



Science Meeting – Scientific Report

The scientific report (WORD or PDF file - maximum of seven A4 pages) should be submitted online within two months of the event. It will be published on the ESF website.

Proposal Title: Downstream from the Digital Humanities

Application Reference N°: 5422

1) Summary (up to one page)

The meeting 'Downstream from the Digital Humanities' was convened in Zadar, Croatia to discuss pre-circulated papers on a number of aspects of the topic of scholarly publishing and communication in the digital age. In particular, the group was recruited so as to represent a wide variety of the relevant perspectives required to understand the full ecosystem of scholarly communications, including libraries, funding agencies, publishers and scholars trained in both computer science and traditional humanities. In addition, the group was comprised of both senior scholars (preparing and presenting the papers) and early stage researchers (commenting on the papers and guiding the discussions). In general, the format worked very well, the discussion was focussed and enlightening, and we made substantial progress in our understanding of the issues presented on day 1. The group was encouraged throughout to capture the aspects of the discussion they felt were most important or interesting on post it notes on the walls.

On the second day, the group organised the notes from the previous day into two groups of issues: ones around the types of output the digital humanities was encouraging and ones about systemic or policy issues. Each group then set about creating a workshop output around these issues, with the first group creating a tabular listing of outputs matched to their form of communication and validation and the second creating an outlines which has since been fleshed out into a substantial position paper on the nature of communication and publication issues in the digital humanities.

2) Description of the scientific content of and discussions at the event (up to four pages)

The content of the scientific discussions revolved around the issues raised in the eight pre-circulated papers, the titles of which are listed in section 4a below. The issues indicated here were manifold, but can be separated into roughly four categories:

hierarchies and their effect on the individual scholar; the shifting role of institutions and systemic changes in the macroenvironment of scholarship; changes in scholarship and in scholars themselves, including the way in which they view (or are trained to view) their role in society; and finally technical issues (by far the most distinct to digital humanities, and also by far the smallest).

I. Hierarchies and their effect on the individual scholar

Scholarship does not arise in a vacuum, but rather within a complex ecosystem of ideas, people, structures, institutions, markers of esteem (like acceptance at a high profile conference or invitation to sit on a Board) and punishments (like denial of promotion). *In the current climate, many of the wider social drivers toward digital forms of communication and publication of and about scholarship come into direct conflict with the still dominant traditional modes of rewarding that scholarship.* Many of our communal norms for quality are actually proxies dependent on the old model for their relevance: for example, journal impact factors only apply to journal articles, and publisher reputations only apply to books.

This crisis of conscience in scholarly evaluation hits the digital humanities particularly hard: The Catch 22 of new forms of scholarly output is that one wants to feel assured one's work will be recognised, but that recognition is contingent, generally, on a certain familiarity and critical mass of accepted examples. Scholars applying digital methods should not be 'punished' for making this choice by the requirement that their digital work be accompanied by a traditional interpretive essay (with that essay being the only part of the output seen and reused by the community). A renewed requirement to deepen our understanding of what we expect from scholarship is created not just by new methods, but by the **new objects** produced by scholars in Digital Age: Books, journals, blogs, collaborative texts (wikis), databases, algorithms, software, encoding, maps, images, videos, schemas, documentation).

Digital humanities is also shaped by the ecosystemic requirement to share parts of our epistemic process, rather than just its outputs. We foster the scholarship of others by sharing data, sharing tools, sharing algorithms, indeed even by sharing our search results and the connections we make in our research process. But again, the inability to protect, document, or prove the value of these productions not only neutralises any incentive, it actively disincentivises the productive circulation of knowledge at the early phases of our research. The slow production cycle of disciplines still focussed on the monograph as the most important form of output means that 'discoveries' need to be protected for a long time, awaiting the appearance of the work in a validatable form.

To solve this problem would require some very fundamental shifts in the scholarly ecosystem, requiring a return to the first principles of scholarship and the value it brings, clear analysis of the incentives the system provides and to whom, and innovative thinking about how recognition can be given for activities falling outside of the traditional categories. This is no small matter, but its benefits would accrue far beyond the digital humanities. One of the primary issues that will need to be addressed in this context is that of what we consider authoritative in a scholarly sense. Authority comes in many forms: peer review is regarded as an authoritative process for evaluating quality, in spite of its many flaws and the wide range of practices the term represents. Individuals may have a certain authority within a community, either on the basis of their reputation or their position in an organisation. Funding agencies and certain journals or publishers also embody a certain authority. What is flawed about this model is that it focusses on the embodiments of the ability to pass a valid judgement over scholarship, rather than

the innate requirements for the ability to make this judgement. How can new modes of communication gain access to these same symbols of authority, easing their path to acceptance as scholarship?

