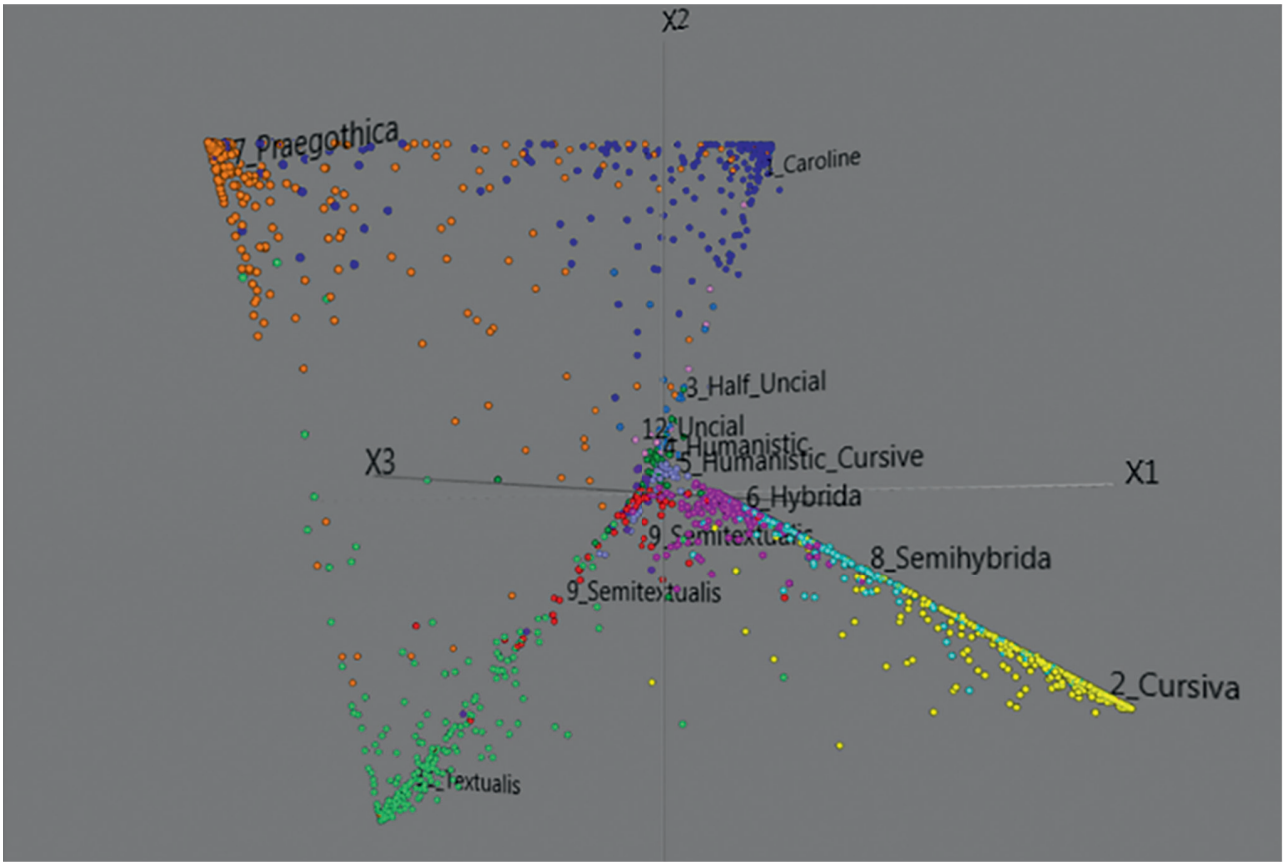
mc Nº 15 2020

manuscript cultures

Hamburg | Centre for the Study of Manuscript Cultures ISSN 1867–9617



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG
CENTRE FOR THE STUDY OF MANUSCRIPT CULTURES
DFG

Publishing Information

Natural Sciences, Technology and Informatics in Manuscript Analysis
Edited by Oliver Hahn, Volker Märgner, Ira Rabin, and H. Siegfried Stiehl

Proceedings of the third *International Conference on Natural Sciences and Technology in Manuscript Analysis* and the workshop *OpenX for Interdisciplinary Computational Manuscript Research* that took place at the University of Hamburg, Centre for the Study of Manuscript Cultures, on 12–14 June 2018.

Editors
Prof. Dr Michael Friedrich
Universität Hamburg
Asien-Afrika-Institut
Edmund-Siemers-Allee 1/ Flügel Ost
D-20146 Hamburg

Tel. No.: +49 (0)40 42838 7127
Fax No.: +49 (0)40 42838 4899
michael.friedrich@uni-hamburg.de

Prof Dr Jörg Quenzer
Universität Hamburg
Asien-Afrika-Institut
Edmund-Siemers-Allee 1/ Flügel Ost
D-20146 Hamburg
Tell. No.: +49 40 42838 - 7203
Fax No.: +49 40 42838 - 6200
joerg.quenzer@uni-hamburg.de

www.csmc.uni-hamburg.de

Editorial Office
Dr Irina Wandrey
Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstraße 26
D-20354 Hamburg
Tel. No.: +49 (0)40 42838 9420
Fax No.: +49 (0)40 42838 4899
irina.wandrey@uni-hamburg.de

Layout
Miriam Gerdes

Cover
Image of ICDAR2017 Tensmeyer's distance matrix (axes 2 and (1 and 3)), see article by Dominique Stutzmann, Christopher Tensmeyer and Vincent Christlein in this volume.

Translation and Copy-editing
Amper Translation Service, Carl Carter, Fürstenfeldbrück

Print
AZ Druck und Datentechnik GmbH, Kempten
Printed in Germany

ISSN 1867–9617

CONTENTS

ARTICLES

Article

On Digital and Computational Approaches to Palaeography: Where Have we Been, Where Are we Going?

Peter A. Stokes | Paris

Although palaeographers have been using technological developments since the discipline's beginning, nevertheless the last decade has seen an enormous increase of and very rapid advances in computational approaches to the field. This seems likely to prove highly transformative and disruptive and has not gone unnoticed by palaeographers. Indeed, debate has been ongoing for many years now about the nature of palaeography as an 'art' or a 'science', as well as the role (if any) of quantitative or other 'hard' evidence in palaeographical studies.[1] Originating over a century ago, if not longer, this debate has resurfaced with the ready availability of digitised images of manuscript pages, along with relatively powerful personal computers and developments in machine vision, artificial intelligence, and related topics. At the same time, recent years have also seen increasing interest from computer science and related fields; this is demonstrated by increasing recognition of historical materials at conferences such as the International Conference for Handwriting Recognition (ICFHR) and the International Conference for Document Analysis and Recognition (ICDAR), by centres like the CSMC, and by the recent creation of a chair in digital and computational humanities applied to the study of historical writing in France. Although fears of a split into 'digital' and 'non-digital' palaeography now seem thankfully unwarranted,[2] nevertheless real questions remain about how the many different approaches can and should best be integrated, ensuring that all parties gain through fruitful dialog and that none simply supplies services or 'answers' to the other. The objective of this paper is therefore to review these developments, focussing not on technical advances *per se* but rather on the 'view from the (Digital) Humanities', making comparison with previous reviews to suggest some points around where we have been and where we might be going.[3]

1. A (very) brief overview of digital and computational work to date

Significant work in the last decade or more has been done on topics such as automatic analysis of handwriting for dating, localising and (especially) writer identification, script classification, and layout analysis. This work has typically drawn on the very large volume of digitised material and has developed and applied methods in machine vision and artificial intelligence to their analysis. Recently, good progress has been made in fields such as line detection, Handwritten Text Recognition (HTR), wordspotting of handwritten documents, as well as automatic layout analysis and the dating of script.[4] More specifically aligned to palaeographical research is work on the characterisation of script.[5] Success has also been achieved in aligning images of text to pre-existing transcripts, as well as the identification of fragments of manuscripts from the same original document, and the identification of specimens of script likely to have been written by the same individual.[6] Also relevant here are methods in Machine Vision such as automatic feature extraction, whereby the software determines automatically what aspects of the image are significant. This raises the possibility that such features could be meaningful and useful

[1] The discussion has been summarised briefly by Stokes 2015; relatively recent contributions include Costamagna et al. 1995 and 1996; Gumbert 1998; Derolez 2003, 1–9; Ciula 2005; Sirat 2006, 309ff and 493; Canart 2006; Davis 2007; Schomaker 2007; Stokes 2009; Stokes 2012; and Stutzmann 2011, among others.

