

New
Publication Cultures
in the Humanities

***Exploring the
Paradigm Shift***

Péter Dávidházi (Ed.)



Amsterdam
University
Press

New Publication Cultures in the Humanities

New Publication Cultures in the Humanities

Exploring the Paradigm Shift

Edited by Péter Dávidházi

Amsterdam University Press

Cover design: Maedium, Utrecht

Lay-out: Crius Group, Hulshout

Amsterdam University press English-language titles are distributed in the US and Canada by the University of Chicago Press.

ISBN 978 90 8964 564 7

e-ISBN 978 90 4851 971 2

NUR 600

© P. Dávidházi/ Amsterdam University Press B.V., Amsterdam 2014

All rights reserved. Without limiting the rights under copyright reserved above, no part of this book may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the written permission of both the copyright owner and the author of the book.

Every effort has been made to obtain permission to use all copyrighted illustrations reproduced in this book. Nonetheless, whosoever believes to have rights to this material is advised to contact the publisher.

*In memory of Irma Vogel (1954-2011)
who worked for this project to the last*

Table of Contents

Preface	9
Exploring Paradigms and Ourselves	
<i>Péter Dávidházi</i>	
 The Digital Enterprise: Views Philosophical, Historical and Personal	
Digital Humanities	21
Foundations	
<i>Jacques Dubucs</i>	
Looking Forwards, Not Back	37
Some Ideas on the Future of Electronic Publications	
<i>Gudrun Gersmann</i>	
The Dynamics of Digital Publications	47
An Exploration of Digital Lexicography	
<i>Claudine Moulin and Julianne Nyhan</i>	
Too Much of a Good Thing?	63
Or, A Historian Swamped by the Web	
<i>Luca Codignola</i>	
 Changing Models for Textual Editing in Electronic Publication	
Electronic Textual Criticism	91
A Challenge to the Editor and to the Publisher	
<i>Gábor Kecskeméti</i>	
Computer-assisted Scholarly Editing of Manuscript Sources	99
<i>Andrea Bozzi</i>	
Electronic Media and Changing Methods in Classics	117
<i>Bernhard Palme</i>	

Cutting Edge

New Means of Access, Evaluation and Funding

Publication Practices in Motion	131
The Benefits of Open Access Publishing for the Humanities	
<i>Janneke Adema and Eelco Ferwerda</i>	
The Future of Publications in the Humanities	147
Possible Impacts of Research Assessment ¹	
<i>Milena Žic Fuchs</i>	
ERIH's Role in the Evaluation of Research Achievements in the Humanities	173
<i>Ferenc Kiefer</i>	
Performing Excellence in the Humanities	183
The Funding Initiative 'Opus Magnum' of the VolkswagenStiftung	
<i>Vera Szöllösi-Brenig</i>	
Contributors	203
Index	209

Preface

Exploring Paradigms and Ourselves

Péter Dávidházi

Publication cultures are systems of social practices held together by the need to publish texts and defined by the dominant paradigms of publishing. Appropriated from Thomas Kuhn's theorization of scientific revolutions, *paradigm* here refers to a coherent pattern of communal practice. For Kuhn, a paradigm is either a model derived from actual scientific practice – that is, an accepted example from which a particular tradition of scientific research springs – or, in a broader sense, it “stands for the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community” (Kuhn 10, 175). It is this latter meaning that can be more useful when we seek to understand the far-reaching implications of the current changes in scholarly publishing; indeed, investigators often resort to it without spelling it out or being aware of its use. For example, it seems to have been employed by members of the Modern Language Association (MLA) Ad Hoc Committee in 2000, when they published their classic paper, *The Future of Scholarly Publishing*, aiming to analyse “the widely perceived crisis in scholarly publishing” and to recommend viable solutions (Ryan et al. 172-186); it was their holistic, if latent, notion of a comprehensive paradigm that enabled them to make valid recommendations for all sorts of scholars, funding agencies, university departments, libraries, publishers and administrations. Kuhn's multifaceted term is certainly not simple; nor is its application to disciplines outside the natural sciences obvious – though, as has been shown, it is certainly worthwhile (Masterman, cf. Gutting). The essays in the present volume are talking neither about the paradigms of science, nor of any discipline in the humanities, but rather present paradigms of publication cultures. Hence our subject calls for a further reinterpretation of the term, modifying its theoretical and practical implications. Consequently, what we mean by a paradigm is a set of concepts, habits, technologies, institutional norms and regulations, which together govern, directly or otherwise, all our procedures in publishing.

It is the ruling paradigm that controls the latent implications and tacit assumptions of what is meant by the verb *publish*. It is only by analyzing such paradigms that we can hope to understand the diverse connotations of that verb, far beyond its seemingly obvious sense; that is, to make any

kind of text (a piece of writing, a musical score or something diagrammatic) available in the public domain. That seemingly obvious sense of the word, however, has always been only a fraction of its total meaning. To *publish* comes from the Old French *puplier*, which, in turn, derives from the Latin *publico, publicare*, a verb that could mean not only to make something public, but also to adjudge it to public use, to declare it the property of the state and even to confiscate it. Characteristically, these obsolete connotations have often been revived by some of the modern publication cultures we know. When I started to publish, four decades ago, the meaning of the verb (and its cognates such as the Hungarian *publikál*) in the region of Central and Eastern Europe implied and foregrounded an unavoidable element of state-controlled authorization, even if the criteria of that authorization were much more lenient than they had been in the 1950s. For example, there was no longer an emphasis on, say, the ideological purity of scholarly works on nineteenth-century literature. While the publication of a text has never been a purely practical or technical endeavour, in the 1970s in Hungary it required a symbolic act of political power performed by a publishing house of the *state*. It entailed authorization in the manifold sense of official approval, legitimation and appropriation; it restored to the range of meanings carried in 'publication' the obsolete element of 'confiscation' and, indeed, some age-old implications of *publicare*. In a world almost hysterically sensitive about the dissemination of any information (including something appearing in as rudimentary a format as a stenciled handout for students of English grammar), the entire infrastructure of publishing was still so jealously monopolized by the ruling party that there was only a narrow margin left for the illegal and risky alternative called *samizdat* publishing. An official, if unwritten, ban on a text or its author was so forceful that no publishing house in the country could ignore or violate it. It was not until much later, around 1989, that the unauthorized Western publication of a blacklisted work no longer resulted in severe punishment for the author at home. Compare that regime of publication to the wide variety of possibilities available today in the same region, and you begin to see some conspicuous differences between the respective paradigms at work in the two epochs.

Comparisons like this are small, first steps towards a systematic exploration of our subject; yet, they are indispensable and their details should be subjected to close scrutiny. A paradigm cannot be ascertained only by a facile and often unwarranted assumption of coherence. Therefore, one should heed the warning of Alfred North Whitehead (whose forgotten *Science and the Modern World* was once hailed as the most important

philosophical book on science since Descartes' *Discourse on Method*) that science needs more than just "a general sense of the order of things." It also requires "the habit of definite exact thought," methodologically speaking "the priceless habit of looking for an exact point and sticking to it when found" (Whitehead 23). A mathematician and philosopher, Whitehead was interested in the changing worldviews or conceptual schemes underlying the major changes in the history of science; hence, he was looking for exact points, i.e. the dominant preoccupations, defining the mentality of each period. Our own questions are, what kind of order and what kind of exact points should we be looking for in publication cultures when trying to understand what a paradigm is like and what a paradigm shift is about. It is only through well-focused analyses that we can hope to arrive at 'exact thought.' Most of the papers in this volume, then, are meant to sharpen our view of key elements hitherto neglected. For example, the paper by Milena Žic Fuchs, which tries to reveal something as vitally important as the possible impacts of research assessment on the future of publication in the humanities, starts with the sobering observation that one can find only "fragmentary data or insights on the effects of assessment mechanisms in a small number of disciplines," because we still have very few "extensive, in-depth analyses showing more precisely the effects that the evaluation, whether of institutions or individual researchers, have had on publication cultures." We badly need these in-depth analyses, and once the minutiae of publishing fall into a new pattern, into a composite picture of details we can truly explore, then the task will be to reveal its overall historic significance to our present moment.

Publication cultures have always had their own history, and there is nothing unprecedented about abrupt shifts either. Different epochs have been dominated by different paradigms and there have been exciting periods of transition between them, provoked by a major discovery or other social factors that challenged the established paradigm and called for a new one. The underlying hypothesis of this collection of essays is that we are witnessing one of the most decisive paradigm shifts in history. As was recently noted by Jerome McGann, it can be compared in magnitude to the fifteenth-century printing revolution that the Renaissance thrived on. Yet, if we want to measure how radical a shift the new digital technology is driving in the humanities we need to consider the inexorable fact that "the entirety of our cultural inheritance will have to be reorganized and re-edited within a digital horizon." (McGann 2010) Thus, we must revise the rationale of a publication culture that has been taken for granted during several centuries of relative stability. Veiled by the present upheaval, a

number of simultaneously existing, even competing, paradigms are at work and our age is still trying to come to terms with its own possibilities, in pursuit of a master paradigm that will likely incorporate or marginalize the others without eliminating them altogether. The time has not yet come to describe that master paradigm, but we can explore the paradigm shift itself, its phenomena and our responses to it. As the new modes of publishing are bound to affect the future of research and our scholarly communities are not yet ready to live up to the new challenge, the goal of this collection is to facilitate a more profound understanding of what this paradigm shift is about and to reveal how to make the most of it in the humanities. At a time when severe budget cuts are jeopardizing the development, if not the sheer survival, of the humanities in most European countries, it is vital for us to find potential resources hidden in the emerging new paradigm.

At the same time, these papers detect and analyze the probable causes of our own reluctance to make best use of the digital turn, urging us to overcome our difficulties, be they technical, psychological, or both. Although the new digital world no longer looks as formidable as it used to, when its newly discovered technology elicited knee-jerk or even “neo-Luddite” reactions (Himmelfarb), its pace of development is still frightening enough for many a traditional-minded scholar. From the first group of papers to the last, this book is meant to fulfil its dual function: to facilitate the understanding of the changes and to try to dispell this paralyzing sense of fear by mediating between (to revive a dead metaphor) the cutting edge of science and the thin skins covering some of our best colleagues. Such mediations are vitally important because, as Peter Givler, executive director of the Association of American University Presses, pointed out in 2000, scholars at work today have to cope not only with the usual difficulties of their scholarship but also with developing new vehicles for the publication of their results (Ryan et al. 180-181). Moreover, as the young can master the use of the new electronic devices more quickly and easily (on the whole) than their elders, the digital turn has tended to aggravate the generation gap, always latent in the scholarly community, and the temptation to look at professors of immense traditional knowledge as mere fossils of the past has become greater than ever. Hence we, authors and editor, are convinced that no examination of our rapidly changing publication cultures can suffice without due attention to the human problems involved, because it is vital to minimize the inevitable collateral damage caused by the triumphant progress of the Digital Humanities.

The exploration of an ongoing paradigm shift requires the courage to experiment with the transitory; to ask hypothetical, even counterfactual, questions and to take the risk, greater than usual in research, of being proven

wrong. Facing this challenge, the present collection of essays unwaveringly follows the logic of questions asked by various groups of scholars in our team. The first major unit, titled *The Digital Enterprise: Views Philosophical, Historical and Personal*, enquires into the general nature of the new modes of publishing. This preliminary section focuses on the transition from paper-based to digital publication in the humanities, devoting papers to the philosophical foundation of Digital Humanities (Jacques Dubucs), to the exciting promises of electronic periodicals (Gudrun Gersmann), to the dynamics of digital publication as exemplified by Digital Lexicography (Claudine Moulin and Julianne Nyhan) and to the provocative question of whether the Web could be overwhelming for the historian or any scholar (Luca Codignola). From various angles, these papers highlight what is at stake now and demonstrate that the digital enterprise, far from being a merely technical issue, is a unique opportunity in the humanities: we can both preserve and renew our ideals of distinction by integrating a great (though in bytes relatively small) cultural heritage in the huge upsurge of miscellaneous digital information. Arguing that digitization is not immaterialization but rematerialization, Jacques Dubucs maintains that the enhanced reproducibility of works of art (now far beyond what it was in Walter Benjamin's time) need not be paralyzing. Moreover, it is the mission of the Digital Humanities to avert the dangers of interpretive anarchism. In a paper suggestively titled "Looking forward, Not Back: Some Ideas on the Future of Electronic Publications," Gudrun Gersmann argues that the publishing culture that has dominated academic publication for a century, whereby a few authoritative scholarly journals publish the work of a few distinguished specialists in their field, is no longer unchallenged, mainly because the voice of experts has now been joined by an ever-growing chorus of opinions. Furthermore, the old vertical and hierarchical communication is being replaced by a new communication both horizontal and fluid, transforming the self-image of the humanities. The case study offered by Claudine Moulin and Julianne Nyhan analyzes the digital remediation of the German and Luxemburgish dialectal lexicography, showing how this kind of material may benefit from and contribute to a global information space, such as the emerging Semantic Web, and why it is vital to train some Humanists to understand such new digital tools well enough to bridge the gap between the respective communities of traditional scholarship and e-science. Luca Codignola's paper investigates how profoundly and in what stages the profession of historians has changed with the advent of the Web and other technological inventions. It takes a long, cool look at the trajectory of the ensuing methodological developments in order to ascertain their

diverse, complicated and mostly (but not altogether) beneficial impacts on the quality of historical research and writing.

The second unit, *Changing Models for Textual Editing in Electronic Publication*, analyzes the ways that editing, one of the oldest practices in the Humanities, is being altered by new media, which seeks to replace the old ideal of perfection (supposedly realized, or at least targeted, by the definitive edition) with an ideal of continuous perfectibility (accomplished by the new digital editions). This section starts with a succinct demonstration of how new and much-needed electronic textual criticism challenges both the editor and the publisher (Gábor Kecskeméti), how effectively the computer can assist the editing of manuscripts (Andrea Bozzi) and how the electronic media is changing methods in classical philology (Bernhard Palme). Gábor Kecskeméti highlights the difference between the steady and manifold usability of the digital edition, volatile or evanescent as it may seem, and the dead-end of its traditional paper-based counterpart, especially if it has no renewable electronic basis. Addressing scholars who work with some of the earliest sources, the papers by Andrea Bozzi and Bernhard Palme illustrate how the newly invented technical devices can be applied to a wide range of textual problems in fields as diverse as late Aegyptian papirology and modern philology, and how they can sustain and even improve fields of study instead of threatening their survival. Explicitly or otherwise, all these papers endorse the reorientation of research as a fundamentally collective and (by definition) unfinished enterprise.

The third unit, *Cutting Edge: New Means of Access, Evaluation and Funding*, charts the possible advantages of some brand new or recently discovered tools of our trade. Here, we learn about the great benefits of Open Access publishing for humanities scholars who make the shift from print to digital publication (Janneke Adema and Eelco Ferwerda); about the far-reaching impact of the new means of research assessment (Milena Žic Fuchs); about the possible roles of ERIH, the European Reference Index for the Humanities, which was initiated with due caution by the European Science Foundation's Standing Committee for the Humanities (SCH) (Ferenc Kiefer, one of the founding fathers of ERIH); and about an important funding scheme that has been masterminded and implemented in order to provide a financial basis for the sustained scholarly work required by great monographs (Vera Szöllösi-Brenig). The OAPEN (Open Access Publication in European Networks) project is shown against a background of transition phenomena: the current communication and publication practices of scholars just getting acquainted with their new possibilities; the slowly (or quickly?) eroding authority still attributed to printed books as opposed to e-publication; the

increasing consumption of online scholarly information and the growing number of scholars who accept Wikipedias, not uncritically, but at least as a supplement to traditional sources of information. ERIH could be listed here as well: its great importance and its highly controversial status was indicated by the clash of metaphors traded at our SCH meetings and our 2009 Budapest conference, one of the metaphors intimately endearing, the other apocalyptically frightening: ERIH is a lovely baby of ours, dirty enough to need a good bath, yet not to be thrown out with the bathwater; but, it is also a highly dangerous invention, a potential loose cannon or even a weapon of mass destruction. The two papers devoted to the subject in our volume abstain from such extreme analogies but painstakingly analyze the prospects of this device together with the proper function and possible benefits or hazards of bibliometrics in general, knowing that the future of the Humanities is at stake. They help us realize that bibliometrics, all too frequently and too crudely used today to measure the 'impact' that constitutes a norm of 'fundability' in several European countries, is an instrument that can be refined and which should be used together with complementary methods to foster high quality publication in the Humanities.

The fourth paper in this unit concerns the problem of how to protect the future of monographs; that is, books written by one person, based on sustained research and focusing on a single, usually complex, subject. Though traditionally one of the most important publication genres in the Humanities, so much so that we could easily summarize the history of practically every discipline in the Humanities by enumerating the most important monographs as its milestones, the monograph has been unduly ignored recently by both the funding and the monitoring schemes of most European countries. Thus, it has become increasingly difficult to secure the adequate institutional, financial, even psychological conditions needed for its survival. One of the anomalies of the genre's academic position in Europe remains similar to one diagnosed by the MLA Ad Hoc Committee in the US in 1999-2000; the monograph is the 'holy grail' for those waiting for tenure or promotion; yet, publishers, in many ways constrained, cannot find the means for assessing and publishing them (Ryan et al. 172-186) and the number of monographs purchased by libraries nowadays is less than one fifth of what it was in the 1970s (Greco and Wharton 2008, cf. Adema and Ferwerda in the present volume). It seems that foundations have patterned their scheme of support on the requirements of the natural sciences: they tend to favour collaborative research. Likewise, most instruments developed for quality assessment of scholarly production have failed to take monographs into consideration, let alone give them adequate weight. This problem has been

aggravated by the institutional pressure to produce quantities of measurable output; scholars have to write short pieces for conferences where they are given 20 minutes to speak: enough for a paper-to-be but not conducive to writing a book-length study. In countries where the 'sabbatical' is not yet established, it is especially difficult for a scholar to escape from teaching duties for a semester in order to prepare or complete a book. In search of remedies for these ills, our last paper, by Vera Szöllösi-Brenig, examines the funding initiative 'Opus Magnum' from the VolkswagenStiftung, a promising device to save the endangered academic species that an optimistic analysis recently called a rising phoenix (Steele 2003).

Though all the authors of this volume are related to the humanities, the perspective they jointly offer is interdisciplinary, combining the insights of philosophy, linguistics, historiography, literary scholarship and lexicography, knowledge acquired as university professors, researchers, editors, publishers or representatives of a founding agency. True interdisciplinarity, however, can only be achieved by the masters of each discipline. Hence, the authors of this volume have been selected with care from among the eminent practitioners of diverse professions in many regions of Europe. My initial idea for this project, though considerably narrower in focus, was conceived in 2008 at a brainstorming session of the Standing Committee for the Humanities (SCH) at the European Science Foundation (ESF) and the proposition was soon accepted. The first contributors, then, were the volunteering members of that Committee, selected and delegated by their respective countries. But we also needed to invite further researchers from outside the ESF, via networking, because of their special expertise. Although the SCH has always been devoted to a 'bottom-up' science policy – that is, to letting the researchers themselves take the initiative in proposing new projects, and to confine its own role to assessing, ranking and selecting the proposals for whatever financial support was available – it sensed a growing need for a supplementary 'top-down' strategy; that is, taking the initiative, not least because members of the SCH were themselves scholars with insights of their own. Their Science Policy Briefing of 2011, published under the title *Research Infrastructures in the Digital Humanities*, succinctly yet amply documents their initiative to bridge physical and digital research infrastructures, to secure their sustainability, to evaluate the outputs of digital research and to educate new generations of scholars so that they master the computing skills required. In the meantime, we were also eager to study the specific manifestations of the paradigm shift in the context of our respective disciplines, and each of us had the opportunity to write papers based on individual research (e.g. Dávidházi 2013). Thus, our project

has evolved through these phases: adopted by the ESF in 2008 and adapted by an international team of scholars, it was developed at the ESF 'Strategic Workshop' held at the Hungarian Academy in Budapest on 27-28 November 2009; then, at the 'Humanities Spring,' a forum for young European researchers organized in Maynooth on 9-11 June 2011 to inspire their common manifesto on the changing publication cultures in the Humanities (Kelleher and Hoogland 2012); and finally, through the papers collected in the present volume. Meanwhile, from 2008 to 2014, our subject developed faster than practically any other we have ever dealt with, so we had to learn faster than ever, revising and updating our papers several times until the very last minute and, let the tormented editor sigh, far beyond the appointed deadline.

Finally, I want to thank all those who helped this project from its inception; first and foremost, the fellow members of the ESF Standing Committee for the Humanities, a scholarly community I was part of between 2004 and 2009. This book is indebted to Milena Žic Fuchs, the chair of the Committee from 2009 to 2012, who has staunchly supported the project from the beginning. I am grateful to the previous staff of the ESF office, especially for the inventiveness of Monique van Donzel and Rüdiger Klein, for the help received from the angelic and unforgettable Irma Vogel (1954-†2011) as well as to its recent staff, especially Arianna Ciula and Claire Rustat-Flinton, both of whom worked for the project at crucial stages of its development, and Nina Kancewicz-Hoffmann, who secured the financial support through difficult times of institutional transition. The 2009 Budapest conference, the occasion for the first draft of several papers in this volume, was carefully looked after by Villő Denke, at that time one of the secretaries at the Hungarian Academy of Sciences. A dialogue with Carol Chillington Rutter, fine Shakespeare scholar and ever-helpful friend, helped to clarify some of my ideas for this preface.

A special tribute is due to the former chair of the Committee, Gretty Mirdal. A clinical psychologist supervizing the treatment of refugees who had been victims of torture, yet also a professor of psychology deeply versed in literature, she was our guiding light from 2004 to 2009. Many of us will always be grateful for having been transformed by her transcultural wisdom, tactful benevolence and quiet serenity. This book would not have materialized without her inspiring presence in our thoughts. Her leading role in the Committee when we were composing our 'Position Paper 2007' was epitomized by its starting point that "self-reflection is at the root of the humanities" (ESF 2007, 5), and the legacy of this insight shaped the present project as well, causing us to realize that we cannot explore paradigms without exploring ourselves.

Bibliography

- Dávidházi, Péter. "Redefining Knowledge: An Epistemological Shift in Shakespeare Studies," *Shakespeare Survey* vol. 66, ed. Peter Holland, Cambridge: Cambridge University Press, 2013, 166-176.
- ESF Standing Committee for the Humanities. *Position Paper 2007*, Strasbourg, 2007. www.esf.org.
- ESF Standing Committee for the Humanities. *Science Policy Briefing 42: Research Infrastructures in the Digital Humanities*. September 2011. www.esf.org/publications/science-policy-briefings.html.
- Greco, Albert N. and Robert Michael Wharton. "Should University Presses Adopt an Open Access Business Model for all of their Scholarly Books?" *ELPUB. Open Scholarship: Authority, Community, and Sustainability in the Age of Web 2.0 – Proceedings of the 12th International Conference on Electronic Publishing held in Toronto, Canada 25-27 June 2008*, eds. Leslie Chan and Susanna Mornati, 2008.
- Gutting, Gary (ed.). *Paradigm and Revolutions: Appraisals and Applications of Thomas Kuhn's Philosophy of Science*, Notre Dame and London, University of Notre Dame Press, 1980.
- Himmelfarb, Gertrude. "A Neo-Luddite Reflects on the Internet." *Chronicle of Higher Education* vol. XLIII n. 10 (November 1, 1996) A56. Web.
- Kelleher, Margaret, and Eva Hoogland (eds.). *Changing Publication Culture in the Humanities: Young Researchers Forum ESF Humanities Spring 2011*, Strasbourg: Ireg, 2012. www.esf.org/fileadmin/Public_documents/Publications/Changing_Publication_Cultures_Humanities.pdf.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*, Second, enlarged edition, Chicago: The University of Chicago Press, 1970.
- Masterman, Margaret. "The Nature of a Paradigm," in Imre Lakatos and Alan Musgrave (eds.) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 1970, 59-89.
- McGann, Jerome, (ed.). "Introduction. Sustainability: The Elephant in the Room." *Online Humanities Scholarship: The Shape of Things to Come. 26-28 March 2010, University of Virginia-Connexions*. 8 May 2010. Web. Accessed 18 July 2011. <http://cnx.org/content/col111-99/1.1/>.
- Ryan, Judith, Avelar, Idelber, Fleissner, Jennifer, Lashmet, David E., Hillis Miller, J., Pike, Karen H., Sitter, John and Lynne Tatlock. "The Future of Scholarly Publishing," *Profession* 2002, 172-186 or www.mla.org/pdf/schlrlypblshng.pdf.
- Steele, Colin. "Phoenix Rising: New Models for Research Monograph?," *Learned Publishing* 16 (2003) 111-122. Web.
- Whitehead, Alfred North. *Science and the Modern World*, Harmondsworth: Penguin. 1926. 23.

**The Digital Enterprise:
Views Philosophical, Historical and Personal**

Digital Humanities

Foundations

Jacques Dubucs

We must keep in mind some numerical data when we evoke the transition from the paper to the digital age. In particular, the following contrast speaks for itself:

1. All the books ever written represent 50 billion bytes.
2. The information produced in 2006 represents 150 quintillion (150×10^{18}) bytes. That is to say, during 2006 alone, the world produced three million times the informational content of all the books ever written.
3. Things continue in this way at high speed: the only internet track of May 2009 has generated 500 billion bytes.

Thus, our paper-based heritage is already a tiny fraction of what the human race has produced and this fraction decreases, relatively, every day. Viewing these data, the conception of a digitization enterprise should be thought of and considered by humanists as enlarged. The narrow acceptance of the project – the view that it is merely a technical process of converting our paper-borne heritage into electronic form – is dramatically insufficient. To paraphrase Clemenceau's famous words about war and militaries, digitization may be too serious a thing to be left to the digitizers alone. Scholars must face the issue and understand it as one of the most important problems they have to deal with and, as I will argue, as a real opportunity to renew their practices and disciplines.

To start with some affective considerations, which, after all, matter too, digitization is a major source of anxiety for the Humanities. Not least because the world of computers was radically strange to scholars, many of whom spent a long time using computing machines and word processors often in quiet ignorance of the underlying processes. Moreover, there is some resistance from scholars, long familiar with browsing, annotating and cherishing their books, to iPads and other devices, which allow similar practices and may even evoke similar feelings in the process. Deeper than that, though, they are disturbed by the idea that, along with the possible evanescence of the final printed output, the material signs of distinction and *auctoritas* may also disappear. For digitization is also a transition from a universe of tangible books or academic journals, whose recognizable appearance

evidences interest in content and the dignity of authors, to an undifferentiated situation of Big Commensurability, deserted by the immediate material marks of reliability and reading-worthiness. The dramatic issue is: what is the future of philological probity and other humanistic ideals in a world where the ultimate measure is that of the byte, according to which, millions of contentless electronic messages have greater weight than the Quixote? To cope with this central question, we should perceive digitization not as a technical issue, but as a whole process of appropriately integrating the tiny hand of our cultural heritage in the ocean of digitally-borne information.

1. The Digital Empire

It is a commonplace that the ways of doing and communicating science are rapidly changing under the ubiquitous influence of computers. The widespread use of these machines results in an epistemic enhancement. This enhancement has several aspects. First, and most manifestly, it consists of an extension of our native capacities as paper readers: we can now be acquainted with many more items than those we previously had access to. Overcoming the necessity of travelling to libraries and of waiting for book delivery, our new databases and repositories lead us to a position of consulting and browsing more material than before. In this way, we are doing approximately the same job but at higher speed, at a wider range and in greater comfort. Second, and more importantly, this enhancement may be a qualitative augmentation of the scope of scholarship, giving us access to features of cultural heritage that we were not equipped to detect in previous times. The frequency of occurrences of given words in wide corpora, the dynamics of their evolution or the regularities of their coexistence with other words are typical features of this kind. Third, the treatment and analysis of such massive data often involves conversion of the modalities of our work; for example, when we are led to visualize these data by means of curves or other iconic ways of representation. In such cases, the very task of reading cannot be done by the human eye and it is left to machines: we simply deal with the result of their work, using a radical approach to our cultural heritage that was previously unavailable.

To grasp the meaning of the transformation just described, scholars should firstly realize that their field is not as specific as they once believed, and that the Digital Turn currently affects the whole of scientific enterprise in a similar way. The Galilean ideal of reading Nature in mathematical language; namely, of understanding and foreseeing it in an analytical way, by means of appropriate differential equations, is now in question. To provide some

limited examples, the dynamics of fluids has become a largely computerized discipline, not only because we use computers to solve numerically equations that are analytically unsolvable in most cases (this use of computers corresponds to the extensive meaning of ‘epistemic enhancement’ above), but also because the detail of a turbulent flow cannot be suitably represented and analyzed except via a super-computer.’ The same consideration of the complexity of data issue applies, even more eloquently in the present context, to the ‘reading’ of the human genome, which can obviously not be done by man alone: the database currently used encompasses more than five times the amount of information contained in the Library of Congress. To sum up, the most intriguing features of the Digital Humanities (DH) are currently shared with the classical domains of the hard sciences. One simply has to ask, in the general landscape of this New Kind of Science, what is the specific object of the humanistic brand of this general evolution?

2. *Philology Aufgehoben*

The best way of characterizing the Digital Humanities is probably to contrast them with the tradition of the past half-century in ‘continental’ philosophy. People in this trend used to describe their activity as ‘interpretive’ or ‘hermeneutic.’ Considering that the Masters of Suspicion (Marx, Nietzsche and Freud, to quote some of them) had definitely deconstructed the traditional idea of philosophy as a variety of the pursuit of truth, they have emphasized the interpretive nature of any enterprise of understanding. The salient issue, in this perspective, is the adjudication of the standards of correctness for the interpretations. Either one faces the problem seriously, and one must, therefore, be prepared to admit an exception to the move beyond True and False. Or, one assumes integrally the relativist motto of hermeneutics, and one has to make interpretation itself subject to interpretation, not to a justifiable verdict of correctness or incorrectness. The second horn of the dilemma has been clearly preferred, positing the idea that any interpretation is legitimate, insofar as it is expressive enough of the idiosyncratic situation of the interpreter: briefly expressed, interpretation becomes a genuine act of creation by the reader. No surprise, then, that in these conditions philology or ‘objectivist’ scholarship were dismissed in recent times: their relegation is on a par with the fashionable idea that the meaning of a text is just what the reader attributes to it:

The relegation of writing to the indeterminate and endlessly transforming processes of textual dissemination is a by-product of Saussurian linguis-

tics and some of the structuralist theories built upon it. In privileging the structures of speech over those of script, it displaced the older, text-based, philological, diachronic study of language, in favour of purely synchronic analysis – how people talk now. This shift in attention away from the study of historical process makes it easy to conclude that we cannot really presume to recover an authorial voice at all, or an intended meaning, from the written or printed records of it. We are left only with synchronic structures, and the conventions which regulate their meaning as we read. It follows, of course, that if the meaning we read is entirely a function of the structural relations within the verbal sign system which constitutes a text, then it is not something inherent which can be expressed at all. Meaning is not what is meant, but what we now agree to infer (McKenzie 42-43).

3. Rematerialization

Digitization is often regarded as immaterialization. The presiding spirit of the assimilation is simply this: books are solid and available, whereas their digitized counterparts only enjoy an intermittent and mediated existence, at the moments when, and on the condition that, the relevant electronic devices are powered up. Although it corresponds significantly with daily superficial experience and feeling, the immaterialization view is wrong in many senses. First, digital information is not floating above the material world; rather, it is always physically implemented, albeit in another manner than paper information (if this was not the case, the nagging question of the so-called obsolescence of the electronic storage devices would disappear). Second, and deeper, digitization, when strictly and adequately conceived and deployed, locates the texts in the landscape of a fine-grained ontology, closer to the usual material ontology than the paper publication does. This point, which is of fundamental significance for the Digital Humanities, needs some explanation.

Currently, a copy of Twardowski's famous essay 'Actions and Products' is on my desk, open at the initial page, and I am reflecting on the ideas defended in this text and on their relevance to the topic of the Digital Humanities. It appears as if the book - I mean, my copy of the book, the book as a physical object – connects me with the thoughts Twardowski had a century earlier. Now, the question is that of the nature of the relation. No miracle, of course, is involved, because the relationship between myself, as a material creature, and the products of the mental activity of a philosopher I have never met is mediated by the book. I am physically related to my copy of that book, the inked inscriptions on the book impress my retina, I recognize in those

inscriptions familiar words and sentences and, as a competent English-reader, I access the linguistic meaning of those sentences, which delineate the thoughts Twardowski had in his head at the time he wrote them. This is a well-known story: our capacity to access the thoughts of past thinkers is to be decomposed in (i) our perceptual capacity to access the token of the words they drew on paper; and (ii) our linguistic capacity to access the meaning of those words (now considered, not as concrete inscriptions or ‘tokens,’ but as linguistic ‘types’). Nevertheless, crucial details of this familiar story differ, according to whether we are dealing with printed books or digital devices.

To say that I have the text of Twardowski’s essay at hand is somewhat misleading. Of course, this assertion is always misleading, except when uttered by the happy few who have the privilege to leaf through Twardowski’s manuscript, contained in the Library of the Institute of Philosophy in Warsaw (Fig. 3). As for the other readers, the author has not marked the signs they read himself. The physical product of Twardowski’s writing activity in Lvov lies in a blue-grey dossier T.16 in Warsaw library and nowhere else and, strictly speaking, we cannot talk of it being displayed on the pages of a printed book or on the screen of a computer. Nevertheless, the assertion about my possession of his text is particularly unacceptable when I am reading a copy of the book. The point is not that Twardowski never had any physical contact with this object – this is trivial – but rather that the material chain that connects the both of us and which allows me to access his thoughts has been seriously broken.

Twardowski had a wide command of many languages beyond Polish, but he had never written anything in English and, therefore, he has not marked anywhere any sequence of signs equiform to those in front of me. The text I have on my desk is just a translation of the Polish original and the relevant causal chain has, therefore, been broken or dashed by the translation process. And what if we suppose that the book on my desk is the Polish version (*Twardowski, Wybrane pisma filozoficzne*), republished in 1965 in Warsaw along with other writings by the author. The problem remains the same: Twardowski never collected his works in this way and the logical structure of this book as well as its physical shape cannot be considered as genuinely derivative from his intentions and activity. This, then, is the point where paper and digital publication crucially differ. This difference may be explained by Fig. 2.

The traditional circuit (in red on the figure) of paper publication is rather simple, albeit sometimes complicated to execute. One or another of the sources (S_1 , S_2) are selected and a subsequent book B_1 is published, providing its reader with words and sentences that are type-equivalent to their counterparts in the sources. The organization and the structure of this material are left to the appreciation of the publisher and they are not supposed to be



Fig. 1: Twardowski's archive, Library of Warsaw Philosophy Institute.

similar to the organization and structure of the sources. *A fortiori* the same applies to the physical shape of the original, which is neither conserved, nor recoverable from the printed book. In the case of a scholar edition, an *apparatus criticus* separate from the text refers to the source and discusses the other printed versions on this basis. With some rare exceptions (the best example of which may be *Kritik der reinen Vernunft*, whose 1781 and 1787 editions differ significantly, specifically in relation to some well-known passages), only one source is displayed in the main text and the others are simply referred to in footnotes. In other words, the object one has in hand, the printed volume, is just one book, in the sense of it being the output of the publication process from a given source. The relationship between B_1 and B_2 , which are considered as material objects, is not the physical cohabitation in one mundane item, but rather their proximity on the shelves of our libraries.

The situation is thoroughly different with Digital Humanities editions, from the perspectives of their relationship to their sources and also of their mutual relations. The circuit (in blue on the figure) differs from the previous one in terms of the interposition of a new level between the source and the final output; namely, that of the canonical encoding TE of the properties of the source, which extend beyond its literal content properly said. Far from being a continuous recording indifferent to the fine structure (page-breaks, etc.) of the source or even to its physical appearance, this operation keeps track of everything potentially relevant in the source,

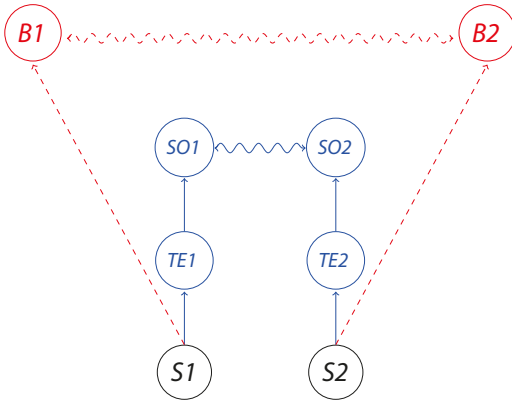


Fig. 2.

making, in principle, any consultation of the original by the reader pointless. The process can be extended to the manuscript sources, as shown in the example of Twardowski's essay. The first lines of the manuscript in Fig. 3 are encoded as follows in XML/TEI standard:²

```
<teiHeader>
<text>
<body>
<pb n="1"/>
<fw place="top-center" type="pageNum">71</fw>
<head>
<title>Fonctions et produits</title>
</head>
<p>
<lb n="1"/>Quelques remarques
<subst>
<add place="above">des</add>
<del>touchant les</del>
</subst>
confins
<lb n="2"/>de la psychologie, <del>la</del> grammaire
et <del>la</del> logique
</p>
<p>
<lb n="3"/><hi>&]167;1.</hi> En comparant deux expressions
</p>
```

[Page 13]71

Fonctions et produits

Quelques remarques des touchent les confins
de la psychologie, la grammaire et la logique

§1. En comparant deux expressions comme ~~appareillées (sic) jointes de en paires appariées (sic)~~ juxtaposées en paires comme celles-ci : + marcher-marche, courir-course, sauter-saut, crier-cria, chanter-chant, parler- ~~parie~~ parole penser-pensée, errer-erreur, juger-jugement, tourner-tournure, nous on voit que la première signifie une fonctions action ; quelle est la significations de la seconde en rapport à la signification de la première – c'est la question, dont s'occuperont les considérations suivantes.

Fig. 4: Screen View of the French Version of Actions and Products after TEI Encoding.

This code integrates not only the author's words, but the concrete way he wrote them on his page (e.g. the centring of the title is reflected on line 4, the deletion of the articles before the words *grammaire* and *logique* is recorded on line 15, etc.). In short, the TE level keeps track in coded format of every feature of the source, including the material aspects that are neglected in the traditional edition. This long-established method is more or less guided by the famous slogan WYSIWYG: What You See – you typist, on your screen – Is exactly What You Get in printed form. From the Digital Humanities standpoint, this result is unacceptably poor, because the editorial output is only what one sees in the main course of the text, obliging us to add separately a lot of paraphernalia to deal with the remaining text. One could say that, in this case, the leading slogan is the opposite of WYSIWYG; namely, WYKIWAW: What You Keep – you, editor – Is exactly What the Author Wrote. Moreover, the connectedness of the screen outputs SO (Fig. 2) for different sources or variants is physically enforced by the versatility of electronic devices: suitable hyperlinks are enough to ensure the possibility, for the reader, to navigate between the corresponding windows without changing his environment in the slightest way.

One can summarize the difference between the two publishing schemes as follows: the new scheme firmly re-establishes the chain between the author and the product of his activity, which was threatened by the traditional way of proceeding. This restoration has wide consequences, often unperceived or misapprehended.

4. The End of the *Reader-shift*

Prima facie, the enterprise of digital edition involves a new avatar of the technical means of reproduction of the cultural works Walter Benjamin analyzed in his most influential essay ‘The Work of Art in the Age of Mechanical Reproduction’ on the philosophy of culture. Benjamin’s train of thoughts is well-known: the quality of the original presence – of the “aura,” in his terms – of a cultural item is always deprecated, even destroyed, by its reproduction. The impetuous development of reproduction technics in modern times induces a “sense of the universal equality of all things” (223), which is a mark of the contemporary mass culture.

Benjamin even seems to have put his finger, in a premonitory manner, on the Big Commensurability characteristic of the Digital Age, with his description of a certain erosion of the very distinction between readers and writers:

For centuries a small number of writers were confronted by many thousands of readers. This changed toward the end of the last century. With the increasing extension of the press, which kept placing new political, religious, scientific, professional, and local organs before the readers, an increasing number of readers became writers – at first, occasional ones. It began with the daily press opening to its readers space for “letters to the editor.” And today there is hardly a gainfully employed European who could not, in principle, find an opportunity to publish somewhere or other comments on his work, grievances, documentary reports, or that sort of thing. Thus, the distinction between author and public is about to lose its basic character. The difference becomes merely functional; it may vary from case to case. At any moment the reader is ready to turn into a writer (‘The Work of Art in the Age of Mechanical Reproduction,’ 231-232).

From this perspective, the cultural impact of the contemporary Web might be viewed as a magnification, at a huge scale, of the effects that the daily press had a century ago; namely, the loss of distinction, authority and sacrality of the written work.

As far as written works are concerned, Benjamin’s analysis of the effects of reproducibility can be challenged. The first objection is that if reproduction itself is to be incriminated, then the fatal turn should have been taken, not in contemporary times, but centuries ago by Gutenberg. This conclusion, which Benjamin seems ready to accept – “printing, the mechanical reproduction of writing [...] is merely a special, though particularly important,

case [of the phenomenon examined here]" ("The Work of Art in the Age of Mechanical Reproduction' 218-219) – does not fit well with the background of the whole essay; namely, that of a mutual harmony between capitalism and mechanical reproducibility: the anachronism was unacceptable, even for one ready to accept that "the perspective of world history" (Loc. cit.) affords a certain liberty with historical accuracy. Besides, Benjamin claims insistently that the possibility of traceable ownership should be considered as one of the characteristic marks of an original, not a reproduced, piece of culture:

Even the most perfect reproduction of a work of art is lacking in one element: its presence in time and space, its unique existence at the place where it happens to be. This unique existence of the work of art determined the history to which it was subject throughout the time of its existence. This includes the changes which it may have suffered in physical condition over the years as well as the various changes in its ownership. The traces of the first can be revealed only by chemical or physical analyses, which it is impossible to perform on a reproduction; changes of ownership are subject to a tradition which must be traced from the situation of the original ("The Work of Art in the Age of Mechanical Reproduction,' 220).

Books can, of course, be such items, as lyrically described by Benjamin himself in a short essay dedicated to book collecting called 'Unpacking My Library'. To conclude, printed books, or at least some of them, should not be considered as inauthentic replicas of "auratic" sources, but rather as original pieces of culture by themselves. This changes everything, because:

1. the significance of such books can no longer be viewed as only derivative of a first, inaugural (handwritten) act of creation.
2. there is no compulsory reason to consider different editions, or even different copies of the same edition,³ as just redundant tokens or realizations of the same abstract type.

Once we have rejected, in the domain of written culture, Benjamin's motto of the absolute privilege of the punctual, inaugural *hic and nunc*, the binary opposition between the 'auratic' start-point and the replications of this sacrosanct origin vanishes in the same way. The site of textual authority can no longer be equated with the only Ur-Text. The sources of reading-normativity are distributed in the subsequent proliferation of versions or variants and their readers are, therefore, not left to appreciate and to judge the erratic circumstances of their activity. In sum, one can escape the

dramatic consequence of Benjamin's conception in the domain of textual culture; namely, that people remote from the origin are obliged, and then free, to adopt whatever interpretation of their cultural legacy. To avoid this interpretive anarchism, according to which 'everything goes' for the contemporary reader, is precisely the aim of the Digital Humanities.

5. Interpreting Texts and Cultural Artefacts: The DH Stance

As has been remarked (Fitzpatrick), current practitioners of the Digital Humanities deal with such a number of technological and institutional issues that one might consider this to be making (archives, tools, methods or consortia), rather than interpreting. Some (including Alvarado) evoke the temptation of simply describing the distinctive nature of the Digital Humanities by pushing forward the practical idiosyncrasies of the people working in the field – who have a minimal acquaintance with the traditional Humanities, they use computers, they collaborate in an interdisciplinary way, and so on. This was a plainly desperate move, analogous to that of defining chemistry as merely the scientific field of those who manipulate retorts and test tubes.

The foundations of the difficulty lies, beyond the fact that Digital Humanists are supposed, to reuse Sokal's phrase, to "transgress the boundaries," in the inconsistency of the following data:

1. The hope, or wish, or certainty, that nothing should prevent the Digital Humanities from dealing with digital production as well as with paper-borne heritage.⁴
2. The fact – the Big Commensurability analyzed above – that, in the digital realm, no watertight distinction between authors and readers can survive: in digital environments readers become prolix and give free and public play to their verbosity, becoming authors as well and, moreover, indiscernibly so.
3. The normative thesis that humanistic scholarship rests on a fundamental asymmetry between authors and commentators, the latter trying to grasp the meaning of texts and being apt to be right or wrong in doing so.

In sum, and paradoxically, the Digital Humanities seems to make sense only as far as paper-legacy is concerned, that means in the domain of computer-aided Humanities or, as one used to say, the 'computing Humanities.'

The solution to this difficulty might be found in analyzing the canonical author/reader asymmetry more deeply than usual. Given the lack of clarity in the very notion of the 'meaning of a text' – while we have at hand a reasonable

analysis of the meaning of a word, or of a sentence, it is highly questionable whether a rigorous counterpart to this is available for a text today – the simplest way is to start at the sentential level. According to the traditional picture, the author uses a certain sentence σ to express, or to refer to, some abstract entity A , which is σ 's meaning. The duty of the reader is to recognize that σ has that meaning. He is reading σ as meaning B , and he is right if $B = A$.

As simple as it may appear, this picture raises a series of difficulties. The first lies in the assumption of the pre-existence of the entity targeted by the author; as it were, an inert, sleeping entity waiting to be woken up by someone's statement. This view is not entirely indefensible: it has been notoriously advocated by Bolzano, who conceives his '*Sätze an sich*' in this way, i.e. sentences that are what they are independently of any speaker who is able to give them a linguistic form in his or her idiom. Nevertheless, one would prefer, *ceteris paribus*, to dispense with this metaphysical view of language, where authors are rarely considered either active or creative. As soon as we dismiss this kind of conception, we should also abandon the deciphering model of interpretation: readers are not trying to decode an author's message, for there is nothing for an author to encode.

Turning to more mundane entities and prosaic analysis, one could simply distinguish, (following Twardowski in 'Actions and Products'), between two kinds of products of a mental activity such as thinking. The first is a psychic product, the thought. These thoughts, and psychic products in general, never last. Far from being eternal *Gedanke an sich*, they ontologically depend on the thinker and they do not survive him or even his particular relevant thinking action. On the other hand, mental activity may be the (partial) cause of physical products, such as spoken or written sentences, which survive its non-lasting psychic products. In other terms, the only enduring products of mental activity are its material traces. Twardowski calls these traces psycho-physical products, because the people who accede to them may have a similar mental activity and produce similar psychic products for themselves. That is, roughly presented, the Twardowskian explanation of the way that 'thoughts' are subtracted from the transitoriness of one's mental life and potentially transmitted to others: by the mere survival of the material traces that often accompany and express psychic activities.

Returning to the question of author/reader asymmetry, it appears in this frame in another light. This asymmetry is essentially that of anteriority, in both a temporal and a causal sense: authors leave traces for their readers that can be interpreted by them to occasion similar mental episodes as those that governed the production of these marks. This 'naturalistic' account of the transmission of culture is, however, compatible with the normative

component that is currently, and rightly so, felt to be an essential part of humanistic scholarship. All readers are not on the same line: one in a position of being acquainted with the trace and its material details is the best equipped to understand the author. That does not forcefully mean that the author's contemporaries are the only ones able to do this. Contrary to the spoken discourse, which, in order to be understood, often requires hearing *in praesentia*, the written discourse, when suitably established and appropriately transcribed (metadata are crucial) can survive for a long time from the moment of its material production.

Far from being a renouncement of the traditional Humanities, DH may be viewed as scholarship put in the right perspective: neither platonic thoughts jointly targeted, nor abstract texts independent of their material realizations, but rather normed mental activity directed towards inscriptions and governed by them.

Notes

1. On the general scope of this extension of human capacities by computing machines, see Humphreys.
2. The facsimile of the manuscript, as well as the XML/ TEI transcription, are extracted from the repository of the Archives of the Łvov-Warsaw School www.elv-akt.net/.
3. As an example, compare the two following editions of the latin translation of Diophante's *Arithmétiques* by Claude-Gaspard Bachet, Sieur de Meizirac: the 'original' published in Paris in 1621, and the re-edition of 1671 in Toulouse, augmented by Fermat's notes. Or, to come to the nub of the issue, the specific significance of Fermat's copy of the original, with the marvellous annotation stating Fermat's "last theorem" as well as the impracticability of writing its demonstration within the margin of the copy.
4. On 'Electronic Literature'; namely, digitally- borne writing, not printed literature digitized afterwords, (cf. Hayles).

Bibliography

- Alvarado, Raphael, C. "The Digital Humanities Situation." The Transducer, WordPress Blog, 2011. Web.
- Benjamin, Walter. *Illuminations. Essays and Reflections*. New York: Schocken Books. 1969. Print.
- . "The Work of Art in the Age of Mechanical Reproduction," in Benjamin, *Illuminations. Essays and Reflections*, 217-251. 1936. Print.

