

The Concept of “Web Science” in the Social Realm: Building Bridges between a new Interdisciplinary Field and the Cultural “Wealth of Networks”

Michael Dick

Ryerson and York Universities
Toronto, Ontario, Canada
m2dick@ryerson.ca

Abstract: Discourse in the field of “cyber culture” largely does not take into account the major shift in constituent technology that has begun to advance the Web from one based solely on human-understandable hypertext documents to one based on machine-understandable data. Such innovation includes the refinement of new search engine technology to mine data in Web services applications (the “Deep Web”) coupled with the desire to annotate data with mark-up languages that facilitate greater interactivity and infer meaning within either user-created knowledge representation models (“folksonomies” as a part of “Web 2.0”) or more rigid ontological structures (part of the “Semantic Web” or “Web 3.0”). In this paper, I consider this overall evident and predicted shift from a “Web of documents” to a “Web of data” to be the central element in the creation of the next-generation of the Web, and the recent drive to study it within an integrated framework known as “Web Science”. Central to this shift is the need to reconsider not only the cultural aspects of the medium, but also the interactions between cultural theory and technical texts. I conclude that with the emergence of certain new technology the entire concept of intellectual property, and more specifically where value ultimately lies in terms of the creation of cultural product, is also changing. Within, I thus focus on alternative frameworks (namely the work of Yochai Benkler) to conceptualize knowledge production, in order to re-examine issues of Web-enabled participatory culture. In order to highlight new cultural paradigms, opportunities and challenges, I discuss how the concept of “social production” may foster a “cultural democracy” that transcends traditional hegemonic conditions that encumber publics.

Keywords: ICTs; public sphere; Semantic Web; social production; Web Science; World Wide Web

Acknowledgement: I wish to acknowledge the Social Sciences and Humanities Research Council of Canada for their financial support of my studies. I am also grateful to Dr. Michael Murphy, my supervisor at Ryerson University, for his enduring support of my work. Finally, my thanks to two anonymous reviewers for their helpful suggestions in the peer review process. This is a modified version of a paper presented at the 10th Annual International and Interdisciplinary Conference of the Association of Internet Researchers (AoIR), which was held October 7-11, 2009 in Milwaukee, Wisconsin, USA.

While working as a research scientist at CERN in the late 1980’s, Tim Berners-Lee conceived of a user-friendly method for representing and sharing scientific data over the Internet. He thus developed the concept of “hypertext” and the HyperText Markup Language (HTML), which served as the *lingua franca* of what he called the “World Wide Web” at the time of its public release in 1991. This invention, for which Berners-Lee did not seek protection under our traditional notion of intellectual property (i.e. a patent), indeed began to facilitate linkages across the disparate nodes present on the wide-area “interconnected network of networks” (i.e. the

Internet), and to this day embodies the same decentralized nature. Although it may seem odd to begin a research paper on the cultural aspects of the Web with such a forthright statement of fact, doing so serves to remind both reader and researcher of two key ideas that are often lost in the annals of “progress” when it comes to the Web: first, that the Web is a very-public realm that rests on the underlying, and often-privatized or privately-exploited, technological protocols of the Internet (e.g. “walled gardens” like “IPTV”, and even one’s home connection as furnished by a corporate ISP); and second, that outright regulation of the Web itself (either through

corporations, government or both) is inherently difficult as a result of, and is even counterintuitive to, this socially-constructed decentralism.

This said, we know the Web is not an altogether “uncontrollable” space: various levels of governments around the world have, rightly or wrongly, applied measures of regulation to the Web and content that is diffused thereupon; the Domain Name System (DNS) is effectively managed by the Internet Corporation for Assigned Names and Numbers (ICANN) and the respective governing bodies of the various top-level domains; and organizations like the World Wide Web Consortium (www.w3.org) set quality-related standards for Web coding and markup languages, as well as accessibility. Indeed, much of this work (especially from the W3C and related entities) has formed the impetus for the new and related discipline of “Web Science” or “Webology” (managed by the “Web Science Trust” under the leadership of Tim Berners-Lee, see www.webscience.org) that has been created to study and understand the myriad implications of the Web. At the same time, various aspects of technological convergence have enabled multiple forms of user-generated content to grow in popularity in the online environment, and this overall concept of the democratization of the Web – known by several names but most importantly for the purposes of this paper as “social production” (see Benkler, 2006) – holds the potential to fundamentally shift how we create cultural product and how we understand intellectual property, particularly in non-Communist market configurations.

Although the Web has now become – in light of the above – the venue of choice for so many forms of communication, commerce and cultural dissemination, I argue that both academic and professional communities have a tendency to either hyperbolize its long-term sociocultural potential, or ignore these considerations outright. Achieving a balance in this area is at the core of this paper and, as will be shown, is the intended, long-term goal of movements to study new iterations of the Web, such as Web Science. However, work done in this area to date often lacks core theoretical paradigms through which dis-

course over the evolving online environment can grow and prosper intellectually. More specifically, and as an example of this point, I posit at the outset that a fundamental flaw exists in the Web Science research agenda, in that the available literature on this topic is disproportionately focused on the technical composition of the underlying coding itself, rather than the various sociocultural issues resulting from our interactions with such coding – the very material that was intended to be at the core of this new discipline.