Along side this issue of how we understand scholarship in emergent formats is the concomitant issue of how we give credit for work done. **Citations practices** have to evolve. The ongoing tradition of quoting an original source, rather than the edition or digital facsimile you may actually have consulted, gives *short schriff* to the hard work of scholarly editors and resources, particularly digital ones. With the inclusion now in style guidelines for the citation of all sorts of works and formats, as well as tools like Zotero to make this process easier, there is no longer any reason for this complete elision of the point of access we use to research materials. New/Digital forms of scholarly output may need to include recommendations for users as to how the resource can be cited (be that in a monograph or within a software code), but we also need authoritative confirmation of the importance of this practice.

And of course, how do we counteract the insidious, transitional misgivings we still seem to have about digital sources not being 'real', and scholarship conducted in a virtual environment some being less worthwhile because it is viewed as having been somehow 'easier' than traditional modes scholarship, involving travel and discovery among dusty records without the assistance of Google translate or our digital camera? Again, if we are to make progress in supporting the scholarship that is appropriate for our age and our disciplines, we will need to return to the primitives of knowledge creation and value those, rather than the romantic vision and symbolic authority of our of accepted proxies.

II. The shifting role of institutions and systemic changes in the macroenvironment of scholarship;

The role the scholarly publisher, traditionally our primary 'gatekeeper' for the validation and production of scholarly resources, once had is splintering. The physical production of tangible book objects was only a small part of the process, so the reduction in importance of this stage in the process alone does not in any way mean that all points in the chain from author to market are being adequately covered in the new landscape.

The acceptance process was and still is a powerful signal of perceived quality, a proxy on which we seem, in spite of our slightly bad conscience, reliant upon. The editing function and rights clearance also must still exist. The creation of a durable object was easy with a book, and much harder with a web publication, tool or software. And the marketing and selling functions also should not be underestimated as a part of scholarly dissemination, in particular as audiences are becoming multiple and varied, from the small community of specialists to works with vast popular as well as scientific interest

All too often, the discussion of the emerging role and responsibilities of these 'gatekeepers' becomes overdetermined by concerns of the cost of providing access to scholarly materials. More and more scholarly materials are now available online (whether created as a digital native object or not), and some research methods (such as those based on data-mining techniques) and collaborative relationships are contingent and reliant upon this availability. Furthermore, even within a largely digital ecosystem, less established researchers or researchers from less affluent countries or institutions may have substantially less access to materials as users. It is therefore of greatest benefit from a researcher perspective to have them **as widely accessible as possible**. There are both ethical and economic arguments for the provision of greater access to scholarship, but we need also to be wary of the turn of the current discussion to APCs (article processing charges) as a solution to the imperatives to provide wider

access to scholarship: while that might ease the situation on the user side, we could easily create a different risk, that is that publication in the best journals will become tied to the author's ability to pay, rather than only the quality of the scholarship.

Finally, we must also give due consideration to policies on the sharing and reuse of research data, rather than just research publications. Open access to data and source code will improve the cycle of innovation and creation in the fields, increase transparency, encourage collaborations, facilitate reuse, and generally create greater value from research than is currently possible. But researchers must know both where and what they can share as well as how they can protect their work. Researchers are now making more active choices about the licensing of their content but there has not been an informed enough debate about the impact of these choices. For example, CC licenses do not work for open software/code, and there has been little consideration of the interoperability of licensing schemes in the digital environment.

The macro level issue lurking behind this status is the fact that copyright laws in Europe must evolve before advanced digital methods can become commonplace. Concerns about access and reuse hamper a lot of scholarship that might be pursued with digital methods. A clear, unified approach to 'fair use' in the digital age is an absolute requirement for progress, and must be prioritised at the highest levels within government, business and the legal community. At the current moment, there is no clear baseline for even the most straightforward work, while at the same time, new digital methods and modes of dissemination/publishing are only increasing in complexity. Text and data mining is a good example: Is the right to read the right to mine? Are researchers currently flouting copyright law and is it worth the risk? Will we see a situation in Europe where such approaches migrate and cluster in certain countries to take advantage of clearer statutes?