- . "Unpacking My Library: A Talk about Book Collecting," in Benjamin, *Illuminations. Essays and Reflections*, 59-67. 1931. Print.
- Dubucs, Jacques and Wioletta Miskiewicz. "Logic, Act and Product," in Giuseppe Primiero and Shahid Rahman (eds.) *Acts of Knowledge: History, Philosophy and Logic: Essays Dedicated to Goran Sundholm*. London: College Publications, Oct. 2009. 73-94. Web.
- Fitzpatrick, Kathleen. "The Humanities, Done Digitally." Gold. Web.
- Gold, Matthew K. (ed.). *Debates in the Digital Humanities*. Minneapolis: Minnesota University Press, 2012. Web.
- Größes und vollständiges Universal-Lexicon aller Wissenschaften und Künste*. 64 vols. Leipzig: Johannes Heinrich Zedler, 1731-1754. Web.
- Hayles, N. Katherine. *Electronic Literature: New Horizons for the Literary*. Indianapolis: Notre-Dame University Press, 2008. Print.
- Humphreys, Paul. *Extending Ourselves. Computational Science, Empiricism, and Scientific Method*. Oxford University Press, 2004. Print.
- McCarthy, Willard. *Humanities Computing*. New York: Palgrave Macmillan, 2005. Print.
- McGann, Jerome. "Our Textual History." *Times Literary Supplement* 5564 (20 Nov. 2009): 13-15. Print.
- . *Radiant Textuality: Literature after the World Wide Web*. New York: Palgrave Macmillan, 2001. Print.
- McKenzie, Donald Francis. *Bibliography and the Sociology of Texts*. Port-Chester, NY, USA: Cambridge University Press, 1999. Print.
- Twardowski, Kasimierz. "On Actions, Products and Other Topics in Philosophy," in Johannes L. Brandl and Jan Wolenski (eds.) *Poznan Studies in the Philosophy of the Sciences and the Humanities*. Amsterdam and Atlanta: Rodopi. 1999. Print.
- . "Actions and Products: Comments on the Border Area of Psychology, Grammar and Logic," In Johannes L. Brandl and Jan Wolenski (eds.) *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 103-102. Amsterdam and Atlanta: Rodopi, 1999. Print. Rpt. of "O czynnościach I wytworach." Lvov. 1912.
- . "O czynnościach I wytworach." Polish original text, together with French and German translations by K. Twardowski www.elv-akt.net/ressources/archives.php?id_archive=1. Lvov. 1912. Print.
- . *Wybrane pisma filozoficzne*. Warsaw: Panstwowe Wydawnictwo Naukowe. 1965. Print.
- Wolfram, Stephan. *A New Kind of Science*. Champaign, Illinois: Wolfram Media. 2002. Print.

Looking Forwards, Not Back

Some Ideas on the Future of Electronic Publications

Gudrun Gersmann

Things have changed dramatically during the last decade. Ten to fifteen years ago the electronic world seemed to be reserved for geeks, nerds or scientific outsiders. In the Nineties many people had no idea of how quickly the 'new media' would change their personal and professional lives. I recall a friend, who – at one point in the 1990s – was astonished to hear that I wanted to buy a hard disk for my home computer. Why would you want to buy that, he asked me: "You will never write enough books in your life to fill a hard disk."

However, in the new media have, indeed, entered our daily lives. Using a PC to write texts, send e-mails, take, store and pass on photos, download bibliographical data or read e-journals is now a routine activity for many people in large parts of the world: In spring 2011, a German newspaper published an interesting article on 106-year-old Lilly Strugnell, who is willing to become the oldest Facebook user in the world – and, she is not only online working on her 'profile,' but also has plans to open a Twitter account. Now, Lilly wants an iPad 2 for her 107th birthday; she told the reporter: "Twitter sounds very interesting and I can watch Coronation Street online."

State of the Art

Even if we, defenders of the Humanities, sometimes long for the for the ivory tower, the debate on new forms and the future of academic publications is part of the fundamental process of social change, which is calling the old familiar models of literary production and reception into question and replacing them with new ones. We cannot deny the impact of these developments: when I travel on the underground in Paris, I see more and more people using their iPad, iPod touch or other reading devices to read books and newspapers. The triumphant advance of e-books appears to be unstoppable. This observation also applies to the academic book market.

Unfortunately, in my eyes, the academic system has, until now, not yet faced these challenges. Even if today – in contrast to the past – German

universities show a certain recognition of the need to familiarize students and junior researchers with specialized online services, there is still a long way to go, especially in the Humanities. Special priority programmes for example involving cooperation between representatives of various disciplines, ranging from history to literary studies and archaeology, are still a rare phenomenon in Germany at the moment; they are limited to a handful of universities such as Trier, Cologne, Würzburg and Frankfurt.

What is the attitude of scholars in the Humanities to electronic publications? Where is resistance still encountered? What expectations are there of an online publication? To find out what our fellow specialists expect we, at the German Historical Institute, prepared a questionnaire some time ago. The intention was to ascertain the interests and wishes of researchers in their roles as authors and publishers. The questionnaires were sent to the heads of various special Humanities research units and projects. While those that were returned did not supply any statistically utilizable results, they nonetheless provided an interesting snapshot of opinions.

The outcome of the survey can be summarized as follows: large numbers of historians no longer have any fundamental reservations about electronic publishing. On the contrary, most researchers – and especially junior researchers – are only too willing to have their work published electronically. In contrast to the situation with print publications, which often take years to produce in a publishing house and frequently only appear in limited editions, the researchers' hope is that electronic publications will give them faster and greater visibility in their field. The junior researchers surveyed repeatedly referred to the attractiveness of online publications, particularly with respect to application procedures. In the field of reviews, journals and conference proceedings, too, online publications are clearly now an established and integral part of specialist publications. Open Access (OA) extends the scope of their work by ensuring worldwide access and retrievability. Researchers also benefit from direct access to publications, improved search mechanisms across collections and ease of use. OA publishing offers a more effective and sustainable approach to the dissemination of scholarly knowledge. This means increased accessibility and dissemination leading to better research and greater benefits for society at large. Has the dream of Kofi Annan, who describes the future as the future of young people, "who want to shape the global world in a sustainable way and for the wellbeing of all," just become a reality?

Problems and Challenges

a) Peer Review / Impact Factor

Nevertheless, there are still many battlefields: for example, the respondents to our questionnaire frequently addressed the problem of quality standards. There were repeated demands for a peer review of electronic texts and for a guarantee of their long-term availability. However, online journals in the Humanities have one serious handicap compared to print journals. Since they do not yet have an impact factor, there are considerable worries that online publications count for nothing in application procedures. Here, again, we will have to wait and see how the debate on the furnishing of online journals with impact factors develops.

Certainly these are not the only worries and questions OA raises. There are also concerns about the growing flexibility and liquidity of information as a result of the possibility to copy, re-use and adapt. Only when we succeed in solving these issues in a satisfactory manner, will OA publishing allow for true development of Digital Humanities and social sciences.

b) Publishers and Authors

When we discuss the future of academic publications, we should certainly not forget the publishers. Although the role of the publisher in the publishing chain may remain relatively stable (choice, quality assurance and coordination of production), the structural position of publishing in the economic model may change dramatically. In the medium term, we believe that academic publishing should try to evolve into a service sector providing services to scholars, faculties and academic consortia, taking care of specific tasks and roles in knowledge creation and distribution. In this scenario, these activities could replace the present role of the investor and risk taker in the market for academic information, yet reaffirming one of the basic principles of the chain of scholarly publishing: the great importance of quality assurance and efficient dissemination/access as essential to good research.

In Germany – but also in other European countries – there has been a controversial discussion on open access in recent years. Subsequently, we can see some light at the end of the tunnel: even if some editors remain sceptical towards OA-concepts, there are a number of hopeful initiatives. Some European publishing houses, for example, recently joined forces in a new project called Open Access Publishing in European Networks (OAPEN),

which now consists of five European university presses (Amsterdam, Göttingen, Florence, Lyon, Manchester and Copenhagen) and two universities in the Netherlands, and is open to new partners. Jointly, the members have digital publishing programmes, conduct OA experiments, make use of digital repositories, publish in different European languages, have a worldwide distribution network (including the US) and cooperate closely with university libraries. At the request of the European Commission (EC), these partners have developed a joint proposal in the category of 'Targeted Projects' in the EC's eContentplus programme.

One of the aims of OAPEN is to develop and implement an OA publication model for academic books. The project also aims to achieve a sustainable approach to improve quantity, visibility and usability of high-quality OA content. It aims to foster the creation of new content by developing future-oriented publishing solutions. In order to expand the content of the online library and achieve critical mass, OAPEN is also aggregating content from other scholarly institutions that are interested in publishing high-quality monographs and are open-minded towards OA. It also welcomes institutes with a publishing aspiration or scholars in charge of a series.

c) Active Engagement in E-publishing

On the one hand, online publications are considered to be a good and economical alternative to print publications. On the other hand, most fellow specialists show no inclination to become active on the internet, for instance as bloggers or critical commentators. In my view, we are still a long way away from a concept of 'networked Humanities.' For the most part, the use of internet publications is still limited to what one might call passive consumption in private. A playfully active approach to texts on the internet continues to be viewed with mistrust. There is obviously still a great fear of being looked at sceptically by other experts as a result of quick, pointed and sometimes perhaps even 'flippant' commentaries posted online.

Taking all this into account, it comes as no surprise that online Humanities journals are still little more than print journals transposed to an electronic medium. Though the benefits of an internet publication, such as the possibility of rapid global availability, are exploited, no use is made at all of the potential offered by online writing or, indeed, of collaborative writing.

To give you an illustration of what I mean, let me take *Francia*, the journal issued by our institute in Paris, as an example.

Francia was founded in 1973 and has undergone a thorough overhaul over the past three years. In the process, *Francia* has been transformed from a



Fig. 1: Homepage of Francia Online (<http://www.perspectiva.net/content/publikationen/francia>).

purely print journal into a hybrid journal with both an online and print section. While essays continue to be published in a print edition, we have shifted all the reviews to the online edition. Book reviews are published four times a year online in open access.

The realignment of *Francia* has gone hand in hand with a retro-digitization project undertaken jointly with the Thorbecke-Verlag publishing house and the Bavarian State Library in Munich. Since November 2008, all the previous volumes of *Francia*, for the years 1973 to 2006, have been made available online free of charge for everybody. With the help of a two-year ‘moving wall,’ all future editions will be successively digitized and made available to researchers.

While the institute in Paris wanted to give the journal greater ‘visibility,’ that was not the only reason for the rigorous restructuring of the familiar publication channels. High up on our list of considerations were the Francophone researchers in countries with a poor library infrastructure, who are glad to be able to download specialist texts free of charge from the internet. Our initiative met with a broad positive response. Nearly all of the approximately 1,000 *Francia* authors whom we asked for permission to reprint their articles before publishing them online welcomed and supported the initiative. Only one reviewer – a 90-year-old gentleman – politely requested to be spared any requests for reviews in the future. At his age, he no longer saw any prospect of him learning to use a computer!



Fig. 2: Homepage of Perspectivia.net (<http://www.perspectivia.net/>).

Given the positive responses we have received, we are planning to join forces once again with the Centre for Electronic Publishing (ZEP) of the Bavarian State Library in the years ahead to carry out the retro-digitization of other series of publications issued by the German Historical Institute, such as the *Beihefte der Francia*, the *Instrumenta* and *Pariser historische Studien*. They will subsequently be published on www.perspectivia.net, the institutional repository of the institutes grouped together in the Foundation of German Humanities Institutes Abroad.

Though *Francia* has been successfully restructured and is regarded benevolently by the journal's 'regular customers,' there is no escaping the fact that, if we take a critical view of things, – like the publishers of other online specialist journals – we have not even begun to tap into the potential offered by electronic presentation and publication. There is no question yet, in this context, of collaborative writing processes. The process of publishing the *Francia* reviews is much too conservative for that. The content management system we use for reviews is not accessed directly by the authors, but via a member of staff at the institute, who is simultaneously responsible for the correct handling of the assessment process. While we permit readers' commentaries, we do not publish them straightaway, but

wait until an editorial check has been carried out. This helps us to ensure the quality of *Francia* as a publication, but it does not liven up the texts! Under such conditions a discussion between the reviewer/author and the reader is largely ruled out.

Is there an alternative, a counter-model to 'traditional' journals like *Francia*? I would say: yes, there is. In my opinion – and maybe this will be the future of e-publications – we are currently at a crossroads. Day by day we are confronted with users who, as digital natives, come from a culture whose laws and codes of conduct differ from those in the classical world of the Humanities. The best example of the incipient changes in our use of the media is the success of Wikipedia. This internet encyclopaedia has come to symbolize a dynamic, collective knowledge store compiled on a voluntary basis with the help of self-organized systems. At the same time, Wikipedia provides an exemplary illustration of collective, networked knowledge generation based on the principle of 'swarm intelligence' or the 'wisdom of many.' In the case of Wikipedia, collective writing is not just invoked in theory but also put to the test in practice. In contrast to classical specialist journals, in which experts raised their voices and acted as opinion-leaders, the people who write for Wikipedia no longer attach any importance to their individual authorship. This is not least reflected in their renunciation of any mention of their own names. Most of the authors use made-up names, although that should not be construed as a lack of seriousness. The team of observers ensures that these cover-up tactics do not lead to any misuse of texts: a Wikipedia article does not need any prior peer review procedure. The Wikipedia community ensures that errors do not remain on the internet for long. Every Wikipedia author is obliged from the very beginning to accept a kind of voluntary self-control. Among the most important tools used by Wikipedians are special lists, which enable a close watch to be kept on particularly interesting or vandalism-prone articles – anything to do with Nazi history, for example. In addition, changes made by newcomers are not visible to the reader immediately, but only after they have been looked at by an experienced author. Studies made of the quality of Wikipedia articles have given the Wikipedia project fairly good marks. Though – or perhaps precisely because – there is a lack of any traditional quality assurance mechanisms and though experts, who are normally as much in demand in the Humanities as elsewhere, play no role whatsoever, the articles that appear are by no means inferior to those in comparable encyclopaedias produced in accordance with traditional criteria. The principle of 'control by a group,' which ideally is well networked, carefully observes the installation of every single text on Wikipedia and responds immediately if problems are spotted, seems to work quite well.

recensio.net
 Rezensionen für die europäische Geschichtswissenschaft

Rezeptionen, Präsentationen, Themenrevue, AUFKAB, UMS UMS, RITZACHEN, FRAGEN

rezensieren

Neueste Rezensionsteile online

- Reviews in History, 2011, September
- Bohemia, Zeitschrift für Geschichte und Kultur der böhmischen Länder 51 (2011), 1
- Jahrbücher für Geschichte Osteuropas / Jgo.-revue, JGO 59 (2011), H. 3

Neueste Rezensionen online

- Dutch Culture in the Golden Age
- Filozofická fakulta Masarykovy university
- Popular Culture and the Public Sphere in the Rhineland 1800-1850

Neueste Präsentationen von Monographien und Aufsätzen

- Das KZ Lublin-Majdanek und die Justiz
- Die Koggenbüchsen in der Forschung
- Europa der Bürger?

Neueste Präsentationen von Internetressourcen

- e-Spania. Revue interdisciplinaire d'études hispaniques médiévales
- Diacronia. Studi di storia contemporanea
- EGO | Europäische Geschichte Online

Teilnehmende Zeitschriften

Mélanges de la Casa de Velázquez, Nouvelle série | Archiv für Sozialgeschichte | Recherches sur l'islam et sur l'Éthiopie | Germano-Polenica, E-Newsletter der Kommission für die Geschichte der Deutschen in Polen e.V. | Reviews in History | Frühaugen-Info | Revue d'histoire du XIXe siècle | Jahrbücher für Geschichte Osteuropas / Jgo.-revue | German Historical Institute London Bulletin | Südosteuropa, Zeitschrift für Politik und Gesellschaft | Lietuvos Istorijos Metraštis | Südost-Forschungen, Internationale Zeitschrift für Geschichte, Kultur und Landeskunde Südosteuropas | Besuche und Forschungen, Jahrbuch des Bundesinstituts für Kultur und Geschichte der Deutschen im östlichen Europa | Zeitschrift für Germanologie/Forschung (ZfG) | Quellen und Forschungen aus Italienischen Archiven und Bibliotheken | byzantinische Forschungen | Bohemia, Zeitschrift für Geschichte und Kultur der böhmischen Länder | Lithuanian Historical Studies | Francia-Parsavia | Zeitschrift für Balkanologie | Ungarn-Jahrbuch, Zeitschrift für interdisciplinäre Hungarologie | Jahrbuch zur Liberalismus-Forschung | Recensio Moskva | Il Mattino di Storie

DFG recensio.net ist ein von der Deutschen Forschungsgemeinschaft geförderter Gemeinschaftsprojekt folgender Institutionen:

BSB Bayerische Staatsbibliothek, DHP HA, INSTITUT FÜR EUROPÄISCHE GESCHICHTE

Fig. 3: Homepage of Recensio.net (<http://www.recensio.net/front-page>).

Prospects for the Future

What might new forms of collaborative writing look like? What publications are conceivable under the general heading of 'networked humanities'? In Munich, we are currently in the process of developing a major project funded by the German Research Foundation under the heading recensio.net (www.recensio.net), which is moving in exactly this direction.

On the one hand, recensio.net has a classical component in that it collects, as it were, the reviews published in regional and national historical journals and makes them available online.

From our point of view, however, the more important and innovative aspect of recensio.net is what you might call a 'Web 2.0' idea. In contrast to *Francia*, the intention is to actively incorporate authors and readers into the publication and commentary process. We very much hope to be able to offer 'living texts.' Authors of dissertations, for example, can 'self-announce' their books on recensio.net. The idea is not to provide a platform for self-advertising, but rather space for a brief, form-like presentation of the core propositions. The plan is that readers should, firstly, comment on or criticize these presentations in brief or at length. In a discussion forum of this kind we envisage the author as not being merely an 'object,' but also as being in a position to provide feedback himself on the remarks made by the readers, so that – ideally – a genuine debate can unfold on propositions and findings.

Recensio.net is an experiment. We are well aware of the many reservations that still exist concerning interactive and collaborative concepts. However, we are convinced that this method of ›reviewing‹, which has long become well established outside the purely academic book market, harbours potential for the reasons I have just given.

An additional 'value-added' aspect will be that authors are able to present not only monographs on the platform, but also essays they have published in anthologies or journals. This is an attempt to help counteract the typical 'anthology review,' which has always been problematic in that it rarely deals with issues at the individual essay level, though academic debates often take place precisely at the anthology, i.e. essay, level.

To sum up I would like to formulate two propositions. What I have said so far is certainly ambivalent. A look at e-journals and portals in the Humanities shows, on the one hand, how strong the presence of the 'old world' of print publications still is in terms of implementation and use even in the 'new electronic world.' On the other hand, the success story of Wikipedia and other similarly designed projects shows that we are still deeply involved in a process of change. The interest that new Web 2.0 initiatives like recensio.net have met with – particularly on the part of junior researchers – could be a sign that the generation of digital natives is no longer content to passively consume internet publications, which is still largely the case at present, but is rather prepared to embrace the opportunities and the risks offered by collaborative working.

In the long term, I feel, this will result in more than a coexistence of two different sets of writing and reception habits on the internet. In my

opinion, the idea and practice of collaborative writing will have a considerable impact on the understanding of academic research that we have had hitherto. In contrast to earlier times, the production, distribution and reception of specialized knowledge will, firstly, no longer be understood solely as the task of recognized experts and prominent individual scholars or publishing houses who can claim authority in their discipline by dint of their academic merits. To put it bluntly, the traditional system of a handful of established, authoritative specialist journals, in which the 'VIPs' in the field presented their articles and new editions were eagerly awaited, has had its day. The reason for this is that the judgement of experts has now been joined – indeed, if not replaced – by a democratic process of opinion forming that will gradually claim ever more space and put its stamp on mutual communication. The consequences of this process of transformation for the self-image of the Humanities will require investigation in greater depth, but it is my firm conviction that we are talking here about a genuine revolution. The hierarchically structured 'vertical communication' of the past will be replaced by a 'horizontal, fluid communication,' in which academic laymen and experts stand side by side on an equal footing. The logic behind Wikipedia is very persuasive. Why should only a 60-year-old full professor of history be allowed to write articles for an encyclopaedia? Is not a local historian, who has taken a keen interest in the history of his village over many years, better qualified to publish articles about it? In Wikipedia, value is attached to the academic, who writes excellent articles on matters to do with his specialist field, as well as to the enthusiastic amateur, who shows a painstaking love of detail in writing about areas for which no academic training is required. No greater importance is attached to the texts of the one than to the texts of the other.

Recent years have shown that, after a hesitant beginning, more and more debates and articles are being published in mailing lists and blogs – and no longer in the specialist journals. From my point of view, a mailing list like *H Soz Kult* now has a greater degree of interpretative power in the field of history than any monograph or specialist journal. Let us be curious.

The Dynamics of Digital Publications

An Exploration of Digital Lexicography

Claudine Moulin and Julianne Nyhan

Introduction or ‘Why Should Humanists Care about Digital Publications?’

“I have not established my name stamped on bricks as my destiny decreed”
remarked Gilgamesch to Enkidu
(The Epic of Gilgamesh 72).

Since the earliest times, memory, learning, imagination and the technologies (in the sense of *technê*) that aid their recording, communication and study have been deeply interwoven. The roles played by technologies in reflecting, shaping and informing our interpretation of the world have been highly complex. Think of the technology of the book, for example, and the political, social and intellectual significance of the transition from the scroll to the codex both in and for the Christian tradition. Nevertheless, it is important to state that technologies are not the only or even driving force behind such changes and scholars, such as Winner (1999), have made clear the inherent limitations and myopic nature of the technological deterministic model. Indeed, De Smedt has brought out the complex interplay between human intellect, culture and technology in producing new knowledge:

The telescope was invented in 1608 and was initially thought useful in war. Galileo obtained one, improved it a little, and used it to challenge existing ideas about the Solar System. Although a magnificent new technology in itself, the telescope was hardly a scientific tool until Galileo used it to create new knowledge (De Smedt 2002, 99).

While the limitations of the theory of technological determinism are clear, the refrain that ‘the computer is just a tool’ is equally problematic, and for a number of reasons. Prominent among them is the implicit categorization of a tool as a neutral agent that is used to affect a more important end. In the Humanities, process is held to be as important as outcome, if not more so, and the aim is rarely to solve problems, once and for all, but rather to

rediscover and reinterpret them in order to ask new and better questions. Thus, such a statement implies that there can be no role for the computer in the intellectual work of the Humanities. It has deeper cultural resonances too:

Just because the separation between thinking and making is long-standing and well-entrenched doesn't make it a good idea. At various times in the past, humanists have been deeply involved in making stuff: Archimedes, the Banu Musa brothers, da Vinci, Vaucanson, the Lunar Men, Bauhaus, W. Grey Walter, Gordon Mumma. The list could easily be multiplied into every time and place [...] (Turkel 2008, par. 5).

Perhaps most problematic about the statement is the way it can be used to proscribe the boundaries of our *Wissensräume*, i.e. knowledge spaces, whose description, transfer and even production are profoundly based on symbolic systems (Ash 2000, 239). Describing the computer as 'just a tool' implies that the spaces that it is conceived in, developed in and even operate in need not concern the Humanities. Indeed, it implies that the nature of a given tool is conceptually and philosophically independent from the society that develops it and uses it, as well as from the processes it is used in and the results it helps to create.

Early commentators on the effects of computing technology on the Humanities, and society more generally, often discussed such changes in highly emotive ways, whether it was the self-termed "neo-luddite" claims of Himmelfarb (1996), the profound elegy on the death of reading by Birkerts (2006) or Perelman (1993) who hailed computing technology to be the greatest invention since fire. Humanists in particular are able to draw on their understanding of history to contextualize such statements and know that all of these scenarios are highly unlikely to exist outside of the world of the page that they are captured in. Notwithstanding such hype, since the 1990s, in particular, which saw the arrival and exponential uptake of the World Wide Web and the making available of the Mosaic browser, the application of computing technologies to Humanities scholarship, and the primary and secondary sources that it concerns itself with, has been ushering in many changes.¹ Scholars from across the Humanities have open to them new ways to access, search, interconnect and visualize primary and secondary sources; to disseminate and access research findings, as well as to communicate and collaborate with one another. Many effects of such changes not only remain undiagnosed but, furthermore, are little understood. We find ourselves in the most nascent stages of identifying, analyzing, theorizing and responding to the profound changes that the application of computing

to our cultural heritage and learning is bringing about. For example, how are the ways we communicate, think, read and write altering? How are the information-gathering patterns of Humanities scholars changing? What of the issues of digital preservation and obsolescence? And this is to say nothing of the changes that are resulting from research that attempts to involve the computer as partner rather than slave. As McCarty (2010) observed in his inaugural lecture, "I celebrate computing as one of our most potent speculative instruments, for its enabling of competent hands to force us all to rethink what we trusted that we knew." Despite the embryonic stage we are at we need have no doubts about the significance of such changes for the Humanities: "As with the renaissance sped forward by the printing revolution of the fifteenth century, digital technology is driving a radical shift in humanities scholarship and education. The depth and character of the change can be measured by one simple but profound fact: the entirety of our cultural inheritance will have to be reorganized and re-edited within a digital horizon" (McGann 2010, 'Introduction'). In March 2010, a three day conference entitled *Online Humanities Scholarship: The Shape of Things to Come* took place at the University of Virginia. The conference was aimed at the Humanities in general and sought to explore how online Humanities research may be developed and sustained. In the introduction to the subsequent proceedings McGann wrote:

But as we all know, online scholarship is still practised by only a tiny fraction of our humanities faculties. The absence of a broad professional involvement has been long-lamented and variously explained: steep learning curve, entrenched habits, lack of available time and resources, wariness at the volatile character of the new technologies. And all of those explanations are pertinent. But equally pertinent is the general failure of scholars who use digital media to give clear explanations of its critical research value (McGann 2010, 'Introduction').

In this paper,² we will present two case studies of research that has been carried out on the digital remediation of German and Luxemburgish dialectal, regional and historical lexicography. In doing so, we will take some initial steps towards setting out what we hold the critical value of this research to be. The focus of the following reflections is not mere image digitization; rather, we aim to reflect on a host of new research findings that can be created via the scholarly evaluation, interpretation, semantic annotation and subsequent analysis of research material and data using computational methods. To close we will briefly address the "absence of broad profes-

sional involvement” noted by McGann. By exploring how the German and Luxemburgish lexicographical material described here might both benefit from and enrich a global information space, such as the emerging Semantic Web, we will argue that Humanists need to engage with such developments so that their unique knowledge and learning may inform them.

Creating and Identifying New Lexicographical Constellations: The Case of German and Luxemburgish

The European vernacular languages are characterized by a linguistic tradition in which dictionaries generally belong to the first forms of documentation and codification of a language. The dictionary corpus of a language (here, we shall begin by focusing on German) includes old lexicographical work in medieval manuscripts and old prints. From the nineteenth century onwards, dictionaries begin to document in a more modern sense standardized language forms, dialectal variants or historical stages of a language (Moulin 2010). Today, the manifold developments in the field of the Digital Humanities and e-philology have enabled new settings for the creation of lexicographical resources, ranging from the use of computer-based methods in the conception of dictionaries to the design of born digital lexicographical projects. Furthermore, the possibility of linking completed as well as ongoing lexicographical projects so as to form a complex and ‘intelligent’ multidimensional network of dictionaries becomes manifest. Though challenges are multiple, a project on German dictionaries at the University of Trier³ has been playing out this scenario by building up in the past ten years a network of different sorts of lexicographical data on the German language. One of the first steps in creating such a lexicographical network of the German language was built by the dictionaries of the (south) western German language area (Fournier 2003; Moulin 2010; Hildenbrandt and Moulin 2012). The dialect landscapes of the Rhineland, the Palatinate and the Euroregion of Saarland, Lorraine and Luxembourg are covered by completed broad-scale dictionaries, originally published in printed form and mainly based on the lexicographical tradition of dialect dictionaries as they have been developed since the middle of the nineteenth century.

The layout of these printed dictionaries varies from project to project. As the lemmatization and hierarchical order of the headwords have different realizations in the print dictionaries, the lexical matching of the different linguistic systems of these conjoining regions can only be examined and compared when using digital versions with appropriated encodings and

annotation standards. Such a system then enables complex enquiries, such as a full-text search through all the underlying materials or specialized search for specific detailed information in the dictionaries enclosed in the system. Explicit links from the dictionaries themselves to other dictionaries and sources can be rendered as hyperlinks, and this is also true of the indirect cross references, which can only become apparent as a result of various statistical algorithmical and e-philological algorithms (Burch and Fournier 2004; Burch and Rapp 2007). Meanwhile, the digital edition of the monumental Grimm brothers' 'German Dictionary' (www.dwb.uni-trier.de), the network of Middle High German dictionaries (www.mwv.uni-trier.de) and some others like the Goethe Dictionary (<http://gwb.uni-trier.de/de/>) have also been completed and can also be linked with the dialectal lexicographical material. In essence, a substantial lexicographical information system has been created (www.woerterbuchnetz.de) that allows cross-referencing and interlinking between the dictionaries. As a result, for example, dialect and historical vocabulary can be accessed and analyzed via standard language metaforms.

The digital network of dictionaries 'woerterbuchnetz.de' was developed in several stages, starting from the available printed versions of extant dictionaries, enabling the continuous growth of the system also with digitally-borne material and ensuring the long-term, cross-platform availability of the data produced (Hildenbrandt and Moulin 2012). A machine-readable full-text version of each dictionary in the network was created. The geographical maps and diverse illustrations, charts and images contained in the material were also digitized and linked to their respective headwords. In the machine-readable version, the individual elements of each article were labelled in SGML/XML using TUSTEP scripts applied to complex layout information (e.g. headwords, grammatical information, phonetic transcriptions, semantic information, citations and quotes, geographical information concerning survey data, etc.). The markup is standardized and conforms to the internationally accepted guidelines of the Text Encoding Initiative (TEI). Dynamically generated and classified network information can be accessed in the online version: both the lemma on which the cross reference is based and the reference target are displayed. Interlinking the dictionaries, as in the transformation of existing cross references between the dictionaries to hyperlinks and the algorithmical generation of new semantic cross references at a virtual level, are of a high lexicographic potential, as they enable the idea of "rendering semasiological dictionaries onomasiologically useful" (Reichmann 1986: 176, our translation; in the original "*semasiologische Wörterbücher onomasiologisch nutzbar zu machen*").

Such a dictionary network can only be set up with the aid of algorithmic methods drawn from the field of information retrieval and grid-technology. In comparison with traditional, printed monographical publications, the assets of the digital linkage of the lexicographical network lie in the fact that principally *all* existing lexicographic sources and materials (such as underlying sources, texts, geographical maps, images, etc.) can be included and also be augmented with further components, such as other dictionaries, primary texts including the *belles lettres* literature or new maps.

Moreover, the network of dictionaries demonstrates how such digital works can become the central pillars of additional tasks such as text analysis. In the case of the Virtual Research Environment TextGrid (www.textgrid.de), the network is reconceptualized as a Webservice, thus enabling other software applications to act on the data contained in it. Via TextGridLab, which provides a single point of entry to the Virtual Research Environment itself, it is possible to implement an integrated search on all dictionaries, whether from the standalone Dictionary Search Tool or by double clicking on the required word from within the TextGrid XML Editor.

Such additional steps are, for example, undertaken in the *LexicoLux* project at the University of Luxembourg, which addresses the documentation of the Luxemburgish language, the youngest extant Germanic language.⁴ Its lexicographical tradition goes back to the nineteenth century and arises thus at almost the same time as Luxemburgish emerges as an independent language (Gilles and Moulin 2003). The first lexicographical source was published in 1847, the second in 1906. The most extensive lexicographic documentation to date was published in five volumes between 1950 and 1977 (*LWB*). In creating a dynamic and multi-directional network, the *LexicoLux* project not only documents the Luxemburgish language and its metalinguistic reflection from its beginnings in the nineteenth century, but the system also provides an extendable basis for further linguistic analysis and lexicographical documentation of the modern forms of the language itself. The headwords of different dictionaries can be browsed using a standardized meta-lemma list of Luxemburgish, which is necessary as the different dictionaries also use three distinct orthographic systems documenting different stages of orthographical reforms. Moreover, the linguistic components of the data are contextualized by the linking of the Luxemburgish lexicographical data with that of the neighbouring dialect varieties in the Grande Région; namely, by integrating the Luxemburgish data in the Trier network of regional dictionaries included in the woerterbuchnetz.de. A linking to similar digital dictionaries of the French language (e.g. www.atilf.fr) can also be considered, documenting

the roman elements of the Luxemburgish language. Political and national borders can thus be surpassed and common linguistic and lexicographical structures, phraseological parallels and cultural-historical features better analyzed. The principle of dynamical interoperability also extends to external digital resources (including audio and video). For example, the survey locations in the dictionaries can be linked in to the Digital Linguistic Atlas of Luxemburgish (www.luxsa.info) and its sources.

The fundamental conditions and necessary prerequisites for such complex information systems as the lexicographical networks presented here are the availability of open access data, the high quality in-depth annotation of this data using international standards in the digitization of text, maps, images and artefacts, and the interoperability of the systems created, plus a genuine willingness to engage in interdisciplinary collaboration within the academic community in the Digital Humanities.

Digital Lexicography in the Global Information Space

An indication of the huge amount of information that is available on the Web can be gained from Google's announcement (Alpert and Nissan 2010) that in 2008 the number of unique URLs that it processed had reached 1 trillion (and this is to say nothing of the 'Deep Web' that remains inaccessible to most search engines). At present, keyword searches implemented via a search engine and its underlying algorithms are the most common way of retrieving information likely to be relevant to a query. However, as set out by Antoniou and Van Harmelen (2008), there are significant, commonly encountered problems associated with search engines. These include the case when a search engine may return pages relevant to the key words given, but additionally include many thousands of pages of moderate to low relevance; where the requested pages are not found, for example because results are highly sensitive to vocabulary; or where results are spread over numerous pages and so a number of queries must be carried out (1-3). Furthermore, once a number of hits have been obtained, the steps involved in filtering such information by, for example, selecting documents of particular relevance, tracing recurring themes in documentation, or making links between documents, can be a laborious task. The result is that it is not currently possible to use a search engine to answer either a highly complex or highly precise question. To take a rather simple example, it is not possible to retrieve websites that refer to the German personal name *Achim*, rather than to the place name *Achim* without considerable

human intervention both in the selection and combination of keywords and subsequent filtering and evaluation of results. To move from there to a more complex query, such as 'all encyclopaedias printed in Paris during the eighteenth century (whether in French, German or English, etc.) that mention the place name *Achim* in the body of an article' is practically impossible. Relevant to the problem too is the quasi-'embarrassment of riches' that we are currently facing. Ever increasing amounts of data are becoming available every day (through, for example, mass digitization initiatives, such as, inter alia, Europeana, Haithi Trust and the Internet Archive); however, much of it is not machine accessible. As a possible solution to such issues, in what is now regarded as a seminal paper, Tim Berners Lee et al. (2001) set out their vision for the Semantic Web.

Spearheaded by the World Wide Web Consortium (W3C), the Semantic Web research community has been working towards the realization of a Semantic Web that will not replace the current day web, but rather extend it. A central aspect of this research involves defining and making explicit the relationships that exist between data on the web. This can be done by describing such information in a formal way and with the complete consistency and explicitness that the computer requires. The most complex and fine-grained approach to expressing such definitions and relationships is the process of developing ontologies.

A term often used in philosophy, in the computational context referred to here, ontology is usually defined in line with Gruber (1993). The term is used to describe formal, explicit and re-usable conceptualizations of a specific domain, such as an ontology of linguistic terms. An ontology describes not only the concepts of a domain but their properties and relationships too. An important aspect of Semantic Web research, they have been developed by many disciplines across the Sciences for some time in order to enable the representation and interchange of the semantic knowledge on a large scale. In the area of linguistics and lexicography, perhaps the most well-known example is WordNet. An argument for the benefits of extending XML encoding with ontologies in order to model natural languages has been convincingly made by Canfield (2009) on the basis of the highly inflected Navajo language of North America. He concluded that this approach enables enhanced search of the lexicon, interoperability and, in turn, can be shared with the emerging Semantic Web. The research being carried out in the project '*Wechselwirkungen zwischen linguistischen und bioinformatischen Verfahren, Methoden und Algorithmen,*' a collaborative project between the University of Trier, the *Institut für Deutsche Sprache* (Mannheim) and the Julius-Maximilians-Universität Würzburg, in which linguists, computer

scientists and bioinformatic scientists seek similarities between the human genome and language change, indicates the kind of highly innovative multidisciplinary work that such approaches can enable.⁵

Since 2009, another research thread in the Semantic Web community has been gaining ground, that of Linked Data, which is a set of “best practices for publishing and connecting structured data on the web” (Bizer 2010 p. 1; for best practices see Tim Berners-Lee 2006). Linked data is about assigning a unique name (a URI) to a piece of data and making it accessible with HTTP, thus making dereferencable URIs in order to lay important prerequisites for making information recoverable, capable of being shared and re-purposed on demand. When a URI is dereferenced, useful information should be returned, in RDF/XML. RDF uses XML syntax and consists of *resources*, *properties* and *statements*. Each RDF file contains a number of RDF Triples, which specify a subject, predicate and object. For example, *TextGrid* (**subject**) is a *collaborative project* (**predicate**) of the *University of Trier* (**object**). Where HTML is used in the traditional web to create links between documents, linked data establishes typed links between things in the web of data. In turn, this information can be linked to from other datasets, so once data has been prepared and exposed according to linked data principles it can both draw on and be used by other disparate and distributed pieces of data or collections of data in the “global data space” (Bizer 2010 p.1) in a fully automated way. As more linked data becomes exposed the store dynamically becomes more comprehensive. This is one of the key benefits of linked data over other approaches to efforts to combine data from disparate sources such as Mashups, microformats, etc. “In summary, linked data is about using the Web to create typed links between data from different sources. The result, which we will refer to as the Web of Data, may be more accurately be described as a web of things in the world, described by data on the web” (Bizer 2010 p.2).

From the perspective of the Humanities, it is not difficult to see the possible benefits of establishing connections between linked data sets and much finer-grained and complex domain-specific ontologies. To look again at our encyclopaedia example, it should be possible to use a search engine to retrieve a list of the dictionaries that were printed in Paris during the eighteenth century (whether in French, German or English, etc.) that mention *Achim* (the place name and not the personal name) in the body of an article. If ontological information were available that describes, for example, the domain vocabulary of geographical place names, it should furthermore be possible to use logic-based languages capable of ‘reasoning’ and automatically identifying new connections in the datasets at hand.

However, it is not clear that this is how the Semantic Web will develop. Most interestingly, Hausenblas (2009) has observed that linked data focuses mostly on RDF, linking and HTTP, rather than ontologies and inferencing and is thus a simplified approach to the Semantic Web. “This simplification – just as the Web simplified the established academic approaches of Hypertext systems – lowers the entry barrier for data provider, hence fosters a wide-spread adoption” (‘Linked Data Principles’).

For Man and Machine: The Changing Role of the Digital Dictionary

But what of the role of dictionaries, whether as individual works or realized as the new lexicographical constellations (described above) in this emerging information space? As information retrieval and querying technologies – regardless of whether they are based on the approaches of the Semantic web or Linked Data communities described above, or other approaches such as data mining etc. – become more sophisticated, will the dictionary become obsolete? In short, might changing publication practices in the Humanities bring about the decline of the dictionary?

Dictionaries have existed for more than 4000 years; indeed, no literate society has existed that has not created lexicographical works (Hausmann 1989). At the most generic level, dictionaries are of immense and continuing importance to scholars of all disciplines and levels, as well as to lay people, as reference works. For the lay person they are a trusted repository of authoritative knowledge; for the scholar, dictionaries are a crucial resource, particularly in the Humanities, where scholars must often navigate multiple languages. Throughout history, dictionaries have allowed segments of cultural heritage to be identified and navigated; for example, by providing access or pointers to sources crucial to the word or idea under discussion. It is notable too that dictionaries have long been consulted, not only to locate information or references, but to reassemble information and hypothesize about it too (see, for example, Ó Corráin, Donnchadh (2001)). Dictionaries are also far more than simple word books or mere repositories of linguistic information. They are as much witnesses to and repositories of cultural and societal information as they are sources and objects of learning and also knowledge spaces (see, *inter multa alia*, A. Reddick 2010 and Hausmann *op cit.*). Dictionary research is increasingly identifying the many ways that dictionaries are implicitly and explicitly interlinked with one another and the languages and societies that they sought to reflect and sometimes influ-

ence. In this way, even dictionaries that are by today's standards insufficient or, for example, reflect a questionable zeitgeist, can in and of themselves remain important historical witnesses (see, for example, Rapp 2011).

We hold that the changing publication cultures of the Humanities have the potential to reinforce and deepen this central role of the dictionary while also creating completely new roles for lexicographical information. In relation to this latter point, that dictionaries are a sine qua non of research infrastructures has been convincingly argued by McArthur. Drawing on Kronick's theory of "information media" as consisting of two categories, primary (e.g. textbooks, periodicals, scholarly journals, etc.) and "secondary" (e.g. indexing systems, abstracting systems, etc.) he writes:

[...] it is clear from [Kronick's] categorization that any comprehensive secondary system must also contain works of lexicographical and general reference. Indeed, dictionaries and encyclopaedias are classic examples of how we abstract information from primary sources of various kinds and marshal that information in terms of some kind of indexing and pointing system (McArthur 1988, p.7).

Indeed, over the past years we have seen dictionaries gain a new audience: machines, computational systems and networks. Just as humans rely on dictionaries to understand, access and navigate cultural heritage so, too, do machines. Presently, dictionaries play a crucial role in many kinds of computational systems; for example, multilingual data processing and interchange (see, inter alia, Dietrich 2010); machine translation (see, inter alia, Aljlayl et al. 2011); and information retrieval (see, inter alia, Gotscharek et al. 2009). While sophisticated search engines deliver an abundance of information, one of the greatest on-going challenges of modern day information technology is to make this abundance manageable. Accordingly, even the most sophisticated search engines will not, for the foreseeable future, offer viable or superior alternatives to authoritative scholarly dictionaries and thoroughly machine-readable dictionaries (such as those described above) play a multifaceted role in strategies developed to manage this abundance. A concrete example of the role of dictionaries in such systems is provided by Smith et al. (2001) in their discussion of the process they have used for the disambiguation of geographical names in the Perseus Digital Library. They note that:

Before either identification or disambiguation could proceed, we gathered the knowledge sources used to make the categorization and disambiguation decisions. Perseus uses some knowledge sources, such as the Getty

Thesaurus of Geographic Names or Cruchley's gazetteer of London, that were purpose-built for geocoding. We captured other information, such as lists of authors or the entries in the *Dictionary of National Biography*, as a by-product of constructing the digital library as a whole. In total, the gazetteer used for name identification and disambiguation contains over one million place names (Smith and Crane, 2001).

So, it is clear that the changing publication cultures in the Humanities open new possibilities for how dictionaries may be used now and in the future by both humans and machines. However, this is, as yet, not fully realized. It is interesting to note that many of the problems that afflict hard-copy dictionaries can also be noticed in their digital surrogates. Since the early days of the World Wide Web, a critical mass of dictionaries of all kinds and formats has become available online, ranging from the scanned images of dictionaries made available by mass digitization initiatives, such as the Internet Archive (<http://www.archive.org/>), to the scholarly, retrodigitized and enhanced editions of historical dictionaries researched by Digital Humanities projects; for example, *Das Wörterbuchnetz* (www.woerterbuchnetz.de/). Yet, the varying use of standards, technologies and publishing strategies (e.g. server side transformation to HTML with underlying XML not being made public) used by these works mean that discovery, search, reuse and systematic research of individual digital dictionaries is often very difficult, and certainly impossible at the aggregate level. This is not the case with *Das Wörterbuchnetz*; however, such projects are still at the cutting edge of what has been realized in the domain of digital lexicography. As a result, much of the potential of digital dictionaries also remains fragmented and under-exploited by both humans and machines. We hold that dictionaries, viewed from their totality, rather than as individual work, represent a currently fragmented infrastructure and corpus. One possible outcome of the semantic and structural interweaving of successive dictionaries within a particular language will be the emergence of a dense matrix of cultural, linguistic and bibliographical information that is accessible to both Human and machine as never before.

Conclusion

In conclusion, we argue that new digital publication opportunities are not a substitute for the print publication of a book, but create wholly new publication forms that simultaneously transcend and replenish traditional book

publication. The formation of new publication cultures is a most dynamic and multi-layered process that requires a careful balance to be struck between innovations and tradition and calls for new kinds of methodological and theoretical skills to be brought into play, so that knowledge may be created, explored and interpreted in new ways, also involving the scope of new research infrastructures.⁶

Notes

1. Notwithstanding the impact of the internet, the application of computing to the Humanities is not a new or recent development. The field presently known as Digital Humanities aims to use “information technology to illuminate the human record, and [bring] an understanding of the human record to bear on the development and use of information technology” (Schreibman et al. 2004, xviii). Historiographically, it traces its most immediate origins to 1949, when Fr Roberta Busa began work on a digital index variorum of the complete work of St Thomas Aquinas in order to investigate the manifestation of ‘presence’ in his oeuvre (Busa 1980, 83).
2. We wish to express our gratitude to Andrea Rapp (Universität Trier) and James Cronin (University College Cork) for their helpful comments and suggestions about this paper. The paper reflects on the whole the situation as it was written for this book in 2010 and was updated punctually for the final publication in 2014.
3. www.woerberbuchnetz.de.
4. The *LexicoLux* project aims to make the existing dictionaries of Luxembourgish available electronically and to link them with other materials; see <http://infolux.uni.lu/lexicolux/>.
5. See www.sprache-und-genome.de.
6. Space will not allow us to expand here on the issues of Research Infrastructures in any comprehensive sense; instead, we point the reader to the ESF Science Policy Briefing on Research Infrastructures in the Humanities (Moulin, Nyhan et al. 2011) and Ciula et al. (2013) where these issues are explored in much detail.

Bibliography

- Alijayl, Mohammed and Ophir Frieder. “Effective Arabic-English Cross-Language Information Retrieval Via Machine-Readable Dictionaries and Machine Translation,” in Henrique Paques, Ling Liu, and David Grossman (eds.) *Proceedings of the Tenth International Conference on Information and Knowledge Management (CIKM '01)*, ACM, New York, NY. 2001. 295-302.

- Alpert, Jesse and Nissan Hajaj. "We Knew the Web was Big" [Weblog]. *The Official Google Blog. Google*: 25 June 2008. Web. 11 February 2014.
- Antoniou, Grigoris and Frank van Harmelen. *A Semantic Web Primer*. 2nd ed. Cambridge Mass.: MIT Press. 2008. Print.
- Ash, Mitchell G., "Räume des Wissens." *Berichte zur Wissenschaftsgeschichte* 23 (2000) 235-242. Print.
- Busa, Roberto. "The Annals of Humanities Computing: The Index Thomisticus." *Computers and the Humanities*, 14 (1980) 83-90. Web. 11 February 2014.
- Berners-Lee, Tim, Hendler, James and Ora Lassila. "The Semantic Web." *Scientific American*, 284(5) (2001) 34-43. Web. 11 February 2014.
- Berners-Lee, T. (2006). "Linked Data – Design Issues" (27 July 2006). Web, 11 February 2014.
- Birkerts, Sven. *The Gutenberg Elegies: The Fate of Reading in an Electronic Age*. New York: Macmillan. 2006. Print.
- Bizer, Christian, Heath, Tom and Tim Berners-Lee. "Linked Data – The Story So Far." *Linked Data. Spec. issue of International Journal on Semantic Web and Information Systems*. 5(3) 1-22. Web. 25 August 2010.
- Burch, Thomas and Johannes Fournier. "Vom Buch zur Elektronischen Publikation – Textdigitalisierung auf der Basis von SGML/XML," in Alexander Mehler and Henning Lobin (eds.) *Automatische Textanalyse. Systeme und Methoden zur Annotation und Analyse natürlicher sprachlicher Texte*. Wiesbaden: VS Verlag für Sozialwissenschaften. 2004. 265-283. Print.
- Burch, Thomas and Andrea Rapp. "Das Wörterbuch-Netz: Verfahren – Methoden – Perspektiven," in Daniel Burckhardt, Rüdiger Hohls and Claudia Prinz (eds.) *Geschichte im Netz: Praxis, Chancen, Visionen. Beiträge der Tagung.hist 2006*. Berlin, 2007 (Historisches Forum 10, Teilband I), 607-627. Web. 11 February 2014.
- Canfield, Kip. "Modeling the Lexicon with Ontologies". *Digital Humanities 2009 Abstracts*. Maryland: MITH. 2009. Web. 11 February 2014.
- Ciula, Arianna, Nyhan, Julianne and Claudine Moulin, "Science Policy Briefing on Research Infrastructures in the Digital Humanities: Landscapes, Ecosystem and Cultures." *Lexicon Philosophicum: International Journal for the History of Texts and Ideas*, 1 (2013) 287-294. Web. 11 February 2014.
- De Smedt, K. "Some Reflections on Studies in Humanities Computing." *Literary and Linguistic Computing* 17.1 (2002) 89-101. Web. 11 February 2014.
- Dietrich, Michael, Weissmann, Dirk, Rech, Joerg et al. "Multilingual Extraction and Mapping of Dictionary Entry Names in Business Schema Integration," in *Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services (iiWAS '10)*. ACM, New York, NY, 2010. 863-866.
- Fournier, Johannes. "Vorüberlegungen zum Aufbau eines Verbundes von Dialektwörterbüchern." *Zeitschrift für Dialektologie und Linguistik* 70 (2003) 155-176. Print.
- Gilles, Peter and Claudine Moulin. "Luxembourgish," in Ana Deumert and Wim Vandenbussche (eds.) *Germanic Standardizations. Past to Present*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 2003. 303-329. Print.
- Gruber, T.R. "A Translation Approach to Portable Ontology Specifications." *Knowledge Acquisition* 5 (1993) 199-220. Print.
- Gotscharek, Annette Gotscharek, Reffle, Ulrich, Ringlstetter, Christoph and Klaus U. Schulz. "On Lexical Resources for Digitization of Historical Documents," in *Proceedings of the 9th ACM Symposium on Document Engineering (DocEng '09)*. ACM, New York, NY. 2009. 193-200.
- Hausmann, Franz Josef. "Die gesellschaftlichen Aufgaben der Lexikographie in Geschichte und Gegenwart," in Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand and

- Ladislav Zgusta (eds.) *Dictionaries*. Vol. 1. Berlin, New York: Walter de Gruyter. 1989. 1-19, here p 1.
- Hausenblas, Michael. "Exploiting Linked Data to Build Web Applications," *IEEE Internet Computing*, 2(4) (2009) 68-73. Web. 11 February 2014.
- Hildenbrandt, Vera and Claudine Moulin. "Das Trierer Wörterbuchnetz. Vom Einzelwörterbuch zum lexikographischen Informationssystem." *Korrespondenzblatt des Vereins für niederdeutsche Sprachforschung* 119 (2012) 73-81. Print.
- Himmelfarb, Gertrude. "A Neo-Luddite Reflects on the Internet." *Chronicle of Higher Education* vol. XLIII n. 10 (1 November 1996) A56. Web. 11 February 2014
- McCarty, Willard. "Attending from and to the Machine." Inaugural lecture. Kings College London. 2 February 2010. Web. 3 April 2010. PDF file.
- McGann, Jerome. "Introduction. Sustainability: The Elephant in the Room." *Online Humanities Scholarship: The Shape of Things to Come: Proceedings of the Mellon Foundation Online Humanities Conference at the University of Virginia, 26-28 March 2010*, ed. Jerome McGann. Houston: Rice University Press. Web. 11 February 2014.
- McArthur, Tom. *Worlds of Reference*. Cambridge: Cambridge University Press. 1988. 24-26.
- Moulin, Claudine. "Dialect Dictionaries – Traditional and Modern," in Peter Auer and Jürgen Erich Schmidt (eds.) *Language and Space. An International Handbook of Linguistic Variation, Volume I. Theories and Methods*. Berlin, New York: De Gruyter 2010, 592-612. Print.
- Moulin, Claudine, Nyhan, Julianne, Ciula, Arianna et al. "ESF Science Policy Briefing 43" *Research Infrastructures in the Humanities*, Strasbourg 2011. Web. 11 February 2014.
- Ó Corráin, Donnchadh. "Some Cruxes in Críth Gablach". *Peritia* 15, 2001. 311-320.
- Perelman, Lewis J. "School's Out: The Hyperlearning Revolution Will Replace Public Education." *Wired Magazine* (March-April 1993). Web. 11 February 2014.
- Rapp, Andrea. "Prof. Andrea Rapp on Tools, the Added Value of Networked Data and the Purpose of Quality Control." www.textgrid.de/en/press-kit/rapp.html. 25 August 2011.
- Reddick, Allen. "Past and Present in Samuel Johnson's Dictionary of the English Language." *International Journal of Lexicography* 23.2 (2010) 207-222. Print.
- Reichmann, Oskar. "Die Onomasiologische Aufbereitung Semasiologischer Dialektwörterbücher. Verfahrensvorschlag und Nutzen," in Hans Frieberthäuser (ed.) *Lexikographie der Dialekte. Beiträge zu Geschichte, Theorie und Praxis*. Tübingen: Niemeyer. 1986, 173-184. Print.
- Sandars, Nancy K. *The Epic of Gilgamesh*. New York: Penguin Classics. 1972. Print.
- Schreibman, S., Siemens, R. and J. Unsworth. "Introduction," in Schreibman et al. (eds.) *A Companion to Digital Humanities*. Oxford: Blackwell 2004.
- Turkell, William J. "A few Arguments for Humanistic fabrication." [Web blog entry] *Digital History Hacks: Methodology for the Infinite Archive* (2005-08). 21. November 2008. Web. 10 February 2014.
- Winner, Langdon. "Do Artifacts Have Politics?" *Daedalus*, Vol. 109, No. 1, Winter 1980. Reprinted in *The Social Shaping of Technology*, edited by Donald A. MacKenzie and Judy Wajcman (London: Open University Press, 1985; second edition 1999).