The primary goal of this paper is thus to identify ways in which cultural economy studies perspectives on the Web can inform the future work of initiatives like Web Science or, perhaps more accurately, to outline new paradigms that can inform the overall cultural economy studies approach to the Web. My underlying assumption is that if this material is contextualized within the most appropriate discourse, new opportunities for both researchers and end-users may result. To this end, I will also argue that Yochai Benkler’s concepts of “social production”, the “networked information economy” and the “networked public sphere” (all derived from his seminal text *The Wealth of Networks*) amply address the challenges and opportunities of the Web and thus warrant further exploration from this perspective (although relevant, I am opting to focus on Benkler here rather than the work of other futurists, such as Lawrence Lessig and Clay Shirkey). This paper, an early theoretical piece in what I am planning to be a long-term and in-depth research agenda regarding Web Science and new Web ICTs, is both a literature review and an exploratory essay focused on extending this emerging discipline towards the enablement of comm.-unitarian ideals regarding the creation of cultural product. My aim is not necessarily to criticize the work that *has* been done in this area – as it flows largely from the initial architects of the Web – rather I am interested in forging new and original ground where possible.

My thesis closely follows the structure of this paper: first, I will outline in the next section further details about the evolution of the Web and Web Science, and the underlying major advancements in Web technology (namely the Semantic and Deep

Webs) that are shifting the importance and value from documents and other forms of cultural product to the underlying data itself; part two will then present key ideals of Benkler's perspectives on the Web and the communitarian social production it enables in the cultural sphere, linking this where practical with the aforementioned discussion on new Web-related ICTs, data and relationships between the data that can be exploited by the end-users; finally, part three will return to Web scholarship, suggesting ways in which academics and practitioners can extract utility from Benkler's work in linking key tenets of Web Science with cultural economy studies and cognate disciplines going forward.

1. Studying and Understanding the Web of the Future

In order to better understand Web Science as a method of encapsulating the next-generation of the Web, and so that I may provide a framework through which I can approach the constituent material within cultural economy studies, it is important that this first section offers both some background information on the discipline and technology. Accordingly, in this section I will discuss: a) documents related to the Web Science Trust; and b) the present shift towards the "Web of Data" (encompassing both the Semantic and Deep Webs).

1.1. A framework for Web Science

The first ever conference on Web Science was held in March 2009. Here, scholars from a wide variety of academic, professional and personal backgrounds convened in Greece to discuss this burgeoning new academic field, developed largely by the recently-formed "Web Science Research Initiative" (WSRI, now the "Web Science Trust"). According to their website, this congress marked the first significant opportunity for scholars to dialogue on the core tenets of the Web Science discipline: namely to better understand what the Web is and how it transforms human relations, and to "engineer its future" in ways that benefit society (Web Science Trust, 2010). Many of the conference's scheduled keynote speakers are indeed heavily-involved

in the Trust's work, and are familiar faces in the areas of computer science and the Web. They include: Sir Tim Berners-Lee, self-professed "inventor" of the World Wide Web, Director of the World Wide Web Consortium and Co-director of the Trust; and the oft-published James Hendler, Nigel Shadbolt, Danny Weitzner and Dame Wendy Hall. News of this conference had been sparse within the academic community. However, it followed from nearly three years of publicity that was first directed outside academia. This is evidenced by headlines like "A Science of the Web Begins" that have graced the news and editorial columns of publications like *Scientific American*. In this one example alone, Web Science is touted as a manufactured "new branch of science", as a method of "analyz[ing] the Web in a scientific way", and as a panacea for studying the "vast emergent properties" arising on the Web in a way that counteracts the "piecemeal and inefficient" efforts of others to do the same (Fischetti, 2006). Since then, numerous other papers have been authored for both academic and professional communities that outline the broad research framework of the Web Science discipline – many of which will be surveyed in this paper. Further still, a peer-reviewed journal has been established exclusively to cover developments in this field (*Foundations and Trends in Web Science*); finally, graduate degrees in Web Science are now offered at the University of Southampton (via a concurrent research partnership with the Oxford Internet Institute) and at Rensselaer Polytechnic Institute.

The seminal document of Web Science remains, at this time, a set of proceedings that emerged from meetings held by what was then the WSRI in 2005 and 2006. Entitled "A Framework for Web Science", the text consists largely of a discussion of engineering issues that are central to the Web's decentralized nature (Berners-Lee et al., 2006). The technical specifics of these issues are beyond the scope of this text, but they flow from the detail that will be presented on the underlying Web-related ICTs. Although subsequent chapters do focus on issues of governance and managing the socioeconomic impact of the Web, these do not progress beyond broad generalizations that are ex-

emplified by developments at the coding level to ensure the security, privacy and trustworthiness of the data itself. For example, the authors outline some eventual goals for the discipline with respect to policy (i.e. the need to avoid regulation by the State to maintain a sense of democracy online), usage of the technology in a way that optimizes social benefit, and outreach activities; this said, an actual action plan for achieving these goals (or at least an agenda for further research) is not included. In other words, there is an assertion overall that forming this new discipline is required in order to advance our understanding of the Web, yet this concept document does not actually address anything on a macroscopic level.

