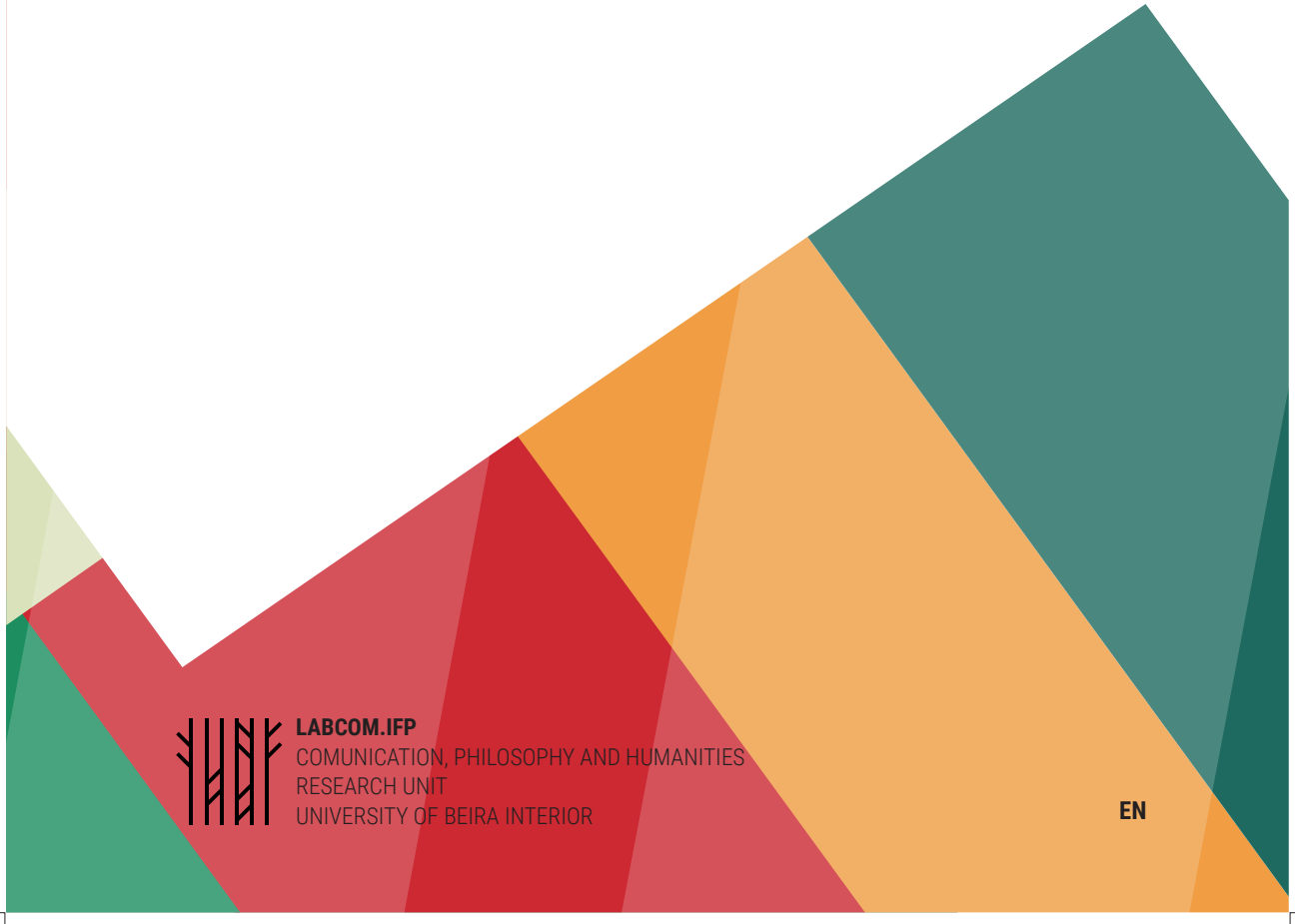
COMMUNICATING AND EVALUATING SCIENCE

Anabela Gradim and Catarina Moura (Org.)



LABCOM.IFP
COMUNICATION, PHILOSOPHY AND HUMANITIES
RESEARCH UNIT
UNIVERSITY OF BEIRA INTERIOR

EN

COMMUNICATING AND EVALUATING SCIENCE

Anabela Gradim and Catarina Moura (Org.)

LABCOM.IFP
COMUNICATION, PHILOSOPHY AND HUMANITIES
RESEARCH UNIT
UNIVERSITY OF BEIRA INTERIOR

Technical Specification

Title
Communicating and Evaluating Science

Organization
Anabela Gradim and Catarina Moura

Collection
LabCom

Series
Research in Communication

Direction
José Ricardo Carvalheiro

Graphic Design
Cristina Lopes
Catarina Moura and Sara Constante (Cover)

ISBN
978-989-654-238-2 (paper)
978-989-654-240-5 (pdf)
978-989-654-239-9 (epub)

Legal Deposit
395070/15

Print
Print-on-demand

LabCom.IFP Books
www.labcom-ifp.ubi.pt

Beira Interior University
Rua Marquês D'Ávila e Bolama.
6201-001 Covilhã. Portugal
www.ubi.pt

Publishing copyright authorizations from both articles and images are exclusively the author's responsibility.

Contents

Introduction 1

Part I
ACADEMIC INTEGRITY
AND RESEARCH & EVALUATION INDICATORS

Predatory Publishers are Destroying the Integrity
of Scholarly Communication 11
Jeffrey Beall

By Albert Einstein, the Editor, you, and me 33
Allen W. Wilhite and Eric A. Fong

Scientific Output in Humanities and Social Sciences:
Problems and Alternatives 49
João Costa

Portuguese universities on the rankings: strategies for improvement 59
Paula Pechincha, António Marques and José António Sarsfield Cabral

The evaluation of the scientific production
in biotechnology, law and arts: proposal of a model 89
Sofia Fernandes

Part II
OPEN ACCESS CHALLENGES FOR SCIENCE POLICY

Predatory publishers and open access models 113
Anabela Gradim

OpenAIRE and the communication of science:
the Open Access infrastructure for research in Europe..... 127
Pedro Príncipe

I know you know I quote: a strategic vision of publications in science 135
João Fernando Ferreira Gonçalves

Wine-growing science and technique – 30 years of challenges..... 149
J. Silvestre; S. Canas; J. Eiras Dias

Part III
THEORIES, METHODS AND CASE STUDIES
IN SCIENCE COMMUNICATION

Investigation in, about and through Art from the point of view
of its publication: epistemological and validation aspects..... 157
Francisco Paiva

Graphic argumentation: diagrammatic modeling
in the communication of science..... 177
Irene Machado

Mapping digital methods: where science and technology studies and
communication studies meet? 203
Chiara Carrozza and Tiago Santos Pereira

In need of a better communication of mathematics in the press..... 237
Susana Simões Pereira, José Manuel Pereira Azevedo
and António José de Oliveira Machiavelo

Notes on the authors..... 249

Mapping digital methods: where science and technology studies and communication studies meet?

Chiara Carrozza and Tiago Santos Pereira[1]

Introduction

The fast growing field of 'digital studies' is densely inhabited by approaches, artifacts, jargons, devices and technologies. This paper aims to explore the extent that different notions and approaches related to digital technologies that are gaining relevance in the social sciences identify different scholarly communities. In particular, it also aims to pinpoint the role of Digital Methods[2] in this bigger picture. The specific interest for Digital Methods follows from the authors exchanges with members of the research team of the Médialab-SciencesPo (hereafter just Médialab), a research center created in 2009 by Bruno Latour, specialized in developing and applying digital tools for social research. The connection with Médialab developed during the FCT Exploratory Project 'The importance of being digital: exploring digital academic practices and methods'. This project - involving a small research team coming from Science and Technology Studies, Anthropology, Communication and New Media – seeks

1) The authors gratefully acknowledge the support granted by the Fundação para a Ciência e a Tecnologia (FCT) to the project "The Importance of Being Digital. Exploring digital scholarship and digital methods" (EXPL/IVC-ESCT /1509/2013).

2) In this paper we capitalize the expression Digital Methods to refer to a specific research agenda in the broader context of digital studies; the wording "digital methods" in small letters refers instead to the keyword used by scholars to specify the topic of their paper or to the methods themselves.

[Communicating and Evaluating Science, pp. 203 - 236]

to explore the transformations that 'being digital' entail for research practices and scholarship in general. The project is developed around two main aims. On the one hand, it aims to explore the role of digital technologies and social media in the academic profession. In this respect, the objects of research are both the actual transformation of research practices and the researchers' perceptions associated with the use of digital media in their work. On the other hand, the project explores the relatively new and fast-growing area of digital methods and tools for social research. An initial objective of the project was to connect these two areas of inquiry by developing some ideas about how digital methods could support or enable the analysis of scale and trajectory of digital academic practices. The project was organized around two distinct sets of training activities held in Portugal - one about digital scholarship and the other about digital methods - involving leading scholars in these fields. The series called Oficinas Digital Methods (ODM)[3] took place in Coimbra in Autumn 2014 and involved three members of the research team based at Médialab. Its format was based on a pair of two-day workshops, held within a two week period, and aimed at the illustration of a set of digital tools and their application to a wide set of research questions. Fifteen researchers were invited and proposed to experiment the tools developed and/or used by Medialab on their research questions or data.