[2] For an expression of these concerns see Stokes 2009.

[3] Previous reviews include Hassner et al. 2013 and Hassner et al. 2015.

[4] Examples of such work are given by Sudholt and Fink 2016, applied to Cuneiform by Rothacker et al. 2015, and recent work for the project HIMANIS; Chen and Seuret 2017, Seuret, Stökl ben Ezra and Liwicki 2017, and Xue Yu's recent success at ICDAR 2017; Christlein, Gropp and Maier 2017; and Kestemont, Christlein and Stutzmann 2017, respectively.

[5] See, for example, Kestemont, Christlein, and Stutzmann 2017.

[6] See, for example, Fischer et al. 2011a and 2011b; Stutzmann et al. 2015; Sadeh et al. 2015; Wolf et al. 2011a; Wolf et al. 2011b; and work by Schomaker et al. such as He, Wiering and Schomaker 2015, respectively.

to palaeographers, perhaps even highlighting new areas on which to focus. Indeed, there are standard techniques in machine vision which can be helpful here, such as the so-called 'Bag of Words' which automatically builds a set of visual elements which make up the image; these visual elements are established by the machine and so may not be meaningful to scholars, but in principle they are very similar to the graphic elements referred to by palaeographers such as 'wedge', 'foot-serif', 'hook' and so on, and some work has been done on making these automatic 'visual words' meaningful to palaeographers as well.[7] Similarly, Deep Learning has recently been applied to the classification of medieval script (among many other things), and not only was this technique successful in itself, but it was also possible to 'look inside' the system to see what elements of the script it found distinctive.[8]

Another important development is the degree to which these methods and techniques are becoming freely available for use by other projects, through free or Open Source software but also through web APIs, a noteworthy example of which is the suite provided by the DIVA group at the University of Fribourg.[9] This is potentially a significant boon to palaeographers with some understanding of digital methods but without the resources or expertise to implement their own code, and it may also go some way towards addressing the need for benchmarking and standard algorithms.[10] This refers to the need in computer science to have standard sets of data which different groups can use in order to establish meaningful comparisons between algorithms. However, the very existence of these benchmarking datasets tends to encourage people to produce algorithms that are suited to that data, and so it is important that the datasets themselves give a good representation of the range of manuscript material. For this reason increasing numbers of such datasets are becoming available to cover a wider range of historical material, as for instance in competition datasets produced by the IRHT for use at ICDAR.[11] This is badly needed, as the 'standard' datasets that are typically used for benchmarking are those such as the George Washington Papers which comprise images of documents written by the famous eighteenth-century general.[12] Although a useful example, his writing bears very little resemblance at all to medieval manuscripts from five hundred or a thousand years earlier.

In addition to this, important work is also continuing on how to better present and interact with palaeographical knowledge in tractable ways through symbolic representation and modeling, interface design and UI/UX, visualisation, and so on. An example of this is DigiPal and its successors, particularly Archetype, which focussed on structured descriptions of handwriting which researchers could enter into software. This involves manually drawing annotations on images of handwriting and entering descriptions of the letters. Because the database already contains information about the components or essential elements of letters (for instance that **b**, **h** and **l** all contain ascenders), it therefore becomes very easy for researchers to find and compare forms in ways that are palaeographically meaningful (for instance searching for examples of ascenders by a particular scribe or from a particular region). The emphasis is firmly based on knowledge creation through experimentation, exploration and visualisation, as well as the communication of evidence to support the resulting argument.[13] This approach has proven successful insofar as the methods and results are now increasingly used by palaeographers and others, and have been successfully applied to many different writing systems.[14] This aspect of discovery, analysis and communication relates directly to larger questions in Digital Humanities and beyond about how one represents expert knowledge in systems that are tractable to the computer, connecting to areas and technologies such as ontologies, formal modelling, Linked Open Data and the Semantic Web. In this respect other important projects on modelling manuscripts and documents are relevant here, particularly ORIFLAMMS but also

[7] Examples here include Wolf et al. 2011; Hassner et al. 2015; and Kestemont, Christlein and Stutzmann 2017.

[8] Kestemont, Christlein and Stutzmann 2017; compare also Rajan Sampath 2016 for a different approach.

[9] Eichenberger et al. 2015; Garz et al. 2015; Würsch, Ingold and Liwicki 2016; other examples include Sudholt and Fink 2016 and Kestemont, Christlein and Stutzmann 2017.

[10] Hassner et al. 2013, and see also discussion below.

[11] Kestemont, Christlein and Stutzmann 2017.

[12] Washington Database 2016, and compare those in Kestemont, Christlein and Stutzmann 2017. For this need see also Hassner et al. 2015, 118, among others.

[13] Brookes et al. 2015.

[14] Brookes et al. 2015; Stokes 2017.

Europeana, Biblissima, IIIF, CRMtex and others.[15] These tend to focus on the document more than the script, but some formal ontologies are starting to emerge for describing writing itself. ORIFLAMMS, for instance, established important groundwork towards a substantial ontology for script, work which is still continuing.[16] CRMtex adds a small extension to the CIDOC-CRM model for museums and objects: it comprises entities TX1 Written Text, TX2 Writing, TX3 Writing System and TX4 Written Field, as well as properties such as TXP1 Used Writing System.[17] More complex are the ontologies developed for the IDIOM project to model Mayan hieroglyphics which include not only the distinction between graphs and signs but also represent graphs that are derived from other graphs, allographic relationships between signs, reading hypotheses through sign functions, confidence levels of different readings, and so on.[18] Although not (yet) expressed as a formal ontology, a further model which focuses more on the graphical aspect of writing and written communication is that developed for the DigiPal project which includes structural relationships between allographs and graphs, relations between different graphs written by the same scribe, and so on.[19] Similarly, models and ontologies for manuscripts as objects have been developed as part of several projects. One of these is an event-based model for Hebrew manuscripts which seeks to model the object's history and its cultural context.[20] Another ontology has been published by the Biblissima team which extends FRBRoo with entities such as Binding, Annotation, Digital Surrogate, Electronic Edition and more, with relationships such as Used as Source, Used as Copy, Used Style of Script, and so on.[21] IIIF, on the other hand, is less an explicit model and more a standard for interoperability of images, including not only the images themselves but also regions, basic manipulation such as rotating, but also relationships between images such as sequences of images representing consecutive pages in a book.[22]

This is of course a small overview of just some of the work that is underway towards modelling books and writing, but the promise of such developments is starting to be seen. Although by no means using the most advanced computational methods, IIIF in particular is nevertheless transforming the way in which scholars in the Humanities are working with manuscripts today. The projects together provide (among other things) stable protocols for addressing and manipulating images of manuscripts and other cultural heritage over the Web, in a system which is becoming increasingly widely used by libraries, archives and other cultural heritage institutions. This means that we are now able to refer unambiguously to images of manuscript pages and to regions in those images, and to access the images directly from many different repositories. For the Humanities researcher this means access to material and – really for the first time – the ability to easily compare images from different institutions in the same software. It also responds very directly to the need for ready and open access to data which has been noted in the Dagstuhl events and elsewhere.[23]

2. Some Continuing or Future Directions?

The necessarily very brief survey above points to some of the main directions in recent research, at least as seen by the present author. It also alludes to some issues and gaps, many of which have been recognised for some time but which still require further work. Without claiming completeness or indeed originality, the following issues are relevant to the Humanities and seem to this author likely to become increasingly important in the near future

2.1 Algorithmic Accountability

In terms of future developments, an important question that has often been raised with regards to computational methods and has long been discussed in Digital Humanities is the need to be able to interpret and understand algorithms and their approaches.[24] This approach formed the premise of DigiPal

[15] For examples see Stutzmann 2013 (as well as Stutzmann 2012); Biblissima [n.d.]; IIIF [n.d.]; Murano and Felicetti 2017; Felicetti and Murano 2017; and Zhitomirsky-Geffet and Prebor 2016.

