

Open Science, Digital Humanities e il problema della valutazione della ricerca digitale

Fabio Ciotti, AIUCD - Università di Roma "Tor Vergata"

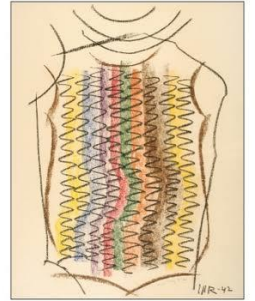
Digital humanities: definition and roots

- DH is the scholarly field defined by the intersection between computational methods, digital resources and humanities scholarship
- A long history: the foundations are commonly recognized in the work of father Roberto Busa on the *Index Thomisticus*
- Then the field has grown slowly but relentlessly for decades mostly in the domain of linguistic and literary computing, history and archeology, under the name of *Humanities Computing*



Digital Humanities: the Next Big Thing

- In 2004 with the publication of the *Companion to Digital Humanities* from Wiley/Blackwell the field undergoes a tremendous expansion in terms of practitioners, funds, status in the academy and broader cultural and societal impact
- Over the last couple of decades the varied field of Digital Humanities have become a major phenomenon both in academic and cultural debate: the Next Big Thing
- Over the years, the community of DH scholars have become so broad and multifaceted, that it is hard to find a shared set of intellectual commons in the field

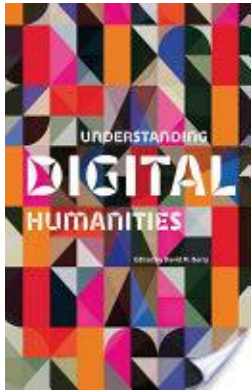


A COMPANION TO
DIGITAL
HUMANITIES

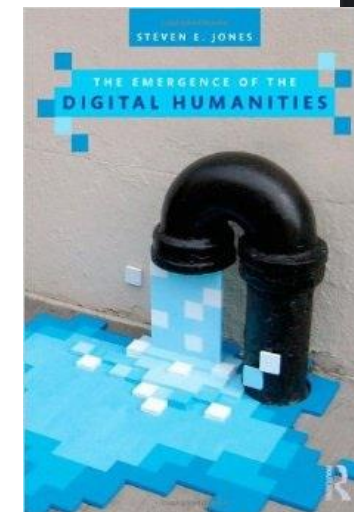
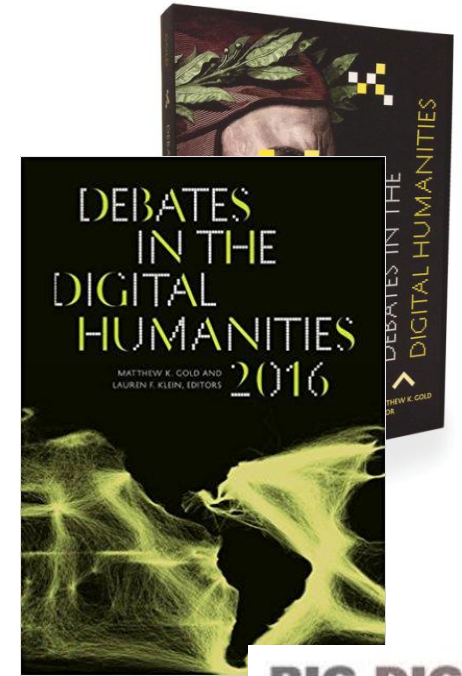
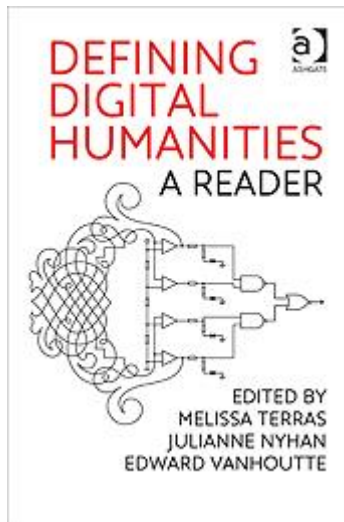
EDITED BY SUSAN SCHREIBMAN,
RAY SIEMENS AND JOHN UNSWORTH

 Blackwell
Publishing

Defining *Digital Humanities*



- The debate over the problem of the disciplinary self-definition is central in DH:
 - The day of DH since its first edition in 2009 asks: 'How do you define DH?'
 - Digital Humanities (Burdick Drucker Lunenfeld Presner Schnapp)
 - Debates in the digital Humanities I e II (<http://dhdebates.gc.cuny.edu/>)
 - Defining Digital Humanities (Terras, Vanhoutte, Hyan eds.)
 - Understanding Digital Humanities (Berry ed.)
 - Big Digital Humanities (P. Svensson)
 - The Emergence of the Digital Humanities (Steven E. Jones)
 - ...