This paper presents a preliminary attempt to apply some of the tools we have been experimenting with during the Oficinas in investigating the emerging field of digital social research, and in mapping this field in terms of theoretical backgrounds and research topics.

Methodology

We built up a corpus of bibliographic references in the Thomson Reuters' platform Web of Science (WoS). We limited the search to the Core Collection and within it, to the 'Social Sciences Citation Index (SSCI)' and 'Arts &

3) See the blog of the project 'Being Digital' for details about the event, http://bedigital.hypotheses.org/432

Humanities Citation Index (A&HCI)', in order to avoid the inclusion of strictly technical papers, not related to the social sciences or humanities. Only articles were included, to consider the types of publications typically associated with new contributions. The final corpus was composed of 1,492 items (extracted as 'Full records' and including cited references), resulting from merging searches into 'Topic' (abstract, keywords and title) on the basis of a list of 18 search terms related to methodologies (ex. 'hyperlink network* analysis'; 'co*link analysis'); tools ('issue*crawler'); social media ('twitter'); approaches ('webometric*'); buzzwords ('big data'; 'digital humanities'). The list of search terms was defined from an initially larger list by analyzing the results of each of the search terms of the starting list and then deciding which terms to keep and which to drop on the basis of the relevance of the results . For example, the search terms 'scraper' and 'scraping' – which refer to a computer software technique for extracting information from websites crucial for digital social research – returned 203 items, but we soon realized that for most of the items these terms refer to archaeological work and research, and, therefore we decided to drop the terms from our indexing list.

The dataset was subjected to a process of cleaning of Authors' name (AU, in WoS abbreviation) and Authors' Keywords (DE, in WoS abbreviation) using the tool OpenRefine[4] (formerly Google Refine), designed to clean and transform datasets. The analysis was performed using three main tools: Sciencescape[5] (a tool for scientometrics working with Scopus and Web of Science files), Table2Net[6] (a tool to convert a table, or a dataset, into different kinds of networks, such as unipartite, bipartite and citation) and Gephi[7] (a network analysis and visualization software package allowing for visual network analysis). These three tools, which are designed to be integrated/combined with each other have all been developed by Médialab researchers and developers; in the case of Gephi, it is being designed by a larger partnership including SciencesPo, that is now a partner of a not-for-

4) http://openrefine.org/

5) http://tools.medialab.sciences-po.fr/sciencescape/

6) http://tools.medialab.sciences-po.fr/table2net/

7) http://gephi.github.io/

profit corporation, the 'Gephi Consortium', including Linkfluence, WebAtlas, and Quid. All the tools are available as open source.

We used the tool in an explorative and inductive fashion. This choice was not only motivated by our recent interest in digital methods scholarship. Rieder-Rohle (2010) discuss digital methods as a set of tools that in several cases are best suited for 'data exploration', meaning that by rendering certain aspects, properties, and relations visible, these tools can offer new perspectives on the phenomena under study. In this respect, even if these are not adequate to verify or falsify hypotheses, they still generate knowledge about the analysed data.

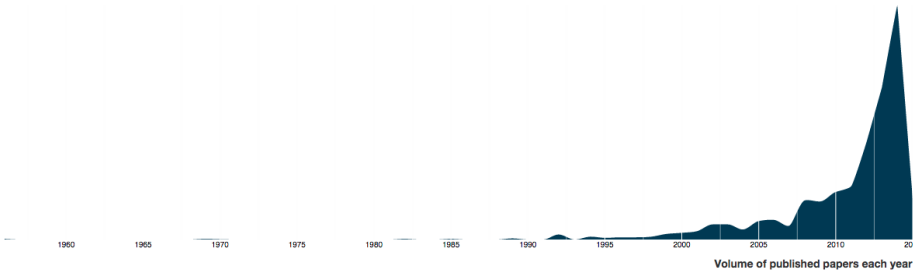
Venturini et al. (2014a, p. 16) note that in developing digital methods for the social sciences, the well known problem of 'experimenter regress' emerges: because of the novelty of these methods, the claims are difficult to ground and both concepts and results appear 'unclear'. There is a risk that visual artefacts aesthetically appealing or intuitively convincing may self-validate. Finding stable criteria to evaluate the artefacts is particularly difficult for a novice in the field. As a solution, Venturini et al. (ibid.) suggest comparing the results of the tools with results obtained in traditional research, in looking for consistency or agreement among them. We adhere to this approach: the considerations we propose in this paper are the result of the triangulation between visual network analysis of the scientific fields around digital social research and a more traditional literature review, enriched by our personal exchanges with leading scholars in the field. In attempting 'visual network analysis' we rely on Venturini et al (2014b) tentative framework based on three variables: position (related to the spatialization of the network and aimed at the identification of clusters), size (related to ranking the nodes, and aimed at identifying *hubs* and *authorities*) and hue (related to coloring the nodes according to defined categorizations, aiming at making further dimensions of analysis visible).

Results of the tool-based analysis

Sciencescape analysis

Digital social research (as defined in our methodology) is a recent trend, and an expanding one, reflecting the novelty of the topic. Most of the papers on this subject have been published since 2010 (cf. Figure 1).

Figure 1 - Papers in the area of digital studies per year



The core of the papers focus on ‘data visualisation’, ‘big data’, ‘webometrics’ and ‘social media’, as identified by their keywords (cf. Figure 2). As expected, these central keywords configure an area which is largely determined by what ‘the digital turn’ is based upon, i.e. digital data and the specific methodologies applied to its analysis, and in particular to its visualisation, which digital technologies have also come to develop. Only after these keywords comes ‘digital humanities’, which are linked to more specific contents. ‘Digital methods’ as a keyword, the starting point of our analysis, is relatively residual, only emerging in the last few years, reflecting the developing identity of this field.

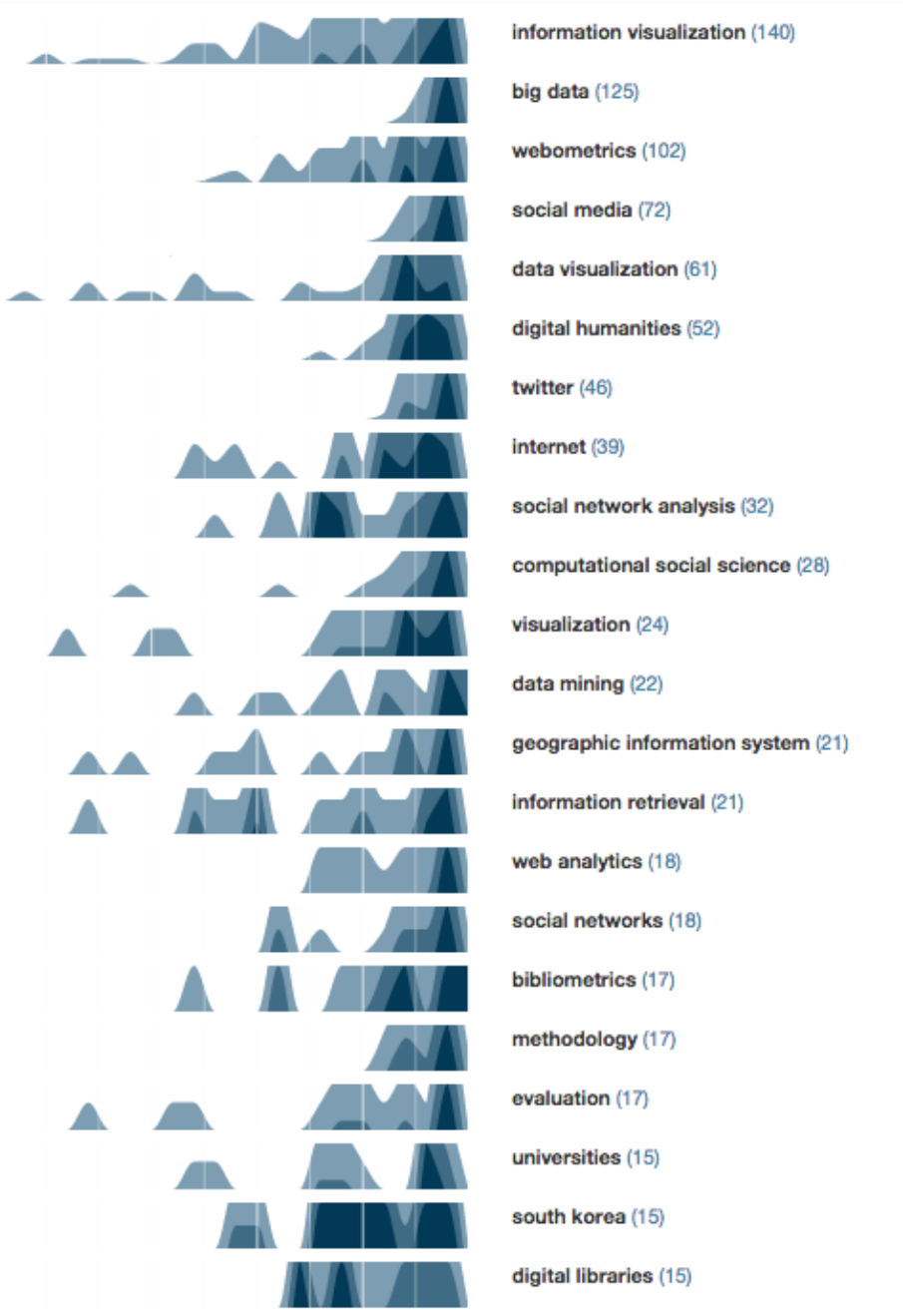


Figure 2 - Top keywords in digital studies

It should be noted that while 'internet' (and related technologies) were prominent in the beginning of the 2000s, 'information visualisation' and 'webometrics' came to lead the area afterwards, until the more recent emergence of 'big data' and 'social media' as central keywords in digital research in the social sciences and humanities (cf. Figure 3). It appears that the possibilities created by social media and related big data are only now being tapped into, and one might expect this will lead to a significant development of the field in the coming years, supported by the visualisation technologies (and related contributions to the development of the research communication processes) that have been a central focus until now.

