

Stine Lomborg (ed.)

14

Network Analysis: Methodological Challenges

**Skrifter fra Center for Internetforskning
Papers from The Centre for Internet Research**



Network analysis – methodological challenges

Edited by Stine Lomborg

Department of Media, Cognition and Communication

University of Copenhagen, Denmark

E-mail: slomborg@hum.ku.dk

Skrifter fra Center for Internetforskning

Monograph Series from The Centre for Internet Research

Aarhus, Denmark 2012

Published by The Centre for Internet Research, Aarhus, 2012.
Editorial board: Anja Bechmann, Jakob Linaa.

Stine Lomborg (editor): Network analysis - methodological challenges

© The authors, 2012

Printed at SUN-TRYK. Fællestrykkeri for Sundhedsvidenskab og Humaniora,
Aarhus University.

Cover design: Thomas Andreasen

ISBN: 978-87-91163-14-2

ISSN: 1601-5371

ISSN: 1601-538X (electronic version)

The Centre for Internet Research
Department of Aesthetics and Communication
Helsingforsgade 14
DK-8200 Aarhus N
cfi_editors@imv.au.dk
tel.: + 45 8942 9200
fax: + 45 8942 5950
www.cfi.au.dk

Issues from the Monograph Series, The Centre for Internet Research

- | | |
|--|---|
| 01 Mark Poster: Print and Digital Authorship | 08 Henrik Bødker: The Changing Materiality of Music |
| 02 Niels Ole Finnemann: The Internet — A New Communicational Infrastructure | 09 Anja Bechmann Petersen: Mediediffusion |
| 03 Wolfgang Kleinwächter: Global Governance in the Information Age | 10 Rune Dalgaard (ed.): Lev Manovich and the Language of New Media |
| 04 Jesper Tække & Berit Holmqvist: Nyhedsgrupper set som selvorganiserende interaktionssystemer | 11 Jesper Tække: Chat as a technically mediated social system |
| 05 Niels Brügger & Henrik Bødker (eds.): The Internet and Society? Questioning Answers and Answering Questions | 12 Niels Brügger: Website Analysis. Elements of a conceptual architecture |
| 06 Anne Ellerup Nielsen: Rhetorical Features of the Company Website | 13 Rikke Frank Jørgensen: ICT as a Tool for empowerment in Uganda |
| 07 Jakob Linaa Jensen: Den virtuelle politiske offentlighed — et dansk casestudie | 14 Stine Lomborg (ed.): Network analysis — methodological challenges |

Table of contents

Preface: Network analysis – methodological challenges <i>Stine Lomborg</i>	5
We have the data – what do you want to know? <i>Vidar Falkenberg</i>	8
Network analysis in web archives <i>Niels Brügger</i>	22
Hyperlink networks as social metaphors and boundary drawers <i>Anders Koed Madsen</i>	32
Moving into the flow: Using a network perspective to explore complexity in internet contexts <i>Annette N. Markham</i>	45

Preface: Network analysis – methodological challenges

In April 2011 the Centre for Internet Research hosted a seminar entitled 'Network analysis - methodological challenges'. The aim of the seminar was to begin addressing the methodological and analytical opportunities and challenges of studying the internet through the prism and tools of network analysis. Four invited speakers presented very different conceptual reflections on and uses of network analysis as part of ongoing research within the field of internet studies. This anthology presents the outcome of the seminar's rich discussions of some of the core conceptual, methodological and analytical issues raised in the approaches to network analyses of the internet.

Network analysis has a long, interdisciplinary history. Its research programme developed at the juncture of sociological theory of social relationships and mathematical graph theory (Wasserman & Faust 1994), and is marked by a distinct focus on the *relationships* between entities (individuals, groups, nations or - in the context of the internet - websites, elements on given websites [e.g. images and textboxes], users). That is, relations are at the centre of network analysis, and these are conceptualised in terms of 'nodes' (entities) and 'links' (connections and patterns of activity among nodes). For instance, links between nodes are seen as channels for the transfer and flow of resources (information, money, affect etc.).

Network analysis is particularly useful for making visible the structural connections between nodes, whereas it says very little about the semantic level, that is, the qualities and the content that characterise the structural connection between nodes. A key synergy between sociological network analysis and mathematical graph theory is the development of methods for computing, calculating and visualising social networks in graphs.

The internet as an area of study lends itself very well to network analysis. With the internet, the network structures between individuals, organisations

and so forth are made visible in a different and often very concrete way through hyperlinks between actors, websites, web elements and so on. Different types of and techniques for network analysis have been applied and continue to be developed in internet research (for a review, see Park & Thellwall 2003).

The long-winded sociological research tradition of social network analysis has been very productive in terms of describing internet-based social ties in the context of everyday life. This line of study has mainly focused on measuring individual users' social networks, as expressed through the use of a range of contact points, face-to-face and technologically mediated (e.g. Haythornthwaite 2000; Hogan 2009), and on examining how the individual maintains different types of ties (strong/weak) with alters with a view to their function and significance at the individual level (e.g. Ellison et al. 2007; Wellman & Haythornthwaite 2002).

Another main application of network analysis in internet studies is web sphere analysis (Schneider & Foot 2005). Web sphere analysis is the study of a collection of dynamically defined digital resources spanning multiple web sites deemed relevant or related to a central theme or issue. For instance, following a thematic delimitation of the web, Schneider & Foot (2005) have studied the US presidential campaign on the web; and Moe (2011) has studied issue-driven temporary web spheres in Norwegian blogs.

In addition to such well-developed theoretical and analytical frameworks, various types of software tools are being developed to assist researchers in mapping, harvesting and analysing link patterns, activity clusters and so forth on the web. This anthology engages some of the pathways within network analysis in a critical discussion of opportunities and challenges for the collection of network data and for specific types of analysis within the network perspective in internet research.

Stine Lomborg, February 2012

LITERATURE

- Ellison, N. B., Steinfield, C., & Lampe, C. (2007) The benefits of Facebook 'friends': Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4), pp. 1143-1168.
- Haythornthwaite, C. (2000) Online personal networks. Size, composition and media use among distance learners. *New Media & Society*, 2(2), pp. 195-226.
- Hogan, B. J. (2009) *Networking in Everyday Life*. PhD Thesis, University of Toronto.
- Park, H. W. & Thelwall, M. (2003) Hyperlink Analyses of the World Wide Web: A Review. *Journal of Computer-Mediated Communication*, 8(4).
- Schneider, S. M. & Foot, K. (2005) Web Sphere Analysis: An Approach to Studying Online Action. In *Virtual Methods: Issues in Social Research on the Internet* (pp. 157-170), ed. C. Hine. Oxford: Berg Publishers.
- Wasserman, S. & Faust, K. (1994) *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.
- Wellman, B. & Haythornthwaite, C. (eds.) (2002) *The Internet in Everyday Life*. Malden, MA and Oxford, UK: Blackwell Publishing.

We have the data – what do you want to know?

Vidar Falkenberg

Retriever, Norway

E-mail: vidar@retriever.no

Abstract: An entire business has developed in order to satisfy the commercial request for media monitoring and analysis. Every day, enormous amounts of data from publishers, social network sites, bloggers and microbloggers are harvested and archived. The potential for analysing the aggregated data is huge and mostly untouched, not just from a commercial perspective, but also as a source for research on media, society, democracy and digital life in general. This paper identifies some of the steps necessary to ensure the quality and relevance of an archive that can support both commercial and general research needs.

Keywords: network analysis, content analysis, media monitoring, web archiving

NETWORK ANALYSIS: A COMMERCIAL RESPONSE

The purpose of this text is to offer a commercial and practical response to the task of network analysis. Network analysis, as explained by Stine Lomborg in the introduction to this volume, is an approach with a long, interdisciplinary history, where relationships between entities are the key concept. These entities can be individuals or groups of people, websites or elements of websites. As the relationship is the pivotal component, links and patterns of activity comprise

network analysis' area of study. The internet, Lomborg explains, lends itself very well to the study of linkages, because of hyperlinking and the physical and metaphorical networks that constitute the internet. Although the main outcome of network analysis is to make the structural connections between entities visible - little is inferred about the content of their interactions - content analysis is nevertheless strengthened by visualising the relations, and vice versa. In this paper, I discuss commercially collected data as a source for scholarly research through network analysis of social media. I work for Retriever, a Norwegian/Swedish media monitoring company owned by the news agencies TT and NTB. Retriever provides news and social media monitoring, applications for editorial research and media analysis to both the private and the public sector in the Nordic region. Our vision is to supply relevant information, better decision-making and deeper insight, and we believe the industry as well as the research community would benefit from closer cooperation.

The question of how to perform network analysis on social media content raises a number of new questions regarding privacy, anonymity and copyright, technical limitations, commercial potential and, finally, methodical issues like *validity* and *representability*. How do we address this from a commercial perspective? We also want to combine our knowledge of analysing traditional media content with automated analysis of activities that include new types of content and relationships between creators.

A SHORT HISTORY OF MEDIA MONITORING

The business of monitoring media spans at least several decades and first started to gain momentum with the combination of a growing number of media outlets and a stronger need to pay attention to communication about companies, topics, competitors, markets or other areas of interest. The business has grown alongside the advent of communication departments in both private and public sectors, and even marketing departments and customer relations departments are relying on media monitoring. Instead of buying and reading

every newspaper in circulation, it became necessary to have someone do that for you. For a long time this business was performed by people who read through printed newspapers and manually cut out, copied and mailed relevant pieces of paper to each customer. This service meant that every person in a company could be informed about the content of the media - without having to read everything themselves. Naturally, the manual effort involved in identifying keywords made the monitoring part both expensive and inaccurate: expensive because of the amount of work required, inaccurate because humans make mistakes when tasks are tedious.

Digitalisation and online publishing

From the late seventies and the following decades computers were incorporated in newsrooms at more and more levels of the production cycle. A side effect of the digitalisation of the production of texts was that it was now possible to archive, index and retrieve content through computers. Journalists could thus more easily access what had previously been written about a company or a subject, both in their own newspaper and, later on, in other newspapers.

