Schriften des Instituts für Dokumentologie und Editorik — Band 3

# Kodikologie und Paläographie im digitalen Zeitalter 2

_____

# Codicology and Palaeography in the Digital Age 2

herausgegeben von | edited by

Franz Fischer, Christiane Fritze, Georg Vogeler

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Malte Rehbein, Patrick Sahle

2010

BoD, Norderstedt

# Artefacts and Errors: Acknowledging Issues of Representation in the Digital Imaging of Ancient Texts

Melissa M. Terras

> The shadows of artefacts would
> constitute the only reality people
> in this situation would recognize.
>
> <div align="right">Plato, the Republic, Book 7, 515c.</div>

## Abstract

It is assumed, in palaeography, papyrology and epigraphy, that a certain amount of uncertainty is inherent in the reading of damaged and abraded texts. Yet we have not really grappled with the fact that, nowadays, as many scholars tend to deal with digital images of texts, rather than handling the texts themselves, the procedures for creating digital images of texts can insert further uncertainty into the representation of the text created. Technical distortions can lead to the unintentional introduction of 'artefacts' into images, which can have an effect on the resulting representation. If we cannot trust our digital surrogates of texts, can we trust the readings from them? How do scholars acknowledge the quality of digitised images of texts? Furthermore, this leads us to the type of discussions of representation that have been present in Classical texts since Plato: digitisation can be considered as an alternative form of representation, bringing to the modern debate of the use of digital technology in Classics the familiar theories of mimesis (imitation) and ekphrasis (description): the conversion of visual evidence into explicit descriptions of that information, stored in computer files in distinct linguistic terms, with all the difficulties of conversion understood in the ekphratic process. The community has not yet considered what becoming dependent on digital texts means for the field, both in practical and theoretical terms. Issues of quality, copying, representation, and substance should be part of our dialogue when we consult digital surrogates of documentary material, yet we are just constructing understandings of what it means to rely on virtual representations of artefacts. It is necessary to relate our understandings of uncertainty in palaeography and epigraphy to our understanding of the mechanics of visualization employed by digital imaging techniques, if we are to fully understand the impact that these will have.

## Zusammenfassung

Die paläographische, papyrologische und epigraphische Forschung geht davon aus, dass das Lesen eines beschädigten und radierten Textes Unsicherheiten mit sich bringt. Bislang haben wir uns jedoch noch nicht ausreichend mit dem Umstand auseinandergesetzt, dass heutzutage, wo sich viele Forscher eher mit digitalen Bildern von Texten als mit dem Text selber befassen, die Erstellung von digitalen Bildern ebenso zu Unsicherheiten in der Repräsentation des Textes führen kann. Technisch bedingte Verzerrungen können zu unbeabsichtigten 'Artefakten' in den Bildern führen, die sich auch in der Textrepräsentation niederschlagen. Wenn wir schon den digitalen Surrogaten der Texte nicht trauen können, können wir es dann ihren Transkriptionen? Wie gehen Wissenschaftler mit der Qualität digitaler Bilder von Text um? Das führt zu der Diskussion über die Repräsentation von Texten, die in der Altphilologie seit Plato geführt wird: Digitalisierung kann als eine alternative Form der Repräsentation gelten, welche die klassische Diskussion um Mimesis (Imitation) und Ekphrasis (Beschreibung) in die moderne Diskussion über die Nutzung digitaler Technologien einbringt. Die Umwandlung von visueller Evidenz in explizite Beschreibung von Information, gespeichert in einem Computer in voneinander klar getrennten sprachlichen Ausdrücken, zeigt all die Schwierigkeit von Umwandlungen, die als Ekphrasis verstanden worden sind. Die Fachgemeinschaft hat bislang noch nicht darüber nachgedacht, was es für ihr Forschungsfeld, praktisch wie theoretisch, bedeutet, von digitalen Bildern abhängig zu sein. Fragen nach Qualität, Kopie, Repräsentanz und Substanz sollten Teil unseres Dialoges werden, wenn wir digitale Surrogate dokumentarischen Materials benutzen, auch wenn wir ein Verständnis von dem, was es heißt, sich auf virtuelle Repräsentationen von Artefakten zu verlassen, gerade erst konstruieren. Wir müssen also unsere Vorstellung von Unsicherheit in Paläographie und Epigraphik mit einer Vorstellung von der Mechanik der Visualisierung mit digitalen Bildgebungsverfahren ergänzen, wenn wir denn ihre gesamten Auswirkungen auf die Forschung verstehen wollen.

## 1. Introduction

Constructing readings of ancient documents is a difficult, complex, and time-consuming task, often involving reference to a variety of linguistic and archaeological data sets, and the invocation of previous knowledge of similar documentary material. Due to the involved reading process, it is difficult to record how the final interpretation of the document was reached, and which competing hypotheses were presented, adopted, or discarded in the process. It is also difficult to acknowledge and present the probabilities, and uncertainties, which were called on to resolve a final reading of a text. It is assumed,

across all aspects of palaeography, codicology, papyrology, and epigraphy[1] (which share a central core of identifying aspects and making sense of documentary material, despite their individual focus on media or form) that a certain amount of uncertainty is inherent in the reading of damaged and abraded texts. Indeed, the Leiden system (and its related markup technique, EpiDoc) allows for the encapsulation of this uncertainty, and acknowledge that uncertainty is a critical part of the reading of ancient texts. Yet we have not really grappled with the fact that, nowadays, as many scholars tend to deal with digital images of texts, rather than handling the texts themselves, the procedures for creating digital images of texts, or relying on computational systems to aid the process of reading ancient texts, can insert further uncertainty into the representation of the text created. This is becoming a general problem in all fields of manuscript studies, and requires further focus and concentration.