Figure 3 - Top Keywords in digital studies' papers per year

This thematic approach is also reflected in the journals which have published the identified papers. 'Scientometrics', information science journals, or visualisation and information processing journals have been the main publishing areas. However, it is clear that in recent years other specialized journals have embraced this digital turn and emerge as relevant publication outlets for this work. Journals such as 'Literary and Linguistic Computing', 'New Media and Society', 'Information Communication and Society', 'Historical Social Research', 'PS – Political Science & Politics' or others in health studies fields, show that digital methods are reaching beyond the initial, more technical development of methodologies to a recent more thematic use of their potential, including in the field of communication studies (Figure 4).

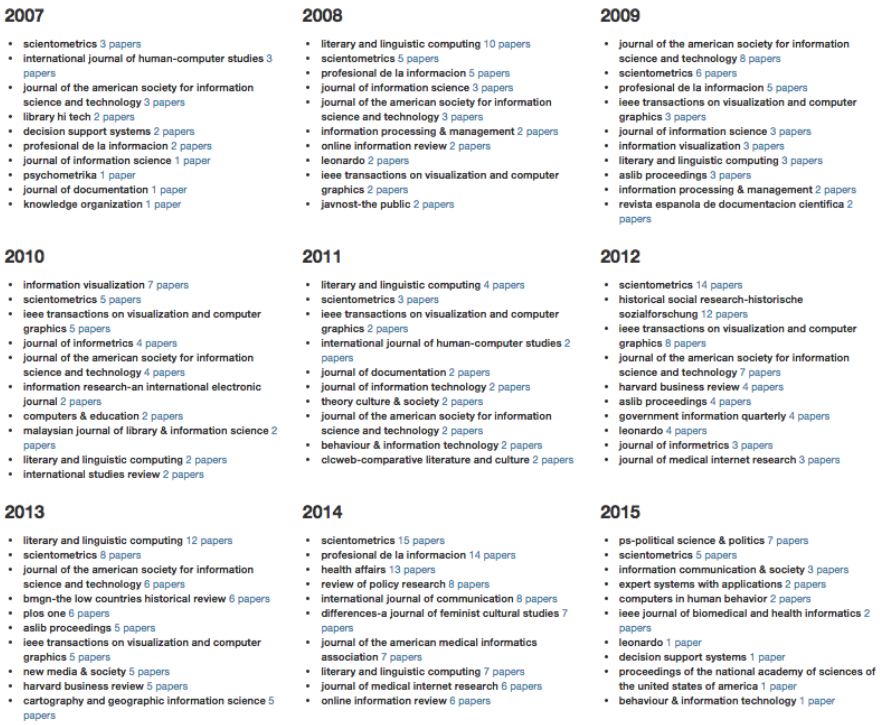


Figure 4 - Digital studies Source Journals per year

Despite this recent trend, it is clear that the area is still developing mostly around the development of methodologies for data processing and visualisation, rather than in its use in wider social sciences research. An analysis of a set of almost 6000 papers citing the digital research dataset, as we defined it, shows that these citing papers largely follow the earlier pattern of the publications, i.e. focusing particularly on journals in the fields of information sciences and scientometrics. However, when looking at keywords, the picture changes slightly. While keywords such as 'visualisation', 'bibliometrics' and 'internet' are dominant, 'social media' also beginning to have significant relevance. 'Social media' emerges as the central keyword in the last few years, perhaps revealing that social media data is a primary signal of a digital turn, competing with traditional, official statistics as a primary source for social science research, and suggesting that 'digital methods' are having significant impact beyond the specific dedicated publications. This turn is also reflected in the significant number of papers citing the set of papers in the corpus we identified.

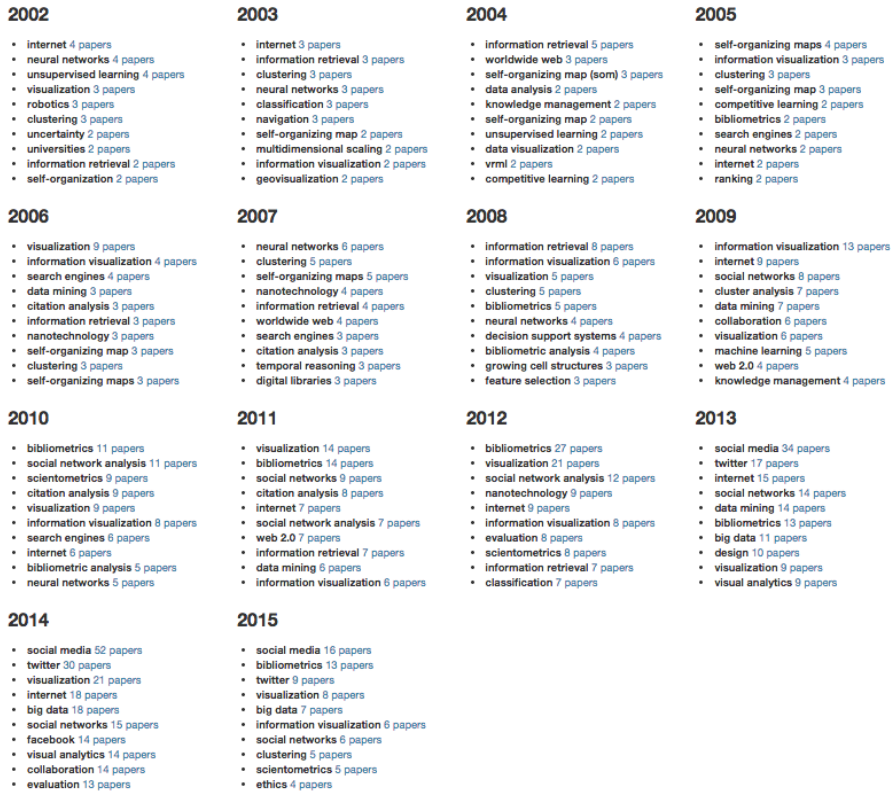


Figure 5 - Keywords in papers citing digital studies dataset

Core of Digital Methods and Digital Humanities

Since our analysis highlights the importance of the areas of information science, visualisation and scientometrics in our digital studies' set of papers, we decided to particularly focus on those papers which reflect a more strict view of digital methods, as related to science and technology studies (STS). We call this new subset 'digital methods core' (DMCore). In doing so we selected 59 papers among the initial set, which are identified by topics such as 'digital methods', 'controversy mapping' (and related terms), or specific tools (such as 'sciencescape'). In addition, we added 194 papers which cite these articles

(excluding self-citations, already included in the initial 59 papers), and that, as such, are expected to reflect the impact of those papers. The number of citations reflects a significant influence of the DMCore, namely taking into account that these papers have mostly been published in recent years, hence with a limited time frame for citation influence.

While the number of papers in this set remains relatively low for a more extensive analysis, it shows that there is a breadth of research that finds intellectual inspiration on digital methods and that goes beyond the specific development of digital technologies and visualisation. Indeed, what is striking in this data compared to the wider set is on the one hand, the diversity of keywords involved, with 'data visualisation' or 'social media' being just two among others, and its relation to STS (and actor-network theory (ANT)) in particular (cf. Figure 6).

