

Scientific Report

Meeting of the NeDiMAH Cross-Cutting Working Group on Scholarly Publishing 25-26 July 2012

1. Event Summary

While it is clear that many of the barriers to more widespread acceptance and proliferation of digital methods in humanities research are internal to the community, many others are not. Scholars must make a calculated decision when choosing to embark on a digital project, not just about their research questions and how best to address them, but about their careers, their institutions and their scholarly record. In spite of a long recognition of the value of digital scholarly outputs, many institutions and national systems still struggle to judge the merit of such outputs and credit their creators accordingly. This WG will seek to assemble an authoritative body of knowledge regarding best practice as defined to date in the evaluation of digital scholarship and its outputs, and through this, advance knowledge, a policy position, and/or at least a best practice resource for institutions to use in supporting scholars.

The goals of the WG are ambitious, but in many ways less concrete than those of the parallel, methodological WGs in NeDiMAH. We therefore propose to pull together this small-scale workshop for the core members of the WG as an opportunity to create a strategy for our activities over the course of 2013-2014. In particular, we will be developing a common understanding of the impact of digital methods on scholarly publishing (particularly important as our group is both interinstitutional and intersectoral.) More concretely, we will plan the focus and locus of the events we will host and the publications we hope to see arise from the WG's activities, in particular:

- two/three events, to be planned for 2013 and 2014
- taxonomy of digital scholarly output formats, and suggested guidelines for judging their quality.
- special issue of a DH-focused journal or book series
- charter on scholarly outputs in the arts and humanities

2. Scientific Discussions at the event

The discussions were quite wide ranging and lively: however, we were able by the end of the two days to consolidate our ideas around two particular key issues.

Key Issue 1: The changing role and locus of 'research gate keepers' in supporting the outputs of new methodologies (eg. "how do we publish scholarship driven by digital methods")

1.1 Revisiting the idea of "publishing" - It became clear in our discussion that the 'crisis in scholarly publishing,' diagnosed most prominently by

Stephen Greenblatt back in 2002, has continued to morph is its reaction to scholarly practice. From the point of view of our WG, the current aspect of this evolution most problematic from the point of view of delivering scholarship based on technical methods is the splintering of the role the scholarly publisher once had. The physical production of tangible book objects was only a small part of the process, so the obviation 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 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. Finally, with the democratisation of publishing itself, came also a raft of difficulties in understanding who was reading what and why. Usage metrics are complex and often flawed, in part because what we know (and what we need to know) about reading in books is not comparable to what we know (and need to know) about reading online. Some tools do exist, but in general, we need more and better statistics about on-line reading, differentiating the access and the reading indicators (as PLOS does), providing article level metrics. Services like COUNTER may provide a starting point, but it too is not quite fit for purpose, being overly library-oriented. In an ecosystem where traditional publishers (with and without their own on-line presence) coexist with independent peer reviewers (like NINES), self-publishers (from individuals to universities), and everything in between, a new understanding of the scholarly communications 'value chain' and the best practice for forging all of its links is a fundamental requirement. This new understanding should be able to encompass all forms of publishing from the traditional to the *avant garde*, utilising the strengths and mitigating the weaknesses of each.

1.2 Reliability and Sustainability. 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 s/he needs it. We have no such guarantees for the objects being created now, as neither libraries, universities, presses 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. Open access is a good strategy for this in many cases, with copies maintained at institutional level, national or by an pan European organisation like Dariah, but will have its limitations as there is a reliance on 'not for profits', lack of semantic encoding, or insufficient sophistication applied in archiving required.

As partial steps toward this goal, scholars require a non-commercial (which is not to say free) **Unique ID Agency**, allowing massive usage, and agreed guidelines for publishing textual and non-textual material, and static and dynamic contents.

1.3 Access. More and more scholarly materials is now available online (whether created as a digital native object or not). Some research methods (such as those based on data-mining techniques) are contingent and reliant upon this availability, and it is therefore of greatest benefit from a researcher perspective to have them **as widely accessible as possible**. Open access doesn't mean free, and key elements of this development would be to create business models for this mode of publishing, and understand when it is inappropriate (for example, in cases where copyright or confidentiality may prevent any publication if open access is the only option). But access is an issue which goes beyond the parameters of the debate around the deposit of scholarly research into trusted public or institutional repositories. Access to materials also encompasses issues of **conservation** (for it is to the analogue originals that many people want access, with the digital surrogate being just that, a surrogate), and **linguistic** availability to scholars who may not have mastery over the language of a particular discourse. While these issues may be beyond the reach of a project with its basis in digital methods, their impact must be recognised and incorporated (if only at a background level) into our discussions.

1.4 Copyright laws must evolve. 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.

Key Issue 2: Ensuring that digital outputs can be understood as scholarship (eg. "how do we evaluate scholarship driven by digital methods") (publication vs. communication across the 'languages of scholarship' and audiences for scholarship).

2.1 Evaluation and Appreciation of Digital Objects as Scholarship. 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... The old proxies of press and journal reputation will not assist us in appreciating these highly influential new forms of scholarly communication, so a part of the solution must lie in an enhanced need for **EXPLICIT METHODOLOGY**:

documented and, therefore verifiable. All too often, technology, once applied to a problem, retreats into the 'black box' and recedes from the discussion. This, however, undercuts the desire for rigorous, repeatable scholarship. The ideal scholarly output would allow others to manipulate the same data and to verify a colleague's results, or produce new knowledge with the same data. To make it possible, the underlying framework, including the modeling of a database and the data themselves, should be made public.