Further adding to the ambiguity surrounding this discipline (and as corollaries to some of the general goals of the movement discussed above) are some of the initial responses to Web Science, particularly in academic literature where I would argue rigor is indeed lacking. As an example, we can consider the Web Science Trust's approach to "awaken Computer Science to the interdisciplinary possibilities of the Web's socially embedded computing technology" by conducting existing Web-related research under the guise of Web Science (Berners-Lee, Hall, Hendler, Shadbolt, & Weitzner, 2006; Shneiderman, 2007; Hendler, Shadbolt, Hall, Berners-Lee, & Weitzner, 2008). As well, "pitches" to other fields outside the purview of Computer Science are convincing to say the least: examples include using Web Science to understand the Web's implications to technology enhanced learning (O'Hara & Hall, 2008), and boldly investing in Web Science research to find applied solutions to problems like identity theft, so as to "protect our future" (Shadbolt & Berners-Lee, 2008).

Again, Web Science literature is very general in its applications to society. Although it must be remembered that scholarship in this field (or under this specific moniker) is relatively new, I see little reason why an organization like the Web Science Trust – which is essentially dedicated to promoting an integrated research agenda on the implications of the Web – has elected not to actually review and integrate literature from peripheral areas of inquiry into their initial

concept document. This, of course, is in spite of the fact that promoting such work ("peripheral" in the sense that it examines implications of things normally under the purview of Computer Science) is a hallmark of what Web Science hopes to eventually represent.

1.2. "Raw Data Now": The Deep and Semantic Webs

Regardless of the broad eventual goals of Web Science, some focus on the nature of Web coding itself is certainly not unwarranted. To a certain extent, Web Science is an extension of the work of the W3C, in that it is an attempt to encapsulate research done on various micro-level aspects of the Web and its technologies, so as to encourage other work to be conducted within its framework. Here, I am attempting to add to this dialogue by offering much needed theoretical perspectives; however, it must be noted that much of this work is based on an underlying shift from a "Web of documents" to a "Web of data" that is present in this research community and indeed within Web Science. The full technical specifics of the components of this "Web 3.0" can be loosely-classified as belonging to either the Deep Web or Semantic Web concepts, though again, a full explanation of each is beyond the scope of this paper.¹ In essence, what is important to understand is that the Web is no longer evolving as a medium based solely on hypertext documents, rather its future will be based on raw data (and metadata) that is annotated with new markup languages that allow for greater "intelligence" to be encoded within. For example, a search for "Jaguar" would no longer be as ambiguous to even the best

¹ For discussions on the technical specifics of the Semantic Web (encompassing the shift to XML and RDF-based markup languages, coupled with ontology models written in the Ontology Web Language or OWL), I would direct the reader to the following sources: Warren & Davies (2008) for an overview; Miller (2004) for details on the development of W3C standards regarding Semantic markup languages like RDF and OWL; and Shadbolt, Hall, & Berners-Lee (2006) for further information on ontologies, RDF Schema and linkages with the provision of Web Services. Also, for use cases of the Semantic Web in business and engineering, I recommend Feigenbaum, Herman, Hongsermeier, Neumann, & Stephens (2007) and Murphy, Dick, & Fischer (2008) respectively.

search engines in future, since the data itself can be encoded in such a way that the contextual meaning of the term (is it a car, an animal or an operating system?) is evident to search engines, Web agents and the like. With this, the Web is thus no longer just syntactic, it is semantic too.

The Semantic Web was formally introduced in Berners-Lee, Hendler, & Lassila (2001), although mention of it as the “ultimate destiny” of the Web can be traced back to Berners-Lee and Fischetti (1999) and their text *Weaving the Web*. Some of the benefits of its eventual widespread deployment include improved business-to-business collaboration and enhanced corporate Web services (such as intelligent online travel agents that actually book – with alleged certainty – based on the travel preferences and parameters of the user). It is very important at this stage to remember that a controlled vocabulary is lacking in this field: depending on which literature one reads, the concept of the Semantic Web may be referred to as “Web 3.0”, the “Data Web” or the “Web of Data” interchangeably. In any event, it can ultimately be conceived as leading-edge Web-related ICT no matter how it is termed.

Actually realizing the vision of the Semantic Web can effectively proceed in two ways, as discussed in Greaves (2007) and Heath (2008): on the one hand, existing and newly-created data can be annotated within an ontological framework using a controlled vocabulary that is common to a specific industry (e.g. all parts suppliers agree on a knowledge representation scheme where “a catalytic converter is a part of the exhaust system”); on the other, individuals can annotate data as they see fit and, in doing so, create “folksonomies” (a portmanteau of “folk taxonomies”). The latter case has already entered the mainstream through numerous “Web 2.0” applications that rely on collective knowledge management, namely the social production that is at the root of a Wiki, a social networking site or an aggregator of user-generated video content. Even with tools that allow a certain degree of automation, creating highly-defined ontologies is indeed a daunting task, most notably because data remains “trapped” in the various Web services databases of organizations, and this material

must be “mined” in order to produce the ontologies required to proceed (Geller, Soon, & Yoo, 2008). Strategies for accomplishing this are centred on “scrubbing” XML databases (within this “Deep Web”) so that popular search engines can properly annotate the found data with Semantic markup languages (Wright, 2008).