Legal issues go beyond copyright as well: researchers are underinformed as well about the privacy implications of their work. This is an emerging issue that will affect access to content but also make it necessary for disciplines to revisit ethical research practices in relation to the use of digital methods.

III. Changes in scholarship and in scholars

Another eddy in the waters of the 'Downstream' results from an increasing need to differentiate between the two differing processes of communication and publication, defined as the difference between making public your data and results (communication) and submitting them to some sort of peer review or other sort of verification by the scholarly community of your results (publication), which may or may not include editing, enriching and enhancing work by the publisher. In particular under the pressures of the rubric on impact and the need to justify research spend in a publicly-funded system, the need for both of these modes is increasing clear, and the relationship between them increasingly muddy. The issues of evaluation and marketing are implicated here, as well as the question of publication format and what to make available (best practice, include citation, code, XML) - form and content are both very much in play in the current environment, which creates particular challenges and opportunities. Without a clear understanding of the interplay between communication and publication, confusion will reign where confidence is sorely needed. Does it make a difference to the status of a digital edition if the author is an institution, such as a library, rather than an individual? Are the curation aspects of that edition a part of or separate to the publication aspects? Is a publication held behind a paywall actually published at all? How should we view the continuous nature of digital publications: if a project is continually updated (generally

agreed to be an advantage of the digital) then when is it published? How can it be 'fixed' or indeed assessed? How should we view the publication of data, of software, of blog posts and Twitter archives?

But we are not only moving toward a different paradigm of communication, but also toward different paradigms of knowledge creation. Collaboration is a term that has come to mean many things in the current environment, from co-creation and co-authoring¹ to the casual sharing of information and validation of others' results that has always occurred within the scholarly communities. Knowledge sharing paradigms are perhaps still primarily imagined as unidirectional processes, flowing from expert to novice, but in reality, the complexity of the research questions being tackled today is such that knowledge is increasingly densely networked, partial and reliant upon multiple intelligences to reach conclusions. This move toward greater integration between disciplines should not, indeed can not, be forced, but when it does occur it should be able to be validated and rewarded. Rewarding collaborative work is more than just an issue of deciding how much credit should go to how many people, however. Collaboration also brings a cross-fertilisation of methodologies, which is productive for enquiry, but creates tensions in a system where senior colleagues may be asked to evaluate the work of others whose epistemological frameworks have been defined according to a foreign idiom (critical theory, at least, was text – but software?) As such, the collaborations at the heart of the digital humanities tear at the fabric of the disciplines and many of the institutional structures that support and organize scholars and scholarship – hardly safe or solid ground. A better understanding of what the various actors in the system, including potential industry and non-academic partners, 'want' and what they 'do' would go some distance to addressing these tensions.

Technical issues

There is a lot of concern in the community about the reliability of digital scholarly outputs: after all, how do we evaluate, indeed how do we even reference, what we can't 'fix'? The guarantor of a book's durability is established in the institution of the library. The existence of multiple copies of a physical object (beginning from the point when the age of print was established) means that these collections provide a perhaps less than systematic but, in the end, relatively trustworthy guarantee that things held as important in their own age will likely be available to the future scholar somewhere when he needs it. We have no such guarantees for the objects being created now, as neither libraries, universities, presses, research centres or national agencies have a clear (funded) mandate to ensure these objects remain accessible, in their current formats and in migrated formats into the future. This fear that resources could disappear, wholly or in part, diminishes the coinage of the digital output, and addressing this difficulty will be a part of the process of ensuring their equal status with traditional publications. Self-archiving is a good strategy for this in many cases, with copies maintained at institutional level, nationally or by an pan European organisation like DARIAH, but will have its limitations if there is a reliance on 'not for profits', lack of semantic encoding, or insufficient sophistication applied in archiving.