[16] Stutsmann 2013, 85 and 89–90. For the the ontology by the end of the project see Stutzmann 2017, 17.

[17] Murano and Felicetti 2017; Felicetti and Murano 2017.

[18] Gronemeyer and Diehr 2018.

[19] Brookes et al. 2015. 'Allograph' and 'graph' are here used in the palaeographic/forensic sense, where 'allograph' refers to different ways of writing the same letter (such as roman **a** and italic *a*), and 'graph' refers to a single instance of a letter on the page. See further Davis 2007, 254–55.

[20] Zhitormsky-Geffet and Prebor 2016.

[21] Biblissima [n.d.].

[22] IIIF [n.d.].

[23] Hassner et al. 2013; Hassner et al. 2015.

[24] For further discussion see Ciula 2005; Davis 2007; Sculley and Pasanek 2008; Stokes 2009; Clement, Steger and Unsworth 2009; among others.

(now Archetype), the basis of which was the insight that although techniques in machine vision and machine learning are increasingly successful, they typically are opaque to researchers, and are therefore difficult to interpret or verify.[25] However, if palaeographers cannot interpret the statistical results or understand their limitations then it becomes difficult or impossible to use the software. Furthermore, many of the computational systems listed above must be trained on a 'Ground Truth', namely a relatively large corpus of material that is taken to be known and without doubt, but the very assumption that such a ground truth exists is problematic in palaeography, where the quantity of known material is relatively small and where so much doubt often surrounds this content.[26] Rather than going away, the problem of understanding, inherent bias and therefore trust in algorithmic approaches has instead become increasingly prominent and is now widely recognised, particularly under the rubric of 'algorithmic accountability'. The Association for Computing Machinery (ACM) US Public Policy Council and Europe Policy Committee, for instance, published a joint 'Statement on Algorithmic Transparency and Accountability' which notes in words similar to those of Davis from a decade earlier that 'there is … growing evidence that some algorithms and analytics can be opaque, making it impossible to determine when their outputs may be biased or erroneous'.[27] The ACM Council goes on to observe that '[d]ecisions made by predictive algorithms can be opaque because of many factors, including technical (the algorithm may not lend itself to easy explanation), economic (the cost of providing transparency may be excessive, including the compromise of trade secrets), and social (revealing input may violate privacy expectations). Even well-engineered computer systems can result in unexplained outcomes or errors, either because they contain bugs or because the conditions of their use changes, invalidating assumptions on which the original analytics were based.' Although work on historical documents does not have the societal implications of the cases discussed by the ACM, nevertheless machine-generated features and highly computational methods are often not meaningful to humans and particularly to those in the Humanities. This difficulty is further compounded by the financial pressures that surround this work, meaning that the underlying software and algorithms must often be proprietary and so become commercial secrets rather than being fully Open Source. This is less of a problem if the results are easily verifiable by a human observer: in the case of HTR, for instance, it is normally very easy to see if the machine's results are valid or not, and in this context how the results are obtained is relatively unimportant. However, the problem is significantly greater in cases where judgement is required, such as whether two samples are by the same scribe or not, or indeed in a broader context what sentence should be delivered to an offender, and in this case it is essential that one can 'look inside the machine' in order to understand and judge the validity of underlying assumptions and the impact of the biases that inevitably underlie them, without which these results cannot sustain the necessary scepticism and critical questioning that is the foundation of the academic and scientific method. Important progress is already being made here but more work remains to be done.[28] The problem of 'looking inside the machine' is famously difficult, particularly for very deep neural networks. Even if we can see the final layer(s) this is only a small part of the whole computational process, and it remains very unclear how we can go about understanding the process more fully. As always, there is also still the risk of misunderstanding, particularly for those without expertise in the algorithms: statistics, visualisations and other information can be misleading, for instance by misunderstanding or misusing the confidence levels in a classifier. Granted this opacity is nothing new, as human palaeographical expertise has long proven very difficult to articulate, with some even arguing that it cannot be taught but can only be 'acquired' through years or decades of exposure to the materials.[29] It does, however, raise the question to what extent these new computational methods are simply reproducing the problems of the old manual ones and how, if at all, we can get beyond this.

[25] Davis 2007 n. 27; Stokes 2009; Hassner et al. 2013; Stokes 2015; and Hassner et al. 2015, 113–4, where the question is usefully rephrased as one of trust rather than the 'black box'

[26] Hassner et al. 2013.

[27] ACM 2017.

[28] See especially Kestemont, Christlein and Stutzmann 2017 for recent work applied directly to handwriting. A useful discussion of these systems from the point of view of a palaeographer is also given by Smit 2011.

[29] For this argument see, for example, Pfaff 1977, 104, discussed by Derolez 2003, 1.

2.2 Interdisciplinarity and combining approaches

Closely related to accountability in manuscript studies is the field's inherent interdisciplinarity. This has already been discussed at length elsewhere[30] but always needs emphasising: that manuscripts are extremely complex objects that can be approached along infinitely many dimensions, whether textual (philological, literary, linguistic, historical), cultural (as art, as status-object, as reflection of learning and world-view), visual (palaeographical, art-historical) or material (codicological, archaeological), from the 'hard' sciences (chemistry, physics, biology), as images (machine vision, pattern-matching), as three-dimensional objects (3D scanning, Reflectance Transformation Imaging), and many, many more. This – thankfully – has led to many interdisciplinary studies of manuscripts, many of which have been highly successful. Such work in turn inevitably leads to challenges in communication and, more significantly, in finding true collaboration such as research questions that are relevant to all parties and finding an appropriate balance of power and trust.[31] For instance, in practice 'collaboration' can in fact be one group providing 'solutions' to the other group's 'problems'. This carries the implication that this is a one-way process, with one group holding the power and having nothing to learn from the other, a risk that is not far from one identified by Gillian Beer, namely 'becoming merely disciples because not in control of a sufficient range of knowledge'.[32] The converse is also a risk, and is also observed in practice, namely that 'the problems preoccupying those working in another discipline may sometimes (initially, arrogantly) seem quite simple – because we are not familiar with the build-up of arguments across time that has reached this moment of dilemma'.[33] Fortunately the experience of this author has been that both of these risks have largely been avoided in the case of manuscript studies, not least because so many in the field have committed to the hard work that is required for true collaboration, and so fears raised by the author in 2009 of the possible split into two disciplines – 'digital' and 'non-digital' palaeography – seem happily misguided, but nevertheless this work must necessarily continue. In either case, though, is the question of trust once again, which Beer phrased as 'the matter of competence … Is this a raiding party? Is there time to question and to learn? *How much must be taken on trust?*'[34] Although perhaps not an issue of trust or the 'raiding-party', nevertheless even different computer-based approaches rely on different types of expertise, each of which (usefully) has different strengths and weaknesses and is therefore better adapted to different questions. For instance, research in machine vision, clustering and large-scale statistical approaches have produced very good results with relatively large amounts of data and more-or-less established Ground Truth, giving a type of palaeographical 'distant reading' or helping one find a way into a large and complex corpus. Fields like knowledge representation and knowledge elicitation seek to model and represent palaeographical content in 'symbolic' ways that are closer to those of the domain experts, tending towards more of a 'close' reading of a defined and relatively manageable corpus. User Interface and User Experience (UI/UX) is another relevant domain, looking more at visualisation and communication with the user, and so on. Relatively little attention has been paid to interface design in the field of document analysis, but there are obvious areas of potential difficulty here, for instance an interface reflecting the model from one discipline being unfamiliar, difficult to understand or even misleading from those of other disciplines. Such difficulties are often observed in complex search forms which might reflect the underlying data structure but which do not necessarily reflect the users' understanding of their materials, for instance. In this sense, then, the interface risks perpetuating disciplinary and conceptual differences, rather than helping to bring disciplines together. Furthermore, it is very often the case that off-putting, unfamiliar or inefficient interfaces are not tolerated by target users: rather than learning to use seemingly difficult systems, people tend to simply abandon the software and go back to their previous practices. Once again, then, combining these different domains of research is an important step that needs to continue, even if the disciplinary nature of academic research and its need for specialised research and publications may sometimes make this difficult in practice.