Too Much of a Good Thing?

Or, A Historian Swamped by the Web

Luca Codignola

I recently asked a former student of mine, now a brilliant Irish PhD with a Canadian MA and an Italian BA: “So, what do you think of Google Books?” In response, he sent me a PDF of Leopold von Ranke’s original German version of *History of the Popes* (1834-6). Without Google Books, he explained, he would not have been able to read it, except by travelling to one of the libraries that possessed this rare work. My former student is a regular exploiter of these new technologies. More than once I have e-mailed him asking for what I thought were obscure Irish periodicals in the field of ecclesiastical history, only to discover, from his response, that these were fully available through the internet. A peer of mine, a specialist in nineteenth-century American travel literature, is similarly enthusiastic, albeit somewhat frustrated. She has spent years trying to trace obscure travel books in several libraries around the world; now, most of these are available, in their entirety and for free, through a single website (see, for example, American Libraries, www.archive.org/details/americana).¹

This is exactly the point made by an American historian of my generation, John K. Thornton, in the preface of his *magnum opus*, which he had begun fifteen years previously: “During the 2000s [...] other developments helped shape the way the book grew. Gallica [...] began offering vast number of books for download, and I discovered that books I had once been able to consult in large libraries were available on my computer in a matter of minutes.” JSTOR and Google Books followed suit soon thereafter, making available to him yet more “sources I had only been able to read in summer jaunts to the Library of Congress, overseas in Paris or Lisbon, or perhaps at Harvard University.” This “flood of new material” had unexpected consequences, in that it altered his “approach to the text” and made him “turn increasingly to primary sources.” Meanwhile, the printed medium and even paper itself were almost abandoned: “[E]verything I used or needed was lodged in my computer” (Thornton xi-xiii).

In my prime, about forty years ago, I was an enthusiastic user of 6”x4” paper index cards. These grew by the day and at home they came to occupy several rows of library stacks, neatly stored in grey metal boxes. With the arrival of my own first IBM personal computer in late 1988, I paid another former

student of mine to transfer all the information contained in my paper index cards into new, identical but digital index cards. For a while I also kept the old paper index cards, then I discarded them to save space. (Needless to say, I have regretted that move about a dozen times, but not more than that.) The transfer to the new technology had been painless, though a bit costly, and the saving of time from then onwards astonishing. At the last count, my computer stores almost 45,000 files of various sizes, all manually cross-indexed; the equivalent of about 50 grey metal boxes of paper index cards. I can carry them with me at all times and revise and search them at ease. As I was born in the era of the typewriter, the computer represented, for me, a major step in terms of word and data processing. It did not, however, fundamentally change my way of being a historian.

About fifteen years later, the advent of the World Wide Web transformed my relationship with my computer, until then a one-to-one man-machine exercise, into a personal connection with a worldwide network that provided me with instant information, drawing upon a seemingly endless data bank. Almost overnight, all that pounding on my typewriter first and a computer later became obsolete. Had the web been available when I started my academic life, all those years, indeed decades, of entering my data into my own personal data bank would have been reduced to weeks. Any proficient person with access to the web might have done it quicker and more efficiently. A real revolution seemed to take place around me; a technological change comparable to that of the sixteenth century, when the invention of the printing press, via the earliest printed bibles that emerged in 1455 from the Mainz shop of Johannes Gensfleisch, alias Gutenberg, had made redundant thousands of *amanuenses*, the monks that over centuries had painstakingly copied and recopied unique manuscripts – rendering those carefully preserved treasures useless in the process. The words they contained ceased to be accessible to a lucky few, one at a time, in only one place. Instead, they became available to everybody, at the same time, in different places. The main difference between the two revolutions was that the printing press was only recognized as revolutionary long after it was already in place. Conversely, the digital revolution is taking place before our eyes and we are very much aware of it – although, as with all real revolutions, their long- and even medium-term consequences are completely unpredictable and allow for tentative explanations only several generations after their conclusion. Indeed, so much is being said and written about the digital revolution that simply keeping track of what is going on is a professional endeavour – and that profession is not mine. My profession

is to be a historian and my question is, how much and in which way has the historical profession changed on account of the existence of the web?²

In mapping the chronology of the digital revolution, we are dealing with three main stages. The first preceded the advent of the computer on the historiographical scene. Although historiography has undergone various modifications from Herodotus to the twentieth century, the recipe has basically remained the same. Fundamentally, the historian starts with a question, looks for sources wherever they may be and in whatever form they may present themselves, and then comes up with an honest, balanced and fair answer, acknowledging documentary gaps and reasonable alternative solutions. In spite of the well-recognized usefulness of unwritten sources (oral, visual, archaeological, statistical, etc.), the historian's world is the world of the written word, normally made available to him in archives and libraries. One of the reasons why, for a long time, most history was political history was that historians did not have time to look further than their main national archives. Whereas, for example, North American archival institutions often employed a clerk, usually a historian by training, in London, Paris or Rome, in order to identify and copy documents of major interest, only a lucky few historians could hire secretaries and copyists who worked for them away from home and mailed them the results of their endeavours.

From this point of view, a most significant change was the introduction of mechanical facsimile reproduction, such as microfilms, microfiches and especially photocopies. (The first photocopier was commercially produced by Xerox Corporation in 1958.). Late in the 1960s, instead of the extended visits to archives and libraries away from home normally enjoyed during sabbatical years or summer holidays, the usual procedure for any historian looking for primary sources became a short stay in a repository, the rapid selection of material by skimming documents or books, the insertion of a slip of paper indicating the selected page(s), and the sending of the whole folder to the institution's photocopying office. The real, in-depth examination came later, at home. Photocopying significantly altered the amount of reading that could be performed by any individual, allowing for, among other things, more time for the historian to examine different kinds of sources (Gilbert 521). According to one commentator writing in 1971, this was particularly true of the historian dealing with late modern or contemporary history. For her, the photocopy machine had become "an indispensable tool" simply on account of "the vast increase in the number of records that flowed into the archives in the nineteenth century and the enormous increase in the flow of books into the libraries that accompanied it" (Hanham 512).

The second stage of the digital revolution coincided with the advent of the computer onto the scene of the historical profession. It was almost invariably described as a revolutionary tool. In 1971, *Daedalus*, the journal of the American Academy of Sciences, devoted two full issues to an assessment of the state of historiography. Whether it was intended or not, when the editor first conceived these issues, the appearance of the computer, “the full significance of which is only just becoming apparent,” was acknowledged by most *Daedalus* contributors. They regarded it as an instrument of “almost limitless capacity.” One of them explained: “The correlation of numerous variables affecting large masses of data, assembled on a uniform basis, is precisely what the computer can do best; it is also what is most laborious, and in many cases virtually impossible, for even the most mathematically-minded of historians working without electronic aids” (Stone 1971, 71-72).³ All contributors agreed that in certain areas of historical study, such as in historical demographic and “much economic data,” the computer had “made possible work that could never have been done.” They also agreed that computers enabled “the creation of data banks from which it is possible to draw factual data [...] with a minimum of effort.” This outcome was “enormously facilitated simply by feeding data into a computer programmed to analyze it” (Hanham 510).

One of the contributors to the *Daedalus* assessment was less enthusiastic. According to British military historian Harold J. Hanham, who provided the most original and perceptive contribution by far, while the computer had made “the construction of economic models easier and have encouraged the expansion of the numerically-based branches of history,” he explained, “as yet little of this has rubbed off on most historians.” He subscribed to the contrasting view that “the greatest innovation of recent times ha[d] been the xerox machine, not the computer.” In his view, the photocopier had enabled the historian “to make cheap, readable, and more or less permanent copies of printed matter and of manuscripts,” eliminating, “almost at a blow the heaviest task of the historian – the copying of books and manuscripts.” According to Hanham, the computer had not revolutionized historiography. It had “done little more than carry a stage further the analysis of complex data which began long ago.” Its “chief benefit” had, until then, been “psychological – the very possibility of breaking away from what had come to seem a rather limited set of historical techniques ha[d] given a great boost to [the historians’] morale” (Hanham 511-512, 519). One must recall that in 1971, when the *Daedalus* collection was published, the world was still a relatively long way from the computer as we now know it. IBM introduced its first personal computer a decade later (1981); there were another three

years (1984) to go before Apple launched its Macintosh, the first successful mouse-driven computer with a graphic interface; Toshiba T 1100, the first portable computer, was launched in 1985, but it was a bulky and heavy machine for which real portability was still a dream.

In the mid-1990s, the third stage of the digital revolution took off, coinciding with the appearance and the rapid use of the internet technology. This generalized instant connectivity between individuals via both electronic mail (e-mail) and the World Wide Web (WWW, or simply 'The Web'), officially created in 1991.⁴ In 1996, personal and laptop computers (and later developments such as tablets) were joined by mobile telephones ('smartphones') as physical platforms for acquiring and sharing data. We are now some fifteen years into that phase and some reflections on the impact of this new technology on the writing of history and learning about it can be attempted. Before we address the Web issue directly, however, let us recap some of the main elements of the immediate pre-Web state of affairs.

Historians of the pre-Web generation started with library catalogues and, through them, made their ways to books and learned journals. The creation of a bibliography, the first step in any research, consisted of the painstaking sieving of other historians' footnotes and bibliographies. This was a process that usually included an initial explosion of titles and references, followed by a drying-up phase in which one realized that most items originated from or pointed to a limited number of master publications. A scholar could claim to have produced an original contribution to knowledge when he had written something that added to, or criticized, these master publications. Aside from reading it all, the real and first problem confronting the historian was to find a hard copy of the publications listed in one's preliminary bibliography. Even in the most advanced library networks, simply to know in which library a book or a journal was physically available represented no mean issue. Hence, the necessity to travel to some core libraries of international reputation for both its contents and accessibility. There one's time could be used more efficiently. Similar problems were encountered when hitting the second step of any historical research – the finding, retrieving and processing of archival and documentary evidence, except that in most instances archival documents are unique by definition and could not be accessed in any location except the original one. Hence, the significance of the 'xerox revolution' so well described by Hanham, including the use of other physical means of reproduction, such as microfilms and microfiches. Taken all together, these new technical supports partially solved the problem of uniqueness

and made it possible to literally take home a vast amount of material for future use at a more leisurely time.

In the 1970s, the availability of reproductive facilities coupled with the new computer technology seemed to make it possible to fulfill the librarian's dream, i.e. to reunite all books and journals under one roof. Physically, the trend towards the 'one big comprehensive library' had proved to be impossible to fulfill, so that the fragmentation of the various repositories, including the rare books libraries, had been recognized as a fact of life that could partially be amended through a comprehensive card catalogue. From that perspective, computers not only made the virtual amalgamation of card catalogues possible, but actually solved most physical constraints by simply doing away with the bulky card catalogues of yesteryear. The paper cards were copied onto new virtual cards, accessible either in the library itself or from wherever there was a cable connection. (In fact, the networking of library and office computers can be regarded as a first step in the direction of the global networking later provided by the internet.) Most libraries destroyed their card catalogues in the process. Some also discarded their more space-consuming original items after microfilming them. Harvard College Library, one of the best libraries in the world, began to destroy its newspaper collection as early as 1970.

The contributors to the *Daedalus* collection of 1971 were well aware of the revolutionary transformation of the library world that was taking place before their eyes. They were all enthusiastic about it: "Every page of every book on record in a great information retrieval center, each big library linked to the center, each library user free to type the code for the very page of the very book he wants? Technically such a project is already feasible." Here is how the new system worked:

[T]he engineers have designed [...] a system that points to the way in which information-retrieval can be made of immediate use to historians. A great collection [...] has been abstracted, indexed, and photographed. A code-book has been produced for each user. A journalist who wishes to know, for instance, what great floods have occurred in recent years has only to look up the appropriate code, dial it, and his machine will give him an abstract of the information available in the files. From that abstract he can select a particular article, and by dialling its code can secure a reproduction on the screen before him. Clearly, such system is extremely useful to the journalist or the historian who wants to get at the basic facts about a particular subject. [...] there is no limit to possible extensions across the country (Hanham 513, 515-516).

New avenues opened up for archival material as well: “Facsimiles of all manuscripts can be assembled in one place making possible a previously unknown exactitude in establishing the relation among various drafts of a manuscript, in placing a fragment in its appropriate context, and in the dating of letters or memoranda.” Furthermore, in order to avoid the bulkiness of book collections that is usually associated with the publication of historical sources material, “[w]ould it not be possible to achieve the same purpose if microfilms of the collected and edited materials were accessible to libraries?” (Gilbert 521, 523)

About a generation later, historians appear to be much more sceptical. Firstly, they are all too aware of the unfulfilled hopes engendered by the sudden trend towards quantitative history and what came to be known as Cliometrics; that is, the application of econometric and mathematical techniques and methods to the study of history. That application had indeed been made possible by the advent of the computer on the historical scene. “It is painful to admit,” British historian Lawrence Stone regretted in 1971, “that the advent of a technical gadget should dictate the type of historical questions asked and the methods used for solving them” (Stone 1971, 72). Hanham, however, warned that no forecast could be made, as normally new technologies “must await a new generation that takes their use for granted because it has moved on to other novelties” (Hanham 512). Hanham was right. In spite of the opinion of French historian François Furet, who described quantitative history as being “used so sweepingly that it cover[ed] almost everything,” allowing the “new panorama of data” to shake “the premises of [the historian’s own] profession” (Furet 151, 154), that trend had a comparatively short life. In the end, it only influenced the closed circuit of economic historians.⁵

Secondly, and perhaps even more significantly, simply by looking at their filing cabinets overflowing with unread photocopies, or at their office bookshelves stacked with piled-up cardboard boxes containing sealed microfilms, historians have also become aware that too much material might just be as dangerous as too little. For one thing, it adds to the researcher’s sense of guilt for wasting so much time since her last visit to the archives. For another, it creates the frustrating sentiment that everything is ‘out there’ for anyone to catch, in a virtual Borgesian Library of Babel (Borges), but that there is too little time to process it all. At any rate, the availability of too much information makes it impossible, or at least blurs one’s ability to distinguish between relevant evidence and casual leftovers from history. In a way, the formidable growth of available primary sources engendered by the new technology – in the form of photocopies, microfilms, microfiches

and the computer – has simply extended the methodological problems of late modern and contemporary history to all the chronological brackets of historiography. One can argue that the real problem facing today's historian is not so much that of finding and assessing his sources, but rather that of managing increasing masses of unprocessed information that quite simply swamp him and make him gasp for simplicity. In fact, this overload of information seems to allow for the writing of two kinds of books only: the very general (interpretive, impressionistic, oversimplified) and the very focussed (new evidence, narrow set of questions, limited public). Anything in between – a good synthesis based on a sound coverage of the available primary documentation – becomes very difficult to write and hence very rare.

Although most historians probably regret having overestimated their ability to process accumulated documentation, none that I am aware of nurture nostalgic hopes for a pre-xerox golden age. In fact, from a methodological point of view, facsimile reproduction facilities and the advent of the computer have simply enlarged the opportunities for knowledge without significantly changing the way historians read, teach and write. As all professional historians well know, the simple accumulation and analysis of data is not enough. When confronted with their material, their main challenge and responsibility remains that of organizing their materials and “think[ing] creatively about them” (Graubard xi). At the core of the matter is the unavoidable realization that one's time is limited and that a human being is not a computer. Whereas the latter can absorb a potentially unlimited amount of pieces of information, the former is limited in what she can do with her material. Here is, back in 1971, Hanham again: what if, in an ideal world, “xerox copies were supplied free of charge and were readily obtainable from all countries in the world, from all collection of archives?” (In a way, this is what happened later through the extensive sharing of library and archival resources via the web.) “The answer, I am afraid – Hanham continues – is that at the moment most of us would not find the writing of history any easier. The human body is as much a data processing machine as a computer, [but] we have not yet found easy ways of training it to absorb large quantities of data.” The bottom line is that “the workaday historian must read and read and read again.” In a way, American Atlantic historian Philip D. Curtin was expressing a similar realization of the historian's limitations: “I know only a few of the more common languages [...] With a good knowledge of Russian or Armenian or Chinese, this study could have been considerably more authoritative [...] On the other hand, to have taken the time necessary to learn all relevant

languages would have meant that it would never have been written at all” (Hanham 519; Curtin x).⁶

We have now reached the third and, so far, last stage of the digital revolution, that of instant and potentially unlimited connectivity through the Web, e-mail, tablets and mobile telephones, including RIM’s Blackberries (introduced in 1999) and the following waves of smartphones.⁷ Let us quickly do justice to e-mail and mobile telephones by saying that they have both been described as ‘tyrannical’ instruments. E-mail, in particular, has been accused by John Freeman, a literary critic who now edited *Granta Magazine* until 2013, of many misdeeds. It requires one’s constant attention, eschews face-to-face conversation, interrupts one’s train of thoughts and makes it more difficult to sift relevant information. Furthermore, its rapidity and the typically condensed nature of its messages facilitate a huge percentage of unread or misunderstood messages (Freeman; also Day). The process through which e-mail connects the two correspondents is comparable to the throwing of a myriad of meaningless words onto a screen in the same fashion as, in the old days, we would have thrown a bottle of ink onto a sheet of paper, hoping for a coherent message to appear and to be understood as such. There is, however, no way back. Instant connectivity is here to stay in spite of its critics, in the same way as in the late nineteenth century automotive technology won its battle against the nostalgic proponents of horse power and manure.⁸ If anything, the debate nowadays seems to be on whether the Web should be free or there should be a way of making users pay for access. A book by former *Wired* editor-in-chief (2001-2012), Chris Anderson, forcefully suggests that writer Stewart Brand’s famous cyberlibertarian declaration, “Information wants to be free” (1984), is still proving its full value. The cost of getting information is lower and lower, getting nearer and nearer to zero, so that more people have access to it and have become themselves providers of information.⁹ (More on this further on.) For the time being, let us recall that Malcolm Gladwell, a *New Yorker*’s popular culture critic, is of a different opinion. He points to the enormous cost of Google infrastructures and to the fact that YouTube, the apparently very successful Google activity, “has so far failed to make any money” for its umbrella company (Gladwell 82; Auletta).¹⁰

This is, of course, food for thought for mass mediologists and macro-economists. With regard to historians, let us recall the main question from which we have started – how much and in which way has the historical profession changed on account of the existence of the web? In his 1971 *Daedalus* article on the new historiography, Hanham warned that it normally takes one generation to assess the impact of a new technology. Time is up,

then, for some tentative answers. Here is one. The impact of the Web on the historians who reached their professional maturity before the advent of the Web or in its early stages has been, in my view, relatively modest.¹¹ There has, indeed, been a revolution, but only in the sense that the photocopier too had produced a revolution in the writing of history through a sudden increase in the magnitude of accessible sources and the rapidity in the accumulation of data. The historian still starts with a question and then looks for sources wherever they may be. It is true that mature historians using the web have modified their initial habits. Their first step is no longer likely to be in the direction of the library's card catalogue. This has generally been replaced by a rapid search through the Web to find basic information – facts, dates, names, biographies, texts, statistics, et al. What used to be gathered from reference works such as encyclopedias and dictionaries, normally available in any library's open stacks, is now found in a Wikipedia entry or in a similar finding aid. For the mature historian, however, this is simply a more convenient way to find books and journals, or to take advantage of archival sources that have been reproduced and made available on the Web. Their research methodology has not been substantially altered. The critical approach to sources, which was the basic element of their professional training, is still there in full.

This limited impact is well exemplified by the debate about the Google Books digital library project, introduced at the Frankfurt Book Fair on 6-10 October 2004 as Google Print. This project scans printed texts (mainly books) and converts them into a digital database (Roncaglia 52-122). As we know, Google has already signed a number of contracts with significant libraries throughout the world and the project is now fully operative, albeit far from completed. The sheer magnitude of this project – its final objective is that of making all books ever printed available through the web¹² – reminds us of the famous Arthur C. Clarke's 1953 science fiction short story, 'The Nine Billion Names of God,' a story that makes the world end when all the possible names of God have been identified and written down.¹³

Indeed, American linguist Geoffrey Nunberg's 2009 devastating critique of Google Books focussed on how clumsily, incompetently and unreliably Google engineers were doing their job. The result, according to Nunberg, was "a mishmash wrapped in a muddle wrapped in a mess," on account of hundreds of thousands of "endemic" errors regarding book dates, classification, authors, editors, writers of introductions, etc. (Nunberg). For all his polemic vehemence, however, Nunberg did not question the nature of the project itself, only its initial implementation. Just like in the 1970s, the way to go was the amalgamation of card catalogues, so now there is no other

way but the digitization of the printed heritage of humankind. This is what American cultural historian Robert Darnton believes too, in spite of his own fear that the rush to transform books into bytes might engender the loss of irreplaceable details: “Whatever the future may be, it will be digital” (Darnton 15; also Negroponte).

If the impact of the Web on mature historians is relatively modest, then on the younger or would-be historians, let alone students of history or users of historiography such as journalists, that impact is much more pervasive. All professors are well aware of the reluctance of their students to go beyond the Web in the search for their materials. What journalists and politicians used to say about television – what is not on television does not exist – journalists and students now say about the Web – what is not available on the Web does not exist. This is, of course, sheer laziness. But far from being an occasional attitude, laziness seems to be a feature that is intimately linked to a medium as easy to manoeuvre as the Web. The fact that distance from the original source of information does not represent an obstacle any longer often goes hand in hand with the idea that knowledge can be accessed from anywhere. Although quality is still an issue (see Nunberg’s critique above), mass digitization of books via Google, Microsoft, Carnegie Mellon University et al., has already made millions of books available for viewing online. One of the consequences of this ease in getting to ‘the real thing’ is that the consultation in loco of books and learned journals, let alone of archival documents, is no longer a real option. In fact, research trips have become less necessary and, at the same time, funding for research has reached an all-time low throughout the Western world. Whereas extended visits away from home were viewed as an entitlement for would-be scholars of the baby-boom generation, these have now become a luxury for the younger scholars born a generation later. In this framework, one cannot but partially excuse the laziness of some younger Web enthusiasts and lament the reluctance with which many European institutions are adjusting themselves to the so-called Digital Humanities – let alone the ‘user unfriendliness’ of all too many libraries and archives, especially in southern Europe. Rome’s National Library (Biblioteca Nazionale), for one, still hands out only three items per day.

A second consequence of this detachment from their *real* sources, as opposed to *virtual* sources, coupled with the myth that all items can be accessed from any place in the world – something that is still far from true, especially with regard to archival documents – is that younger scholars normally lack the shock of recognition and the physical impact with a historical artefact – say, a broadsheet that was distributed when the American

War of Independence was fought in the streets of Boston, complete with ink stains, corrections and thorn ends; or a series of letters written by one person that we follow for a period of his life, until the correspondence ceases and we realize that person had died. Tristram Hunt, a British historian now involved in politics and broadcasting, thus lamented the deal made by Google and the British Library to make 250,000 books ubiquitous and then available on everybody's screen:

[I]t is only with manuscript in hand that the real meaning of the text becomes apparent: its rhythms and cadences, the relationship of image to word, the passion of the argument or cold logic of the case. Then there is the serendipity, the scholar's eternal hope that something will catch his eye. Perhaps another document will come up in the same batch, perhaps some marginalia or even the leaf of another text inserted as a bookmark. There is nothing more thrilling than untying the frayed string, opening the envelope and leafing through a first edition in the expectation unexpected discoveries (Hunt 2011).

Indeed, all this is lost on the computer's screen.

The third consequence is directly linked to both the rapidity of access and the consultation-at-home option – the process of becoming steeped in the past, an essential feature of the historian's work, is lost when it becomes too quick and easy and the flickering of images replaces years of patient sedimentation. Finally, travelling to another country or city, inconvenient and costly as it might be, allows a spatial and psychological recognition and apprehension of the physical environment in which certain events took place, let alone the human contact with new colleagues and the making of friendships, personal and professional, that may accompany a person throughout her life (Brandão 670-671).

The most serious methodological threat to the learning and the writing of history, however – a real historian's ordeal – comes from a new concept of what is a source and what use historians can make of it. (Note that internet lingo always employs 'resources,' not 'sources,' when referring to its riches.) Sources, of course, do not consist only of what in popular parlance is referred to as 'documents' (i.e. handwritten items), but also of printed matters, such as books, articles, etc. In their sources, historians identify bits and pieces of evidence, upon which they build their narratives, such as a date, a name, a law, a commercial transaction or a personal letter exchange. They also derive more intangible elements from which they provide contexts and interpretations, such as points of view, nuances, feelings, innuendos or falsi-

ties. In post-Web users, commonly referred to as ‘digital natives’ or the Net Generation, this second feature seems to be lost – or at least unconsciously, but irremediably, fading away.¹⁴ This new attitude originates in the very top; that is, in the most successful entrepreneurs of the digital world who have invented Microsoft, Apple, Google, YouTube, etc. and now preside over the hardware and the software of their search and processing engines. They treat sources as data banks;¹⁵ their only preoccupation, aside from profit, is how to proceed in these data banks in the quickest and simplest way. As Sergey Brin, co-founder of Google, phrased it with regard to the Google Books project: “There is fantastic information in books. Often when I do a search, what is in a book is miles ahead of what I find on a Web site” (in Rich).¹⁶ This apparent appreciation of the value of a book is, in fact, its reduction to the level of a provider of information, i.e. of data to be entered in the string of keywords required, for example, by the Google search engine.¹⁷ As pointed out by Nunberg, however, “we are sometimes interested in finding a book for reasons that have nothing to do with the information it contains” (Nunberg). Even when we only read a few pages, through the architecture of the book, its cover page, its table of contents, its index, its publication history and its footnotes, we get a perception of its coherence, of the relationship between its parts and of its context. Similarly, with an article, the context of its publication is often as significant as the contents of the article itself – the prestige of a learned journal, the position of the article in a newspaper, its proximity to related or unrelated matters, and so forth. A good part of all this is lost in the digital rendering of an originally written text. In any case, to have it contextualized is an extra effort that requires a rather discerning reader.

Unconsciously or not, students of the Net Generation eschew complexity for simplicity and go straight for the ‘useful’ detail, wherever that is. Usually, this is to be found in several entries, whose hierarchical order is determined by the search engine according to criteria on which the researcher has no knowledge or control. (The Google search engine, for example, does not actually search the internet, but it searches the index that the company itself has previously made of the internet and that has then stored in its own servers together with a copy of searches previously made by users.) At best, in the words of cultural historian Roger Chartier, this operation is leading to a universe “*de fragments décontextualisés, juxtaposés, indéfiniment recomposables, sans que soit nécessaire ou désirée la compréhension de la relation qui les inscrits dans l’oeuvre dont ils ont été extraits*” (Chartier). At worst, this is a process of copy-and-paste that resembles plagiarism and is completely devoid of any expert filtering or scholarly assessment. (Just apply

this same intellectual process to that of a person who decides to cure her serious illness by looking up her symptoms on the internet, only to realize how catastrophic the results may be.) Worse still, the rapid pace of any Boolean search, together with the still imperfect readings provided by the so-called Optical Character Recognition technology and the like (Milligan), leads to a crescendo of steps of exponentially-growing levels of intellectual sloppiness. The first step in this crescendo is the use of one search engine only (Google is clicked three billion times a day), and of only one, two, or exceptionally three keywords. The second step is the reading of the first page of results, rarely going past the second, and normally limiting oneself to a quick glance over the entry's title, short summary and link. If one ventures beyond the title and summary and actually accesses the item, his viewing time will average 56 seconds. Nunberg confirms that, in this irresponsible approach to knowledge, it is considered sufficient "just to find a chunk of a book that answers our needs and barrel into it sideways" (Nunberg).¹⁸

Take Wikipedia (launched 2001), still the most popular web resource where *wiki* (Hawaiian for *quick*) replaces *ency* (from *enkyklios*, Greek for *circular* and *general*). In late 2001, 21-year-old Simon E. Pulsifer began to contribute to the English-language edition of Wikipedia. As of September 2007, he had edited over 94,000 entries and had started more than 2,000 articles. Meanwhile, his astounding productivity had made him the subject of many human-interest stories in newspapers, magazines and television shows. Pulsifer comes from a family of scholars and has a university education of his own (<http://en.wikipedia.org/wiki/Simon_Pulsifer>; also Gaudiano, Shimo, Gulli, Grossman). Yet 24-year-old Ryan Jordan ('Essjay'), another Wikipedia contributor, reportedly edited thousands of entries, before being exposed for not being the tenured professor he claimed to be (Schiff, also http://en.wikipedia.org/wiki/Essjay_controversy). The Web might be "the great equalizer, the last bastion of democracy" (Sabrina), but participatory democracy – Wikipedia purports to be "the free encyclopedia that anyone can edit" – is not the equivalent of a peer review system and accessibility is, indeed, not the same thing as scholarship.

Even though he is not concerned with scholarship, but only with accurateness and consistency, British-American Web entrepreneur and expert Andrew Keen (a BA in history and an MA in political science) has been a passionate enemy of the Wikipedia approach. He describes it as a typical example of the unreliability and the sloppiness of the new Web culture: "Since Wikipedia's birth, more than fifteen thousand contributors have created nearly three million entries in over a hundred different languages – none of them edited or vetted for accuracy." Keen compares Wikipedia's popularity

with Encyclopaedia Britannica. The latter's website provides authoritative and non-amateurish information, but is far from being a favourite among Web users. (Incidentally, the Encyclopaedia Britannica website requires a user fee.) Keen also takes issue with the reliability, ranking and possible manipulations of the Google search results and of those of other less popular search engines. In fact, his *The Cult of the Amateur*, notorious as it is among his fellow bloggers, is a heavy-handed book-length attack on the egalitarian and the participatory philosophy of the Web. According to Keen, in the hands of a community of digital utopians the new user-generated Web culture has willfully blurred, if not blatantly extinguished, the difference between fact and opinion, quality and pulp, authoritativeness and blurb, professionals and "noble amateurs" – an explicit reference to the utopian believers in the 'noble savage' myth (Keen 4, 35-36).¹⁹ We are heading back towards the new Dark Ages, Keen warns. Web pioneer but fellow pessimist Jaron Lanier fully shares Keen's information nightmare – everybody jots and skims and nobody understands:

[My] words will mostly be read by nonpersons – automatons or numb mobs composed of people who are no longer acting as individuals. The words will be minced into atomized search-engine keywords within industrial cloud computing facilities located in remote, often secret locations around the world. [...] The vast fanning out of the fates of these words will take place almost entirely in the lifeless world of pure information. Real human eyes will read these words in only a tiny minority of the cases," (Lanier ix)

most likely simply to scan, rehash, and misrepresent them in the quickest and sloppiest way.²⁰ The general sense is that of a self-perpetuating cycle of rehashing what, at best, is general knowledge, which adds nothing new but the mistakes.

In his book, global historian Luke Clossey describes the "topological" approach to history – from topology, the study of the non-quantitative properties of geometric space. He offers as the best known example of such an approach the iconic map of the London Underground, where what is important is the relationship between the various points, not their exact distance (Clossey 90; also Henderson et al.). The fragmented universe of the Net Generation is similarly topological – bits and pieces of knowledge that are completely unconnected and need not refer to any authorial or authoritative source. This learning process is similar to that of a person who tries to know the city of London by using the Underground and exiting into open space at random –

one station today, another tomorrow – without ever walking in between them. Indeed, in the end, and after a very long time, that person may theoretically get to know the entire city. During the long process of acquaintanceship, however, her only real familiarity would be with uncontextualized and unconnected nuggets of information casually added one on top of the other, without any reference to the treasure of overall knowledge. Hypertextuality, of course, is what distinguishes a Web document from a classical, linear document. The latter applies the Aristotelian rhetorical canons of writing (introduction, treatment/development/demonstration, conclusion), assembles its material in a sequential order, and has a beginning and an end. The former, through its links, drives its user in the opposite direction and invites him to abandon the original document in order to jump to another Web document that may present itself in a variety of formats – another written text, a picture, a video, a music, etc. There is no logic linearity, no chronological sequentiality, no departure or arrival points and no contents hierarchy. These hypertextual features of a Web document are normally regarded, of course, as an improvement over the more limited opportunities offered by a static, written document. Yet, hypertextuality also substantially adds to the fragmentation of knowledge and the subjectivity of its apprehension. Whereas pre-Web students grew with the idea that there was an objective corpus of knowledge and that they should get acquainted with it in the best possible way, digital natives believe that there are as many corpora of knowledge as there are subjective options to be exploited in the Web.²¹

Self-training and self-learning on the part of students, based on a complete reliance on Web technology, is now being embraced by many universities as a cost-saving measure – online content means, in the end, fewer faculty members. It is also being promoted, since 1996, by the New Technology Foundation in the United States.²² This group's approach to higher education encourages the creation of even more distance between students and faculty by transforming a professor's traditional leading role into a role of facilitator and motivator and, in fact, consigning the students to the anarchy and the unreliability of the Web. A special issue of the Québec popular magazine, *L'Actualité*, lamented the backwardness of the universities of the province while extolling the modernity of the technological ways of learning: "*Pourquoi leurs étudiants suivraient-ils en classe le cours d'un professeur ennuyeux quand, en direct dans Internet, il peuvent écouter gratuitement les cours d'un expert d'un prestigieux établissement européen ou américain?*" (Beaulieu). Once again, in the current approach to knowledge fostered by leading Web entrepreneurs and thinkers, knowledge is viewed as an immense data bank whose key is represented by the rapidity of access,

not by an interpretive guidance through the complexity of contexts: “*Comment être professeur d’université quand vos étudiants peuvent à tout moment, en classe, sur leur portable, contre-vérifier votre enseignement dans le site Web d’une autre université?*” In this overall perspective, the fact that these professors sometimes find themselves speaking to “un mur de 30 portables” becomes irrelevant (Grégoire 28, 32). Furthermore, the sheer immensity of this data bank is somewhat overpowering in both the inability on the part of most users to sift the relevant information theoretically at their disposal, and in the wasting of time that is intimately linked to the very concept of unlimited searchability. The YouTube network, created as recently as 2005 by Jawed Karim, Chad Hurley and Steve Chen, now contains over a billion videoclips and is accessed over a billion times a day. Theoretically, a single person could spend five times the length of her entire life just looking for the ‘relevant’ item. This might be an entertaining pastime for most and may enhance artistic creativity in some. For would-be historians, however, this curiosity-driven haphazard itinerary among user-generated videoclips is the equivalent of spending one’s professional lifetime in reading readers’ letters to the editor of a newspaper – and nothing else.

The dangers inherent in this intellectual process of overloading are multiplied by the number of providers of information. Not everybody agrees, however. Seth Godin, who describes himself as “a writer, a speaker and an agent of change,” rejoices: “When there are thousands of people writing about something, many will be willing to do it for free.” In a world “where there is room for anyone to present their work, anyone will present their work.” According to Godin, there is no need for elaborate mechanisms to monitor and judge the quality of content (peer review, for example), because web space is infinite and “[i]n a world of free, everyone can play.” Critics of Keen’s and Lanier’s catastrophic views are a majority among Web writers and subscribe to Godin’s opinion. They treasure the amateur approach because, according to them, rank-and-file users are more creative than professionals and stimulate experts to participate in the debate and to share their knowledge with the unwashed. The amateur approach, they also believe, allows for the constant vetting of such instruments as Wikipedia and does not purport to represent any final truth – as books normally do. Furthermore, they conclude, was not the pre-Web professional culture as much flawed as the web culture – with fewer chances of being corrected and revised in such a quick way? (see for example Lessig). At any rate, some *impromptu* contributions “might even be really good,” Godin adds, suggesting that “[a] good book review on Amazon is more reliable and easier to find than a paid-for professional review” (Godin).²³

Well, not quite. Aside from the unedited, sloppy and singularly careless prose favoured by bloggers, even some more 'learned' contributions might be really bad or not represent a state-of-the-art opinion. An Amazon review of a book on the origin of New England Puritanism is not likely to be as good or significant as a review that is printed in *The William and Mary Quarterly* or published through the specialized blog H-Atlantic – although the former is certainly cheaper and easier to find. By making free, ubiquitous and quick the most important features of any search, Net Generation members are led towards the relinquishing of any hierarchical value judgement. In doing so, they also separate knowledge from scholarship. This may be good enough for a person looking for the best spa resort in the neighbourhood (although it is not); it might be somewhat excusable in journalists (who are pressed for time and may correct themselves the next day); but it is not acceptable in historians, scholars or students. "Neatness is for historians," Godin concludes, but the real world is "messy" (Godin). Actually, it is just the opposite. An overly confident and unfiltered reliance on fragments drawn from the Web and recomposed according to one's preconceived needs and hypotheses makes for 'neat' stories. But the historian's role must be, on the contrary, that of showing the complexity and more often than not the inexplicability of what happened in the past.

Undoubtedly, Keen's critique on the "cult of the amateur" strengthened my awareness of my own lack of professionalism in the debate surrounding the technology, meaning and future of the digital revolution. Common sense, however, based on past experience with similar movements, seems to point in one direction only: digital knowledge is here to stay, and so are amateurs messing around with what knowledge 'should be,' if only it were left in the hands of professionals. Let me then restrain myself to my own field of expertise in order to answer my original question – how much and in which way has the historical profession changed on account of the existence of the Web?

Mature historians have profited greatly from the existence of the Web. Primary and secondary sources are now available to everybody, everywhere, at any time, to a degree that was quite simply unthinkable four decades ago, when, in 1971, the *Daedalus* debate on the future of historiography was convened. Websites and digital journals allow almost instantaneous scholarly exchanges of the kind that would have been impossible in the age of the printed word. (Professional book reviews in printed scholarly journals, one may recall, appear one to two years after the publication of the book.) Yet, even among professional historians, this undoubtedly positive trend must be weighed against a problem that can be summarized

by William Shakespeare's proverbial sentence, "[t]oo much of a good thing" (Shakespeare). In a world where how to allocate what is scarce is the main issue, the problem of the internet is just the opposite – abundance. There is a pervasive yet frustrating sense that there is 'too much to do, too little time,' as when the White Rabbit meets Alice and mutters "Oh dear! Oh dear! I shall be too late!" before entering Wonderland at the risk of drowning in a "pool of tears" (Carroll). This information overload has extended the methodological difficulties of late modern and contemporary historians to all chronological brackets of historiography, multiplying them by a factor equal to the pieces of information at one's theoretical disposal. Still, professional historians have been trained to deal with this new challenge. They know – or should know – how to combine a good initial question with the relevant sources and an original conclusion.

As the coming of age of younger historians or students of history coincided with the advent of the Web, their psychological approach to sources and their methodological awareness was shaped in ways that are not to be found among their older colleagues.²⁴ On the one hand, they are less likely to suffer from the White Rabbit syndrome. They favour simplicity and usefulness and eschew complexity. They conceptualize knowledge as an assemblage of nuggets of information. They presume that the information they want is hidden away, yet available, somewhere in the digital universe. They believe that the real challenge is to enter the right keyword sequence into their Web search engine. On the other hand, they postulate a fragmented universal knowledge. They trust that a chaotic Web goes hand in hand with an infinite potentiality of apprehension. They avoid sequentiality and authoritativeness in favour of the somewhat casual way of hypertextuality and subjectivity. In both instances, going digital means changing the way we think and that is not an option for the younger generation as it might have been for their elders. In sum, this seems to be a major anthropological (and perhaps neurological) change whose medium- and long-term consequences are yet unfathomable.

Historians, of course, old and young, are not immune to this change. Yet, the Web has not altered the fundamental objective of the historical profession, to discover what happened to men and women of the past and to tell it to the men and women of the present. Neither have the rules of the game (sometimes referred to as professional standards) changed. Historians should still seek objectivity, avoid biases, acquire the necessary competences, acquaint themselves with previous historiography and examine their sources critically. This professional recipe applies to both traditional sources and to the Web. With regard to the Web, however, historians must,

first, be wary of the enthusiasm that derives from this newly-found – and God-sent – intellectual abundance. Secondly, they must strive to assess the Web-generated contents for their transparency and the reasons behind their production. Thirdly, and perhaps most importantly, they must refrain from being distracted by and wasting their precious time in chasing technological novelties and semantic instruments that are not theirs. (Remember the White Rabbit syndrome.) In due time, these novelties and instruments might even become useful.

Notes

1. All websites mentioned in this article were last accessed on 20 August 2011, except when otherwise stated. The author wishes to thank Matteo Sanfilippo (Università della Tuscia), Vittorio Zambardino (*L'Espresso group*), and Matteo Binasco (Università di Genova), his 'brilliant student,' for their most useful comments on early versions of this chapter.
2. Books and articles (digital or otherwise) that try to describe the new internet environment for the historian, suggest solutions and fathom consequences are innumerable. Just like printed bibliographies of previous generations, however, most of the new 'instruments' or 'guides' meant for the digital needs of newer generations become obsolete the moment they reach their readers. Vitali and Genet and Zorzi are worth reading.
3. Gilbert cautioned that "we are only beginning work in [...] the use of the computer for historical research" (Gilbert 526).
4. Simply put, the internet is the infrastructure, or the medium; the Web is one of the applications that uses it, together with the exchange of music and video files, software updates, e-mail, instant messages, VoIP telephony, audio and video streaming, etc.
5. See Stone's later reflections on quantification and the use of the computer in historiography (Stone 1979).
6. Hanham believed that books and journals would have remained the historian's basic medium of communication "for a long time," but that "we must be prepared eventually for innovations of presentation" – a prophetic hint at the Microsoft Power Point presentations of the next generation (Hanham 517). To make the story even more complicated, reading too is a process that is far from being clear. Robert Darnton himself warns that "reading remains mysterious," as "[r]eading itself has changed over time" (Darnton 201-202).
7. Science fiction tells us that there is but one next step: "Beam me up, Scotty! There's no intelligent life here." That was, you may recall, *Star Trek's* Capt. James T. Kirk ordering his chief engineer, Montgomery Scott, to disintegrate and transport him back to the starship *Enterprise* from some God-forsaken

rotating piece of rock in a galaxy far, far away. For the history of the franchise, see http://en.wikipedia.org/wiki/Star_Trek.

8. Mailing lists, newsgroups, and social networks such as Facebook or Twitter, significant as they are, are specific applications of basic communication principles inherent in the connectivity allowed by the Web, e-mail and mobile telephony. The difference, if any, between the World Wide Web and Web 2.0 (coined 2004) is still controversial, whereas the transition towards 'cloud-computing' is a more recent development.
9. Stewart Brand's declaration first appeared in *Whole Earth Review*, 46 (May 1985), 49, which itself originated (as *CoEvolution Quarterly and Whole Earth Software Catalog*) in the post-beat and hippy counterculture as an educational tool. The book is Anderson's *Free* (2009), which was enthusiastically supported by specialized bloggers Godin and Cuban.
10. Auletta points out that YouTube lost US\$500m in 2009.
11. This assessment does not include a number of historians who had been trained with pen and paper around the time of World War Two, whose technological options were represented by typewriters and carbon copies. They adjusted with some difficulty to the onslaught of the xerox revolution, but were devastated by the arrival of the computer. Some wasted the last potentially productive years of their professional careers trying to cope with the basic notions of computer literacy that condescending students imparted to them – ASCII characters and symbols, DOS files and directories, save and delete keys, and the like.
12. According to Google's own estimate (August 2010), books in print are 129,864,880. This number does not include microfilms, microfiches, audio-books, maps and grey documents, but takes into account various editions of the same work. The story of the Google Books project, including the legal battles waged against it, is narrated in *Google and the World Brain*, a 2013 documentary film directed by Ben Lewis.
13. Here is a summary of the story's plot. The monks of a Tibetan lamasery seek to list all of the numerous names for God, since they believe the Universe was created in order to note all the names of God and once this naming is completed, God will bring the Universe to an end. Three centuries ago, the monks created a writing system in which, they calculated, they could encode all possible names of God. Writing the names out by hand would take another 15,000 years. The monks wish to use modern technology in order to finish this task more quickly. So they rent a computer capable of printing all the possible permutations, and they hire two Westerners to install and program the machine. After three months, as the job nears completion, the two computer engineers fear that the monks will blame the computer, and by extension its operators, when nothing happens. Thus, they delay the operation of the computer so that it will complete its final print run just after their scheduled departure. After their successful departure on ponies, they pause on the mountain path on their way back to the airfield, where a plane

- is waiting to take them back to civilization. Under a clear, starlit night sky they estimate that it must be just about the time that the monks are pasting the final printed names into their holy books. This is when they notice that “overhead, without any fuss, the stars were going out” (summary adapted from (http://en.wikipedia.org/wiki/The_Nine_Billion_Names_of_God)).
14. Italian historian Rolando Minuti seems to believe that the solution to this problem is the availability of new guides (in digital form) specifically meant for the use of would-be historians of the newer generations (Minuti).
 15. This utilitarian approach is also recognizable in the logical way in which the internet has organized its information; that is, in the tree-like structure, via files and folders (formerly, ‘directories’) that tend to make it very difficult to even think that knowledge could be organized in a different way.
 16. On the early days of the company, see *Vise*; *Vise and Malseed*.
 17. Roger Chartier agrees with this view, but he also emphasizes the implications that come with the fact that we read all texts, whatever their kind and function, through a screen: “*Est ainsi rompue la relation qui, dans toute les cultures écrites [...] liait étroitement des objets, des genres et des usages. C’est cette relation qui organise les différences immédiatement perçues entre les différents types de publications imprimées et les attentes de leurs lecteurs, guidés dans l’ordre ou le désordre des discours par la matérialité même des objets qui les portent [...] Dans le monde de la textualité numérique, les discours ne sont plus inscrits dans les objets, qui permettent de les classer, hiérarchiser et reconnaître dans leur identité propre*” (Chartier). It is a fact that, before internet, no information could be acquired without some form of physical support.
 18. The internet, as we know, is a global phenomenon. Twenty-seven per cent of the world’s population (1.8 billion people) use it every day (2011), while the rest of the world population would very much like to join in.
 19. Encyclopaedia Britannica’s website, www.britannica.com, originated in Britain but is now based in the United States.
 20. Lanier describes himself as “a philosopher and a computer scientist who spent his career pushing the transformative power of modern technology to its limits” (Lanier, back cover).
 21. Very much in line with Keen and Lanier is another internet specialist, Nicholas Carr, who stressed the anthropological and neurological effects of the internet on the human race, whose members are enhancing their faculties for quick reaction to multimedia impulses while shedding their abilities for individual and abstract reasoning – a process that is (was) deep and slow, as well as sequential and consequential (Carr 2008, 2010).
 22. This is the way New Tech Network, a group based in Napa, California, describes their ‘difference’: “[S]tudents decide how to allocate their time, team roles, and how to collaborate, and even have a voice in campus leadership and policy. Traditional management tools such as hall passes and class bells are a thing of the past. [...] All classrooms have a one-to-one computing ratio. With access to Web-enabled computers, every student becomes a

self-directed learner who no longer needs to rely primarily on teachers or textbooks for knowledge” (www.newtechfoundation.org/about.html). This site was last accessed in 2010. The new site is slightly different in wording, but does not substantially change the way the New Tech Network pictures itself. See www.newtechnetwork.org/, last accessed 20 February 2013.