The content was archived, and a similar system could be used to monitor the content on a daily basis. Technologically it was necessary to follow a number of paths that have subsequently proven expensive and ineffective - for instance, scanning and OCR-ing printed paper - but on a commercial level, it became clear that the monitoring business could no longer be separated from the publishing business. Newspapers took control over monitoring companies by giving them exclusive rights to the content to a provider they were interested in. Newspapers could also ensure additional profits due to legislation that made it necessary for monitoring companies to pay for the rights to redistribute content produced by newspapers. Monitoring companies that were not owned by newspapers thus had to claim a higher price from their customers or choose not to pay for the re-use of content. Newspaper associations in several countries have been or are disputing this issue, regarding both printed and online content. A separate challenge has been to adjust copyright agreements between

publishers and journalists in order to allow re-use of content on other platforms than the ones the content was originally published on.

At the same time, as the rise of the internet made publishing on the web easy, the situation became even more complex for traditional news media. Content published on the web was mostly made available for free, but some newspapers opposed the searching and indexing that followed the development of search engines. In Denmark, it is enough to mention the Newsbooster case, infamous for the term *deep linking*. In fact, the problem is deeper than just direct links to individual articles and revolves around questions of copyright and newspapers' ambition to take control of media monitoring as a separate business model. Copyright legislation, the digitalisation of newspaper production, the fact that virtually every printed newspaper also has a presence on the internet and the maturing of web harvesting techniques all contribute to the present situation, where the majority of Scandinavian monitoring companies are owned or controlled by media companies. Technologically, automated monitoring is now more cost effective and of a higher quality than manual monitoring, although there are still some limitations due to inadequate data formats or lack of error handling. In general, machines are better at running through the enormous amount of content produced every day to identify keywords and relevant material, while humans still outperform machines in detailed semantic evaluation of media texts. This is particularly true for smaller languages in the Nordic region, and it is also a genre issue, as short status message texts on for example Twitter and Facebook are filled with humour, irony, slang and abbreviations. Looking for keywords alone is not enough, as they may be used in different contexts and bear different meanings. The variety of expressions makes writing algorithms difficult, because there are so many new combinations of words that may convey a given sentiment. The precision of automated sentiment analysis of a full unit of texts needs a higher level of accuracy in order to be trustworthy, or it is necessary to limit the automation to specific pre-defined phrases, leaving the overall sentiment assessment unanswered. Instead of trying to replicate what humans do better, network

analysis and text mining might be a more productive method for automated analysis of large quantities of data.

A second wave of online content

Some years after the burst of the dotcom bubble, blogs had advanced from being a niche phenomenon to a source of increasing importance and awareness in mainstream media. Blogs were different from newspapers, but in essence not more different than it was logical to include blogs in monitoring alongside online newspapers. With the advent of social network sites and microblogging, the still fairly limited commercial attention towards blogs and user-generated content increased dramatically, and in terms of content production the amount of data reached extreme proportions. The hours of video uploaded to YouTube every minute is still increasing, and the number of Facebook users is growing with new countries and new groups of users within each country. The enormous volumes of wisdom, rants and colloquial small talk on discussion forums and newsgroups are at the same time becoming interesting, as yet another form of user-generated content to which it is necessary to pay attention. Forums and newsgroups are probably underrated in the wake of the massive attention given to Facebook and Twitter. Forums and newsgroups are not as new and exciting, but generally the content is more accessible through web search, prolonging the lifetime of each piece of information and ensuring that ‘how to’ questions and product feedback normally feature on the first page of Google’s search results.

The monitoring business’ playing ground has changed radically. There are no longer exclusive rights to content, and the automated archiving of everything from every publication is, in most cases, not an option, because of rate limited access (explained below) and the sheer volume of data. The logic of monitoring is moving from pulling (every available message from a limited set of publishers) to filtering (relevant information being pushed from any content producer). Furthermore, new competitors that do not belong to and are thus not limited to the traditional publishing industry emerge. For commercial players, a necessary reaction to the new situation is to explore the emerging market of mixing media

monitoring and its content analysis with new forms of monitoring and with network analysis. This way, they may be able to profit on their advantage of having a broad set of sources and present a wide-ranging solution for media monitoring and analysis.

Analysis: what media monitoring has done

Primarily, media monitoring has been dedicated to finding and displaying relevant news content and, later on, with the digitalisation, developing automated quantitative analyses like counting the number of articles on a given subject or company. This is what the systems are built for and what the customers are used to paying for. In recent years, communication departments and marketing departments have used media coverage as a way to measure their performance and, more specifically, as a tool to report to the management the value of PR and marketing in building and maintaining a reputation and attracting new customers. Documenting and reporting features have been extended extensively to accommodate this demand and were a natural next step for media monitoring companies, as most or all media outlets were being covered. With the data at hand, semantic analysis of media content became possible and affordable and has been a field of growth within the business.

The first steps of network analysis and monitoring of other kinds of content from new sources have taken a similar approach. Content is harvested, then indexed and searched and finally analysed manually, but the new possibilities create a demand for both a higher level of abstraction and a deeper semantic analysis of the content, which must be performed with the help of computers.

Analysis: what media monitoring is doing

The move from identifying and displaying content towards a more analytical approach is realised through the mentioned qualitative or semantic analysis for traditional media monitoring. In these reports, texts are analysed at both a semantic and a discursive level, individually and collectively. Each text is read in detail and coded according to predefined rules and content-specific variables.

The analytical process takes place both in coding and in the act of exploring the data. The task is to find relevance and extract vital from irrelevant content, and humans are effective and precise in this context. Experiments have been made with for instance automated sentiment analysis, but for shorter texts and especially for languages other than English algorithms are not mature enough to be trusted. Getting it right 70 percent of the time is often not enough, so the time-consuming labour of reading everything is still necessary. Obviously, using the computer as a tool has raised the level of efficiency when it comes to performing qualitative analysis, and searching, filtering and pivoting are powerful software features used for analytical purposes. The digitalisation of content also means that consuming data and the actual coding process are quicker and more economical as they can be done in one interface. Still, the results are limited to the semantic content of each text and of the aggregated patterns of all texts within the scope of the given analysis.

Analysis: what else can media monitoring do?

The next step is to look beyond the content itself and try to find meaning and relevance in the data behind the data and between the data - network analysis, in other words. This fits neatly with an ambition to look at the bigger picture and to combine the traditional media picture with the activities in the world of user-generated content. It also mirrors the unexploited potential of the wisdom hidden in the aggregated data of media archives. Each customer gets tailored reports that answer certain questions, but there are still unexploited resources yet to be analysed that may open up for new perspectives on a broader scale. The practical issues, quality and the ethics of such efforts all need to be addressed to fulfil this ambition, but from a commercial perspective, the potential outcome justifies the necessary investments.

ANALYSIS OF DATA – IN PRACTICE, METHODS, ETHICS

In this section, some of the more general practical challenges are presented.

There are a lot of minor, site-specific technicalities that may be interesting at the present time and in a more specialised context, but in this paper these hopefully short-term obstacles are less relevant than the more permanent and general challenges. Furthermore, I discuss two directions for the near and the more distant future of researching the data in question. One is the possible use of the data already archived, and the other is the adjustments that can be made to future archiving in order to overcome current shortcomings. In both cases, most of the challenges apply.

Practical challenges

Although the technical limitations companies experience today are based on the empirical trial-and-error of trying to develop a system to monitor sites like Facebook and Twitter, which undergo frequent and radical changes in infrastructure and purpose, it is possible to identify some of the general questions that will continue to challenge the development. The monitoring of printed newspapers and, subsequently, online newspapers has developed during several decades, and because of the strong connection between publisher and gatherer, acquiring the content often includes custom-built delivery formats and dedicated manpower to ensure quality and stability. With content from other sources, the task is quite different. The content has a different structure and style, the frequency of updates is unpredictable and often much more intense, and the number of contributors is ever-growing. In effect, it is a challenge to get relevant data about the authors in order to qualify the impact of the message. Furthermore, the global nature of many social media sites means that language identification or geographical placement is vital to ensure the relevance of the content in each case. A general deficiency, compared with content from newspapers, is the possibility of indexing and searching the text with suitable operators and filters. Using the often sparsely documented APIs - enabling developers to access the content of a system without having to scrape the output of the web page - relies upon methods and formats defined by the system in question. Often, the API is not intended for searching the system, but

as a way to interact with the data for third party developers who want to build applications that mirror or extend the functionality of the system.

Yet another essential challenge is the request limit that most systems impose on their users in order to avoid breakdowns due to unrestricted access. These limits are necessary, especially as systems gain popularity and attract more users and third party developers. If the limit is ignored, the system may blacklist the IP address of the violator temporarily or permanently. The restrictions also prohibit data mining companies or search engines from pulling the data as often as they want, and there may be delays between the time when something is first posted and the point when it is available in the system requesting the data. It could also result in lost data, if the activity exceeds the number of delivered results between each query. A possible solution is to ask for a whitelisting that allows a more liberal number of requests; in this way the relationship between content producers and data miners might come to resemble the relationship between traditional news media and monitoring companies. At the same time it opens up business opportunities for resellers of content and enables content producers to monetise on their products in new ways, while their content is made more accessible. The success of social media services is heavily dependent upon an active community of third party developers extending the reach and use of the original service.

Methodological challenges

All of the practical challenges mentioned above are interesting from a technical point of view and, to some extent, as an indication of how content producers and monitoring companies may have to cooperate. For the purpose of this paper, the challenges related to the quality of the data are interesting, as the main question is to examine whether the collected data can be used for more general research purposes. The challenges are similar to and an extension of the issues pertaining to research on digital text archives. In order to serve as more than just specific data to answer ad hoc queries from single companies, traditional scientific standards apply to ensure research validity:

representability, reproducibility and reliability to name just a few (Deacon 2007). The operationalisation of each concept is explained below. For commercial customers, these issues are normally not imperative for ad hoc reports, but as a measurement for improving the services of media monitoring companies; the standards and methods necessary for enabling sound research are useful guidelines and objectives. Commercially, the outcome of such an effort is justified if the core product is improved to allow a higher market share or increased prices, or if a new product enables entrance to a new market. More generally, the outcome of an investment in methodological soundness is beneficial, insofar as the data can be proven to be reliable enough to serve as material for further research.