Defining *Digital Humanities*

- “Going meta” it is a typical stance in any new inquiry domain, in particular when it points at becoming an institutionalized discipline
- But it is not only (meta)theory!!!... It’s also institutional and existential materiality:
 - Where you do DH? centers, departments, laboratories...
 - What you teach in DH? The DH syllabus problem
 - Who is doing DH? The problem of education, recruitment and career of DH scholars
 - What is good DH scholarship? The problem of DH research products and projects assessment

Digital Humanities as a Fuzzy Galaxy

- **The Core: Humanities computing/Informatica Umanistica**
 - autonomous meta/trans-discipline, transcending traditional disciplines, focused on theory and methodology
- **The Disk: Digital Disciplines**
 - (Inter-)Disciplines which are fundamentally changed through digital media and computational methods (e.g. Computational Linguistics, Digital History, Digital Literary Studies)
- **The Outer Rim: “Trad” Disciplines with D(H)**
 - (Multi-)Disciplines which incorporate digital methods resources and tools and consider them a part of their traditional canon



McCarty/Short Methodological Commons

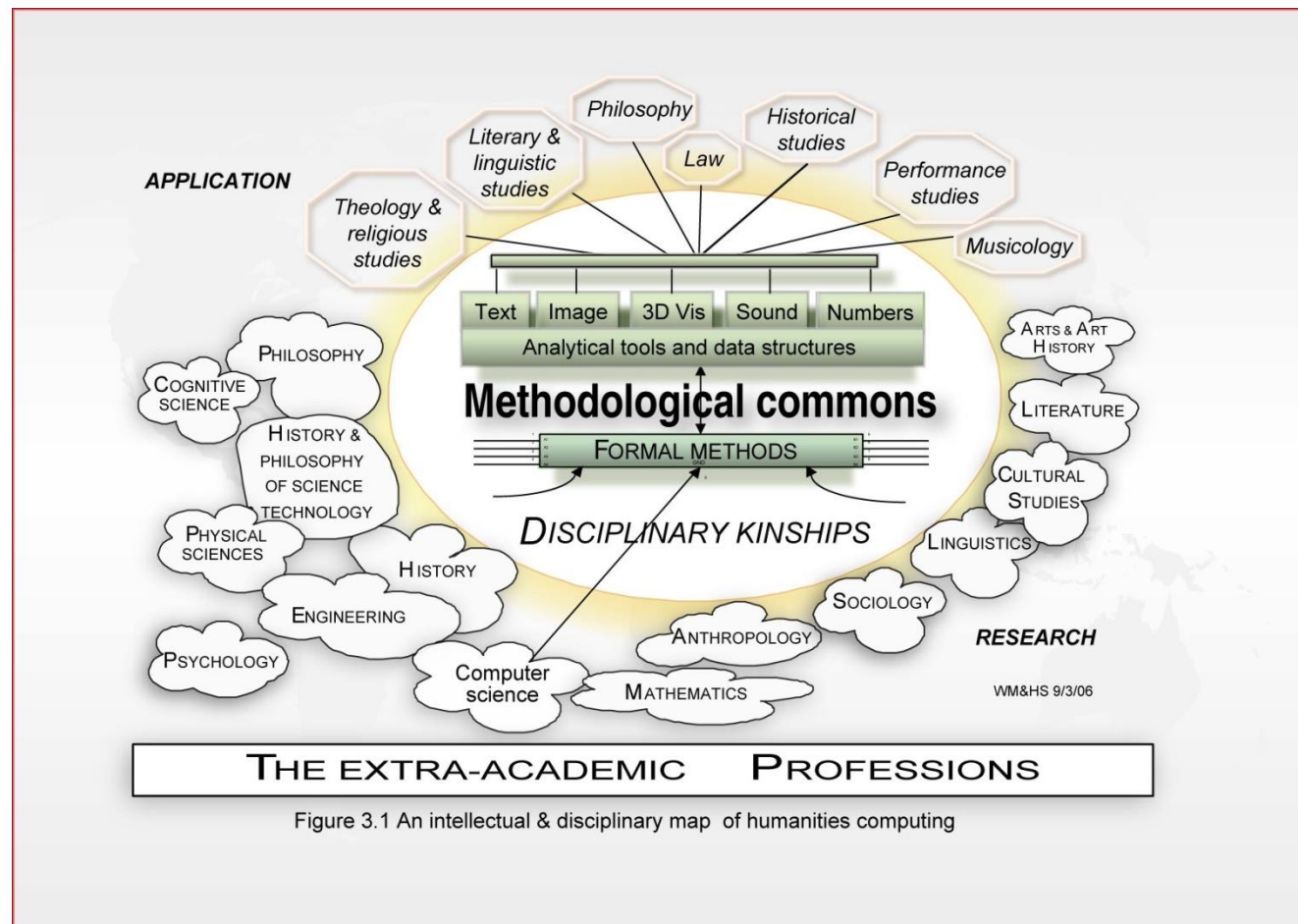


Figure 3.1 An intellectual & disciplinary map of humanities computing

DH: methods and tools

- Two main paradigms:
- Human based computing
 - non-formal (free form annotation)
 - quasi-formal (text encoding, database, GIS)
 - formal/semantic (formal ontologies and Semantic web)
- Machine based computing paradigm
 - Quantitative and statistical/probabilistic approaches
 - Semi-quantitative approaches (eg. rule based Natural Language Processing)
- It's actually a continuum not a neat division...
- Methods are implemented in tools (more and more often developed inside DH scholarship and research activity)