23. Godin’s autodescription is in <http://sethgodin.typepad.com/about.html>.
24. A difference must be made between the first Web cohort, that born in the decade from 1975 to 1985, and the second one, born between 1985 and 1995 (and later). Members of the first cohort still enjoyed access to both the pre-Web and the post-Web facilities. They also profited from more traditional schools of history that, at least in the Western world, still included methodology and historiography at the pre-doctoral level. Before proceeding further, one had to master the state of the art and to know how her predecessors had arrived there. Conversely, members of the second cohort were educated in a higher education system that standardized knowledge by limiting the number of pages to be read or words to be written and, in doing so, encouraged students simply to borrow from the Web what others (such as Wikipedia contributors) had already prepared for them. Of course, for both cohorts much of what they received was determined by faculty’s level.

Bibliography

- Anderson, Chris. *Free: The Future of a Radical Price*. New York: Hyperion. 2009.
- Auletta, Ken. *Googled: The End of the World as We Know It*. New York: Virgin Books. 2010.
- Beaulieu, Carole. “Bye-by école du 17^e siècle,” *L’Actualité*, XXXIV, 18 (15 November 2009), 9.
- Borges, Jorge Luis. “La biblioteca de Babel,” in *El Jardín de senderos que se bifurcan*, Buenos Aires: Editorial Sur. 1941.
- Brandão, José Antônio. “No Grants, No Travel, No Excuses: Researching and Writing Early North American History in the Digital Age,” *RiMe. Rivista dell’Istituto di Storia dell’Europa Mediterranea*, 4 (2010) 663-672. http://rime.to.cnr.it/rivista/n4/2010/rivista_2010/RiMe_04_2010.pdf.
- Carr, Nicholas G. “Is Google Making Us Stupid? What the Internet is Doing to Our Brain,” *Atlantic Magazine* (July/August 2008). www.theatlantic.com/doc/200807/google, last accessed 20 February 2013.
- Carr, Nicholas G. *The Shallows: What the Internet is Doing to Our Brains*, New York: W.W. Norton & Company. 2010.
- Carroll, Lewis. *Alice Adventure’s in Wonderland*. 1865. Chapter I.
- Chartier, Roger. “L’avenir numérique du livre,” *Le Monde* (26 October 2009).
- Clossey, Luke. *Salvation and Globalization in the Early Jesuit Missions*. Cambridge, New York: Cambridge University Press. 2008.
- Cuban, Mark. “Free vs Freely Distributed,” *Blog Maverick, The Mark Cuban Weblog* (30 June 2009). <http://blogmaverick.com/2009/06/30/free-vs-freely-distributed/>.
- Curtin, Philip D. *Cross-Cultural Trade in World History*. Cambridge, New York: Cambridge University Press. 1984.
- Darnton, Robert. *The Case for Books: Past, Present, and Future*. New York: PublicAffairs. 2009.

- Day, Elizabeth. "How the Tyranny of Email Keeps us in Touch – but Drives us All Apart / A new book from the Editor of Granta argues that while Texting and Emails Facilitate Easier Communication they also Reduce Human Contact. Is Technology to Blame or Are We," *The Observer* (1 November 2009), 20.
- Freeman, John. *The Tyranny of Email: The Four-Thousand-Year Journey to Your Inbox*. New York: Scribner. 2009.
- Furet, François. "Quantitative History," in Stephen R. Graubard (ed.) "The Historian and the World of the Twentieth Century." Special issue of *Daedalus*, C, 2 (Spring 1971), 151-167.
- Gaudio, Nicole. "Inside the World of Wikipedians, There's Drama, Politics and Love." *USA Today* (27 February 2006).
- Genet, Jean-Philippe and Andrea Zorzi (eds.). *Les historiens et l'informatique: un métier à réinventer*. Rome: École française de Rome. 2011.
- Gilbert, Felix. "Post Scriptum," in Stephen R. Graubard (ed.) "The Historian and the World of the Twentieth Century." Special issue of *Daedalus*, C, 2 (Spring 1971), 520-530.
- Gladwell, Malcolm. "Is Free the Future?" *The New Yorker* (6 & 13 July 2009), 80-84.
- Godin, Seth. "Malcolm is Wrong," *Seth Godin's Blog* (30 June 2009). http://sethgodin.typepad.com/seths_blog/2009/06/malcolm-is-wrong.html.
- Graubard, Stephen R. (ed.). "The Historian and the World of the Twentieth Century." Special issue of *Daedalus*, C, 2 (Spring 1971), 1-532.
- Grégoire, Isabelle. "Génération piton! / Ils sont nés avec Internet. Ils ont grandi dans un monde sans fil... Pour eux, les universités se transforment. Voyage au coeur d'une révolution." *L'Actualité*, XXXIV, 18 (15 November 2009), 28-39.
- Grossman, Lev. "Simon Pulsifer: The Duke of Data." *Time* (26 December 2006).
- Gulli, Cathy. "Meet Mr. Know-it-all: Simon Pulsifer." *Maclean's* (15 August 2006).
- Hanham, Harold. "Clio's Weapons," in Stephen R. Graubard (ed.) "The Historian and the World of the Twentieth Century." Special issue of *Daedalus*, C, 2 (Spring 1971), 509-519.
- Henderson, Jarrett, Marriott, Brandon, Vélez, Karin and Kenneth Mills. "A Round Table on Luke Clossy's *Salvation and Globalization in the Early Jesuit Missions* (Winner of the Wallace K. Ferguson Prize, Canadian Historical Association, 2010)." *Histoire sociale/Social History*, XLV, 90 (November 2012), 393-409.
- Hunt, Tristram, "Online is Fine, but History is Best Hands On," *The Observer* (3 July 2011).
- Keen, Andrew. *The Cult of the Amateur: How Today's Internet is Killing Our Culture*. New York: Doubleday/Currency. 2007.
- Lanier, Jaron. *You Are Not A Gadget: A Manifesto*. London: Allen Lane. 2010.
- Lessig, Lawrence. "Keen's *The Cult of the Amateur*: BRILLIANT!" http://lessig.org/blog/2007/05/keens_the_cult_of_the_amateur.html (31 May 2007).
- Milligan, Ian. "Illusionary Order: Online Databases, Optical Character Recognition, and Canadian History, 1997-2010." *The Canadian Historical Review*, XCIV, 4 (December 2013), 540-569.
- Minuti, Rolando. "Internet e il mestiere di storico. Riflessioni sulle incertezze di una mutazione." *Cromohs*, 6 (2001), 1-75. www.cromohs.unifi.it/6_2001/rminuti.html. French transl. *Internet et le métier d'historien*. Paris: Presses Universitaires de France. 2002.
- Negroponce, Nicholas. *Being Digital*. New York: Knopf, 1995.
- Nunberg, Geoffrey. "Google's Book Search: A Disaster for Scholars." *The Chronicle of Higher Education* (31 August 2009). <http://chronicle.com/article/Googles-Book-Search-A/48245/>.
- Rich, Motoko. "Google Hopes to Open a Trove of Little-Seen Books." *The New York Times* (4 January 2009).
- Roncaglia, Gino. *La quarta rivoluzione: Sei lezioni sul futuro del libro*. Rome, Bari: Editori Laterza. 2010.

- Sabrina (8 June 2007). www.lessig.org/blog/2007/05/keens_the_cult_of_the_amateur.html.
- Schiff, Stacy. "Know It All: Can Wikipedia Conquer Expertise?" *The New Yorker* (31 July 2006), 36-43.
- Shakespeare, William. *As you Like It*, written 1599/1600, first published 1623, Act IV, Scene I (Rosalind).
- Shimo, Alexandra. "Prolific Canadian is king of Wikipedia." *The Globe and Mail* (4 August 2006).
- Stone, Lawrence. "Prosopography," in Stephen R. Graubard (ed.) "The Historian and the World of the Twentieth Century." Special issue of *Daedalus*, C, 2 (Spring 1971), 46-79.
- Stone, Lawrence. "The Revival of Narrative: Reflections on a New Old History." *Past and Present*, 85 (1979), 3-24.
- Thornton, John K. *A Cultural History of the Atlantic World, 1250-1820*. Cambridge, New York: Cambridge University Press. 2012.
- Vise, David A. *The Google Story*. London: Macmillan. 2005; rev. as David Vise and Mark Malseed, *The Google Story: For Google's 10th Birthday*, New York: Delta Trade Paperbacks. 2008.
- Vitali, Stefano. *Passato digitale. Le fonti dello storico nell'era del computer*. Milan: Bruno Mondadori. 2004.

**Changing Models for Textual Editing
in Electronic Publication**

Electronic Textual Criticism

A Challenge to the Editor and to the Publisher

Gábor Kecskeméti

The Hungarian Academy of Sciences (HAS) was originally established in the early nineteenth century for linguistic and literary studies, including laying the foundation of research in Hungarian literary history. Studying national classics of Hungarian literary history has been the Academy's mission ever since, so HAS played and plays an essential role in the research into classic authors' life-works.¹ Scholarly text editions are made of the oeuvres of national classics or compiled from certain periods of literary history on the basis of genre. Most of these critical editions are produced by the Institute for Literary Studies of the HAS, even though faculties of literary history from prominent Hungarian universities contribute to the work. This work has been supervised by a HAS committee created for this purpose: the Textological Committee. The Committee was established in 1960 (Klaniczay) and has been operating ever since without interruption. Its functions include formulation of obligatory norms and methodological recommendations for critical text editions of Hungarian national classics; coordination and supervision of such works in different workshops; approval of plans for text editions; and accreditation of series and individual volumes as scholarly editions. In short: the Textological Committee provides quality assurance for text editions of Hungarian national classics. The efforts of the committee resulted in several hundred volumes of critical text editions in the past decades;² thus, the most important authors in Hungarian literary history are available for study in reliable editions of high standard that are based on carefully considered uniform principles. In the beginning, it was the publishing house of the HAS, Akadémiai Kiadó, which published all critical editions.³ In the 1990s, four other scholarly publishing houses joined in. These critical editions provide the basis for all of the popular text editions that aim for a wider audience. These are published by several other commercial publishing houses.

So it seems that everything is alright with text editions of works of Hungarian literary history: a series of critical editions have been published with reliable text, in uniform structure, with scholarly apparatus, extensive commentaries and explanations. We would have every reason to be satisfied if the objective of the work was the same as the original goal of the Textological Committee: the production of editions in print.

However, in the twenty-first century, we cannot overlook the need for the availability of critical editions in electronic form. And the current methods and technology used in Hungary are only partially suitable for this.

Of course, a digital version can be made of any printed edition, even afterwards. In its simplest form, this is not more than a facsimile stored in a digital image format. The PDF file format became the standard in the last decade, and it is a more appropriate solution than a simple image format, because two layers can be generated in it: the facsimile appears on the user's screen, while beneath it there is a hidden character format text layer in which search operations can be performed. This text layer is usually produced automatically through optical character recognition, and usually it is not proofread even once, so its quality is not satisfying. Nevertheless, it somehow extends the user's possibilities in handling the text.

In the case of books published in the last decade – if we are lucky – there is no need to scan the facsimile or for making a character recognition. The PDF file can be produced from the publisher's file that contains the book's layout, preserving the final step of desktop editing before printing the book. This layout file has unique value: besides having a harmonious, consistent and functional format, it is usually only this file that preserves the final state of the text. Proofreading and printing approval are based on this file; final corrections and additions are usually made in this file only, so this is the only point in the technological process that records the state of the text intended for printing. From this file one can produce a single-layered PDF file that stores the text only in character form. Yet its appearance, i.e. the layout, the formats applied, the lines and pages are identical to the printed book and this text is suitable for search operations without concession.

However, the latest document accepted by the Textological Committee in 2004 calls electronic texts produced this way "digitised editions," and clearly differentiates these from "digital" or "electronic" editions (Debrecezeni and Kecskeméti). This differentiation is absolutely justified and, in my opinion, we cannot aim for less in the twenty-first century than producing critical editions as truly digital editions. A digital edition is designed for the electronic medium from the very start and utilizes all the possibilities of this medium. No paper equivalent of a truly digital edition is conceivable – its "way of life," its philosophy, its principles of handling the text are all designed exclusively for the electronic environment. It offers much more complex possibilities beyond static reading of the body text and its notes. The electronic medium provides special dynamic functions of handling the text, like queries based on any criteria, functions for ordering or displaying the text in special ways. In order to open up the way for these features, the text must be structured

in an orderly fashion. The means for creating these structures are markup languages. The Textological Committee prescribes markup language structure as an obligatory requirement for electronic critical editions.

A truly digital edition cannot be created through mechanical conversion of a printed edition. The layout files produced and preserved by publishing houses are almost inadequate – even in an optimal case – for serving as bases for digital editions. The layout editor's work means a lot of added value in a paper edition, but its primary purpose goes against the clear structure of the text. Layout editors traditionally create harmonious appearance through interventions that are carried out by breaking the structural unity of the text. If there is a need, certain parts of the text receive unique direct formatting; smooth transition from line to line or from page to page is achieved by inserting 'hard' characters or breaking up paragraphs; the unity of spacing that also expresses structure is modified individually; elements that belong together functionally are separated and processed differently, etc. Even if the file before layout editing had some functional structure, this structure disappears from the layout file and the PDF file created after it. (Let us consider: the layout editor may break the functional connection between the location of reference and the footnote, and such structural connections necessarily disappear from a PDF file designed inappropriately.)

Thus, the organized technology necessary for producing digital text editions must be present from the first moment of the work in the text management practice of the specialists who work in the research process.⁴ Such an undertaking cannot expect this kind of professionalism only from the publishing house, who join the work in the publishing phase. In the twenty-first century, even the research stage of a critical edition must be highly organized electronically. In a long-term endeavour that is carefully planned, employs a wider circle of contributors and specialists and aims for producing series, it is inappropriate if scholars enter their texts in general-purpose office word processors the way they can, and apply their own unique systems of notations in the process. I believe, a well-established enterprise for producing critical editions cannot exist today without a tailor-made or very highly customized specialized computer program, which means that some money needs to be invested in programming. This includes training and consultation services for the participants, which would help them make technological decisions in a unified, coherent manner. Such a text-entry application can be integrated into a database management system that records metadata for the text. Optimally, the corpus would be uploaded to a server on a monthly basis, even if parts of it are produced on

separate stand-alone computers. This means that all scholars in the project would have continuous access to the whole corpus. This would provide substantial help already in the research phase in terms of having a clear overview of the material. In the publishing phase, producing all kinds of output from the same material, from a printed book to an indexed database, would be possible. The new challenge for the publishing house is to create body text and apparatus of a specialised text edition from a complex system of databases and markup language structures of the research phase through systematically reducing them by a series of correct conversion steps.

Unfortunately, current textological practice in Hungary is very far from this. In my experience, not only individual professionals but participants of large, collective endeavours, too, use some version of Microsoft Word almost exclusively, disregarding the fact that this general-purpose word-processing application that is optimized mostly for office use does not provide all the functions indispensable for scholarly handling of text. An apparatus for critical text editions would require, for instance, the possibility of handling several different types of notes, while Word only allows two kinds; namely, footnotes and endnotes. Of course, there are several ways and degrees of utilizing the features that this program does provide and this depends on the degree of expertise. It is much better to connect automatically numbered endnotes to the references than typing the currently actual numbers and giving them upper index format manually, and then recording the corresponding notes completely independently in manually numbered paragraphs at the end of the file or in a separate file. If we miss recording the structural connection between the location in the text and the corresponding note, then we give up the possibility of achieving a near-professional state of the edition through automated conversions during a functional processing of the text. The editor of the leading periodical of literary studies in Hungary still receives manuscripts in which the line of thought is interrupted sometimes in the middle of a sentence or even in the middle of a word by some empty lines and a manually centred page number, and the sentence or the word is continued in a next paragraph. What happened was that the author who uses his or her computer as if it was a typewriter reached the bottom of the page at this point on his own display. (See Kecskeméti.)

In the conventions of recording text that result in ways of giving a clear overview of all aspects of content and form, the application of styles should have a special role. Applying paragraph styles is a rare event in Hungarian textological practice, and I have not seen an author's file in the last fifteen years in which the author used character styles for highlighting. In each and every case, standard italics, bold and other such direct formatting were used.

Individual formatting and the application of styles may be equivalent if it is only a matter of 'What will appear on the typist's display?' The moment the file enters the technological line of professional text processing there are potential dangers in individual character formatting. For example, if we achieve functional order or aesthetic unity of the text through paragraph formatting in some later stage, many applications remove individual character formatting the scope of which is more than half of the current paragraph. This way of operating the application is understandable; as it is assumed that earlier formatting, aimed at displaying the whole paragraph without expertise, is no longer needed when we produce a consistent look for the whole text through the more professional way of paragraph formatting. However, in many cases, the formatting that disappears was entered laboriously by a non-expert author and has functional meaning in the structure and interpretation of the text. Restoring such formatting needs subsequent interventions that are cumbersome and have the potential of introducing new errors. Furthermore, individual character formatting will never provide a proper basis for markup conversion. For instance, italics may have many different kinds of functions within the same text (such as highlighting titles, names, foreign words, unique expressions or quotations, or it can serve as typographical clarification, etc.). On the other hand, any number of character styles can have the same appearance, so we can create separate character styles for all the above functions, and these styles can be processed in an automated way and converted into markups. This is a fairly extensive conversion; there is a considerable added intellectual value involved. Essentially, it means transformation of the linear text into a structured database or expert system. But if we have to use Word files as points of departure, the application of styles seems like an appropriate first step on the way.

In international considerations of preparing electronic text editions and their technological realization, the usual subject of discussion is the – formerly SGML-, today XML-based – DTDs of the Text Encoding Initiative and the adaptations of these (Cover and Robinson; Sperberg-McQueen; Driscoll; Durusau; Lavagnino; Barney; in Hungarian: Zsoldos-Demjén). I am afraid that the situation is different in the field of Hungarian textology. The extent of expertise in informatics is such that colleagues who have been working on critical editions for years need to be taught how to enter correct typographic quotation marks, how to enter special characters and how to turn off autocorrecting, which makes entering critical text impossible. Hungarian textology can be proud of the decades of its book series of critical editions. But if we think we can rest content after we have published a printed edition in the belief that now the preservation of the work is ensured, then we are wrong.

We have a single layout file in the publisher's possession, which preserves the final text of the printed edition. Compared to this state of the text all other files kept by those who had worked on the text in the research process are out-of-date, since they represent the input of the process of publishing. From the layout editor's file a static representation of the text was produced in PDF. The chances are good that – because of the quick technological development – in some three years new page setting programs will not be able even to open the publisher's current file or to interpret it properly. Thus, we have not created anything that we may consider a standard electronic basis either for passing on the published text unimpaired or for using it in new ways. We have created something and we have nothing in our hands.

Just like for the whole of national cultural heritage, it is true for the texts of Hungarian literary history that creating their digital versions and publishing them on the internet must be a principal aim of national culture politics. At the same time, it is also an essential disciplinary interest. Making the texts available and searchable is basically the only tool for ensuring equal cultural and scholarly opportunities in the digital world of the twenty-first century. One of the most important challenges the Textological Committee is now facing is to care about the retrospective publication of the digitized versions of existing paper editions and, at the same time, secure that the possibilities offered by digital technology will be considered when future critical editions are prepared and published. I am convinced that in case we have reliable critical editions, we must definitely use them as the base for quality digital publishing. Their conversion needs a content exploration and semantic analysis of the text and the apparatus, the result of which is recorded according to some kind of a syntactic convention. The XML based syntactic recommendations of TEI seem to be the most widely accepted technological choices today in projects that aim at digitally processing big volumes of texts. The XML files can simultaneously ensure the dynamically generated content during the query, the support of advanced query options in the functionally formed field structure, the flexibility of the visual arrangement and formatting, and the long-term conservation and optional future conversion of the text database organized by them on a semantic basis. In my opinion, the real strength of the syntax of TEI lies in the possibility of coding the critical and interpretational apparatus together with the main text. The dynamic text generation done real time opens the way to the documentation of the text history, variants and interpretations as well, which can be easily and freely visualized functionally as needed, with the help of colours, pop-ups, typographic tools due to CSS, PHP, Javascript and further technologies. These technological solutions are absolutely suitable

for serving as a basis for a new kind of cooperation enabling the national community of professionals to join the edition of texts and the development of a new nation-wide homepage for the publication of the texts with primary importance in the literary history of Hungary. Given the fact that all the texts written in Latin are, throughout Europe, evidently included in national literary history, or in a broader sense in the textual part of national cultural heritage, these technological choices open up the possibility of cooperation even for the international community of scholars to contribute to a promising project of philological and, at the same time, digital expertise.

Notes

1. Summary of the nineteenth-century beginnings and the recent series of text editions: Szilágyi.
2. Bibliography of the critical text editions supervised by the Textological Committee can be found on the webpage of the committee (<http://textologia.iti.mta.hu>). This bibliography contains 461 volumes at the present time. There are only 47 items amongst them with a link to an electronic version. In other words, approximately 90 per cent of the critical text editions are available in paper editions only.
3. Akadémiai Kiadó National Company was founded after the socialization of publishing houses and presses, in 1950. The activity of the first three decades of the firm is summarized by Köpeczi. The company had come near to an economical and technological smash before the next change of the political system, by the middle of the 1980s. On this crisis as it concerned the critical text editions, see R. R.; Herman et al. The publishing house has been an incorporated company since 1996, now a co-property of the HAS and Wolters Kluwer of the Netherlands.
4. A great number of significant publications discuss the full technological process of electronic textual editing by now. For example's sake: Finneran; Sutherland; Burnard, O'Brien O'Keefe, and Unsworth; Deegan and Sutherland; and the articles of special issue *Historical Perspectives on Digital Editing* of the periodical *Textual Cultures* (vol. 7, n. 1, 2012), especially Earhart.

Bibliography

- Barney, Brett. "Digital Editing with the TEI Yesterday, Today, and Tomorrow." *Textual Cultures* 7.1 (2012) 29-41. Print.
- Burnard, Lou, Katherine O'Brien O'Keefe and John Unsworth (eds.) *Electronic Textual Editing*. New York: Modern Language Association of America. 2006. Print.

- Cover, Robin C. and Peter M. W. Robinson. "Encoding Textual Criticism." *Computers and the Humanities* 29 (1995) 123-136. Print.
- [Debrecezeni, Attila and Gábor Kecskeméti]. "Alapelvek az irodalmi szövegek tudományos kiadásához" (Principles to scholarly editions of literary texts). *Irodalomtörténet* 85 (2004) 328-330. Print.
- Deegan, Marilyn and Kathryn Sutherland (eds.) *Text Editing, Print and the Digital World*. Farnham, UK: Ashgate. 2009. Print.
- Driscoll, M. J. "Levels of Transcription," in Lou Burnard, Katherine O'Brien O'Keefe and John Unsworth (eds.) *Electronic Textual Editing*. New York: Modern Language Association of America. 2006. 254-261. Print.
- Durusau, Patrick. "Why and How to Document Your Markup Choices," in Lou Burnard, Katherine O'Brien O'Keefe and John Unsworth (eds.) *Electronic Textual Editing*. New York: Modern Language Association of America. 2006. 299-309. Print.
- Earhart, Amy E. "The Digital Edition and the Digital Humanities." *Textual Cultures* 7.1 (2012) 18-28. Print.
- Finneran, Richard J. (ed.) *The Literary Text in the Digital Age*. Ann Arbor: U of Michigan P. 1996. Print.
- Herman József et al. "Merre tartson az akadémiai könyvkiadás?" (Where is the way to go with the book-making of the academy?). *Magyar Tudomány* 93 (1986) 467-478. Print.
- Kecskeméti, Gábor. "Filológiai problémák a magyarországi irodalomtudományi kutatásban és a számítógépes kézirat-előkészítésben" (Philological problems in Hungarian literary research and in computer-prepared manuscripts), in Emil Hargittay (ed.) *Bevezetés a régi magyarországi irodalom filológiájába* (Introduction into the philology of old literature in Hungary). 3rd ed. Budapest: Universitas, 2003. 132-146. Print. Klaniczay, Tibor. "A textológiai munka problémái" (Problems of the textological work). *A Magyar Tudományos Akadémia I. Osztályának Közleményei* 19 (1962) 339-349. Print.
- Köpeczi, Béla. "Az akadémiai kiadás utolsó harminc éve" (The last thirty years of the publishing activity of the academy). *A Magyar Tudományos Akadémia I. Osztályának Közleményei* 32 (1981) 45-51. Print.
- Lavagnino, John. "When Not to Use TEI," in Lou Burnard, Katherine O'Brien O'Keefe and John Unsworth (eds.) *Electronic Textual Editing*. New York: Modern Language Association of America. 2006. 334-338. Print.
- R. R. "Az elnökség napirendjén: az akadémiai könyv- és folyóiratkiadás időszerű kérdései" (Some actual questions in relation to publishing books and periodicals of the academy on the agenda of the officers). *Magyar Tudomány* 93 (1986) 230-233. Print.
- Sperberg-McQueen, Michael C. "Textual Criticism and the Text Encoding Initiative," in Richard J. Finneran (ed.) *The Literary Text in the Digital Age*. Ann Arbor: U of Michigan P. 1996. 37-62. Print.
- Sutherland, Kathryn (ed.) *Electronic Text: Investigations in Method and History*. Oxford: Clarendon P. 1997. Print.
- Szilágyi, Ferenc. "Akadémiánk 'Nemzeti Könyvtár'-terve 1850-ből s jelenlegi szövegkiadásunk" (The plan for a series with the title National Library of the Hungarian Academy of Sciences in 1850, and our recent text editions). *Magyar Tudomány* 88 (1981) 186-194. Print.
- Zsoldos-Demjén Anetta. "Digitális szövegrögzítés TEI irányelvek alapján" (Digital text markup on the basis of TEI principles), in Gábor Kecskeméti and Réka Tasi (eds.) *Filológia és textológia a régi magyar irodalomban: Tudományos konferencia, Miskolc, 2011. május 25-28.* (Philology and textology in the old Hungarian literature). Miskolc: Miskolci Egyetem BTK Magyar Nyelv- és Irodalomtudományi Intézet. 2012. 491-503. Print.

Computer-assisted Scholarly Editing of Manuscript Sources

Andrea Bozzi

Introduction

Over the last months, various reports have appeared in the media regarding the creation of large digital libraries implemented by public institutions and private companies working in collaboration. More recently, many of the problems connected with copyright seem to have been resolved or have found feasible solutions. In the meantime, new tools for the dissemination and reading of *e-books* have started to appear on the market.

A new form of publication appears to be emerging: on the one hand this important issue tends to be oversimplified in the media, especially when the advent of these new technologies is presented as the first step towards the disappearance of the traditional paper book; on the other hand, the general reading public is not informed on another aspect of the problem, which probably does not have the same importance for them as it has for the specialists. What changes in study methods should scholars dealing with linguistic, philological and literary texts expect? In this contribution, I will try to show how it is possible to keep alive (and possibly improve) fields of study that have a long tradition and which are apparently threatened by the advent of radical transformations in the creation and dissemination of information.

My intervention will concentrate on the specific aspect of textual criticism. I realize this is a discipline that could be defined as being 'very exclusive,' as the scholars are not numerically equivalent to the community of people working in other Humanities disciplines; for example historians, philosophers, or those dealing with the history and criticism of literature. However, if we consider that textual criticism covers a very large period (Ancient, Medieval and Modern times) and many languages, there is also an increase in the population of specialists. If we add the specialists working on the philology of ancient printed texts and those studying the manuscripts of modern and contemporary authors, we realize that we are dealing with a very respectable community, which deserves to be helped to access the new digital world. An important element that needs to be considered is the inexplicable reluctance on the part of this community of scholars to

accept the use of new technology. For this reason, it is our responsibility not only to enhance the study and development of suitable technology, but also to promote appropriate and effective actions able to convince sceptics that these technological tools are crucial to accompany the disciplines of philology and textual criticism into the new field of Digital Humanities.

The Model

I have called the model I am using for this community of textual critics 'system for digital philology,' which can be summarized as follows. Firstly, it is necessary to adopt shared standards at the international level for text labelling, so that the elements contained in the texts (format values, style, capitals, onomastics, etc.) can be recognized by internet browsers or can interoperate with other data available on the Web. Even the annotations added to the text (as, for example, the lemma for each wordform or their part of speech, the indication of variants, etc.), produced either manually or using NLP¹ tools, must respect specific standards. This element currently has great importance since the standards have a crucial function for the creation of research infrastructures enabling the members of the same community of users and of other communities operating in the human sciences to share the resources made available on the Web.

Furthermore, a digital philology application should be provided with a series of fundamental research tools within a modular architecture. The system is developed by adding to the core of basic tools increasingly specialized ones designed to meet particular needs. This developmental strategy can customize the system to the many philology types and, at the same time, can be specialized for each one. Some of the basic tools are represented, for example, by an indexer, a concordance program, a statistic analyzer, an image enhancement system, etc.

The digital document, eventually supplied with additional information and annotated standard markup systems, is produced on a server so that it can be queried via the Web. In this phase, the administration of accesses is important because data visibility could initially be limited only to some members of the community (collaborative editions) and later become public (final publication). XML encoding of the documents and annotations allows for a second form of dissemination, that of traditional paper, at any time by adopting appropriate XML data transformers, also in view of the print-on-demand or ebook-controlled distribution. I shall not focus here on the standards issue; instead, let us browse through the basic tools for

scholarly editing of digital documents, bearing in mind the prospect of a digital philological system.²

We are convinced that an appropriately designed scholarly editing module of the computational philological system meeting the specific needs of classical and medieval philology, especially in the field of Romance languages, will also respond to other types of philology; for example, Greek and late Aegyptian papyrology, epigraphy, palaeo-Slavonic philology, philology of ancient printed texts, and – with appropriate adaptation – the philology of texts of modern and contemporary authors. The basic criteria are the following:

- linear transcription of a single source;
- positive apparatus for recording the variants of the collated sources;
- specific area of the apparatus for storing the readings selected or proposed by the critical editor;
- automatic generation of the *textus constitutus*;
- automatic generation of the text of all the other reviewed and collated sources;
- computer-assisted assessment of the variants and man-machine user interface to hypothesize stemmata resulting from the apparatus data.

From the Model to the Textual Criticism Web Application

The application that I am presenting here is aimed at achieving these ambitious targets through an architecture based on interconnected modules. In other words, it works with a nucleus of components for the treatment of both text files and digital image files, which form the core of the system. According to the specific needs, from time to time a number of programs are added both for the management of images (enhancement, segmentation, pattern recognition, etc.) and of text (natural language processing, information extraction, data mining, electronic editing, etc.).

1. *Textual levels.*

The main principle is the structural subdivision of the textual data to be represented in a digital system: the data can include images (of the document containing the text), text (namely, the diplomatic or interpretative transcription of the text contained in the images), or extra- and paratextual information. The first group (extratextual information) comprises the references of the text in the page (page number or manuscript folio, number of paragraph, comma, section, running titles in the case of a dictionary

or volume of encyclopaedia, etc.). The last one (paratextual information) includes annotations, apparatuses and bibliography.

This structure, consistent with the encoding and markup systems adopted at an international level for textual data, and referred not only to texts (see, for example, the markup language of the *Text Encoding Initiative-TEI*,³ now available also in xml version), places the different types on different levels: for instance, a number representing a date within a work or a critical note studied and edited by a philologist are elements belonging to information levels different from the text proper.

Therefore, the application takes these different levels into account, provided they have been appropriately marked, so that they may be exploited by the user who is performing search operations. The user will first need to select the option activating the function expressed by these distinctive elements, which ensure better results than those that would be obtained without these distinctions.

A typical example is that of indexation, which can, for example, produce a separate list of words that are read in current titles, as well as lists of linguistic forms of the text proper, as long as the distinction between the two classes of data (current text and title) has been marked appropriately. Otherwise, we would obtain a single, all-comprehensive list of inflected forms with no possible distinction between the two sets.

These aspects have been widely analyzed by the research community, which for many years now has been involved in problems of markup and language encoding; however, it is necessary to consider the level of distinction that must be attained in order to maintain a substantial balance between allocated resources and expected results. In fact, the problem of markup between the information levels that a written text can contain is strictly connected with the assessment and personal requirements of the user who is performing the electronic processing of a text. In other words, there is no universally valid criterion according to which all the theoretically feasible levels of a text should be identified and featured. As previously said, these depend on the sensitivity, purpose and depth of analysis of the user who studies, prepares and eventually produces a text on the Web.

On the other hand, the designer of a new generation information system for text processing should provide for representing and using any textual, extratextual or paratextual element designed to achieve the results expected.

In order to clarify this aspect of the problem, let us look at an example relating to the electronic processing of the corpus of Ancient Latin Grammars,⁴ making it assume a general value. A typical feature of this archive

is represented by the large number of citations of works by Latin authors commented on by the Grammarians from the point of view of linguistic usage on the basis of which they establish normative principles. The citation phenomena are very interesting and numerous⁵: ancient titles and passages of works to be taken as models are explicitly quoted by the grammarians. In this case, it is necessary for researchers of history of the Latin language and grammar to recognize that the titles, texts and names of the authors (e.g. Cicero, Quintilianus, etc.) quoted by the ancient grammarians should be classified and encoded in categories different from the ones used for the texts written by the quoting authors (e.g. Donatus and Priscianus).

A philologically oriented application should allow us to distinguish these intertextual phenomena using them during the phases of indexation and running of the search options. A query slot, for example, should allow us to:

- indicate whether the string of characters to be searched in the archive belongs to an entire word, or whether it forms the initial, central, or final part of the word;
- indicate whether the search should be made on the entire texts or on one of the levels into which the text has been divided (title of the works, citations of works in prose, citations of works in poetry);
- indicate whether to assign a distinctive value to the accents/breathing spirits or capital letters;
- indicate whether a string of characters should be searched in co-occurrence or in alternative to a second string of characters, activating specific Boolean inclusion or exclusion operators;
- indicate which interval of words should be considered as maximum threshold in the search of two strings of characters with Boolean operators.

We could summarize this section as follows: on the one hand, there is not a universal criterion establishing which levels, depending on subjective competence and evaluation, are theoretically present in a text; on the other hand, a specialized system should allow for categorization of these levels, adoption of standardized markup tools, and the possibility to use them by processing modules aimed at producing multiple results and allowing very detailed queries.

2. *Critical apparatus and annotations.*

Particular attention should be given to the treatment of critical annotations associated with a text. Annotations represent a particular problem as they contain different types of data. In philological works studying texts that

have been handed down by several witnesses, textual readings (variants) are often an alternative to those referred to by a source that, for various reasons, has been considered as the basis of comparison with all the others. The class of variants is followed by another element, represented by the name, generally an abbreviation, specifying the source from which the single variants have been extracted. Together with the indication of the source, there is often an explanation of the reasons according to which the critical editor accepts or rejects a variant and the comments that are useful to understanding any individual conjectures.

I do not intend to describe other details that could be useful here; for example, in the field of genetic criticism to implement technological components for the preparation of electronic editions of works, manuscript or printed, where the author has intervened at different times, introducing marginal notes, interlinear corrections, eliminations.⁶

Let me stress that a philological application will meet the requirements described above, making available a series of modules able to use suitably identified and encoded levels of text according to the international standards of the sector. Furthermore, the different processing components included in a modular structure will either interact or work separately, one from the other.

A typical case is represented by the lemmatization module for Latin,⁷ now ready to be included among the services offered by our philological application. Another case is formed by the module organizing information on the critical apparatus, now in the implementation phase. Activation of the module will make it possible to associate, for each element of the text (single words or entire periods), the variants transmitted by the collated sources.

3. *Collaborative scholarly editing on the Web.*⁸

A particular problem concerns the possibility for users of our philological application to realize collaborative projects in which the competences can be shared to achieve a common result.⁹ In an early phase of development, the technological components will be focused on encouraging shared annotations of images and texts aimed at the publication of critical editions. However, this sector could also be flanked by other collaborative activities, concerning not only the editorial component, but also the joint production of digital contents by means of computational or manual tools analyzing the system from a linguistic and philological point of view. This is the case of semantic or syntactic information added to the texts to enrich the possibilities of interoperability with the data made available by other communities of scholars.

The following example (still a hypothesis of work, not yet analyzed in detail) shows how to represent this need, and which solutions could be adopted.

Let us suppose that a community studying the works of a seventeenth-century scientist, for example Galileo Galilei, is working on “primary information” (manuscripts, printed works, texts of the national edition, drawings, etc.) available in digital format; each researcher belonging to the community can associate “secondary information” (critical notes, studies, bibliographies, etc.) of his own competence. The information system first checks whether the semantic or, better, the ontological classes containing the new information have been assigned correctly and, after verifying authorization of the users (only reading, only writing, or both), makes the information available (and shareable) to the entire community. An annotation that has not been classified as such will not be associated with the relative passage and will not be included in the list of all the other annotations eventually assigned by the other members of the community. It is up to the scientific head, represented by only one person or by a committee of various members, to validate the interventions and to allow their publication on the Web, either public or restricted to the participants.

In other words, our philological application allows the same community of scholars to share large amounts of consistent information also distributed on different servers, using the logical structure according to which the information has been organized. Therefore, a search function launched by a member of the community activates a navigation process among the data made available to the others, to everybody’s advantage.¹⁰

So far, and with only rare exceptions, critical editors have shown scarce propensity towards collaborative editorial work, as the phases of interpretation of the text and evaluation of errors and variants are strictly connected with the knowledge and sensitivity of the single editor. Exchanges of opinions have always existed, in particular for more demanding projects set up by different groups of competence and work. However, these experiences do not correspond to an effective collaborative work made possible only by a Web structure and application.

Experimentations for the international community of the Galilei scholars were started in the autumn of 2009, on the occasion of the celebrations for the 400 years since the first astronomical use of the telescope by the great Pisan scientist. Other joint initiatives will be the project concerning the works of Gerolamo Cardano and, it is hoped, the project relating to Antonio Vallisneri.

As already mentioned, the system prefers editorial works relating to the history of scientific thought, but experiments will also be performed on literary works in poetry and prose. In collaboration with the Società Dantesca Italiana, for example, tests are being done in an attempt to import in the application data already available in digital format so that editions, images, comments and indexes can be consulted organically, overcoming the limits of paper editions, on the one hand, and unreliable ones already available on the Web, on the other.

4. *The application as a tool for computational philology.*¹¹

On the basis of the experiences gained by the creation of experimental programs at Istituto di Linguistica Computazionale – Consiglio Nazionale Ricerche (ILC-CNR) within the framework of national and international research projects,¹² a number of specialized electronic publishing components on the Web are being designed and will shortly be integrated in the application as additional modules. Let us now examine the most important principles at this stage of development:

- the principle according to which a text to be included in the system should correspond to the text contained in the digital image of the source, which can be associated with and shown together with the text;
- Bédier's principle, according to which the best text is the one considered as such for ecdotic reasons: the scholar compares it with the readings transmitted by the other witnesses;¹³
- the principle according to which – in the early phase of data storage – even the venial errors are recorded in the apparatus; namely, the forms (single words or expressions) that will not contribute to shaping the history of tradition and the relation between manuscripts and that, for this reason, are not considered useful to the *constitutio textus*;¹⁴
- the principle regarding ancient printed documents, according to which even the typos should be recorded at an early stage, and then eliminated at a later stage of the editorial process, in the same way as the venial errors of the copyists;
- the principle that all marginal notes, both of manuscripts and ancient printed texts, are encoded as paratextual information so that the linguistic forms contained in the notes can be indexed separately, but connected to the sentence (or the single word) they are or can be referred to;
- the principle that – wherever possible – an automatic or semi-automatic link is created between the words of the transcribed text and the corresponding words of the digital image in which they appear.¹⁵ The level of

usefulness of this tool is inversely proportional to the clarity of reading of the text on the image. Furthermore, the tool becomes absolutely essential when the material support is highly damaged (e.g. rough copies of letters obtained by carbon copy and realized by means of mechanical typing systems; papyrus fragments; all those cases in which there has been partial evanescence of the ink).

The prototypes of the philological application so far implemented at ILC, some characterized by stand-alone applications (BAMBI, DIPHILOS), others Web-based (for example FAD),¹⁶ as well as the tests performed on medieval manuscript sources, ancient printed documents, papyrus fragments and other textual documents of different periods, can now share strategies, methods, and standards for the new philological application, which is an integrated information system on the Web.

The advantage of this development strategy is also in the fact that this application becomes the scholarly editing component of the even vaster system for text management, but it can also work independently. In both cases, we are provided with a tool that is simultaneously a book-support and a digital scriptorium for consultation, study and source editing. The user is provided not only with a magnifying glass to read the documents more easily and a writing tool to transcribe and annotate the text; he or she is also able to select an option from a menu in order to create indexes and apparatuses, connect one document with another, and is assisted in the (paper or web) preparation and publication of a critical edition.

As far as philological research is concerned, it allows the user to include annotations within the text and the images, using a simple and very easy-to-use selector. In fact, within the selected part of text or image, a double annotation can be inserted in a specific box.

The first type of annotation consists of a linguistic, stylistic, or other type of evaluation note of the text during the study phase. Moreover, the user who has established an ontological classification (in other words, a semantic-conceptual classification) of the documents studied, can use the same classification (appearing in a pop-up menu) to annotate parts of the text and/or of the images.¹⁷ This function is extremely useful for retrieving all the portions dealing with the same subject not only in a specific text, but also in a large corpus of texts. The method is now used in a project for the management of documents concerning archaeological excavations in Egypt (the Ippolito Rosellini project); a specific ontology is being studied for the management of a digital archive of the manuscripts of the great Genevan linguist Ferdinand de Saussure.¹⁸

The second type of annotation makes the management of variant readings possible. Let me explain this aspect of the problem in further detail. A wide range of interesting developments can be predicted for the near future, even if a long experimental phase in close contact with the critical editors will be necessary. The indexation program allows us to:

- access the alphabetical index of wordforms occurring in the transcription file. When one item is selected, all the relative variants and annotations are displayed;
- access the alphabetical index of variants (when represented by more than one word, only the first one is indexed). When one item is selected, all the relative words linked to the variant are displayed;
- automatically produce the text of all the witnesses and the one proposed by the critical editor, merging only the text transcribed with the apparatus information.

The information recorded in the critical apparatus is linked both to the text and to the images: in this way, the scholar (and, generally speaking, any user) is always allowed to browse through the digital sources.

However, if the computational system is supplied with a second annotation level of the apparatus variants, the user is likely to receive some hypotheses concerning the distribution of the witnesses over time and space, and therefore can be assisted in the creation of a *stemma codicum*.

For this reason, the apparatus module will also be developed in a new perspective, very interesting from both a methodological and experimental point of view. To allow the system to make a stemmatic hypothesis, a procedure was implemented that asks the user not only to store information in the apparatus area, but also to compare the readings and evaluate their mutual relationships. In this way, we avoid the variants being detected and classified by specific algorithms used in other European Research Centres. At the same time, we will not cluster the variations on the basis of statistical methods.¹⁹ This system is more complex than the one used for variants storage in the apparatus module. It is designed as follows:

- First phase (typology and variants weight). When at least one of the manuscripts taken into consideration differs from the others, the scholar should:
- a) indicate the readings of all of the manuscripts comparing them in pairs, and indicating the correspondence of each manuscript with all the others. In this way, the system records all the correspondences between manuscripts;
 - b) compile a “typology of variants,” which looks like a table of different kinds of variants. The specialist is given the possibility to weigh these

divergences according to a scale of values that he or she has established. In the example used for validation of our system, this classification ranges from a maximum value of dissimilarity of +5 to a maximum value of similarity of -5.

- c) choose the relationship existing between each pair of variants, from a previously compiled list, which is the most suitable on the basis of the textual tradition and the kind of text that is being edited. Recording of the variants and attribution to each pair of the type of existing relationship is made possible by a dialogue box, which appears as many times as is necessary in order to complete all the pairs of variants. Only one example: the numerical value (-5) is given by the presence of the same mistaken reading transmitted by two codes (*conjunctive errors*) or by two manuscripts having the same linguistic expression, which the others have missed.

The second phase of this module is represented by a man-machine interaction system aimed at showing all the possible relationships between the codes, according to the apparatus information. The summations yielded by the system, on the basis of the numerical values attributed to the variants indicated by the philologist in the apparatus, are interpreted by the MDS (multidimensional scaling) algorithm, which places the collated codes in a three-dimensional space. The distances between the codes are proportional to the summations calculated on the apparatus values. The information contained in the tables can be accessed through the point-and-click function, while the information concerning each single manuscript can be obtained by simply clicking on one of the spheres. This method also allows us to make queries. By selecting the shadow of two objects a database query is performed and the result is the annotation relative to pairs of manuscripts, making it possible to see how each variant appears in both texts and what type of relationship exists. The manuscripts selected are graphically united by a line between them, which displays the dissimilarity value of each manuscript.

The last step of the stemmatic module is represented by the *stemma codicum* construction. The graphical representation of relations between the codes obtained using the Multidimensional Scaling Algorithm suggests to examine the possibility of obtaining a graphical representation of dependencies between codes. The same information contained in the apparatus, with the addition of specific algorithm rules allows the system to:

- make classifications;
- establish dependencies consistent with the classifications;

- establish a dialogue with the user who can confirm or refuse the proposals for classification or dependence.

The algorithm has been implemented so that every time the computer finds a pair of manuscripts with minimum distance (very similar manuscripts), these are placed in relation to each other. The computer proposes all the relations between the two manuscripts compatible with the information contained in the critical apparatus and in the *conspectus codicum* (for example, it will consider the possibility of the two codes being *descripti* only if this is compatible with the relative dates). In the absence of information (for example, where chronological indications are missing), the user is asked what type of dependence may exist between them (whether they are *descripti* documents or derive from the same *interpositus*). According to the philologist's answer, the proximity table is updated in a number of different ways. If the philologist suggests the existence of an *interpositus*, the graphics are automatically updated, inserting a new line that links the two close manuscripts to the same *interpositus*.

Therefore, the two provisional *stemmata*

archetype – ms. A – B *descriptus*

and

archetype – ms. B – A *descriptus*

(where the two manuscripts are one the *descriptus* of the other) are cancelled. Otherwise, if the philologist considers it possible for the codes to be *descripti*, the system will maintain active and visible on the graph only the relations in question. In order to have a clear, simple view of the structure of the relationships, which may become increasingly complex as the number of manuscripts increases, the stemma can be rotated and zoomed. A more articulated decision-making structure is expected in the future. Bear in mind, however, that some dependencies suggested by the system may only be theoretical and can be cancelled by the user on the basis of his/her experience, using the point-and-click interaction technique. In other words, the man-machine interactive system allows the philologist to accept or refuse each dependency proposed by the system.

In the experimental trials so far conducted on medico-pharmaceutical manuscript documents in Medieval Occitan,²⁰ at the end of the process

a proposed stemma identical to the one assumed without the use of a computational tool was obtained.

Conclusions

Let me now provide a brief overview of what has been presented so far and then return to the observations made in the introduction: we can conclude that the digital revolution has affected – and will be doing so even more in the future – even a very technical and specialistic sector like text philology. The possibility of operating on the growing number of digital images of documents made available to public and private libraries affects the ways in which the results of research are published and disseminated. Unfortunately, the reaction of those who feel threatened by this new culture aimed at realizing and publishing works derived from philological research is still very frequent. However, we think this fear should be removed rapidly because the true risk that the philological disciplines are running is to be marginalized if the methods of analysis of the sources and outcome production are not adapted to Web-compatible digital technology. Therefore, our work should be considered not only as a simple tool for the creation of electronic editions using a computer-assisted *digital scriptorium*. Indeed, our approach to the problem of critical editing in the different fields of philological research is more general, in an attempt to offer a new vision of work, not just a range of technological tools.²¹

Confirmation of the validity of this attitude comes indirectly from all those researchers who are asking whether our system is able to import into the workspace all the digitized sources dealing with the same issue, even if stored on different servers. This general request has been made by scholars who intend to carry out research on the manuscript documents of de Saussure as well as scholars who need to find thematically correlated documents in large libraries such as the Ambrosiana in Milan or the Gregoriana in Rome. In this respect, the ontological approach offers considerable advantages in terms of classification of the data and bibliographical records, compared to the often dishomogeneous and inconsistent subject-lists.

The design of such a system with open source tools, avoiding the payment of royalties by present and future users, represents an even more important reason for hoping for a rapid change of behaviour: the current scepticism of critical editors should be replaced – and fortunately this is what seems to be slowly happening – by a more open attitude towards technological development for the continuation of future studies. Finally, the systems

operating on the Web offer the great advantage of creating for specific communities of users (classical, medieval, romance philologists) a research infrastructure with standards, markup languages and software that can be used by everybody. This will certainly make dissemination and sharing of the results on the part of the international scientific community more rapid than it is today. If we succeed in bringing this process of updating of the technology tools to an advanced and internationally acknowledged stage (which is taking place nowadays, albeit slowly), a second and even better result will have been achieved. Furthermore, we shall have contributed to doing away with the scepticism mentioned in the introduction, and also to training a new generation of scholars able to perpetrate a philological method consolidated over the centuries thanks to the renewal of its technological tools. Once the use of information technology becomes increasingly user-friendly (an event that is highly likely to occur), this process will be facilitated and, to a certain extent, will become almost natural.