In technical terms, issues raised by the digitisation process include distortion caused by lens shape, difficulties in colour management and reproduction, and the unintentional introduction of 'artefacts' into images, which can have an effect on the resulting image. If we cannot trust our means of reproduction of images of texts, can we trust the readings from them? How do scholars acknowledge the quality of digitised images of texts?

Furthermore, this leads us to the type of discussions of representation that have been present in Classical texts since Plato: digitisation can be considered as an alternative form of representation, bringing to the modern debate of the use of digital technology in Classics the well trodden arguments of mimesis (imitation and representation of aspects of the real world in different forms and media). Digitisation is, even, a form of ekphrasis, "a descriptive speech which brings the thing shown vividly before the eyes"[2] converting visual evidence into explicit descriptions of that information, stored in computer files in distinct linguistic terms, with all the difficulties of conversion understood in the process of ekphrasis.

Has the classical community considered what becoming dependent on digital texts means for the field? Are these issues of quality, copying, representation, and substance part of our dialogue when we consult digital images of ancient texts? What measures can be taken to ensure that we understand what we are looking at when utilising digital images of ancient texts? How does the form of the digital image inform or distort

---

[1]  Whilst papyrology is the study of ancient manuscripts, mostly written on papyrus, palaeography is the study of handwriting, and the decipherment and reading of historical manuscripts. Epigraphy deals with Ancient inscriptions, whilst codicology is the study of books, especially manuscripts, as physical objects. All have a fundamental core function of extrapolating meaning and understanding about culture and history from written historical primary sources. Although the aims and foci of each speciality remain distinct, there are various crossovers in methodology and approach which allow for generalisations regarding the study of ancient texts and our approach to uncertainty therein.

[2]  This definition appears in a series of first century CE Progymnasmata from Theon, via Hermogenes. See Webb.

our potential readings of content, and relate to our understandings of uncertainty in palaeography and epigraphy? This chapter aims to explore issues of representation and uncertainty within the reading of ancient texts, with a particular focus on our increasing use of digital images, and how this impacts the papyrology and epigraphical community.

## 2. Papyrology and Computing

Classics as a subject has made much use of information technology (see Crane for an overview). "This tendency can partly be explained with reference to two observations: (1) the complexity of the textual, historical, linguistic, material, and artistic sources that need to be considered in classical scholarship, and (2) the patchy coverage and fragmentary state of many of these same artefacts" (Bodard and Mahony). Most of the uses of computing when reading ancient texts, however, do not turn to advanced computational techniques: like many disciplines in the humanities, the use of computing in Classics is mostly to speed up and enhance access to information which had previously only available in analogue format, through a process of digitising existing resources and making them available online. Deegan and Tanner summarise succinctly the wide range of reasons given for digitisation, and the advantages digitisation of materials can provide for scholars and institutions:

> immediate access to high-demand and frequently used items; easier access to individual components within items (e.g. articles within journals); rapid access to materials held remotely; the ability to reinstate out of print materials; the potential to display materials that are in inaccessible formats, for instance, large volumes, or maps; 'virtual reunification'—allowing dispersed collections to be brought together; the ability to enhance digital images in terms of size, sharpness, colour contrast, noise reduction, etc.; the potential to conserve fragile/precious objects while presenting surrogates in more accessible forms; the potential for integration into teaching materials; enhanced searchability, including full text; integration of digital media (images, sounds, video, etc.); the ability to satisfy requests for surrogates (photocopies, photographic prints, slides, etc.); reducing the burden of cost of delivery; the potential for presenting a critical mass of materials. (32–33)

Classical scholars were swift to recognise the benefits of digitisation, and began to address major issues of computational infrastructure in the 1970s with large, fairly centralized efforts which have since become the central starting point for many aspects of the scholarly research of antiquity (Crane). These projects included David Packard's Ibycus system, the *Thesaurus Linguae Graecae*, the *Database of Classical Bibliography*,

the *Bryn Mawr Classical Review*, the Duke Databank of Documentary Papyri, and the Perseus Project. By the close of the 20th Century, many online projects devoted to delivering images, transcriptions, and notes regarding papyri had been set up. A project such as APIS, the Advanced Papyrological Information System (2007a), demonstrates how digitisation, and the drawing together of disparate existing knowledge sources, can aid those engaged in reading ancient Texts. APIS describes itself as

> a collections-based repository hosting information about and images of papyrological materials (e.g. papyri, ostraca, wood tablets, etc) located in collections around the world. It contains physical descriptions and bibliographic information about the papyri and other written materials, as well as digital images and English translations of many of these texts. When possible, links are also provided to the original language texts (e.g. through the Duke Data Bank of Documentary Papyri). The user can move back and forth among text, translation, bibliography, description, and image (APIS 2007c).

APIS currently hosts over 30,000 different records, and 23,000 individual images from 27 major institutions (providing online links to digital images, where available, at their host site if they are not part of the APIS system themselves). With the provision of an intuitive search function, scholars can search across a wealth of texts, and can often have access to various high resolution images of each document, allowing them to download them and access them remotely, negating the need to travel, visit, and handle the document themselves.