[30] See especially Hassner et al. 2013 and Hassner et al. 2015 as two examples among many.

[31] This has been discussed particularly by Hassner et al. 2013, and Hassner et al. 2015, among others.

[32] Beer 2006.

[33] Beer 2006.

[34] Beer 2006; my emphasis.

2.3 Multigraphism and a Conceptual Reference Model

Another recurring question is the need for a precise, transversal ontology and conceptual reference model (CRM) for handwriting. The need for this has been recognized for some time,[35] but its importance is becoming increasingly evident as research towards digital resources for the field is largely resulting in different and incompatible models. The advantages of sharing and interconnecting palaeographical data are evident, but at the same time the likelihood of a single universal terminology is remote and not necessarily desirable: as Petrucci argued, every palaeographical terminology is based on a particular view of scribal practice and all such views are potentially valid,[36] and this argument suggests we should encourage rather than limit the proliferation of analytic frameworks. Nevertheless, the need for a consistent and standardised way of at least referring to manuscripts is evident, and indeed significant progress has already been made here.[37] However, as work to date is increasingly demonstrating, a successful model needs clarification of more fundamental concepts. What *is* writing, after all, and how can we accurately distinguish between its many aspects: visual, stylistic, phonological, functional, chemical, mechanical, semantic, kinetic, graphematic, allographic, and so on? What *exactly* is a script, and how does one define the boundaries between one script or another, or decide if two samples of writing are in the same script or not? As discussed in Section 1 above, some work is beginning to appear in this area but much more is needed. A true CRM should also account for different writing systems, and indeed interest in transversal and multigraphic approaches is growing rapidly: how, for instance, can one compare Arabic and Hebrew written by the same scribe? All this suggests the need for a formal CRM that can provide formal precision while functioning meaningfully across all the world's writing systems. Clearly it is not possible to create a complete model that will fully capture all aspects of all the world's written communication: indeed this is not even necessarily desirable, as one of the values of modelling is that it simplifies, allowing one to focus on particular aspects which are intractable in the domain being modelled.[38] The question, rather, is whether it is possible to create a model that is sufficiently broad to be useful to palaeographers and others and to allow genuine transversal study of scripts. The answer to this question is unclear at present. To raise just some of the challenges that it presents, researchers in both palaeography and computer science are now recognising the difficulties of multigraphism, namely cases where individuals or cultures simultaneously use entirely different scripts, alphabets, or even writing systems. Most computational methods have been developed at least theoretically independently of any given script or writing system, and examples in practice include the same software working successfully across Hebrew, Tibetan, Old English and Greek, for instance, or Word Spotting in Latin and Arabic.[39] In these cases, however, the work has generally been applied to only one script at a time, rather than to different scripts simultaneously. However, many – perhaps almost all – cultures wrote and still write in different scripts and alphabets or writing systems at the same time. This problem is beginning to be recognised from informatics and also from palaeography.[40] Indeed, this is becoming so much a 'hot topic' that a competition on the subject will be run at ICFHR 2018, and at least one large grant on the subject has recently been submitted for funding by the European Research Council. However, the challenges are significant, particularly in modelling, as (for instance) it requires a clear distinction between a letter's normal physical form (the allograph), its given instance on the page (the graph), but also its function whether graphematic, phonetic, semantic and so on. For instance, two given forms that look identical such as **H** and **H** would normally be assumed to represent the same grapheme but may not: if we write **H**ABEAM and I**H**ΣOYΣ then it becomes clear that the first **H** is the Latin letter (Unicode U+0048) and the second **H** a Greek Eta (Unicode U+0397).[41] A scholar interested in palaeographically comparable forms in a Greek-Latin multigraphic context would presumably want to find examples of both allographs, and this requires the capacity to search by form rather than graphematic function.

[35] See, for example, Stutzmann 2011 and Hassner et al. 2013.

[36] Petrucci 2001, 70–71.

[37] Examples of relevant projects addressing this question include MESA, Biblissima, Pinakes, TRAME, SNAP DRGN and Medium, among others.

[38] McCarty 2004.

[39] Hassner, Wolf and Dershowitz 2013, and Sudholt and Fink 2016, respectively.

[40] Examples from informatics include Singh et al. 2014; He, Wiering and Schomaker 2015; Bertolini, Oliviera and Sabourin 2016; Ahmed et al. 2017; Ubul et al. 2017. Discussions from palaeography include Petrucci 2005, p. 53; cf. also Cavallo 1990 and Smith 2004, p. 429; de Robertis 2012, p. 223; de Robertis 2013, p. 17; cf. also Ceccherini 2012, 237 (but contrast that of De Gregorio 2000, pp. 19–2); Dewartes 2013; Radiciotti 2006; Stokes 2017; and Stokes 2018.

[41] Stokes 2018, drawing on Burgarski 1993.

In other cases, however, scribes might use a Greek Φ in an otherwise Latin context for the phonetically equivalent **F**, or the Latin symbol **&** (meaning *et*) for the semantically equivalent *and* in an English context or *ocus* in Old Irish. These complexities multiply significantly in other more complex writing systems, for instance logosyllabographies such as Sumerian or Mayan where signs can function semantically or phonetically depending on context, and where signs can take significantly different forms, even incorporating elements of one sign within another, and so on.[42]

3. Concluding notes

As mentioned above, it must be emphasised once again that the preceding discussion is just one of many views on the field, from one of many possible directions. Certainly there is much to do, and many challenges still to face; indeed, it still seems largely the case that digital and computational approaches to palaeography have yet to reach their full potential in terms of having tangible impact on new results in palaeographical scholarship. Nevertheless the feeling of this author is one of optimism. There is much evidence of genuine and productive collaboration, with many examples of people making a real effort to listen, understand, learn from and gain from each other. The genuinely collaborative workshops to date have shown that the combination of digital/computational with palaeography is not one of providing tools for use or answers to questions in a one-way and uncritical flow, and neither is it providing ready access to texts or visual reproductions of manuscript pages. Rather, it is profoundly changing the way that people think in palaeography and – I hope – in computer science and informatics as well. What does it mean to work with an image instead of the object, and how do the results of our research change as a result? How does a 'distant' view of a manuscript or corpus change our view of it, and how much does this matter? What does the instability of text and writing mean for computer classification? Once again, what *exactly* is a letter anyway, and what is its relation to the texts, images, and physical and digital objects that transmit it? Asking new questions, seeing new points of view, and reflecting on these and how they change our very world-view: these are the real gains of Digital Humanities, and this is where we need to focus in future.