Notes

1. By NLP we intend Natural Language Processing Systems, which, thanks to the use of Computational Linguistics programs (morphological and morphosyntactic analyzers, natural language parsers, extractors of meaning from a text, etc.), contribute to enhancing the value of information of a text, thus increasing the possibility of responding to complex queries. The standard value is represented by the fact that the annotations introduced by NLP systems follow criteria that are shared at an international level, making it possible for archives and corpora implemented at different sites to interoperate.
2. Very interesting and updated information about digital tools for literary studies and philological activities can be found in the Huygens Instituut KNAW (a research institute for text edition and textual scholarship of the Royal Netherlands Academy of Arts and Sciences): www.e-laborate.nl/en/.
3. See www.tei-c.org/Guidelines/P5.
4. For a detailed description of the project and results, see A. Bozzi, *Aspetti e problemi di spoglio elettronico di un archivio testuale: il caso dei Grammatici Latini antichi*, in "Euphrosyne", 31, 2003, pp. 533-550.
5. See, e.g. D. Bamman, G. Crane, *The Logic and Discovery of Textual Allusion*, in LREC 2008 Proceedings (Workshop W22 on *LaTeCH-Language Technology for Cultural Heritage Data*), Marrakech, 2008, pp 1-8 (available online at www.lrec-conf.org/proceedings/lrec2008/workshops/W22_Proceedings.pdf). See also P. Mastandrea, *Gli archivi elettronici di Musisque deoque. Ricerca intertestuale e cernita fra varianti antiche (con qualche ripensamento*

- sulla tradizione indiretta dei poeti latini*), in L. Zurli, P. Mastandrea (eds.), *Poesia latina. Nuova e-filologia. Opportunità per l'editore e per l'interprete*, Roma, Herder, 2009, pp. 41-72 (ISBN 9788889670477).
6. The issue of digital philology in the field of genetic criticism now boasts a very large literature and many experiments: some interesting examples of this are represented by the HyperNietzsche project (see P. D'Iorio, *HyperNietzsche*, Presses Universitaires de France, Paris 2000) and by the Samuel Beckett genetic edition project (see www.beckettarchive.org/introduction.jsp). From a more general point of view, see P. Delany, G.P. Landow, eds., *Hypermedia and Literary Studies*, MIT Press, Cambridge 1991 and, more recently, P.M. de Biasi, *Génétique des textes*, Biblis-CNRS editions, Paris, 2011. For other important projects see, for example, D. van Hulle, *Compositional variants in modern manuscripts*, in "Linguistica Computazionale" (XX-XXI) pp. 513-27, Istituti Editoriali e Poligrafici Internazionali, Pisa-Roma 2004; H.W. Gabler, *Computer-aided critical edition of Ulysses*, in "Bulletin of the Association for Literary and Linguistic Computing", VIII (1980), pp. 232-48.
 7. The system was originally designed and developed to produce the lexical index of the Latin Grammarians at the Institute for Computational Linguistics of the CNR in collaboration with the Department of Linguistics and Classics of Turin University. In recent years, the system has undergone considerable transformations consistent with the development of technologies. The version compatible with the web, titled LemLat, has just been released. A provisional, purely demonstrative version can be consulted and used at the address <http://www.ilc.cnr.it/lemlat/>. For further information, see the section "Documents" on the site as well as A. Bozzi, G. Cappelli, *A Project for Latin Lexicography: 2. A Latin Morphological Analyzer*, in "Computers and the Humanities", 24 (5-6), Kluwer Academic Publishers, Dordrecht, 1990, pp. 421-426.
 8. A. Babeu, *Rome Wasn't Digitized in a Day: Building a Cyberinfrastructure for Digital Classicists*, CLIR Publications, Washington (DC), 2011, pp. 307 (available on-line at www.clir.org/pubs/reports/pub150), ISBN 978-1-932326-38-3, offers an updated overview on the international projects related to collaborative scholarly editing.
 9. See G. Crane, B. Seales, M. Terras, *Cyberinfrastructure for Classical Philology*, in "Digital Humanities Quarterly", 3 (1), 2009.
 10. Concerning the advantages represented by the conceptual or ontologically based information retrieval systems, see N. Ruimy, S. Piccini, E. Giovannetti, *Les Outils Informatiques au Service de la Terminologie Saussurienne*, in "Congrès Mondial de Linguistique Française (Lyon, 2012)", www.ilc.cnr.it/saussure_prg/papers/2012_CMLF.pdf.
 11. The distinction between electronic editing and computational philology is examined in A. Bozzi, *Edizione elettronica dei testi e filologia computazionale*, in Stussi A., *Fondamenti di critica testuale*, Il Mulino, Bologna, 2006, pp. 207-232. A more detailed study of the problems connected with the critical

- edition of a text carried out by means of information systems appears in a note written by Paolo Chiesa regarding the edition of the Dante Alighieri's *Monarchia*: see P. Chiesa, *L'edizione critica elettronica della Monarchia: La filologia informatica alla prova dei fatti*, in "Rivista di studi danteschi", VII, 2, 2007, pp. 325-354.
12. The first experiment of computational philology was realized for the European project BAMBI, see A. Bozzi. (ed.), *Better Access to Manuscripts and Browsing of Images. Aims and Results of an European Research Project in the Field of Digital Libraries (BAMBI LIB-3114)*, CLUEB, Bologna, 1997. Further methodological information can be found in A. Bozzi, *New trends in Philology: A Computational Application for Textual Criticism*, in Zampolli A., Cignoni L. (eds.), "Linguistica Computazionale", XVI-XVII Special Issue, Istituti Editoriali e Poligrafici Internazionali, Pisa-Roma, 2003, pp. 47-77. The description of an application for classical philology and, in particular, for Greek papyrology, can be found in A. Bozzi, *Digital Documents and Computational Philology: The Digital Philology System (Diphilos)*, in Marco Veneziani (ed.), *Informatica e Scienze Umane. Mezzo Secolo di Studi e Ricerche*, Firenze, Leo S. Olschki Editore, 2003, pp. 175-201.
 13. J. Bédier, *La tradition manuscrite du Lai de l'Hombre: réflexions sur l'art d'éditer les anciens textes*, Romania, LIV, pp. 161-196 et 321-356.
 14. This information is useful as the program for generation of the indexes, which connects text and apparatus data, allows us to obtain a complete picture of the situation, even during the phases of production of the edition - and not only *a posteriori* - of the linguistic coating of each witness. Following a more careful study made possible by consultation of the indexes, errors could help identify the geographical and cultural environment in which a copyist worked. The information relative to the venial errors can eventually be discarded at a later stage.
 15. *Multitext Homer* is one of the leading projects in this field. See C. Dué (ed.), *Recapturing a Homeric Legacy: Images and Insights from the Venetus A Manuscript of the Iliad*, in "Hellenic Studies", 35, Center for Hellenic Studies, Harvard University, pp. XVI, 168, 2009, ISBN 9780674032026.
 16. See A. Bozzi, A. Raggioli, *Tecnologia digitale negli Istituti Culturali: Un case study*, in M.M. Mattioda, F. Bruera (ed.), *Itinerari Multimediali Umanistici*, Alessandria, Edizioni dell'Orso, 2003, pp. 23-42.
 17. Named entities are an example of ontological annotations. See, e.g., A. Babeu, D. Bamman, G. Crane, R. Kummer, G. Weaver, *Named Entity Identification and Cyberinfrastructure*, in "Proceedings of the 11th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2007)", pages 259-270, Springer Verlag, pp. 2007-09.
 18. Details about the de Saussure project and the technological components developed so far are available on M.P. Marchese and D. Gambarara (eds.), *Guida per un'edizione digitale dei manoscritti di Ferdinand de Saussure*, Alessandria, Edizioni dell'Orso, 2013.

19. Interesting activities in computer-aided stemmatology are developed at the Sorbonne University in Paris by Jean-Baptiste Camps: see his contribution presented at the 27th International Congress of Romance Linguistics and Philology (Nancy, 2013) available, under registration, at the following web address: www.atilf.fr/cilpr2013/actes.php.
20. Interesting results are described in M.S. Corradini, *Formalisation des variants à des fins computationnelles: vérification de l'hypothèse expérimentale sur un texte occitan*, in D. Billy et A. Buckley (eds.), *Études de langue et de littérature médiévales offertes à Peter T. Ricketts à l'occasion de son 70^{ème} anniversaire*, Turnhout, Brepols, 2005, pp. 355-368.
21. See the pioneeristic vision expressed by P. Robinson, *Redefining Critical Editions*, in G.P. Landow and P. Delany (eds.), *The digital Word: Text-based Computing in the Humanities*, MIT Press, Cambridge/Mass. – London, 1993, pp. 271-291. More updated information in R. Siemens, S. Schreibman (eds.), *A Companion to Digital Literary Studies*, Oxford, Blackwell, 2008 (available on line at www.digitalhumanities.org/companionDLS/).

Electronic Media and Changing Methods in Classics

Bernhard Palme

At first sight, Classical Studies may not seem an obvious field of application for advanced computer technology. Focused on Greek and Latin literature, Classics appears to be concentrated on texts, books and book production in an established and conservative way. However, Classical Philology, Ancient History and a wide range of auxiliary disciplines like epigraphy, papyrology and numismatics have been among the first in the Humanities to systematically develop and use electronic tools.¹ This is partly due to the vivid interest of some leading scholars of the 1980s and 1990s in computer science and partly also due to the favourable circumstances that led a number of influential individuals in the computer business (like David W. Packard) to study Greek or Latin.

Although heavily text- and book-oriented, Classical Studies has undergone a radical change in the last thirty years with the introduction of the extensive use of electronic tools. Nowadays, carrying out research without these tools not only seems outdated but, indeed, impossible. The following essay will address the question of how electronic media have transformed Classical Studies and the way in which Classicists ask questions and approach their topics. Extensive use of electronic media has not only influenced the development of the field, but has changed the entire methodology employed and, currently, it is the catalyst for a revolutionary transformation of the publication culture.

For approximately 5,000 years our culture of writing remained basically unchanged. Papyrus, parchment or paper, sometimes clay or potsherds were used as writing materials for everyday correspondence and bookkeeping as well as for literature, historiography or other expressions of collective memory. Writing was done by hand – occasionally until the beginning of the twentieth century, if typesetting was expensive or difficult because “exotic” characters (e.g. hieroglyphic, cuneiform, Coptic, Greek) had to be used.² Printing has always been expensive, but dissemination without print was difficult, if not impossible. Consequently, the decision to print (or not to print) has often been a kind of a selection process as well. Printing remained expensive even after the PC found its way onto the desks of scholars. At least since the 1990s Classical scholars, like many of their colleagues in the

Humanities, have been expected to do the typesetting and the layout for their manuscripts in order to produce a “camera-ready copy” for photomechanical reproduction. Since the invention of PDF and programmes like Photoshop, publishers also expect scholars to prepare any illustrations too. Fancy words like “desktop publishing” should persuade scholars to accept the fact that they have to perform a second, time-consuming job alongside their scientific research: transforming their manuscript into a book. It is noteworthy that the number of publications has increased significantly since typesetting is done by scholars themselves.

The first step made by the Classics into the digital world was marked by major efforts to make the texts of Greek and Latin authors available in electronic form. The corpus of ancient and medieval texts is extensive in both languages, but this process was almost certainly helped by the fact that the corpus was more or less complete, new discoveries and additions rather marginal and thus the shape of the enterprise calculable. A milestone, in fact almost a revolution, was the first venture (in 1972) to get Greek literature into an electronic version: the *Thesaurus linguae Graecae* (TLG), offering a comprehensive library of Greek literature, from Homer to the fall of Byzantium.³ Similar projects followed with Latin, Coptic and other ancient languages.⁴ Although all Classical texts were available in electronic form since the late 1980s, this did not effect the production and distribution of printed text editions. The electronic versions were, apparently, regarded as an additional research tool rather than a text book for reading. The most important innovation was that the TLG could be fully searched for every single word and, indeed, even a sequence of characters. This was a major achievement for studies in textual criticism as well as comparatistic approaches. Moreover, each search in the TLG automatically revealed a complete list of references from Homer to the late Byzantine authors, thus immediately showing the appearance, frequency or disappearance of a word or phrase. The search program also offered the context, comfortably displayed on the screen and thus encouraged linguistic analyses. A search in the TLG can be done within a few seconds (or minutes, twenty years ago) – a big advantage compared to several days of painstaking and laborious checking of references in dozens of lexica and hundreds of possibly relevant text editions.

Making all classical literature available in a searchable form had two far-reaching consequences. First, philologists and linguists are not only able to do much more research in shorter time, they also dare to face complex scientific questions, which nobody would have approached before the electronic tools were available, because it would have taken an enormous

amount of time to discover and check all relevant sources. Suddenly, it became possible to study, for example, the semantic evolution of specific words with a simple search in the database and by rearranging the hits according to chronological order. The same holds true for various kinds of historical investigations.⁵ The amount of texts available and that are searchable in a reasonable time has increased significantly thanks to the electronic version of classical authors, inscriptions and papyri; this opened up various approaches that had previously seemed out of reach.⁶

Second, for text editors, databases were an equally huge step forward. Identifying tiny fragments of papyrus or parchment was virtually impossible, and generally fragments could be identified only if they contained rare words or phrases – and if the editor had spent half his lifetime reading classical texts in order to memorize significant passages. With the TLG and similar electronic tools an investigation that once would have taken an experienced philologist two weeks (and perhaps would have still been unsuccessful) can now be undertaken (and probably successfully) by an undergraduate student in a moment. Fragmentary texts, hitherto neglected because they could not be identified, now attract philologists attention.

What we are gaining from such small fragments of ancient literature on papyrus may not always be sensational in terms of textual criticism, but under certain circumstances the identification of such a fragment opens up much wider perspectives. For example, a tiny scrap of parchment proved to be the only extant piece of evidence for Amphilochius of Iconium surviving from antiquity (otherwise the writings of this church father are transmitted only by much younger medieval manuscripts).⁷ We learn that Amphilochius' Greek writings were read and copied in Coptic Egypt even after the Arab conquest.

Similar tools for neighbouring disciplines were developed and showed similar effects in their respective fields. The Packard Humanities Institute (PHI) financed a text database for Greek inscriptions.⁸ Comparable enterprises for Latin inscriptions are going on in Germany⁹ and at Duke University (North Carolina) a database was set up for the non-literary papyri: the Duke Databank of Documentary Papyri (DDBDP).¹⁰ Since these tools have become available, the number of corrections and piecing together of fragments have increased significantly, improving the basic texts for all historical or linguistic analyses. Literary papyri may likewise be consulted in a database from the Centre de Documentation de Papyrologie Littéraire of the Université de Liège¹¹ or the Leuven Database of Ancient Books¹² – to cite some prominent examples from papyrology. Similar examples could be drawn from other fields of Classical Studies as well.

With the complete corpus of ancient texts accessible in electronic form, research no longer depends on large libraries – at least to a degree. In the beginning, the electronic media provided the original texts, but not (or not yet) the commentary, the apparatus criticus or the editor's remarks. Meanwhile, many academic journals are available online, providing digitized back issues as well as current editions.¹³ Thus, it became easy to go back to the editio princeps and read the full editions and commentaries on screen. Moreover, comprehensive digital libraries – like JSTOR¹⁴ – offer immediate access to almost all journals, an increasing number of books and primary sources and, in addition, often provide full-text searches. On the other hand, one disadvantage of digital publications still remains: it remains difficult to keep more than a handful of “books” open on the screen simultaneously. However, as computer screens get bigger, this problem may be solved in near future. At the moment, alas, it looks as if the academic world prefers to work with the printed copy *and* the digital version online, enjoying the advantages of both.

Supplementary electronic “instrumenta” (tools) soon followed the text-oriented databases. For example, the Heidelberger Gesamtverzeichnis der Papyrusurkunden Ägyptens (HGV) provides the user with an exhaustive catalogue of Greek and Latin documentary papyri in a critical database that also has useful links to further tools.¹⁵ Within seconds, one may orientate himself on the status quaestionis in textual criticism and available translations as well as the date and origin of each single papyrus text; electronic links direct you to digital images. With the steadily growing number of electronic tools, specific navigators and overarching “portals” became an indispensable help for the user.¹⁶ A growing mass of data stimulated a widening of the scope of the tools. Comprehensive and systematic online catalogues followed an holistic approach and managed to include all extant sources and may function as directories to the scholarly resources and tools as well.¹⁷

After the texts, the next big step relates to electronic images. Within a few years the quality of digital images has reached such high resolutions that we may enlarge a single letter to the size of our screen – an incredible help for reading or checking difficult or damaged sections of a text. Once, scholars had to order black-and-white photographs, pay for them and wait half a year for them to arrive in the mail. Today the specialist checks a doubtful reading within a few minutes on his screen. In most cases the digital images are provided for free on the homepages of the public institutions (universities, libraries, archives, museums) housing the original objects. Again, this has had considerable consequences for our way of working. Checking edited

texts has become easier, corrections more numerous and there have been significant improvements of texts and interpretations. Paradoxically, by making images available, the texts have become even more central.

In papyrology – to stay with the example already cited above – a revolution started with the Advanced Papyrological Information System (APIS),¹⁸ a well organized project to digitize all (!) papyri of the various collections in the US. The revolutionary aspect lay in the including both unpublished and published texts. Initially, this principle encountered widespread criticism among European papyrologists, especially from those colleagues who were in charge of papyrus collections and thereby controlled access to unpublished material. After twenty years of APIS' "Open Access Policy" the critical voices fell silent. Virtually no negative experience was reported (like "pirate" editions being made without permission of the institution owning the text; or two scholars working simultaneously on the same text without knowing it). On the contrary, nowadays it is possible for an expert to survey tens of thousands of documents on the World Wide Web in various collections all over the world, searching systematically in "virtual collections" for a specific genre of text.¹⁹ Meanwhile, most European collections have also started digitization projects. A specialist for Hellenistic poetry (let us say) may go through the electronic files of at least thirty collections in Europe and North America in order to identify relevant, but still unpublished papyri. A specialist may work on them and publish them much faster and much better than anyone else. Thus, even in their initial phase of application, electronic media transformed Classical Studies into a much more international enterprise. The Open Access Policy, adopted by most (but still not all) academic and cultural institutions, clearly promotes a democratization of knowledge and accessibility of cultural heritage.²⁰ During the past few years, most European papyrus- and manuscript collections followed the American example.²¹ The opening up of the storerooms brought and will bring up many more exciting texts and thus stimulate research.²²

After only a few years, the impact of the electronic media on our way of working cannot be overlooked. It may be true that these media indirectly favour specialism, but, on the other hand, this gives hope that, in future, papyrologists will be able to edit more than the approximately 150 texts that an expert on papyri can currently do in his lifetime. Given the fact that in the past 120 years some 80,000 papyri have been published, but more than a million still remain to be deciphered, increased editing speed seems highly desirable.²³ Thus, digital images definitely helped classical scholars to work more systematically and efficiently.

It is worth mentioning another aspect of these developments. Until now, palaeographical studies have been a narrow field in Classics. It was simply too time-consuming and expensive to order a substantial number of photographs (often of questionable quality) from all over the world and then to decide which are suitable for palaeographical analyses and argumentation. Digital images, however, have resulted in a significant revival of palaeographical studies. Indeed, comparison of images and searches for parallels have only become possible on a larger scale thanks to thousands of freely available digitized images online.

Another topic is online catalogues and lexica. Preparing a lexicon is an extremely time-consuming task. Modern scientific communities and changing policies on financing academic institutions, making scholars increasingly dependent on external fundraising, are not favourable pre-conditions for long-term enterprises or ambitious projects producing, let us say, the third volume of a dictionary covering the letters from K to M in five years. On the other hand, electronic databases have the considerable advantage of not needing to be perfect from the outset. It is work in progress, acknowledged as such and treated as such by colleagues. Certainly, scholars are grateful for access to such databases and similar tools also in their preliminary form, let alone the fact that, for example, prosopographical lists are steadily enriched by new material and any printed version would soon be outdated. Collections of all kinds of sources, materials and prosopographical or geographical data no longer appear in printed paper versions, but are now set up as online databases. Such enterprises are best done in electronic version in the Wiki-system – and indeed the first Wiki databases of high scientific standard are evolving.²⁴ Although such databases are often kept up to date quite conveniently, others have been discontinued and remain nothing more than a torso now the funding has stopped. Sadly, in certain circumstances, the sustainability of electronic tools – usually created with an enormous amount of time and effort – have proved more vulnerable than printed media.

Lexica may be different from databases on prosopography or similar subjects. What we get here in electronic form is merely a mechanical, unreflected upon collection of data – not an intellectually reworked, systematic explanation of a word and its various meanings. In this respect, one potential danger emerges. The command of masses of data may, all too easily, lead to (and, in some cases, already led) collecting, but no deeper rethinking of information. Collecting and providing steadily growing masses of data might provoke (as Rüdiger Klein remarked at a conference in Budapest) a kind of “hunter-gatherer humanities.”

With increasingly complex digital tools, journals and books in the background it seems only a short step to electronic publishing. Indeed, the first activities have also begun in the Classics that rely heavily on durable, unmodified texts. Experience is gathered by publishing minor corrections and observations to texts only in an electronic form,²⁵ and, at the moment, a number of Classical periodicals not only put PDFs of their printed issues online, but also prepare for online publication or for the hybrid form of simultaneously printed and electronic publication. Computer programs like the *Classical Text Editor* – originally designed for scholars working on critical editions including commentary and apparatus criticus to prepare a camera-ready copy for print – were adopted for work on electronic publications as well.²⁶ Although hardly any Classical scholar would like to miss printed books, the field appears technically and mentally ready for electronic publishing.

Unlike the technical conditions, the legal and administrative conditions are far from being clear and prepared for the decisive step into the Digital Humanities. On the one hand, national and European research organizations expect that all papers arising from publicly-funded research are placed immediately in open access archives with as few restrictions as possible.²⁷

On the other hand, many academic publishers ask the exact opposite from their authors. An author may not be allowed to deposit his contribution in repositories for several months or years (embargo period),²⁸ or he may not even be allowed to post it at all, not even on his own homepage.²⁹ A scientific author who, for example, wishes to publish a conference paper in an edited volume may find himself in a serious conflict of legal interests between the obligation to publish open access and the archiving policies of the publisher. However, compared to the situation in the hard sciences, where a handful of powerful publishers have a monopoly over the best-ranked journals and charge vast fees to access research papers,³⁰ the problems in the publication culture of the Humanities may still seem limited and manageable.

A completely new method of dissemination and discussion of scientific papers was put forward by the Princeton-Stanford Working Papers in Classics.³¹ It started some years ago as an experiment and, although this platform is still limited to a rather close circle of scholars, it has already proved very successful. Papers are posted on a website for discussion even before they are published and printed. As everybody may comment on the paper, its final, printed version sometimes becomes significantly improved or enlarged. This form of discussion will probably increase and may have a major influence on the publication culture in the near future.

In addition to the many positive impacts of these developments, there are some problematic aspects that cannot be ignored. Firstly, electronic

publications are not usually ranked in any bibliometrical index and thus do not bring credits to authors. In particular, young scholars without tenured positions may be cautious about investing much time and effort on electronic tools, databases or digitization projects, as there may be disadvantages in terms of being in competition with colleagues who publish their research results in traditional printed form. Although many academics have become aware of the problem, a common method of evaluating electronic publications is still a desideratum. Secondly, we face technical storage problems. Who will be responsible for the repository infrastructure and who will pay for the expensive long-term storage of digital data? In a globalized scientific community the storage and technical update of data need to be done only once and may then be shared all over the world, but this still requires organization. Another, major problem is the long-term reliability of all electronic media. All branches of historical, philological and social studies need a safe and permanent basis as point of reference for further scholarly discourse. Most experts are sceptical that electronic media will ever be able to provide this firm basis. Too often, websites move from one electronic address to another or simply disappear (“Page not found”).³² Visions of a “single global archive of academic literature and data” are still to come to fruition.³³

In addition to all their praiseworthy advantages (like widening our perspectives, enabling us to approach more complex questions or tasks, commanding vast masses of data, working much faster) electronic media leave considerable space for the printed hard copy, standing on the bookshelf and serving as our fixed point of reference, untouched by changing computer systems and internet addresses. Along with the electronic tools, printed books and journals may serve as reliable media – at least as long as libraries are not suffering the same fate as the Alexandrian library and all other libraries of the ancient world.

Notes

1. In 2004, the American Philological Association (APA) had already defined its policy with a “Statement of Computer Support” on its homepage: <http://apaclassics.org/search/node/computer> (Accessed 30 March 2014).
2. For example, some of the oldest papyrological editions were reproductions of handwritten copies: *Aegyptische Urkunden aus den Königlichen Museen zu Berlin, Griechische Urkunden I*. Berlin: Weidmannsche Buchhandlung, 1895 and *II*. 1898. Print; Wessely, Carl. *Studien zur Palaeographie und Papy-*

- ruskunde III, Griechische Papyrusurkunden kleineren Formats*. Leipzig: Verlag von E. Avenarius, 1904. Print.
3. The TLG successively extended the canon of texts as well as the search possibilities: www.tlg.uci.edu/ (Accessed 30 March 2014). Follow the link to “Project History” for detailed information on the technical and scholarly development of the TLG.
 4. Greek and Latin authors are easily accessible via the *Perseus Digital Library*: <http://perseus.mpiwg-berlin.mpg.de/> (Accessed 30 March 2014).
 5. Müller, Matthias. *Alte Geschichte online. Probleme und Perspektiven althistorischen Wissenstransfers im Internet* (Computer und Antike 6). St. Katharinen: Scripta Mercaturae Verlag, 2003. Print. — Hainzmann, Manfred, and Schäfer, Christoph (ed.). *Alte Geschichte und Neue Medien. Zum EDV-Einsatz in der Altertumforschung* (Computer und Antike 5). St. Katharinen: Scripta Mercaturae Verlag, 2000. Print.
 6. Schäfer, Christoph, and Meier, Angelika (ed.). *Computer und antike Texte. Wortrecherche, Konkordanz- und Indexerstellung mit Volltextdatenbanken* (Computer und Antike 1). St. Katharinen: Scripta Mercaturae Verlag, 1993. Print. Schäfer, Christoph, and Krüpke, Florian (ed.). *Digitalisierte Vergangenheit. Datenbanken und Multimedia von der Antike bis zur frühen Neuzeit* (Philippika. Marburger altertumskundliche Abhandlungen 5). Wiesbaden: Harrassowitz Verlag, 2004. Print.
 7. Mitthof, Fritz. “Bruchstück einer Homilie des Amphilochos von Ikonion aus frühbyzantinischer Zeit (Or. VI 1-3)”. *Zeitschrift für antikes Christentum* 4 (2000) 209-217, photo on p. 211. Print. A digital image of the fragment is available on the homepage of the Austrian National Library: http://aleph.onb.ac.at/F/RT4C7D7J1JQ1YA3S224BG7IX3JKDRHXVCPMLTSRE-7FY7KUYM-02852?func=full-set-set&set_number=035390&set_entry=000001&format=999 (Accessed 30 March 2014).
 8. Greek inscriptions are collected by a project of The Packard Humanities Institute (PHI): <http://epigraphy.packhum.org/inscriptions/> (Accessed 30 March 2014).
 9. The *Epigraphische Datenbank Heidelberg*: www.uni-heidelberg.de/institute/sonst/adw/edh/index.html (Accessed 30 March 2014) contains more than 67,000 critically revised ancient inscriptions, originally published in 14,000 printed media, while the *Epigraphische Datenbank Clauss - Slaby* has even more, with more than 458,000 inscriptions from 2,500 publications: www.manfredclaus.de/ (Accessed 30 March 2014). Links for other epigraphical databases may be found at *Epigraphical Documents in tei xml* (EpiDoc): <http://epidoc.sourceforge.net/> (Accessed 30 March 2014).
 10. Conveniently accessible via the navigator *papyri.info*: www.papyri.info/ (Accessed 30 March 2014).
 11. CEDOPAL: <http://promethee.philo.ulg.ac.be/cedopal/index.htm> (Accessed 30 March 2014).

12. LDAB: www.trismegistos.org/ldab (Accessed 30 March 2014), currently containing more than 14,000 Greek, Latin, Coptic, Demotic, Syriac and other literary texts.
13. A typical example is *The Bulletin of the American Society of Papyrologists*, appearing in print since 1963. The website makes all issues of the *BASP* available in electronic form: <http://quod.lib.umich.edu/b/basp/> (Accessed 30 March 2014).
14. JSTOR, launched in 1995 by the Andrew W. Mellon Foundation, contains digitized issues of approximately 2,000 academic journals: www.jstor.org/ (Accessed 30 March 2014).
15. HGV: www.rzuser.uni-heidelberg.de/~gvo/ (approximately 60,000 entries). (Accessed 30 March 2014).
16. A good example is *Clio online – Fachportal für die Geschichtswissenschaften*: www.clio-online.de/site/lang__pl/mid__11526/ModeID__1/40208253/default.aspx (Accessed 30 March 2014), providing links to other major databases and portals; cf. Schäfer, Christoph. “Vernetztes Wissen,” in: Wirbelauer, Eckhard (ed.). *Oldenbourg Geschichte Lehrbuch. Antike*. München: Oldenbourg Verlag 2004. p. 481-492. Print.
17. For Classics, the interdisciplinary portal *Trismegistos*, once focused on Egypt, currently keeps developing into such an encyclopaedic tool by “expanding its geographical scope to the Ancient World in general”: www.trismegistos.org/ (Accessed 30 March 2014).
18. APIS: www.columbia.edu/cu/lweb/projects/digital/apis/about.html, since 2013 also available at <http://papyri.info/> (approximately 20,000 images provided), (Accessed 30 March 2014).
19. Van Minnen, Peter. “The Future of Papyrology,” in: Bagnall, Roger S. (ed.). *The Oxford Handbook of Papyrology*. Oxford: Oxford UP 2009. 644-660, esp. 649. Print.
20. Science Europe defined Open Access in a Position Statement from April 2013: “Open access, as defined in the Berlin Declaration, means unrestricted, online access to peer-reviewed, scholarly research papers for reading and productive re-use, not impeded by any financial, organisational, legal or technical barriers. Ideally, the only restriction on use is an obligation to attribute the work to the author” (page 1): www.scienceeurope.org/uploads/PublicDocumentsAndSpeeches/SE_OA_Pos_Statement.pdf (Accessed 30 March 2014). Cf. the initiative of the European Commission on Research and Innovation: <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1294&lang=1> (Accessed 30 March 2014).
21. In Germany, a comparable effort to link the digital databases of numerous papyrus collections under one navigator has been started in Leipzig: “*Das Papyrus Portal*”: www.papyrusportal.de/content/below/start.xml?XSL.lastPage.SESSION=/content/below/start.xml (Accessed 30 March 2014).
22. For similar approaches in epigraphy, see: *Concordia*: <http://concordia.atlantides.org/> (Accessed 30 March 2014).

23. Van Minnen, Peter. "The Millennium of Papyrology (2001-)" in: Palme, Bernhard (ed.). *Akten des 23. Internationalen Papyrologenkongresses* (Papyrologica Vindobonensia 1). Vienna: Verlag der ÖAW, 2007. p. 703-714.
24. Bierweiler, Jan, and Scholz, Martin. *Altertumswissenschaftliche Multimedia-Produktion mit den Autorensystemen Authorware und Director von Macro-media* (Computer und Antike 7). St. Katharinen: Scripta Mercaturae Verlag, 2003. Print.
25. Staying with the papyrological example, one might quote the *Bulletin of Online Emendations to Papyri (BOEP)*, first issued in 2012: www.uni-heidelberg.de/fakultaeten/philosophie/zaw/papy/projekt/bulletin.html (Accessed 30 March 2014).
26. *Classical Text Editor* version 8.08 provides features for digital publication: <http://cte.oeaw.ac.at/> (Accessed 30 March 2014).
27. Cf. the Position Statement of Science Europe (cited above, note 20): "Science Europe wishes to encourage the European Commission, national governments, research funding and research performing organisations and other stakeholders across the world to adopt this approach to Open Access and to actively nurture collaboration in this area" (page 4).
28. A typical (and comparatively generous) regulation for journal articles and contributions to edited volumes reads, e.g. "The author is also allowed to archive the 'version of record,' i.e. the article or contribution as published including copy editing and lay out (in PDF) on the author's personal webpage and on the author's departmental website upon publication, and in institutional repositories or subject repositories no sooner than 12 months after publication date" (Leuven University Press, Open Access Policy: http://upers.kuleuven.be/sites/upers.kuleuven.be/files/page/files/LeuvenUniversityPress_OAPolicy.pdf), (Accessed 30 March 2014).
29. Several important publishers are restrictive about the publisher's version, but allow postings of the author's version "He/she (sc. the author) may not use the publisher's version (the final article), which is posted on SpringerLink and other Springer websites, for the purpose of self-archiving or deposit. Furthermore, the Author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website": www.springer.com/open+access/authors+rights?SGWID=0-176704-12-683201-0 (Accessed 30 March 2014).
30. Cf. George Monbiot, "Academic publishers make Murdoch look like a socialist", *The Guardian*, 29 August 2011: www.theguardian.com/commentis-free/2011/aug/29/academic-publishers-murdoch-socialist.
31. The purpose of this collaborative project is to "make the results of current research [...] available in advance of final publication": www.princeton.edu/~pswpc/ (Accessed 30 March 2014).
32. Online publications, however, are much more flexible than other storage methods. During the 1990s a number of progressive book publications were

- accompanied by discs or CDs containing the basic data or sources of the analytical or hypothesis-driven discussion in the book. Needless to say, these media are barely readable only a few years after their publication.
33. Björn Brembs, 2011. What's Wrong with Scholarly Publishing Today? II. www.slideshare.net/brembs/whats-wrong-with-scholarly-publishing-today-ii (Accessed 30 March 2014).

Cutting Edge
New Means of Access, Evaluation and Funding

Publication Practices in Motion

The Benefits of Open Access Publishing for the Humanities

Janneke Adema and Eelco Ferwerda

Introduction

The internet holds a myriad of benefits for science as a whole and more specifically for the Humanities. To some, the dream of a universal online library containing all the texts in the world seems more conceivable than ever before. In this narrative, global access to information, from any place and at any time, offers the potential to improve the quality of research, the efficiency of science, and the speed of communication.¹

Despite these potential benefits of the online environment, the uptake of digital communication and publication practices in the Humanities has been slow. But this situation is changing. The availability of e-books in libraries is on the rise, though this has yet to answer the growing need of scholars to be able to access monographs online. (*Textual analysis of open-ended questions in e-book national observatory survey* 7, 11; Milloy 3; Swan 41) Nevertheless, many books still appear only in printed form and, although screen reading is on the rise, most scholars still prefer to read their books on paper when studying them in detail (L. Brown et al. 9; Rowlands & Fieldhouse 15; Heath & Robey; Adema & Rutten 52-53; Tenopir, Volentine, and King). Scholars in the Humanities appear to be standing on a threshold – on the one hand, they are eager to enter the digital realm (where some forerunners have already gone and are experimenting with new possibilities) and, on the other hand, they are held back by their norms and values as well as accepted institutional practices. These practices – mixed with a good amount of scepticism and fear concerning online publishing – are inhibiting them from taking advantage of the possibilities that the digital world offers. Any experimentation that is taking place with new practices more attuned to the digital environment is occurring only very slowly. How this will work out is uncertain. One thing *is* certain, though: when it comes to dealing with the internet, the Humanities are very much in a transitional phase.

This paper will show how Open Access publishing² can aid these Humanities scholars in transition – from the English professor who hardly touches a computer (except for the occasional email or to search for something in

an online library catalogue) to the digital humanist building collaborative virtual environments to present and communicate the newest version of her (and her groups') data sets. In particular, Open Access publishing will serve the scholar that finds himself positioned somewhere in between these two extremes: the scholar who occasionally reads and browses a book online, and increasingly publishes in hybrid journals; the scholar who finds that her communication with her peers is increasingly taking place in an online environment; the same scholar who is afraid that her work might be stolen or plagiarized in the online world and who is sceptical about how authority, quality and integrity can be maintained in digital publishing; and who, moreover, is all too aware that tenure and promotion committees still judge a book by its (printed) cover.

As Christine Borgman argues, we must create an information infrastructure that supports scholarship in all its multiple forms of communication. At the same time, we need to balance the interests of the many stakeholders in the scholarly communication system, since "any new system that we might develop will work better if it works for the interests of the scholars and other stakeholders than if it works against these interests" (Borgman 233). It is therefore crucial for publishers to cater to the needs of Humanities scholars in all their old and new forms of communication and publication, and to follow them as their practices change and develop. As we will argue below, Open Access publishing does exactly this. Moreover, the move towards Open Access models for the monograph is essential for the development of Humanities research at both the *transitional phase* from print to digital as well as during the period after this paradigm shift has been completed.

Current Communication and Publication Practices in the Humanities

When describing the "traditional profile" of the Humanities scholar, the first thing that comes to mind is the image of the lone scholar working extensively, for years, on his *magnum opus*: a thick and very specialized research monograph. Although this ideal image is increasingly being contested as a generalization, both from a historical viewpoint – where collaborations, forms of co- and sub-authorship, and networks of scholars and texts have always been part of science, even in the Humanities (Cronin, Shaw & La Barre; Cronin; Birnholtz) – and from a practical viewpoint, given that both the reading and writing of articles as opposed to monographs has been on the rise in the Humanities (Adema & Rutten 54). However, as many studies

show, this traditional image remains a fairly good representation of the work practice of the Humanities scholar, even in the digital age (Jennifer Wolfe Thompson 121-136; Adema & Rutten 51-55; Cronin, Shaw & La Barre 856).

As Jennifer Wolfe Thompson asserts, Humanities scholars still mostly work alone, following an “individual approach” in which authorship, individual responsibility and the attribution of a text are considered very important. The monographic format is preferred, as it allows scholars to develop multiple complex arguments and narratives. In addition to the monograph’s multifaceted and complex nature, another feature that appeals to Humanities scholars is the possibility that the monograph offers for extensive analysis of large sets of (primary) sources, whereas journal articles serve more as a means to develop critical dialogues (Williams et al. 74). As many reports and recent research have shown, the monograph also remains essential for the career perspective of Humanities scholars, as it is a prerequisite for promotion to senior academic posts. For these scholars, the main objectives in publishing are to communicate with their peers and to enhance their careers, with financial rewards or royalties playing only a minor role (Steele, “Phoenix rising” 115-116; Swan & S. Brown; Kennan & Kautz 4; Adema & Rutten 54).

When scholars conduct a thorough study of a monograph, they still prefer to consult a printed edition. Print also holds more status, as there is a strong cultural attachment to the printed form in the Humanities. However, this situation is starting to change. Previously, many researchers considered electronic-only publications as the equivalent of publishing something without peer review. Recent research has shown that electronic and openly available publications are not necessarily seen as being of lesser quality than printed publications (Adema & Rutten 51-55). Scholars’ attitudes in this respect are thus shifting. Nonetheless, many researchers still feel that the internet is not a good place to find authoritative material because of the large amount of poor-quality information, as evidenced in the CIBER/UCL interviews (Williams et al. 17-18).

The Rise of the Digital Humanities

Although the Humanities scholar still closely reflects the traditional field profile as outlined above, change is in the air. This is most noticeable in the consumption of digitally and online available scholarly information. As we remarked upon in the introduction, Humanities scholars increasingly find and consult publications online. The Ithaka report has likewise observed

a marked rise in online informal communication between scholars. As scholars increasingly start to rely on these informal communication channels, the Ithaka report predicts the boundaries between what constitutes formal and informal publication will continue to blur. The status of online communication channels is also being reconsidered. As Alma Swan notes, blogs and wikis are increasingly perceived as places where scholars can find and share high-quality scholarly material (Swan 12).

However, while the consumption of online available information has risen, when it comes to the production side, things have not yet changed as dramatically. Furthermore, these developments are not equally distributed among all fields in the Humanities. Indeed, in some fields they are still highly contested (Adema & Rutten 6).

Nevertheless, a group of forerunners in the Digital Humanities, Humanities Computing and similar fields is already using digital applications and online publication and communication possibilities extensively. The Ithaka report foresees a future in which collaborative and virtual knowledge and research environments will become more widespread in the Humanities, as will new experimental formats, multimedia and enhanced publications, versioning³ updates, and data mining (L. Brown et al. 4). These changes in scholarship and research practice will, however, only come about if the necessary applications and services are readily available and widely used and adapted.

How can Open Access publishing assist Humanities scholars who stand before, on and have stepped beyond the digital threshold? As we argue below, Open Access makes sense for *all* Humanities scholars, regardless of the transition phase they are in. Moreover, it is through Open Access publishing that publishers can best guide Humanities scholars through this transition.

How Open Access Publishing is Beneficial to the Scholar who Hardly Uses the Internet

For the most traditional group of Humanities scholars, Open Access publishing offers the possibility of a revival of the monograph. Through their experiments with Open Access book publishing, publishers are searching for new (business) models for monograph publishing, as the current model for book publishing in the Humanities has come under sustained pressure.⁴ The spike in journal prices over the last decades and the subsequent cuts and re-allocations in libraries' budgets have led to a decline in monograph sales. This has had major

consequences for scholars, as publishers have been forced to focus more on general (trade) books, limiting the opportunities for scholars to publish highly specialized monographs. This has had a detrimental effect in recent years on the variety and accessibility of specialized academic books in the Humanities (Greco & Wharton, 154; John B Thompson 93-94; Steele, "Phoenix rising" 111-112; Waltham; Willinksy; Steele, "Scholarly Monograph Publishing in the 21st Century").⁵ The consequences for scholars have been even more dire, as the publication of monographs remains essential for Humanities scholars to acquire tenure or advance their careers (Bernius et al. 105).

New solutions based on Open Access publishing aim to tackle this problem and search for models that increase the dissemination and accessibility of specialized Humanities books. Open Access models should bring an end to the traditional small print runs of scholarly monographs. Proponents believe that Open Access book models could open up scholarly book content to a wider audience and to new services and applications. Open Access publishing could also lead to a more efficient and cheaper publishing system through innovations in the publishing model and the use of new digital techniques, infrastructures and tools. Furthermore, in many cases, Open Access publishers advocate the idea that publishing should be considered an integral part of the research costs itself, which opens up new sources of funding for the publication of monographs (Greco & Wharton, 158).⁶ All in all, Open Access publishing offers a renewed chance for a revival of the monograph, both in printed and online formats.

How Open Access Publishing is Beneficial to the Typical Humanities Scholar Shifting from Print to Digital

The average Humanities scholar stands on the threshold between the traditional and digital worlds, with an affection and need for both the printed book and its electronic counterpart. For these scholars, Open Access publishing has much to offer. Firstly, most Open Access book experiments focus on the so-called hybrid model in which an online Open Access edition is supplemented with a printed edition, often a print-on-demand version (Adema 7-8). Researchers can thus choose to browse through, search, mine and adapt the online version; or, if deemed essential for a more thorough study, to obtain a printed version (either personally or through a library). Online versions also allow researchers to explore digital publications deeper by running (multilingual) full-text searches over collections, chasing hyperlinks to original sources, or annotating PDFs after downloading them.

Open Access publishing can also play a role in changing scholars' scepticism concerning the quality and trustworthiness of online publications. One way Open Access publishers can do this is by being fully transparent with regards to their peer-review policies and by stressing the fact that they adhere to the same quality standards as in the past. Open Access initiatives can also push forward the discussion on new standards and rules in terms of establishing the quality and ensuring the integrity of the text if one wishes to do so. The same holds true for scholars' concerns pertaining to copyright and digital preservation. Fears concerning these issues need to be addressed, and this can only be done by discussing and implementing new standards based on digital practices.

How Open Access Publishing is Beneficial to the Digital Humanist

For the Humanities scholar already immersed in the digital world, openly accessible information is a must. In order to conduct experiments with large amounts of data or to explore open research, liquid publications, visualizations, collaboratories (virtual collaboration environments) and wikis, the free availability of information for users in various global settings is a necessary precondition. Information then becomes much easier to mine and reuse; it can become truly interactive, offering a variety of possibilities for new scholarly methods and forms of analysis in the Humanities. Open Access publishing also enables publishers to connect more closely with their (potential) authors by offering them services already during the research phase. With the shift in focus in the digital realm from product to process, publishers can aid scholars by setting up (qualified and branded) research environments, which can also offer multiple publication paths (both formal and informal). In this way, Open Access publishing becomes just one of an assortment of publishers' services aimed at facilitating scholarly communication up to and after the moment of publication.

New Models to Assist the Humanities Scholar

In the experiments currently being conducted with new Open Access models for books, two models stand out. One is based on an author-pays scheme and is developed by the Open Access in European Networks (OAPEN) project. The argument for having the author pay for publication is that funding for

publications should be integral to research funding in the Humanities, as otherwise publishing models for such books will probably never be sustainable. Through this model, OAPEN wants to ensure and improve the availability of books both online and in print. Another model, advocated by the Open Humanities Press, is based on a (academic-led) library-press combination, in which the importance of quality assurance and branding is stressed. This approach aims to ensure that online and openly available publications adhere to the same quality standards as printed publications, thus making it easier, or at the very least more acceptable, for Humanities scholars to experiment with the online publishing of monographs.

OAPEN

The OAPEN project, which started in September 2008 and was completed in 2011, aimed to develop an Open Access publication model for peer-reviewed books in the humanities and social sciences, the area of academics in which most monographs are written. The project was backed by a consortium of seven university presses and two universities⁷ and was supported by the European Union for a period of thirty months.

The goals of OAPEN were to improve the accessibility and impact of humanities research and to create a large, freely available, online collection of current books in European languages in various fields within the Humanities. Most importantly, OAPEN wanted to engage all stakeholders – i.e. authors, researchers, funding agencies, libraries and publishers – in order to make this transition towards a new business model for books as easy and smooth as possible. OAPEN therefore tried to adapt to the system as it was changing, even as its objective was to transform and improve this system without alienating the various actors involved in the creation, production, dissemination and preservation of academic books. OAPEN's strategy in this respect was also very focused on making the benefits of Open Access publishing visible to the various stakeholders and to actively engage them. In their view, all stakeholders in the scholarly communication process can benefit from Open Access monographs. Authors receive a much larger outreach for their work through the internet. They have a worldwide potential audience through which they can increase their visibility and their impact (Alma Swan). Funding agencies also gain from the wider dissemination of research, as it increases the return on their investment. Researchers have the advantage of full-text search and unrestricted access to academic texts. They can search across platforms and also make

connections across platforms, potentially creating a truly linked data web. Open Access publishing enables libraries to improve their services and to answer their users' needs in a better way, particularly the needs of students and researchers. And finally for publishers, Open Access allows them to be more effective in disseminating knowledge, which is still one of their main functions. With the help of these stakeholders, OAPEN wanted to develop common funding models, common standards and standardized metadata schemes to improve the retrievability of texts. An OAPEN library of freely available HSS books was built, as well as a publishing platform that can be used by academic presses.

The central belief behind the OAPEN project was that publishing should be an integral part of research funding and that publishers should develop Open Access publishing as a service to the scholarly community. What is clear from research conducted by the OAPEN project into current experiments of various Open Access book initiatives (Adema) is that none of these initiatives are able to sustain their businesses based on print-on-demand sales alone. Therefore, in most cases, some form of support is necessary, be it institutional support, temporary grants or infrastructural support. The remarkable thing about the OAPEN study was that, at that time, it found no examples of a press publishing Open Access books for a publication fee under a so-called author-pays scheme, which is one of the more common business models for Open Access articles. And this is exactly the model that OAPEN intends to pursue for Open Access books in the Humanities. To promote this model among different research funders, OAPEN is creating and testing a model that consists of guidelines – guidelines for calculating the cost of an Open Access book, for peer review, for quality assurance, for producing the technical standards for the metadata, and for the licensing of the content.

In the OAPEN model, the online version (the Open Access edition) is free. Other editions can be sold. Catering to the needs and current practices of scholars and libraries, a printed edition can also be provided via print-on-demand. The OAPEN model is based on identifying the costs of publishing, all of which should be made as transparent as possible. The costs to make the Open Access edition (such as peer review, editing and typesetting) are separated from the costs to produce other editions (such as printing, distributing and marketing and sales). The publication fee is based on the costs for the Open access edition. The publication fees charged by academic publishers could come from a mix of institutions that would include research funders, universities, and libraries or library consortia. OAPEN thus sees publishing Open Access books as a collaboration between publishers and funders. In sum, OAPEN's aim is to change the business

model for monograph publishing, making it more sustainable. This transition towards OAPEN's new, more digitally oriented and intrinsically open publishing model will assist Humanities scholars in overcoming their fears and uncertainties concerning digital publishing and research practices. Not only does the OAPEN model offer services through its online library (including increased accessibility and search possibilities), it also offers scholars the possibility to *publish* their books online in a trusted, branded and integrated setting. And from this initial structure and platform, the groundwork has been laid for experiments with enhancements, multimedia works and other new formats. OAPEN thus remains open to incorporating new possibilities that arise in the digital medium.

After the close of the project in 2011, OAPEN continued its activities as a foundation, with support from a number of academic institutions in the Netherlands, including the Netherlands Organization for Scientific Research (NWO). In 2014, OAPEN aims to establish the OAPEN Library as a Deposit service for Open Access books, with a number of launching participants in various European countries.

Open Humanities Press

Open Humanities Press (OHP) is another example of a press experimenting with Open Access book publishing. Founded in 2006 by "Open Access journal editors, librarians and technologists" (Jöttkandt 4), OHP is an international Open Access publishing collective in critical and cultural theory. In contrast to OAPEN, the focus of OHP lies less on creating a new publishing model and more on removing negative perceptions that still exist concerning Open Access and online publishing.