Overall, creating the “Deep Semantic Web” (perhaps a more deserving entity of the “Web 3.0” moniker) requires researchers to focus on both creating deep ontologies and new methods of adding intelligence to Web data.² Going forward, I assert the most important point to stress about new Web technologies that are emerging within this framework is the fact that value (in both sociocultural and related terms) is moving towards being inherent within the raw data itself and not necessarily in the final product (be it a document, a video, etc.) Again, the “Web of data” holds that data is of greater importance than the appearance or expression of that data. It is no small surprise then that in a recent talk presenting “Web 3.0” at a TED Conference, Tim Berners-Lee invited attendees to chant “Raw Data Now!” so as to solidify this concept. This interjection is, in my view, an important foundational concept.

2. The Challenge to “Cyber Culture”

Analyzing the opportunities and challenges presented by such new Web ICTs can begin in the area of cultural economy studies, by considering how this desire to study the medium as a whole (inherent in the theoretical framework of the Web Science movement) relates to core literature in the field of medium theory. Using the Toronto School (and more specifically theories of McLuhan and Innis) as

² The potential of a true “Web 3.0” – that is the Semantic and Deep Webs combined – is especially outlined in two articles by James Hendler (2008 & 2009). In the former, the author focuses on how existing platforms like Joost and Twine are leveraging new, efficient search techniques using this technology and how these examples are fostering further growth (i.e. are mitigating the “chicken and egg” conundrum that has, until this point, been common in the development of new Web technologies). In the latter, Hendler examines new query languages (e.g. SPARQL) that are related to seeking information from semantically-annotated databases and web-sites.

an example³, this section will consider: a) contemporary theoretical responses to medium theory that take into account the implications of the Web as a medium (and meta-medium); and b) how this can be related to the study of new Web technology.

2.1. Theoretical Perspectives on the World Wide Web

Much scholarship in recent years has been dedicated to understanding “cyber culture”, and has, to some degree, been focused on extending the theories of McLuhan and Innis to the case of the Web and the online realm. With the former, this relates especially to notions of medium theory, and in the latter to the idea of the “bias of communication” (Burnett & Marshall, 2003). In some circumstances, such discussions regarding the Web as a medium rely on an underlying assumption that the technology directly impacts society. Examples of such determinist thought include: assertions that a “Web society” has emerged predicated on “connected intelligence” and “massive global networking” (De Kerckhove, 1997); and the idea that society has been altered by the Web in terms of knowledge sharing, the creation of a grander sense of community and new collective aspirations as a result (Weinberger, 2002). Additionally, several histories of the Web have been written using McLuhan’s notion that media are “extensions of man” (see for example Rowland, 2006 and Weinberger, 2002). My intent with these theorists is not necessarily to prove or disprove how they conform to determinist or constructivist ends, rather to present their applications to the online environment to understand how areas for studying new Web ICTs, like Web Science, may extract utility from them.

To this end, beginning with perspectives that relate to Innis’s work will prove useful. For

³ I am focusing in this section on scholarly thought related to the Web using theories extrapolated into the “digital age” and online environment and building off the historical development of the medium with respect to commonly-cited determinist versus social construction theories. For a detailed timeline of important events related to the development of the Internet, hypertext and the Web, see Berners-Lee & Fischetti (1999).

instance, Zhao (2007) argues that the Internet (and assumedly the Web) is both space and time-biased: although the Web is a key driver towards globalization and reducing geographic divides in a virtual sense, so too does it enable citizens (and publics of all sorts) to construct new notions of democratic governance and data sharing in the long-term (such as through Wikis and peer-to-peer networks) that are arguably important to deconstructing traditional spatial, temporal and power-related constraints. Further still, by applying a “six dimensions framework” (that represents a taxonomy for conceptualizing all media technology and considers the morphology, scalability, synchronicity, directionality, mode, connectivity and throughout of the medium), Shifman and Blondheim (2007) conclude that the Innisian perspective (with respect to both political economy and thoughts on space-time bias) cannot, in actuality, be effectively applied to the Web because it, in itself, represents not one distinct medium but a “meta-medium” that encompasses traditionally separate media like printed matter, various audiovisual channels, and so forth. Instead, since the Web serves as a delivery vehicle for such media (and since it may indeed deconstruct both space and time parameters) the corollary extending from the arguments above is that attempts to extend determinism and medium theory to the Web are fleeting – we remain in control because the meta-medium can thus reflect such varied “extensions” of its users and, if managed sufficiently, the breaking down of bias can serve society well in mitigating the historically imperialist tendencies of other forms of communication.