DH: data and digital resources

- DH research **use** data and **produce** (enriched) data
- «Raw» data... in (Digital) Humanities do not exist
 - data building involve theory and modelling
 - Problem => Data Model => Data => Analysis Model => Analysis => Explanation**
 - Even in massive textual digitization in facsimile/non supervised OCR text you have embedded implicit models
- DH methods transform data and enrich them
- Semantic oriented approach is based on the modeling of complex human interpretations and annotations of the data through formal languages: creating and processing **Rich Data**
- DCH institutions also produce rich data in the form of collections with metadata (often in cooperation/synergy with DH scholarship)

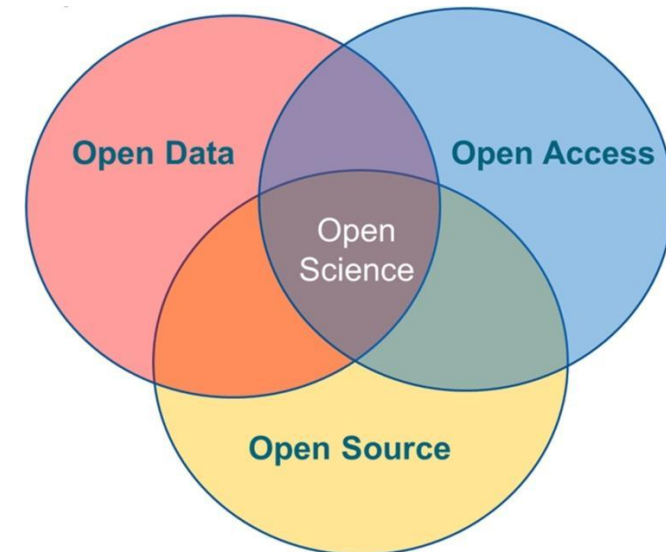
DH an Open science

On the base of this overall description of the field of Digital Humanities we can say that



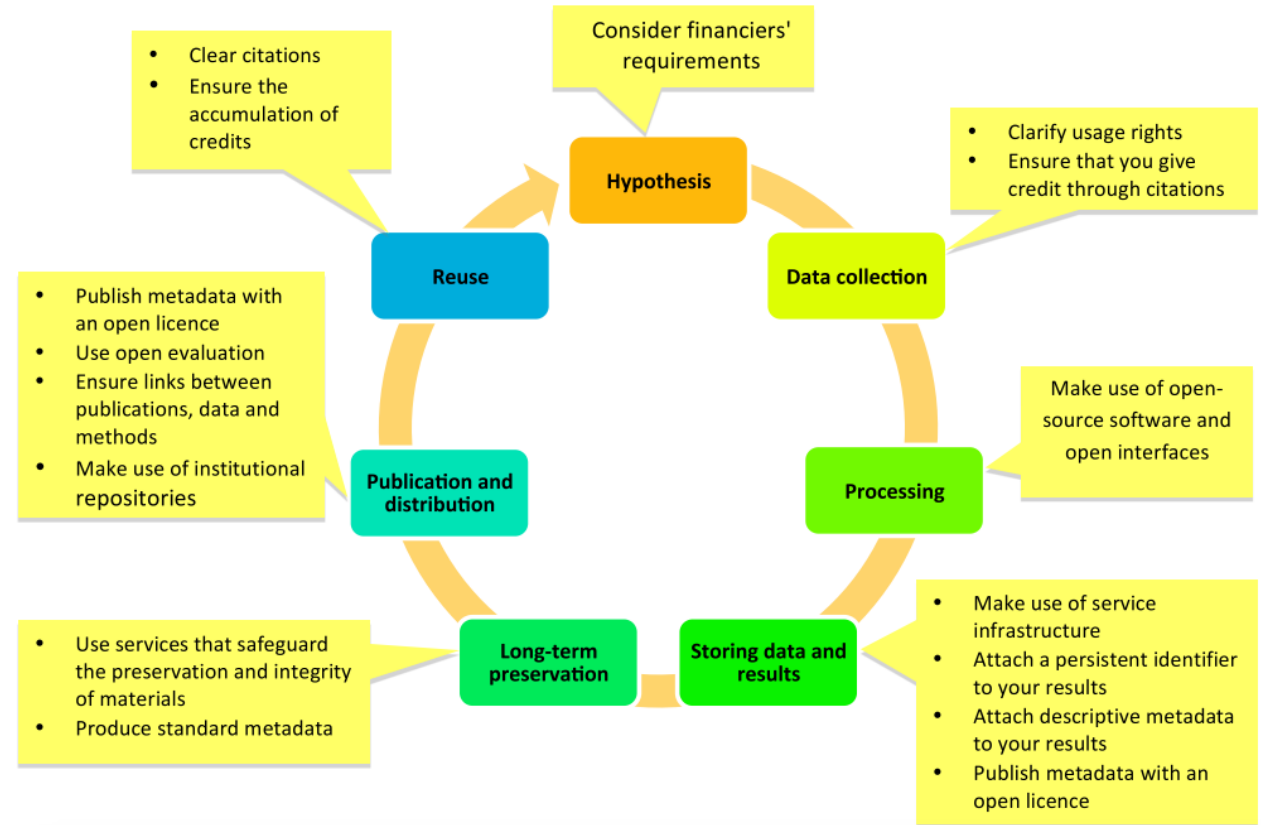
Open Science is a foundational pillar of DH