The question of validity is an overall precondition for any form of research on a set of data. In brief, validity refers to the extent to which a concept, conclusion or measurement is well-founded and corresponds accurately to the real world (Deacon 2007). In the context of network analysis of structural data from social network sites, the validity has to be founded in accurate harvesting and indexing and, subsequently, in a representation that corresponds to the original systems, both in visual appearance and in structural context. This representation, *the representability*, does not necessarily imply a complete mirroring of all data, but enough data has to be present in order to keep *the nodes* and *the content of the nodes* meaningful. To analyse a user profile on Twitter, for instance, access to both the connections to other profiles as well as individual tweets is required. *Reproducibility*, or to what degree repeated measurements under unchanged conditions show the same results, is dependent upon how well the archiving system is able to recreate a corresponding representation of the part of the world it represents. Ideally, archiving data from for instance Facebook should not be dependent upon where or when the archiving process starts. Data from a Fan Page Wall on Facebook should thus appear identically in different archives (apart from temporal differences), as the access to the data in principle is uniform. This is similar to a single edition of a printed newspaper that is identical no matter whether it is archived

physically in Denmark or in Australia. In practice, though, the dynamics of web material makes it difficult to create identical copies, since the act of archiving implies a reproduction, rather than just pulling a copy out of circulation (Brügger 2005). Furthermore, the ever-changing world of social network sites is challenging the possibilities of repeating the archiving process. With reliability, the issue is to ensure that searching the data produces consistent and replicable results over time. Computerised searches are in general fail-safe and reduce the possibilities of human errors, but changes in either the storage system or the retrieval system may cause inconsistencies between otherwise identical searches. To achieve a reliable and valid data set, it is necessary to pay close attention to the design of the archiving system, changes in the world it is meant to represent and changes to either the back-end or the front-end of the system itself.

Ethical challenges

The ethical challenges are not to be underestimated, although this concern is rarely mentioned among customers or as a decisive factor in the monitoring business' development process. To some extent, the ethics involved, regarding the question of privacy and the legal aspect of copyright and ownership of content on the web, are fairly straightforward from the perspective of media monitoring companies. In short, the restrictions from each content producer or network have to be obeyed to use the APIs and respect the terms of service. However, the ethics of monitoring online content is more problematic as there might be a difference between what is considered public from a technical point of view and what the people posting updates think of as public (boyd 2010).

Should commercial companies take extra caution? Is it necessary to question whether we get more data than people want to give us? The purpose of media monitoring is to reduce the burden of following every media outlet every day and, similarly with social media monitoring, to enable a dialogue between companies and users. The question is still, if we are creating a tool for the benefit of end users, in parallel to how journalists and researchers may use

media archives to uncover misuse of power, or if we are empowering already influential entities like the government, media and commercial companies? As with many tools, the use of archived media content is not restricted to one singular, predefined purpose. A hammer is a great tool when building a house, but is equally suitable for smashing a window or beating someone. Basically, the purpose of a commercial media archive is to support already existing tasks like finding relevant news and to enable analyses that would be impossible or at least highly impractical without this general tool. A system gathering data from social media has the same outlook; the only difference is that it contains publicly available information from or about individuals instead of news items intended for traditional publishing. The questions monitoring companies are facing are whether it is acceptable to entrust commercial companies like Facebook and Google to inform their users about the availability of data entered into the system, and whether it is acceptable to augment this data with, for instance, data from other publically available systems. There are no easy answers to these questions, and public awareness of these issues may alter the choices at any point in time. For instance, the default privacy settings on Facebook are changing constantly and often without users' knowledge. There are also large cultural differences in what is considered private and public information, and global services may not be the best advocates for respecting these limits (Ess 2009). Additionally, the aspirations of companies like Facebook, Google and Apple are to be the hub of people's life online, and they have made great efforts to make it easy to re-use personal information across websites. While it is practical to maintain only one or a limited set of profiles on the net, it is not given that everyone wants to be the same person in real life and online at all times. The possibility of anonymous action is vital for many groups of people. This is not merely a concern for minorities or people who are suppressed, but also to ensure a more diverse public debate. There are many negative effects of being anonymous on the web as well, but removing this possibility entirely is not the solution. The ethical challenge for social media monitoring is deciding where to draw the line between what is publicly available

from one system and which elements can, technologically, be transferred from one system to another (e.g. Andrejevic 2011; Arvidsson 2011; Hearn 2011; Kennedy 2011).

CONCLUSIONS – COMMERCIAL CHALLENGES AND POSSIBILITIES

The biggest challenge of them all is in many ways the commercial issue. Before it is possible for a commercial service to ensure research validity, there has to be a motive for making the necessary effort to build a system capable of serving general research needs. Normally, the development process involves a ‘show me the money’ moment. There has to be a business case which, for the monitoring business, has been that each customer pays for his or her own data. So far, the willingness to pay up front for aggregate data has been absent or at least marginal, and the incentive for monitoring companies to strengthen the potential for comparative analysis has been equally small. Customers pay for monitoring, archiving and accessing the archive, the one-stop solution, the reporting features and the analytical opportunities. The professionalisation of both the monitoring business and the responsible departments using these services is increasing the likelihood of offering a solution more capable of comparative analysis. Without the money, either up front or as a secure future profit, the investments necessary to enable data valid for general research are unlikely to materialise by themselves. Close cooperation between the industry and the academic community is essential to overcome the practical, methodological and ethical challenges mentioned above and, not least, to respond to any subsequent challenges that will arise along the way. From a commercial perspective, such cooperation is beneficial for a number of reasons: it will strengthen the product or service, it will expand existing markets and open new ones, and the relation to the public will improve as the ethics of doing these analyses is made more explicit. For the academic community, it is valuable that more data is made available for research, especially data that

otherwise has limited access, and there is a great encouragement for publicly supported research to be able to affect how ethics is considered in for-profit research. A collective benefit is the mutual sharing of knowledge enabled by such cooperation, turning improved network analysis into a win-win-win situation for companies, researchers and the public.

LITERATURE

- Andrejevic, M. (2011) The work that affective economics does. *Cultural studies*, 25, pp. 4-5.
- Arvidsson, A. (2011) General sentiment: how value and affect converge in the information economy. (April 19, 2011). Retrieved from:
<http://ssrn.com/abstract=1815031>
- boyd, d. (2010) *Making sense of Privacy and Publicity*, SXSW. Austin, Texas, March 13.
- Brügger, N. (2005) *Archiving websites. General considerations and strategies*. Centre for Internet Research, Aarhus University.
- Deacon, D. (2007) Yesterday's paper and today's technology: Digital newspapers archives and 'push button content analysis'. *European Journal of communication*, 22(1), p. 5.
- Ess, C. (2009) *Digital Media Ethics*. Polity.
- Hearn, A. (2010) Structuring feeling: Web 2.0, online ranking and rating, and the digital reputation economy. *Ephemera: theory & policies in organization*, 10(3/4).
- Kennedy, H. (2011) *Net Work: ethics and values in web design*. London: Palgrave MacMillan.

Network analysis in web archives*

Niels Brügger

Aarhus University, Denmark

E-mail: nb@imv.au.dk

Abstract: This article discusses some of the methodological issues which may arise from performing network analyses on archived web material by the use of analytical software. The methodological discussion takes as its point of departure the case of the political network on the web in relation to the Danish parliamentary elections in 2011, 2007 and 2001.

Keywords: web, web archiving, history, network analysis, analytical software

Despite the fact that network analysis has gained currency within internet studies in the last decade, historical network analysis of the web has not played a very important role.

Since the late 90s, many countries have established extensive national web archives, but these archives do not seem to have attracted much attention from the internet research community, especially not with regard to the use of analytical software for network analysis. Therefore, the methodological issues which may arise from performing network analyses on archived web material have not yet been discussed. This article sets out to outline some of the fundamental methodological challenges related to doing historical network analyses, based on material in web archives and on the use of analytical software.

* An extended version of this article was presented at the 8th Conference on Applications of Social Network Analysis, Zürich, Switzerland, 14-16 September, 2011.

The methodological discussion will take its point of departure in a case study: an analysis of the political network on the web which was available to voters in the Danish parliamentary elections in 2011, 2007 and 2001. This study is based on a sample of online political websites and search queries in relation to the Danish parliamentary elections 2011, as well as archived web material from 2007 and 2001 in two web archives: the national Danish web archive Netarkivet [the Net archive] and the American Internet Archive. However, the primary focus will not be on the empirical dimension of the case, but rather on the methodological issues related to studying the case.

DANISH PARLIAMENTARY ELECTIONS 2011, 2007 AND 2001

The analytical design of the case study has been constructed with a view to answering the following question: How did the available Danish web look to a voter who wanted to use it for political interaction in relation to the Danish parliamentary elections in 2011, 2007 and 2001?

The point of departure for the analysis is an imagined voter's possible behaviour, that is, the ways s/he presumably acts when setting out to find nodes for possible political interaction (information, debates, opinions etc.) of relevance to the national parliamentary elections. It is supposed that the imagined voter would: 1) visit political parties' websites, 2) visit individual candidates' websites, 3) visit parties and candidates on other 'web presences', for instance Facebook, Twitter, YouTube etc., 4) search on Google. The network analysis thus attempts to map the network of websites and web pages which this imagined voter is very likely to come across, if s/he follows the links from these four categories of starting points.