More than anything else, we much embrace the fact that digital outputs are inherently **MULTILINGUAL**, not only in the sense that they may engage different human languages, but also different semiotic systems (eg a map or visualisation protocol), machine languages (eg software code, algorithms or search queries) or knowledge organisation frameworks (eg relational database structure or underlying ontologies). Like the proverbial iceberg, much of the work of scholarship and indeed of its final communication may be invisible to the uninitiated, that is to the monoglot in the scholarship landscape. Yet most of these languages do have accepted and stable norms for evaluation within their 'language communities': how do we make this knowledge accessible to the audience or evaluator of a piece of scholarship without the administrative and financial burden of ever-replicating layers of peer review?

2.2 Communication vs. publication, Part of the difficulty we perceive to be coming 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.

2.3 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 shrift 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. Should standards such as the MLA style, or any other of the myriad options

developed for specific disciplines, include a reference to site of access of a resource? How do we ensure we fully cite collaborative, non- traditional work? Do we need to reassess the demarcation between reference works (like bibliographies) and primary works? Between primary and secondary works?

2.4 Expressing and understanding collaboration. No one ever promoted an editor to full professor on the basis of their work on another author’s book, and yet the need for our collaborators, across disciplines and sectors, is increasing such the emergence of such a practice seems not just possible, but probable. But there is deep discomfort with co-authorship in many places in the academy, in spite of its critical place in the digital methodological approaches and their diverse outcomes.

3. Assessment of the results and impact of the event

The primary aim of this meeting was to establish a cross-sectoral dialogue which would inform the activities of the group over the next two years. In this context, we decided that the core group would need to expand to include at least one further member, representing the perspective of a research funding agency. We also discussed in depth the kind of information about research methods that would be required from the other WGs in order to inform our work more effectively. Most importantly, however, we established a programme of actions to be delivered over the next couple of years as vehicles for the extension of our conversation to the wider community. We believe this to be a robust set of goals able to bring significantly increased awareness to the issues we are tasked to address and advance. While a programme of actions may not in and of itself indicate future impact, the programme of work arising has huge potential significance for our understanding of the methodological implications of the changing face of publishing, and certainly these actions could not have been defined without the meeting held.

Goal	Action
Invite discussion in the community through a ‘thoughtful’ collection of invited essays (potentially using Commentpress to trigger broad peer engagement?)	TOPIC EXAMPLES: The many things a publisher does/did and how this workflow is unravelling; open is not free; the languages of digital scholarship; knowledge organisation; scholarship/evaluation; best practices? Useful Metrics?
Understand Key Issue 1: The changing role and locus of ‘research gate keepers’ (publishers, peer reviewers, self-publishers, etc) in supporting the outputs of new methodologies	EVENT-BASED ACTION: Use this as framework for a large meeting - after NeDiMAH GA Oct/Nov 2013. 15 Minute invited pieces from a number of perspectives, to be potentially released as a PUBLICATION (Special journal issue?)
Provide basic support in terms	Best practice docs and annotated series of

of best practice and resources	links to be published on NeDiMAH site: 'How to' guidelines, also the taxonomy of output types; Possibly also as a print publication, eg in multiple versions for multiple roles
Understand Key Issue 2: Ensuring that digital outputs can be understood as scholarship (publication vs. communication across the 'languages of scholarship' and audiences for scholarship)	EVENT-BASED ACTION: TBD

Annex 1: List of Participants

Dr Jennifer Edmond, Chair of the Working Group (IE)

Dr Susan Schreibman, Trinity College Dublin and NeDiMAH Steering Group (IE)

Dr Linda Bree, Cambridge University Press (UK)

Dr Marin Dacos, Centre pour l'edition numerique ouverte (FR)

Professor Claudine Moulin, University of Trier (DE)

Dr Franjo Pehar, University of Zadar (HR)

Annex 2: Programme of the meeting

25 July 2012

2:00 Arrivals and Introductions

2:15 Discussion of the role of the Scholarly Publishing WG: what is "scholarly publishing anyway?"

3:00 Overview of Nedimah Project (S Schreibman)

3:30 Work so far in this space – state of the art (national/sectoral perspectives)

- Examples of good practice

4:00 Agenda setting for Day 2

6:00 Dinner, Pig's Ear

26 July 2012

10:00 Identifying a landscape of outputs

- what is a "scholarly publication/output," revisited? static and dynamic; nature/locus of knowledge creation
- information for the template - what do we need to know about something to know what it is? I

- 11:00 Events (2013, 2014)
- locations, piggyback opportunities, focus, formats
- small or large (7k, ca. 10 people/20k, ca 20-30 people)
- workshop, summerschool, panel at larger event...
- 12:00 Publications (or outputs?)
- locations, piggyback opportunities, focus, formats
- 1:00 Lunch
- 2:00 Public engagement/Process/Charter/Toolkit(s) – what it could be/should be/shouldn't be
eg allourideas.org
- cycle of academic service (like book reviews)
- what kind of imprimatur would have credibility
- 3:00 AOB: Is this something we need a research infrastructure for?
- 4:00 close