2008

- **technology** 1 paper
- **academic dishonesty** 1 paper
- **moral judgment** 1 paper
- **college students** 1 paper
- **patient satisfaction** 1 paper
- **content-filtered speech** 1 paper
- **provider-patient communication** 1 paper
- **voice tone** 1 paper
- **abstraction** 0 paper
- **cheating** 0 paper

2009

- **geometry** 1 paper
- **urban mass transport** 1 paper
- **expertise** 1 paper
- **science and technology studies** 1 paper
- **physician-patient communication** 1 paper
- **universities** 1 paper
- **cartographic animation** 1 paper
- **competency groups** 1 paper
- **digital libraries** 1 paper
- **doric order** 1 paper

2010

- **qualitative research** 2 papers
- **auditory modality** 1 paper
- **participation** 1 paper
- **nature** 1 paper
- **option pricing** 1 paper
- **politics** 1 paper
- **web-based learning** 1 paper
- **levy processes** 1 paper
- **cartography of controversies** 1 paper
- **pedagogy** 1 paper

2011

- **participation** 2 papers
- **science** 2 papers
- **hybridisation** 1 paper
- **cheating** 1 paper
- **plagiarism** 1 paper
- **nature** 1 paper
- **politics** 1 paper
- **internet** 1 paper
- **learning design** 1 paper
- **ipcc** 1 paper

2012

- **web cartography** 4 papers
- **digital methods** 2 papers
- **cheating** 2 papers
- **option pricing** 2 papers
- **learning design** 2 papers
- **levy processes** 2 papers
- **methodology** 1 paper
- **controversy** 1 paper
- **technology** 1 paper
- **environmental governance** 1 paper

2013

- **climate change** 5 papers
- **policy capacity** 5 papers
- **digital methods** 3 papers
- **cartography** 3 papers
- **canada** 3 papers
- **actor-network theory** 3 papers
- **plagiarism** 2 papers
- **adaptation** 2 papers
- **data visualization** 2 papers
- **ethics** 2 papers

2014

- **digital methods** 4 papers
- **social media** 3 papers
- **twitter** 3 papers
- **websites** 3 papers
- **actor-network theory** 3 papers
- **controversy** 2 papers
- **technology** 2 papers
- **methodology** 2 papers
- **digital humanities** 2 papers
- **gis** 2 papers

2015

- **science and technology studies** 3 papers
- **technology** 2 papers
- **idiot** 1 paper
- **social media** 1 paper
- **twitter** 1 paper
- **materiality** 1 paper
- **big data** 1 paper
- **ethnography** 1 paper
- **temporality** 1 paper
- **work** 1 paper

Figure 6 - Keywords of core digital methods papers per year

The main source journals of this specific set reflects also that 'digital methods' is not limited to methodological publications, but rather appears to influence social science research across a wide range of fields (Figure 7). The

diversity of journals also appears to reflect the diversity of case studies which have used the digital methods approach.

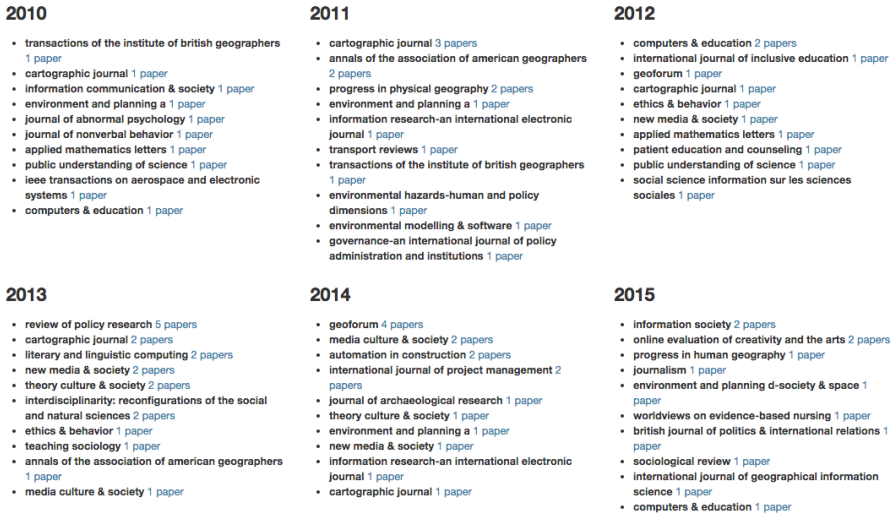


Figure 7 - Source journals of core digital methods papers per year

A similar analysis was developed regarding research in ‘digital humanities’. A core group of papers was identified with ‘digital humanities’ as a topic (175 papers), reflecting a larger and more well established research area, when compared to ‘digital methods’. However, while better established, the ‘digital humanities’ appears to have a lower impact, with fewer papers citing these(112) than the digital methods core set. This may in fact be a result of the clearer definition of the area, and of its lower diversity. This is reflected in the structure of the publications, with more focused source journals dominated by ‘Literary and linguistic computing’ and with top journals including other journals in the digital humanities. Moreover, this is also reflected in the keywords of this core set. As reflected in Figure 8, ‘digital humanities’ has been the top keyword for the last 5 years, along with other strongly related keywords in the top list.

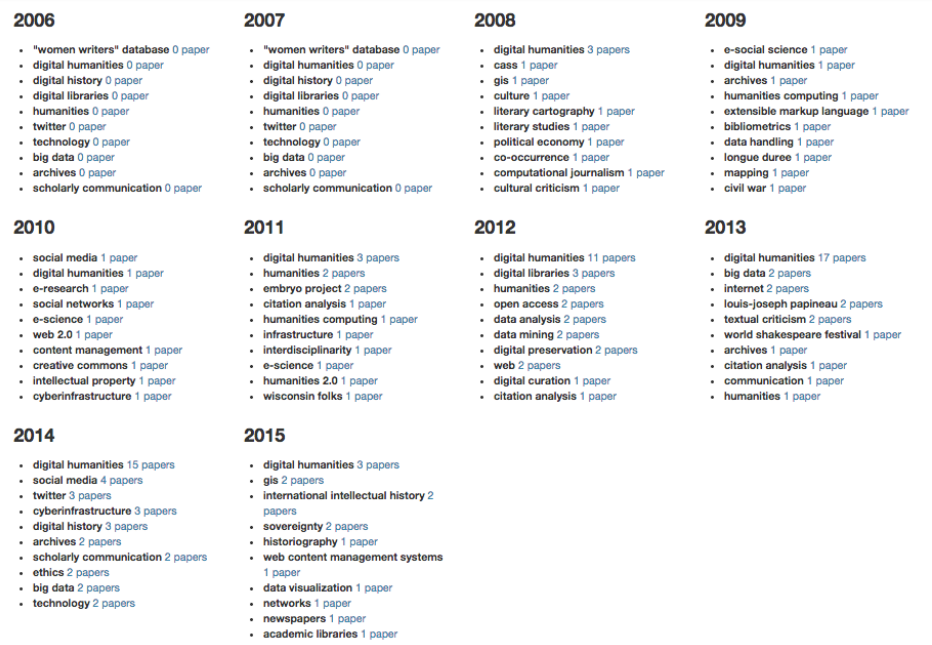


Figure 8 - Keywords of digital humanities papers per year

The analysis of this data reflects the digital methods field as an emerging, dynamic, field, with potentially wide impact well beyond the more strictly defined area. This is clearly distinct from the characterization of the field of 'digital humanities' which appears to have a more narrow impact, and a more institutionalized development.

Visual network analysis though Gephi

By exploring some of the variables of the main dataset (1,492 papers), we attempted to analyze visually the connections between the items composing to our dataset in terms of authors (AU), author's keywords (DE), journals (SO) and journals' area (SC) (the latter field was recoded into Open Refine to reduce the

options). We selected three visual artifacts for the analysis and the discussion (next section).

The first one is a bipartite graph[8] of AU and DE, which has been elaborated to visualize only the giant component[9] and that takes into account only authors and keywords whose occurrence is greater than 1. Being a bipartite network, by selecting an author's node, we visualize all his/her keywords, while clicking on a keyword's node we visualize all the authors using the same keyword. The graph has been spatialized using Force Atlas 2 algorithm whose principle is that linked nodes attract each other and non-linked are pushed apart (see Jacomy et al. 2014). The spatialization makes the position of the nodes meaningful because force-vector algorithms assure that the distance among nodes is roughly proportional to their structural equivalence, that is to say the number of neighbors that they have in common: in other words two nodes are close if they are directly connected to the same set of nodes (Venturini et al 2014b).

The calculation of the modularity[10] (setting the resolution to 2, in order to get the bigger communities) identifies 6 communities of authors and keywords, represented by the different colors in Fig. 9 (size represents the degree of the nodes).

8) In a bipartite graph the nodes are divided into two categories, and each edge connects a node in one category to a node in the other category (Easley and Kleinberg 2010, p. 279).

9) The giant component is a connected component of the graph that contains a significant fraction of all the nodes (Easley and Kleinberg 2010, p. 30).

