Mode 2 in Action: Working Across Sectors to Create a Center for Humanities and Technology

Sally Wyatt

eHumanities Group, Royal Netherlands Academy of Arts and Sciences

Abstract

This article examines recent developments in Amsterdam to establish a Center for Humanities and Technology (CHAT). The project is a collaboration between public research institutions and a private partner. To date, a White Paper has been produced that sets out a shared research agenda addressing both humanities and computing challenges. The article begins with a brief summary of "Mode 2" knowledge production. After providing a fuller description of the White Paper, the extent to which CHAT can be considered an instance of Mode 2 is assessed. Some of the challenges in cross-sector and cross-disciplinary work are presented.

Résumé

Cet article examines les récents développements concernant l'établissement d'un centre pour les sciences humaines et la technologie ("Center for Humanities and Technology - CHAT") à Amsterdam. Il s'agit d'une collaboration en des institutions de recherche publiques et un partenaire privé. Au jour d'aujourd'hui, une feuille de route ("White paper") a été produite afin d'établir un calendrier de recherche s'attaquant aux challenges dans les sciences humaines et l'informatique. Cette article commence avec un aperçu de la production de connaissances dite de "Mode 2." Après avoir fourni une description approfondie de la feuille de route, il continue sur une évaluation de la

Sally Wyatt is the Programme Leader of the eHumanities Group of the Royal Netherlands Academy of Arts and Sciences (KNAW), PO Box 94264, 1090 GG Amsterdam, The Netherlands. Email: sally.wyatt @ehumanities.knaw.nl

Wyatt, Sally. (2015). Mode 2 in Action: Working Across Sectors to Create a Center for Humanities and Technology. *Scholarly and Research Communication, 6*(4): 0401216, 8 pp.

portée avec laquelle CHAT peut être qualifié d'instance de Mode 2. Enfin, certain challenges spécifiques au travail en milieu multi-sectoriels et multi-disciplinaires sont mis en avant.

Mots clés
Collaboration; Digital humanities; Knowledge production; Public-private partnerships

Introduction

In this article, I describe efforts that have been undertaken in Amsterdam in recent years to establish a Center for Humanities and Technology (CHAT). CHAT is a joint effort of the University of Amsterdam (UvA), the Free University of Amsterdam (VU), the Royal Netherlands Academy of Arts & Sciences (KNAW), and International Business Machines Corp. (IBM). In 2014, a White Paper (Wyatt & Millen, 2014) was published, outlining the CHAT research program. It was the result of intensive collaboration between humanities scholars and computer scientists working within both public and private sector research environments. In this article, I use the notion, first put forward by Michael Gibbons, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow (1994), to understand the CHAT initiative: there has been a shift from Mode 1 to Mode 2 knowledge production. The first part of the article outlines what "Mode 2" means. The focus then turns to CHAT, and how it can be understood as an instance of Mode 2. The article concludes with a reflection on some of the difficulties faced when humanities scholars and those working in computing and information sciences collaborate, across disciplines and across public and private organizations.

Mode 2 knowledge production

"Since about 1980, there has been another scientific revolution."[1] This is an idea that has gained traction in both academic and research policy circles. The argument is that science has fundamentally changed; that the change extends beyond scientific practice itself and is intrinsically related to wider sociocultural changes. This is sometimes referred to as a shift from Mode 1 science to Mode 2 (Gibbons et al., 1994; Nowotny, Scott, & Gibbons, 2001). When the first book about it, *The New Production of Knowledge*, appeared in 1994, it attracted much positive attention in science policy circles, but during the first decade of the twenty-first century, it began to receive more criticism, particularly from philosophers and historians of science and technology (Forman, 2007; Mirowski & Sent, 2008; Shinn, 2002).

Mode 2 knowledge has five key elements. First, it is generated within a context of application. This is different from Mode 1, in which pure science is produced in university-type settings and only later applied in industrial/social settings. Digital technologies, genomics, and nanotechnologies are all examples of goal-directed knowledge produced by university/industry/government collaborations. Second, Mode 2 is transdisciplinary, in that a range of theoretical perspectives and practical methods are mobilized to produce new knowledge, whereas in Mode 1 science is advanced within clearly demarcated disciplines. Third, Mode 2 knowledge is produced in a variety of sites, not only in research laboratories, and there is more diversity in the types of knowledge produced. Fourth, research can no longer be characterized as an

"objective" investigation of the natural (or social) world. Fifth, notions of accountability have been radically revised, and as a result, widely agreed upon criteria for determining the quality of science are no longer available.