2.2. The “Message” of the Meta-Medium

Discussions about the Web and Web Science are informed by medium theory and related discourse because of the interest in studying the varied qualities of this ICT. But as has just been shown, the key value in connecting the Web to such dialogue is also rooted in understanding what the Web is not. This is perhaps most true when we consider one of the key tenets of medium theory, namely McLuhan’s idea that “the medium is the message” coupled with the notion that

media can enable a “global village”, and how it is applied to the online environment (see Levinson, 1999). Indeed, if we look upon the Web as a meta-medium, then is it capable of embodying a message unique to the type of content being transmitted, or does it embody some sort of consolidated message that reflects the sum of its parts? I would argue that the Web transmits, above all else, a unique message – one of a “cultural democracy” versus the democratization of culture. Analogies between the latter and a “global village”, and between the former and a “world bazaar”, can be drawn (Vacker, 2000). Essentially, what this means is that the Web as a decentralized and democratic medium can, when left unregulated and uncensored, provide a cultural delivery vehicle that is more equitably accessed. Granted, steps must be taken to ensure people have access to ICTs in general (i.e. that they have access to the means of production, encoding and reception of the message), but the Web is nonetheless a conduit that can allow for a greater degree of free expression than the traditional, offline media by design.⁴ If near-universal access to content from predominantly developed nations can be achieved through offline ICTs in creating our “global village” of today, then it is the online ICTs like the Web that will allow all “villagers” to more easily contribute to a global marketplace of ideas tomorrow.

2.3. Problematic Theories of the Knowledge Economy

The aforementioned discussion can also be re-framed with respect to the so-called “knowledge economy”, thus allowing us to create linkages between it and cultural economy studies (which I argue is essential seeing that a focus on the “cultural economy” is in vogue). Here, I will draw heavily upon Yochai Benkler’s text *The Wealth of Networks* (2006) to illustrate my points. Overall, what is most interesting about Benkler’s work is his

⁴ This dialogue is at the core of social construction of technology (SCOT) theory, whose proponents are more fully-reviewed in Filchy (2007); additionally, this author puts forth a related ideology known as the “socio-technical frame of reference”, which is also useful in understanding the communications and cultural economy studies perspectives that foreground Benklerian perspectives.

decidedly different take on the characteristics of our present economic era, and this aspect requires further amplification. To be sure, several other well-known political and/or cultural economists have written at length about the role of ICTs (and especially the Internet and the Web) in shaping today’s “knowledge economy”, but I assert this is seldom differentiated from how we understand the implications of offline mass media, and as a result, fallacious (and potentially hegemonic) arguments are extrapolated into the realm of the “cultural economy”. For example, literature about the Web that is written in the tradition of McLuhan (as previously discussed) begets an economic argument like Thomas Freidman’s (2007) in which the technology serves as a global “flattener”, and thus operates under the false view that it is actually creating social and economic opportunity for all who have access to it.

In this vein, the Web is just another enabler (though perhaps a more efficient one) of the “global village”, wherein culture is democratized simply in the sense that it can be diffused to more places and to more people. Of course, this view is problematic because it does nothing to address the dynamics of cultural production itself – “the message” can thus be transported in new ways, but this “one to many” model (resembling the broadcasting system) ignores the fact that major socioeconomic divides persist between the Global North and the Global South. To turn a phrase, “the world is flat” in the sense that those who already control the means of cultural production can maximize resources by exploiting new markets via the use of the Web, while others can only “benefit” by supplying less expensive outsourced labour. However, the “world bazaar” concept – wherein all “global villagers” have the potential to reach their publics and thus contribute to the ongoing dialogue between the Web’s effect on society and vice versa – offers a focal point upon which the cultural and technological aspects of this emerging Web environment can be balanced with alternative viewpoints grounded in cultural economy studies. Indeed, the argument I would fashion from all of this is that this new online paradigm is one of a

“cultural democracy” (meaning opportunities for marginalized or “offside” publics), rather than the aforementioned “democratization of culture”. But this is not simply achieved by understanding how the determinist model of the “knowledge economy” is problematic, it merely puts us in a better position to appreciate the novelty of Benkler’s forthcoming argument and why it is so important to focus on the macro-effects of raw data and collaboration.

2.4. Locating Cultural Value in the Networked Information Economy

The subtitle of Benkler’s work is “how social production transforms markets and freedoms”, and logically it is this notion of “social production” that enables what Benkler terms the “networked information economy” (NIE) and the “networked public sphere” (NPS, or in this case, what I have called more generally the “cultural democracy” – albeit confusing at times, there are many ways to express similar concepts in this discourse). Benkler’s approach to foregrounding this new cultural economic sphere is, in some respects, quite different from the theoretical counterpoint I have provided above. This said, seeing that an entire chapter of his text is dedicated to lamenting “the trouble with mass media” (p. 176), I suggest I have at least captured the essence of his arguments pertaining to the status quo. But I have yet to explain what is at the core of the culturally-democratic ideals of the NIE, at least as compared to the “industrial information economy” (i.e. today’s “knowledge economy”) that is slowly being displaced. Here, Benkler offers an explanation for what, actually, enables the “cultural democracy” to flourish and thus what constitutes the NIE and the phenomenon of “social production”: first, more liberal access to the means of cultural production (i.e. “the declining price of computation, communication, and storage”); and second, the actions of individuals across decentralized networks like the Web in the pursuit of “cooperative and coordinate action” carried out through “nonmarket mechanisms that do not depend on proprietary strategies” (p. 3).