Likewise, sites such as the Vindolanda Tablets Online (n.d.) devoted to the documents found at one particular fort on Hadrian's Wall, provide much greater access to both images, transcriptions, translations, and notes regarding documents, allowing cross referencing and in depth scholarly analysis of one particular set of documents to be undertaken without having to gain access to the original artefacts in the British Museum nor depend on print volumes in which photographic provision can sometimes be limited. A recently developed companion site, Vindolanda Tablets Online II provides updates to the collection, and provides a web service where the existing information concerning the tablets can be searched in finer detail than in the original site. The Vindolanda websites are based on print volumes, with the online equivalent extending their functionality and increasing the volume of information available regarding texts. Other sites (such as the Inscriptions of Aphrodisias project, or Inscriptions of Roman Tripolitania) have adopted the online medium to extend and expand the remit of the print editions they are based on, using a sophisticated blend of content specific markup (EpiDoc) behind the interface to allow searching, and scholarly analysis, both in and across the individual collections.

It is not the place here to survey every website which provides online versions of print volumes of ancient texts, or high resolution images of papyri, ostraca, epigraphical

and other documents.[3] The point to be made is that in a short space of time, much work in papyrology has moved from consideration of the physical document itself, or print surrogates (which, although previously commonly used have their own limitations in print and image quality) to working with relatively high resolution images of the documents, provided by online scholarly editions through databases. When Roger Bagnall, in his seminal 1995 introductory text "Reading Papyri, Writing Ancient History" opened with the phrase "Papyrology has tended to one of the most resolutely technical and positivistic disciplines of antiquity" (vii) he was not concerned with web based technologies which were then in their infancy. Likewise, when H. C. Youtie commented in a lecture in 1957 that

> the distinguishing characteristics of the scholarship of the twentieth century has been its dependence on papyri, and papyrology, like archaeology, epigraphy, numismatics, and mediaeval palaeography, has become a permanent adjunct to the technical equipment of classical scholars and ancient historians (267)

he was not concerning himself with Information Technology. As Brunner said in a 1993 chapter covering "Classics and the Computer, the history of a relationship":

> The field of Classics encompasses quite a few centuries; yet there are few distinct periods within this history that can be said to have witnessed changes as rapid and fundamental as through brought about by the entry of the field into the electronic world. A mere two and a half decades after "electronic machines" first found mention in an AP publication, few (if any) members of the Association remain uninvolved in, and unaffected by, computing. (28)

Papyrology is a field which revolves around the resolution and consideration of uncertainty within ancient texts. We may acknowledge this in textual transcriptions of documentary material, but do we acknowledge how the dependence on digital surrogates can affect and inject other modes of ambiguity, uncertainty, and representation in the images of documents that we now attempt to read?

## 3. Uncertainty, Ancient Texts, and Computing

The process of reading ancient documents is traditionally undertaken by an expert such as an epigrapher, papyrologist, or palaeographer. The expert will use their accumulated knowledge combined with external resources to piece together an interpretation of each

---

[3] The Digital Classicist Wiki hosts a list of projects undertaking relevant research, including those who provide digitised surrogates of papyrological material. A list of papyri sources is also given in Bernhardt.

ancient document. Such interpretation can be a long process, and it can be difficult for experts to maintain a record of the decisions made whilst undertaking their reading (Youtie 1963). This is important when defending their interpretation, sharing their hypotheses with other experts, or breaking off in the process of reading an ancient text and hoping to pick up the thought process in another reading session.

When undertaking tasks which depend on complex reasoning, there is often an element of probability that needs to be addressed. Unlike mathematical models which can deal with, say, different levels of dosage of medication on a patient, those reading ancient documents are often faced with a range of uncertainties with few or no prior models on which to base reasoning, and little framework in which to test these hypotheses. Ambiguity is a feature of the readings of ancient texts, even in their published version. This is not a critique of papyrologists: published versions are open to correction, and merely detail the extent to which the author has resolved the reading of the text at that moment in time. Bowman and Tomlin provide examples of readings which have changed dramatically between different versions of published texts. Tomlin demonstrated how the correct reading of a tablet was achieved by rotating the tablet through 180 degrees and rereading the text. However, the fact that this is accepted practice indicates that uncertainty about the reading of ancient text is seldom exhausted. Reading resolves around the resolution of ambiguity through prediction, prior knowledge, reasoning about the characteristics of the documents, and the head-on addressing of uncertainties.

For over a century, those reading ancient texts have tried to encapsulate their reasoning process in the resulting published transcripts of the texts in question. Due to the costs involved, producing a facsimile of a text proved prohibitive (Mahoney), and scholars became dependent on the use of set of signs and brackets in transcriptions to signify textual features such as lost or supplied characters, damage to the text, the expansion of abbreviations, etc. Grenfell and Hunt were consistent in their use of signs and brackets to describe the state and reading of texts from Oxyrhynchus, influencing a generation of scholars. Bidez and Drachmann published a pamphlet examining the different customs of using bracket and sign conventions in editions. Van Groningen used this analysis to suggest a unified system for marking editions (1932a). However, the International Congress of Orientalists agreed upon a different system which strongly resembles the papyrological system used by Grenfell and Hunt. It is this system that is now known as the Leiden System (Van Groningen 1932b) which aims to capture various characteristics of a text, widely adopted in print publications for all types of ancient texts.

A commonly used symbol within the Leiden system is the under-dot, used to represent uncertainty. However, this is one of the most confusing concepts to represent in such a transcription. There is no way to measure the extent of uncertainty (for example, where the reader is a little trepidatious that their interpretation is correct, or marking

that the letter in question is unreadable). Practice varies between papyrologists, with some using the underdot to represent a broken letter which is certain, others using it only when identification is in doubt. The authors of the Tabulae Vindolandenses I and II (Bowman and Thomas 1983, 1994) tend to use it both when identification is uncertain and when to show that there is no doubt: using the dot to show that the letter is broken. An analysis of the Leiden Markup of the Vindolanda ink text indicates that 9.9% of the letters in the ink text were marked as being broken (Terras 71): the identification of a proportion of these will be problematic. Uncertainty is therefore an important issue to address when building computational systems to aid papyrologists. Unfortunately, encapsulating uncertainty in computational systems is not a straightforward process.