There are many ways of critiquing this. Logically there is no clear reason why these five elements should operate together. Empirically it could be argued that at the beginning of the twentieth century industrial chemistry, for example, already existed in a context of application, it was transdisciplinary, it was produced in university and industrial settings, and so on. Philosophically one could say that attention to the external validity of science is not new. Social scientists have always been aware of the distinction between internal and external validity, and know that results obtained under artificial conditions such as experiments do not reveal much about how human beings will behave in the wilds of social life, a situation of extreme complexity and uncertainty. From very different positions, both Roy Bhaskar (1979) and Bruno Latour (1987) have clearly demonstrated that external validity is essential to the practice and legitimacy of science.

Terry Shinn (2002), an historian of science, has argued that Mode 2 "legitimates a neo-corporatist vision of the world" (p. 608), and that this is why it has been popular with science and education policymakers as well as university managers. The original authors – Michael Gibbons, Helga Nowotny and others (1994) – are keen to stress, however, that Mode 2 does not challenge the role of universities as independent, autonomous institutions existing for the common good.

Despite its normative and empirical limits, Mode 2 provides an interesting frame for understanding the discussions about the establishment of CHAT, as it draws attention to application, transdisciplinarity or interdisciplinarity, the multiplicity of sites of knowledge production, and the difficulty of defining quality criteria. In the next section, I outline the CHAT process and White Paper, and then assess it in light of Mode 2.

The CHAT White Paper

The promise of digital scholarship for the humanities has been articulated many times over the years. In her "call to action" for humanities scholars, Christine Borgman (2009) argues that the transformation of the field will require new ways to create, manipulate, store, and share the many kinds and huge quantity of research data. Just as important, she adds that new publication practices, research methods, and collaboration among researchers will be required. She argues that we need to go beyond improving access to data and knowledge, through digitization projects, in order to consider what kinds of new knowledge can be created using advanced analytic instruments and techniques.

As part of a Dutch effort to transform humanities scholarship, several workshops were held between 2011 and 2014 in Cambridge, MA and Amsterdam. Participants included researchers from the Royal Netherlands Academy of Arts and Sciences (KNAW), VU University Amsterdam, University of Amsterdam (UvA), the Netherlands eScience Center (NLeSC), and International Business Machines Corp. (IBM). One goal was to reach a common understanding of some of the most important challenges within "digital humanities" (DH).[2] Perhaps the most significant challenge identified was the

Wyatt, Sally. (2015). Mode 2 in Action: Working Across Sectors to Create a Center for Humanities and Technology. *Scholarly and Research Communication, 6*(4): 0401216, 8 pp.

3

need to acquire, represent, and archive humanities data in a way that is easily accessible to a broad range of scholars.

The workshops spurred wide-ranging conversations about specific projects and the instruments and research practices that were used. Much discussion centred on the kinds of humanities research currently underway using state of the art instruments, and what innovations would be possible and desirable in this area. Often the conversation focused on how new forms of humanities scholarship offer new understandings of human behaviour and potential for great societal influence and impact.

The White Paper is structured into three main sections. Section 1 discusses current challenges and opportunities for *humanities* scholarship. During the preparatory meetings for the White Paper, participants were invited to prepare "use cases," examples of research problems where new developments in computer techniques might offer some solutions. These form the background for Section 2, which describes the *core technologies* that are critical to the future of humanities scholarship. While much progress has been made in recent years in areas such as text analytics and cognitive computing, many technological challenges remain. Section 3 lays out some of the *infrastructural* challenges, including the social/collaborative infrastructure, and new forms of training and education for humanities scholars.[3]

OPPORTUNITIES FOR THE HUMANITIES

The development of instruments for, among others uses, text mining, pattern recognition, and visualization (see next section) have potential benefits for the way in which humanities research is conducted and for the questions researchers will be able to ask. The White Paper identifies five important opportunities for the humanities. The first relates to understanding changes in meaning and perspective, over time and across groups. The ways in which humanities scholars understand historical and current objects will change as new sources come to the surface, that may affect the scholar's own theoretical position and value system. Furthermore, current issues and concerns largely inform understanding of the past, and concepts and ideas can also have different meanings across time, space, and social groups.