Effectively, this is to say that increasingly-distributed access to production tools, coupled

with the adoption of the public Web as a means to diffuse content and collaborate on the production thereof, constitutes the new information production system par excellence. And here, value is to be found more so in the raw data itself – and in how users work with that data (i.e. how they “socially produce” content) – than in the “finished product” so to speak. This is to say, the process is more important than the product; or, in some cases, the process reflects the product itself. To this end, Benkler presents several mainstream examples of products of the NIE, chief among them Wikipedia, open source software (Linux especially) and SETI@Home, a super-computer created from the leftover cycles of personal computers linked via the Web (p. 5). With each of these, and especially with the first two examples, the importance of the manipulation of data by various groups of people – along with the inherent relationships and peer review process that comes from the collective nature of development, refinement and so forth – is paramount to their respective successes as cultural objects, along with their respective utilities as such (Wikipedia, for instance, derives much of its use value from its “read/write” nature).

Although not a primary focus of his work, I would add examples of social production from the audiovisual realm to Benkler’s exploration, especially the user-generated, streaming content (or “IP Video”) that has become the cornerstone of various social networking applications and video aggregator sites. The fact that some of the material in these forums hardly fits any of the more lofty goals for social production that we might like to set vis à vis democracy and social change (even though this is a subjective measure, I concede that clips of cats riding robotic vacuum cleaners do little to further such goals) is arguably less important than the possibilities that are at least created by such means. Indeed, a recent documentary co-produced by the National Film Board of Canada underscores applications of what I would perceive to be the “Benklerian tradition” to audiovisual cultural product: entitled *RiP: A Remix Manifesto* (2009, Dir. Brett Gaylor), the film focuses on the political economy of cultural “mash-ups”, which are effectively new products created from the sampling of

existing, often copyrighted, material. In addition to exploring the implications of such social production on this domain (especially the music industry and the debates over piracy that were spawned by it), the film itself is an example thereof, since viewers are invited to create their own “remix” of it to share via the Web for possible inclusion in further documentary media (see www.ripmix.com).

3. Opportunities for the Cultural Democracy

As mentioned, the existing framework for new Web research (and especially Web Science as it has been established to date by the Trust) is lacking in foundational sociocultural paradigms that can be leveraged to achieve recognition as a truly interdisciplinary field of work. In order to contextualize the various social implications of the Web, I assert that such work would be wise to examine the potential cultural implications of peripheral research, and that my evaluation of Benkler’s ideals (coupled with the criticism of other perspectives that I have presented so far in this paper) is a logical place to start. As a result, this final section will make preliminary connections between the concepts of social production and the NIE, while offering opportunities for further research both within and outside of the Web Science movement towards the realization of the networked public sphere (NPS) using the new Web technologies under development. The goal here is to highlight how the aforementioned paradigms can be applied in the pursuit of a more democratic Web, and thus the opportunities (and related challenges) that have arisen, are arising and likely will arise in future.

3.1. Social Production and its Relation to the “Web of Data”

The most evident connection between research on new Web technologies, Web Science, and Benklerian perspectives can be found in a mutual focus on data and its importance in the production process. Within research on new Web-related technologies (presently under examination by the Trust), it

was shown in Part One that an ultimate goal is the creation of the Semantic Web or “Web of Data”. The enhanced autonomy that such intelligent agents could provide through semantic annotation could enhance social production, by putting more powerful markup (i.e. Web language) tools in the hands of the general public. Additionally, Web Science may be able to use these developments to inform some of the fundamental challenges Benkler sees in progressing social production from a subversive, “non-market” activity to one that ultimately realigns the core economic system to a true “cultural democracy” – that is the NIE giving way to a broader NPS that positions social production at the forefront of market activity.

In order to realize the NPS, Benkler outlines five specific design characteristics that are effectively required: universal intake; the ability to filter for potential political relevance; the potential to filter for accreditation; the means to synthesize public opinion; and lastly, the requirement to remain independent from government control (pp. 182-185). Many of these are actually simple to address at least in building a preliminary research agenda for future work in the area of Web Science (although perhaps deceptively so, as I shall discuss shortly). For instance, the means to furnish “universal intake” of social production have either already been addressed in this paper (i.e. increasingly inexpensive consumer production equipment and broadband Internet access) or largely fall outside the domain of the Web Science Trust’s general mandate (though this is not to say that issues of Web access and the digital divide are not equally-important areas of concern). Similarly, based on the research and development that has thus far been conducted on next-generation Web technologies, the ontology-driven framework of tomorrow’s online environment is already focused (though once again, not explicitly as a result of Benkler) on addressing requirements to filter for relevance and contextual legitimacy, so as to ensure accuracy and trustworthiness (i.e. accreditation), and to synthesize data.