Computational systems depend on resolving "real world" situations into exact numerical strings. The ordinary, or "real" world of our senses, exists in a continuous flowing stream of signals across time and often space. An ancient document—or a photograph of the document—exists in analogue, where a varying signal represents a continuous range of values. In order to record, copy, transmit, or analyse such a complex signal using computational power, it is necessary to translate this into a form which is more simple, predictable, and processable. All telecommunication systems have one thing in common: the information to be sent is converted into signals which can be transmitted, and reassembled on reception, to be converted into something we can perceive as a fair copy of the original. Digital systems are those which rely on a sequence of discreet numeric values, rather than the unconstrained and continually varying qualities of analogue signals. Numeric values are used in digital systems for processing, display, transmission, and input: often sampling values from analogue sources in a process called "digitisation".

The most common digital systems are those used in computing and electronics, which rely on the binary numeric system. This is a system which represents all numbers using only two symbols, typically 0 and 1. These zeroes and ones are known as binary digits, or more commonly as the shortened derivation: "bits". Strings of bits can represent text, images, sound, and moving images: as the information to be represented grows more complex, more bits are required to represent it, and more complex mechanisms are used to store, display, and process the information contained within the data stream.

Providing numeric, textual, image, sound, and video based data in digital format, whether they have been translated from an analogue signal into bits, or "born digital" by being created with computational technologies in the first place, has various advantages. These strings of bits can be easily replicated, transmitted, accessed, and processed. Saving the data in a structured, predetermined format means it may be device independent, and can be transferred from system to system with minimal problems. Data can be manipulated by dedicated computer programs, allowing new versions of the information to be generated. Data can also be processed: mathematically sorted through to show hidden relationships, new arrangements, different views, and

expanded, contracted or concatenated knowledge. Human eyes and ears can sometimes distinguish between continuous analogue signals, and bit by bit digital approximations. Digital media are at their most effective when their constituent parts—samples—are not detectable by human senses.

The digitized images of ancient texts delivered by internet technologies to our desktops therefore have a more complex relationship to their original manifestation than we may like to consider, given that the visualization reproduced on the screen is often so seductive:

> A digital representation of an artefact is a representation of certain relevant characteristics of the artefact. It is not the original and complete artefact, nor even a metonymy or simulacrum of the complete artefact. It is only a representation of some "relevant characteristics". (Arnold 127)

Further to just producing realistic representations of ancient texts, the development of new imaging techniques such as multi spectral imaging, image processing algorithms which can remove noise, character recognition tools which can propose combinations of strokes as possible letters, and tools which can search databases to match word fragments to relevant words, grammar, and orthographies, depend on building explicit representations of knowledge which computational systems can work with. Mathematics underpins all image processing and Artificial Intelligence systems, depending on concrete algorithms and defined representations of information. The humanities scholar's grasp of uncertainty when approaching damaged and abraded texts, written in a foreign, ancient language, is not something that maps easily onto computational systems (or our expectations of the results computational systems can produce). It is hard to provide enough real-world knowledge encapsulated computationally which can reflect the amount of contextual information necessary to undertake the reading of ancient texts (Terras). Additionally, many of these relatively new and emergent technologies which supply encouraging results and novel readings of documents, such as multi spectral imaging as applied to ancient texts (see Ware et al., Bearman), are not benchmarked and thoroughly tested to ensure their methodologies are robust, and that the new images created from the process have a mathematically sound relationship to the original artefact.

## 4. Quality and Quality Assurance in Digitised Collections

If scholars are using digital surrogates to produce readings of ancient texts, it is imperative that these images are of high quality. But what makes a "good" digital image of an ancient document? How do we know the digital representations that we

work with are fit for purpose? How can we assess image quality, and trust the resulting representation?

> The creator of the image has the purpose of communicating it through a suitable channel to one or more observers. Every element of the chain from creator to receiver affects the quality of the image, and hence the effectiveness of the communication process. Noise in the channel, however, may degrade the transmission, and the characteristics of both the origination medium and reproduction medium limit the overall ability of the system to render the image (size, density range, colour gamut, etc). Viewing conditions in each case affect the perception of the person viewing the image. (Macdonald and Jacobsen 352)

Many digitization projects either produce their own, or adhere to established, guidelines for the production of digital images. APIS, for example, produce their own guidelines for contributing organizations which details file format, resolution, the use of colour targets to allow calibration of colour, lighting, and even file naming (APIS 2006). The Library of Congress produces up to date "Technical Standards for Digital Conversion of Text and Graphic Materials" (Library of Congress), which many digitization projects consult (and should be the first port of call for those considering a digitisation project, as they provide the minimum, and best practice, specification for digitisation processes). In the UK, the Joint Information System's Committee, which provides leadership in IT services for Higher Education, provides a free advisory service, JISC Digital Media, for those creating and dealing with digital material (2010). Given that "few standards govern the creation and use of digital images, and in a world of multiple stakeholders and multiple perspectives it may be difficult to agree on a uniform approach that suits all circumstances" (Kenney 24) it is imperative that those undertaking digitisation programs consult guidelines and carry out benchmarking procedures to ensure quality control of the digitised output (see Kenney for an overview).

