2011

Social justice and an information democracy with free and open source software

C. Sean Burns, *University of Missouri*

Information, Society and Justice, Volume 4 No. 2, December 2011: pp 19-28
ISSN 1756-1078

Social justice and an Information Democracy with Free and Open Source software

C. Sean Burns

Abstract

This paper includes some thoughts on the implications of proprietary software versus free and open source software with regards to social justice, capital, and notions of an information society versus an information democracy. It outlines what free and open source software is and why it is important for social justice, and it offers three cases that highlight two salient themes. This includes a case about preference ordering & decision-making and two cases about knowing and knowledge.

Keywords: social justice, information democracy, Open Source Software

Introduction

In an enlightening comparison between the philanthropic roles played by Andrew Carnegie and Bill Gates for public libraries, Stevenson (2010) asks, "where are today's battles over the ownership and control of the new productive technologies unfolding" (p. 253)? We might respond that many of these battles take place between proprietary and free and open source (FOSS) software models. The battles occur in our homes, our public and academic libraries, our schools, our governments, our places of employment, and elsewhere. Choosing a side means we either promote "capital's requirements for production and consumption" (p. 251) or a distribution that is both communal and equitable.

We can think of these battles being for *information democracy* instead of being for the more neutral *information society*. Doctor (1992) defines information democracy as

> a sociopolitical system in which *all* people are guaranteed the right to benefit from access to information resources. Information democracy deals with empowerment, with ensuring that people have the tools they need to participate in the decision-making structures that affect their daily lives (p. 44).

Faculty of Social Sciences & Humanities, London Metropolitan University,
Ladbroke House, 62-66 Highbury Grove, London N5 2AD, United Kingdom
Website: www.londonmet.ac.uk/isj

Doctor further highlights parts of a stable and healthy information democracy. Among these, the first is the realization that there is a relationship between society and technology and that they interact with each other substantially. The second concerns access---that information democracy means "providing access to informational resources and the means for people to use that access effectively (p. 48). The third part draws on the fact that a democracy depends on "an informed and empowered people" (p. 52). If proprietary models dominate use, then, to appropriate Doctor's words, we will continue to find "an intensified centralization of tools of control, that increase power is being vested in an oligopoly and that there is a growing disparity between the affluent [...] and the economically disadvantaged" (p. 51). This implies that when capital shifts to the few, which by definition occurs with proprietary software models, if capital controls access to information and knowledge, then it further tightens the reigns on information and knowledge.

Pyati (2009) suggests that the open source software model does "provide an opportunity and opening for libraries to re-envision alternatives to the dominance of corporate, capitalistic modes of software development" (p. 218), but he notes that the model "may not be quite the democratizing technology that many of its fervent advocates claim it is" (p. 218). While that is a matter of debate, and the situation might be more subtle, we do not argue that FOSS is a sufficient condition, but we do argue it is a necessary one. We can also go one step further. We can argue that FOSS re-envisions an alternative that is ethical since it displaces a system where knowledge, trust, and validity are purchased and guaranteed based on a system of financial transactions.

Since information and communication technologies (ICTs) are a real, defining part of the information society, we have an obvious concern about the supply of these technologies and the developing infrastructure to support their use across the world. Because of this concern, we call the fuzzy and problematic line separating those with access from those without access the digital divide (Warschauer, 2002). However, the concern is not simply whether we promote the information society by supporting universal access for everyone. This does not sufficiently guarantee that such societies are fair, free, democratic, and just. That is, promoting the 'information society' in no way guarantees or causally determines social justice or an information democracy. While there are a number of necessary conditions before such societies can exist, this paper argues that, with regard to ICTs, an essential ingredient for a socially just information society is the wide adoption and active promotion of free and open source software. The World Summit on the Information Society's "Declaration of Principles, Building the Information Society: A Global Challenge in the New Millennium" (ITU, 2003) lays out a set of principles to promote such communities. However, these principles only mention FOSS once (Principle #27). This paper seeks to explicate FOSS' significance. Particularly, it illustrates how this class of software incorporates the principles needed for such societies and how using FOSS supports an equitable, free, democratic, and socially just community development.