10) Modularity is one measure of the structure of networks or graphs, designed to measure the strength of division of a network into modules (also called groups, clusters or communities), see https://en.wikipedia.org/wiki/Modularity_%28networks%29

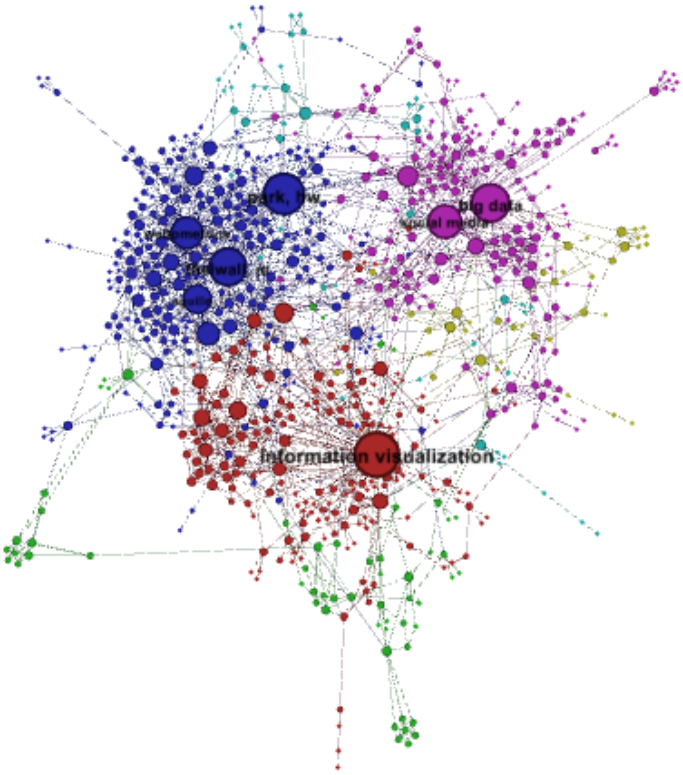


Figure 9 - Bipartite graph AU&DE, overview

There are basically three main clusters of AU&DE. The first one, which is also the one that shows a stronger internal cohesion, appears more dense than the others and develops around the keyword ‘Webometrics’ with the scholars Mike Thelwall, Han Woo Park, Isidro Aguillo and José Luis Ortega as central authors, not only locally, but also within the bigger picture (Figure 10). Several methodological approaches for the study of digital data are located in this cluster, such as ‘hyperlink network analysis’, ‘co-link analysis’, ‘informetrics’, ‘scientometrics’, ‘bibliometrics’. ‘Internet’ and ‘world wide web’ are relevant keywords of the cluster.

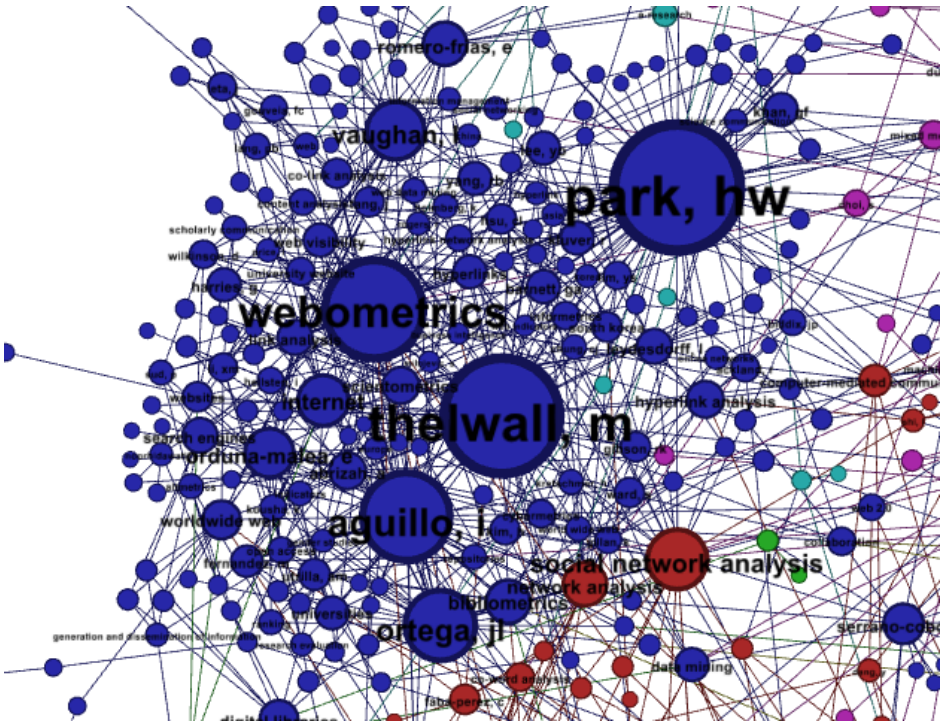


Figure 10 - Bipartite graph AU&DE, 'Webometrics' cluster

A second cluster (Figure 11) is dominated by the keyword 'Information visualization' and appears less dense than the previous one with at least two visible sub clusters: one thematic, around 'data visualization' (below the node 'Information visualization') and one that is characterized by the frequency of authors with Spanish surnames (on the left side of the picture; 'Spain' is also an important keyword of the cluster). The central author is the scholar Félix De Moya Anegon, founder of the Scimago group, who is connected with several scholars that have collaborated with him - and also thematically, with several keywords ('scientific output', 'impact factor', 'co-citation') related to the study of the scholarly communication system and of scientific information databases. This sub cluster bridges the 'information visualization' cluster with the 'webometrics' cluster, and keywords such as 'social network analysis' and

'network analysis' connects the more influential scholars of the two sections of the graph.

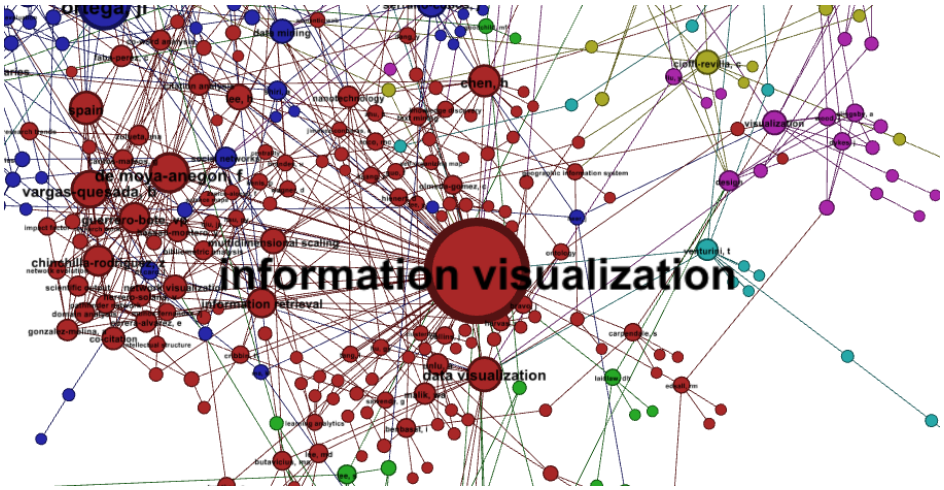


Figure 11 - Bipartite graph AU&DE, 'Information visualization' cluster

A third cluster develops around the keywords 'big data' and 'social media' (Figure 12). It appears more disperse, compared to the previous ones, reflecting the more recent emergence of these keywords in digital research in the social sciences and humanities (see Figure 3; see also the previous section with Sciencescape analysis). 'Twitter' as medium and, to a lesser extent 'open data' as topic of analysis and 'computational social sciences' as methodology, stand out as other relevant keywords of the cluster, while there are no authors clearly emerging as central nodes. The specific keyword 'digital methods' is located in this cluster.

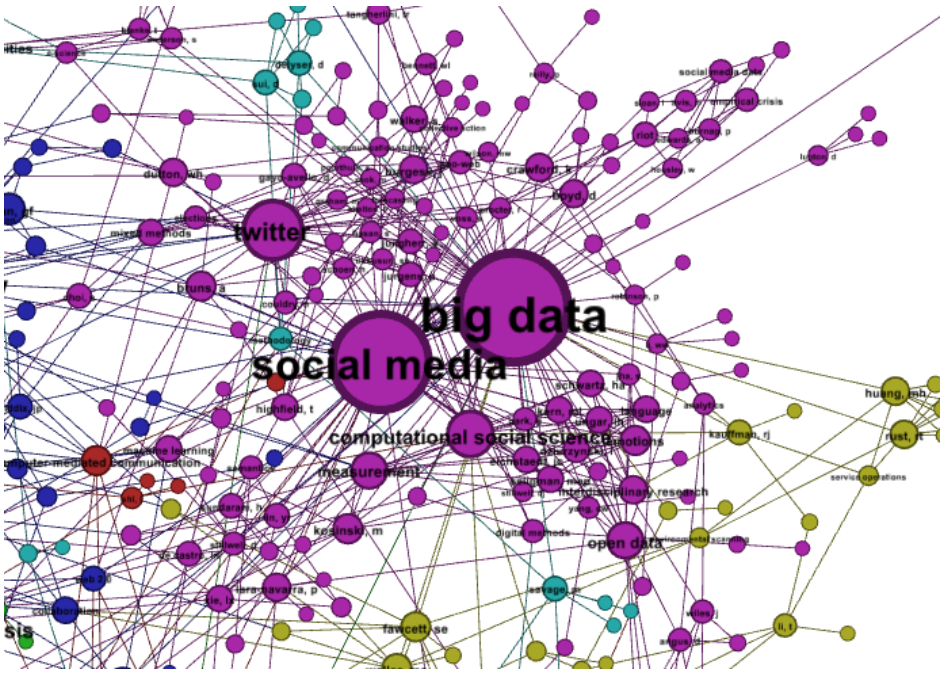


Figure 12 - Bipartite graph AU&DE, 'Big Data-Social Media' cluster

A second bipartite graph representing AU and SO (with SC as attribute) shows the relationships between authors and journals in terms of one connected cluster in a central position - dominated by 'Computer Science and other' journals - and a great number of small clusters disconnected from the central one. By focusing on the central cluster, we can visualize the main authors and the main journals (by modifying the size of the nodes according to indegree and outdegree measures[11]). While in terms of authors we find in this central cluster the main scholars of the 'webometrics' and 'information visualization' clusters of the previous graph, in terms of journals (Figure 13) 'Scientometrics' and the 'Journal of the American Society for Information Science and Technology' dominate the cluster, with a relevant role of other journals classified in the same

11) In the case of a directed network (where ties have direction), it is possible to calculate separate measures of degree centrality, namely indegree and outdegree: indegree is a count of the number of ties directed to the node and outdegree is the number of ties that the node directs to others, see

category of 'Computer Science and other' such as 'Information Processing and Management', 'Aslib Proceedings', 'Journal of Information Science', 'Journal of Documentation'; other important source journals, of the category of 'computer science', are the journals 'Information Visualization' and 'IEEE Transactions on Visualization and Computer Graphic'; 'Profesional de la Informacion', classified by WoS as 'Information Science and Library Science', is also relevant.