Taking these considerations into account, the analytical software should start with the following seed lists of URLs: 1) the websites of all political parties already in Parliament (nine parties in 2011) as well as their alternative web presences on Facebook, Twitter, Flickr, YouTube and blogs; 2) the websites of individual candidates who are members of Parliament when the elections are

called (179 individuals) and their alternative web presences; 3) the first 100 search results of a search on Google on the words 'folketingsvalg 2011' ('parliamentary elections 2011').

The parliamentary elections for 2011 took place on 15 September. The first step of the analysis was a network analysis of the online web, using the software IssueCrawler; analyses were made every month, and at shorter intervals as the elections approached. The second step would be to perform a network analysis of the archived material from the same points in time in Netarkivet and the Internet Archive (this has not been done yet).

Previous parliamentary elections were held in Denmark on 20 November 2001, 8 February 2005 and 13 November 2007. Since the Danish web archive did not undertake event archiving in relation to the elections in 2005, these elections have not been selected. However, as part of the pilot project which was the forerunner for the founding of Netarkivet, an event archiving of the elections in 2001 was done, which is why these elections may be included. But since a Google search of the past cannot be recreated, it is impossible to include the Google network in relation to the elections of 2001 and 2007.

Based on this material, the aim of the analysis is to study the developments of the network on both the short and the longer term. On the one hand, the study examines the development of the network during the months leading up to elections, and, on the other hand, it identifies the differences and similarities of the networks between 2001, 2007 and 2011. The overall design of the analysis is outlined in figure 1 below.

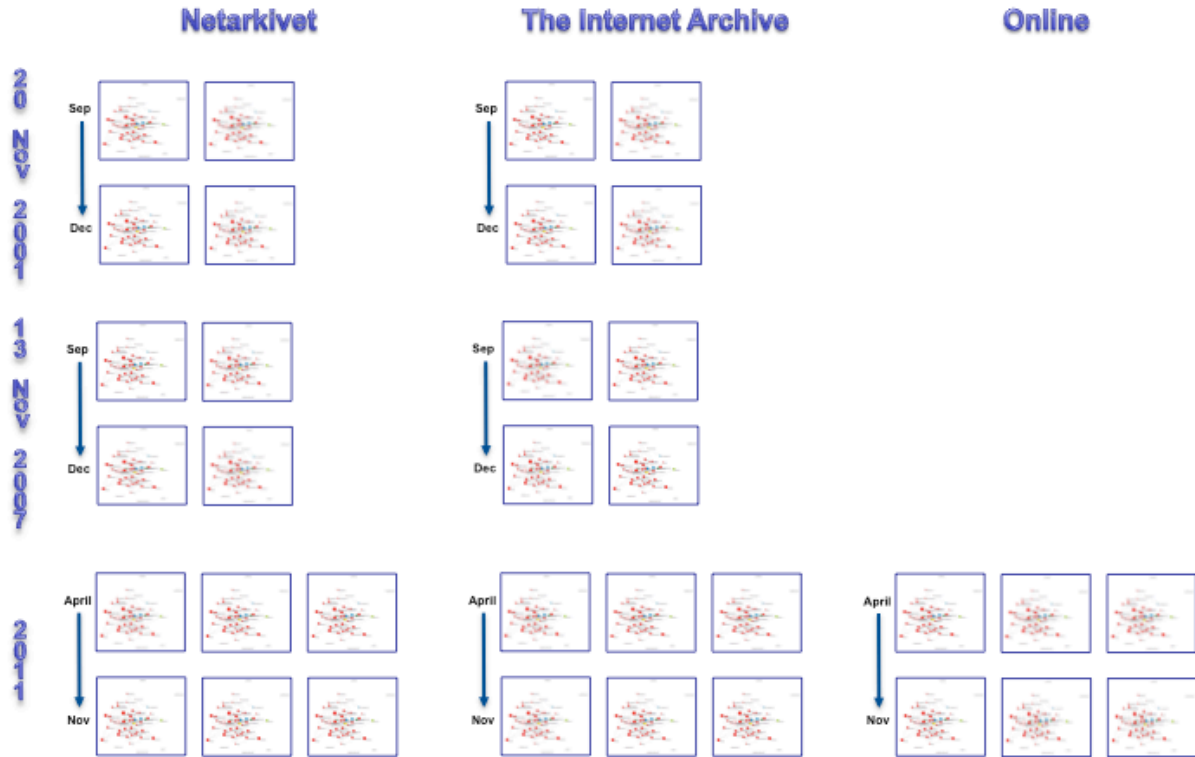


Figure 1: Study design: Danish parliamentary elections 2001-2011.

ARCHIVED WEB MATERIAL: THE CHALLENGES

When trying to make historical network analyses of the web based on material in web archives and by the use of analytical software, the major challenge is that, to a certain degree, we no longer have access to the web as it actually looked online. Instead we have representations of it in the form of archived web material. In this sense the object of study is constructed by the web archive, and it is very hard to determine to what extent it is actually identical to or different from what was once online. For an introduction to web archiving and to the specific nature of the archived website, see Brügger (2005, 2009, 2011); Brown (2005); and Masanès (2006).

In addition to the general problem of not having access to the web as it actually looked online, the more detailed challenges which arise from the use of archived web material can be divided into two main groups: constitutive challenges and practical challenges.

Constitutive challenges related to archived web material

Constitutive challenges are almost always there to some degree and they can be considered an inherent part of using archived web material.

Incomplete and too complete

When performing network analysis of the online web, the web is per definition identical to what is online at a certain point in time. The online web is always complete.

But web archives tend to differ significantly from the online web in terms of completeness. On the one hand, the archive is in most cases incomplete – something may be missing – but, on the other, it may also be too complete – there are simply too many versions of 'the same', some of which may be duplicates. And if this material is used unreflectively for a network analysis, the result of the analysis may turn out to be biased, for instance if a specific website and associated links have simply not been archived, or if three versions of the same website and associated links are present in the archive.

Temporal inconsistencies in the archived network

A network analysis of the online web maps the link structure as it appears at a specific point in time. On the online web there is a high degree of temporal consistency between link structure and network analysis of this structure.

In a web archive this temporal consistency between link structure and network analysis cannot be taken for granted. If all entities of the network – normally hyperlinking websites or web pages – are not from the same point in time (or close to it) the whole link structure – and thereby the whole network – may have changed in the period from when the first entity was archived to the last one was; in many cases, this can be an interval of several days or weeks. Thus, the whole web sphere may very well be temporally inconsistent.

In the case of the parliamentary elections in Denmark, this problem can be seen in the relation between what is archived in Netarkivet and in the Internet Archive. In Netarkivet, the political parties' websites are all archived with the same intervals, but this is not the case with material in the Internet Archive. It

is, therefore, very likely that the networks may look different in these two archives.

As a consequence of this, the network scholar has to choose between two different approaches. Either s/he should use web material which has been archived within a very short interval, with the advantage of minimising the temporal inconsistency, but with the possible disadvantage that the number of URLs in the network has diminished. Or s/he should choose material archived over a longer time interval, thus accessing more URLs for the network, but at the expense of a higher degree of temporal inconsistency.

Spatial inconsistencies in the archived network

When performing network analysis of the online web, the spatial extension of the analysis on each website can usually be specified; that is, how deep the analysis of the website is to be, in terms of how many levels of web pages the web crawler is to include below a website's start page.

In a web archive this spatial consistency is not necessarily present. It is not unusual for different strategies for web archiving to apply different settings in terms of crawl depth. A so-called snapshot strategy, where, for instance, a whole top level domain such as .uk is archived normally, archives very few levels, whereas a selective strategy, where only a limited number of websites is archived, may archive more in depth. And if all entities of the network are not archived in the same depth, the result may be a spatial inconsistency in the network – the ones archived in several levels may contain more hyperlinks – and as a consequence their possibility of appearing more important in the network may increase, and as a whole the analysis of the network may be biased.

In the case of the Danish parliamentary elections, the selective harvest of the political parties' websites in Netarkivet is probably very deep, whereas the Internet Archive usually only archives one to two levels.

Practical challenges of using archived web material

The minute a network scholar enters a web archive, s/he faces the constitutive challenges outlined above. Whereas these constitutive challenges are an inherent dimension of using archived web material, the practical challenges can vary, depending on the concrete construction and content of the web archive in question.

Access to the archive

The question of access is not an issue with regard to the online web, but it may very well be in web archives. Some archives are online and allow free access, others are not online, but have to be consulted onsite, on computers exclusively dedicated to this, and yet others are only open to researchers (some online, some onsite).

For instance, access to Netarkivet is very restricted – it is only open to researchers (online) and only by application. In contrast, the Internet Archive allows open online access.

The analytical software and the archived web material

Most of the analytical software for network analysis is produced with a view to making analyses of the online web. However, one cannot be sure that the software runs on archived web material.

In the case of the parliamentary elections in Denmark, IssueCrawler has been used to map the networks on the online web in the period leading up to the elections, but this software does not run on archived material.

The collections of the archive

The online web is not a collection, but a web archive is – it is based on a number of choices as to what to archive from the online web. As mentioned above, the web archive may be incomplete as well as too complete, but in addition to this general characteristic each web archive may contain a variety of collections, and these collections may have been integrated in the archive from different sources and archived with different types of software and archival strategies. Thus, the collections in the archive are not necessarily

homogeneous.

In the case of the Danish parliamentary elections, the network analysis based on the material in Netarkivet will primarily be based on material from 2007 and 2011, archived by the use of snapshot, selective and event strategies. But, in addition to this, other collections from Netarkivet will be used: 1) web material archived in connection with the elections in 2001 – this material was archived as part of a pilot study, and it was archived very unsystematically and without documentation; 2) old versions of the political parties' websites, which the Danish parliament has archived and which will be handed over to Netarkivet. These different collections will probably vary, and using them as part of the network analysis will probably constitute a challenge in terms of completeness as well as temporal and spatial consistency.

Technicalities: robots.txt

The file robots.txt is a small text file which can be included on a website, and it tells web crawlers, search engines and the like how to behave. It is used especially if the website holder wants to communicate to a web crawler that s/he does not want the website to be copied or indexed and that s/he therefore denies the crawler access.

