ARQ consortium. All but two of the virtual manuscripts were shared in this way.

Colleagues at Illinois's College of Fine and Applied Arts, in partnership with the National Center for Supercomputing Applications, adopted as their particular focus the virtual manuscripts' decorative and illustrative content, concentrating on the twin manuscripts displayed in Paris in Spring 2010. Art historians working on early fifteenth-century iconography, artists and artistic schools in Paris long ago identified

⁹ Froissart manuscripts are today housed in libraries across two continents, from Texas to Toulouse.

¹⁰ David Cooper, and Colin Dunn (Scriptura Ltd).

the hand responsible for the primary decoration of Besançon BM, ms. 864 and Paris, BnF f. fr. ms. 2664 as being that of a disciple of the Rohan Master. The artist responsible for the miniatures of Besançon BM, ms. 865 and Paris, BnF f. fr. ms. 2663, on the other hand, was for many years thought to be a mediocre follower of the Master of the Berry Apocalypse, so called after a copy of the Apocalypse illustrated for the Duke of Berry and housed today at the Pierpont Morgan Library in New York (under ms. shelfmark M.133)¹¹. More recently, Inès Villela-Petit has argued that the artist, properly identified as the Boethius Master, deserves to be clearly distinguished from the Master of the Berry Apocalypse (“Deux Visions”). The illustrations to Besançon, Bibliothèque Municipale, ms. 864, meanwhile, are to be attributed to a forerunner of the Rohan Master, named the Giac Master for having illustrated a Book of Hours for Jeanne du Peschin, dame de Giac (Villela-Petit, “Les Heures”). DID-ARQ’s starting point, therefore, was four manuscript volumes whose miniatures, initial letters and decorative borders were entrusted to two artists’ workshops: those of the Giac and Boethius Masters. These artists were particularly favoured by bookseller De Liffol: their handiwork, and the often elegant penmanship of the scribes given the job of copying the texts, are an eloquent testimony to the remarkable activities of book trade artisans in Paris during the first quarter of the fifteenth century. The research questions that Illinois elected to explore included the following:

1. How does the application of computer algorithms to the analysis of portrayal of the human face in the manuscript miniatures help scholars to refine the parameters of discriminating features traditionally used by art connoisseurs for

¹¹ This was for many years the view imposed by the eminent scholarship of Millard Meiss (chap. XI, p. 360 et sq.).

characterising the distinctive handiwork of individual artists such as the Giac and Boethius Masters?¹²

2. What do these computer techniques and their application to the image data reveal about the hands responsible for secondary decoration (e.g. initials and marginal decoration) of these manuscripts?

3. To what extent might our new e-Science techniques assist scholars to refine current knowledge of the human presence behind such broad-brush labels as ‘Giac Master’ or ‘Boethius Master’?

4. Do our procedures suggest the presence and activity of more than one individual actively at work behind these labels?

The Sheffield operation, meanwhile, concentrated on the copying process and its outcome: the writing or script constituting the narrative of the four volumes. This extensive text was copied from exemplars by at least two teams of scribes to whose workshops the unbound quires were sent by Pierre de Liffol.¹³ Scholars tend to call these scribes ‘A’, ‘C’ or ‘G’, so anonymous are they for the most part.¹⁴ Scholars have

¹² The scientific findings of our Illinois colleagues are to be published separately (paper submitted to *Digital Humanities Quarterly*); in broad terms, two key iconographical elements generally used by art historians as an index of authorship were identified for scrutiny: (i) the faces of queens, kings and other figures within the illuminations; and (ii) representations of armour.

¹³ Unbound quires were also sent out to the illustrators; their circulation and the ‘piece work’ character of their gradual preparation prior to receiving the attentions of bookbinder and bookseller are key aspects of the book culture of this period, as evidenced in *Patrons, Authors and Workshops. Books and Book Production in Paris around 1400*, G. Croenen and P. Ainsworth (eds), Peeters, “Synthema” 4 (Louvain Paris- Dudley MA, 2006). See also Anne D. Hedeman, *Translating the Past. Laurent de Premierfait and Boccaccio’s De casibus*, J. Paul Getty Museum (Los Angeles, 2008).