However, it can be difficult to assess the quality of digital images themselves, even if there are assurances from the creators that guidelines have been followed. Many things are lost in the sampling of an image to create a digital representation: it can be difficult to ascertain size, physical characteristics, texture, and the accuracy of colour. Although certain measures can be taken to ensure the capture process is as accurate as possible, including adequate documentation (see MacDonald and Jacobsen), many projects do not comply with these. For example, the images within the Vindolanda Tablets Online do not have a colour target supplied with them, meaning that scholars could not calibrate their monitors to best look at the images—if they indeed understood why and how they should do so. Some, but not all, of the images presented in Inscriptions of Aphrodisias are supplied with a measurement bar—partly because many of the images were decades old before digitisation—and this can lead to confusion about scale.

It is good practice to inform users of the digitisation process, and how and when the primary material was digitised. Vindolanda Tablets Online does this (2003), as does Inscriptions of Aphrodisias (Bodard and Spence), although the documentation requires some searching for on both of their websites. It can be difficult within APIS to see where and how the digital record, and digital image, of each individual papyrus was created, as this technical metadata is not recorded about each entry in the database (APIS 2007b). This lack of easy access to documentation regarding the digitisation process can create problems for scholars who are dependent on the outputs of digitisation projects to undertake their own original research, as studies have shown that digital humanities resources tend to be poorly documented and therefore not be trusted by users:

> In the absence of technical documentation, it was impossible to reuse files [. . . ] Although users require procedural documentation, about the status and completeness of sources, and selection methods, this is often difficult to locate [. . . ] and shows that this makes reuse of digital resources almost impossible. (Warwick et al. 33)

Furthermore, the persuasive nature of the visualisation and display can mean we do not stop to question the very nature of digital images of historical artefacts:

> its labor of production has been concealed and therefore bears less evidence of authorship, provenance, originality, and other commonly accepted characteristics attributed to physical objects. For these reasons the digital object's materiality is not well understood. (Cameron 70)

There is a lot that can go wrong in the creation of digital images that are used for representation of historical artefacts. Colour, in particular, is a thorny issue, with many issues such as illumination, and the differences between the way the human eye and computer systems record colour (see Hunt for an introduction). There are technical issues with the sampling and representation process that depend on the nature of device characteristics that determine special resolution, and effects that are created dependent on the type of camera lens used for image capture (see Holst for an overview). Quantifying the accuracy of imaging equipment—whether analogue or digital—is not a simple mathematical task (Keelan). Assessing image quality is usually dependent on human observation, and this is a subjective notion dependent on observer-based quality judgements (Engeldrum). Additionally, conversion from one format to another, sampling and resampling, and compressing file sizes can result in unintended visual effects, or "artefacts" becoming obvious in images, where information is deleted and inserted awkwardly into the resulting representation.

It is rare that those utilizing digital images for scholarly research would stop to consider the mechanics which produced that image, or their accuracy or veracity. It

is difficult, even with adequate documentation, for those undertaking digitization to express the capture conditions and relationship of the surrogate to the digitized item in question. Images may be from a trusted scholarly source—but did those capturing the images understand and comply with the technical issues necessary to best represent that ancient object in the modern digital realm?

## 5. Digitisation and Theory

It may be useful to pause and consider the theoretical frameworks that exist for dealing with surrogates, or representations, of the objects that we wish to study. The reading of ancient texts when faced with image based source material is an ekphrastic task: that is, in general, papyrology attempts to produce a faithful textual description of image based material. Ekphrasis, "the verbal representation of visual representation" (Heffernan 297) can be traced back to the legendary Shield of Achilles in the Iliad, and stems from ancient poetics and rhetoric. The impossibility of describing the visual within the textual realm is at once both recognised, and discounted by the "ekphrastic hope" (Mitchell 152) where the impossibility of ekphrasis is overcome in imagination, metaphor, and sense discovery from the image itself. The difference between text and image is both celebrated and feared, acknowledged and discounted: can the description of an image in textual form ever match up to the visual sensory perception of seeing that image in the flesh?

> The utopian figures of the image and its textual rendering as transparent windows onto reality are supplanted by the notion of an image as a deceitful illusion, a magical technique that threatens to fixate the poet and the listener. (Mitchell 156)

Ekphrasis can be used as a tool to focus "the interarticulation of perceptual, semiotic, and social contradictions within verbal representation" (180). The literature regarding ekphrasis is large, with each type of visual representation "such as photography, maps, diagrams, movies, theatrical spectacles [… carrying] its own peculiar sort of textuality into the heart of the visual image" (181). Can images of ancient texts, and the description of the texts we visualise therein, be discounted from the problems understood and articulated by the task of describing, accurately, in text the visual nature of image based material?

Furthermore, digital images can be viewed as fundamentally ekphrastic: digital image data itself generally consists of a list of values of the colours of individual pixels, and associated instructions for computing applications regarding how this data should be displayed, written in a data stream in a computer file. This draws us back again to the notion of representation in the digital realm, and the reduction of our sensory

experience back to zeroes and ones. What is lost in this ekphrastic translation? What accuracy is compromised through this point by point sampling, and expression into explicit numerical (if not textual) values? Our images of text become numbers: we translate them back to images, then to transcriptions. Are we careful enough that nothing is changed or lost in the process?