[42] Daniels 1990; Gronemeyer and Diehr 2018.

REFERENCES

ACM US Public Policy Council and Europe Policy Council (2017), *Statement on Algorithmic Transparency and Accountability*. <https://www.acm.org/binaries/content/assets/public-policy/2017_joint_statement_algorithms.pdf>.

Ahmed, Ahmed Abdulla, Harith Raad Hasan, Fariaa Abdalmajeed Hammed, and Omar Ismael Al-Sanjary (2017), 'Writer Identification on Multi-Script Handwritten using Optimum Features', *Kurdistan Journal of Applied Research*, 2: 178–185. doi: 10.24017/science.2017.3.64.

Archetype (2018), *Archetype: An Accessible and Evolving Toolset to Empower Digital Research* (London: King's College). <https://www.archetype.ink>.

Beer, Gillian (2006), 'Dame Gillian Beer's Speech on the Challenges of Interdisciplinarity' (Durham: University of Durham). <https://www.dur.ac.uk/ias/news/annual_research_dinner/>.

Bertolini, Diego, Luiz S. Oliveira, and Robert Sabourin (2016), 'Multi-script Writer Identification using Dissimilarity', *23rd International Conference on Pattern Recognition (ICPR)*. <https://www.etsmtl.ca/getattachment/Unites-de-recherche/LIVIA/Recherche-et-innovation/Publications/Publications-2016/Bertolini_icpr_2016.pdf>.

Biblissima [n.d.], 'Ontologie', *Biblissima Documentation*. <http://doc.biblissima-condorcet.fr/ontologie>.

Brookes, Stewart, Peter A. Stokes, Matilda Watson, and Debora Marques de Matos (2015), 'The DigiPal Project for European Scripts and Decorations', in Aidan Conti, Orietta da Rold, and Philip Shaw (eds), *Writing Europe 500–1450: Texts and Contexts* (Cambridge: D. S. Brewer; Essays and Studies n.s., 68), 25–58.

Bugarski, Ranko (1993), 'Graphic Relativity and Linguistic Constructs', in Robert J. Scholes (ed.), *Literacy and Linguistic Analysis* (Hillsdale, NJ: Erlbaum), 5–18.

Canart, Paul (2006), 'La Paléographie est-elle un art ou une science?', *Scriptorium*, 60: 159–185.

Cavallo, Guiglielmo (1990), 'Écriture grecque et écriture latine en situation de "multigrafismo assoluto"', in Colette Sirat, Jean Irgoin, and Emmanuel Poulle, *L'Écriture: le cerveau, l'œil et la main* (Turnhout: Brepols), 349–362.

Ceccherini, Irene (2012), 'Poligrafia nel Quattrocento: Sozomeno da Pistoia', *Medioevo e Rinascimento*, 26: 237–251.

Chen, Kai, and Matthias Seuret (2017), 'Convolutional Neural Networks for Page Segmentation of Historical Document Images', *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery), 101–106.

Christlein, Vincent, Martin Gropp, and Andreas Maier (2017), 'Automatic Dating of Historical Documents', in Hannah Busch, Franz Fischer, and Patrick Sahle (eds), *Kodikologie und Paläographie im digitalen Zeitalter 4—Codicology and Palaeography in the Digital Age 4* (Norderstedt: Books on Demand), 151–164.

Ciula, Arianna (2005), 'Digital Palaeography: Using the Digital Representation of Medieval Script to Support Palaeographic Analysis', *Digital Medievalist*, 1. <http://www.digitalmedievalist.org/journal/1.1/ciula/>.

Clement, Tanya, Sara Steger, and John Unsworth (2008), 'How Not to Read a Million Books', *Seminar on the History of the Book* (Rutgers University, New Brunswick NJ, 5 March). <http://www.people.virginia.edu/~jmu2m/hownot2read.html>.

Costamagna, Giorgio *et al*. (1995), 'Commentare Bischoff', *Scrittura e Civiltà*, 19: 325–348.

—— *et al*. (1996), 'Commentare Bischoff', *Scrittura e Civiltà*, 20: 401–407.

Daniels, Peter T. (1990), 'Fundamentals of Grammatology', *Journal of the American Oriental Society*, 119: 727–731. doi: 10.2307/602899.

Davis, Tom (2007), 'The Practice of Handwriting Identification', *The Library*, 7th series, 8: 251–276.

De Gregorio, Giuseppe (2002), 'Tardo medioevo greco-latino: manoscritti bilingui d'Oriente e d'Occidente', in Francesco Magistrale, Corinna Drago, and Paolo Fioretti (eds), *Libri, documenti, epigrafi medievali: possibilità di studi comparativi. Atti del convegno internazionale di studio dell'associazione italiana dei paleografi e diplomatisti, Bari 2–5 ottobre 2000* (Spoleto: Centro italiano di studi sull'alto Medioevo), 17–135.

De Robertis, Teresa (2012), 'Digrafia nel Trecento: Andrea Lancia e Francesco di ser Nardo da Barberino', *Medioevo e Rinascimento*, 26: 221–235.

—— (2013), 'Una mano tante scritture: Problemi di metodo nell'identificazione degli autografi', in Nataša Golob (ed.), *Medieval Autograph Manuscripts: Proceedings of the XVIIth Colloquium of the Comité International de Paléographie Latine* (Turnhout: Brepols), 17–38.

Derolez, Albert (2003), *The Palaeography of Gothic Manuscript Books from the Twelfth to the Early Sixteenth Century* (Cambridge: Cambridge University Press).

Dewartes, Thomas (2013), 'Polygraphisme et mixité graphique: Note sur les additions d'Arias (1060–1070) dans l'Antiphonaire de León', *Territorio, Sociedad y Poder: Revista de Estudios Medievales*, 8: 67–84.

DigiPal (2010–14), *Digital Resource and Database of Palaeography, Manuscripts and Diplomatic* (London: King's College). <http://www.digipal.eu>.

Eichenberger, Nicole, Angelika Garz, Kai Chen, Hao Wei, Rolf Ingold, and Marcus Liwicki (2015), 'DIVADesk: A Holistic Digital Workspace for Analyzing Historical Document Images', *manuscript cultures*, 7: 69–82.

Felicetti, Achille, and Francesca Murano (2017), 'Scripta manent: a CIDOC CRM Semiotic Reading of Ancient Texts', *International Journal of Digital Libraries*, 18: 263–270. doi: 10.1007/s007.

Fischer, Andreas, Emanuel Indermühle, Volkmar Frinken, and Horst Bunke (2011a), 'HMM-Based Alignment of Inaccurate Transcriptions for Historical Documents', *International Conference on Document Analysis and Recognition.* doi: 10.1109/ICDAR.2011.20.

—— Volkmar Frinken, Alicia Fornés, and Horst Bunke (2011b), 'Transcription Alignment of Latin Manuscripts Using Hidden Markov Models', *Proceedings of the 2011 Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery), 29–36.

Garz, Angelika, Nicole Eichenberger, Marcus Liwicki, and Rolf Ingold (2015), 'HisDoc 2.0: Toward Computer-Assisted Paleography', *manuscript cultures*, 7: 19–28.

Gronemeyer, Sven, and Franziska Diehr (2018), 'Organising the Unknown: A Concept for the Sign Classification of not yet (Fully) Deciphered Writing Systems Exemplified by a Digital Sign Catalogue for Maya Hieroglyphs', in J. Girón Palau and I. Gallina Russell (eds), *Digital Humanities Book of Abstracts* (Mexico City: La Red de Humanidades Digitales). <https://dh2018.adho.org/en/organising-the-unknown-a-concept-for-the-sign-classification-of-not-yet-fully-deciphered-writing-systems-exemplified-by-a-digital-sign-catalogue-for-maya-hieroglyphs/>.