Some context might be helpful. Given the financial capabilities of corporations like Microsoft and Apple, which entail certain marketing advantages, many people may not be aware that there is a choice of software models based on open copyright licensing. For instance, a number of operating systems use open models, such as the manifestations or distributions of GNU/Linux (commonly called Linux but more often referred to by

the name of the distribution: cf., Debian, Ubuntu, Fedora, Red Hat, SUSE, etc.). However, with regards to operating system availability, most people seem only aware of what is found on 'PCs' and Macs. The term PC, regularly used to describe computers running Microsoft Windows, stands for the generic term 'personal computer' or some variant. The term's dominant association with computers running Microsoft Windows is evidence of successful marketing. In a world where the public commonly sees a binary choice (PCs or Macs), a simple categorical syllogism illustrates the hegemony: *All non-Macs are computers that run Microsoft Windows. All PCs are non-Macs. Therefore, all PCs are computers that run Microsoft Windows*. Fortunately, the effect is becoming less common with specific applications. Microsoft's Internet Explorer Web browser is less prominent now that other Web browsers, such as Mozilla's Firefox and Google's Chrome, both open source browsers, have grown in popularity. Unfortunately this is not the case with other software applications. For example, most seem to equate word processing, spreadsheet, and presentation software with Microsoft Office despite the headway achieved by open products such as OpenOffice and LibreOffice.

In light of these issues, in this paper I outline what free and open source software is and why it is important for a socially just information society. While not exhaustive, I offer three hypothetical cases that highlight two salient themes. This includes a case about preference ordering & decision-making and two cases about knowing and knowledge. These cases are generalized and rather coarse, but they are close enough to the kind of mundane circumstances we can easily take for granted.

What is Free and Open Source Software?

Corrado (2005) notes there are three crucial and relevant 'open' models: open access, open source, and open standards. While this paper focuses on the open source model as it applies to software, all three should be recognized as other, necessary, and sometimes interdependent components. Open access' importance lies in its potential to make available scholarly research, and some other genres, to a broader array of the public. This publishing and archiving model liberates literature from subscription-based, digitally protected pay walls. Open standards allow for interoperability across platforms, promote document exchange, insure document archival, enable cross-platform communication, and prevent vendor lock-in (Almeida, Oliveira, and Cruz, 2011). Although, and unfortunately, many proprietary applications exist and are used on the Internet, the Internet as a platform is itself an example of open standards.

I may refer to, as the same thing, open source, free software, free and open source software, or simply FOSS. The naming has its own history and the terms emphasize different philosophical perspectives. The joint phrase 'free and open source' is a later description of what was originally called, and still is, 'free software.' Richard Stallman, founder of The Free Software Foundation, developed the principles of free software in the early 1980s. The term refers not to the cost to obtain the software. While it is often a cost-free product, and many proponents often focus on this, 'free' refers to the user's liberty and freedom. As The Free Software Foundation puts it, "Free software is software that gives you the user the freedom to share, study and modify it. We call this free software because the user is free" (Free Software Foundation, 2011). Thus, free software acknowledges and spotlights the user's integrity and freedom, or the user's moral agency.

By implication, it suggests that proprietary software limits or denies a person's integrity and freedom. Specifically, free software refers to four freedoms, numbered 0 to 3, all of which focus on a person's ability to do some purposeful action. These freedoms are:

- The freedom to run the program, for any purpose (freedom 0).
- The freedom to study how the program works, and change it so it does your computing as you wish (freedom 1). Access to the source code is a precondition for this.
- The freedom to redistribute copies so you can help your neighbor (freedom 2).
- The freedom to distribute copies of your modified versions to others (freedom 3). By doing so you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this. (GNU Operating System, 2011).

The term 'open source,' applied in this context, arrived late into the 1990s. Due to the philosophical and idealistic connotations of the term 'free software,' the Open Source Initiative (OSI) adopted the 'open source' term to appeal to those who might be persuaded by more practical and financial interests in free software (Open Source Initiative, n.d.). Stallman (2009) notes that the open source model focuses more on a software development approach while the free software perspective advocates a social movement. In theory and in practice, they are not mutually exclusive camps: free software advocates and open source developers. This paper uses the juxtaposed and unofficial term, free and open source software, to encompass both meanings.

Three Cases

Case 1: Preference Ordering & Decision Making

Many of our decisions show our preferences. For example, all else being equal, if we choose an orange and not a candy bar, we show a preference (taste, nutrition, craving, etc.), at least at a certain time, for an orange and not a candy bar. With some cases, our preferences contain specific assumptions or our preference-based decisions have certain implications. The assumptions may include our value systems, our beliefs, our attitudes about the world and the implications may involve the effects these decisions have on the world. These effects may have moral consequences and in such cases, we would way we have moral responsibility for them.