¹⁴ There are of course notable exceptions; see for instance M.-H. Tesnière, ‘Les manuscrits copiés par Raoul Tainguy : un aspect de la culture des grands officiers royaux au début du XV^e siècle’, *Romania* 107 (1986), pp. 282-368. Wherever possible, the transcriptions of the *Online Froissart* include marked-up indications of shifts in hand from one copyist to another (see also: *Online Froissart*, “Apparatus”, Codicological Descriptions).

none the less gradually built up an idea of particular scribal ‘personalities’ by singling out and describing the distinguishing features of their particular hands, and have been able in this way to adduce significant characteristics on which to found conclusions: particular ways of executing certain sequences of letters, such as the ligatures ‘th’ and ‘ch’, the ending ‘-ent’, or certain ways of writing the letters ‘a’ and ‘r’ (against the models furnished by contemporary bookhands taught to all apprentice scribes). DID-ARQ’s archive comprises 10 virtual manuscripts, each of which contains 300,000 words or more, all copied in a bookhand known to specialists as *littera cursiva libraria*. Early in 2011 the Sheffield DID-ARQ team began to test some of the hypotheses just referred to against this digital dataset, applying algorithms referred to below. The key research question (broadly formulated here) underpinning research theorization and practice in this domain was:

How might one adduce pertinent e-Science methodologies for the interrogation of such a large-scale database, the better to explore, characterise and circumscribe several particular manifestations (individual scribal hands ‘A’, ‘C’ or ‘G’...) of an attested early 15th-century bookhand (*littera cursiva libraria*) (see Stiennon 284-5; Brown)?

Initial palaeographical study suggested that it might prove significant to compare letter and word clusters (across our many hundreds of digitized folios) of writing by a postulated scribe ‘C’, for instance, using semi-automated definition of the perimeters of the letter and word shapes; this, it was thought, should help to generate augmented – and more objective – evidence towards the assignment to a particular scribe of responsibility for a given section of a given manuscript (‘X’). Once that

scribe's manual 'idiolect' had been so defined, it should become possible to search for his/her activity in other manuscripts ('Y' or 'Z'). There was already some preliminary evidence to suggest that this was happening across the codices of the 'Pierre de Liffol' corpus, but it was our expectation that finer and more accurate methodologies based on the virtual manuscripts might confirm the hypothesis and account for it on the basis of more scientifically convincing evidence. Other potentially interesting areas for electronic investigation might include the *ductus* of the written text (the movement and direction of hand and pen as they make their upward and downward strokes and curves) as realised by a particular scribe, the overall neatness of the particular folio, the characteristic recourse by a scribe to abbreviations, and his or her recurring deployment of consistently similar spelling patterns. Ways to explore these electronically are still under review. The Sheffield DID-ARQ team has begun its investigations by attempting to extract a 'digital fingerprint' from the data using Polygonal Models and Shape Recognition. Sobel edge detection (from NCSA's Image2Learn API) was applied to source images accessed from commonly-shared samples mounted on the consortium's Medici image library. Line segments were then fitted to edge map data using the expectation-maximization (EM) algorithm. Shape recognition algorithms were subsequently applied to polygonal models to identify letters, words, symbols and patterns. The algorithms will be run in due course on the NCSA's Petascale High Performance Computing (HPC) machine, Blue Waters, or on the NSF (National Science Foundation) Teragrid HPC resources, since the computation takes many hours of computing time. At time of writing, results are beginning to emerge which appear to offer some solid confirmation of the potential

for identifying a steadily more objective digital fingerprint for our hitherto rather shadowy medieval copyists.¹⁵

All of the above has only been made possible because of digital resources originally obtained for the *Online Froissart* (Ainsworth and Croenen). Digital photography of the Stonyhurst and Toulouse manuscripts was supported by a Leverhulme Research Fellowship (Peter Ainsworth, 2005-2006). Grants were also obtained from Yorkshire Universities Gift Aid (2006) and the Worldwide Universities Network (2007 and 2008) to support digitisation of manuscripts at the Bibliothèque Royale, Brussels, and the Bibliothèque nationale de France, Paris. In 2010 the Harry Ransom Research Center at Austin, Texas, and the Bibliothèque Royale Albert 1er, Brussels, kindly gave their permission to the project to display their digital images (respectively Austin ms. 48 and Brussels BR mss II 88 and IV 251) alongside those from Besançon, Stonyhurst and Toulouse.