OHP is not only academic-led but also (and perhaps more so than other, more traditional presses) academic-focused. Its philosophy is one of advocacy and of making clear to researchers and other stakeholders what the benefits of Open Access are. Most of all, OHP wants to battle the negative perceptions that pertain to online publishing. Digging into the origins of these negative perceptions, the founders of OHP asked themselves: what creates this perception that the internet is "not currently a suitable publication medium" for serious scholarship in the Humanities, and why is it that "this [perception] continues despite the existence of numerous excellent online journals, both toll-access and open access, in many of [their] fields?" (Jöttkandt 1). Battling this perception serves two goals: first, it makes experimentation with new business models truly possible, which can help

solve the current publishing crisis in the Humanities; and second, it paves the way for further experiments in scholarly communication (Jöttkandt 3-4).

As OHP argues, it is crucial for publishers to follow the practices and perceptions of scholars. This is where Open Access publishing can be beneficial:

As a Steering Group, we felt that the most effective way of achieving these goals, that is, the most efficient way to change people's attitudes and behavior towards internet publishing, is not to ask them to change at all. People do not easily change established patterns of behavior, therefore to be successful technology must adapt to them. In the online publishing space, this would mean that mainstream OA journals, that is, OA journals that are online for reasons other than the inherent theoretical interest of the digital medium itself, should ideally be indistinguishable from their print counterparts, or at least indistinguishable from the electronic copies of these print journals that our colleagues are increasingly demanding. (Jöttkandt 4-5)

Establishing notions of trust and quality in electronic and Open Access publishing are essential to persuading the scholar to step over the threshold into the digital realm. OHP wants to make it easier for those scholars who are not that involved or interested in the online world to make this transition:

OHP's strategy is thus to use an existing model from humanities publishing, the one that our colleagues understand best, namely, a publishing house or "press." We felt that this was still the most powerful metaphor we could employ to gain the trust of our communities, especially that of colleagues who have no interest in digital media per se. (Jöttkandt 5)

OHP wants to counter this perception problem by establishing a strong brand around its online and openly available products that is first and foremost trustworthy. All of its publications are peer reviewed and academically certified by OHP's renowned independent board of international scholars (Jöttkandt, Kimball, & Willinsky). Furthermore, OHP actively advocates university executives to rethink their hiring and admission policies and to accept the legitimacy of online digital publications (Jöttkandt & Hall 4).

But beyond creating a trustworthy, reliable, high-quality system for those scholars still sceptical about online information, the OHP's Open Access initiative also strives to cater to the needs of the digital humanist, who relies on Open Access to develop his work. It is only by making the resources he builds his quantitative analysis on openly available that the academic's work becomes truly innovative (Jöttkandt 7). One of the founders of OHP, Sigi

Jöttkandt, posits that in order to make use of the power and potential of the internet as a network – in order for sites of genuine potential, innovation and interaction to arise – we must have no barriers to access:

What we at OHP would like to see more explicitly foregrounded in the discussion around the Digital Humanities, however, is something that strikes me and other members of the Steering Group as absolutely fundamental: the need to make the scholarly materials these tools are created for freely and openly accessible. I'll go out on a limb here and say that without free and open access to these materials, the majority of the innovations of the Digital Humanities will remain purely aesthetic. This is to say, the Digital Humanities will have built a tremendously fascinating instrumentarium but the internet's genuinely transformational promise will have been missed, largely as a result of our failure to understand the full implications of the digital medium itself (Jöttkandt 6).

Furthermore, Open Access and online publication should not be seen as a threat to notions of authority, trust, legitimacy and quality. Gary Hall, another founder of OHP, argues that it should be seen as a chance to start a (much-needed) discussion on these issues in order to make the transition to the digital world easier and smoother:

From a humanities point of view, however, the digital mode of reproduction, with its loosening of much of the stability, permanence and 'fixity' of texts, promises to place us in a position where we are again called on to actively respond and make such judgements and decisions. In this respect, one could argue that the shift to online open access publishing offers us a chance to raise precisely the kind of responsible questions regarding the authority and legitimacy of scholarly writing, authorship, peer review and intellectual property, that we really should have been asking all along (Hall 13).

Whereas, on the one hand, OHP focuses on creating a trustworthy environment that mirrors the paper world in order to make it easier for sceptical Humanities scholars to make the step to the online world; on the other hand, as a publisher it aims to be part of the movement of scholars that is at the forefront in its embracing of the digital world and in its rethinking of the paper notions upon which our scholarly communication system is still built:

Now, the way most electronic academic publications have attempted to address this issue to date is by imitating their paper counterparts: in their

'page' layouts; their publication of material in the form of 'papers' written in a linear, sequential form; their reliance on International Editorial Boards of established academics who have already proven themselves in the 'paper' world; but most especially in their peer reviewing processes. They do so because they need to try to reassure the University about something that is still relatively new by demonstrating that they are providing recognizable forms of quality control. And of course what I've just described is precisely the strategy Sigi, David and I, and the rest of the project's Steering Group are pursuing with Open Humanities Press. Because if we didn't, we'd risk not being taken particularly seriously. But we might also ask: to what extent can we continue to insist that digitally reproduced texts conform to the standards, hierarchies and protocols of the paper world? (After all, when it comes to academic journals this format is over three hundred years old now.) (Jöttkandt & Hall 4).

The Future of Openness Within the Humanities

Both the OAPEN model and the Open Humanities Press initiative show the potential benefits Open Access book publishing can bring to Humanities scholars. At the same time, we need to be wary of the direction the current discourse on openness seems to be taking. Openness in many ways can be seen as a floating signifier (Laclau 129-155), a concept without a fixed meaning and one that is easily adopted by different political ideologies. In this case, Open Access runs the risk of being overtaken by a neoliberal rhetoric that seems to be increasingly apparent within academia, one focusing on transparency and accountability and on the measurement and evaluation of research and research results as part of an audit culture (Hall *Media Gifts*). For Open Access to be beneficial for scholars and society at large, the focus should remain on the values most cherished by the scholarly community: accessibility and quality. Although the search for a sustainable business model is of the utmost importance, the adaptation of Open Access as simply another choice model by commercial publishers (offering it as an option only if the author asks for it) or as a way to further advance profits by charging above proportional author fees, is not the way forward. Openness is admirable but we should be aware of the negative side effects. At the same time, we should be wary of arguing for the Open Access exchange of research results within a world in which the delivery platforms (especially mobile platforms) are increasingly closed off as part of their business models. Open Access will only flourish within an open web, without proprietary standards and extensive DRM regimes.

Open Access should be promoted to governments and institutions as a system that can create more value for society (Houghton) and as a model that needs extra investment to experiment with new models, to make sure that the cost of providing Open Access will not reduce the availability of funding for Humanities research and will not come out of already decreasing research budgets. In sum, Open Access offers possibilities for change within the system but while we are promoting it, we should be wary of wrong uses and adaptations.

Conclusion

To assist the Humanities scholar in making the transition from print to digital, it is of the utmost importance for publishers to create a system that improves the availability of books, both in print and online, while at the same time providing the highest level of quality to ensure a trusted setting. Open Access publishing can be the solution to the problems currently faced by the print-based subscription model, while at the same time increasing the accessibility and availability of scholarly monographs for all. At least for the time being, a hybrid model should be pursued, which would offer both a printed and an online edition. As new possibilities continue to be developed in the digital realm, a flexible and modular system that is open and adaptable to changes in scholarly practices is preferred. The main focus should thus be on scholars and their needs, avoiding the development of narrowly focused, top-down constructions. Connecting not only scholars and texts but also scholars and institutions is essential to improve the quality of research and to make it more effective and efficient. Publishers should be transparent about their aims, policies and procedures in order to make the system robust and fair. Only if these issues are taken into consideration and the necessary conditions are met, the true potential of the digital realm will be able to unfold.

Notes

1. The ideal of the universal library has a long history, from the Library of Alexandria to Paul Otlet's Mundaneum and Vannevar Bush's Memex. It has more recently been expressed in an article by Kevin Kelly called 'Scan this Book.' However, one should not forget that this ideal also knows a dystopian

- counterpoint, envisioned most clearly in Borges' description of the total library.
2. Open Access publishing, also known as Gold Open Access, is a way of delivering research output via the publishing of Open Access journals or books. This is opposed to Green Open Access, where the research results are deposited in Open Access repositories after publication. (See Peter Suber's Open Access overview: www.earlham.edu/~peters/fos/overview.htm).
 3. A "version" is a digital object (in whatever format) that exists in time and place and has a context that can be described by the relationship it has to one or more other objects. Definition used by the Version Identification Framework: www2.lse.ac.uk/library/vif/index.html.
 4. There are many Open Access activists who believe that Open Access and online publication may be the only sustainable model for the monograph (Jensen).
 5. This is confirmed by Greco and Wharton who estimate that the average number of library purchases of monographs has dropped from 1500 in the 1970s to 200-300 at present. Thompson estimates that print runs and sales have declined from 2000-3000 (print runs and sales) in the 1970s to print runs of between 600-1000 and sales of in between 400-500 nowadays.
 6. Greco and Wharton show that the vast majority of presses post financial losses every year, even when they are subsidized.
 7. The OAPEN partners were Amsterdam University Press, Georg-August Universität Göttingen, Museum Tusulanum Press, Manchester University Press, Presses Universitaires de Lyon, Firenze University Press, University of Amsterdam and Leiden University.

Bibliography

- Adema, Janneke. *Overview of Open Access models for eBooks in the Humanities and Social Sciences (OAPEN 2010)*. OAPEN, 2010. Print.
- Adema, Janneke, and Paul Rutten. *Digital Monographs in the Humanities and Social Sciences: Report on User Needs*. OAPEN, 2010. Print.
- Bernius, Steffen et al. "Open Access Models and their Implications for the Players on the Scientific Publishing Market." *Economic Analysis & Policy* 39.1 (2009): 103-115. Print.
- Birnholtz, Jeremy P. "What Does it Mean to be an Author? The Intersection of Credit, Contribution, and Collaboration in Science." *Journal of the American Society for Information Science and Technology* 57.13 (2006): 1758-1770. Web.
- Borgman, Christine. *Scholarship in the Digital Age: Information, Infrastructure and the Internet*. Cambridge MA: MIT, 2007. Print.
- Brown, Laura et al. *Ithaka Report. University Publishing in a Digital Age*. 2007. Print.
- Cronin, Blaise. "Hyperauthorship: A Postmodern Perversion or Evidence of a Structural Shift in Scholarly Communication Practices?" *Journal of the American Society for Information Science and Technology* 52.7 (2001): 558-569. Web.

- Cronin, Blaise, Debora Shaw, and Kathryn La Barre. "A Cast of Thousands: Coauthorship and Subauthorship Collaboration in the 20th Century as Manifested in the Scholarly Journal Literature of Psychology and Philosophy." *Journal of the American Society for Information Science and Technology* 54.9 (2003): 855-71. Print.
- Greco, Albert N, and Robert Michael Wharton. "Should University Presses Adopt an Open Access [Electronic Publishing] Business Model for all of their Scholarly Books?" *ELPUB. Open Scholarship: Authority, Community, and Sustainability in the Age of Web 2.0 – Proceedings of the 12th International Conference on Electronic Publishing held in Toronto, Canada 25-27 June 2008 / Edited by: Leslie Chan and Susanna Mornati*. 2008. Print.
- Hall, Gary. "The Impact of the Humanities: or, What's Next for Open Access." 3 April 2008: n. pag. Print.
- Hall, Gary. "On the limits of openness III: open government." *Media Gifts* 12 December 2010. Web.
- Heath, Malcolm Michael, and David Robey. "E-Publication and Open Access in the Arts and Humanities in the UK." *Ariadne* 54 (2008): n. pag. Print.
- Houghton, John. "Open Access – What are the Economic Benefits?: A Comparison of the United Kingdom, Netherlands and Denmark." June 2009.
- Jensen, Michael. "Scholarly Publishing in the New Era of Scarcity." 2009: n. pag. Print.
- Jöttkandt, Sigi, and Gary Hall. "Beyond Impact: OA in the Humanities." 13 February 2007: n. pag. Print.
- Jöttkandt, Sigi, Shana Kimball, and John Willinsky. "The Role of Libraries in Emerging Models of Scholarly Communications." 13 Oct. 2009: n. pag. Print.
- Jöttkandt, Sigi. "Free/Libre Scholarship: Open Humanities Press." 3 Apr. 2008: n. pag. Print.
- Kelly, Kevin. "Scan This Book!" *The New York Times* 14 May 2006. Web. 25 March 2010.
- Kennan, Mary Anne, and Karlheinz Kautz. "Scholarly publishing and Open Access: Searching for Understanding of an Emerging IS Phenomenon." *Proceedings ECIS 2007 – The 15th European conference on Information Systems*. St. Gallen, Switzerland, 2007. Print.
- Laclau, Ernesto. *On Populist Reason*. London: Verso Books, 2010. Print
- Milloy, Caren. *JISC National Ebooks Observatory Project – Ebooks project first user survey a4 final version*. JISC. Print.
- Rowlands, Ian, and Maggie Fieldhouse. *Information Behavior of the Researcher of the Future. Trends in Scholarly Information Behavior*. A British Library / JISC Study, 2007. Print.
- Steele, Colin. "Phoenix Rising: New Models for the Research Monograph?" *Learned Publishing* 16 (2003): 111-122. Web.
- . "Scholarly Monograph Publishing in the 21st Century: The Future More Than Ever Should Be an Open Book." *Journal of Electronic Publishing* 11.2 (2008): n. pag. Print.
- Swan, Alma. *Key Concerns Within the Scholarly Communications Process. Report to the JISC Scholarly Communications Group*. Key Perspectives, 2008. Web. 25 March 2010.
- Swan, Alma, and Sheridan Brown. *Authors and Electronic Publishing: The ALPSP Research Study on Authors and Readers Views of Electronic Research Communication*. West Sussex, UK: ALPSP: Association of Learned and Professional Society Publishers & Key Perspectives Ltd., 2002.
- Tenopir, Carol, Rachel Volentine, and Donald W. King. "Article and Book Reading Patterns of Scholars: Findings for Publishers". *Learned Publishing* 25.4 (2012): 279-291. Web.
- Textual Analysis of Open-ended Questions in E-book National Observatory Survey*. UCL: CIBER, 2008. Print.
- Thompson, Jennifer Wolfe. "The Death of the Scholarly Monograph in the Humanities?" *Libri* 52.2002: n. pag. Print.

- Thompson, John B. *Books in the Digital Age: The Transformation of Academic and Higher Education Publishing in Britain and the United States*. Cambridge: Polity Press, 2005. Print.
- Waltham, Mary. *JISC: Learned Society Open Access Business Models*. JISC, 2005. Print.
- Williams, Peter et al. "The Role and Future of the Monograph in Arts and Humanities Research." *Aslib Proceedings* 61.1 (2009): 67-82. Web.
- Willinksy, John. "Toward the Design of an Open Monograph Press (DRAFT)." 28 September 2008: n. pag. Print.

The Future of Publications in the Humanities

Possible Impacts of Research Assessment¹

Milena Žic Fuchs

Introduction

During the last decade we have witnessed lively debates on two interlinked issues: how to ensure a higher level of visibility for Humanities research outputs, and how to establish assessment mechanisms for Humanities research. The aim of this paper is to give an overview of existing attempts in both directions, with special emphasis on possible impacts of evaluation mechanisms on publication cultures characteristic of the disciplines that traditionally comprise the research domain of the Humanities.² Discussions and proposals on the above issues, both in academic circles and in funding bodies across Europe, clearly indicate the need for providing more visibility for Humanities research. At the same time, many stress the need for better grounded assessment mechanisms. However, opinions on what is the best way to go forward vary a great deal and, in principle, not much is said about the possible effects such types of research assessments could have on the way Humanities scholars will publish in the future and to what extent evaluations of different kinds may affect the research output landscape.

The Context

Publication cultures or publication practices in various disciplines across the domains of research may change for a variety of reasons. Thus, technological breakthroughs and innovations – and here we primarily refer to various forms of internet publishing – have without a doubt left their mark on the way research results are “published” as well as on the level of their accessibility. The phenomenon of Open Access is more than worth mentioning, a phenomenon that has opened up many debates and issues. Developments in this new form of research output have no doubt already impacted publication cultures and may do so even more in the future.

However, this topic remains outside the scope of this paper, which deals with traditional research outputs in the Humanities.³

On the other hand, assessment mechanisms and systems used in disciplines across the domains of research can be said to have had a very specific impact on publication practices. However, to date, extensive, in-depth analyses showing more precisely what the effects that the evaluation, whether of institutions or individual researchers, has been on publication cultures, are still few and far between.⁴ One usually finds fragmentary data or insights on the effects of assessment mechanisms in a small number of disciplines, but comprehensive analyses covering all domains of research have yet to be produced; analyses that aim to showcase, from an “historical” perspective, how evaluation systems have impacted publication cultures over time.

Assessment mechanisms or instruments have, generally speaking, become “an accepted fact” especially in the so-called hard sciences. These instruments are often based in principle on the belief that research assessment should be done using simple and objective methods. These methods are most often subsumed under the term *bibliometrics*, referring more specifically to citation data and/or impact factors. The philosophy behind such assessment instrument is succinctly formulated in “Citation Statistics – A Report from the International Mathematical Union” (Adler et al. 3):

The drive towards more transparency and accountability in the academic world [...] has created a “culture of numbers” in which institutions and individuals believe that fair decisions can be reached by algorithmic evaluation of some statistical data; unable to measure quality (the ultimate goal), decision-makers replace quality by numbers that they can measure.

Although the prevalence of bibliometrics in assessment of scientific research is the dominant and only method used in many disciplines today, voices can be heard that doubt its often declared “objectivity.” Furthermore, arguments are put forward in favour of using multiple methods in research assessment (e.g. Martin 343-362, Carey et al. 84-89, Adler et al. 1-26). The gist of these criticisms directed towards the exclusive use of numbers in assessing research is also voiced in the above mentioned Report of the International Mathematical Union (Adler et al. 2): “The sole reliance on citation data provides at best an incomplete and often shallow understanding of research – an understanding that is valid only when reinforced by other judgments. *Numbers are not inherently superior to sound judgments.*”

Whether such pleas will result in more flexible approaches to assessment of research, especially in the so-called hard sciences, remains to be seen. At present, we are witnessing lively discussions as to the true worth of bibliometrics, especially the Journal Impact Factor approach, and much thought is being given to possible evolved systems of peer-review. Reflections on the many dilemmas facing evaluating research in all domains can be found in the January 2014 online issue of *EuroScientist* (Euroscientist, n. pag.). In this issue of *EuroScientist* one finds topics from “old” debates, such as the Journal Impact Factor, to discussions of altmetrics, which can be seen as a new “hot” topic in research evaluation. We also find the claim that a culture change in evaluation is in progress. If such a culture change does occur, then it is only natural that it will take time for scientists and scholars to adjust. Even more importantly, adjustments will have to be made by funding agencies, universities, as well as other institutions that rely on evaluation procedures. In this respect, a document that has triggered much discussion in leading journals, as well as the broader research community, is the San Francisco *Declaration on Research Assessment* (DORA). Many reactions to DORA reflect the main intent of that declaration; that is, the over-reliance on the Journal Impact Factor in research assessment. DORA, together with opinions provided by leading journals, indicate that at least a rethinking is under way, if not a radical culture change (Žic Fuchs 107-115).

The issue of quality of research in contrast to the bibliometric approach relying on quantitative measures has been repeatedly resurfacing in recent decades. In this context, it is worth quoting Nobel Prize Laureate for Chemistry Richard R. Ernst (Ernst 90): “We are deeply convinced that human ingenuity and creativity are beyond all conceivable quantitative measure. [...] The present hype of bibliometry made it plainly obvious that judging the quality of science publications and science projects by bibliometric measures alone is inadequate. [...] Start reading papers instead of merely rating them by counting citations!” How to “capture” quality in assessment procedures without a doubt should be the ultimate goal, but the question of course is how this can be achieved, especially in an environment of assessment practices that are still primarily geared towards citations and Journal Impact Factors. Issues pertaining to quality are sometimes taken a step further. Thus, for instance, Randy Schekman, Nobel Prize Laureate, brings into question the true role of the big three journals *Science*, *Cell* and *Nature* by accusing them of promoting their brands “in ways more conducive to selling subscriptions than to stimulating the most important research” (Schekman n.pag.).

Along with all of the above, new issues and views on what comprises evaluation are surfacing. Thus, not surprisingly, one encounters titles such as *Is*

there Currently a Scientific Revolution in Scientometrics? (Bornmann 647-648). Bornmann (647-648) argues that scientometrics is undergoing a taxonomic change; in other words, a revolution. This specifically applies to the notion of *impact*, which is increasingly understood more and more broadly, not simply implying scientific impact, but many different kinds of impact, especially those that come under the rubric of societal impact. Indeed, the notion and calibration of “societal impact” is becoming a pivotal issue, and is surfacing especially in institutional assessment exercises (Bornmann 647-648):

For example, the UK funding bodies have decided that the overall framework for assessing impact in the 2014 Research Excellence Framework (REF) is as follows: “The impact element will include all kinds of social, economic and cultural benefits and impact beyond academia, arising from excellent research, that have occurred during the period 1 January 2008 to 31 July 2013” (Higher Education Funding Council for England, 2011, p.4) [...] New tools for measuring impact will play an important role in measuring societal impact.

But yet again, assessing “new tools” for measuring societal impact opens up a myriad of questions as to what societal impact actually is and how indicators for assessing it can be determined. There are also warnings as to the potential dangers of using “social benefits” as an assessment criteria for basic research. More specifically, Henk Moed (Moed n.pg.) states that “social benefit” cannot be measured in a politically neutral way, and this view should be taken seriously.

Voices such as the above cited should at least be seen as a warning that assessing research is a far more complex endeavour than traditional bibliometrics based methods would have us believe. It is also a warning for the domain of Humanities research, that when and if bibliometric methods are introduced, they should be seen as only a part of the complexity of evaluating research and not its sole basis. In other words, assessment in the Humanities should at best be *bibliometrics-informed*, not *bibliometrics-led*.

The above brief overview of different opinions indicates, to say the least, that there are those that see the mainstream of assessment practices as being inadequate. Manifold criticisms of the “traditional” bibliometric methods of evaluation open up a multitude of questions, but time will only tell in which direction these lines of thinking will lead research evaluation in the future.

When we turn to the Humanities, an overview of the different publication outputs traditionally found in the various disciplines indicates a complex picture, particularly if viewed from a bibliometric perspective or from the perspective of the so-called hard sciences. The two fundamental

characteristics that mark this complexity are the diversity of publication outputs and the diversity of languages in which these publications appear.

The awareness that Humanities research output is not, as is the case in hard sciences, predominantly based on journal articles is slowly but surely gaining recognition. Scholars working in the different disciplines of the Humanities publish their research findings not only in the form of journal articles, but also in monographs, chapters in books, “collections” such as revised editions, collections of data, dictionaries, grammars, etc. In some disciplines excellence is measured by scholarly monographs; in fact, the monograph can be considered to be the golden standard and at many European universities the rank of Professor cannot be attained without at least one published monograph. The wide range of publication outputs is further “complicated” by the fact that all these different published forms appear in national languages. Furthermore, it can also include older variants of the language in question or various dialectal forms.

Needless to say, the roughly outlined complexity of Humanities research outputs has been at the core of discussions of the adequacy of evaluating research (whether of individuals or institutions) on the basis of WoS, SCOPUS or Google Scholar. It has often been stressed that they are primarily geared towards the hard sciences and thus primarily based on journals and, moreover, journals characterized by a strong bias towards the English language. Thus, it is not surprising to find claims that, for instance, the Arts & Humanities Citation Index (A&HCI) provides poor coverage for the Humanities (Leydesdorff et al. 5), or that classifications of proceedings papers and reviews that seem to work fairly well in the hard sciences illustrate a profound misunderstanding of research and publication practices in the Social Sciences (Harzing 23).

In line with the above, an excellent, visually clear overview that shows coverage of different domains of research in the Web of Science and Scopus, as well as the relationship between books and journals is provided by Sivertsen (“Appropriate Coverage of [...]”, n.pag.). The data refers to the coverage of scholarly publications in Norway from 2005 to 2012; however, the ratios and relationships between the different domains of research reflect what one could call general coverage by the Web of Science and Scopus, at least in rough approximation.

The specific nature of Humanities research outputs, visually reflected in Figure 1, has been repeatedly stressed by Humanities scholars as well as those bibliometric experts writing on these complex issues and with a deep understanding of the specific nature of Humanities research.⁵ The issue of the importance of books and chapters in books has also been repeatedly highlighted, or rather the fact that journal articles in some disciplines play a

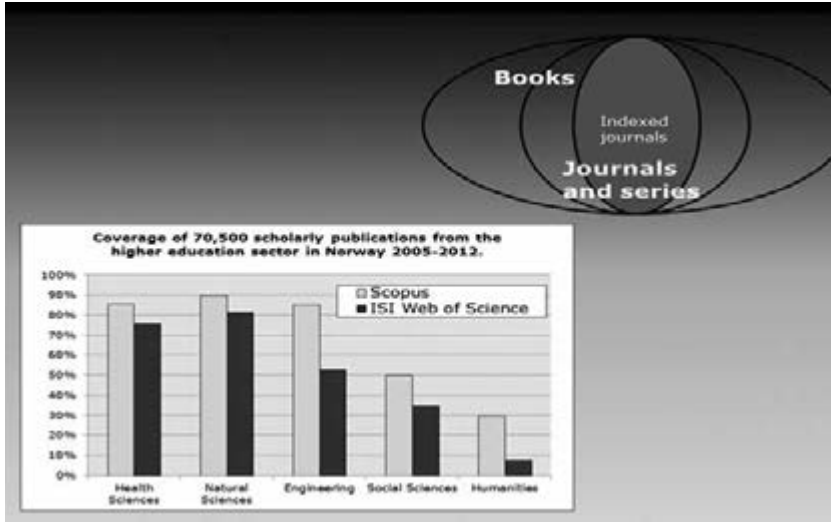


Fig. 1: Siversten (“Appropriate Coverage of [...]”, n.pag.).



Fig. 2: Siversten (“Appropriate Coverage of [...]”, n.pag.).

lesser role than books or book chapters. (Nederhof “Bibliometric Monitoring of Research Performance” 84-86; and Nederhof “Books and Chapters are not to be Neglected” 734-735).

Again, a visually indicative overview is provided by Sivertsen (“Appropriate Coverage of [...]”, n.pag.):

However, despite repeated warnings about the specific nature of Humanities research outputs, unfortunately one still finds a lack of understanding in the wider academic community of the issues involved and we still encounter attempts to evaluate Humanities research on the basis of databases that do not reflect their true nature, thus distorting research results. This, in turn, can have devastating effects for evaluation of individual scholars or institutions. It is important to realize that there is still a long way to go in order to reach, as Sivertsen and Larsen point out, “a well-designed and comprehensive citation index for the Social Sciences and the Humanities (which may have) many potential uses, but has yet to be realized” (567-575). Thus, more concentrated effort is needed to showcase the diversity of Humanities research outputs and their inability to fit neatly into standardized bibliometric frameworks used in other domains of science. However, in all fairness, the awareness of the differences inherent in the Humanities and Social Sciences publication cultures, in comparison with other domains, is growing. Both the WOS and Scopus are paying increasing attention to further developing the Social Science and Humanities (SSH) coverage in order to meet the demands of the above mentioned complexity (Moed et al. 10).

Apart from the above mentioned general trends, attempts to enhance the visibility of Humanities research, and in some cases also pave the way for possible evaluation systems for SSH, have appeared during the last decade or so. It is worth examining these endeavours and assessing them as possible “contributions” to future visibility or evaluation mechanisms. The aim of what follows is not an in-depth bibliometric analysis of Humanities research; rather, it is an attempt to address wider questions of a more conceptual kind, based on lessons learned from initiatives that have come from the Humanities and/or the Social Sciences. These initiatives stem from different sources, they have or had different aims and aspirations; nonetheless, they are instructive in conceptualizing possible future developments.

ERIH (European Reference Index for the Humanities)

In order to deal with the diverse nature of the Humanities research, in 2002, the European Science Foundation (ESF); specifically, the Standing

Committee for the Humanities (SCH), launched the *European Reference Index for the Humanities* (ERIH). From the very beginning,⁶ the main aim of ERIH was to identify and gain more visibility for top-quality European Humanities research in journal-based publications, regardless of language. In other words, the main aim was to showcase the multilayered diversity of Humanities research within disciplines and research communities in all European languages. It should also be clearly stated that, from the start, monographs, book chapters, edited volumes, etc. were planned to be included in order to fully reflect the nature of Humanities research outputs, their multilingual diversity as well as the diversity of research traditions.

Today, the ERIH can be seen as an historical fact; but, what is far more important is that it is, and should be seen as, a work in progress. More precisely, in 2007/2008 the Standing Committee for the Humanities published the so-called Initial Lists covering the following disciplines: Anthropology, Archeology, Art and Art History, Classical Studies, Gender Studies, History and Philosophy of Science, Linguistics, Literature, Musicology, Oriental and African Studies, Pedagogical and Educational Research, Philosophy, Psychology, Religious Studies and Theology.⁷

The Initial Lists covering the 15 disciplines are the outcome of a complex procedure that began with input from ESF Member Organizations, who provided initial proposals for journals to be included in the ERIH, according to the guidelines provided by the ERIH Steering Committee.⁸ Consequently, Expert Panels were set up and their remit was to define scope, analyze input and produce draft lists. The draft lists underwent a further round of consultations involving ESF member organizations, European-level as well as national subject associations and specialist research centres. On the basis of the feedback received, the lists were revised and then calibrated and harmonized by the ERIH Steering Committee and, finally, validated by the Standing Committee for the Humanities and the HERA Network Board.⁹

The above is an outline of the procedures behind the ERIH Initial Lists, procedures that involved literally hundreds of scholars from different disciplines from all parts of Europe. The brief outline does not do justice to the effort that has gone into the ERIH so far, nor does it provide true insight into the complexity of the system of procedural steps. However, it does show that ERIH has been primarily a *peer-review* driven process. Peer-review is at the heart of the Expert Panels' work and it also permeates other procedural steps in which scholars made key-value judgements.

Further details on how the Initial Lists were developed is of less interest than the reactions to the ERIH that erupted after the Initial Lists were published. These reactions can be classified as, on the one hand, extremely

loud criticisms and, on the other, a much quieter support for the project. Some of the criticisms stemmed from misunderstandings about the true nature and aim of the ERIH. Despite this, some of the criticisms had to be taken very seriously. Part of the criticism, whether based on misunderstanding or not, was subsequently incorporated into further development of ERIH on its course to 'phase two'; that is, the so-called Revised Lists. It is neither possible, nor necessary to cover all points and issues that were raised by some of the critics of the Initial Lists, but it is necessary to address the key ones because they can be seen as reflecting different views of the bibliographic-bibliometric divide.¹⁰ However, before we turn to what was one of the key issues, it should be stressed that, apart from critical views, the ESF Humanities Unit (which coordinated the ERIH procedures) as well as the ERIH Steering Committee received numerous messages that expressed positive views and letters from editors, publishers and researchers expressing their wish for their journals to be included in the ERIH lists. This fact is less known and was certainly overshadowed by the avalanche of criticism directed towards the ERIH Lists.

One of the most discussed issues concerning the ERIH was the categorization of journals into A, B and C categories. Criticism directed at this aspect of the ERIH, on one hand, can be seen as stemming from a misunderstanding of what the main aim of the ERIH is; on the other, it had the beneficial effect of triggering the rethinking and the renaming of categories. The main idea behind the renaming of categories was the prevention of possible misinterpretations about what the ERIH is really about, but also preventing further misuses, in the sense of using ERIH categories as a ranking system, a system supposedly reflecting "quality." The Initial Lists, published in 2007/2008, categorized journals into A, B and C, according to the following "definitions":

Category A: high-ranking international journals with high visibility that are regularly cited all over the world and have a strong reputation among researchers in different countries

Category B: standard international journals with significant visibility and influence in the various research domains in different countries

Category C: high-ranking national or regional journals with recognized scholarly significance

The above categorization was, unfortunately, perceived by many editors and researchers as a ranking, although ERIH documents and websites clearly stated that this was not the case. Furthermore, despite the explicit warning, some research councils and research funding bodies began to misuse the

ERIH Initial Lists, ignoring the proclaimed limitations of their use and seeing them as a ranking instead of a categorization based on audience, distribution and reach. Especially worrisome was the misinterpretation of Category C, which was interpreted by some as representing “low quality” journals, instead of understanding it as a category that identifies quality European journals but with limited circulation, primarily due to linguistic reasons. One can readily imagine the negative effects such misuses could have had (or had) if used for promotions, appointments and allocation of research grants. This possible misuse highlights the fundamental misunderstanding concerning the ERIH; that is, that its main aim was to increase the visibility of quality research in the Humanities and *not* to be the basis for an assessment mechanism, especially for evaluating individual research outputs.

The ERIH Steering Committee considered feedback on the A, B, C category names and with the intent of avoiding further misinterpretations and misuses introduced new category names. The purpose of this renaming was primarily to clarify the nature of ERIH categories, and the new names introduced appear in the Revised Lists.¹¹ The main dividing line between the categories was highlighted, i.e. the division between National Journals – NAT (formerly Category C) and International Journals – INT₁ and INT₂ (formerly categories A and B). The definitions are:

National Journals – NAT: European publications with a recognized scholarly significance among researchers in the respective research domains in a particular (mostly linguistically circumscribed) readership in Europe
International Journals – INT₁ and INT₂: both European and non-European publications with an internationally recognized scholarly significance among researchers in the respective research domains, and which are regularly cited worldwide.

Additional elaboration of the categories was published together with the Revised Lists. The above has been cited in order to clarify categorization/ranking issues and to stress that categories primarily indicate influence and scope.¹²

In the period following the publication of the Initial Lists, further steps were also taken by the ERIH Steering Committee: recomposition of Expert Panels on the basis of the panel membership rotation mechanism as well as the integration of online feedback from publishers, editors, European and national subject associations.¹³ Expert Panels revised the Initial Lists based on received feedback and publication of the Revised Lists began in the first half of 2011.

Although the expectations were that the Revised Lists would be a refined version of the Initial ones, –especially due to the fact that feedback from many quarters (both European and national) was taken into consideration – they were not perceived by the academic community as an improvement of the previous ERIH Lists. Once again, an avalanche of criticism was directed towards them. A succinct overview of the reactions can be found in the article “Outraged European Academics Resent “Rankings” in *The Guardian* (“Outraged academics [...]” n.pag.) Professor Michael Worton, Member of the Steering Committee of ERIH stressed again that the aim was never to create a ranking tool, but rather to make high-quality non-English language journals more visible (Ibid. n.pag.). However, the academic community failed “to see any difference in the new system, and [...] (claimed it to be) the same hierarchical ranking.” (Ibid. n.pag.) Worton further points out that the only hierarchy is whether the journal is on the list or not. He says: “The point that very few people seem to get is that being on the list, you already have got your gold star – you are a really good journal.” (Ibid. n.pag.)

The above explication should not be seen only as a clarification of basic issues concerning the ERIH and the criticisms directed towards it, but also as a reflection of the tension that exists between evaluation/assessment issues and scholarly outputs in the Humanities. The so-called misuses of the ERIH, at least in part, show the need often expressed by funding bodies, assessment committees, etc. for a European “evaluation tool” that could not only serve as a kind of “measuring” device but would also help overcome difficulties that arise from the diversity of Humanities research outputs. When one views the reactions to the Initial Lists, and later to the Revised Lists, the criticisms can be roughly seen as coming from two groups: one being those that are against any changes in the overall Humanities research assessment culture, and are especially against “bibliometric” tools and rankings of any kind.¹⁴ Voices such as these come mainly from the research communities, or more specifically from the researchers themselves, as well as from, of course, journal editors. On the other hand, the second group – that is, voices coming from funding and evaluation bodies – express strong opinions that the time has come to take evaluation in the Humanities to a higher level. From the ERIH experience, the question arises whether any kind of journal ratings (in the widest sense of the word possible) are the way forward in order to achieve evaluation criteria in the sense that funding and evaluation bodies would wish. The ERIH was not the only example of categorizing journals, and for insights into other attempts worldwide see “The Controversial Policies of Journal Ratings: Evaluating Social Sciences and Humanities” (Pontille and Torny 347-360).

With the winding down of the European Science Foundation that began in 2012 a major issue that had to be addressed was the future of the ERIH. Namely, member organizations were of the opinion that it was too valuable a database of Humanities' journals to simply disappear or be left in the state that it was in. An agreement was reached between the ESF and the Norwegian Social Science Data Service (NSD) to transfer the ERIH to Norway, and the NSD, together with the ESF, committed themselves to maintain and operate the ERIH database from December 2013 onwards. More importantly, the Standing Committee for the Humanities and its successor, the Scientific Review Group of the ESF, together with the Norwegian partners concurred in the signed Memorandum of Understanding that the "new" ERIH PLUS, as it is now being called, will not feature the previous ERIH categories. This fundamental feature of the so-called ERIH PLUS is a reflection of years of discussions, both within the Standing Committee for the Humanities and the Steering Committee of ERIH. These discussions, needless to say, reflect in many ways the voice of the academic community and show that all criticisms directed towards ERIH were taken very seriously into consideration. It should also be stressed that the removal of categories was motivated by the wish to avoid further misuses of the ERIH as an evaluation tool, and to reinforce the basic aim of fostering the visibility of Humanities research. Another fundamental feature of ERIH PLUS should also be noted, namely, that ERIH PLUS is now also open to journals from the Social Sciences (<https://dbh.nsd.uib.no/publiseringskanaler/erihplus/>).

The SPRU Report: Towards a Bibliometric Database for the Social Sciences and Humanities: A European Scoping Project

In the summer of 2008, a number of funding bodies ESRC/AHRC (UK), ANR (Fr), DFG (De) and NWO (Ne) commissioned a project entitled *Towards a Bibliometric Database for the Social Sciences and Humanities: A European Scoping Project*. It was led by Professor Ben Martin of the Freeman Center, SPRU, University of Sussex. The Report was finalized in March 2010 following regular meetings of the Board, two workshops and, most importantly, it was based on two mini-projects commissioned by the Board and written by Hicks and Wang (1-26) and Moed et al. (1-66)¹⁵.

The main aim of the SPRU Report was to facilitate the creation of a *bibliometric* database for the Social Sciences and Humanities, and the main aim is worth quoting in extenso since it clearly specifies the intended uses of such a bibliometric database:

Such a SSH bibliometric database is intended to serve five main purposes. The first is to permit the construction of indicators needed to demonstrate accountability with regard to the public funds devoted to research. The second is to develop indicators for assessing research excellence. Third, policy makers, research funders and others may wish to use the bibliometric database to provide an overview of SSH research outputs in Europe. Fourth, funders may use it as a means for assessing research capability and for identifying areas in SSH that may require capacity-building. A research output database might permit a ‘portfolio’ analysis of SSH research, contributing to decisions on resource allocation, particularly in the Humanities where there is often little relevant information. Fifth, research councils may use the information provided by the bibliometric database to map emerging areas of (often interdisciplinary) research with an aim of ensuring that they are adequately resourced.

Since the SPRU Report is available online there is no need to go into lengthy explications. However, it should be stressed that the SPRU Report takes into consideration the diversity of Humanities and Social Sciences research outputs as well as the importance of covering the languages of Europe in which these outputs are produced. From this point of view, it continues the main scope of coverage embodied in the ERIH, but it also goes a step further to a third kind of input; namely, the so-called “enlightenment literature,” which refers to those kinds of outputs from Humanities and Social Sciences research that are aimed at the general public rather than to academic peers. This is an ambitious addition and possibly difficult to achieve in practice, but is without a doubt a welcome addition to the traditional research outputs coverage.

The definition of *bibliometrics* is rather broadly conceived in this report and goes beyond “just citations” (*SPRU Report 2*), but is still in keeping with the five main aims set for a bibliometric database of this kind. If compared to the ERIH, the main difference lies in the basic aims. The ERIH aims primarily at enhancing the visibility of Humanities research, whereas SPRU is geared towards a bibliometric database that would serve as a basis for setting up evaluation mechanisms. The ERIH is focused on Humanities research outputs, whereas SPRU intended to cover both the Social Sciences and the Humanities.

The “evaluation/assessment” aim of the SPRU Report coincides with the fact that the project was launched and funded by some of the leading research councils and funding agencies in Europe. Thus it reflects, from their point of view, the need to set up an “accountability mechanism”; that

is, to develop evaluation tools for assessing research quality and impact. This is in contrast to the aims of the ERIH, no matter to what extent these aims were perceived as having been achieved.

The DFG (Deutsche Forschungsgemeinschaft) organized a workshop on the SPRU Report in Berlin in October 2009. Not surprisingly, most of the researchers present were extremely critical of the report on a general level, opposing the idea of any bibliometric tool envisaged for the future. From a more “technical” viewpoint, as a member of the SPRU Board, I stressed a number of points that, in my opinion, still had to be addressed before such a bibliometric project could be finalized even in its conceptual form. Here, I present two of them: 1) the active role of researchers was downplayed in the standard-setting body;¹⁶ and 2) the timeframe for realizing the project was extremely short (six months for journals, one year for books).

In the Final Report (March 2010), these points were taken up and modified to a certain degree. Thus, it was suggested that the standard-setting body be comprised of bibliometricians *and researchers*, but not all disciplines were to be represented. This, in my opinion, still presented a danger because disciplines within the Humanities differ in the order of the importance of research outputs, or the significance certain outputs have within the different traditions of the disciplines. Although extended in the Final Report, the timeframes remained very short (approximately 12 months). To my mind, this was a constraint that could potentially cause unnecessary damage. If a bibliometric database, as described by the SPRU Report, should emerge in the future, it has to be conceived and built as a very solid structure with foundations deeply rooted in the myriad of diversity inherent in both the Humanities and the Social Sciences. The pressure of time and lack of involvement of the research communities themselves could result in oversights with respect to specific disciplines, or of specific national traditions in the disciplines involved.

The explication of the above critical remarks primarily has the intention of showing how difficult and complex any attempt of setting up a bibliometric database is and shows that great delicacy must be applied to all matters relevant for setting up such an infrastructure. Establishing such an infrastructure implies other issues as well, some of which were also brought up at the Berlin workshop. One of these was related to the possible impact of bibliometric tools on publication cultures. And because of the “worrying possibilities” such an endeavour could bring about, the question was put forward whether the Humanities and Social Sciences can be “preserved” from bibliometric assessment tools? This was very strongly endorsed by the representative of the German historians. Such tools or

assessment mechanisms at national levels already exist or are being set up in some countries, (e.g. Norway, Croatia, etc.) or are planned in others. Bibliometric assessment on international or European level seems to be knocking at the door, and one can only wonder if it can, in the long run, be avoided. The SPRU Report has, so far, not generated concrete action in terms of setting up a bibliometric database at European level, but this does not mean that an endeavour, based on the SPRU Report, or some other similar concept, will not become a reality at some point in the future. If this does occur, then it is to be expected that changes will occur in certain aspects of publication cultures, and possibly even research practices.

Both the ERIH and the SPRU Report seen from the perspective of their declared aims also raise questions on priorities: what is more important – visibility mechanisms for Humanities research outputs, or tools that could be used for evaluating research quality and impact? Or does the possibility exist of combining both in such a way that both the needs of researchers and the funding bodies can be fulfilled? There are no easy answers to these questions, and the experience gained both from the ERIH and the SPRU Report should be harnessed in order to pave the best way forward. But whatever happens in the future in this respect, especially within the context of a major rethinking of “traditional bibliometrics,” great care should be taken not to generate unwanted changes in publication cultures, changes that would push researchers and disciplines in possibly unwanted directions.

Assessment Mechanisms and Their Possible Impacts on Humanities Research Outputs

As already stated in the Introduction, documented examples of how assessment tools affect publication culture are not numerous. However, it is interesting to see at least some examples available in order to foresee possible impacts of evaluation mechanisms for the Humanities.

The BMJ (British Medical Journal) in April 2007 brings insights from a group of medical researchers on the “Impact of the Impact Factor in Spain” (López-Cózar et al. n. pag.). Namely, when the national system of evaluating researchers’ production was implemented in 1989, a system in which a bonus was awarded to articles published in journals of recognized prestige in the lists of science fields in the Subject Category Listings of the Journal Citation Reports of the Science Citation Index (ISI), changes in publication behaviour as well as research practices became evident.

Although the reward system had the aim of improving the quality and international visibility of Spanish science the following effects surfaced in a relatively short space of time:

- a) the most immediate effect was mass migration of the best research articles to foreign journals;
- b) the system resulted in increasing neglect of Spanish journals to which researchers rarely submitted their work;
- c) the negative effect of deterioration (the authors say destruction) of Spanish as a language of science became apparent;
- d) many research groups altered their research agendas – often replacing local, regional and national topics with research topics more likely to be received by international journals.

The same kinds of effects produced by bibliometric tools have been reported for economics in Spain. Carmona et al. (5-22) report that the implementation of the national assessment system in economics has resulted in a major shift to international journals indexed in WoS, or national journal lists, showing the same effects on language and research practices as in medicine.

In Croatia, medical journals are more and more frequently published in English resulting in effects on Croatian as the language of medical science. Since researchers prefer submitting papers either to international journals or national Croatian journals published in English, the few remaining Croatian medical journals published in Croatian receive, in principle, papers of poorer quality.

As far as the Humanities and Social Sciences are concerned, concrete, data-based analyses of possible changing publication patterns at national (or language) level have become even more relevant and interesting. In this respect, it is worth stressing the work of Engels et al. (1-18) and Verleysen and Engels (110-143) based on VABB-SHW, a full coverage database of peer-reviewed publication outputs in the Humanities and Social Sciences developed for the region of Flanders, Belgium as part of the Flemish performance-based funding system for University research. For instance, Engels et al. (1-18) in their analysis show considerable differences across the disciplines of the Humanities and Social Sciences in terms of coverage by the Web of Science, but also a steady increase in the number and proportion of publications in English – again varying in percentage from discipline to discipline. The increase in the number of publications in English is, at the same time, reflected in the decrease of publishing in Dutch and other languages. Interestingly, no overall shift away from publishing research in book form was observed for this community of researchers. Verleysen and

Engels (110-143), on the other hand, analyze on the basis of VABB publication patterns in History of Flemish Universities from 2000 to 2009. Their analyses show that compared to other disciplines “[...] historians often publish in Dutch and they attach a greater importance to book publication as opposed to journal articles.” (Verleysen and Engels 141).¹⁷ What analyses such as these show is the diversity found between various disciplines and their publication patterns under the same performance-based funding system for university research. They also show the need to develop full-coverage regional (national) databases for different research communities of different disciplines in different languages. One can only agree with Henk Moed’s view that much more research is necessary in order to truly understand the complexities inherent in the Social Sciences and the Humanities, as well as his warning that statements about the characteristics of SSH as a whole should be avoided (“Research Assessment in SSH” n. pag.)

The above indicates that evaluation systems, whether national or international, can impact publication behaviour, change priority of research topics and even have effects on research language and development of terminology. The possibility of such changes is not only worrisome to scholars in the Humanities and Social Sciences but also to researchers in other domains of research. As already mentioned, more and more frequently we see articles appearing in which authors discuss negative effects of bibliometrically based assessment systems. Questions are being raised about what *metrics* truly show in terms of quality assessment, what their impact is on research practices, and how to deal with these issues in the future. Indicative of this “reassessment” in the hard sciences are six articles published in the *Opinion* section of *Nature* in June 2010.¹⁸ Six commentators present their views on what changes in the metrics systems might ensure that individuals are assessed more fairly. Worth quoting from these commentaries are the following:

John D. West: “Giving bad answers is not the worst thing a ranking system can do – the worst thing is to encourage bad science.”

David Pendelbury: “Metrics are an aid to decision-making, not a shortcut to it.”

But even more interesting in the light of the topic of changing publication cultures is the editorial published in *Nature* in February 2010,¹⁹ entitled *Back to Books*, in which a plea is made to researchers in the hard sciences to go back to writing books. The following argumentation is put forward:

The expansiveness of a book allows sophisticated arguments to be put forward and widely debated; new ideas that cross disciplinary boundaries can more readily be shared and worked through. But if this exhortation is to have any traction, the effort and skill required to write a book needs to be rewarded in the career recognition of scientists who devote time to mastering the art to good effect – a recognition that is commonplace in the social sciences and humanities. It is time to bring the book back into the science mainstream.

I sincerely doubt that Humanities scholars would welcome seeing an editorial entitled “Back to Books” in a prestigious journal at some point in the future.

Possible Future Approaches and Developments

The main aim of the previous sections was not simply to describe or discuss various criticisms that have been directed towards the ERIH Lists or the SPRU Report. The previous discussion only stresses the still very open question of “visibility mechanisms” and “assessment mechanisms” in the Humanities and their possible effects on publication cultures. Both “visibility mechanisms” as well as “assessment mechanisms” can impact the overall research output landscape and, through this, also influence and trigger changes in publication cultures. The question arises whether such developments can be avoided in cases when researchers, or communities of researchers, see them as negative, unwanted developments reflecting influences outside the disciplines themselves, and, their respective ways of producing and publishing research results. The ideal would be to allow unhindered development of publication cultures, development reflecting intra-disciplinary changes and needs for possible “new” outputs and, if possible, to avoid externally induced developments such as those that, for instance, can be called *metrics* induced changes. Whether this is possible in the context of the contemporary push towards evaluation and assessment of research remains an open question.