Consider then that the active and intentional decision to select a particular software model (or any product) is not simply an agent's decision as a consumer, it is an agent's decision as an ethical being, or what we call her moral agency. We can say this because of the information and communication functions inherent in software. Since information and communication are value laden, any tool that modifies, distributes, and bounds information and communication exists within a moral sphere. Therefore, the ability to choose from a range of these products is not enough. It is also important to have available a set of specific products or models that support moral agency. Combined, these conditions support the following proposition: *an agent should be allowed to choose from a range of x that supports her moral agency and the moral agency of others*. The Free Software

Foundation provides the basis for this proposition when they outline what it means to choose and use free software:

To use free software is to make a political and ethical choice asserting the right to learn, and share what we learn with others. Free software has become the foundation of a learning society where we share our knowledge in a way that others can build upon and enjoy (Free Software Foundation, 2011).

We can illustrate this with a hypothetical case. A young adult acquires her first automobile and wishes to know how it operates. As she prepares to inspect the engine she realizes that special locking bolts prevent her from accessing the engine's compartment. She then learns that the compartment can only be opened with a special wrench distributed to authorized mechanics who pay for this wrench after they have acquired a license to use it. Let's say the purpose of these locking bolts is to prevent automobile owners, or others, from seeing the automobile's machinery and other parts. This is an intentional move on the part of the automobile manufacturer to protect its proprietary designs and engineering.

Such a case has ethical consequences. Most obviously, our agent would not be able to learn how the automobile operates. By implication, she would not be able to infer from that one automobile how other automobiles might operate. Furthermore, her learning would be inhibited because someone thought it would protect its financial interests by way of its intellectual property. Additionally, it supports the systematic justification of institutions, profit or non-profit. If our agent desires to learn how automobiles work, she must matriculate into a program that has received, and possibly paid for, the authority to transfer such training and knowledge, which further necessitates that such institutions exist and are accessible to the person (geographically, technologically, financially). In the end, this system has an implicit set of perpetually reinforcing priorities that we may call an ordering of preferences: the financial profits of one agent rank higher than the potential learning of another agent. It has systemic consequences because it maintains and supports an infrastructure that endorses the ranking. Kipp (2005) describes an analogous situation among farmers challenged by multinationals' intellectual property of seeds. She writes that the "main threat to the cycle of struggles is the use of intellectual property laws to subsume collective knowledge into a few multinational companies" (Discussion section, para. 4).

This preference ordering or ranking signifies a moral issue because it limits our ability to learn and make decisions among all relevant alternatives. The agent may need to know that there are 'open hooded' models, where the agent has the freedom to inspect the automobile's insides, and proprietary models, where the hood is bolted shut and the agent is not free to inspect the insides. She should be able to decide among these alternatives. The categorical syllogism at the beginning of this paper suggests that she is prevented or discouraged by market forces and other agendas to decide among a set of relevant alternatives. In other words, the capital system, as it stands, encourages ignorance.

We hope, though, even after acquiring information about the alternatives, that the agent will choose the free and open model. By choosing this model, she supports her own moral agency. That is to say, if our agent has general knowledge of the available options

and proceeds to act on this knowledge, based on the preference ranking (financial interests > agential learning), she becomes responsible for her actions. If she knows there are both proprietary and 'open hooded' models and makes a choice, because of something about the automobile, either its perceived elegance or its popularity, to buy the proprietary model, then we can infer from this the following proposition: *every instance when an agent chooses a proprietary model, the agent sacrifices her learning (education) to the intellectual property owner's financial interests*. Phrased contextually, the agent chooses to subordinate her own freedom to learn to the manufacturer's financial interests. In instances where she has no desire or cause to learn from the product herself, and still accepts the proprietary model, the perpetually reinforcing nature of the system entails that she subordinates the freedom of persons in her community who do wish to learn.