3.2. Towards the Decentralized, Networked Public Sphere

Benkler's final requirement for the NPS, that is independence from government control, raises a key challenge to Web Science and related technology research going forward and thus deserves special consideration. To be sure, much of what Benkler presents in his text may prove challenging to implement in the long-term, given the inertia of governments (not to mention corporate lobbyists) towards changing the existing system of intellectual property protection (which as we know is generally quite restrictive of certain aspects of social production, namely the "remix" concept); moreover, even when IP is less the issue (such as when a collective of independent producers wishes to produce and subsequently modify their own content) financial considerations still remain. To reiterate, Benkler does not adequately explain in his work how social producers are expected to make a living as cultural workers, let alone cover their production expenses (even as amateurs). Assumedly, he is not suggesting people be altruistic, rather if shifting value in raw data and networks continues to proliferate the mainstream, governments may eventually incorporate such philosophies into public policy regarding IP protection and, as a result, affect how these cultural objects can be funded and financially exploited. Again though, the immediate issue is a sort of paradox, wherein policymakers may not take the NIE as seriously without high-quality, meaningful examples of social production, yet access to resources in the decentralized environment is necessary to finance such production (not to mention the livelihoods of those who devote themselves to production). Although integrating this work into Web Science may be useful to the argument here, it alone cannot address the aforementioned situation. Further research on policy in this area is thus essential.

A further paradox requiring discussion stems from the very configuration of the Web itself as a decentralized (and generally unregulated) technology and meta-medium. As referenced earlier, Tim Berners-Lee and the Web Science Trust regard decentrali-

zation as a requisite feature that must be maintained in further developing Web architecture. Somewhat ironic though is the fact that Web Science is viewed as a means to construct (or maintain) this decentralization, which presumably could then be used as an argument for keeping the Internet (and Web) unregulated. Even going back to *Weaving the Web*, a recounting of his invention of the technology, Berners-Lee and Fischetti (1999) reflect on the importance of standards groups like the W3C in setting quality assurance guidelines for Web development; I would argue that, as a consequence, Web Science is designed to carry on this legacy by suggesting best practices through which Web users, businesses and governments can ensure beneficence in their interactions with the online realm, specifically within the emerging environment outlined here.

Nevertheless, this is not to suggest a return to a determinist perspective on the Web: indeed, recent attempts to apply such thinking to Web Science in characterizing it as necessary to manage the supposedly "lifelike" qualities the Web purports have been largely dismissed by the Web Science Trust as hyperbole (see Tetlow, 2007). On the contrary, I posit that the Web Science perspective can achieve a balance in terms of applying just enough regulation to maintain a decentralized topology. As we know, this is the type of environment that is most conducive to social production, and Benkler too thus avoids a determinist approach (this would be no better than the government control which he finds problematic); however, he does not argue for all-out libertarian or anarchistic paradigms instead (p. 16). For him, the State is not completely discounted; rather it can take a role (perhaps along with Web Science) in facilitating this new political and cultural economy – a role not entirely dissimilar from what was thought to enable the Habermasian notion of the public sphere with other media, at least from his perspective. Nonetheless, the degree to which knowledge representation models exist as top-down controlled vocabularies (i.e. ontologies within the Semantic Web), versus user-driven, "crowd sourced" annotations (i.e. "folksonomies" within "Web 2.0" and most social production to date) will need to be

addressed by researchers as a first step towards achieving this balance.

4. Concluding Remarks

The overall purpose of this paper was to introduce the Web Science discipline and related research on new Web technologies that I argue constitute the next-generation of Web technology, and to situate this work within theoretical discussions of the implications of such emerging ICT to the cultural economy studies perspective. My intention was to review and discuss a range of literature that would ground this new body of work in theory, so as to help provide a firmer foundation upon which new ideas can be generated and subsequently ferment. If this work seems tentative it is because this emerging online environment is so new and, despite Web Science's goal of unifying Web research to further social benefits, is lacking in scholarship regarding its theoretical foundations, methodologies and paradigms to which ways to conceptualize the macro-level changes constructed upon, and caused by, the Web can be ascribed. As it is now too focused on coding, research in this area needs such work to inform the discipline so as to illuminate new paradigms and academic discourse, along with the opportunities and challenges that arise as a consequence of such thought. One major conclusion I have thus drawn from my research is that Web and Web Science scholarship should consider its relation to cultural economy studies not in a determinist sense (i.e. that the medium

determines a message and thus impacts those who receive it), rather that studies of the medium would be wise to assume that the Web may effectively deconstruct both spatial and temporal boundaries, and that it has great potential to stimulate the growth of a "cultural democracy" (again, greater agency for counterpublics) rather than the democratization of culture throughout the "global village".

For this to happen, however, very current issues like the regulation of the Web and ways to stimulate the creation of user-generated, socially-produced content must be taken into account. Accordingly, another aim of this paper was to draw connections (or "build bridges") between Web Science and the cultural economy studies approach, embodied especially by Yochai Benkler's views on the value of data and networks. This is to say, by leveraging a focus on raw data and its malleability (especially within collective situations where it can be viewed as social production) as a common ground, I have argued that the underlying and emerging online environment and related technologies, Web Science, and the work of the Web Science Trust can exist symbiotically with Benkler's perspectives. Moreover, such an infusion between these paradigms can, with future work, lend credence to the Web Science discipline as representing an overarching research agenda for the Web itself, both within the realities of the networked information economy, and within the promise of the networked public sphere.