As stated in the Danish legislation for Netarkivet, the archive does not have to respect robots.txt, whereas the Internet Archive accepts a website holder's robots.txt, stating that s/he does not want the website to be crawled.

This means that if a very important website (or parts of a website) wants the crawler to respect robots.txt, this website may be present in Netarkivet, but missing in the Internet Archive (for instance, the major Danish newspaper Politiken wants crawlers to respect robots.txt). Thus, robots.txt can cause two network analyses, performed on material from two different archives, to be different.

CONCLUDING REMARKS

When an earlier and shorter version of this article was presented at the Centre

for Internet Research's seminar 'Network analysis – methodological challenges', a very good colleague of mine remarked: 'With all these challenges, it is not possible to do historical network analysis of the web; you should do something else'. But although the challenges may seem insurmountable, I think it is worth a try. However, one probably has to accept that the image we make of the network of the available web, as it appears in the web archive, may not be as accurate as it would have been if it had been made on the basis of the online web.

One of the main reasons for this is that it is difficult to get an adequate overview of what actually happens when the analytical software is started and whether the software actually analyses what we want it to analyse, no more, no less. Since the web in the web archive is bigger and less consistent than the online web from which it is created, the main challenge is to delimit the web material on which the network analysis is performed in a precise way, based on conscious choices.

Hence, I would draw the conclusion that it is definitely worth a try to develop historical network analysis of the web, despite the challenges outlined above. Especially within internet studies, two sub-disciplines may benefit from engaging with historical network analysis of the web. On the one hand, an established discipline such as network analysis of the web may benefit from some analytical depth by stretching its perspective backwards in time, thus comparing web networks of today with related networks of the past, and, on the other, an emerging discipline – web historiography – could possibly benefit from broadening its conception of what is understood by 'source' and 'source criticism' by integrating network analysis software in the range of methods used.

LITERATURE

- Brown, A. (2006) *Archiving websites: A practical guide for information management professionals*. London: Facet Publishing.
- Brügger, N. (2005) *Archiving websites: General considerations and strategies*. Aarhus: Centre for Internet Research.
- Brügger, N. (2009) Website history and the website as an object of study. *New media & society*, 11(1&2), pp. 115-132.
- Brügger, N. (2011) Web Archiving – between Past, Present, and Future. In *The handbook of internet studies* (pp. 24-42), eds. M. Consalvo and C. Ess. Oxford: Wiley-Blackwell.
- Masanès, J. (ed.) (2006) *Web Archiving*. Berlin: Springer.

Hyperlink networks as social metaphors and boundary drawers

Anders Koed Madsen

Copenhagen Business School, Denmark

E-mail: akm.ioa@cbs.dk

Abstract: This short paper presents reflections on an on-going study, which includes the production of hyperlink-based visualisations of actors engaged in the controversy about synthetic biology. The paper discusses the performativity of such networked visualisations and argues for thinking about them as 'web visions' that vary over time and across entry points to the web. 'Web visions' are seen as a result of a distributed network of actants, including the researcher, and the concept is put forward as an alternative to thinking about hyperlinked networks as representations of something external to them.

Keywords: web visions, network analysis, Google, synthetic biology

INTRODUCTION: WHEN TECHNOLOGIES PROVIDE METAPHORS AND BOUNDARIES

A central lesson to be learned from the history of ICTs seems to be that, ultimately, they have a bearing on the way we represent, conceptualise and think about the world. Studying the Greek tradition for theatres, Annamaria

Carusi provides an example of this in her explanation of the way this communicative venue provided a frame of thought in ancient Greece that made some questions about the world more obvious than others. Because the theatre staged interplays between representation and reality; it had a make-up that fitted questions about the relationship between ideas and representations and these questions were also the ones that dominated philosophical thought at the time, in the works of Plato, for instance (Carusi 2009).

Similar approaches to the relationship between ICTs and social thought can be found in the writings of Harold Innis (2008) and James Carey (2008). Innis and Carey argued that the essential job of any theory of communication is to enable an understanding of the way technologies provide us with metaphors that shape our thinking, of the way they help us draw boundaries around aspects of the world that we may not have great affiliations with and the way they gradually shape the identities of the people using them.

Innis wrote extensively about the printing press and emphasised the way it gave rise to new translations of scientific works, how it helped promote the vernacular, how it solidified national identities and languages and how the development of print detached the written language more and more from the diversity of oral dialects existing at the time. The printed word became a fixed entity, and Innis argued that the printing press inaugurated the kind of standardisation and abstract reasoning that characterises western thought today (Innis 2008, 19). Carey saw this trend continued in the telegraph which, he argued, fostered impersonal and international business relations mediated by management organisations. He highlighted the way it helped reorganise commodity markets by evening them out in space and opening up time as a region of speculation (Carey 2008, 167-169). Accordingly, the metaphors through which society understood communication and markets changed along with the new technology.

Carusi, Innis and Carey thereby link themselves to a pragmatic theory of communication according to which an utterance is an action rather than a mere representation of the external world (Austin 1962). This also means that ICTs

are more than transmitters of information signals, and Carusi captures this nicely when she speaks about such technologies as 'epistemology engines'. The concept denotes the idea that ICTs have a shaping force on our thinking, because they frame the questions we ask. This performative element of ICTs becomes even stronger when she notes that sometimes we use ICTs, not just as a frame for thinking, but as concrete instruments in the process of carrying out thoughts (Carusi, 2009). We simply use technologies to think with!

HYPERLINK NETWORKS AS 'ENGINES OF THOUGHT'

The works of Innis, Carey and Carusi are relevant for a discussion of hyperlink visualisations in three ways; each way is outlined in this section. First, they are relevant, because they highlight the way in which the web increasingly functions as an 'epistemology engine' that frames our thinking about society and how it does that by providing a new kind of empirical data that fits structural explanations of social phenomena such as social network analysis (e.g. studies of social capital [Burt 2004] and creativity [Uzzi & Spiro 2005]). The hyperlinked infrastructure of the web is especially suiting to research strategies that conceptualise the social as a networked structure, and it is therefore no surprise that relationally-oriented social scientists have turned to the web for data to generate understandings of phenomena such as social movements (Benkler 2006) and social intelligence (Johnson 2001).

Secondly, the works are interesting, because they allow us to focus on the way web crawlers are used as concrete tools through which they are harnessed and synthesised into concrete visualisations of social phenomena. When software agents are used to make structural relations visible in scientific controversies it makes sense to think of them as 'controversy engines'. The hyperlink network in figure 1 was produced by the author as part of a longitudinal study of the controversy about synthetic biology, and it serves as an example of the visual make-up of the kind of networks that are increasingly used to make social phenomena tangible in social science departments as well

Type of parameter	Parameter	Explanation of the parameter	Visualisation
Temporal	New websites in the vision	Looks at whether there are new URLs in a specific vision, compared to the previous month, in order to detect the type of fluidity in the given vision.	<u>Size of nodes</u> Big node = New Small node = Recurring
Structural	The existence of clusters and brokers	Looks at the extent to which URLs are organised in clusters, whether these clusters have specific characteristics and whether there are URLs that serve as brokers between clusters. This is done in order to detect the polarisation of the issues and the actors capable of mediating between polarised parties.	Spring-based graphs
Spatial	The geographical origin of the visible URLs	Looks at the geographical origin of URLs in order to geo-locate the issue.	<u>Shape of nodes</u> Round = US Square = UK Diamond = Global Circle in square = Canadian Triangle = Other Europe Plus = Other world
Spatial	The type of organisation	Looks at each URL in terms of the type of organisation it represents in order to detect the kind of actors dominating the issue and the kind of actors that are connected to each other.	<u>Colour of nodes</u> Red = Policy advice, social science or public engagement Blue = Commercial Green = Natural science Yellow = News and magazine Purple = Funding Dark red = Governmental Black = Other

Table 1: Parameters used to colour and shape the visualisations in UCI net.

Aside from deciding on these visual layers, the final appearance of a hyperlink network is influenced by a decision about the way the nodes and links are to be rendered visible in the two-dimensional space that the network software allows for. The network above is shaped according to a ‘spring embedded’ layout in which an algorithm computes the amount of links one has to follow to get from one website to another. This is the path length between two websites and those with the shortest paths are closest to each other in the graph (Hanneman & Riddle 2005). Alternatives to this could, for instance, be to visualise the network in a circle or as clustered around similarities of attributes.

The coloured nodes, directed arrows and distance-based graphs are common and most of the hyperlink networks produced of biotechnological controversies share them. But the similarities on the visual surface can hide fundamental differences in the way they are constructed and interpreted as

well as the purposes they are intended to serve. Hence, even though people engaged in making hyperlink networks may agree on the overall structural framing of social questions, they may have very different takes on the way they use the concrete tools of analysis (Koed Madsen 2012).

Finally, drawing our attention to the way the web provides us with metaphors of the social and the way web crawlers are used as tools reminds us to look at the way the outcomes of communication technologies are interpreted and translated into society. Carey (2008), for example, highlighted how the information circulated by the telegraph was interpreted as valid enough to build markets on. Even though the use and interpretation of digital traces and web crawlers as 'controversy engines' is still quite flexible, there has been a tendency to interpret hyperlink networks as valid representations of something external to the web. The details of the differences between forms of representation can be found elsewhere (Koed Madsen 2012), but a common goal amongst analysts has been to provide valid insights into communication structures, social movements and organisational identities and use such insights to aid decision-making in relation to managing the technologies in question (see e.g. Ackland et al. 2010; Ackland & O'Neil 2011; Thelwall 2009). Digital visualisations are interpreted as representations of something external to them, and according to Carusi et al. (2010) the alternative is to approach them as modes of seeing. Emphasising the latter approach, I suggest to consider these visualisations as 'web visions'.

WHAT IS A 'WEB VISION' OF SYNTHETIC BIOLOGY?