The second opportunity concerns how uncertainty can be represented in data sources and knowledge claims. Changes in meaning and perspective arise from the availability of sources and reference material. A fundamental part of the training of humanities scholars is to learn to question the provenance and representativeness of available sources (Ockeloen, Fokkens, ter Braake, & Vossen, 2013), and to ask questions about what might be missing, whose voices and opinions are included, and whose might be left out. As the data and sources become increasingly digitized, it is important to develop new ways of understanding and representing the nature of the available data and the claims being made. Again, this is a pressing issue for humanities scholars, but is also of wider relevance, especially as techniques for visualization become ever more sophisticated, and as the available data varies in quality.

A third opportunity concerns the relationship between patterns and categories. Some humanities scholarship is based on identifying and explaining the exceptional person,

object, or event, often as a way of opening up bigger questions. But scholarship is also concerned with the search for patterns, trends, and regularities in data. Identifying such patterns can result in the development of categories for further analysis and use, but these categories may then become too rigid, leading later researchers to miss important new patterns or novel exceptions and outliers (Bowker & Star, 1999). Developing instruments to allow for adding multiple categorizations as new data becomes available is important not only for humanities scholars but for all who deal with big data sets. This is especially challenging for historical data sources, where the data is often incomplete and heterogeneous. The ways in which data can be combined and recombined to make categories are important not only for researchers but also for policymakers who might wish to identify meaningful classifications of, for example, occupation, crime, and disease.

Fourth, scholars sometimes aim to understand more latent and implicit dimensions and meanings of text and data, such as irony, metaphors, and motifs. This fits well with current developments in topic modelling that is language-independent, and that is based on stochastic modelling and information theory (Karsdorp & van den Bosch, 2013). Such developments would have applicability in a range of sectors, including courts, marketing, and anywhere where nuance in meaning is fundamental to interpretation and action.

Finally, huge advances have been made in recent years in the "sentiment mining" of contemporary digital material. To date, this characterizes utterances as positive, negative, or neutral. Yet human emotions are much more complex, and are expressed not only in words but also in gestures, expressions, and movements. In addition, linguistic and body language changes across time, gender, ethnicity, nationality, religion, etc., such that it makes sense to talk of "emotional communities," each with specific styles and practices. Humanities sources, including literature and artistic works, provide a rich resource for developing a fuller and more nuanced set of emotional classifications.

Core technologies

Five important areas of technology were discussed and considered important to enable new research in the humanities. The first concerns cognitive computing, central to IBM's current priorities, and is based around Watson (IBM, 2015). Cognitive computing systems collaborate with humans on human terms, using conversational natural language as well as visual, touch, and other affective interfaces. This partnership between human and machine serves to improve discovery and decision-making by augmenting human abilities with technologies that can reason and can learn from vast amounts of information. A second area is network analytics. Contemporary network theory and technologies could extend the scale and scope of existing work and provide a framework for analysis.

Visualization is the third area. Effective access to large, multimodal data sets requires new visualization instruments for the interactive discovery of meaning across time and integrating multiple modalities. Progress is also needed in the underlying analytics on which these multilayered visualizations are produced. Finally, new approaches are

Wyatt, Sally. (2015). Mode 2 in Action: Working Across Sectors to Create a Center for Humanities and Technology. *Scholarly and Research Communication, 6*(4): 0401216, 8 pp.

5

needed to communicate the stories these visualizations reveal to audiences at all levels of visual literacy.

The fourth area is text and social analytics. Current text analytics enables researchers to perform computation of attributes of text, including determining word and n-gram frequencies, pattern detection, theme identification, information extraction, and association analysis. Progress has been made in linguistic and lexical analysis of text, topic extraction and summarization, and natural language processing (NLP) of meaning and associations within the text. But much remains to be done, as these computation techniques are often fragile and incomplete, and require significant customization for each corpus.