Gumbert, J. Peter (1998), 'Commentare "Commentare Bischoff"', *Scrittura e Civiltà*, 22: 397–404.

Hassner, Tal, Malte Rehbein, Peter Stokes, and Lior Wolf (eds) (2013a), 'Computation and Palaeography: Potentials and Limits', *Dagstuhl Manifestos*, 2: 14–35. doi: 10.4230/DagMan.2.1.14.

———, Lior Wolf, and Nachum Derschowitz (2013b), 'OCR-Free Transcript Alignment', *12th International Conference on Document Analysis and Recognition (ICDAR)*, 1310–1314. doi: 10.1109/ICDAR.2013.265.

—— Robert Sablatnig, Dominique Stutzmann, and Ségolène Tarte (2015), 'Digital Palaeography: New Machines and Old Texts', *Dagstuhl Reports*, 4: 127–128. doi: <10.4230/DagRep.4.7.112>.

He, Sheng, Marco Wiering, and Lambert Schomaker (2015), 'Junction Detection in Handwritten Documents and its Application to Writer Identification', *Pattern Recognition*, 48: 4036–4048.

HIMANIS (2018), *The HIMANIS Project: Indexing the Trésor des Chartes Registers*. <http://www.himanis.org>.

IIIF [n.d.], *IIIF: International Image Interoperability Framework*. <http://iiif.io/>.

Kestemont, Mike, Vincent Christlein, and Dominique Stutzmann (2017), 'Artificial Palaeography: Computational Approaches to Identifying Script Types in Medieval Manuscripts', *Speculum: A Journal of Medieval Studies*, 92: S86–S109. doi: 10.1086/694112.

McCarty, Willard (2004), 'Modeling: A Study in Words and Meanings', in Susan Schreibman, Ray Siemens, and John Unsworth (eds), *A Companion to Digital Humanities* (Oxford: Blackwell), 254–270.

Medium (2016), *Medium: Répertoire des manuscrits reproduits ou recensés* (Paris: IRHT). <http://medium.irht.cnrs.fr>.

MESA [n.d.], *MESA: Medieval Electronic Scholarly Alliance*. <http://www.mesa-medieval.org>.

Murano, Francesca, and Achille Felicetti (2017), *Definition of the CRMtex: An Extension of CIDOC-CRM to Model Ancient Textual Entities*, version 0.8 (CIDOC-CRM). <http://www.cidoc-crm.org/crmtex/sites/default/files/CRMtex_v0.8.pdf>.

Petrucci, Armando (2001), *La descrizione del manoscritto: storia, problemi, modelli* (Rome: Carocci).

—— (2005), 'Digrafismo e bilettrismo nella storia del libro', *Syntagma: Revista del Instituto de Historia del Libro y de la Lectura*, 1: 53–75.

Pfaff, Richard W. (1977), 'M. R. James on the Cataloguing of Manuscripts: A Draft Essay of 1906', *Scriptorium*, 31: 103–118

Pinakes (2016), *Pinakes: Textes et manuscrits grecs* (Paris: Institut de recherche et d'histoire des textes). <http://pinakes.irht.cnrs.fr>.

Radiciotti, Paolo (2006), 'Il problema del digrafismo nei rapporti fra scrittura latina e greca nel medioevo', *Nea Rômê*, 3: 5–55.

Rajan Sampath, Vinodh (2016), *Quantifying Scribal Behaviour: A Novel Approach to Digital Paleography* (PhD Dissertation, St Andrews: University of St Andrews). <http://hdl.handle.net/10023/9429>.

Rothacker, Leonard, Denis Fisseler, Gerfrid G. W. Müller, Frank Weichert, and Gernot A. Fink (2015), 'Retrieving Cuneiform Structures in a Segmentation-free Word Spotting Framework', *Proceedings of the International Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery). <http://patrec.cs.tu-dortmund.de/pubs/papers/Rothacker2015-RCS.pdf>.

Sadeh, Gil, Lior Wolf, Tal Hassner, Nachum Dershowitz, and Daniel Stökl Ben-Ezra (2015), 'Viral transcript alignment', *13th International Conference on Document Analysis and Recognition (ICDAR)*, 711–715. doi: 10.1109/ICDAR.2015.7333854.

Schomaker, Lambert (2007), 'Advances in Writer Identification and Verification', *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR)*, 2: 769–773.

Sculley, D., and Bradley M. Pasanek (2008), 'Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities', *Literary and Linguistic Computing*, 23: 409–424.

Seuret, Matthias, Daniel Stökl ben Ezra, and Markus Liwiki (2017), 'Robust Heartbeat-based Line Segmentation Methods for Regular Texts and Paratextual Elements', *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery), 71–76.

Singh, Pawan Kumar, Arafat Mondal, Showmik Bhowmik, Ram Sarkar, and Mita Nasipuri (2014), 'Word-Level Script Identification from Handwritten Multi-script Documents', *Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA)*, 551–558. doi: 10.1007/978-3-319-11933-5_62.

Sirat, Colette (2006), *Writing as Handwork: A History of Handwriting in Mediterranean and Western Culture*, (Turnhout: Brepols).

SNAP DRGN [n.d.], *Standards for Networking Ancient Prosopographies* (London: King's College). <http://snapdrgn.net>.

Smit, Ginna (2011), 'The Death of the Palaeographer? Experiences with the Groningen Intelligent Writer Identification System (GIWIS)', *Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde*, 57: 413–425.

Smith, Marc H. (2004), 'Les «gothiques documentaires»: un carrefour dans l'histoire de l'écriture latine', *Archiv für Diplomatik*, 50: 417–465.

Stokes, Peter A. (2009), 'Computer-Aided Palaeography: Present and Future', in Malte Rehbein, Patrick Sahle, and Torsten Schaßan (eds), *Kodikologie und Paläographie im Digitalen Zeitalter—Codicology and Palaeography in the Digital Age* (Norderstedt: Books on Demand), 309–338. <http://kups.ub.uni-koeln.de/2978/>.

Stokes, Peter A. (2012), 'Palaeography and the "Virtual Library"', in Brent Nelson and Melissa Terras (eds), *Digitizing Medieval and Early Modern Material Culture* (Arizona: Center for Medieval and Renaissance Studies), 137–169.

—— (2015), 'Digital Approaches to Paleography and Book History: Some Challenges, Present and Future', *Frontiers in Digital Humanities*, 2:5. doi: <10.3389/fdigh.2015.00005>.

—— (2017), 'Scribal Attribution across Multiple Scripts: A Digitally-Aided Approach', *Speculum: A Journal of Medieval Studies*, 92: S65–S85.

—— (2018), 'Modelling Multigraphism: The Digital Representation of Multiple Scripts and Alphabets', in J. Girón Palau and I. Gallina Russell (eds), *Digital Humanities Book of Abstracts* (Mexico City: La Red de Humanidades Digitales), 292–296. <https://dh2018.adho.org/modelling-multigraphism-the-digital-representation-of-multiple-scripts-and-alphabets/>.

Stutzmann, Dominique (2011), 'Paléographie statistique pour décrire, identifier, dater...: Normaliser pour coopérer et aller plus loin?', in Franz Fischer, Christiane Fritze, and Georg Vogeler (eds), *Kodikologie und Paläographie im Digitalen Zeitalter 2—Codicology and Palaeography in the Digital Age 2* (Norderstedt: Books on Demand), 247–277. <https://kups.ub.uni-koeln.de/4353/>.

—— (2012), 'Modélisation des signes graphiques (1)', *Écriture médiévale et numérique* [project blog]. <http://oriflamms.hypotheses.org/921>.

—— (2013), 'Ontologie des formes et encodage des textes manuscrits médiévaux: Le projet ORIFLAMMS', *Document numérique*, 16: 81–95. doi: <10.3166/DN.16.3.69-79>.