(or at least it should be...)



Open science: a new vision of scholarship

- *to make the primary outputs of publicly funded research results – publications and the research data – publicly accessible in digital format with no or minimal restriction” (OECD, 2015:7), but it is more than that. Open Science is about **extending the principles of openness to the whole research cycle** (see figure 1), fostering sharing and collaboration as early as possible thus entailing a systemic change to the way science and research is done*



Why Open Science is foundational to DH?

- Because both DH methods and data/resources are digital artifacts that must be
- Co-created (collaborative and interdisciplinary research)
- Published and disseminated
- Shared and reused
- Preserved for a long term
- **Assessed and criticized**
- **Socially valuable and engaged**

DH and open access: a twin relationship

- This deep entanglement is demonstrated e.g. by the historical preference of DH scholarship for Open Access and Open Content research outputs dissemination
 - Most of the national and international DH journals are diamond open access, some gold (APC based). The only exception is DSH, an OUP journal that is moving to OA after PLAN-S
 - Many DH scholars have personal research Blogs Wiki or repositories where they publish in open format the results of their research activity
 - DH scholars are at the forefront in the experimentation of open monographs
 - It's not a case that the DH community has not raised concerns toward PLAN-S like many other Humanities communities (but beware of APC!!)

The role of research infrastructures

- An RI in hard science is a facility or a service that is actually distinct from the content and aims of research, though they can be the output of the research (the CERN LHC is an example)
- A RI in the Digital Humanities is
 - a community and all the community building services implied
 - a set of basic services and institutions that enable the archiving and dissemination of research outputs
 - a set of advanced services that are the reification and standardization of previous research activities in
 - methods-> tools or frameworks
 - domain object digital representation --> data structures, data languages, standards (TEI)
- RI (DARIAH, CLARIN, OPERA, E-RHIS) are strategic to enable Open Science in (Digital) Humanities

DH assessment: the role of Open Science

- The assessment of Digital Humanities scholarship and research cannot be limited to the evaluation of (traditional) publications
- DH produces (mainly) a complex set of digital objects as research products: blueprints and design documents, prototypes, reports, data sets, models, tools, enriched data and visualizations
- Assessing DH research must take into account that complexity: a framework for DH assessment must have access to all those research products in an open environment

DH assessment: the role of Open Science

- DH assessment requires:
 - an open and shared ontology of the roles and functions performed by the researchers involved in a DH project
 - a set of minimum quality requirements (standards compliance, interoperability, scientific control, coverage of the domain, etc.) to assign a digital resource the status of research product
 - an open metric based on these requirements
 - a convergence of different expertise
- All in all these requirements are much better guaranteed in an **open peer review** context
- NB. A deep change is needed in the policies of evaluation agencies, and in the humanities scholarly community

DH and society

- DH can have an impact on the wider social context beyond scholarship
- The role of DH in engaging the public in the scientific endeavor and in leveraging citizen scholarship
- The role of DH in Public Humanities and in contributing to a reflective society
- The role of DH in contributing to a knowledge based economics

DH and society

- DH can foster the societal impact of Humanities at large
- “Digital Humanities (and the funders which support them) have a motivation to think about and articulate what they contribute to society. The question is not whether the professors in the Digital Humanities or traditional departments of Literature and History do scholarship of higher quality. The question is why society supports the study of the Humanities at all and, if so, at what level and in what form. The Digital Humanities are important because new digital media and automated methods enable all of us in the Humanities to reestablish the social contracts upon which we always must depend for our existence” [Gregory Crane, *Why we need an independent Digital Humanities*]