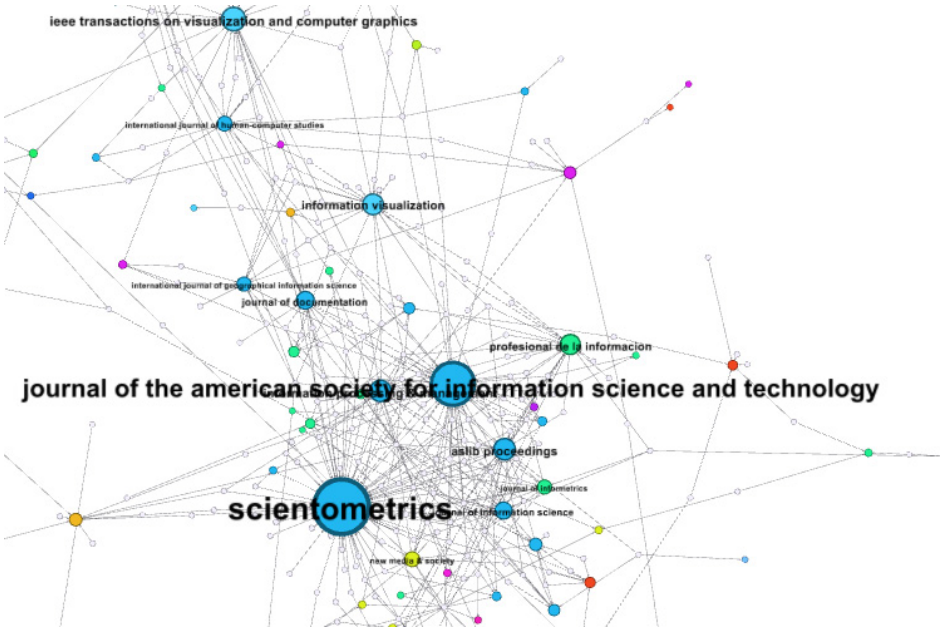


Figure 13 - Bipartite graph AU&SO(SC), main cluster's journals

The last figure (Figure 14) represents the network of papers using the same keywords, that is a monopartite graph of keywords (DE) linked by papers (TI) (the graph takes in consideration nodes whose occurrence >1). The main elements of interest of this graph for our discussion are that the papers using the keyword 'digital methods' do not employ keywords located in the 'webometrics' cluster but are mostly connected to the semantic word of the other two main clusters of 'information visualization' and of 'big data'. With respect to this last cluster,

there are many keywords used in association with digital methods: these refer to methodology ('data collection'; 'digital data'), digital platforms ('Twitter'; 'Facebook'), more consolidated scientific areas ('computational social sciences'; 'digital humanities') but also more traditional research topics in the social sciences ('labour', 'collective action'); interestingly, the very keyword 'big data' is not directly connected to 'digital methods'. On the right side of the picture, 'digital methods' shows connections with a series of keywords clearly related to the field of Science and Technology Studies (STS), such as 'Actor-Network Theory', 'public understanding of science', or 'scientific controversies'.

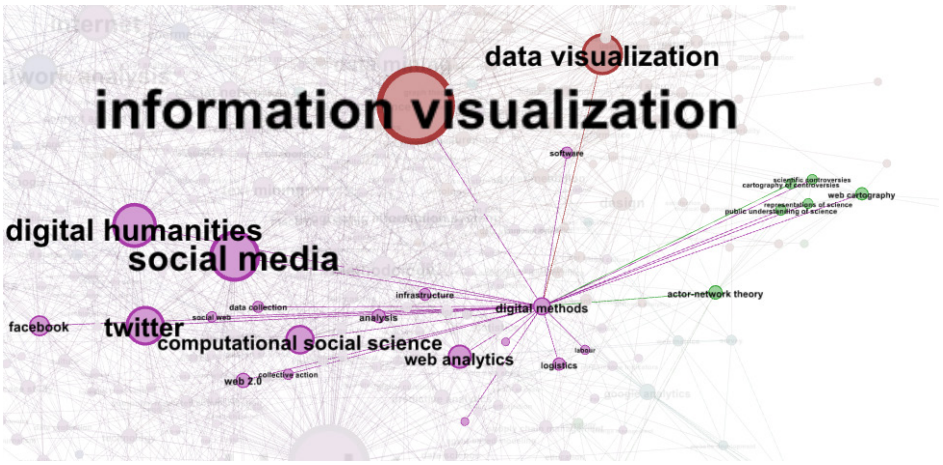


Figure 14 - Normal graph DE by TI, focus on 'digital methods'

Discussion

From the analysis of the dataset through Sciencescape and the visual network analysis we derived some preliminary tracks for the analysis.

1. Digital Methods seems to be located between three distinct areas/ disciplines a) Science and Technology Studies; b) Information visualization and c) Media Studies. .
2. A cluster emerges around other methodologies to study the web such as 'webometrics'; 'social network analysis' and 'hyperlink network analysis'. This appears as the most structured cluster, with the lead authors of the wider area and with strong internal interconnections in terms of keywords and journals of publications. The keywords around 'digital methods' seem particularly far/disconnected from this cluster.
3. The buzzword 'Big Data' is at the center of one of the main clusters, attracting approaches, topics and devices broadly connected with 'media studies'; although 'digital methods' appears connected to several nodes of the 'Big Data' cluster, the two keywords are not directly connected.

How can we make sense of these observations? Do the visual representations and analysis presented above resonate somehow with the literature that we can explicitly trace back to Digital Methods?

First of all we need to characterize Digital Methods in terms of research centers and scholars. In less than one decade, several initiatives and research centers committed to developing and exploring the possibilities offered by digital technologies for social research have emerged in Europe[12]. Among them, it is possible to identify four leading centers. The pioneering center is the 'Digital Methods Initiative' (DMI) in Amsterdam created through the initiative of Richard Rogers, an American Professor of New Media and Digital Culture based at the University of Amsterdam. In Amsterdam, the group of new media scholars led by Rogers have been developing methods, techniques and tools since 1999, starting with the Net Locator and later, the Issue Crawler, a tool for hyperlink analysis. Besides being a New Media research group in Media Studies, DMI

12) According to Ruppert, Law and Savage (2013, p. 29), the list should include: "the National Centre for e-Social Science (NCeSS, UK), now the Manchester eResearch Centre (MeRC); the Digital Methods Initiative (DMI, Amsterdam); the Oxford e-Research Centre and Oxford Internet Institute (UK); the Bartlett Centre for Advanced Spatial Analysis (UK); the Centre for Research on Socio-cultural Change (UK); the medialab (Sciences Po, France); and the eHumanities Group at the Royal Netherlands Academy of the Arts and Social Sciences (KNAW)".

offers a PhD program as well as a well known short-term, intensive, training on Digital Methods in Europe (the Winter and Summer schools that take place every year in Amsterdam).

A second 'hub' is represented by Médialab-SciencesPo, founded in 2009 by Bruno Latour and explicitly inspired by the methodological proposal of the Actor-Network Theory (ANT); among the team members, Tommaso Venturini, Associate Professor at the Institut d'Etudes politiques (IEP Paris) has intensely collaborated with Latour and led the research activities of the center. Joining a multi-disciplinary team composed of social scientists, designers and computer engineers, Médialab applies digital methods for the analysis and visualization of controversies ('mapping controversies', see Venturini 2010 and 2012), extending earlier work on controversies developed at the Centre for the Sociology of Innovation, by Latour, Callon, and others. Médialab has partnered with DMI and other research teams in several projects, such as MACOSPOL (Mapping Controversies on Science for Politics, an EU-funded collaborative research project) and E-MAPS (an EU-funded collaborative research project that builds on the results of MACOSPOL and aims at assessing the opportunities and risks of online communication for the public debate of technoscientific issues).

From a design perspective, another relevant actor in the network of research centers focused on Digital Methods is Density Design – which also has a strong partnership with Médialab, with some researchers associated to both the centers – a research laboratory within the Design Department of the Politecnico di Milano. Density Design specializes in the domain of visual representation, including theoretical and epistemological reflections on visualizations and analysis of their cognitive underpinnings.

Finally, a fourth hub considered is the Goldsmiths College, from the University of London, in the UK, and in particular, the Centre for the Study of Invention and Social Process (CSISP), an interdisciplinary research centre based in the Department of Sociology. The center, offering undergraduate and graduate training in Digital Sociology, is directed by the STS scholar Noortjie Marres, that was part of the team led by Rogers (with whom she has often collaborated) that built the Issuecrawler and is currently developing another tool called Issue Mapping Online.