This section reflects on the central characteristics of 'web vision analysis' by highlighting all the choices that constantly had to be made in the process of producing the visualisation in figure 1. 'Web vision analysis', I have recently proposed, comprises a way to move from the logic of representation to a case study logic that takes as its point of departure the fact that every networked visualisation of social dynamics on the web is inevitably a construct that can

never be found on the web (or in the world) (Koed Madsen 2012). Instead of aiming for this construct to be as representative of social dynamics as possible, the approach of a 'web vision analysis' suggests working deliberately with the performativity of visualisations and taking advantage of the way different information filters on the web demarcate the social. A 'web vision' is accordingly conceptualised as the range of actors, themes and documents that are made visible to the user when entering the web through a specific information filter.

The network depicted in figure 1 is an example of such a vision. It is produced by querying Google.co.uk for 'synthetic biology', de-personalising the search and following hyperlinks from the first 20 search results with software called the Issue Crawler. These 20 webpages are the 'seed sites' of the crawl and the crawler was set to follow their outlinks with a depth of two. Thus, the final visualisation includes all the webpages that the seed sites link to as well as all the webpages these 'secondary webpages' link to. In addition, the crawler was set to remove all websites that did not receive at least two inlinks in order to minimise the visualisation and only make the most central websites visible. After receiving the network from Issue Crawler the network was cleaned by removing all websites that did not mention synthetic biology, and the visualisation was made more readable by using UCI net to colour, shape and size its central elements according to the attributes in table 1.

The purpose in listing these choices is that the final visualisation of the 'web vision', as seen in figure 1, is a construct to which no phenomena on the web can be said to correspond. When looking in more detail at the process of constructing the vision it is evident that several actants are engaged in creating it. These actants are interesting to analyse, because they provide insights into the conditions for producing the visualisation. Arguably, if the researcher is aware of these conditions, they can be used actively to construct interesting performative rather than representative visualisations.

The first actants to take notice of are the people or organisations whose takes on the controversy are exhibited on the websites in the network. These

are persons and organisations that are actually caught up in the controversy about synthetic biology and they have chosen to communicate something about the subject. One of the ways in which they influence the visualisation is by leaving hyperlinks as a way of communicating their views in the public sphere. The meaning of such links have been discussed in internet research for some time and scholars disagree on whether they can be interpreted as recommendations on the part of the webpage forging the link, or whether that interpretation is too ambitious (see e.g. Park 2003; Thewall 2009). Since links can be forged for a broad variety of reasons, the approach of 'web vision analysis' takes a conservative stand and interprets them as associations on the part of the web pages forging the links. The important point is that even though such associations may have different motivations, they play an important role in distributing visibility and thereby in creating a field of vision for the person browsing the web. For instance, in figure 1 we can see that organisations such as the Synthetic Biology Standards Network (synbiostandards.co.uk) and the Synthetic Aesthetics Project (syntheticaesthetics.org) are central actants (i.e. centrally located and densely connected), and the digital traces left by these actors are quite fundamental in generating the final 'web vision'. This direct influence of involved organisations and other social actors makes the depiction of the controversy in figure 1 different from depictions of the same controversy based on methods where people and organisations are prompted to take a stand on an issue by a journalist or a researcher. The direct involvement, however, does not necessarily make them more representative than other depictions.

The visualisation in figure 1 is just as much a consequence of Google's PageRank algorithm, which is the second actant. It works by harnessing the crowd intelligence of web users, and its search rankings are based on parameters such as the number of in-links to a website, the words on the site and the way people move around on the web (Brin & Page 1998; Halavais 2009). Therefore, surfers, and their behaviour, are the third actant contributing to the make-up of the map. What they search for and how they jump from link to link is shaping visibility.

The fourth actant is the Issue Crawler, which by no means can be characterised as neutral visualisation software. The moment the Issue Crawler is put to use, the map starts taking the form seen above. The settings of the crawler naturally shape the map. For instance, above we saw how the setting of the crawler erases websites that are not connected by at least two links from the visualisation.

Finally, we have seen that the researcher has a responsibility for cleaning, colouring, shaping and sizing the map. It has been part of the hype of digital traces that the researcher does not directly influence the data collection, as is the case when conducting surveys and focus groups, for instance. But this does not mean that the researcher is distant and passive. Several choices are made in the process of constructing the network above: the choice to de-personalise the search; the decision to remove websites that do not mention synthetic biology; the decision to clean the data for ‘noise’ such as the ever-recurring links to flash players; and the manipulation of the map in terms of adding colours, shapes and node sizes according to a priori determined parameters. These choices are highly influential on the composition of the network visualisation.

It is important to emphasise that it is *not* possible to avoid making these choices. This is why the ‘web vision’ approach takes the performativity of all these actants into the heart of the analytical strategy instead of thinking about them as biases to be done away with. The implication for empirical studies of hyperlink networks is therefore to select the actants that influence the final visualisation on the basis of case study logic rather than a logic of sampling (Flyvbjerg, 2004). Visualisations are approached as performative cases rather than representations.

A classic case study strategy is to choose paradigmatic cases of the object of study. If the object of study is the web as a public sphere, it may therefore be a useful move to visualise the paradigmatic case of what meets the public web surfer looking for information on synthetic biology at a specific time and place, rather than aiming at making representative visualisations of the discussion of this controversy on the web. The visualisation above is born out of

such a research strategy and it is created by following the links of the sites that Google.co.uk deems most relevant to the controversy. Because this filter is the most heavily used in the UK, it serves as a paradigmatic case of the vision that the UK public is presented with when querying the web. This vision can then be compared to paradigmatic cases of visions in other countries, for instance, by following Google.com while holding parameters such as the settings of the crawler and the personalisation constant. Knowing the roles that the crawler and the personalisation play in performing the visualisation enables the researcher to use them strategically in the research design.

Two other classic strategies of case study research are to look for 'most likely cases' or 'least likely cases'. An example of this logic could be that if the Republicans vote in favour of an increase in income taxes in the USA, one can infer that the Democrats have also supported this increase, because the Republicans are the least likely to support it. In the longitudinal study that I am currently undertaking on 'web visions' of synthetic biology, I utilise this analytical strategy by using the visions generated by following Wikipedia as a 'least likely' case to produce visions with high fluidity and change. Because the filter of relevance in Wikipedia is based on consensus, it is less likely to make new actors visible from month to month than, say, Google Blog Search, which is another case followed. Google Blog Search relies on filtering strategies and media types that makes its 'web visions' quite fluid from month to month, and it is therefore a good example of a 'most likely' case to produce fluid visions. Fluidity in themes and actors in the Wikipedia visions are therefore good indications of a dynamic controversy, whereas stability in the Blog Search visions indicates a dead controversy. These examples illustrates that knowledge of the actants that influence the production of a specific 'web vision' can function as a basis for learning valuable lessons about the world.

'WEB VISIONS' AS HEURISTIC AND NORMATIVE ENGINES

The approach of 'web visions' accepts the fact that all maps are performative

and their potential value must therefore lie in their heuristic value rather than in the extent to which they represent something on the web. This means that their value has to be judged according to the context in which they can be used. The 'web vision' of Google.co.uk discussed in this paper is interesting, because it is paradigmatic of the demarcation of the controversy about synthetic biology that meets the average UK web user.

The performativity of visualisations does not devalue the potential of 'thinking with the web' or using it as a 'controversy engine' and it certainly does not set hyperlink networks apart from other attempts at making controversies tangible and manageable. All of these attempts provide society with social metaphors for what a controversy is and with boundaries of the controversy in question. The difference lies in the metaphors, boundaries and the dialogues and socio-technical set-ups that each attempt creates. The performativity in the process of constructing hyperlink networks must always be explicit, and this is why the approach of 'web visions' proposes to make the choices in this process the heart of the visualising strategy.

LITERATURE

- Ackland, R., Gibson, R., Lusoli, W. & Ward, S. (2010) Engaging With the Public? Assessing the Online Presence and Communication Practices of the Nanotechnology Industry. *Social Science Computer Review*, 28(4), pp. 443-465.
- Ackland, R. & O'Neil, M. (2011) Online collective identity: The case of the environmental movement. *Social Networks*, 33, pp. 177-190.
- Austin, J.L. (1962) *How to do things with words*. London: Oxford University Press.
- Benkler, Y. (2006) *The wealth of networks - How social production transforms markets and freedom*. London: Yale University Press.
- Brin, S. & Page, L. (1998) The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7), pp. 107-117.
- Burt, R. (2004) Structural Holes and Good Ideas. *American Journal of Sociology*, 110(2), pp. 349-399.
- Carey, J.W. (2008) *Communication as culture: Essays on media and society*, revised edition. New York, N.Y.: Routledge.
- Carusi, A. (2009) Philosophy engines: technology and reading/writing/thinking philosophy. *Discourse*, 8(3).
- Carusi, A., Novakovic, G. & Webmoor, T. (2010). Are digital picturings representations? *EVA2010 Conference Proceedings*.
- Flyvbjerg, B. (2004) Five misunderstandings about case-study research. *Sociologisk tidsskrift*, 12(2), pp. 117-142.
- Halavais, A. (2009) *Search Engine Society*. Cambridge, UK: Polity Press.
- Hanneman, R. & Riddle, M. (2005) Introduction to Social Network Methods. University of California Riverside, CA. Retrieved from

<http://faculty.ucr.edu/~hanneman/>

- Innis, H.A. (2008) *The Bias of Communication*, 2nd revised edition. University of Toronto Press.
- Johnson, S. (2001) *Emergence: the connected lives of ants, brains, cities, and software*. New York, N.Y.: Scribner.
- Koed Madsen, A. (2012 forthcoming) Web-visions as controversy-lenses. *Interdisciplinary Science Review*, 37(1).
- Latour, B. (2005) "From realpolitik to dingpolitik or how to make things public". In *Making things public - Atmospheres of democracy* (pp. 14-44), eds. B. Latour & P. Weibel. Cambridge, MA: MIT Press.
- Marres, N. & Rogers, R. (2005) "Recipe for tracing the fate of issues and their publics on the web". In *Making Things Public - Atmospheres of Democracy* (pp. 922-936), eds. B. Latour & P. Weibel. Cambridge, MA: MIT Press.
- Park, H.W. (2003) Hyperlink network analysis: A new method for the study of social structure on the web. *Connections*, 25(1), pp. 49-61.
- Thelwall, M. (2009) *Introduction to webometrics: Quantitative web research for the social sciences*. San Rafael, CA: Morgan & Claypool.
- Uzzi, B. & Spiro, J. (2005) Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, 111(2).