The manuscript tradition of the *Chronicles* is a particularly rich quarry for research into many aspects of the period (social, political and military history, book production, literature and art history), but research on the manuscripts has to date been hampered by difficulties in comparing the original materials, dispersed to libraries across Europe and in the USA. The *Online Froissart* offers access to the manuscript tradition of the first three Books of Froissart's *Chronicles*. It delivers

¹⁵ A workshop scheduled for 06 June 2011 at Sheffield's Humanities Research Institute is set to explore the latest findings and results: <http://www.sheffield.ac.uk/content/1/c6/12/03/81/DID-ARQ%20Workshop%20Programme.pdf>

complete or partial transcriptions of all 113 surviving manuscripts containing these Books, a new translation into modern English of selected chapters from each of these, several complete high-resolution reproductions of illuminated manuscript copies, and a range of secondary materials. These include codicological descriptions, an index of persons and places, historical and textual commentaries, scholarly essays, a glossary and some commentaries on the manuscript illustrations. Also provided are a number of advanced tools with which to unlock the riches of the resource: a collation tool allowing word-by-word comparisons across witnesses, a search engine for simple and more complex queries, and a transcription viewing mode allowing users to go straight to definition entries in the online *Dictionnaire du Moyen Français*.¹⁶ The DMF benefits in return from the very considerable lexicological material provided via the *Online Froissart's* transcriptions. Additional bilateral collaboration is underway towards the production of fully lemmatised glossaries.

Another innovative feature is the *Online Froissart's* dedicated manuscript viewer (*Virtual Vellum*, developed by Peter Ainsworth and Mike Meredith) for manipulating the electronic facsimiles. From May 2011 users will be able quickly to pinpoint episodes from each of the three Books of the *Chronicles*.

The *Online Froissart* project started in October 2007 with two complementary teams based at the Universities of Sheffield and Liverpool, led by Peter Ainsworth and Godfried Croenen. It was launched on 31 March 2010 at the end of more than two years' intensive work and is updated regularly (the latest update being scheduled for late May 2011). Work continues beyond the funded period, with

¹⁶ Thanks to a programme of workshops held over the period 2008-2011 and jointly funded by the British Academy and Centre National de la Recherche Scientifique. The project brought together the *Online Froissart* (Universities of Sheffield and Liverpool), *Christine de Pizan Queen's Manuscript* (Universities of Edinburgh and St Andrews) and the *Dictionnaire du Moyen Français* (Université de Nancy 2, Laboratoire ATILF: Analyse et Traitement Informatique de la Langue Française). Project title: 'Middle French and Other Medieval Vernacular Dictionaries'.

fresh material being added on each occasion. The project website is hosted by the Humanities Research Institute at the University of Sheffield,¹⁷ while the online collaborative environment and associated tools used to share content developed at both universities are hosted at the University of Liverpool.

Much of the expertise underpinning the *Digging into Data* project discussed at the beginning of this paper originated from research and development begun on the *Online Froissart*, which in turn benefited from Leverhulme and other grants supporting the high-resolution photography required. Neither of these electronically-empowered projects would have happened, however, without the primary resources (the real manuscripts of parchment, paper and ink curated by their equally real conservators and librarians), intensive codicological and palaeographical study of which, conducted by ‘lone scholars’ (Peter Ainsworth, Godfried Croenen and their PhD students and postdoctoral colleagues) alone provided the intellectual and academic foundation for the more pervasively collaborative, interdisciplinary work described here.¹⁸ As the Arts and Humanities look to the future and its emerging new challenges, these two kinds of (complementary) academic endeavour – imperfectly termed the individual and the collaborative – need to be kept in thoughtful balance.