The final challenge concerns search and data representation. Modern information retrieval is attempting to shift from document retrieval to more meaningful units such as answers, entities, events, discussions, and perspectives. Advances in this area will enable humanities scholars in exploration and contextualization tasks.

ORGANIZATIONAL OPPORTUNITIES

Humanities scholars have a long history of engagement with computational technologies (Bod, 2013). Yet adoption of advanced analytical instruments and methods remains limited. Not all sources are available digitally, and there remain barriers facing those scholars who work with material that has not yet been digitized, and for those working with material that was converted early with optical character recognition (OCR). Tools are developed for particular projects, but due to lack of long-term funding, are not maintained and thus quickly become out of date. Similarly, investments in cyberinfrastructures are initiated but not always sustainable in the long run (Bulger, Meyer, de la Flor, Terras, Wyatt, Jirotka, Eccles, & Madsen, 2011; Dombrowski, 2014; Wouters, Beaulieu, Scharnhorst, & Wyatt, 2013). Those involved in CHAT have identified a number of organizational mechanisms aimed to prevent such concerns from becoming major obstacles. These include the following: implementing the lessons of previous experiments in digital humanities; improving awareness of the potential and availability of computational instruments and methods; promoting policies for the preservation of computational instruments and data for future researchers and for the digitization of analogue research material; engaging with a range of potential partners in the cultural heritage sector and creative industries; and contributing to policy debates about the future of humanities and the role of computational technologies.

Conclusion: CHAT as Mode 2?

CHAT has ambitious aims. It wants to enable humanities scholars to both contribute to and take advantage of developments in computational technologies, not only to address questions and challenges in their own research fields and disciplines, but also to pioneer new forms of scholarship that bring together humanities and computational ways of thinking. To do so, the CHAT White Paper recognizes that it is important to keep a dual focus. As described above, CHAT aims to develop new computational instruments, methods, and approaches that can be used across a range of research questions and disciplines. Equally important, however, is to understand how

Wyatt, Sally. (2015). Mode 2 in Action: Working Across Sectors to Create a Center for Humanities and Technology. *Scholarly and Research Communication, 6*(4): 0401216, 8 pp.

researchers can make effective use of such innovations in order to develop new research questions, stimulate cooperation between academic, industry, and other public partners, and meet societal challenges. In many ways, CHAT is clearly an instance of Mode 2 knowledge production, as described earlier. CHAT explicitly aims to generate knowledge in contexts of application, across multiple disciplines and sites. Moreover, a key starting point is the recognition that knowledge is partial and uncertain. All of these elements certainly combine in this instance to make the evaluation of quality very difficult.

The process of achieving CHAT specifically and Mode 2 more generally is not straightforward. To simply invoke Charles Percy Snow's (1959) "two cultures" does not fully capture the complexity of the process. There are indeed differences in epistemic cultures (Knorr Cetina, 1999) or disciplinary matrices (Kuhn, 1970) between computer and information scientists on the one hand and humanities scholars on the other. The computational and algorithm-driven ways of thinking of the former are not easily compatible with the hermeneutic tradition of the latter. But there are also differences between those working in public and private sector organizations around working practices, patterns of thought, expectations about the pace of change and the future, and what constitutes success and quality. Treating reflexivity as a generative resource (Kaltenbrunner, 2015) in the production of knowledge and of knowledge infrastructures could help to meet these future challenges.

Acknowledgements

Notes

1. No precise date has been given, unlike Virginia Woolf's (1924, p. 4) announcement of modernism, that "on or about December 2010 human character changed."

2. There are many terms in circulation: digital humanities, e-humanities, computational humanities, data-driven research, fourth paradigm, big data, etc. The choice often reflects subtle differences in emphasis, which vary between linguistic and disciplinary communities as well as over time. I start from the assumptions that all research and scholarship has already been changed by the widespread availability of digital tools for finding, collecting, processing, analyzing, and representing data of all types, and that the promise of change itself drives both research policy and practice (Wouters et al., 2013).

3. The remainder of this section has been adapted from the White Paper (Wyatt & Millen, 2014).

References

Bhaskar, Roy. (1979). *The possibility of naturalism*. Brighton, UK: Harvester Press.