¹ <http://www.crash.cam.ac.uk/blog/post/of-coauthoring>

3) Assessment of the results and impact of the event on the future directions of the field (up to two pages)

If there is one thing that the event demonstrated clearly, it is the need for a sustained and focussed further engagement on these topics. The perception of the group was clear: activity aimed at defining and securing the 'downstream' space has not occurred at a pan European level in the way it has, for example, in North America. Activities have occurred at the national level, and in isolated pockets under the auspices of organisations such as DARIAH or LIBER, but for the most part, there is no clear set of European best practices for managing and fostering the advanced production of the digital humanities.

Going forward, the group has committed to producing a few concrete output, which will be targeted toward instigating wider debate and discussion toward consensus on these crucial issues. First, the scientific content of the discussions will be consolidated into a position paper, able to appear either as a stand alone document or as a preface to a larger collection of theoretical essays, bringing together the European and North American experiences and examples. This may be progressed via a further, follow-up event, or via targeted recruitment of key, senior contributors to the volume. Second, the first draft of the taxonomy of research outputs will be refined and released. This taxonomy, created as a part of the work of the meeting's second day, lists both the types of research communication one might expect in a fully-fledged system of scholarly communications emerging from digital methods, as well as the primary manners by which these outputs are being communicated/disseminated and validated. This important work has already been presented at a meeting of the Scientific Committee for the Arts and Humanities of Science Europe, at that agency's request, as it clearly meets a need within the ecosystem for a starting point for developing validation pathways for new forms of scholarly communications.

Clearly conversation about these systemic issues needs to be encouraged at the European level, and the NeDiMAH-funded event had established a clear template for how to manage its fluid boundaries. As the participants return to their usual contexts – be they academic, agency or institutional – each brings with them a greater awareness of the complexity of the issues. We will build upon this solid basis throughout the final phase of the NeDiMAH network and beyond.

4) Annexes 4a) and 4b): Programme of the meeting and full list of speakers and participants

Annex 4a: Programme of the meeting

29 May 2014

9:15 Registration

9:45 Welcome, Introduction of participants

10:00 Overview and introduction to the papers: Each session will be organised as follows: *Authors' introductory remarks (5 minutes each); Responses from the Discussants (15 minutes) Open Discussion (35 minutes)*

10:15-11:15 SESSION 1: Habits of Scholarship and their Impact Downstream:
Scholarly Communications in Cyberspace (Borjes)
Barriers to data driven innovation in Europe: the case of text and data mining (Reilly)

Coffee Break

11:45-12:45 SESSION 2: Collaboration and Sharing of DH Results and Outputs (O'Connor/Kamposiori)
Data and DH: Possible Methods for Extending the Digital Reach of Literary and Historical Humanities Research (O'Connor)
Facilitating scholarly communication in art history: the role of personal collections (Kamposiori)

LUNCH

14:30-15:30 SESSION 3: Forms of Expression
Reimagining the Scholarly Edition (Schreibman)
Survey and Analysis of Basic Social Sciences and Humanities Research at the Sciences Academies of Europe (Leatham)

Coffee Break

16:00-17:00 SESSION 4: Impact and Audiences (Edmond, Holzer, Pehar)

Academic Publishing: New Opportunities for the Culture of Supply and the Nature of Demand (Edmond)
Access and re-usability of digital content as preconditions for the Digital Humanities (Holzer)

19:00 Workshop dinner, Five Wells Slow Food Restaurant

30 May 2014

9:30-10:00 The Scholarly Communications Ecosystem, a review of the working map. What do we make? How do we validate it as scholarship? How do we share and communicate it? What other issues does it raise?

10:00-11:00 Rotating discussion groups (with static moderators)

+

Coffee Break

11:30-13:00 Synthesis and general discussion; Future steps for the NeDiMAH working group and for the working papers group.

Closing Lunch, followed by optional walking tour of Zadar

Annex 4b: Full list of speakers and participants

Maria Manuela Borges, University of Coimbra
Jennifer Edmond, Trinity College Dublin
Angela Holzer, Deutsche Forschungsgemeinschaft
Christina Kamposiori, University College London.
Camilla Leathem,
Alex O'Connor, Trinity College Dublin
Susan Reilly, LIBER
Susan Schreibman, NUI Maynooth
Franjo Pehar, University of Zadar
Bianca Gualandi, Open Book Press
Emma Clarke, NUI Maynooth
Francesca Morselli, Trinity College Dublin
Eva Kekou, University College London