Cases 2a and 2b: Knowing and Knowledge

The first case highlights an underlying preference ordering or ranking inherent in the proprietary model. We saw how such a model restricts a person's freedom to learn from the world around her. We also saw how if an agent chooses a proprietary model and not a free and open source model, the agent willingly gives up this important freedom. In addition to this, however, proprietary and free and open software models have knowledge implications, which have their own unique moral and practical considerations. To illustrate, let's use a case that is particularly important to researchers.1

A biomedical statistician participates with researchers in evaluating the effects of a certain drug. Let's say that such experiments are expensive to design and implement. A double-blind trial begins with a control group and an experimental group, and the drug and placebo are randomly assigned to the two groups. After the researchers collect their measurements, the biomedical statistician uses a statistical application to analyze the drug's effects. During the analysis, the statistician notices a few outliers. Since the experiment is expensive and because human lives are at risk (due to the potential side effects of the drug or because the disease is especially harmful), the stakes are high. In fact, outliers signify that people died or suffered severe side effects as a result of the drug.

As a statistician, our agent is concerned with statistical and mathematical models, in designing those models, testing them, refining them, understanding them and their applicability to the world, and peer reviewing them. This is the statistician's real work. It is the work of a scholar. Thus, if our agent is thorough, she will want to know the equations and calculations involved in constructing the model. Without a thorough understanding, without an ability to test the methods, our agent and others can never have true confidence in the trial's results. Essentially, we argue that the reliability, the validity, the truthfulness and the soundness of the experiment's methods depend on the statistical models used to explain the results. The agent will need to confirm these models were derived correctly. If the statistician uses a proprietary statistical application, she will be denied access to the exact methods the software application uses and how the software applies these methods. She will be denied transparency and, hence, knowledge.

1 This case is inspired by a presentation delivered by Dr. Frank Harrell Jr., a bio-statistician at Vanderbilt University. The presentation can be seen at http://blip.tv/rtalks/user-2010-vid-1-4013447.

Now consider an implication pertinent to the field of library and information science. Scholars conduct a tremendous amount of research to find optimal measures of relevance. These studies draw from examinations of users, systems, and subject domains. In some way, each contributes to a growing understanding of what it means to 'sort by relevance.' However, the proprietary databases and on-line public access catalogs that so many public and academic libraries license protect the inner workings of their software programs. Consequently, there is no straightforward way to decide whether the relevance algorithms used by these databases serve the best interests of the librarians' communities. Aside from the difficulties with 'relevance' itself (Mizzaro, 1997; Saracevic, 2007), Bade (2007) confirms the secrecy of these algorithms:

"While common determinants of relevance ranking are known, the exact nature of the formulae used remains largely unknown to the public since these are valuable intellectual property for their owners" (p. 831).

In other words, we do not know how these proprietary products 'sort by relevance.' When so much of what we know comes from the documents retrieved by these databases or the books and journal titles retrieved by on-line catalogs, if we do not have a definitive understanding of what 'sort by relevance' entails, the problem warrants mistrust because it invalidates the methods used to access and retrieve information and knowledge.

We may have warrant to criticize general, for-profit search engine companies for their relevance implementations. Not only are so-called *objective* measures at stake, but the implementations affect how users interpret "relevant" results (Pan et al., 2007). Indeed, in the face of profit motivations, we have every reason to remain vigilant in our analysis and critiques of these services. We have to make sure such companies really "do no evil." However, integrated library systems such as those by Innovative Interfaces and Sirsi Dynix and bibliographic databases such as Scopus and Web of Knowledge are provided by for-profit companies also. Scopus is a product of Elsevier, a company that recently earned a 36% profit margin on its journals (Monblot, 2011), and Thompson Reuters owns Web of Knowledge. Yet we criticize Google and ignore the potential conflicts of interests underlying these other corporations. It may be that these products do attempt the best possible relevance systems, but at worst this is merely speculation and at best a matter of indirect empirical inquiry.

The way to be truly sure that an on-line catalog or a database has true potential to meet the needs of the community is to use free and open source systems. These systems provide access to relevance and other retrieval algorithms, which we know are actually non-objective since all algorithms derive from an author (e.g., a computer programmer) (Hjørland, 2010). Sadler (2009) illustrates the constructive nature of an algorithm when she describes the University of Virginia's creation of their open source on-line catalog, Project Blacklight:

> the real power of this project is in the shift of power. Given that the OPAC is the primary method of patrons' interactions with library collections, it makes sense to locate control over its design as close to the assessment of local user needs as possible. This is a significant departure

> from OPAC management as it has been generally practiced in recent memory, which assumes that there can be a single interface that will be good enough for most users, and that this interface must be managed centrally, usually by commercial ILS vendors. By taking local control over the behavior of OPAC interfaces, librarians and patrons can define how objects in the catalog should behave, resulting in greater responsiveness to user needs and increased user satisfaction (p. 58).