Digital images are also surrogates, and have a complex relationship to their analogue equivalent which theorists are only now beginning to start to question:

> The digital historical object can exist in many realms and perform many roles that go beyond representation, interpretation, education, documentation, and archive. Indeed its analogonic role is potentially diverse [...] the status of copies from nondigital originals still remains ambiguous [...] A range of expanded meaning, material characteristics, and behaviours emerge as representing a particular configuration of space, time, and surface, sequence of user activities—a particular formal material and user experience. (Cameron 68)[4]

It is tempting to embrace the value of material authenticity and adopt a repugnant stance to reproductions of primary historical material. Both Walter Benjamin and Jean Baudrillard have argued that mechanical reproduction and simulations pose a threat to the real object of focus, leading to the loss of its "auraic, iconic, and ritualistic qualities" (Cameron 50). Do students currently learning palaeographic methods focus on digital surrogates? Is there a need to ensure that the "Google generation" (JISC 2007) who will be the professors of papyrology of the future understand the relationship the surrogate has to the original? Have they even handled original texts? Does this at all matter? In his chapter in this volume Peter Stokes remarks specifically on this subject

> basic skills in handling original materials, in reading, transcribing, editing and understanding these objects are central to [textual] studies. The question that remains is therefore twofold. First, how can digital tools be used to better teach traditional skills. Second, a question much less frequently raised, is how the teaching of traditional skills should or could itself change as a result: how and to what extent should digital content be explicitly introduced into the curriculum for the study of [...] manuscripts? (Stokes)

It is of course nonsensical to suggest that the benefits brought by digitisation to the palaeographical community should be rejected for some notion of the material superiority of original documentary materials. However, it may be helpful to give ourselves some distance from the digital image of ancient texts, and consider the

---

[4]  This whole paper provides a good overview of available theoretical discussions regarding the relationship of the digital surrogates to their analogue counterpart.

implications of our dependency on digital images for scholarly research. By doing so, we may reach an understanding of the "modality and materiality of digital historical objects" as "new roles and a set of defining characteristics emerge beyond their role as servant to the 'real' as representation, presence, affect, experience, and value" (Cameron 70).

## 6.  Dependency on the Digital: When Surrogates become Primary Sources

It is worth saying here that although it is a possibility, I am not aware of any published documentary material that has been read erroneously (and published, and refuted) due to faults in the digitization process. I am not aware of any digitised versions of ancient texts that are so faulty that artefacts and errors within the digital image are obvious[5]. I am aware that the *intention* with most cultural and heritage digitisation projects is to provide the *best* digital representation of that object that is possible. The intention is to replace the need for travel across countries and continents to see a scrap of papyri the size of your hand, and to facilitate research. I do not mean to criticise the efforts of those undertaking digitization (nor the particular projects named above).

It is also worth saying that the problems in ensuring that we understand the nature of digital representations of primary sources also apply to other disciplines. Any historical or literary research which depends on image-based primary resources is now facing the same turning point, as scholars turn to the digital as a convenient means to view and access a wide variety of digitised content. Some issues will hit fields harder than others (colour reproduction in digital images, for example, should generally be more of an issue for the art historian than the papyrologist), but nevertheless, the problem is similar: scholars become trained in the tools and methodologies entrenched in their own discipline, but seldom are educated in the technical underpinnings which allow the primary sources they depend on from various points across the world to magically appear on the computer screen in their office. Understanding digitisation is then a particular extension of digital literacy:

> the confident and critical use of Information Society Technology (IST) for work, leisure and communication. It is underpinned by basic skills in ICT: the use of computers to retrieve, assess, store, produce, present and exchange information, and to communicate and participate in collaborative networks via the Internet (European Parliament and Council)

combined with information literacy:

---

[5]   Or, in the case of Google Book's digital image of the first page of a Victorian edition of Plato's *Euthyphron*, where the fingers of the digitization operative are clearly visible in the scan (Cohen).

knowing when and why you need information, where to find it, and how to
evaluate, use and communicate it in an ethical manner. (CILIP)

Those utilising digital image resources of primary textual material have an obligation
to their academic discipline to understand the nature of the resources they are basing
readings, transcriptions, and translations on. Additionally, scholars should be more open
when publishing readings of texts in articulating both the sources and methodologies
used when consulting digital resources, and their use of digital surrogates in scholarship.
This will have the added benefit of providing "evidence of value" of costly digital
resources, which are battling at the moment to prove that they are essential services to
academic communities, and deserve further funding (AHRC). Likewise, those providing
digital images of ancient Texts, or any historical documents, have an obligation to
fully document, describe, and elucidate the process by which the digital surrogates
were created. Scholars must be encouraged to use this documentation, to ensure that
their research is based on as authentic a representation as a primary text as possible:
otherwise the readings they generate from them simply cannot be trusted.

## 7.  Conclusion

It would be folly to suggest that we should return to pre-digital dependencies on the
physical document and print based surrogates when trying to transcribe, read, study,
and understand ancient documentary material. The affordances of digital media increase
productivity, reduce travel time and cost, and provide in-depth and detailed information
regarding individual texts. Vast collections of images of ancient texts are accessible to
scholars from their own desk—and although few make this explicit in the methodologies
published in research papers and monographs, many are now dependent on online
databases and databanks.

However, just as palaeographers, epigraphers, papyrologists and codicologists
are educated and trained in textual mores of the ancient world, those scholars
dependent on the digital environment should ensure that they understand the
representations of artefacts that they base their research upon. Those undertaking
digitisation projects should be aware of the minimum acceptable technical standards
and adequate documentary approaches for the digital representations created (see
above for references). Those producing digital surrogates of primary historical texts
should produce adequate documentation that is easily available regarding the technical
procedures involved in capturing images of the artefacts. Only by fore-fronting the use
of both standards and documentation of these standards can we produce robust digital
resources that can stand up to academic scrutiny.