In this respect, it is interesting to note that the impact of the ERIH Initial Lists as a visibility mechanism has been documented and analyzed. Moed et al. (“Spru Report” n. pag.), for instance, show that the ERIH Initial Lists have had an impact on the overall coverage of Humanities journals in WoS and SCOPUS. More specifically, during 2008/2009 WoS included 1500 so-called “regional journals”, while SCOPUS increased its coverage by 2250

so-called “regional journals.” However, the above numbers refer primarily to journals from the “A” and “B” categories of the ERIH Initial Lists, while “C” or “national” journals lagged far behind.

On one hand, the addition of a large number of Humanities journals in WoS and SCOPUS can be seen as enhancing the visibility of Humanities research outputs through journals in well-known and well-established databases. On the other hand, since the greater number of journals included do not come from the National Journals Category (or “C Category”), we are again faced with the bias of high-ranking journals primarily published in English. Once again, such a selection of journals provides only a partial view into the diversity (linguistic and otherwise) of research production in the Humanities throughout Europe.

The question is, of course, what other avenues could be pursued in order to achieve an even higher degree of visibility for Humanities research outputs. In order to address such larger issues, on 31 March 2010 a meeting was held in Brussels on “The Future of ERIH.”²⁰ Under this title, major issues connected to visibility and assessment in Humanities and Social Sciences were discussed. The meeting concluded with a decision to set up a small working group to provide a short report with recommendations on the larger issues.²¹ The resulting report entitled “*Towards comprehensive bibliographic coverage of the scholarly literatures in the Humanities and Social Sciences*” was completed and distributed to the Member Organizations of the ESF in the second half of 2011.

Without going into a detailed account of what the report contains, what must be stressed is its main approach to Humanities and Social Sciences outputs encapsulated in the term *bibliographic*. More specifically, this means a move towards an all-encompassing, integrated European infrastructure that would provide comprehensive bibliographic coverage of Humanities and Social Sciences outputs at European level. At present, full-text databases of publications exist at universities, research centres as well as on the so-called national level. However, these databases present, in most cases, disjointed, non-integrated infrastructures that only reflect small segments of research outputs in different languages and different disciplines.²² Their integration into a searchable and accessible database would provide deep insight into the diversity and richness of scholarly production in the languages of Europe, especially echoing different research traditions in different disciplines. The report goes a step further and outlines a possible way forward for an assessment/evaluation tool that could be developed on the basis of such a broad, all-encompassing database.

The above is a rough outline of the ideas put forward in the report and, needless to say, it will almost certainly trigger new discussions and raise new questions in the research communities about the various disciplines of the Humanities and Social Sciences throughout Europe.

At this point, I would like to stress my personal opinion; that is, that European Humanities and Social Sciences are in need of an integrated infrastructure that would finally showcase the richness and diversity found across all disciplines and all languages covered by these domains of research. Apart from the fact that such a comprehensive bibliographic coverage would enhance integrated, full-text visibility of research outputs it would also:

- produce new synergy in research, connect scholars within disciplines and across disciplines throughout Europe;
- give insights to researchers from different domains, open up avenues for collaboration between domains, and pave the way for multi- and transdisciplinarity, especially connected to the so-called *Societal Challenges* encapsulated in *Horizon 2020 – The Framework Programme for Research and Innovation (2014-2020)*;
- make visible European research at global level, especially necessary for addressing major issues within SSH disciplines themselves, but also issues pertaining to wider topics of global dimension and importance.

Finally, such a comprehensive, integrated database would not only provide true insight into publication cultures characteristic of specific disciplines by simply showcasing them, but would achieve this without exerting external pressures on their future development.

At this point, the idea of a comprehensive database of European research outputs in the Humanities and Social Sciences is only an idea. It should be discussed and further elaborated upon, but always keeping in mind the specific nature of publication cultures that characterize the disciplines of these two domains of research. To my mind, only after such a database is established and implemented, can one seriously turn to developing possible European level “bibliometric” tools for the future. The time is ripe to bring together all those voicing opinions on the necessity of developing and integrating databases. The time is ripe because there is quite a high level of agreement on this issue amongst those that have SSH visibility and evaluation at heart. It is also time not only to encourage the systematic development of databases from the institutional level to the national, and then to the European, but to exert pressure on funders at all levels in order to achieve the high-level aims declared for the European Research Area at present.

Notes

1. My sincerest thanks go to Jadranka Stojanovski, University of Zadar, Croatia, for her collaboration and help in preparing the presentation preceding this paper (ESF Strategic Workshop “Changing Publication Cultures in the Humanities”, Hungarian Academy of Sciences, 2009).
2. Although the Humanities comprise the main focus of this paper, it should be noted that much of what follows can also be relevant for the Social Sciences, although it should be stressed that direct correspondences and especially generalizations should be avoided due to the diversity of publication practices in various disciplines.
3. For issues relevant to Open Access, see Adema and Ferwerda, this volume.
4. One notable exception in this respect is the work of Gunnar Sivertsen and colleagues (Norwegian Institute for Studies in Innovation, Research and Education) who have, for Norway, as well as other Scandinavian countries produced cross-domain analyses of research assessment tools and publication outputs - see for instance (Schneider, (ed.) 1-56), (Sivertsen n.pag.).
5. For an excellent overview and detailed analysis see Moed et al. (1-66) and Sivertsen (“Appropriate Coverage of [...]”, n.pag.).
6. The foundations of what was to become the ERIH were laid in 2001, at a workshop held in Budapest. This workshop was organized with the purpose of discussing the problem of the low visibility of European Humanities research. Its aim was not to produce an additional tool for research assessment, as for instance claimed by Pontille and Torny (4).
7. The following disciplines were also under consideration: Archives, Library and Museum Studies, Media and Cultural Studies, Area Studies.
8. All journals included in the Initial Lists had to meet threshold standards: peer-reviewed selection of articles, active editorial board, openness to unsolicited contributions, publication on time and to an agreed schedule, ISSN numbers as well as other bibliographic requirements.
9. The ERIH lists are a project jointly sponsored by the ESF and the European Commission ERA-NET project “Humanities in the European Research Area” (HERA).
10. For criticisms along these lines see: “Journals Under Threat: A Joint Response From History of Science, Technology and Medicine Editors”, (Andersen et al. 1-4) as well as “Classer, évaluer” (1-4).
11. The Revised Lists were published in 2011 and included the following disciplines: Archaeology, Gender Studies, Linguistics, Literature, Musicology, Paedagogical and Educational Studies, Philosophy, Philosophy of Science, Psychology and Religious Studies. The revised list for Classical Studies was published in 2012, and the Art and Art History revised list in the first part of 2013.
12. Needless to say, other points concerning the ERIH were also criticized; for instance, criteria for the selection of expert panel members, the efficiency

of National Contact Points, the question as to whether peer-review was the best approach, etc. Such issues as a rule provide a basis for differences of opinion, but this does not mean that they have not been addressed in the further development of the ERIH. Of particular importance is the challenge of keeping a balance amongst possible theoretical approaches within disciplines, as well as making sure that subdisciplines are either not under-represented or possibly left out completely, a point that should be seriously considered in the future.

13. According to the ESF Humanities Unit, 84 panel members from 25 countries have discussed 3541 feedbacks.
I would like to thank Nina Kancewicz-Hoffmann, Head of the Humanities Unit, ESF, for input on ERIH procedures and numerical data.
14. Hug et al. address, among other issues, how to deal with scholars' opposition to various methods employed in research evaluation (1-24).
15. The final version of the report (March 2010) can be found at www.vandenbesselaar.net/_pdf/2010%20ESF.pdf. It also contains both mini-projects.
16. Namely, in the so-called August 2009 version of the SPRU Report, the standard-setting body responsible for setting up the main parameters of the bibliometric database was to be comprised only of bibliometricians and library and documentation experts.
17. For differences between the Social Sciences and the Humanities in the realm of book publishing see Verleysen and Engels (2014, 234-240) as well as Verleysen et al. (117-127).
18. *Nature* 465.17 June 2010: 870-872.
19. *Nature* 463 February 2010: 588.
20. The meeting was attended by representatives of 26 Member Organizations, the Chairs of the Standing Committees for both the Humanities and Social Sciences of the ESF, as well as by ESF's Director of Science and Development, Marc Heppener.
21. The working group members were: István Kenesei (Hungary), Pascal Perrin (Belgium), Nigel Vincent (UK), Gunnar Sivertsen (Norway – Chair), Sir Roderick Floud (Chair of SCSS), Milena Žic Fuchs (Chair of SCH), Marc Heppener (Director of Science and Strategy Development, ESF), Nina Kancewicz-Hoffmann (Head of Humanities Unit of ESF), Balázs Kiss (Head of Social Sciences Unit of ESF)
22. Along the same lines, Moed also advocates the stimulation and further development of institutional repositories with the final aim of establishing comprehensive national databases (Moed et al. 10).

Bibliography:

- Adema, Janneke and Eelco Ferwerda. "Publication Practices in Motion. The Benefits of Open Access Publishing for the Humanities". This volume.
- Adler, Robert, John Ewing, and Peter Taylor. *Citation Statistics: A Report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)*. Joint Committee on Quantitative Assessment of Research. (2008): 1-26. Print.
- Andersen, Hanne et al. "Journals under Threat: A Joint Response from History of Science, Technology and Medicine Editors." *Science in Context* 22.1 (2009): 1-4. Print.
- Bornmann Lutz. "Is There Currently a Scientific Revolution in Scientometrics?" *Journal of the Association for Information Science and Technology* 65.3 (2014): 647-648. Print.
- Carey, Alan L., Michael G. Cowling and Peter G. Taylor. "Assessing Research in the Mathematical Sciences." *Gazette of the Australian Math Society* 43.2 (2007): 84-89. Print.
- Carmona, Salvador et al. "From Zero to Infinity: The Use of Impact Factors in the Evaluation of Economic Research in Spain." *IE Working paper* (2005): 05-22. Print.
- "Classer, évaluer", *Annales. Histoire. Sciences Sociales* 6/2008 (63 année), p. 1-4. January 2014. www.cairn.info/revue-Annales-2008-6-page-1.htm.
- Engels, Tim C.E., Tryken L. B. Ossenblok, Eric H. J. Spruyt. "Changing Publication Patterns in the Social Sciences and Humanities, 2000-2009." *Scientometrics* (2012): 1-18. Web. January 2014. www.ecoom.be/sites/ecoom.be/files/downloads/Engels%20et%20al%20changing%20pub%20patterns%20SSH%20Scientometrics%202012.pdf.
- Ernst, Richard R. "The Follies of Citation Indices and Academic Ranking Lists". *CHiMiA* 64.1/2, (2010): 90. Print.
- EuroScientist. "Special Issue on Research Evaluation." *EuroScientist*. 2014. Web. 29 January 2014. <http://euroscientist.com/2014/01/special-issue-on-research-evaluation/>.
- Ha, Tam Cam et al. "The Journal Impact Factor: Too Much of an Impact?" *Annals Academy of Medicine* 35.12 (2006): 911-916. Print.
- Harzing, Anne-Wil. "Document categories in the ISI Web of Knowledge: Misunderstanding the Social Sciences?" *Scientometrics* 94. (2013): 23-34. Web. January 2014.
- Hicks, Diana and Jian Wang. "Towards a Bibliometric Database for the Social Sciences and Humanities: A European Scoping Project." *Mini Project for SPRU Report*. USA: School of Public Policy, Georgia Institute of Technology. (2009): 1-26. Web. January 2014. www.vandenbesselaar.net/_pdf/2010%20ESF.pdf.
- Hug, Sven E. et al. "A Framework to Explore and Develop Criteria for Assessing Research Quality in the Humanities." *International Journal for Education Law and Policy* (in press). Web. January 2014. www.google.hr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&ved=0CEMQFjAE&url=http%3A%2F%2Fwww.sagw.ch%2Fdocs%2Fsaagw%2Flaufende_projekte%2Fwissenschaftskultur_GW%2FIJELP_in_press&ei=GwUTU-PfAcbm7AbLioEw&usq=AFQjCNGLLZZOdzXMfeZKWP2ENw4-3xHQHw&bvm=bv.62286460,d.ZGU.
- Leydesdorff, Loet et al. "The Structure of the Arts & Humanities Citation Index: A Mapping on the Basis of Aggregated Citations among 1,157 Journals." *Journal of the American Society for Information Science and Technology* 62.12. (2011): 2414-2426. Web. January 2014.
- López-Cózar, Emilio Delgado et al. "Impact of the Impact Factor in Spain". *British Medical Journal*. Web. January 2014. www.bmj.com/rapid-response/2011/11/01/impact-impact-factor-spain.
- Martin, Ben R. "The Use of Multiple Indicators in the Assessment of Basic Research." *Scientometrics* 36.3 (1996): 343-362. Web. January 2014. <http://sci2s.ugr.es/hindex/pdf/Martin996.pdf>.

- Moed, Henk F. et al. "Options for a Comprehensive Database of Research Outputs in Social Sciences and Humanities." *Mini Project for SPRU Report*. The Netherlands: Centre for Science and Technology Studies, Leiden University. (2009): 1-66. Web. www.vandenbesselaar.net/_pdf/2010%20ESF.pdf.
- Moed, Henk F. "Research Assessment in Social Sciences and Humanities." *ECOOM Colloquium Antwerp*. Belgium: Centre for Research & Development Monitoring. (2011). Web. January 2014. www.ecoom.be/sites/ecoom.be/files/downloads/1%20Lecture%20Moed%20Ecoom%20Antwerp%209%20Dec%202011%20SSH%20aangepast%20%282%29.pdf.
- Nederhof, Anton J. "Books and Chapters are not to be Neglected in Measuring Research Productivity." *American Psychologist* 44 (1989): 734-735. Print.
- Nederhof, Anton J. "Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: A Review." *Scientometrics* 66.1, (2006): 81-100. Web. January 2014. <http://link.springer.com/article/10.1007%2Fs1192-006-0007-2#page-1>.
- "Outraged European Academics Resent "Rankings"." *The Guardian*. 27 June 2011. n.pag. Web. January 2014. www.theguardian.com/education/2011/jun/27/journals-index-angers-european-academics.
- Pontille, David and Didier Torny. "The Controversial Policies of Journal Ratings: Evaluating Social Sciences and Humanities." *Research Evaluation* 19.5. (2010): 347-360. Web. January 2014.
- Rey-Rocha, Jesús et al. "Some Misuses of Journal Impact Factor in Research Evaluation." *Cortex* 37-4. (2001): 595-597. Print.
- San Francisco Declaration on Research Assessment*. (2012). Web. January 2014. www.ascb.org/SFdeclaration.html.
- Schekman, Randy. "How Journals like Nature, Cell and Science are Damaging Science." *The Guardian*. 9 December 2013. Web. 9 Jan. 2014. www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science.
- Schneider, Jesper W. (ed.) "Bibliometric Research Performance Indicators for the Nordic Countries." *The Use of Bibliometrics in Research Policy and Evaluation Activities* (2010). NordForsk/NORIA-net. Web. January 2014. www.nordforsk.org/no/publikasjoner/bibliometric-research-performance-indicators-for-the-nordic-countries.
- Seglen, Per O. "Why the Impact Factor of Journals Should not be Used for Evaluating Research." *British Medical Journal*, (1997): 314-497. Print.
- Sivertsen, Gunnar and Birger Larsen. "Comprehensive Bibliographic Coverage of the Social Sciences and Humanities in a Citation Index: An Empirical Analysis of the Potential." *Scientometrics* 91.2 (2012): 567-575. Print.
- Sivertsen, Gunnar. *Appropriate Coverage of Scholarly Publishing in the Social Sciences and Humanities*. September 2013. PDF file. Web. January 2014. www.google.hr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CDMQFjAB&url=http%3A%2F%2Fwww.lituanistikadb.lt%2Fdownload%2F216%2F6.gunnar_sivertsen.ppsx&ei=sOQRU6-3G6iv7Aag44CgCQ&usg=AFQjCNHR4Ia7vl-YO4d45bgAv_yx2ShhCpg&sig2=QHPyT87yDU10LWPc_eevvg&bvm=bv.62286460,d.ZGU.
- SPRU Report: "Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project." 2010: 1-55. Web. January 2014. www.vandenbesselaar.net/_pdf/2010%20ESF.pdf.
- Verleysen, Frederik T. and Tim C.E. Engels. "Historical Publications at Flemish Universities, 2000-2009." *Journal of Belgian History XLII* 4. (2012): 110-143. Print.

- Verleysen, Frederik T. and Tim C.E. Engels. "Barycenter Representation of Book Publishing Internationalization in the Social Sciences and Humanities." *Journal of Infometrics* 8. (2014): 234-240. Print.
- Verleysen, Frederik T. et al. "The Objectives, Design and Selection Process of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW)." *Bibliometrics, Use and Abuse in the Review of Research Performance*. (eds.) Wim Blockmans, Lars Engwall, Denis Weaire. Portland Press, 2014: 117-127. Print.
- Žic Fuchs, Milena. "Bibliometrics: Use and Abuse in the Humanities." *Bibliometrics, Use and Abuse in the Review of Research Performance*. (eds.) Wim Blockmans, Lars Engwall, Denis Weaire. Portland Press, 2014: 107-115. Print.

ERIH's Role in the Evaluation of Research Achievements in the Humanities

Ferenc Kiefer

My paper will be organized as follows: In the first part I will say a few words about ERIH, its aims and its possible role in the evaluation of research in the Humanities. In the second part I will discuss the main differences in publication culture between sciences and Humanities, and summarize the main changes in publication practices in the Humanities that have occurred during the last few decades. In the last part I will take up the problem of bibliometrics as an evaluation tool in Humanities research.

ERIH: Its Brief History, Scope and Aims

ERIH stands for "European Reference Index for the Humanities." It is an ESF project that was conceived as constructing a database of journals aiming initially at identifying top-quality European research in the Humanities, published in academic journals in all European languages. The background for launching this project in 2001 was the need felt by European researchers for better databases than the existing ones, which are limited in coverage (even some of the best English language journals are missing from the database and there are even bigger gaps in the coverage of journals in other European languages) and tend to be centred on Anglo-American publications. Research conducted in national (especially in so-called lesser used) languages is either not adequately covered or not covered at all. Furthermore, there are specificities in the Humanities in terms of cultures of publication and traditions of citations, which make it meaningless to work with the evaluation tools used in science. ERIH's main objective was to remedy this defect by providing a more reliable tool for research assessment in the Humanities.

As a first step towards a more adequate database, lists of journals in fourteen disciplines have been compiled and classified into three categories. The international journals include two subcategories, which differ with respect to influence and scope. The first subcategory, called A, covers high-

ranking international journals with high visibility that are regularly cited all over the world and have a very strong reputation among researchers in different countries. The second subcategory, called B, includes standard international journals with significant visibility and influence in the various research domains in different countries. Finally, national or regional journals include standard and high ranking national publications with a recognized scholarly significance among researchers in a particular (typically linguistically circumscribed) readership group in Europe. The main target group of national journals is the domestic academic community. In the case of national journals only European journals are considered. (For further details see the webpage www.esf/erih).

In principle, there should be no qualitative difference between articles published in international and national journals, the difference has to do with the language and with the distribution and the visibility of the journal. That is, the categories A/B and C are not meant to represent any ranking. A national/regional C-category journal may be of the same quality as an international A or B journal. In most disciplines it is expected, however, that scholars publish both in national and international journals. On the other hand, consistently high-quality scholarly content is more typical of A than of B journals and A journals are certainly more visible than B journals. It must be stressed, however, that a paper published in a B journal need not be qualitatively inferior to A-category publications. The ranking concerns the journals and not the articles; yet, the overall higher quality of articles in A journals does provide a certain ranking. The extent to which we accept this ranking will depend not only on our subdiscipline but also on our theoretical bias. But, whatever ranking we accept, it cannot and should not replace peer reviews.

In order to avoid the interpretation of the categories A/B and C as a ranking of the scholarly output, these categories have been replaced more recently by the categories INT₁ (=international 1), INT₂ (=international 2) and NAT (=national). The difference between INT₁ and INT₂ primarily concerns visibility: papers published in a journal categorized INT₁ are more visible than the ones categorized INT₂.

Aspects of quality also become important in the case of the criteria used to decide whether a journal qualifies for inclusion in an ERIH list or not. For all categories, in order to be included, a journal must fulfil normal academic standards of quality, i.e. it must have a quality control policy that governs the selection of articles and it must also fulfil basic publishing standards, such as ISSN, timeliness of publication, active and international editorial board for international journals and an active editorial board for national journals. In

fact, ERIH has already contributed a lot to the improvement of the quality of C (=NAT) category journals, which is, no doubt, one of ERIH's major merits.

The judgement to which category a journal belongs is made by an Expert Panel of peers for each discipline. Great care was taken to select solid scholars with an international reputation as panel members, and to cover as many sub-fields and linguistic areas as possible. It must be admitted, however, that in most human disciplines there has often been a strong bias in favour of English journals, which must be remedied in the future.

As already mentioned, the lists were meant to serve as a database of journals aiming initially at identifying top-quality European research in the Humanities published in academic journals. The lists as they stand are not a bibliometric tool and should not be used as such. But, we believe that it is necessary that peer review in the Humanities be informed by better data. One of the great advantages of ERIH is that the database also includes European research output published in lesser used languages. Such a database did not exist before. It is believed that this approach will enable prompt comparability with other sciences.

To be sure, there exist other quality-based journal databases, such as ISI Thomson and Reuter's Arts and Humanities Citation Index (thomson-reuters.com/web-of-science-core-collection) or, more recently, Scopus (cf. www.scopus.com/scopus/home.url). To date, Scopus is the largest abstract and citation database of peer-reviewed literature, which features smart tools to track, analyze and visualize research. However, not aiming at the broad European Humanities constituency, such databases have not included multilingual European research output in the global context of quality assessment and access. Moreover, their tools tend to be parameterized to the rhythms of knowledge decay in most natural sciences. In the Humanities, the lifecycle of research is much longer than in hard sciences since earlier findings are often reinterpreted rather than superseded and standard works may never become outdated.

So far we have not mentioned the problem of monographs, which – at least in some disciplines in the Humanities – still constitute the most important publication genre. Until now, ERIH has concentrated its efforts to survey and classify journals but the need to include monographs has always been acknowledged. Reviewing and classifying journals is already a difficult task, but to include monographs into this procedure is even more challenging. No doubt, the categories national/regional and international will play a role in classifying monographs as well. The evaluation of books will have to take into consideration at least the following aspects: the international or national prestige of the publisher, the scholarly standing

of the series editor or of the edited volume, and distinction will have to be made between conference proceedings, Festschriften, books whose target readership is not the scholarly community, etc. Furthermore, it must be taken into account whether the publisher requires anonymous reports of the manuscript submitted for publication. Finally, the published reviews of the monograph, too, must be part of the evaluation. Classifying books is thus radically different from classifying journals and it requires a very special methodology. The elaboration of such a methodology will be the task of future research.

Changes in Publication Culture

Since the late 1970s and early 1980s, the classical “von Humboldt” university model, characterized by learning through science and unity of research and teaching, suffered gradual transformation under the influence of changes in social, economic and technological conditions. In addition to the classical mission of knowledge creation and conservation for the next generations and the education of tomorrow’s elite, society gave new tasks to universities. Nowadays, academia is called on to play a more active role both in solving new societal problems and in strengthening economic development. Concrete goals are often formulated in management contracts together with performance indicators to measure their realization.

The publication culture has undergone drastic changes since the beginning of the last century. Up to the mid-twentieth century, Humanities was depicted as disciplines dominated by books, which rely largely upon older literature. In Europe the main publication languages in the Humanities were French and German. Moreover, the individual disciplines did not have as yet a “central journal” or “central journals.” The situation was quite different two decades later as reported by a number of studies (e.g. Nederhof and Zwaan, Nederhof 1996, 2006). It has been shown that journals are of prime importance to scholars in quite a few Humanities fields. Moreover, many Humanities fields are characterized by central or “core” journals. “Core journals” were defined as journals that are well known by the scholarly community, that receive high ratings of scholarly quality, and that are found very useful to one’s own research. Concerning the presumed reliance on older literature it was shown that 36 per cent of the references in core linguistic journals referred to literature published during the past four years. Although this percentage is below that in some of hard sciences, it is comparable to many others.

The shift from books to journal articles may be different in different disciplines. Linguistics is a good example where monographs have become less important; in literary studies, on the other hand, the monograph is still a major publication genre. A similar difference can be observed between, say, psychology and history. Is it possible to draw a general conclusion from such observations? Is the monograph as a publication genre in the Humanities in danger? This may be the case, for example, in literary studies and in history but much less so in linguistics and psychology where monographs had not played such a central role before the mid-nineteenth century either. Consequently, the changes in publication culture did not affect all Humanities disciplines to the same degree. The closer the research methodology in a Humanities discipline is to the research methodology in natural sciences, the more importance will be attributed to journals. (These problems are discussed in more detail in Thompson 2002; as for the role and future of the monograph in humanities research cf. also Williams et al., 67-82).

Scholars in the Humanities serve two publics with their products: research products are directed either to a scholarly public, or to a primarily non-scholarly public. In the latter case we speak of "external knowledge transfer." As a great number of studies have shown, a considerable (sometimes even a major) part of the scholarly outcome in various Humanities sub-fields is directed towards the "enlightenment" of the general public. This implies that Humanities scholars also publish in magazines and newspapers whose target readership is not the scholarly community. Of course, these publications should not be included under the heading "scholarly articles" in lists of publications.

A further difference between publication culture in science and in the Humanities is that in the latter we may encounter "double" publications quite often, i.e. a paper written in a lesser used language published in a C (= NAT) category journal may also be published in a major international language. This state of affairs has some important consequences for evaluation since the number of publications will not adequately reflect the scholarly output.

Yet another aspect concerning the difference between science and the Humanities has to do with the length (in terms of printed pages) of an article. Whereas a mathematical proof may need not more than one page, or a report on an experimental result in neuropsychology may take up just a few pages, it is hardly possible to produce a serious scholarly work in literary studies or linguistics on a few pages. Consequently, size is part of the notion of article in the Humanities but not in science. It does not come as a surprise, then, that in lists of publications we often find two categories,

shorter notes are listed separately from genuine scholarly articles. (It may be noted that most evaluation studies in the humanities are aware of this problem and it was suggested that only publications with a length of more than five pages should be regarded as “substantial” contributions.)

The differences in publication culture between science and the Humanities reflects a fundamental difference between two types of research activity. Humanities scholars use research retrospectively from various time periods, representing often conflicting intellectual positions against which scholars define their views and contributions. In the Humanities it is not research that cumulates, as in the case of sciences, it is rather the scholarly literature that cumulates with a negligible degree of obsolescence.

The differences between the two publication cultures (and between two types of scholarship) were discussed in considerable detail in a study by Wiberley (2003). Wiberley posits five types of humanities scholarship: 1) descriptive bibliography; 2) editing; 3) historical studies; 4) criticism; and 5) theory. He examines their bibliometric characteristics and demonstrates how one can differentiate these various types of scholarship based on bibliometrics and not on subjective impressions. For example, historical studies contain a high percentage of citation older than 20 years, criticism scholarship utilizes much more recent materials and a high percentage of secondary sources, theoretical scholarship cites a high number of very recent materials most of which are theoretical in nature. The five categories are clearly modelled on literary studies and art scholarship and cannot therefore be taken to be general characteristics of Humanities research. However, it demonstrates without any doubt the efforts to make judgements about scholarship more reliable.

What has been said about the general characteristics of humanities research need not be true for all Humanities disciplines, however. At least two Humanities disciplines, theoretical linguistics and cognitive psychology, seem to have more features in common with science than with traditional Humanities research, as far as methodology is concerned. In these two disciplines, articles in journals are more important than monographs, the monographs either summarize earlier research already published or they are textbooks designed for students or interested laymen. The references primarily include recent publications on the topic. Consequently, many articles in theoretical linguistics and cognitive psychology share the fate of articles in some life sciences: their lifespan is not very long. But, we have to accept the fact that Humanities covers a wide range of disciplines and each discipline may have its specific features that must be taken into consideration when evaluating the research output.

It has often been claimed that citation index (impact factor) is not very useful in humanities. De Bot found that there is no relation between the impact factor (as provided by ISI), and the impact of individual articles in the Humanities journals considered. A more recent study (Cope and Kalantzis 58-73) came to similar conclusions and pointed out a number of further difficulties concerning the use of impact factors in the Humanities. Nederhof and Zwaan studied the importance of citations in various Humanities disciplines. They came to the conclusion that the citation indexes are potentially useful for citation analysis in largely internationally-oriented Humanities fields, whereas for nationally-oriented fields, citation data fail to represent a valid picture of the national importance of research, but still may offer an indication of its potential international impact.

In sum, then, we may conclude that some Humanities disciplines may have their own publication practices, and some of these may not differ essentially from hard sciences in this respect.

Bibliometrics and Evaluation in the Humanities

We live in an age of metrics. All around us, things are being standardized, quantified and measured. Scholars concerned with the work of science and technology must regard this as a fascinating and crucial practical, cultural and intellectual phenomenon. In this respect, the Humanities are not an exception. Bibliometric indicators are being required for the evaluation of research output in the Humanities as well. There is a general consensus among researchers in the Humanities that bibliometric indicators, complemented with interviews, questionnaires and other qualitative information, provide a good instrument in research management. It has repeatedly been stressed that quality assessment of research performance can only be made by informed peers. Quantitative results can be used as background information to allow such experts to better form their opinion.

It soon became evident that no unanimously accepted methodology was readily available to highlight Humanities (and social sciences) research activities. Elaboration of a solid methodology, to take into account the specific characteristics of each discipline, turned out to be a matter of concern not only to those in charge of elaborating and implementing research policy, but also to the researchers involved in these disciplines. Discussion focused on research policy and allocation of research funds often led to acrimonious exchanges between "hard" and "soft" sciences protagonists. Researchers in the natural and life sciences have often shown a tendency to assess work

in social sciences and the Humanities within the framework of their own disciplines, not knowing that the evaluation methodology used in hard sciences is completely inadequate for the assessment of research output of human and social sciences.

Only a modest part of bibliometric studies has dealt with the Humanities. One of the recent studies (Hristova 2006) provides an overview of bibliometric studies dealing with the Humanities and discusses their implications for future research. One of the large trends in Humanities bibliometrics research was a tendency to focus on the differentiation between scientific and Humanities scholarship and trying to draw a distinct and accurate portrayal of Humanities research and its characteristics. The differences in publication culture reflect the differences in research activity. The author points out that Humanities scholars differ from scientific researchers in that they work individually, instead of collaborating (which is also shown by the fact that co-authored publications are relatively rare in the Humanities), using materials that span a wide range of years in terms of publication. This also demonstrates that Humanities research is not as susceptible to obsolescence as scientific research. Furthermore, interpretation is paramount in humanists' work because they do not report studies done outside of their writing; their publications constitute the research itself due to the focus on interpretation and analysis of primary sources, such as archives or works of art and literature.

A fair and just research evaluation should take into account the diversity of research output across disciplines and include all major forms of research publications. While journal ranking based on impact can help achieve this in many disciplines, which have formalized criteria for research and publications, the impact factor for Humanities journals is much harder to calculate (if at all) and is less reliable for assessment of research quality. Therefore, a more comprehensive bibliometric approach is in order to establish standard criteria for assessing research quality across each discipline.

A study performed by Moed, Luwel and Nederhof attempts to establish the foundation of a comprehensive bibliometric methodology for assessing the research performance of all scholars within any discipline in the Humanities. The methodology they propose has been successfully tested in the field of law research. The methodology includes several aspects: detailed classification of publications, ranking of journals, and input from scholars on their practices and perceptions. Based on the findings of their study of law research they conclude that a comprehensive and successful methodology should be concerned with developing accurate and discipline-

specific indicators of research performance. They suggest that this can be achieved by collecting data on publication output directly from the scholars, verifying this data electronically and classifying it in meaningful ways to separate the substantial research contributions from the incidental ones. The classified research can then be further weighted through the use of journal ranking derived partly from scholars' perceptions and partly through bibliometric analysis.

Moed, Luwel and Nederhof's methodology constitutes a significant contribution to bibliometric research in the Humanities because it offers a method that utilizes the strengths of quantitative research approaches without ignoring the benefits of qualitative approaches. By collecting information on scholars' output and perceptions, the research methodology involves the community in question and allows them to participate in the formulation of research quality indicators that might not be apparent to an outsider to the discipline. Meanwhile, the actual calculations of research quality remains dependent on a scientific methodology that is both valid and feasible.

It is normally assumed that the international orientation of a discipline is related to the object of research in that discipline. However, we wish to maintain that genuine scholarly research, regardless of the sub-discipline and the object of research, leads to results whose relevance and implications go beyond a purely national viewpoint of interest. Therefore, outcomes of genuine scholarly research, even the ones primarily related to national aspects, deserve to be communicated – in an appropriate form – to scholars outside the country as well.

ERIH, and the wider issues it raises about the definition and identification of excellence in the Humanities, have prompted an ongoing series of reflections about the usefulness of bibliometrics for the Humanities, acknowledging at the same time the need to develop a much wider array of measures of impact. For the time being, however, peer review has been given precedence over a quantitative approach, which remains the standard method used by research communities to identify excellence. Peer review has its advantages and its disadvantages: it can detect originality, but it also has the potential to defend conservative approaches. Generally speaking, however, it is acknowledged that the peer review introduces a measure of comparability into discussions of different national discourses in Humanities scholarship. ERIH relies on the principle of peer review to identify quality through open scholarly debate, and on a lengthy process of consensus building. It is hoped that this may bring us closer to an objective evaluation of scholarly output.

There is a general consensus among researchers and academic authorities that bibliometric indicators, complemented with interviews, questionnaires and other qualitative information, provide a good instrument in research management. It has, however, to be stressed that quality assessment of research performance can only be made by informed peers. This is the point where ERIH's methodology may become important.

Bibliography

- Cope, Bill and Mary Kalantzis. "Evaluating Webs of Knowledge: A Critical Examination of the 'Impact Factor'." *Logos* 21.3-4 (2010): 58-73.
- De Bot, Kees. "Het gebruik van citatie-index gegevens bij de kwaliteitsbeoordeling van publicaties in taalkundige tijdschriften." *Gamma* 11. (1987): 91-100.
- Hristova, Mariela. "Bibliometrics and the Humanities." The University of Texas at Austin. School of Information (2006). Web.
- Moed, Henk, Marc Luwel, and Anton J. Nederhof. "Towards Research Performance in the Humanities." *Library Trends* 50.3. (2002): 498-520.
- Nederhof, Anton J. "A Bibliometric Assessment of Research Council Grants in Linguistics." *Research Evaluation* 6.1.(1996): 2-12.
- Nederhof, Anton J. "Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: A Review." *Scientometrics* 66.1.(2006): 81-100.
- Nederhof, Anton J., and Rolf A. Zwaan. "Quality Judgements of Journals as Indicators of Research Performance in the Humanities and Social and Behavioral Sciences." *Journal of the American Society for Information Sciences* 42.5. (1991): 332-340.
- Thompson, Joseph W. "The Death of the Scholarly Monograph in the Humanities? Citation Pattern in Literary Scholarship." *Libri* 52. (2002): 121-136.
- Wiberley, Stephen. "A Methodological Approach to Developing Bibliometric Models of Types of Humanities Scholarship." *Library Trends* 50.3. (2003): 498-520.
- Williams, Peter et al. "The Role and Future of the Monograph in Arts and Humanities Research." *Aslib Proceedings* 61.1. (2009): 67-82.

Performing Excellence in the Humanities

The Funding Initiative 'Opus Magnum' of the VolkswagenStiftung

Vera Szöllösi-Brenig

Reading is – as the renowned neuroscientist Ernst Pöppel points out – one of the most unnatural activities for the brain. In his opinion, to read is to abuse the brain. Reading is not inherent to human nature, it is a human invention made possible by the flexibility of the brain (Pöppel 40f.) No doubt, books have become the ultimate physical evidence of this invention, and it conquered civilization long before the Gutenberg revolution.

Already in the Old Testament, we find the notion of the “Book of Life.” Those whose names are written in it will have life and spend eternity in heaven, whereas those whose names are not written in it will be cast into the lake of fire. The Book of Revelation mentions the “closed book.” This expression turned into an idiom still used today for everything we do not understand. We are currently in the midst of a paradigm shift: from the book as the main media of reading and writing to digital, with its new challenges and risks for culture and mind. Cultural evolution is accelerating and we do not know its next threshold. It is hardly surprising that this situation has a significant impact on scientific research. The VolkswagenStiftung as a funding body does not dictate the type of publishing media to its grant recipients, but attempts to support the researchers in this paradigm shift. What is currently the best media for different types of publications?

This paper is structured as follows: after a short introduction to the VolkswagenStiftung and its profile, it describes the paradigm shift from paper to digital as experienced in Germany in general and in the Humanities in particular in the year 2009/2010. Thereafter, the paper details some examples of the VolkswagenStiftung's funding portfolio to outline the foundation's policy towards publishing strategies. It particularly addresses the initiative “Opus magnum.” This “portrayed picture” 2009/2010 will be adapted to the current status (2013) where necessary.

The VolkswagenStiftung

The VolkswagenStiftung is an independent, autonomous non-profit foundation under private law. Despite its name, it is not a subsidiary of the respective car company. Its origins are based on a government treaty that settled the controversy concerning the ownership of the Volkswagen Company after the Second World War. In 1961, it was decided to convert the company into private ownership by issuing so-called “Peoples shares.” The proceeds resulting from the sales of these shares are the basis of the capital of the Foundation – in 2009, some €2.3 billion (2013: €2.7 billion). In its approximately 50 years of existence, the VolkswagenStiftung has funded about 30,000 projects with a total of more than €4 billion. With a funding volume of up to €100 million in 2009 (2012: €144 million), it is the largest private science funder and one of the major foundations in Germany.

The foundation’s purpose – as stipulated in the statutes – is to support the Humanities and social sciences as well as science and technology in higher education and research. The foundation’s funding philosophy is to provide systematic stimuli to those approaches and developments that address some of the big challenges of our time. Having identified new fields of research or a new mission, the foundation defines respective funding initiatives in which funds are made available. These initiatives form the framework of its funding profile, which is structured in three categories: support of persons and new structures, thematic and cultural challenges and international focus. By means of this concentration on relatively few initiatives, the Foundation ensures that the available funds are used effectively.

In this way, the VolkswagenStiftung always intends to respond – as appropriately as possible – to the present situation in the realm of research and knowledge production. But: how can this situation in Germany be described with regard to the paradigm shift from paper to digital?

The Paradigm Shift Paper/Digital

Engulfed by the current media revolution with no idea of the future, it is quite difficult to marshal up-to-date and reliable figures. According to the findings of the Allensbach Institute, in 2009, 74 per cent of all Germans had internet access.¹ The increase of the internet access shows astonishing parallels to the rapidly increasing television access starting some 40 years ago.



Chart 1: Schneller, slide 2.

54 per cent of all Germans use the internet every day; an increase of 9 per cent within two years (Schneller, slide 3). However, looking at the European dimension, Germany is not among the top ranking nations in Europe concerning daily internet usage:

	Typical 7 day week internet use (%)	Weekend internet use (5)
Denmark	84	76
Netherlands	81	71
Norway	79	71
Sweden	79	62
Belgium	67	60
France	66	58
UK	64	55
Germany	61	53
Europe	60	51
Spain	53	42
Italy	40	24

Chart 2: Based on Fennah, slide 7.

According to this chart, presented by the European Interactive Advertising Association, Denmark, the Netherlands, Norway, and Sweden are far ahead of Germany. Interesting here is the complete change of lifestyle due to the internet. Life is permeated by the new technologies: job, friends, family, hobbies and finances.

Concerning books and our reading habits, e-books have not yet become widely accepted in Germany. The author and intellectual Umberto Eco emphasizes that the book is like the spoon, the hammer or the wheel: once these things are invented, nothing better is feasible (Tonnac). Nevertheless, in a global perspective, the e-book business appears to be growing rapidly. This fact is indicated, for example, by Google eBooks, which opened in December 2010. Also, in the first half of 2010, Amazon sold three times more e-books than in 2009 and twice the number of hardcovers (Amazon). In China, 79 million readers – this is 5.8 per cent of the Chinese population – already used a reading device in 2008, and this expansion is showing a continuing upward trend. An interesting detail is how e-books appeal in particular to young Chinese people. “Readers below 24 accounted for almost 50% of the group. Unlike reading printed books, reading e-books is seen by the younger generation to be a modern and fashionable activity” (Wong). In general, there seems to be not only a generational, but also a gender aspect. With regard to the Western market, the Swiss media expert Heinz Bonfadelli explains the strategy of the e-book companies to attract more male readers by offering a “male” machine (Bonfadelli).

In November 2009, the Münchner Kreis - National IT summit of Germany, financed by the German ministry of economics, published a new study titled *Prospects and Opportunities of Information and Communication Technologies (ICT) and Media*. They asked German and international experts: “Will the e-book become the standard publication form in Germany?” The answers are interesting: 20 per cent of the experts answered “Yes - in 2024.” About 18 per cent answered: “Yes – from 2025 on.” 20 per cent more answered: “Yes – later than 2030.” But nearly 40 per cent of the experts answered: “No, probably never.” So, regarding this question, the German experts did not agree with their international colleagues.

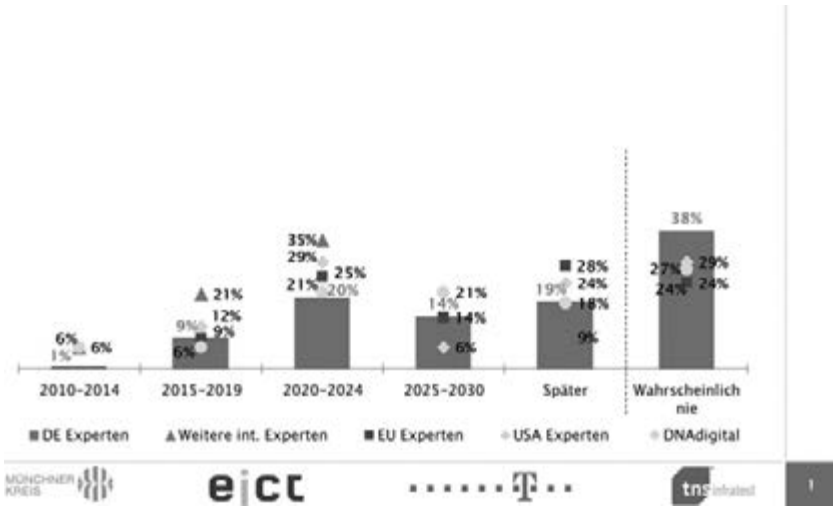


Chart 3: Münchner Kreis 208.

The worldwide book production confirmed this survey result because it continued to increase. According to an extrapolation of Robert Darnton, Director of the Harvard University Library, worldwide there were 700,000 new titles published in 1998, 858,000 in 2003 and 976,000 in 2007 (Darnton). Furthermore, the turnover of the book market is impressive. In 2008, the turnover amounted to \$25 billion in the US, to €23 billion in Europe and to €9.6 billion in Germany alone. In 2009, the sales of digital devices in Germany only reached to 0.1 per cent of this turnover, whereas the phenomenal number of 93,124 new titles was published (Jessen). Critical voices already speak of an overproduction of titles (Güntner 2009). At the same time, media analysis show that our reading habits change: texts are only being read partly and not from the beginning to the end; readers tend to employ a text as a knowledge base (Hamm and Langen). Increasingly, reading turns into data mining.

In German academia, in science and the Humanities, Open Access as an electronic research publication method is widely promoted by German research and funding organizations, such as the Deutsche Forschungsgemeinschaft (DFG), the Max-Planck Society, the German Association of university presidents and the VolkswagenStiftung. They launched an internet platform www.open-access.net to supply information to researchers – technical and practical information as well as policy information.²

At the same time, the average print run in academic publishing houses is significantly decreasing. From 1990 until 2000 this decrease amounted

to 60 per cent (1200 copies/480 copies) (Füssel). Though these figures might not be up-to-date, they demonstrate the trend quite well. No doubt, this decrease in average print run affects mainly the Humanities where the book – the monograph – is still a main publication form. No doubt, either, there is no trend reversal. In addition, the same trend can be identified regarding the purchases by State and university libraries: we find less monographs and less money spent on monographs, more money spent on journals for less or the same amount of journals, especially in the Science, Technology, Medicine (STM)-sector. Exact figures were published for the State and university libraries in Bavaria for the period 1988-1998 under the meaningful title “The journal crises as crises of the monographs” (Kopp). According to these figures, within 10 years, State and university libraries in Bavaria bought 27.18 per cent less monographs and spent 2 per cent less on monographs while the total budget increased slightly and the expenses for journals, especially in the STM-sector, even increased by 200 per cent.

So the aggressive pricing policy of publishers in the STM-journal sector has had its effects. To quote the conclusion: “The exorbitant resource needs in the STM-journal sector produce a reallocation in favor of those disciplines which are the most affected by the cost increase” (Kopp 1826) – which means at the same time a reallocation at the expense of the rest of the disciplines, mainly the Humanities.

At the same time, the number of online dissertations and habilitations in Germany, archived at Deutsche Nationalbibliothek, is increasing in absolute figures (Chart 4) as well as in relation to the total number of dissertations (Chart 5):

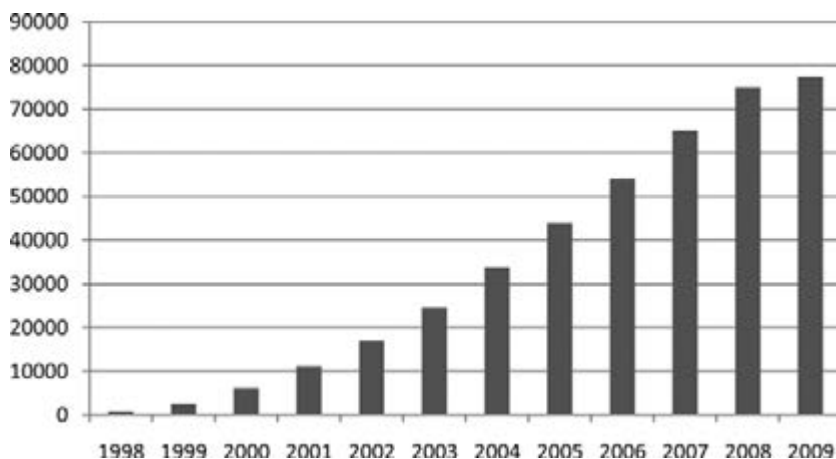


Chart 4: Based on Dissonline (as of September 2009).

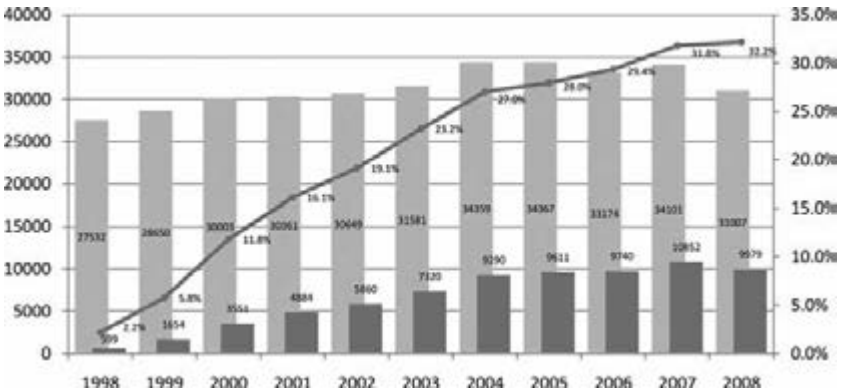


Chart 5: Based on Dissonline (as of December 2008).

But there are very interesting differences between the disciplines: electronic dissertations in biology, chemistry and computer sciences were in the lead, whilst German studies, history, philosophy and law took the rear positions. The following chart is based on the statistics of Dissonline for the year 2008 and concentrates on disciplines important in numbers:

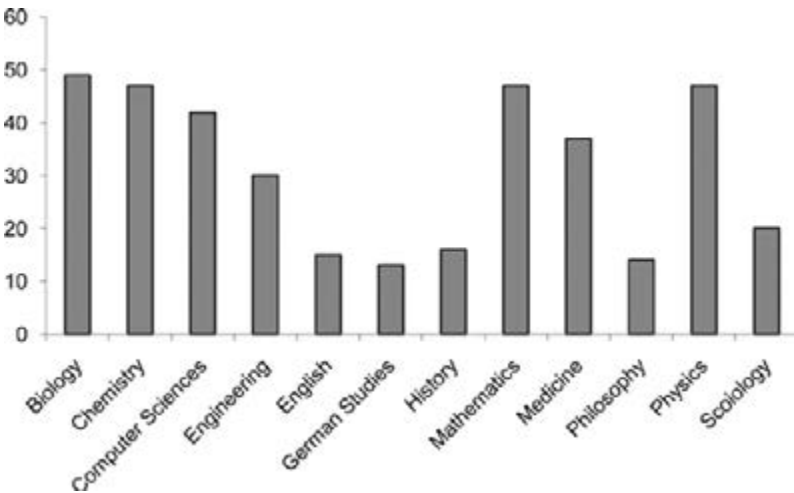


Chart 6: Based on Dissonline.³

Within German academia, there is a discussion about Open Access publications. The following arguments favour Open Access and digital publishing:

- Increased visibility, especially increased international visibility;
- Fast and easy access to information;

- Good searchability;
- Data corpora (dictionaries, encyclopaedia, text editions) as particularly successful applications;
- New digital resources in audio and video;

And last but not least: there is one argument:

- Why should the public pay twice: for the research and then again for its results?

However, it is also important to address the persisting doubts concerning Open Access and digital publishing. Some key points:

- Quality issues concerning text editing;
- Unsolved technical problems of long term storage;
- The lack of a general non-commercial infrastructure to provide access to information;
- Legal concerns regarding copyright issues;
- Questions concerning our reading habits: who is the audience of a text and what is its function?