Conclusion

In a world dominated by information and communication technologies, an information society can only be a socially just society if its peoples have the freedom to understand, learn, and master the tools that build that society's infrastructure. The Free Software Foundation's four freedoms help promote that society's existence, but we need to adopt and pursue those freedoms aggressively. The restrictive end user licenses agreements we so often click-through and accept, because of de facto and industry standards or due to some other perceived benefit, block those freedoms. While proprietary software may promote an information society, it does nothing to promote a socially just information democracy.

So much depends on information. Since information and communication technologies shape, disseminate, store, archive, retrieve how we acquire it, it is essential that we are able "to look under the hood," so to speak. We cannot allow multinationals or other proprietary-leaning institutions, organizations, and corporations to deny us this access. We can reject these limitations and restrictions on our freedoms by not purchasing or endorsing proprietary products---in our homes, in our libraries, in our governments, in our other workplaces, or elsewhere. While it may true that looking under the hood means examining software code, and while software code may be intimidating, esoteric, and foreign to many of us, neither access nor the potential to read and learn should be denied because of some thing's financial interests. People everywhere should have the freedom to learn to read this code and apply the knowledge gained to further their own communities. Software is a tool that must be continually deconstructed and unveiled. In the end, it is simply text.

References

Almeida, F., Oliveira, J., & Cruz, J. (2011). Open standards and open source: Enabling interoperability. *International Journal of Software Engineering & Applications* 2(1): 1-11.

Bade, D. (2007). Relevance ranking is not relevance ranking or, when the user is not the user, the search results are not search results. *Online Information Review* 31(6): 831-844.

Corrado, E.M. (2005). The importance of Open Access, Open Source, and Open Standards for Libraries. *Issues in Science and Technology* 42(Spring 2005). Available at: http://www.istl.org/05-spring/article2.html.

Doctor, R.D. (1992). Social equity and information technologies: Moving toward information democracy. In: M.E. Williams (ed.) *Annual Review of Information Science and Technology* 27. Medford, NJ: Learned Information, 43-96.

Free Software Foundation. (02 June 2011). What is free software and why is it important for society? Available at: http://www.fsf.org/about/what-is-free-software.

GNU Operating System. (13 July 2011). *The Free Software definition*. Available at: http://www.gnu.org/philosophy/free-sw.html.

Hjørland, B. (2010). The foundations of the concept of relevance. *Journal of the American Society for Information Science and Technology* 61(1): 217-237.

ITU (International Telecommunication Union). (2003). Declaration of principles; Building the information society: A global challenge in the new millennium. *World Summit on the Information Society, Geneva 2003 - Tunis 2005*. Available at: http://www.itu.int/wsis/docs/geneva/official/dop.html.

Kipp, M.E.I. (2005). Software and seeds: Open source methods. *First Monday* 10(9). Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1276/1196.

Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science* 48(9): 810-832.

Monblot, G. (2011, August 29). Academic publishers make Murdoch look like a socialist. *The Guardian*. Available at: http://www.guardian.co.uk/commentisfree/2011/aug/29/academic-publishers-murdoch-socialist.

Open Source Initiative. (n.d.). About the OSI. Available at: http://www.opensource.org/history.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). *In Google we trust: Users' decisions on rank, position, and relevance* 12(3). Available at: http://jcmc.indiana.edu/vol12/issue3/pan.html.

Pyati, A. (2009). Open source software and libraries. In: G.J. Leckie and J.E. Buschman (eds.) *Information Technology in Librarianship: New Critical Approaches*. Westport, CT: Greenwood Publishing, 205-220.

Sadler, E. (2009). Project Blacklight: A next generation library catalog at a first generation university. *Library Hi Tech* 27(1): 57-67.

Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science & Technology* 58(13), 1915-1933.

Stallman, R. (2009). Viewpoint: Why "open source" misses the point of free software. *Communications of the ACM* 52(6): 31-33.

Stevenson, S. (2010). *Carnegie's library philanthropy, with a reflection on its relevance to the philanthropic work of Bill Gates* 26(4): 237-257.

Warschauer, M. (2002). Reconceptualizing the digital divide. *First Monday* 7(7). Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/967/888