By exploring the production of the scholars associated with these centers, it is possible to identify three main intellectual traditions and/or scientific debates influencing the development and current frame of Digital Methods, which somehow resonate with the track under 1): a) Actor-Network Theory (an approach developed within the Science and Technology Studies), b) the debate about the performativity of methods (at the intersection of Sociology and the STS); c) the opposition between the 'virtual' and the 'digital' (Media studies).

Actor-Network Theory (ANT), initially developed by Bruno Latour, Michel Callon and John Law, represents a strong root of the emerging scientific community of Digital Methods. As Latour often made clear in his works, the use of the word 'network' in the language of the 'actor-network theory' does not simply designate things in the world that have the shape of a net, but a mode of inquiry – network being used in conceptual terms rather than in concrete ones (Latour 2011). The basic idea conveyed by ANT is that "whenever you wish to define an entity (an agent, an actant, an actor) you have to deploy its attributes, that is its network […] an actor is nothing but a network, except that a network is nothing but actors" (ibid: 5). ANT represents an alternative to the scientific tradition of 'social network analysis', that lies at the basis of other methodological approaches to the study of the Web that we found in the 'Webometrics' cluster (such as Hyperlink Network Analysis, see Park-Thelwall 2003), and in some respect, this could offer a way to make sense and to develop the track 2).

In the long-standing project of ANT, digital technologies stand out for the possibilities they give to 'materialize' a network mode of inquiry. Rogers and Marres (2002: 342) emphasized that "Bruno Latour (1998) argued that the Web is mainly of importance to social science insofar as it makes new types of descriptions of social life possible. According to Latour, the social integration of the Web constitutes an event for social science because the social link becomes traceable in this medium. Thus, social relations are established in a tangible form as a material network connection".

In this respect, Latour and colleagues have been particularly interested in reflecting on how digital traces left by actors inside newly available datasets might help the reformulation of classical questions of "social order" (Latour et al 2012). This exploration is built upon the social theory developed by the

French sociologist Gabriel Tarde (1843-1904) by following an unconventional – at least for the tradition of the social sciences – line of reasoning: instead of asking how the social/collective emerges from the sum of the individuals, the question is somehow inverted: is there a way to define what is durable social order without making the assumption that there exists two separate levels: the micro and the macro, the individuals and the aggregates? A crucial point is that according to Latour and colleagues Tarde's insight never had the possibility to be developed or tested because of the lack of empirical tools adjusted to it: "Our argument is that digital techniques seem to chip away at both ends of what so many social theories take as their indispensable anchors, thereby offering an occasion to illustrate other views of social order" (ibid.: 603). Particularly the research community around the Médialab seems to conceive of digital methods as a route forward, developing social sciences beyond the set of discontinuities that permeate its traditional ways of producing and organizing knowledge: the intensive vs. the extensive (in terms of data); the aggregating vs the situating (in terms of methods); the micro-interactions vs. the macro-structure (in terms of theory), something that 'was an impossible goal just a dozen years ago […] [and now] starts to become more and more realistic as digital technologies spread (Venturini and Latour 2010). In line with the tradition of the Sociology of Scientific Knowledge, Digital Methods are expected to enrich or reframe the classical lines of inquiry about controversies that in the work of the Edinburgh (Barnes and Bloor) and Bath (Collins and Pinch) schools stand out as a crucial engine of science. Typically, controversies are studied when they become visible in the public arena or in specific scientific communities, i.e. after they have emerged: the new set of 'quali-quantitative' digital methods could instead make it possible to follow their process of construction, deconstruction and reconstruction, i.e. to study controversies during their emergence (ibid.). Controversies, therefore, are regarded as a privileged object of inquiry to experiment with the potential of digital methods for social sciences for several scholars working in the field (the two EU funded projects previously mentioned represent the effort to create the context for developing this line of inquiry).

Inspired by ANT's concern for the agency of objects, another debate that seems to be influential for digital social research *a la* Digital Methods is the one

about the performativity of methods. In particular, the research agenda known as 'Social Life of Methods' (Savage 2013) reflects the centrality acquired by research methods in cultural and social debates by approaching them as specific objects of inquiry and exploring them not as a mere 'descriptor' of society, but as something that 'helps to create it anew'[13]. Ruppert, Law and Savage (2013) have recently approached the ways in which digital data and devices are reconfiguring social sciences methods and its very assumptions. One way in which this reconfiguration happens is related to the fact that most of the 'data' used in social research is not directly collected by researchers through specific artefacts such as interviews or questionnaires but emerge as a by-product of other activities; while this is not necessarily new in the history of social research the specificities of *the digital* in this respect deserve to be investigated. In this respect, the authors suggest that "in relation to digital devices, then, we need to *get our hands dirty* and explore their affordances: how it is that they collect, store and transmit numerical, textual, aural or visual signals; how they work with respect to standard social science techniques such as sampling and comprehensiveness; and how they relate to social and political institutions. To tease out these specificities and qualities it is useful to consider, in a historical register, how digital devices compare with other, older, socio-technical devices, and consider the different affordances that they offer in a nuanced manner" (ibid.: 32, our italics). This specific concern for the methodological and epistemological reflections lies at the basis of the particular approach to digital sociology we are discussing, that according to Marres "is not just about theorizing the digital society, and it is not just about applying social methods to analyse digital social life. The relations between social life and its analysis are changing in the context of digitization, and digital sociology offers a way of engaging with this"[14]. Marres (2012) proposes an approach to digital social research as an open-ended process of redistribution of methods among a diverse set of agents, acknowledging the contributions of digital devices, practices and subjects to the enactment of social research. In this

13) http://www.cresc.ac.uk/our-research/social-life-of-methods/

14) See Marres N., *What is Digital Sociology?*, Blogpost on CSISP Online, January 21, 2013, available at http://www.csisponline.net/2013/01/21/what-is-digital-sociology/

perspective, where digital methods are shaped by the social and, at the same time, they do the social, techniques of digital research - such as scraping (Marres and Weltevrede, 2013) – become not merely instrumental tools, but analytic practices.

This perspective resonates with the approach proposed by Rogers, author of the most relevant book-length work on digital methods (Rogers 2013) and initiator of the Digital Methods' research agenda. Rogers focuses on the methodological and epistemological affordances of the dominant devices on the web and aims at taking advantage of them for social research. He suggests *repurposing the methods of the medium* for social research: learning how the dominant devices treat natively digital objects (hyperlinks, tags etc.), which techniques are employed in authoring and ordering information, knowledge and sociality and then thinking along with those devices and treatments so as to recombine or build on top of them (ibid.: 37, our italics).

In clear opposition to the approach transposing traditional social sciences' methods to the 'virtual', Richard Rogers' line of investigation has been oriented to "move Internet research beyond the study of online culture and beyond the study of the users of ICTs only" (ibid.: 4). In situating digital methods in Internet-related research, Rogers (2013) has singled out three main phases: the first one approaches the Web as 'cyberspace', characterizing the 'virtual' as a sort of realm apart from the 'real' that could provide opportunities to redefine politics, identity and even corporality. Starting from the late 1990s-early 2000s, social scientists started to question the earlier approaches to the study of the Web. In order to question the divide between the virtual and the real, scholars started to ground – employing consolidated social sciences methods such as ethnography and interviews – the online by going offline, as in the ground-breaking work on Internet use in Trinidad and Tobago made by ethnographers Daniel Miller and Don Slater (2000) to the study of people using the Internet. Rogers argues that starting from the second part of the first decade of the 2000s a number of scholars have moved from the study of the Internet as the study of the 'online' cultures; he describes this new era in Internet research as: "one that [...] concerns a shift in the kinds of questions put to the study of the Internet. The Internet is employed as a site of research for far more than just online culture.

The issue no longer is how much of society and culture is online, but rather how to diagnose cultural change and societal conditions by means of the Internet. The conceptual point of departure is the recognition that the Internet is not only an object of study but also a source" (ibid.: 21). The key notion elaborated by Rogers to characterize this third phase of Internet-related research is 'online groundedness', referring to a research practice that follows the medium, captures its dynamics, and makes grounded claims about cultural and societal change (ibid. p. 23). Digital methods elaborates and analyzes *digital data* instead of *digitized data*, that is objects, content, devices, and environments (hyperlinks, blog posts, search engines, websites...) that are 'born' in the new medium rather than those that have simply 'migrated' to it (Rogers 2013: 19) (cultural analytics and culturomics are regarded by the author as example of approaches that work with digitized data).

At the same time, Rogers and other scholars of the community of Digital Methods have often highlighted the distinction between 'digital methods' and other approaches that rely on large datasets, and particularly on the so-called Big Data. In a way, Digital Methods could be seen as an emerging research agenda looking for an alternative way of doing social research in the age of Big Data, and this argument provides a way to make sense of our third track for analysis. In contrast with issues of 'correlation' among the huge amount of data available (Bollier and Firestone, 2010) - editor-in-chief of *Wired* magazine Chris Anderson's prophecy that 'the data deluge makes the scientific method obsolete'[15] has become popular - what Digital Methods focuses on is the 'formatted' nature of much of digital data. Keeping central the argument that there is no such thing as 'raw data' (Gitelman 2013), digital methods scholars argue that *practice*, *data* and *analysis* cannot be distinguished in any easy or straightforward way, strengthening the importance of research design in digital research[16]. In other

15) Anderson (2008) argued the provocative case that in an age of cloud computing and massive datasets, the real challenge is not to come up with new taxonomies or models, but to sift through the data in new ways to find meaningful correlations, see Bollier and Firestone 2010: 4.