¹⁷ Credit for designing what is in effect an extremely complex resource is due to Jamie McLaughlin; the overall look of the site is the work of Michael Pidd (both of the Humanities Research Institute).

¹⁸ The present paper could not have been written without the input of our colleagues on the DID-ARQ project from Michigan State University, the University of Illinois at Urbana-Champaign and the National Center for Supercomputing Applications (also at Urbana-Champaign); we are extremely grateful for their support and partnership. For details of the other datasets studied within the DID-ARQ consortium (maps and quilts) see:

http://www.sheffield.ac.uk/hri/projects/projectpages/did_images/datasets.html

Bibliography

- Ainsworth, Peter and Mike Meredith. *Virtual Vellum*. Accessed 04 May 2011. <http://www.shef.ac.uk/hri/projects/projectpages/virtualvellum.html>
- Ainsworth, Peter, and Godfried Croenen, ed. *The Online Froissart*. version 1.2 (2011). Accessed 15 May 2011. <<http://www.hrionline.ac.uk/onlinefroissart>>
- Ainsworth, Peter. "Representing Royalty: Kings, Queens and Captains in Some Early Fifteenth-Century Manuscripts of Froissart's *Chroniques*". *The Medieval Chronicle IV* (Amsterdam/Atlanta: Editions/Rodopi, 2006): 1-38.
- Brown, Michelle P. *A Guide to Western Historical Scripts from Antiquity to 1600*, (London: The British Museum, 1990).
Connected Histories. Accessed 04 May 2011. <www.connectedhistories.org>
- Croenen, Godfried and Peter Ainsworth, eds. *Patrons, Authors and Workshops. Books and Book Production in Paris around 1400*. Synthema 4 (Paris: Peeters, 2006).
- Croenen, Godfried. "Le Libraire Pierre de Liffol et la Production de Manuscrits Illustrés des *Chroniques* de Jean Froissart à Paris au début du XVe siècle". *Art de l'enluminare* 31 (Paris: Editions Faton, 2009): 14-23.
Dictionnaire du Moyen Français. <<http://www.atilf.fr/dmf/>>
- Fearn, Hannah. "Research intelligence – Let's dig a little deeper". *Times Higher Education* (28 April 2011).
- Hedeman, Anne D. *Translating the Past. Laurent de Premierfait and Boccaccio's De casibus* (Los Angeles: J. Paul Getty Museum, 2008).
- Meiss, Millard. *French Painting in the Time of Jean de Berry: The Limbourgs and their Contemporaries*. (New York: Thames and Hudson, 1974).
- Meredith, Mike, with Peter Ainsworth and Tribal. *Kiosque*. Accessed 04 May 2011. <<http://www.shef.ac.uk/hri/projects/projectpages/kiosque/overview.html>>
- Old Bailey Online (Old Bailey Proceedings Online)*. Accessed 04 May 2011. <www.oldbaileyonline.org>
- Simeone, Michael, Jennifer Guiliano, Rob Kooper, and Peter Bajcsy. 'Digging into Data Using New Collaborative Infrastructures Supporting Humanities-based Computer Science Research'. *First Monday* 16.5 (2011).
- Stiennon, Jacques. *Paléographie du Moyen Âge*. 3rd ed. (Paris: Armand Colin, 1999)
Tales from the Digital Archive. BBC Radio 4. 26 April 2011. Radio.
- Tesnière, M.-H. "Les manuscrits copiés par Raoul Tainguy : un aspect de la culture des grands officiers royaux au début du XV^e siècle", *Romania* 107 (1986): 282-368.
- Villela-Petit, Inès. "Deux visions de la *Cité de Dieu* : le Maître de Virgile et le Maître de Boèce". *Art de l'enluminure* 17 (Paris: Editions Faton, 2006): 2-19.
- Villela-Petit, Inès. "Les Heures de Jeanne du Peschin, dame de Giac. Aux origines du Maître de Rohan". *Art de l'enluminure* 34 (Paris: Editions Faton, 2010): 2-25.