Researchers, especially in the Humanities and social sciences, are more or less reluctant to adopt the electronic publication form. In 2005, the DFG published a representative survey on the change of publication strategies in Germany.⁴ The DFG survey asked: For which purpose did you use the DFG publishing grant⁵ (multiple answers possible)?

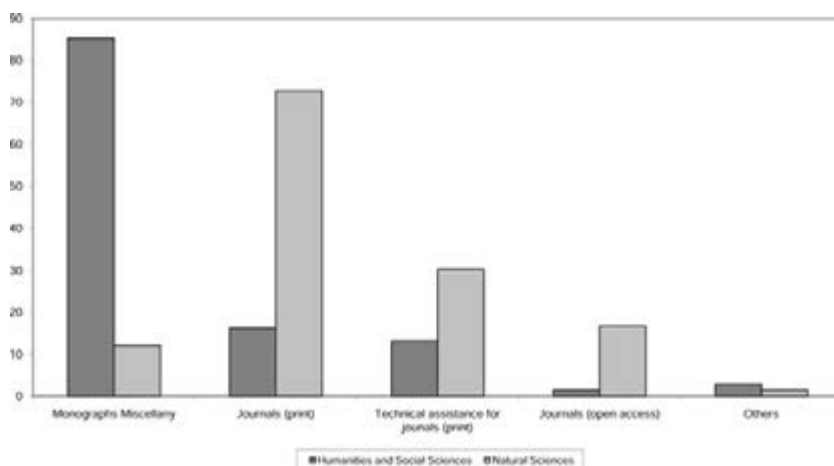


Chart 7: Based on DFG: Table 6.03.⁶

Researchers in the Humanities and social sciences spent more than 80 per cent of their publishing grant on monographs, whereas natural scientists spent it mostly on printed and some open access journals.

When asking: Which publication form do you use either often, or very often, to keep yourself informed in your field? The answers demonstrate that there is still a strong vote for the monograph in the humanities and social sciences:

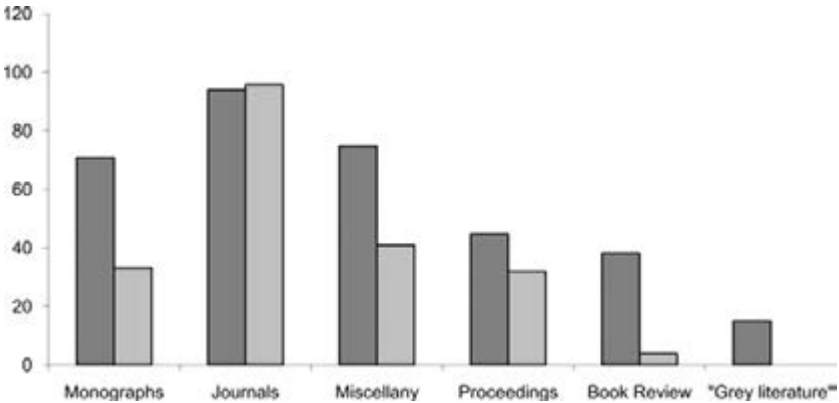


Chart 8: Based on DFG 2005: Table 3.01 (multiple answers possible).

The category of "grey literature" is a specific German phenomenon: that is the German expression for popular science books.

The DFG survey also asked: "Who is the main audience of your publication?" The result is not surprising:

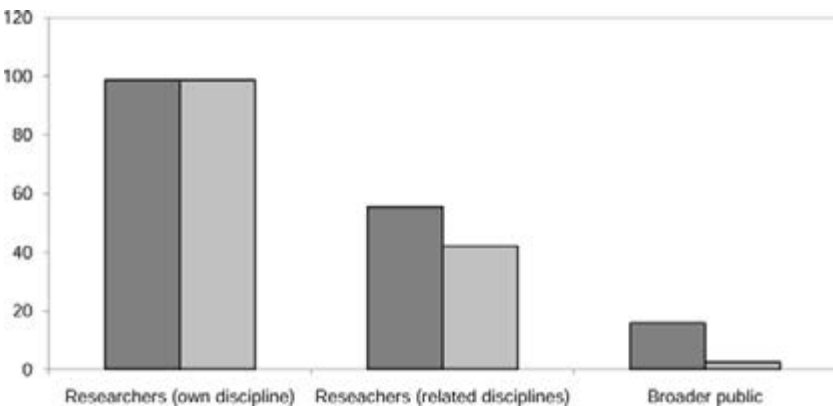


Chart 9: Based on DFG 2005: Table 3.04.

In the Humanities and social sciences, researchers do not only write for colleagues of their own scientific community, but also for experts of related disciplines and for the broader public, whereas in the natural sciences, this is of less importance.

The Initiative ‘Opus Magnum’ and the Funding Portfolio of the VolkswagenStiftung

The focus of the VolkswagenStiftung’s funding initiative “Opus magnum” is monographs with a highly intellectual impact on the own discipline, on related disciplines, and even on the broader public. This initiative was jointly established by the Fritz Thyssen Stiftung in Cologne and the VolkswagenStiftung in 2004/5 and ran until 2011. Starting in 2012, “Opus Magnum” has been continued by the VolkswagenStiftung only. “Opus Magnum” is a funding opportunity for researchers in the Humanities working at German universities. The objective is to offer renowned experts at German universities (professors) the opportunity to write their *magnum opus* by providing an extra sabbatical, which can last from six months to two years. The foundations cover the costs of substitute teaching and a lump sum for other running costs.

In the Information for Applicants, a magnum opus is defined as “a substantial scientific publication which – thanks to its quality and originality – has a significant impact on the development of the respective academic discipline and beyond.”

Up to July 2010, a total of 150 applications were submitted out of which 42 projects were granted a total of €6.35 million (2013: 221 applications/74 projects funded, €8.8 million). Three examples may illustrate the funding initiative and its results: *Bedeutsamkeit – Ontosemiologische Untersuchungen zum Zusammenhang von Zeit und Bedeutung* (Significance – Ontosemiological analysis of the interrelation between time and sense), a two year project granted in 2006 to Jochen Hörisch, professor for Modern German Literature and Media Analysis at the University of Mannheim. This project ended in 2009 and in the same year, his book was published by Hanser Literaturverlag München. Within a few months, until December 2009, it was sold more than 1015 times. Without any doubt this monograph can be called a successful example of the “grey literature”-type.

Second example: The volume *Wozu Kunst? – Ästhetik nach Darwin* (What is Art good for? Esthetics after Darwin) by Winfried Menninghaus, was published in October 2011 by Suhrkamp, a well-known, even popular publisher in Germany. Back then, Menninghaus was professor at the Peter-Szondi-

Institute for General and Comparative Literature at the Free University of Berlin, since 2013 he is funding director of the Max Planck Institute for Empirical Aesthetics in Frankfurt/Main.

Third Example: The magnum opus of Klaus Michael Bogdal *Europa erfindet die Zigeuner* (Europe’s Invention of the Gypsies). This publication has been awarded the distinguished Prize for European Understanding of the Leipzig Book Fair in 2013.

The magnum opus statistics (2004/5-2009) demonstrate a widespread distribution of grants according to disciplines in the Humanities.

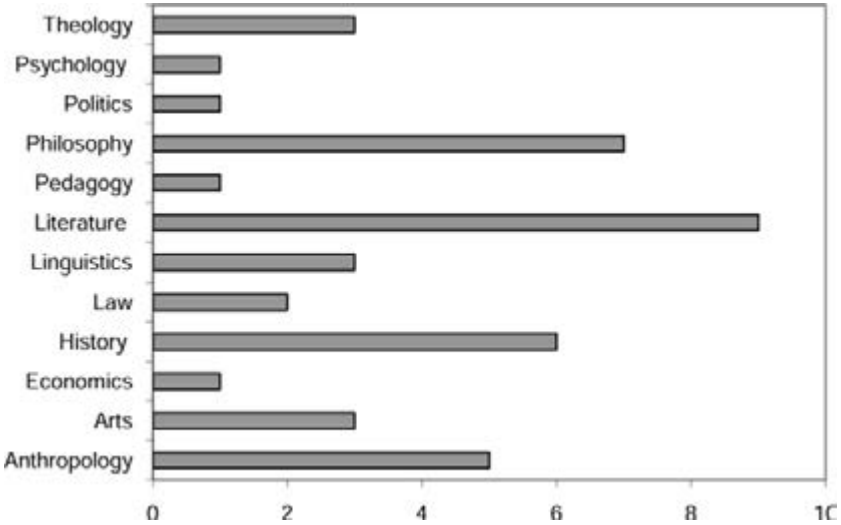


Chart 10: Grants according to disciplines in the “Opus Magnum” funding initiative.

The funding initiative comprehends *Humanities* in its broader sense. Several projects were also granted within disciplines generally attributed to the social sciences, but which, due to the focus of their work, have a quite theoretical or conceptual approach. For example, the political scientist Jürgen Neyer from the Goethe University in Frankfurt/Main has been working on the subject *Macht und Rechtfertigung. Eine deliberative Theorie supranationaler Integration* (Power and Justification. A deliberative theory of supranational integration). The legal scholar Andreas Zimmermann from the Walther-Schücking-Institute for International Law of the University of Kiel has been writing a magnum opus on *Africa and International Law*.

Two details should be noted: First, the funding initiative “Opus Magnum” does not require a print format, an electronic publication is equally possible

since its start in 2004/5. However, applicants have only started in 2013 to discuss electronic publication as a real option. Second, the Volkswagen-Stiftung does not provide publishing grants in this funding initiative. It is assumed that a magnum opus should be commercially viable.

This brings us to the first of five arguments why the VolkswagenStiftung expects to perform excellence in the Humanities through installing and processing the funding initiative “Opus Magnum.”

- (1) The funding decisions are based on an elaborate peer review process as an instrument for quality assurance. There is a 2-level peer review process in the form of a contest. First, all applications are reviewed by two separate experts of the applicant’s discipline who provide us with their evaluation in a written form. Afterwards, the funding initiative’s steering committee carries out a comparative review of all applications. Based on this review process a maximum of ten books are being funded each year. The decision of a publisher to print the magnum opus is the second quality test for the book as well as a subsequent quality test for the peer review process itself.
- (2) In giving renowned researchers in the Humanities the opportunity to write a magnum opus, the VolkswagenStiftung responds to a sustainability issue. Our interest is in the long term availability of a publication due to the “half life of knowledge” in the Humanities. In analogy to the term half life in radioactivity, Heinz Hauffe from the university library Innsbruck defined “half-life of knowledge” as the time after that half of the publications/monographs are no longer in demand. To cite his findings: the half-life of knowledge in medicine is 3.5 years, in physics 4.7 years, but in classical philology 20 years (Hauffe). Until now, a book guarantees, quite easily, a certain sustainability for decades, whereas electronic archiving still has unsolved technical problems.
- (3) There is an epistemic issue: The monograph as an extended publication corresponds to the specific epistemic aspects of the “Geisteswissenschaften”: In contrast to the natural sciences, the text of the publication is not only the “container” of already existing findings, but the foundation of the argumentation itself. The imperative linearity of a text – in contrast to an electronic publication and its possibly circular structure – with a well-defined beginning, the structure of argumentation and a clear end has a constitutive function for the Humanities with their specific epistemic approach (Gradmann 59).
- (4) The consequence out of this epistemic factor is that in academic Humanities’ perception the rule applies: “first book, then journal.” Referring to the argumentation of the German philosopher Pirmin Stekeler Weithofer, in the Humanities new ideas are published first in monographs and then they are

picked up and discussed in journals. “In the priority list of the importance of texts’, monographs rank first in philosophy.” (Stekeler-Weithofer 1).⁷

(5) Last but not least, there is a language issue: Neuroscientists are discovering what philosophers, researchers from the different philologies, and historians know by heart: The styles of thinking and language are so inextricably intertwined that only writing in one’s mother tongue guarantees the quality of one’s thought and argumentation. Stekeler-Weithofer for example stresses the necessity of highest language comprehension to write a paper in philosophy. He even insists that “as a rule the first debates on conceptional suggestions and findings are made in the mother tongue. Only a language of which we have the highest command with its nuances of semantic inferences and pragmatic implicatures provides us with the self-contained and strict language competence necessary for philosophical thinking” (Stekeler-Weithofer). Along the same lines, Konrad Ehlich, specialist of “German as foreign language,” diagnoses a “gnoseological function” of a scientific language (91).⁸

The funding initiative “Opus Magnum” does not require that the project’s product – the magnum opus – is written in German. But it speaks for itself the bulk of applications plan to write the opus in German. In the Humanities it is possible to raise high scientific and public awareness on a national scale. This experience is in accordance with the findings of the DFG survey: “When publishing for or reading in your academic discipline: which is the importance of the following languages?”

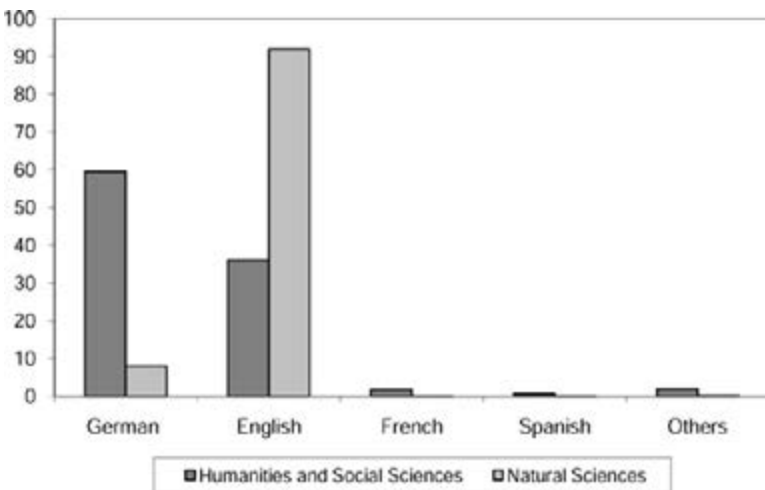


Chart 11: Based on DFG: Table 3.05.

Whereas in the natural sciences the container-language English is the language of publication,⁹ in the Humanities it is the mother tongue German which is of high interest.

For these five reasons we support performing excellence in the Humanities through providing the funding initiative “Opus Magnum.”

But it goes without saying that the VolkswagenStiftung’s funding portfolio has to respond to the paradigm shift from paper to digital and the foundation does not ignore the benefits of electronic publishing. The funding initiative “Documentation of Endangered Languages” – a program which is internationally known by the German title’s acronym DobeS – aimed at multimedia documentations of endangered languages stored in an electronic archive and accessible through the internet. Since 2000, this online archive is located at the Max Planck Institute for Psycholinguistics in Nijmegen, containing digitized video, audio and text data of endangered languages (see dobes.mpi.nl).

With regard to any printed language documentation, these multimedia documentations have an added value: There are the raw data to listen to and to watch; these raw data are complemented with the annotation tiers for transcription, translation and interlinear glossing. Cyclic linking is possible to various resources like dictionaries, encyclopedia, collections of indigenous knowledge, ethno-botany, ethno-zoology etc. Moreover, electronic language data offer the possibility of cross linguistic data analysis.



Picture 1: ELAN Screen shot.

This language archive has been growing continuously. From 1999 until 2012, about 100 languages in the world have been documented. 132 projects were funded out of 288 applications. Within ten years, a system of documentation tools like ELAN, on the one hand, and archival infrastructure, on the other hand, was established. All components are usable free-of-charge in academic research. They can be considered as best practice for eHumanities in general. Since September 2010, the DobeS archive has been integrated in The Languages Archive (TLA), a joint organization of the Max Planck Institute for Psycholinguistics in Nijmegen, the Berlin-Brandenburgische Akademie and the Royal Netherlands Academy of Sciences (KNAW).



Picture 2: The DobeS World.

The VolkswagenStiftung strongly supports the establishment of the Digital Humanities. We are convinced that new research questions will emerge out of the data material. That is why in the final phase of the DobeS funding initiative, researchers were not only invited to submit documentation projects, but also projects which use the DobeS archive for scientific purposes, for example for comparative studies, but also to detect new research questions. In December 2013, the VolkswagenStiftung organized the international Herrenhausen Conference “(Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age.” The aim of the conference was to initiate dialogue between representatives of the Digital Humanities and the “established” Humanities: What kind of new knowledge can we expect? What knowledge is in danger of disappearing? Is there any foundation for a new knowledge culture? (see Szöllösi-Brenig).

On the other hand, despite or just because of the importance of the mother tongue for the humanities, the aspect of international visibility is an important one. The VolkswagenStiftung feels the need to support all efforts to strengthen presence of German researchers on the international scale. It goes without saying that all applicants can ask for publishing grants destined for international publishers. Moreover, from 2006- to 2011, the VolkswagenStiftung ran the funding initiative “Deutsch plus” – A Program for Multilingualism in Teaching and Research. Within the existing context of multilingualism this initiative aimed at attracting more international attention to German as a language of science and to lend more weight to research findings obtained and published in the German language. This program had four components of funding possibilities: multilingual study courses, conferences designed to focus attention on multilingualism, research projects dealing with the linguistic and cultural imprint on academic thinking and production, and – last but not least – the translation of outstanding German academic books and papers in any of the world’s main languages: English, French, Spanish, Chinese and Arabic. The following translations were granted in 2009: *Die Gesellschaft der Gesellschaft* of the German sociologist Niklas Luhmann (translation into English), *Die Zukunft der Verfassung* of Dieter Grimm, law specialist and former judge at the Federal Constitutional Court (translation into English), *Geschichtliche Grundbegriffe. Historisches Lexikon zur politisch-sozialen Sprache in Deutschland* (translation into Russian). After the end of the program, researchers from the Humanities and cultural sciences are invited to include application for translation in their regular applications.

Perspective: Towards a Future of Media Diversity

To sum up: As the VolkswagenStiftung always intends to respond – as appropriately as possible – to present needs in research, it stands up today for diversity in publication forms. It advocates against homogeneity and is in favour for plurimediality in the Humanities: The monograph is *the* publication long form today for comprehensive treatises that invite the readers to immerse themselves in the author’s argumentation chain. Electronic publishing is important for international visibility and quick access. All kind of data collections shall mandatorily be published electronically.

To plan a reasonable publication strategy, the researchers key questions are: Do I want to provide knowledge or information? Is the planned publication a text to be read and understood or is it ideal for data mining? What is

more important: sustainability or quick access? Today, it seems important to retain a balance between paper and digital. But the key issue for tomorrow will be: Will our reading habits change?

Notes

1. Worldwide the number of computers with internet access amounted to 758 million in April 2010, an increase within 2 years of 188 million (Bernau 40).
2. See the policy paper "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities" (2003) of the Allianz der deutschen Wissenschaftsorganisationen.
3. The corresponding figures for the rest of disciplines are omitted.
4. Regarding the situation worldwide, see the results of the Study of Open Access Publishing (SOAP) carried out 2009 - 2011 in 162 countries. Based on an online survey of n=38,358, the most relevant findings of the survey are "that around 90% of the researchers are convinced that open access is beneficial for their research field," however "at the same time, only 8 - 10% of articles are published yearly in open access journals. The origin of this gap is apparently mostly due to funding and to the (perceived) lack of high quality open access journals in particular fields" (Dallmeier-Tiessen S.10f).
5. The DFG publishing grant consists of a lump sum of a max €750 per year. The publication form – print or electronic – can be chosen by the researcher himself/herself.
6. This chart as the following concentrate on humanities/social sciences and natural sciences; the corresponding figures for life sciences and engineering are omitted.
7. In this article, Stekeler-Weithofer discusses the ERIH Initial List.
8. It is not only the Humanities that stick to German as scientific language. In 2007, the "Arbeitskreis Deutsch als Wissenschaftssprache" e.V. (ADAWis) was founded. Its head is Prof. R Mocikat from the Institute of molecular immunology, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH) in Munich. "The primacy of a uniform language use in scientific research will result in intellectual impoverishment" (Mocikat, Haße and Dieter 101). The backslide of German as language of science started long before globalization, at the end of the First World War when international science organisations boycotted texts in German. The total/moral defeat of Germany in 1945 accelerated this development (Glück 41f).
9. The language distribution in natural sciences' publications from 1880 until 2005 is shown in Ammon 2010.

Bibliography

- Allianz der deutschen Wissenschaftsorganisationen. *Open Access. Positionen – Prozesse - Perspektiven*. Oct. 2009. Web. 27 August 2010. www.allianz-initiative.de/fileadmin/openaccess.pdf.
- “Amazon verkauft immer mehr digitale Bücher“. *Frankfurter Allgemeine Zeitung* 21 July 2010.
- Ammon, Ulrich. “Kaum noch ein Prozent Weltanteil in den Naturwissenschaften. Über Deutsch als Wissenschaftssprache“, *Forschung und Lehre* 6/10: 400-402.
- Bernau, Patrick. “WWW. – drei Buchstaben verändern die Welt“, *Frankfurter Allgemeine Sonntagszeitung* 15 August 2010: 40.
- Bonfadelli, Heinz. “Geist und Papier. Erkenntnisse des Züricher Medienforschers. *Die Zeit* 22 July 2010: 40.
- Dallmeier-Tiessen, Sünje, et al. “Highlights from the SOAP Project Survey. What Scientists Think about Open Access Publishing” (2011). Web. 12 March 2014. <http://arxiv.org/ftp/arxiv/papers/1101/1101.5260.pdf>.
- Darnton, Robert. “Verstaatlicht Googles Buchmaschine!“ *Frankfurter Allgemeine Zeitung* 14 October 2009: N5.
- Deutsche Forschungsgemeinschaft. *Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access*. Bonn 2005 Web. 27 August 2010. www.dfg.de/download/pdf/dfg_im_profil/evaluation_statistik/programm_evaluation/studie_publicationsstrategien_bericht_dt.pdf.
- Dissonline. Digitale Dissertationen im Internet. “Statistik. Sammlung, Verzeichnung und Archivierung durch die Deutsche Nationalbibliothek - Angaben zur Statistik.“ Web. 27 August 2010. www.dissonline.de/aktuell/statistik.htm.
- Dokumentation bedrohter Sprachen/Documentation of Endangered Languages. Archive. Web. 13 March 2013. www.dobes.mpi.nl.
- Ehlich, Konrad. “Deutsch als Wissenschaftssprache für das 21. Jahrhundert.“ *Forschung. Politik – Strategie – Management* Fo 3+4/2009: 89-95.
- Fennah, Allison. “Insights into the Evolution of Internet, Mobile, TV, Radio, Newspaper and Magazine Consumption across Europe” (European Interactive Advertising Association). Presentation for Internet World Congress 23rd June 2009. Web. 27 August 2010. www.eiaa.net/Ftp/Conferences/EIAAIWCPresentation%2Epdf.
- Füssel, Stefan. “Von der Medienkonkurrenz zur Mediensymbiose.“ *Wissenschaftspublikation im digitalen Zeitalter. Verlage, Buchhandlungen und Bibliotheken in der Informationsgesellschaft* (Gesellschaft für das Buch; Bd. 7). Ed. Bundesvereinigung Deutscher Bibliotheksverbände et al., Wiesbaden 2001.
- Glück, Helmut 2008. “Deutsch als Wissenschaftssprache“. *Glanzlichter der Wissenschaft 2008. Ein Almanach*. Ed. Deutscher Hochschulverband: 37-43.
- Gradmann, Stefan. “Vom Verfertigen der Gedanken im digitalen Diskurs: Versuch einer wechselseitigen Bestimmung hermeneutisch und empirizistischer Positionen“. *Historical Social Research*, Vol. 29 (2004) No1: 56-63.
- Güntner, Joachim. “Der Buchmarkt im Strudel des Digitalen“. *Aus Politik und Zeitgeschichte APuZ* 42-43/2009: 9-17.
- Hamm, Ingrid and Claudia Langen. “Das Lesebarometer – Lesen und Umgang mit Büchern in Deutschland.“ *Lesen in der Mediengesellschaft: Stand und Perspektiven der Forschung*. Ed. Heinz Bonfadelli and Priska Bucher. Zürich 2002: 43-56. Quoted in Meidenbauer.
- Hauffe, Heinz. “Langfristige Verfügbarkeit elektronischer Medien. Referat am Kolloquium »Speicherbibliotheken - Digitale Bibliotheken«, Graz, 3. April 1997.“ Quoted in Meidenbauer.
- Jessen, Jens. “Hurra, wir lesen noch!“ *Die Zeit* 22 July 2010: 39.

- Kopp, Hans. "Die Zeitschriftenkrise als Krise der Monographienbeschaffung." *Bibliotheksdiens* 34.Jg, (2000) H.11: 1822-27.
- Meidenbauer, Jörg. "Guide Wissenschaftliches Publizieren". *Clio online. Fachportal für die Geschichtswissenschaften*. March 2005. Web. 27 August 2010.
www.clio-online.de/app_support/print.aspx?lang=de&mid=11093&ModeID=1&tabid=40208143&ModId=11093.
- Mocikat, Ralph, Wolfgang Hasse, and Hermann H. Dieter. "Sieben Thesen zur deutschen Sprache in der Wissenschaft." *Forschung. Politik – Strategie – Management* Fo 3+4/2009: 100-2.
- Münchener Kreis e.V. et.al. *Nationaler IT Gipfel. Prospects and Opportunities of Information and Communication Technologies (ICT) and Media, International Delphi-Study 2030*. Stuttgart 2009. Web. 27 August 2010 www.tns-infratest.com/presse/pdf/Zukunft_IKT/International_Delphi_Study_2030_Prospects_and_Opportunities_of_ICT_and_Media.pdf.
- Pöppel, Ernst "Was geschieht beim Lesen?". *Aus Politik und Zeitgeschichte* APuZ 42-43/2009: 40-5.
- Schneller, Johannes. *ACTA 2009. Zentrale Trends der Internetnutzung in den Bereichen Information, Kommunikation und E-Commerce*. Presentation 21 October 2009. Web. 27 August 2010.
www.acta-online.de/main.php?selection=69&rubrik=0.
- Stekeler-Weithofer, Pirmin. "Publikationsverhalten in der Philosophie". *Kritikon*, 25.11.2008. Web. 27 August 2010. www.kritikon.de/article/49.
- Study of Open Access Publishing (SOAP) Web. 27 August 2010. <http://project-soap.eu/>.
- Szöllösi-Brenig, Vera, "Conference Summary: "(Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age" (5-7 December 2013). Web. 13 March 2014. www.volkswagenstiftung.de/fileadmin/grafiken/veranstaltungen/2013/HK_Digital_Humanities/Summary_HK_DigHum_engl.pdf.
- Tonnac, Jean Philippe de. "Die große Zukunft des Buches. Wie lesen wir morgen? Welche Medien werden überleben? Und wie verändert die Technik unseren Lesealltag? Ein Gespräch mit den Schriftstellern Umberto Eco und Jean-Claude Carrière." *Die Zeit* 22.Juli 2010: 41.
- Wong, Tiffany. "Weblog 12 August 2009." *TeleRead. News & views on E-books, Libraries, Publishing and Related Topics*. Web. 27 August 2010.
www.teleread.org/2009/08/12/coming-tiffany-wongs-report-on-e-books-in-china-79m-readers-of-digital-books.

Contributors

Janneke Adema is a PhD student at Coventry University. She is the author of the OAPEN report *Overview of Open Access Models for eBooks in the HSS* (2010) and has published in *The International Journal of Cultural Studies*; *New Media & Society*; *New Review of Academic Librarianship*; *Krisis*; and *New Formations*. She has co-edited a living book on *Symbiosis* (Open Humanities Press, 2011). Currently, she is writing a dissertation on the future of the scholarly book. Her research practice focuses on issues of openness, remix and authorship and on the material production of the book, through which she analyzes the discourse and power struggle surrounding the academic monograph. Her research, both a theoretical and practical intervention into this debate, can be followed, as it develops, on www.openreflections.wordpress.com.

Andrea Bozzi is Director of the Institute for Computational Linguistics of the Italian National Research Council (CNR), Pisa. In 2010 he also chaired the CNR Humanities Department in Rome. He is associated partner to the ERC Advanced Grant 249431. He has coordinated national and international projects in digital philology and digital libraries. His main field of interest is the production of computer-assisted textual criticism systems for philological disciplines, in particular related to ancient documents. His main publications include Bozzi A., Cignoni L., Lebrave J.-L. (eds.), *Digital Technology and Philological Disciplines*, IEPI, Pisa-Roma, 2004; Bozzi A., Cignoni L., Fedele G. "Linguistic tools for navigation in a virtual museum," in *Archeologia e calcolatori*, Suppl. 1 al vol. 17 (2006) *Virtual Museums and Archaeology: The Contribution of the Italian National Research Council*, Edizioni All'insegna del Giglio, pp. 209-220.

Luca Codignola is Professor of North American History at the University of Genoa (Italy) and Adjunct Professor at Saint Mary's University (Canada). His main field of research is the Roman Catholic church in the North Atlantic area in the early modern era. He has also written on the history of early European expansion. Among his most recent books, *Humans in Outer Space: Interdisciplinary Odysseys* (2009, ed., with K.-U. Schrogl); *Le Saint-Siège, le Canada et le Québec* (2011, with G. Pizzorusso and M. Sanfilippo); and *Little Do We Know: History and Historians of the North Atlantic, 1492-2010* (2011, ed. M. Binasco).

Péter Dávidházi, Member of the Hungarian Academy of Sciences, is Head of the Department of Nineteenth-Century Literature at the Institute for Literary Studies of the Research Centre for the Humanities, and Professor of English Literature at Eötvös Loránd University in Budapest. Author of books published in Hungary, England, and the US, including *The Romantic Cult of Shakespeare: Literary Reception in Anthropological Perspective*, Macmillan, 1998. His latest book is *Menj, vándor: Swift sírfelirata és a hagyományrétegződés* (Go, Traveller: Swift's Epitaph and the Strata of a Tradition), Pro Pannonia, 2009. His current research focuses on biblical allusions in Shakespeare, the prophetic tradition in Hungarian poetry, and the methodological changes of recent literary scholarship.

Jacques Dubucs was born in 1953 in Bayonne. He is former fellow of the Ecole Normale Supérieure (Saint-Cloud), Agrégé de Philosophie (1976), PhD in Logic (1978), Doctorat d'Etat on Hilbert's Programme (1984), Assistant Professor of Mathematics at the Ecole Normale Supérieure of Rabat (Morocco) from 1980 to 1985. He works at CNRS since 1985. Currently he is Scientific Director of the Department of Human and Social Sciences at the French Ministry of Higher Education and Research and Chair of the Special Working Group "Societal and Cultural Innovation" of the European Strategy Forum on Research Infrastructures. His main domain of research is non-classical logics and cognitive science, especially the analysis and formalization of actual reasoning. His recent work concerns logic of "feasibility"; namely, formal description of "really or humanly achievable" processes. Another part of his work is devoted to philosophical foundations of probability. His publications include *Philosophy of Probability*, Kluwer Academic Publishers, 1993.

Eelco Ferwerda is director of the OAPEN Foundation, one of the few international initiatives dedicated to Open Access monograph publishing. Before that he managed OAPEN (Open Access Publishing in European Networks) as EU-funded project at Amsterdam University Press. He joined Amsterdam University Press in 2002 as Publisher of Digital Products. He has been involved in electronic publishing since the early nineties and has published about Open Access and Scholarly communication in a number of journals, such as *Serials* and *LOGOS*. He is currently President of the Association of European University Presses. He received the Dutch SURFshare Open Access award in recognition of his work for OAPEN.

Gudrun Gersmann studied History, Romance Languages, German and Philosophy in Bochum, Geneva and Paris. In 1991 she received a doctorate at the Ruhr University Bochum with the thesis subsequently published as *Im Schatten der Bastille: Die Welt der Schriftsteller, Kolporteurs und Buchhändler am Vorabend der Französischen Revolution*, Stuttgart, 1993. After holding a chair for Early Modern History at the RWTH Aachen and later at the University of Cologne, she became Director of the German Historical Institute Paris in 2007. She has been one of the “pioneers” among German scholars in the fields of e-publishing and digital humanities. Her publications include *Forschung und Lehre im Informationszeitalter – zwischen Zugangsfreiheit und Privatisierungsanreiz* (ed. together with Karl Nikolaus Peifer), Berlin, 2007.

Gábor Kecskeméti is the director of the Institute for Literary Studies of the Research Centre for the Humanities of the Hungarian Academy of Sciences (Budapest), and professor of Hungarian literary history and head of the doctoral school in literary studies at Miskolc University. He is the editor-in-chief of the periodicals *Irodalomtörténeti Közlemények* (devoted to the literary history of Hungary) and *Camoenae Hungaricae* (multilingual, devoted to the history of neo-Latin literature in early modern Hungary). He has published widely on the history of rhetoric and on the relations between Hungarian and European literary culture in the early modern period. His most recent monograph: “*A böcsültre kihadott ékes és mesterséges szóllás, írás*”: *A magyarországi retorikai hagyomány a 16–17. század fordulóján* (Rhetorical tradition in Hungary at the turn of the sixteenth and seventeenth centuries, 2007).

Ferenc Kiefer was born on 24 May 1931, in Apatin. He studied mathematics, German and French linguistics. He graduated in German linguistics in 1965. He worked as a computational linguist at the Computing Centre of the Hungarian Academy of Sciences from 1962 until 1973. From 1973 until his retirement in 2001 he was research professor at the Research Institute for Linguistics of the Hungarian Academy of Sciences. At present he is professor emeritus. During his career he held visiting professorships at various universities (Stockholm, Paris, Stuttgart, Aarhus, Antwerp and Vienna). His research areas include the theory of morphology, semantics (esp. lexical semantics), and pragmatics (esp. the semantics-pragmatics interface). He has published over 200 scientific articles and over a dozen monographs in English, Hungarian, German and French. Ferenc Kiefer is member of several learned societies and academies (Hungarian Academy of Sciences (1987),

Academia Europaea (1993), Austrian Academy of Sciences (1995), Honorary Member of the Linguistic Society of America (1996), Honorary Member of the Philological Society of Great Britain (1998)). He received an honorary doctorate from the University of Stockholm (1992), from the Université de Paris 13 (2001) and from the University of Szeged (2006). Since 2003 he is the president of the Comité International Permanent des Linguistes (CIPL).

Claudine Moulin was born in Luxembourg and educated in the United States, Belgium and Germany. Since 2003, she is Professor of Historical Linguistics at the University of Trier/Germany. She is also Scientific Director of the Trier Center for Digital Humanities. Since 2006 she has been a member of the Standing Committee for the Humanities of the ESF. In 2010, she was recipient of the Academy Award of Rhineland-Palatinate, Academy of Sciences and Literature. Her publications range from the study of medieval languages, language change, cultural history to the domain of dialectology and lexicography to Digital Humanities (a. o. Claudine Moulin - Julianne Nyhan et al. *Research Infrastructures in the Digital Humanities*, ESF, SPB 42, Strasbourg 2011).

Julianne Nyhan is Assistant Professor in Digital Information Studies in the Department of Information Studies, University College London. Her research interests include the history of computing in the Humanities and most aspects of digital humanities with special emphasis on meta-markup languages and digital lexicography. She has published widely in Digital Humanities and beyond, most recently *Digital Humanities in Practice* (Facet 2012), *Digital Humanities: a Reader* (Ashgate 2013) and *Clerics, Kings and Vikings: essays on Medieval Ireland* (Fourcourts, at press). Among other things, she is a member of the Arts and Humanities Research Council (AHRC) Peer Review College, the communications Editor of *Interdisciplinary Science Reviews* and a member of numerous editorial and advisory boards. She is also PI of the 'Hidden Histories: Computing and the Humanities c.1949-1980' project. She tweets @juliannenyhan and blogs at <http://archelogos.hypotheses.org/>. Further information is available here: www.ucl.ac.uk/dis/people/juliannenyhan.

Bernhard Palme received his PhD in Ancient History in 1989 from Vienna University. His research focuses on editions and historical analysis of papyri. His numerous publications on documentary texts and the history, economy and administration of the Later Roman Empire include the monographical studies *Das Amt des apaitetes in Ägypten* (1989) and *Dokumente zu Ver-*

waltung und Militär aus dem spätantiken Ägypten (2002). In 1998-2004 he conducted an international project on “New Papyri from Ancient Egypt.” In 2005 he accepted a chair for Ancient History and Papyrology at the University of Vienna. In 2009 he was appointed Director of the Papyrus Collection of the Austrian National Library.

Vera Szöllösi-Brenig is Program Manager at the VolkswagenStiftung, a research funding agency located in Hanover. She studied German and French Literature. She worked for about 15 years as a political journalist at the Bayerischer Rundfunk in Munich and at the Deutschlandfunk in Cologne. Meantime she obtained her doctorate in French Literature with the thesis published as *Die ‘Ermordung’ des Existentialismus oder das letzte Engagement: Künstlerische Selbstfindung im Frühwerk von Claude Simon zwischen Sartre and Merleau-Ponty*, Tübingen, 1995. Since 1999, at the VolkswagenStiftung, she has been responsible for the programs “Key Issues for Academia and Society” and “Documentation of Endangered Languages.” She published on documentary linguistics as well as on art at the interface of science and the public. She is interested in current developments of the humanities and social sciences, especially in Digital Humanities.

Milena Žic Fuchs, Member of the Croatian Academy of Sciences and Arts, is Professor and Chair of Linguistics in the English Department, University of Zagreb. She was a member of the Standing Committee for the Humanities of the ESF from 2005 to 2009, Chair of the SCH from 2009 to 2012, Chair of the Scientific Review Group for the Humanities, ESF, from 2013. She was also a member of the SPRU Project Board (*SPRU Report* [...]) and one of the instigators of the report entitled *Towards Comprehensive Bibliographic Coverage of Scholarly Literatures in the Humanities and Social Sciences*. She was the first scholar to introduce cognitive linguistics into Croatian linguistic circles. Her articles in international and Croatian journals cover research topics in semantics, discourse analysis, cognitive linguistics and the influence of communication technologies on language. She has authored three books: *Knowledge of Language and Knowledge of the World* (1991), a dictionary of acronyms (with Stjepan Babić, 2007) and *Cognitive Linguistics and Language Structures: the English Present Perfect* (2009).

Index

- Advanced Papyrological Information System (APIS) 121
Akadémiai Kiadó 91, 97
Allensbach Institute 184
Amazon 79-80, 186
Ambrosiana (Milan) 111
American Academy of Sciences 66
American Philological Association (APA) 124
Anderson, Chris 71
Annan, Kofi 38
Antoniou, Grigoris 53
Aquinas, Thomas 59
Archimedes 48
Arts & Humanities Citation Index (A&HCI) 151
assessment
 bibliometrics 15, 124, 148-151, 153, 155, 158-163, 166, 179-182
 effects of research assessment 11, 14, 147-149, 157, 173
 European Reference Index for the Humanities (ERIH) 14-15, 153-161, 164-165, 173-175, 181-182
 impact factor 39, 148-149, 161, 179-180
 peer review 39, 43, 76, 136-138, 142, 149, 154, 174-175, 181, 194
 questionable aspects 99, 102, 123, 155
 visibility of research 38, 100, 137, 147, 153-156, 158-159, 161-162, 164-166, 174, 189, 198
Association of American University Presses 12
- Bachet, Claude-Gaspard 34
Banu Musa brothers 48
Bauhaus 48
Bavarian State Library 42
Bédier's principle 106
Benjamin, Walter 13, 30-31
Berlin-Brandenburgische Akademie 197
Birkerts, Sven 48
Bogdal, Klaus Michael 193
Bolzano, Bernard 33
Bonfadelli, Heinz 186
Borghesian Library of Babel (Borges) 69
Borgman, Christine 132
Bornmann, Lutz 150
Bot, Kees de 179
Brand, Stewart 71
Brin, Sergey 75
British Library 74
British Medical Journal (BMJ) 161
Busa, Roberta 59
- Canfield, Kip 54
Cardano, Gerolamo 105
Carnegie Mellon University 73
Carr, Nicholas 84
Carroll, Lewis 81
Centre for Electronic Publishing (ZEP) 42
Chartier, Roger 75, 84
Chen, Steve 79
Clarke, Arthur C. 72
Clemenceau, Georges Benjamin 21
Clossey, Luke 77
commensurability 22, 30, 32
Crane, Gregory 58
critical editions 91-93, 95-96, 104, 107, 123
Curtin, Philip D. 70
- Daedalus* 66, 68, 71, 80
Darnton, Robert 73, 82, 187
databases, databanks 22-23, 72, 94-96, 109, 119-120, 122, 124, 153, 158-163, 165-166, 173, 175
Descartes, René 10
Deutsche Forschungsgemeinschaft (DFG) 187, 190-191
Deutsche Nationalbibliothek 188
digital humanities
 as rematerialization 13, 24
 digital dictionaries 52, 56, 58
 digital lexicography 13, 47, 53, 58, 206
 digital libraries, e.g. JSTOR, Google Books 63, 72, 75, 120
 history 133-34
 'neo-Luddite' reactions to 12, 48
 philosophical foundations of 13, 21
 relation to philology 14, 23, 50, 99-101, 106, 111, 117, 194
Digital Linguistic Atlas of Luxemburgish 53
Diophante d'Alexandrie 34
Dissonline 189
Documentation of Endangered Languages (DobeS) 196-197
Duke Databank of Documentary Papyri (DDBDP) 119
Duke University 119
- Eco, Umberto 186
editing
 and late Aegyptian papirology 14, 101
 challenge for editors and publishers 14, 107, 111, 138, 178, 190
 new methods of, in publishing the classics 92-93, 100-101
 new problems of 14, 104, 107, 121, of manuscript sources 27, 99, 107
electronic journals and periodicals 13-14, 37, 39-40, 42-43, 45-46, 80, 120, 132, 139-140, 155-158, 173-177
Encyclopaedia Britannica 77
Engels, Friedrich 162-163

- Ernst, Richard R. 149
 European Interactive Advertising Association 186
EuroScientist 149
- Fermat, Pierre de 34
 Foundation of German Humanities Institutes
 Abroad 42
Francia 40-43, 45
 Frankfurt Book Fair 72
 Freeman Center, SPRU, University of Sussex
 SPRU report 158-161, 164, 168
 Freeman, John 71
 Freud, Sigmund 23
 funding
 European Science Foundation (ESF) 14,
 16-17, 59, 153-155, 158
 Fritz Thyssen Stiftung 192
 'opus magnum' 16, 183, 192-196
 VolkswagenStiftung 16, 183-184, 187, 192-198
 Furet, François 69
- Galilei, Galileo 47, 105
 Gensfleisch, Johannes 64
 German Association of university presidents 187
 German Historical Institute 42
 German Research Foundation 44
 Givler, Peter 12
 Gladwell, Malcolm 71
 Godin, Seth 79
 Goethe Dictionary 51
Granta Magazine 71
 Gregoriana (Rome) 111
 Grimm brothers' 'German Dictionary' 51
 Grimm, Dieter 198
 Gruber, T. R. 54
Guardian, The 157
 Gutenberg, Johannes 64
- Hanham, Harold J. 66, 69-71, 82
 Hanser Literaturverlag München 192
 Harmelen, Frank van 53
 Harvard College Library 68
 Harvard University Library 187
 Hauffe, Heinz 194
 Hausenblas, Michael 56
 Heidelberger Gesamtverzeichnis der
 Papyrusurkunden Ägyptens (HGV) 120
 historians
 new methods of 13, 72-73, 79
 new generational differences in attitude 38, 73, 75, 79-81, 111, 133
 their data 65-69, 74, 79-81
 Himmelfarb, Gertrude 12, 48
 Hörisch, Jochen 192
 Humanities Citation Index 175
 Hungarian Academy of Sciences (HAS)
 Textological Committee of 91-93, 96
- Hunt, Tristram 74
 Hurley, Chad 79
 Huygens Instituut (KNAW) 112
 HyperNietzsche project 113
 hypertext, hypertextuality 56, 78, 81
- Instituto di Linguistica Computazionale – Consiglio Nazionale Ricerche (ILC-CNR) 106
 interdisciplinarity 16, 32, 53
 International Mathematical Union 148
 Internet Archive 58
 ISI Thomson 175
 Ithaka report 133-134
- Jordan, Ryan 76
 Journal Impact Factor 149
- Karim, Jawed 79
 Keen, Andrew 76-77, 79, 84
 Kronick, David A. 57
 Kuhn, Thomas 9
- L'Actualité* 78
 Languages Archive, the (TLA) 197
 Lanier, Jaron 77, 79, 84
 Larsen, Birger 153
LexicoLux project 52
 Leuven Database of Ancient Books 119
 Library of the Institute of Philosophy,
 Warsaw 25
 Luhmann, Niklas 198
 Lunar Men, the 48
 Luwel, Marc 180-181
- Martin, Ben 158
 Marx, Karl 23
 Max Planck Institute for Psycholinguistics 196-197
 Max-Planck Society 187
 McArthur, Tom 57
 McCarty, Willard 49
 McGann, Jerome 11, 49-50
 Menninghaus, Winfried 192
 Minuti, Rolando 84
 Mocikat, Ralph 199
 Modern Language Association (MLA) 9, 15
 Moed, Henk 150-153, 163-164, 180-181
 monographs
 as an endangered species 15, 144, 175-177, 188
 as the 'holy grail' for promotion 15, 40, 131-135, 139, 151, 191, 194, 198
- Mumma, Gordon 48
 Münchner Kreis – National IT summit of
 Germany 186
- National Journals Category 165
Nature 163-164
 Nederhof, Anton 179-181

- Netherlands Organization for Scientific Research (NWO) 139
- Network of Middle High German dictionaries 51
- new devices
 e-books 37, 99, 131, 186
 iPads 21, 37
- New Tech Network 84-85
- New Technology Foundation 78
- New Yorker* 71
- Neyer, Jürgen 193
- Nietzsche, Friedrich 23
- Nunberg, Geoffrey 72, 75-76
- Old Testament, the 183
- Open Access
 benefits of 14, 38-39, 53, 121, 131-132, 134-138, 140-143, 187
 Gold 144
 Green 144
 Open Access Publication in European Networks, the (OAPEN) 14, 136-139, 142
 Open Humanities Press 139-142
- Packard Humanities Institute (PHI) 119, 125
- paradigms
 Alfred North Whitehead's approach to 10-11
 definitions 9
 in the history of science 9-10
 paradigm shifts 11-12, 16, 132, 183-184, 196
 Thomas Kuhn's approach to 9
- Perelman, Lewis 48
- Perseus Digital Library 57
- Pöppel, Ernst 183
- Princeton-Stanford Working Papers in Classics 123
- 'publish' as a verb
 'publicare', its meaning in antiquity 10
 'publikál', its implications in Hungarian 10
 samizdat 10
- Pulsifer, Simon E. 76
- Ranke, Leopold von 63
- recensio.net 44-45
- reproducibility 13, 30-31
- research
 as collective and unfinished 14, 43, 94
- Reuter's Arts 175
- Rome's National Library (Biblioteca Nazionale) 73
- Royal Netherlands Academy of Sciences (KNAW) 197
- Samuel Beckett genetic edition project 113
- San Francisco *Declaration on Research Assessment* (DORA) 149
- Saussure, Ferdinand de 107, 111, 114
- Schekman, Randy 149
- SCOPUS 151, 153, 164-165, 175
- Shakespeare, William 81
- Sivertsen, Gunnar 151-153
- Smedt, K. de 47
- Smith, David 57
- Società Dantesca Italiana 106
- Standing Committee for the Humanities (SCH) 14, 16, 17, 154, 158, 206-207
- Stekeler Weithofer, Pirmin 194-195
- Stone, Lawrence 69
- Strugnell, Lilly 37
- Swan, Alma 134
- textual criticism
 new perspectives of 14, 91, 99, 101
 Thesaurus linguae Graecae (TLG) 118-119
- Thornton, John K. 63
- Trismegistos* 126
- Twardowski, Kazimierz 24-27, 33
- Université de Liège
 Centre de Documentation de Papyrologie Littéraire 119
- University of Trier 50
- University of Virginia 49
- Vallisneri, Antonio 105
- Vaucanson, Jacques de 48
- Vinci, Leonardo da 48
- 'Von Humboldt' university model 176
- Walter, W. Grey 48
- Whitehead, Alfred North 10-11
- White Rabbit syndrome 81-82
- Wiberley, Stephen 178
- Wikipedia 15, 43, 45-46, 72, 76, 79
- Winner, Langdon 47
- World Wide Web
 and the historian 48, 58, 64, 67, 121
 Semantic Web 54
- World Wide Web Consortium (W3C) 54
- Worton, Michael 157
- Wörterbuchnetz, das* 58
- Zimmermann, Andreas 193
- Zwaan, Rolf 179

The changes we have seen in recent years in the scholarly publishing world – including the growth of digital publishing and the changes to the role and strategies of publishers and libraries alike – represent the most dramatic paradigm shift in scholarly communications in centuries. This volume brings together leading scholars from across the humanities to explore that transformation and consider the challenges and opportunities it brings.

Péter Dávidházi is head of the Department of 19th-century Literature at the Research Centre for the Humanities of the Hungarian Academy of Sciences, and professor of English Literature at Eötvös Loránd University, Budapest.

“This is a timely and judicious collection of essays that examines what is really at stake in the tensions between our established print cultures and the emerging ‘digital humanities.’”

— *Thomas Docherty, Professor of English and Comparative Literature, University of Warwick*

“Combining close attention to material circumstances with a broad cultural and philosophical outlook, the contributors to *New Publication Cultures in the Humanities* offer a bracing vision of scholarly research as an open-ended and collaborative enterprise – a vision that this stimulating collection both advances and exemplifies.”

— *David Damrosch, Ernest Bernbaum Professor of Literature, Harvard University*

“As the Gutenberg Parenthesis is closing, Humanities scholarship that wants its fingerprint to be read for its touch to be felt will have to go digital. This handbook points us in the right direction.”

— *Ortwin de Graef, Dean of Research, Faculty of Arts, Catholic University of Leuven*

AUP.nl