References

- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven, CT and London: Yale University Press.
- Berners-Lee, T., & Fischetti, M. (1999). *Weaving the Web: the original design and ultimate destiny of the World Wide Web*. San Francisco, CA: HarperBusiness.
- Berners-Lee, T., Hall, W., Hendler, J. A., O'Hara, K., Shadbolt, N., & Weitzner, D. J. (2006). A framework for Web Science. *Foundations and Trends in Web Science*, 1(1), 1-130.
- Berners-Lee, T., Hall, W., Hendler, J., Shadbolt, N., & Weitzner, D. J. (2006, August). Creating a science of the Web. *Science*, 5788, 769-771.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001, May). The Semantic Web. *Scientific American*, 284(5), 34-43.
- Burnett, R., & Marshall, P. D. (2003). *Web theory: An introduction*. London: Routledge.
- De Kerckhove, D. (1997). In W. Rowland (Ed.), *Connected intelligence: the arrival of the Web society*. Toronto, ON: Somerville House.
- Feigenbaum, L., Herman, I., Hongsermeier, T., Neumann, E., & Stephens, S. (2007, December). The Semantic Web in action. *Scientific American*, 297(6), 90-97.

- Filchy, P. (2007). *Understanding technological innovation: A socio-technical approach* (L. Carey-Libbrecht Trans.). Northampton, MA: Edward Elgar.
- Freidman, T. L. (2007). *The world is flat: A brief history of the twenty-first century*. Vancouver, BC: Douglas & McIntyre.
- Geller, J., Soon A.C., & Yoo, J. A. (2008, September). Toward the Semantic Deep Web. *Computer*, 41(9), 95-97.
- Greaves, M. (2007). Semantic Web 2.0. *IEEE Intelligent Systems*, 22(2), 94-96.
- Heath, T. (2008). How will we interact with the web of data? *IEEE Internet Computing*, 12(5), 88-91.
- Hendler, J. (2008, January). Web 3.0: Chicken farms on the Semantic Web. *Computer*, 41(1), 106-108.
- Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T., & Weitzner, D. (2008). Web Science: An interdisciplinary approach to understanding the Web. *Communications of the ACM*, 51(7), 60-69.
- Hendler, J. (2009, January). Web 3.0 emerging. *Computer*, 42(1), 111-113.
- Miller, E. (2004). The W3C's Semantic Web activity: An update. *IEEE Intelligent Systems*, 19(3), 95-97.
- Murphy, M. J., Dick, M., & Fischer, T. (2008). Towards the "Semantic Grid": A state of the art survey of Semantic Web services and their applicability to collaborative design, engineering, and procurement. *Communications of the IIMA*, 8(3), 11-24.
- O'Hara, K., & Hall, W. (2008, May). Web Science. *Association of Learning Technologies Newsletter* (12).
- Rowland, W. (2006). *Spirit of the Web: The age of information from the telegraph to Internet*. Markham, ON: Thomas Allen.
- Shadbolt, N., & Berners-Lee, T. (2008, September). Web Science: Studying the Internet to protect our future. *Scientific American*, 299(4), 76-80.
- Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The Semantic Web revisited. *IEEE Intelligent Systems*, 21(3), 96-101.
- Shifman, L., & Blondheim, M. (2007). From the spider to the Web: Innis' ecological approach to the evolution of communication technologies. In R. Wilson, & M. Blondheim (Eds.), *The Toronto School of communication theory: Interpretations, extensions, applications* (pp. 337-353). Toronto, ON: U of T Press.
- Shneiderman, B. (2007). Web Science: A provocative invitation to Computer Science. *Communications of the ACM*, 50(6), 25-27.
- Tetlow, P. (2007). *The Web's awake: An introduction to the field of Web Science and the concept of Web life*. Hoboken, N.J.: John Wiley.
- Vacker, B. (2000). Global village or world bazaar? In A. B. Albarran, & D. H. Goff (Eds.), *Understanding the Web: social, political, and economic dimensions of the Internet* (pp. 211-237). Ames: Iowa State University Press.
- Warren, P., & Davies, J. (2008). The Semantic Web - from vision to reality. In P. Warren, J. Davies, & D. Brown (Eds.), *ICT futures: Delivering pervasive, real-time and secure services* (pp. 55-66). West Sussex, England: John Wiley & Sons.
- Web Science Trust. (n.d.). *Creating a science of the Web (project website)*. Retrieved April 25, 2010 from <http://www.webscience.org>
- Weinberger, D. (2002). *Small pieces loosely joined: A unified theory of the Web*. New York: Basic Books.
- World Wide Web Consortium. (n.d.). *Project website*. Retrieved April 25, 2010 from <http://www.w3.org>
- Wright, A. (2008). Searching the Deep Web. *Communications of the ACM*, 51(10), 14-15.
- Zhao, X. (2007). Revitalizing time: An Innisian perspective on the Internet. In R. Wilson, & M. Blondheim (Eds.), *The Toronto School of communication theory: Interpretations, extensions, applications* (pp. 199-214). Toronto, ON: U of T Press.

About the Author

Michael Dick

Michael Dick is a student in the Joint Graduate Program in Communication and Culture at Ryerson and York Universities in Toronto, Canada. His research interests include: media history; technological convergence; the political economy of newer media and media industries (especially in Canada); ICT and cultural policy; networks; and cultural nationalism. He has held both an Ontario Graduate Scholarship (2009-10) from the Government of Ontario, and a Joseph-Armand Bombardier Canada Graduate Scholarship from the Social Sciences and Humanities Research Council of Canada (2008-09).