However, there are issues regarding digitisation that cannot be resolved in practical
form, and we must begin to build up our theoretical understanding of notions of

digitisation and representation so we can articulate our dependencies and be sure about our methodologies when relying on digital surrogates. In particular, it should be acknowledged that digital images of ancient texts have a complex relationship to their source material. Additionally, digital images created in ways which would never have existed using traditional photography, or human vision, such as multi spectral or infra-red images, should be treated as they are: representations, and surrogates, rather than replicas of original documentary material. Furthermore, more thought should be given as to the computational representational structures that we shoehorn our understanding of image, text, and language into, when experiencing the convenience of online papyrological sources.

The aim of this chapter has been to raise issues of digital representation within the Classics, and particularly within the papyrology community. The rapid computational transformation that has occurred in the field must be followed by the questioning and inquisitive methodology which is applied to trying to understand ancient texts themselves: what does it mean to continually use digital image surrogates to produce readings of ancient texts. If we cannot understand the means of production of the surrogates, can our interpretations ever be robust? Only through becoming digitally and informationally literate can we trust that our images of artefacts are free from artefacts and errors.

## Bibliography

APIS. *APIS Imaging Standards.* 2006. <http://www.columbia.edu/cu/lweb/projects/digital/apis/imaging_guidelines_2006Jan.pdf>.
   *Advanced Papyrological Information System.* 2007a.
   <http://www.columbia.edu/cu/lweb/projects/digital/apis/>.
   *Guidelines for APIS Metadata Contributors.* 2007b.
   <http://www.columbia.edu/cu/libraries/inside/projects/apis/guidelines.html>.
   *About the APIS project.* 2007c.
   <http://www.columbia.edu/cu/libraries/inside/projects/apis/about.html>.
Arnold, David. "Digital Artefacts. Possibilities and Purpose." *The Virtual Representation of the Past.* Eds. Mark Greengrass and Lorna Hughes. Ashgate: Farnham. 2008 158–170.
*ICT in Arts and Humanities Research. Review Activities.* Arts and Humanities Research Council. Reading: University of Reading, 2007.
   <http://www.ahrcict.rdg.ac.uk/activities/review/index.htm>.
Baudrillard, Jean. *Simulcra and Simulation.* Trans. Sheila Faria Glaser. Ann Arbour: University of Michigan Press, 1994.
Bagnall, Roger S. *Reading Papyri, Writing Ancient History.* London and New York: Routledge, 1995.
Bearman, Gregory. "Imaging the Dead Sea Scrolls for conservation purposes." *SPIE Newsroom.* 29th December 2008. <http://spie.org/x32760.xml?ArticleID=x32760>

Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction." *Illuminations: Essays and reflections.* Trans. Harry Zohn. London: Jonathan Cape, 1970 (1st published 1935). 217–252.

Bernhardt, Theodore. *Papyri Pages, Site Links.* 2009. <http://papyri.tripod.com/links-2.html#Collections>.

Bidez, Joseph and Anders B. Drachmann. *Emploi des signes critiques, disposition de l'apparat dans les éditions savantes de textes grecs et latins, conseils et recommandations.* Paris: É. Champion for Union Académique International, 1932.

Bodard, Gabriel and Simon Mahony. "'Though much is taken, much abides': Recovering antiquity through innovative digital methodologies: Introduction to the special issue." *Digital Medievalist* 4 (2008). <http://www.digitalmedievalist.org/journal/4/DCintro/>.

Bodard, Gabriel and Paul Spence. "Technical Preface." *Aphrodisias in Late Antiquity.* London: King's College London, 2004. <http://insaph.kcl.ac.uk/ala2004/about/techpref.html>.

Bowman, Alan and J. David Thomas. *Vindolanda: The Latin Writing Tablets.* London: Society for Promotion of Roman Studies, 1983.

Bowman, Alan and J. David Thomas. *The Vindolanda Writing-Tablets (Tabulae Vindolandenses II).* London: British Museum Press, 1994.

Bowman, Alan K. and Rodger S. O. Tomlin. "Wooden Stilus Tablets from Roman Britain." *Images and Artefacts of the Ancient World.* Eds. Alan K. Bowman and Michael Brady. Oxford: Oxford University Press, 2005. 7–14.

Brunner, Theodore. F. "Classics and the Computer: The History of a Relationship. *Accessing Antiquity: The Computerization of Classical Databases.* Ed. Jon Solomon. Tucson: University of Arizona Press. 1993. 10–33

Cameron, Fiona. "Beyond the Cult of the Replicant: Museums and Historical Digital Objects – Traditional Concerns, New Discourses." *Theorizing Digital Cultural Heritage, A Critical Discourse.* Eds. Fiona Cameron and Sarah Kenderdine. Cambridge MA: MIT Press, 2007.

CILIP. "Information Literacy: Definition." London: Chartered Institute of Library and Information Professionals (CILIP), 2010. <http://www.cilip.org.uk/get-involved/advocacy/learning/information-literacy/pages/definition.aspx>.

Cohen, Dan. "Google Fingers." *Blog Post, Dan Cohen's Blog,* 26th June 2006. <http://www.dancohen.org/blog/posts/google_fingers>.

Crane, Gregory. "Classics and the Computer: An End of the History." *A Companion to Digital Humanities.* Eds. Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell, 2006. <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-4&toc.depth=1&toc.id=ss1-2-4&brand=default>.

*Creating Digital Images.* Excerpt from "Introduction", Alan K. Bowman and J. David Thomas. *The Vindolanda Writing Tablets.* Volume III. London: British Museum Press, 2003: 14. Online: <http://vindolanda.csad.ox.ac.uk/tablets/TVdigital-2.shtml>.