16) See the conversation between the two computer designers Donato Ricci, associated to Medialab-SciencesPo and Density Design and Pedro Miguel Cruz, researcher based at the Computer Design and Visualization Lab (CDV) of the University of Coimbra http://digitalmethods-seminar.org/qa-between-donato-ricci-and-pedro-miguel-cruz/

words, instead of advocating for a data-driven social sciences, digital methods is mostly concerned with *problematizing* online data and devices.

While this approach does not stress the *size* or the amount of data - suggesting on the contrary the relevance of small datasets - it has important consequences in terms of issues of 'information visualization' (an area that emerges as relevant for Digital Methods in our tool-based analyses, see track 1) and for extension of 'science communication'. The move towards digital social research has been accompanied by "the re-emergence of visualization as key to social analysis" (Ruppert et al. 2013, 36), in contrast with the traditional use of numerical and textual devices within the social science apparatus. In this process "visualization becomes a summarizing inscription device for stabilizing and representing patterns so that they can be interpreted" (ibid.).

If information visualization has been primarily seen as a tool for better understanding the complexity of constituted datasets or for making sense of data, Digital Methods approaches visualization as an argumentative device, as *storytelling*. Visualisation has come to be influenced more and more by art and approached as a way to communicate a concern or to talk about our world, rather than just showing data (Manovich 2011, 13).

Rather than as its point of arrival, visualisation is regarded more and more as part of the digital research process: "the graphic is no longer only the 'representation' of a final simplification; it is a point of departure for the discovery of these simplifications and the means for their justification" (Bertin 1983 in Lima 2011, 73). Accordingly, the research process is being reconfigured in order to make room for new collaborative forms of thinking and making involving social scientists, computer scientists and designers. Similarly, science communication may come to be seen less as a practice that comes after the fact – presenting results – but rather as an integral part of digitally embedded research practices, whereby the digital traces created through communication practices are part of the process of understanding the social embeddedness of our knowledge production practices and of the production of new social science knowledge.

Conclusion

In the present paper we proposed to use the tools developed within the field of Digital Methods to map the emerging field itself. In doing so we expected both to explore the mapping methodologies and to present a useful representation of this emerging field. As reminded by Venturini (2012: 798), "[t]he map is not the (observed) territory, neither should it be." The use of such maps derives not from their mirroring power but rather from their readability. In other words, do the digital traces of the field of Digital Methods provide us with a plausible understanding of the field? The current centrality of social media in producing digital traces led us to question whether Digital Methods could thus be an area where Science and Technology Studies and Communication Studies meet.

The paper therefore proceeded through an initial analysis of the digital mapping of the field. The analysis of the bibliometric data revealed that the broad field of Digital Methods is strongly related to webometrics, internet research and information visualization, and more recently to social media and big data, reflecting strong connections with communication studies, and to more classic bibliometric studies of science and technology, and information science.

However, when defining Digital Methods more strictly, in relation to the definition proposed by Rogers (2013) and others, STS emerges more clearly as a relevant influence. Additionally, this core of Digital Methods emerges as more clearly defined and dynamic than Digital Humanities, which has a greater breadth of influence, but less dynamic in the most recent years.

The visual mapping of the different traces from our set of papers contributes to our reading of Digital Methods. The analysis of authors and keywords identifies three main clusters: one around 'webometrics', with greater internal cohesion and structured around more classic methodologies; another dominated by the keyword 'information visualization'; and a third one around the keywords 'big data' and 'social media', where the keyword 'digital methods' is also located. The analysis of keywords (by papers) also links 'digital methods' more strongly with 'information visualization' and 'big data' clusters, connecting both with traditional social science topics as well as with STS topics, through methodological keywords. Interestingly, these readings do suggest that Digital

Methods articulate particularly relevant links between social media, and its analysis in Communication Studies, 'big data' and STS.

The characterization of the emerging area of Digital Methods, and of its most active centers and researchers, developed in the subsequent section, suggests that the mapping results provide resonance of our reading of the field. The work of Latour, Venturini, Rogers, Marres and colleagues suggest that digital methods is strongly linked to ANT, the debate about the performativity of methods and the opposition of the 'virtual' and the 'digital'. In this sense, our own mapping of Digital Methods, and the connections and differentiations it reveals, are a contribution to our understanding of the field. This mapping does not pretend to present the reality of a still emerging area of research, but rather to be a part of our own research process and of our own (and we could say also, of the readers) process of knowledge production through the digital world. It is not only Communication Studies and STS that might connect through digital methods; digital methods also emerges as a mode of communication of different scientific practices.

References

Anderson, Chris. 'The end of theory.' *Wired Magazine,* 16(7), 2008, available at http://www.wired.com/science/discoveries/magazine/16-07/pb_theory

Bertin, Jacques, Semiology of Graphics: Diagrams, Networks, Maps. Translated by William G. Berg. Madison, University of Wisconsin Press, 1983.

Bollier, David, & Firestone, Charles M. *The promise and peril of big data*. Washington, DC, USA: Aspen Institute, Communications and Society Program, 2010.

Easley, David & Kleinberg, Jon. Networks, Crowds, and Markets: Reasoning about a Highly Connected World, Cambridge (UK), Cambridge University Press, 2010.

Gitelman, Lisa (Ed). 'Raw data' is an Oxymoron. Cambridge (MA), MIT Press, 2013.

Jacomy Mathieu, Venturini, Tommaso, Heymann, Sebastien & Bastian Mathieu. 'ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software'. *PLoS ONE* 9(6), 2014.

Latour, Bruno, 'Networks, Societies, Spheres. Reflections of an Actor-Network Theorist'. Keynote Lecture, Annenberg School of Design, Seminar on Network Theories, February 2010, *International Journal of Communication,* 5, 2011, 796-810.

Latour, Bruno, Jensen, Pablo, Venturini, Tommaso, Grauwin, Sebastian & Boullier, Dominique. 'The Whole is Always Smaller Than Its Parts: A Digital Test of Gabriel Tarde's Monads'. *British Journal of Sociology*, 63(4), 2012, 590-615.

Lima, Manuel, Visual Complexity: Mapping Patterns of Information, New York, Princeton Architectural Press, 2011.

Manovich, Lev. "Foreword to Visual Complexity: Mapping Patterns of Information, in Manuel Lima, Visual Complexity: Mapping Patterns of Information, New York, Princeton Architectural Press, 2011, 11–13.

Marres, Noortje, & Weltevrede, Esther. 'Scraping the Social? Issues in live social research.' *Journal of Cultural Economy*, 6(3), 2013, 313-335.

Marres, Noortje. 'The redistribution of methods: on intervention in digital social research, broadly conceived.' *The sociological review.* 60(S1), 2012, 139-165.

Miller, Daniel & Slater, Don. The Internet: An Ethnographic Approach. Oxford, Berg, 2000.

Park, Han Woo & Thelwall, Mike. 'Hyperlink Analyses of the World Wide Web: A Review'. *Journal of Computer-Mediated Communication*, 8, 2003.

Rogers, Richard & Marres, Noortje. 'French scandals on the Web, and on the streets: A small experiment in stretching the limits of reported reality.' A*sian Journal of Social Science,* 66, 2002, 339-353.

Rogers, Richard. *Digital Methods.* Cambridge (MA), MIT Press, 2013.

Ruppert, Evelyn, Law, John & Savage, Mike. 'Reassembling Social Science Methods: The Challenge of Digital Devices', *Theory & Culture & Society*, 30(4), 2013, 22-46.

Savage, Mike. 'The 'social life of methods': A critical introduction.' Theory, Culture & Society, 30(4), 2013, 3-21.

Venturini, Tommaso & Latour, Bruno.'The Social Fabric: Digital Traces and Quali-quantitative Methods'. *Proceedings of Future En Seine*, Cap Digital, 2010.

Venturini, Tommaso, Baya Laffite, Nicolas, Cointet, Jean-Philippe, Gray, Ian, Zabban, Vinciane & De Pryck, Kari. 'Three maps and three misunderstandings: A digital mapping of climate diplomacy'. *Big Data & Society*, July-September, 2014a, 1-19.

Venturini, Tommaso, Jacomy, Mathieu & De Carvalho Pereira, Deborah. 'Visual Network Analysis', working paper, 2014b, available at http://www.tommasoventurini.it/wp/wp-content/uploads/2014/08/Venturini-Jacomy_Visual-Network-Analysis_WorkingPaper.pdf

Venturini, Tommaso. 'Building on faults: how to represent controversies with digital methods.' P*ublic Understanding of Science,* 21(7), 2012, 796-812.

Venturini, Tommaso. 'Diving in magma: How to explore controversies with actor-network theory.' *Public Understanding of Science,* 19(3), 2010, 258-273.