Moving into the flow: Using a network perspective to explore complexity in internet contexts

Annette N. Markham

University of Arizona, USA

E-mail: amarkham@gmail.com

A network perspective, loosened from the bounds of its primary disciplinary trajectories, constitutes a range of techniques and approaches that encourage researchers to move into the flow of culture to find meaning. These need not be tied to a particular theoretical position, but can be used as tools to think with, whether it be through drawing more pictures as a part of the systematic process of analysis or by challenging the very foundations of how we have traditionally conceptualized what we consider the research object.

Keywords: network, qualitative, visualization, ethnography

FROM NETWORK ANALYSIS TO NETWORK SENSIBILITIES

Network analysis offers interesting possibilities to enhance qualitative approaches to the study of contemporary mediated social contexts. When the tools of traditional social network analysis (SNA) are separated from the disciplinary parameters for which they were developed, they offer a beguiling method of extending certain approaches, such as grounded theory or

ethnography, and specifying other approaches, such as actor network theory or practice theory. Specifically, the visualizations emerging from network analysis techniques prompt a range of analytic sensibilities not available through the analysis of text.

In the following pages, I review some of the generative qualities of visual mapping techniques and discuss briefly the reflexive, ethical power of using network analysis to help shift our perspective constantly and radically, by attending to different senses we might use to locate, interpret, and represent that which we might call ‘data’.

Even the earliest network studies focused attention on the idea that *interaction* creates social structures. The premises of network analysis are grounded in general systems theory, whereby structures - whether biological, organizational, or social - are best understood as the result of ongoing and evolving processes of interrelation among various system and environmental elements. The concept of network implies *emergent* (rather than static) structures that shift along with the people whose connections construct these webs of significance. Capturing an image of a network is rather like taking a snapshot of an ever-moving phenomenon, transforming this flow into a somewhat arbitrary object. This aspect of network analysis is not generally highlighted in research reports using network analysis, because the reader/viewer typically only sees the final capture, not the iterative process of creating it from multiple possibilities.

This process-oriented view is revealed in the way that a network researcher might create, view, re-center, move, or animate conceptual or analytical maps throughout the course of a study. The process is much more fluid than that final product might imply. A general network perspective can be characterized as a mode of inquiry that focuses on culture in formation: Humans and/or non-humans interacting and connecting within temporal frameworks to co-construct patterns and structures of meaning, whether ad hoc and temporary or persistent and highly structured over time.

This shift from network *analysis* to network *sensibilities* seems particularly well suited for the study of complex ecologies characterized by convergence, globalization, multiphrenic concepts of identity, and constant shifts in local and global connections. These complex contexts require equally complex tools for thinking. Often this involves disrupting or reconfiguring research methods that are either too rigid or too vague, allowing one to perform, engage, and move with and through data. Indeed, in a literal sense, the goal is to involve more senses in the inquiry process.

In the remainder of this article, I describe how a network sensibility can enhance qualitative research by generating data, adding complexity, and analyze flows versus objects. I conclude with a few remarks about how a network perspective can serve as a catalyst for ethical reflexivity. While I focus primarily on the study of social media or internet-mediated cultural practices, my discussion is not restricted to these environments.

NETWORK SENSIBILITIES AS GENERATIVE TOOL

Most directly, network analysis strategies promote visual mapping of key elements (nodes), connections between them, and the overall structure of the system. This type of visualization can be used in generative ways throughout a study. By generative, I include the processes of generating data, generating organizational strategies for one's data, generating multiple analytic coding schemes or categories, and generating links between levels such as local/global, relational/structural, and so forth. While the focus may be primarily directed toward the phenomenon, it is equally beneficial to use network sensibilities to map one's own conceptual and epistemological standpoints.

Generating data and adding complexity

Adele Clarke (2003, 2005) provides a compelling way to map situations visually, combining elements of grounded theory, actor network theory, and traditional sociological mapping techniques. The key to situational analysis is to use one's

field data to generate more data for analysis. Trying out different *situational*, *relational*, *positional*, and *social arenas* maps of the situation (Clarke 2003, 559-569) can help identify certain general patterns that might not otherwise be noticed. The researcher focuses in serial fashion on particular elements of the situation (a person, an issue, an event, a time period, a theme, a technology or medium, etc.) or notable patterns in larger assemblages. Through this process of analysis, more complicated understandings of the situation emerge. This process cannot help but be deeply iterative with each pass through the data.

This acknowledges the challenge that despite one's goal of identifying a discrete object for inquiry, the object will always be entangled in larger patterns and flows of meaning that operate both at the surface, observable levels and also at less visible, deep structure levels. When adding the premise of swiftly shifting or ad hoc structures, the utility of situational network mapping becomes more meaningful.

This style of mapping adds rather than eliminates complexity, generating additional data for research. This might seem to fly in the face of the goal of narrowing one's research scope to a sensible level, but it highlights a crucial element of qualitative inquiry: seeking depth and complexity in order to reach thick description. Geertz classically described these multiple layers of meaning "winks upon winks." Playing with different possible mappings can help pull this complexity to the surface. This could be accomplished using other methods, but network visualizations serve at least two functions: First, the activity of producing multiple renderings of the context surrounding a phenomenon destabilizes both the context and the phenomenon, an essential step toward shifting to more complex accounts of contemporary culture. Second, multiple layers of visualizations can provide a systematic trace of one's movement through various analytical categories and interpretations.

We most often encounter network maps as the final product of research that focuses on describing large-scale situations (figure 1).

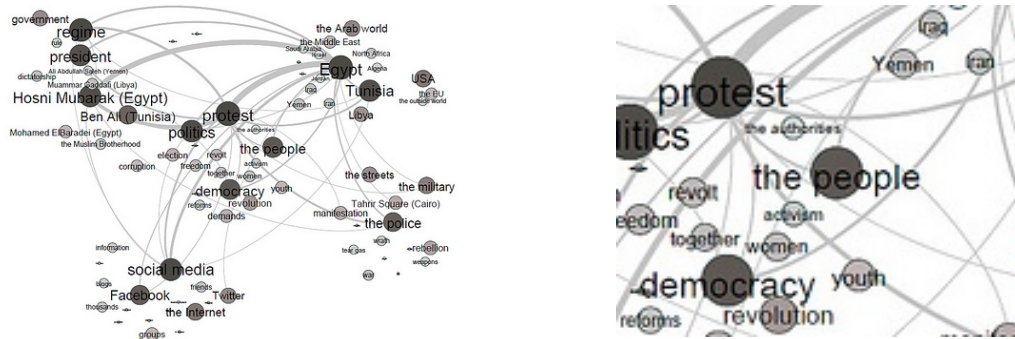


Figure 1: Network map (and close up) overviewing news discourse on the events in North Africa and the Middle East during the first half of 2011. Focus on five thematic clusters in five major Swedish newspapers. Image shows initial rough analysis only. (Lindgren 2011, images used by permission of the author).

Many infographics we see on the internet and in news programs represent the (somewhat) final product of a process of using network analysis to think about, analyze, and represent a phenomenon. These final images do not capture how network analysis works in actuality.¹ We might better identify this by doing our own mapping of maps presented across multiple outlets or changes in mapping over time. Each production will present a different argument about a similar phenomenon. Even a cursory glimpse of the way various stakeholders described or explained the Egypt protests shows the complexity of possibilities.² Daily, if

1 In fact, although some information is offered to explain the meaning of the size of node, thickness of lines, or placement of information, none of these visualizations describe in any detail the methods used to collect, cull, and analyze the data, or the decision process behind the choice and arrangement of particular elements to the exclusion of others. Anders Koed Madsen (this volume) discusses the issue in some detail.

2 Some of this information is no longer available, as cascades of networked information rose and the dissipated over the course of hours or days. For some lingering visualizations and discussions, see for example Boguta's visualizations of Twitter and influence (2012), Paulson's rendering of the protests in a google image search screen grab (2011). For a excellent discussion of various "information cascades," (Bikhchandani, Hirshleifer, & Welch 1992), as well as an interactive visualization, see Lotan, Graeff, Ananny, Gaffney, pearce, and boyd (2011). To get a sense of the difference in focus between information cascades, compare the way timelines are depicted by Al Jazeera (2011), The Guardian (Blight, Pulham, & Torpey

we were paying attention to the situation, we could see a wide range of visualizations, each presenting a partial depiction, as measured by innumerable variables depending on who is doing the mapping and for what purpose. While some focused on speed and diffusion of information, others focused on relative position and power of individuals and/or key stakeholders. Still others traced the geo-located origins of messages and their subsequent travel, mapping the epicenter or apparent impact of tidal waves of information. This list could go on and on. Taken together and over time, these mappings illuminate the power of thinking about situations through a network lens. They also form networks of meaning of their own, not only by virtue of a viewer's experience of them, but also more directly when and if they influence each others' re-renderings over time. We can see that while the focus is ostensibly on an object (social media or protests in the Arab world), the astonishing outcome is that the parts are much more significant and meaningful than the whole, which from an epistemological perspective helps us see that the whole is not just elusive but nonexistent and only ever understood through gross oversimplification or generalization.