Deegan, Marilyn and Simon Tanner. *Digital Futures: Strategies for the Information Age.* London: Library Association Publishing, 2002.

Digital Classicist Projects. London: Centre for Computing in the Humanities, King's College London – Kentucky: Stoa Consortium, University of Kentucky, 2008. <http://wiki.digitalclassicist.org/Category:Projects>.

Engeldrum, Peter.G. *Psychometric Scaling.* Winchester MA: Imcotek Press, 2000.

European Parliament and Council. Recommendation on key competences for lifelong learning. Brussels: European Union, 2005.
<http://ec.europa.eu/education/policies/2010/doc/keyrec_en.pdf>

Grenfell, Bernard. P. and Arthur. S. Hunt. *The Oxyrhynchus Papyri: Part I.* London: Egypt Exploration, 1898.

Heffernan, James. "Ekphrasis and Representation." *New Literary History*, 22.2 (1991): 297–316.

Holst, Gerald C. *CCD Arrays, Cameras and Displays.* 2nd Edition. Winter Park et al.: SPIE Press, 1998.

Hunt, Robert G. W. *The Reproduction of Colour.* Hoboken, NJ: Wiley, 2004.

*Inscriptions of Aphrodisias.* Eds. Charlotte Roueché et al. London: King's College, 2007.
<http://insaph.kcl.ac.uk/iaph2007/index.html>.

*Inscriptions of Roman Tripolitania.* Eds. J. M. Reynolds and J. B. Ward-Perkins. London: King's College 2009. <http://irt.kcl.ac.uk/irt2009/index.html>.

JISC. *Google Generation.* Eds. David Nicholas et al., London: JISC, 2007.
<http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/googlegen.aspx>.

*JISC Digital Media.* [Bristol: University of Bristol] 2010. <http://www.jiscdigitalmedia.ac.uk/>.

Keelan, Brian W. *Handbook of Image Quality.* New York: Marcel Dekker, 2002.

Kenney, Anne R. "Digital Benchmarking for Conversion and Access." *Moving Theory into Practice: Digital Imaging for Libraries and Archives.* Eds. Anne R. Kenney and Oya Y. Rieger. Mountain View (CA): Research Libraries Group, 2000. 24–60.

Library of Congress. *Technical Standards for Digital Conversion of Text and Graphic Materials.* Washington (DC): Library of Congress, 2007.
<http://memory.loc.gov/ammem/about/techStandards.pdf>.

MacDonald, Lindsay and Ralph Jacobsen. "Assessing Image Quality." *Digital Heritage, Applying Digital Imaging to Cultural Heritage.* Ed. Lindsay MacDonald. Oxford: Elsevier, 2006. 351–373.

Mahoney, Anne. "Epigraphy." *Electronic Textual Editing.* Eds. Lou Burnard, Katherine O'Brien O'Keeffe and John Unsworth. New York: Modern Language Association. 2006. 224–240.

Mitchell, W. J. Thomas. *Picture Theory.* Chicago: University of Chicago Press, 1994.

Stokes, Peter. "Teaching Manuscripts in the Digital Age" *Codicology and Palaeography in the Digital Age II.* Eds. Franz Fischer, Christiane Fritze and Georg Vogeler. Norderstedt: Books on Demand 2010. 229–245.

Tawse, Anne. *The Digitisation of Medieval Manuscripts, Success or Failure?* Unpublished MA Dissertation. London: School of Library, Archives and Information Studies, University College London, 2003.

Terras, Melissa. *Image to Interpretation: Intelligent Systems to Aid Historians in the Reading of the Vindolanda Texts.* Oxford: Oxford University Press, 2006.

Tomlin, Roger S. O. "Vinisius to Nigra: Evidence from Oxford of Christianity in Roman Britain." *Zeitschrift fur Papyrologie und Epigraphik* (*ZPE*) 100 (1994): 93–108.

Van Groningen, Bernhard Abraham. "Projet d'unification des systemes de signes critiques." *Chronique d'Egypte* 7 (1932a). 262–269.

Van Groningen, Bernhard Abraham. "De signis criticis in edendo adhibendis." *Mnemosyne* 594 (1932b): 362–365.

*Vindolanda Tablets Online.* Ed. Centre for the Study of Ancient Documents et al. (n.d.). <http://vindolanda.csad.ox.ac.uk/>.

*Vindolanda Tablets Online II.* Ed. Centre for the Study of Ancient Documents and The British Museum [Oxford: University of Oxford], 2010. <http://vto2.csad.ox.ac.uk/>.

Ware, Gene et al. "Multispectral Document Enhancement: Ancient Carbonized Scrolls." *Geoscience and Remote Sensing Symposium. 2000. Proceedings. IGARSS 2000. IEEE 2000 International* Vol. 6 (2000): 2486–2488.

Webb, Ruth. "The Progymnasmata as Practice." *Education in Greek and Roman Antiquity.* Ed. Yun Lee Too. Leiden: Brill, 2001: 289–316.

Warwick, Claire et al. "Documentation and the Users of Digital Resources in the Humanities." *Journal of Documentation* 65.1 (2009): 33–57.

Youtie, Herbert C. "The Papyrologist: Artificer of Fact." *Greek, Roman and Byzantine Studies* 4 (1963): 19–32.

Youtie, Herbert. C. "Papyrology and the Classics." *Zeitschrift für Papyrologie und Epigraphik* (ZPE) 67 (1957): 267–281.