Wyatt, Sally. (2015). Mode 2 in Action: Working Across Sectors to Create a Center for Humanities and Technology. *Scholarly and Research Communication, 6*(4): 0401216, 8 pp.

Bod, Rens. (2013). *A new history of the humanities.* Oxford, UK: Oxford University Press.

Borgman, Christine. (2009). The digital future is now. *Digital Humanities Quarterly*, *3*(4).

Bowker, Geoffrey, & Star, Susan Leigh. (1999). *Sorting things out. Classification and its consequences.* Cambridge, MA: The MIT Press.

Bulger, Monica, Meyer, Eric, de la Flor, Grace, Terras, Melissa, Wyatt, Sally, Jirotka, Marina, Eccles, Kathryn, & Madsen, Christine (2011). *Reinventing research? Information practices in the humanities.* London, UK: Research Information Network.

Dombrowski, Quinn. (2014). What ever happened to Project Bamboo? *LLC The Journal of Digital Scholarship in the Humanities, 29*(3), 326-339.

Forman, Paul. (2007). The primacy of science in modernity, of technology in postmodernity, and of ideology in the history of technology. *History and Technology*, *23*(1/2), 1-152.

Gibbons, Michael, Limoges, Camille, Nowotny, Helga, Schwartzman, Simon, Scott, Peter, & Trow, Martin. (1994). *The new production of knowledge.* London, UK: Sage.

IBM. (2015). *IBM Watson.* URL: http://www.ibm.com/smarterplanet/us/en/ibmwatson/ [September 27, 2015].

Kaltenbrunner, Wolfgang. (2015). *Reflexive inertia. Reinventing scholarship through digital practices* [PhD thesis]. Leiden, NL: Leiden University.

Karsdorp, Folgert, & van den Bosch, Antal. (2013). Identifying motifs in folktales using topic models. In *Proceedings of the 2nd Annual Belgian-Dutch Conference on Machine Learning* (pp. 41-49).

Knorr Cetina, Karin. (1999). *Epistemic cultures: How the sciences make knowledge.* Cambridge, MA: Harvard University Press.

Kuhn, Thomas. (1970). *The structure of scientific revolutions* (2nd edition). Chicago, IL: University of Chicago Press.

Latour, Bruno. (1987). *Science in action.* Milton Keynes, UK: Open University Press.

Mirowski, Philip, & Sent, Esther-Mirjam. (2008). The commercialization of science and the response of STS. In Edward J. Hackett, Olga Amsterdamska, Michael Lynch, & Judy Wajcman (Eds.), *The handbook of science and technology studies* (3rd edition), (pp. 635-689). Cambridge, MA: The MIT Press.

Nowotny, Helga, Scott, Peter, & Gibbons, Michael. (2001). *Re-thinking science. Knowledge and the public in an age of uncertainty.* Cambridge, UK: Polity Press.

Ockeloen, Niels, Fokkens, Anstke, ter Braake, Serge, & Vossen, Piek. (2013). BiographyNet: Managing provenance at multiple levels and from different perspectives. In E. Blomqvist & T. Groza (Eds.), *Proceedings of the Linked Science workshop at the 12th International Semantic Web Conference and the 1st Australasian Semantic Web Conference* (ISWC2013), 21-25 October, Sydney, Australia.

Shinn, Terry. (2002). The triple helix and the new production of knowledge. *Social Studies of Science*, *32*(4), 599-614.

Snow, Charles Percy. (1959). *The two cultures.* Cambridge, UK: Canto Press.

Woolf, Virginia. (1924). *Mr. Bennett and Mrs. Brown.* London, UK: Hogarth Press.

Wouters, Paul, Beaulieu, Anne, Scharnhorst, Andrea, & Wyatt, Sally (Eds.). (2013). *Virtual knowledge. Experimenting in the humanities and the social sciences.* Cambridge, MA: MIT Press.

Wyatt, Sally, & Millen, David (Eds.). (2014). *Meaning and perspective in the digital humanities. A White Paper for the establishment of a Center for Humanities and Technology (CHAT).* Amsterdam, NL: Royal Netherlands Academy of Arts & Sciences (KNAW).