—— Théodore Bluche, Alexei Lavrentev, and Yann Leydier (2015), 'From Text and Image to Historical Resource: Text-Image Alignment for Digital Humanists', in *Digital Humanities 2015: Book of Abstracts* (Sydney: University of Western Sydney). <http://dh2015.org/abstracts/xml/STUTZMANN_Dominique_From_Text_and_Image_to_Histor/STUTZMANN_Dominique_From_Text_and_Image_to_Historical_R.html>.

—— (2017), *Projet ANR-12-CORP-0010 ORIFLAMMS, Programme Corpus données et outils 2012 – Compte-rendu de fin de projet*. <https://oriflamms.hypotheses.org/files/2017/04/Oriflamms-Compte-rendu-final.pdf>.

Sudholt, Sebastian, and Gernot A. Fink (2016), 'PHOCNet: A Deep Convolutional Neural Network for Word Spotting in Handwritten Documents', in *Proceedings of the International Conference on Frontiers in Handwriting Recognition.* <http://patrec.cs.tu-dortmund.de/pubs/papers/Sudholt2016-PAD>.

TRAME (2013), *TRAME: Text and Manuscript Transmission of the Middle Ages* (Florence: Fondazione Ezio Franceschini). <http://trame.fefonlus.it/>.

Transkribus [n.d.], *Transkribus: Transcribe, Collaborate, Share* (Innsbruck: University of Innsbruck). <https://transkribus.eu/>.

Ubul, Kurban, Gulzira Tursun, Alimjan Aysa, Donato Impedovo, Giuseppe Pirlo, and Tuergen Yibulayin (2017), 'Script Identification of Multi-Script Documents: A Survey', *IEEE Access*, 5: 6546–6559. doi: <10.1109/ACCESS.2017.2689159>.

Washington Database (2016), *Washington Database* (Bern: University of Bern). <http://www.fki.inf.unibe.ch/databases/iam-historical-document-database/washington-database>.

Wolf, Lior, Rotem Littman, Naama Mayer, Tanya German, Nachum Dershowitz, Roni Shweka, and Yaacov Choueka (2011a), 'Identifying Join Candidates in the Cairo Genizah', *International Journal of Computer Vision*, 94: 118–135. <doi:10.1007/s11263-010-0389-8>.

—— Nachum Dershowitz, Liza Potikha, Tanya German, Roni Shweka, and Yaacov Choueka (2011b), 'Automatic Paleographic Exploration of Genizah Manuscripts', in Franz Fischer, Christiane Fritze, and Georg Vogeler (eds), *Kodikologie und Paläographie im Digitalen Zeitalter 2—Codicology and Palaeography in the Digital Age 2* (Norderstedt: Books on Demand), 157–179.

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2016), 'SDK Reinvented: Document Image Analysis Methods as RESTful Web Services', *12th IAPR Workshop on Document Analysis Systems (DAS)*, 90–95.

Zhitomirsky-Geffet, Maayan, and Gita Prebor (2016). 'Towards an Ontopedia for Historical Hebrew Manuscripts', *Frontiers in Digital Humanities*, 3:3. doi: <10.3389/fdigh.2016.00003>.

Written Artefacts as Cultural Heritage

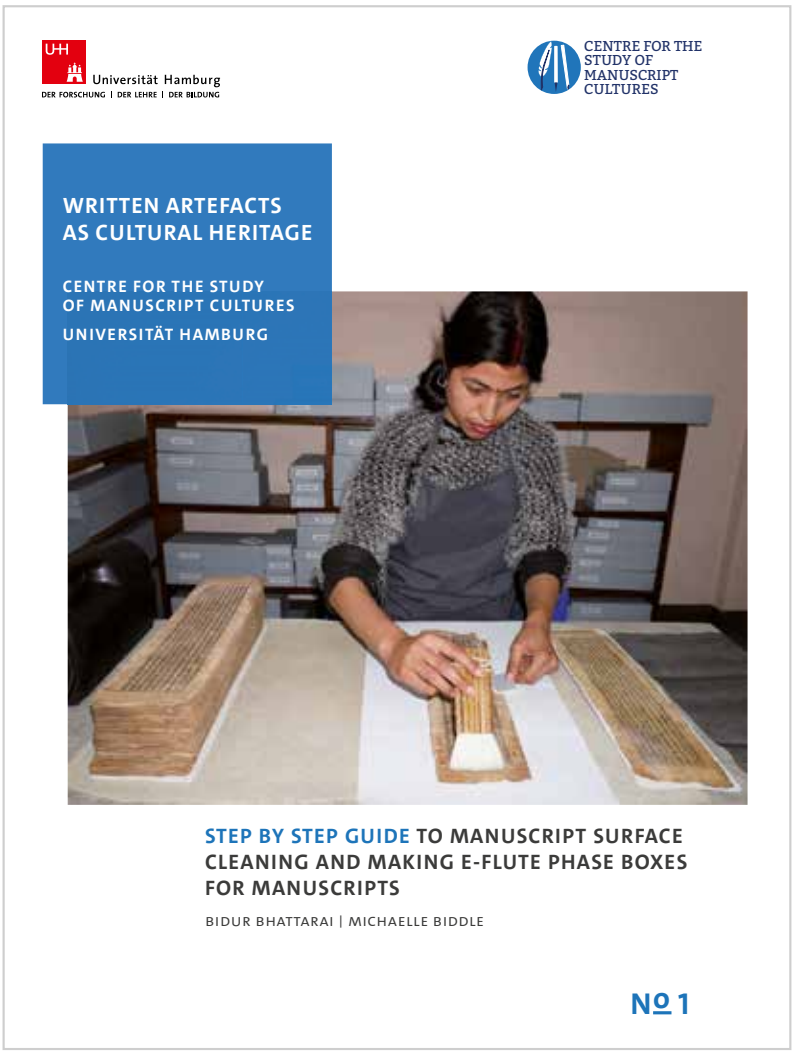
Ed. by Michael Friedrich and Doreen Schröter

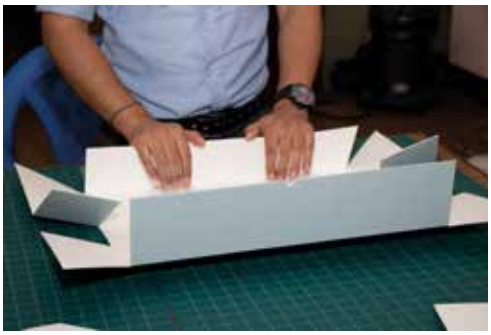
Written Artefacts as Cultural Heritage was established in 2020. The series is dedicated to the double role of written artefacts as representations and generators of humankind's cultural heritage. Its thematic scope embraces aspects of preservation, the identity-defining role of artefacts as well as ethical questions.

The mix of practical guides, colloquium papers and project reports is specifically intended for staff at libraries and archives, curators at museums and art galleries, and scholars working in the fields of manuscript cultures and heritage studies.

Every volume of *Written Artefacts as Cultural Heritage* has been peer-reviewed and is openly accesible. There is an online and a printed version..

If you wish to receive a copy or to present your research, please contact the editorial office:
https://www.csmc.uni-hamburg.de/publications/cultural-heritage.html





manuscript cultures (mc)

Editors: Michael Friedrich and Jörg B. Quenzer
Editorial office: Irina Wandrey

CSMC's academic journal was established as newsletter of the research unit 'Manuscript Cultures in Asia and Africa' in 2008 and transformed into a scholarly journal with the appearance of volume 4 in 2011. *manuscript cultures* publishes exhibition catalogues and articles contributing to the study of written artefacts. This field of study embraces disciplines such as art history, codicology, epigraphy, history, material analysis, palaeography and philology, informatics and multispectral imaging.

manuscript cultures encourages comparative approaches, without regional, linguistic, temporal or other limitations on the objects studied; it contributes to a larger historical and systematic survey of the role of written artefacts in ancient and modern cultures, and in so doing provides a new foundation for ongoing discussions in cultural studies.