Moving beyond the discrete to study process and flow

The study of networks is not just the study of how things are connected. It is a way of rethinking what we identify as the object of analysis. Breaking it down to such a level may seem to oversimplify network analysis, but from a methodological perspective, this actually enables us to build the framework pragmatically from the ground up. Arguably, social media are changing the way we experience the world. What we consider self, structure, and 'the social' are far more temporal and ad hoc than fixed. Whether or not this is the case or anything radically new, social media help us see how our research contexts are not pre-existing milieus but an assemblage of elements "constituted through the connections or articulations among elements" (Balsamo 2011, 15).³ These

2012), or Reuters (2011).

³ This idea cuts across multiple disciplines over the past few decades: That what we consider an

contexts of flow force social researchers to consider the way we have historically conceptualized the object and challenge us to focus on methods for making sense of constantly shifting globalized terrains of meaning.

Despite our acknowledgment that phenomena and research situations ought to be considered more fluidly, qualitative methods are historically designed for use in physically situated, local, fairly static contexts. Traditional analytical tools are object oriented -- focused on those obdurate qualities of the phenomenon that can be identified, parsed, recorded, displayed for other researchers and scrutinized as discrete units of data. It remains easier to focus on the outcomes rather than the processes of interaction.

The sensibilities of network analysis help shift the perspective from object to flow. This can occur at very basic levels of practice, simply by using visualization techniques, such as drawing lines between ideas, which involves senses not inherent in textual analysis. Some of this gets accomplished coincidentally when we include doodles in our notes -- circles around related elements, spaces between elements that illustrate the spatial characteristics of the phenomenon, or connecting lines to denote relationships among people, places, things, or ideas. But because it is not a deliberate practice, it is for the most part not systematized or scrutinized as a method. If we apply McLuhan's argument that the use of any medium extends a particular sense and

organization is a momentary freezing of flux and transformation (Morgan 1986), which we can only identify through retrospective sensemaking (Weick 1969); that space is the crystallization of time (Castells 1996, 411); that the 'individual,' far from being a universal concept (e.g., Strathern 1992), is one that is only understood in terms of relation and interaction (e.g., Blumer 1969); or that both micro and macro elements of 'the social,' such as individuals and institutions, respectively, are nothing but networks (Latour 2005). All of these ideas call for reconsideration of what is conceptualized and captured as the object of analysis. Pushing this further, decades of epistemological discussions challenge us to consider whether "object of analysis" is the best conceptual frame for engaging in what has been called an "anthropology of the contemporary" (Rabinow, Marcus, Faubion, & Rees 2008).

simultaneously amputates other senses to our own analytic tendencies, the value of attending to this traditional mode of research practice becomes clear. Writing by hand, typing on a laptop, drawing, listening; each of these media privilege certain senses (Chandler 1992). The more media we utilize to make sense of a phenomenon, the more we can potentially identify and disrupt our predispositions and limitations in categorizing discursive patterns, people, contextual features, or social structures.

As every context is interwoven with and into incomprehensible and ever changing information networks, it is less and less necessary to identify boundaries and categories in an a priori fashion. Latour and his colleagues (2012) provide multiple instances to demonstrate how it is possible to allow the relevant dynamics to emerge as one surfs these networks: “Instead of trying to simulate and predict the social orders, we wish to acknowledge the limitations of the simulation approach for collective systems and prefer letting the agents produce a dynamics and collect the traces that their actions leave as they unfold so as to produce a rich data set” (Latour et al, 2012, 13; see also Koed Madsen, this volume).

As discourses move more globally, the notion of multi-sited ethnography (Marcus 1998) gives way to multi-sited (and sighted) meaning, an ongoing process of what Rodriguez calls “culturing” (2002). Appadurai’s (1996) notion of ‘scapes’ provides a useful lens for practice. Using the idea of a photographer or painter wanting to capture a ‘landscape,’ we begin to build a sense that depending on innumerable variables, the picture will change. It might be, as we can see in Claude Monet’s Haystacks, the time of year, our situatedness, time of day, or the medium for capturing the moment. This notion is useful as an analytic tool for obvious reasons, in that each iteration causes the researcher to reconceptualize, perhaps radically, the general description of the context and phenomenon, as well as the specific variables influencing the cultural snapshot. Mapping ‘scapes’ helps us envision cultural activity composed of various dimensions of global flow. Each type of scape can focus on a particular type of

information flow, a particular person, a particular moment, and so forth. Almost any word could precede the suffix ‘scape’ to brainstorm types of maps to experiment with. Appadurai (1996) describes ethnoscares, mediascares, financescares, technoscares, and ideoscares, but there’s no need to be limited to these, as they are in this usage simply tools for thinking about multiple sites of meaning through networks and mapping.

NETWORK ANALYSIS AS ETHICAL PRACTICE

When operating within a network perspective, it becomes easier to envision location in relation, or an idea of what it might mean to be “situated.” Playing with networks can help reveal ways of seeing otherwise. The key to maintaining internal consistency and contextual integrity is to constantly rebuild and shift the networks so that different elements can be studied and different nodes centered. Of course, this has the benefit of engendering a more robust analysis, but here, the salient point is that it can help identify the way that one’s analysis is privileging certain standpoints. This becomes crucial when we approach the final stages of the project, when we draw conclusions about what we’ve analyzed and build the argument for particular audiences. During this stage, we’ll emphasize particular connections, eliminating other options. Interrogating one’s own decisions, analyzing conclusions as networks in themselves, provides another level of ethicality.

A network perspective can prompt a methodological and perhaps epistemological approach that better resonates with the study of what is described as cultures of flow. To oversimplify, this involves more visual rendering: Recording and treating as data more of the conceptual and experimental mappings that might otherwise be dismissed as brainstorming ideas, sorting data, or narrowing the scope of the study. It is not just a process that involves mapping, however. It involves a sensitivity to movement and connection, both in the phenomenon and in the researcher’s relationship to this flow. The goal is to embody the perspective of moving with and through the

data, rather than standing outside it as if it can be observed, captured, isolated, and scrutinized outside the flow.

LITERATURE

- Al Jazeera. (2011). Timeline: Egypt's revolution. *Aljezeera*. Retrieved from:
<http://www.aljazeera.com/news/middleeast/2011/01/201112515334871490.html>
- Appadurai, A. (1996). *Modernity at large. Cultural dimensions of globalization*. Minneapolis, MN: University of Minnesota Press.
- Balsamo, A. (2011). *Designing culture: The technological imagination at work*. Durham, NC: Duke University Press.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy*, 100(5), 992-1026.
- Blight, G., Pulham, S., & Torpey, P. (2012). Arab spring: An interactive timeline of Middle East Protests. *The Guardian*. Retrieved from:
<http://www.guardian.co.uk/world/interactive/2011/mar/22/middle-east-protest-interactive-timeline>
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, NJ: Prentice-Hall
- Boguta, K. (2011). Egypt influence network. Retrieved from:
<http://www.kovasboguta.com/1/post/2011/02/first-post.html>
- Castells, M. (2000). *The Rise of the Network Society, The Information Age: Economy, Society and Culture* Vol. I (2nd ed). Cambridge, MA: Blackwell.
- Chandler, D. (1992). The phenemonology of writing by hand. *Intelligent Tutoring Media*, 3(2/3), p. 65-74.
- Clarke, A. (2003). Situational analysis: Grounded theory mapping after the postmodern turn. *Symbolic Interaction*, 26(4), 553-576.

- Clarke, A. (2005). *Situational analysis: Grounded theory after the postmodern turn*. Thousand Oaks, California: Sage.
- Latour, B. (2005). *Reassembling the social: An introduction to actor network theory*. Oxford, UK: Oxford University Press.
- Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (2012). *The whole is always smaller than its parts: A digital test of Gabriel Tarde's monads*. Retrieved from <http://www.bruno-latour.fr/sites/default/files/123-WHOLE-PART-FINAL.pdf>.
- Lindgren, S. (2011). *Beginning analysis of media discourse on the Arab Spring*. [graphic network visualization]. Retrieved from: <http://www.simonlindgren.com/2011/08/12/beginning-analysis-of-media-discourse-on-the-arab-spring/>
- Lotan, G., Erhardt, G., Ananny, M., Gaffney, D., Pearce, I., & boyd, d. (2011). The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions. *International Journal of Communication*, 5, Feature 1375-1405.
- Marcus, G. (1998). *Ethnography through thick and thin*. Princeton, NJ: Princeton University Press.
- Morgan, G. (1986). *Images of Organization*. Thousand Oaks, CA: Sage.
- Paulson, T. (2011). Images of Egypt in revolt. *Humanosphere*. Retrieved from: <http://humanosphere.kplu.org/2011/01/images-of-egypt-in-revolt/>.
- Rabinow, P., Marcus, G., Faubion, J., & Rees, T. (2008). *Designs for an Anthropology of the Contemporary*. Durham, NC: Duke University Press.
- Reuters. (2011). Timeline: The changing U.S. reaction to Egypt's crisis. *Reuters UK*. Retrieved from: <http://uk.reuters.com/article/2011/02/11/us-egypt-usa-reaction-idUKTRE71A64620110211>.
- Rodriguez, A. (2002). Culture To Culturing. Re-imagining Our Understanding Of Intercultural Relations. *Journal of Intercultural Communication*, 5.

Strathern, M. (1992). *After nature: English kinship in the late twentieth century*. Cambridge: Cambridge University Press.

Weick, K. (1969). *The social psychology of organizing*. Boston: Addison-Wesley.

The Centre for Internet Research

The Centre for Internet Research was established in September 2000 with the aim of encouraging research in the social and cultural implications and functions of the internet. More information about the centre and its activities can be obtained from www.cfi.au.dk.

The papers from the Centre for Internet Research

All papers from the Centre for Internet Research can be found on the centres website. As long as in print, copies of the papers may be obtained by contacting cfi@imv.au.dk. Please specify complete address (for the purpose of citation please note that the printed and electronic versions are identical).



Center for Internetforskning

Institut for Informations- og medievidenskab

Helsingforsgade 14 · DK-8200 Århus N

Tel. + 45 8942 9200 · Fax +45 8942 5950

cfi_editors@imv.au.dk · www.cfi.au.dk