Every volume of *manuscript cultures* has been peer-reviewed and is openly accessible:
https://www.csmc.uni-hamburg.de/publications/mc.html

If you wish to receive a copy or to present your research in our journal, please contact the editorial office:
irina.wandrey@uni-hamburg.de

manuscript cultures 14

manuscript cultures 13

manuscript cultures 11

manuscript cultures 10

manuscript cultures 9

manuscript cultures 8

manuscript cultures 7

manuscript cultures 6

manuscript cultures 5

manuscript cultures 4

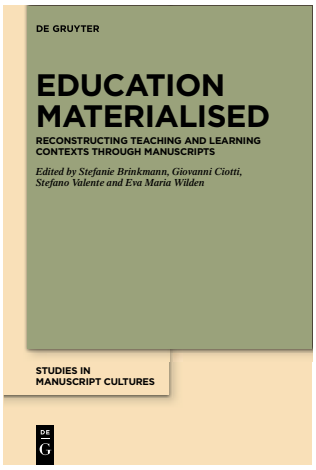
Studies in Manuscript Cultures (SMC)

Ed. by Michael Friedrich, Harunaga Isaacson, and Jörg B. Quenzer

From volume 4 onwards all volumes are available as open access books on the De Gruyter website:
https://www.degruyter.com/view/serial/43546
https://www.csmc.uni-hamburg.de/

DE GRUYTER

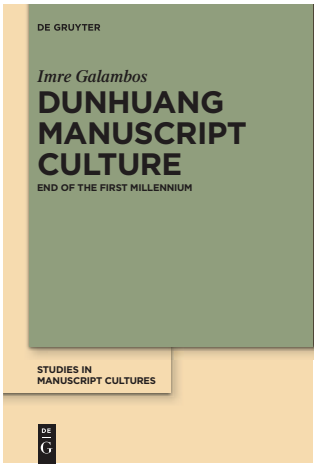
Forthcoming



23 – *Education Materialised: Reconstructing Teaching and Learning Contexts through Manuscripts,* edited by Stefanie Brinkmann, Giovanni Ciotti, Stefano Valente and Eva Maria Wilden

Manuscripts have played a crucial role in the educational practices of virtually all cultures that have a history of using them. As learning and teaching tools, manuscripts become primary witnesses for reconstructing and studying didactic and research activities and methodologies from elementary levels to the most advanced.
The present volume investigates the relation between manuscripts and educational practices focusing on four particular research topics: educational settings: teachers, students and their manuscripts; organising knowledge: syllabi; exegetical practices: annotations; modifying tradition: adaptations.
The volume offers a number of case studies stretching across geophysical boundaries from Western Europe to South-East Asia, with a time span ranging from the second millennium BCE to the twentieth century CE.

New release



22 – *Dunhuang Manuscript Culture: End of the First Millennium*, by Imre Galambos

Dunhuang Manuscript Culture explores the world of Chinese manuscripts from ninth–tenth century Dunhuang, an oasis city along the network of pre-modern routes known today collectively as the Silk Roads. The manuscripts have been discovered in 1900 in a sealed-off side-chamber of a Buddhist cave temple, where they had lain undisturbed for for almost nine hundred years. The discovery comprised tens of thousands of texts, written in over twenty different languages and scripts, including Chinese, Tibetan, Old Uighur, Khotanese, Sogdian and Sanskrit. This study centres around four groups of manuscripts from the mid-ninth to the late tenth centuries, a period when the region was an independent kingdom ruled by local families. The central argument is that the manuscripts attest to the unique cultural diversity of the region during this period, exhibiting – alongside obvious Chinese elements – the heavy influence of Central Asian cultures. As a result, it was much less 'Chinese' than commonly portrayed in modern scholarship. The book makes a contribution to the study of cultural and linguistic interaction along the Silk Roads.

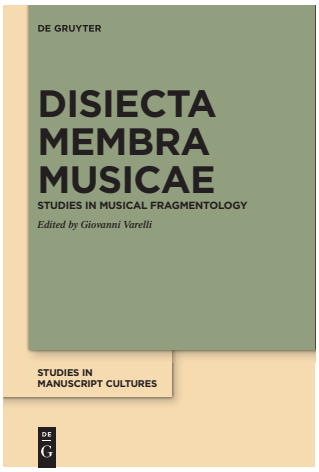
Studies in Manuscript Cultures (SMC)

Ed. by Michael Friedrich, Harunaga Isaacson, and Jörg B. Quenzer

From volume 4 onwards all volumes are available as open access books on the De Gruyter website:
https://www.degruyter.com/view/serial/43546
https://www.csmc.uni-hamburg.de/

DE GRUYTER

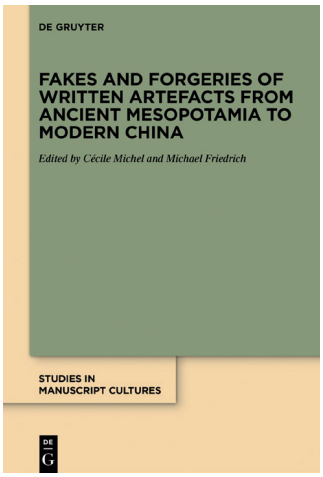
New release



21 – *Disiecta Membra Musicae: Studies in Musical Fragmentology*, edited by Giovanni Varelli

Although fragments from music manuscripts have occupied a place of considerable importance since the very early days of modern musicology, a collective, up-to-date, and comprehensive discussion of the various techniques and approaches for their study was lacking. On-line resources have also become increasingly crucial for the identification, study, and textual/musical reconstruction of fragmentary sources. *Disiecta Membra Musicae. Studies in Musical Fragmentology* aims at reviewing the state of the art in the study of medieval music fragments in Europe, the variety of methodologies for studying the repertory and its transmission, musical palaeography, codicology, liturgy, historical and cultural contexts, etc. This collection of essays provides an opportunity to reflect also on broader issues, such as the role of fragments in last century's musicology, how fragmentary material shaped our conception of the written transmission of early European music, and how new fragments are being discovered in the digital age. Known fragments and new technology, new discoveries and traditional methodology alternate in this collection of essays, whose topics range from plainchant to *ars nova* and fifteenth- to sixteenth-century polyphony.

New release



20 – *Fakes and Forgeries of Written Artefacts from Ancient Mesopotamia to Modern China,* edited by Cécile Michel and Michael Friedrich

Fakes and forgeries are objects of fascination. This volume contains a serie of thirteen articles devoted to fakes and forgeries of written artefacts from th beginnings of writing in Mesopotamia to modern China. The studies empha sise the subtle distinctions conveyed by an established vocabulary relating t the reproduction of ancient artefacts and production of artefacts claiming t be ancient: from copies, replicas and imitations to fakes and forgeries. Fake are often a response to a demand from the public or scholarly milieu, or ever both. The motives behind their production may be economic, political, reli gious or personal – aspiring to fame or simply playing a joke. Fakes may be revealed by combining the study of their contents, codicological, epigraphi and palaeographic analyses, and scientific investigations. However, certain fa mous unsolved cases still continue to defy technology today, no matter hov advanced it is. Nowadays, one can find fakes in museums and private collec tions alike; they abound on the antique market, mixed with real artefacts tha have often been looted. The scientific community's attitude to such objects calls for ethical reflection.

ISSN 1867–9617

Centre for the Study of Manuscript Cultures
Universität Hamburg
Warburgstraße 26
D-20354 Hamburg

www.csmc.uni-hamburg.de