Going Beyond Open Data: Challenges and Motivations for Smart Disclosure in Ethical Consumption

Djoko Sigit Sayogo\textsuperscript{1,2}, Jing Zhang\textsuperscript{3}, Theresa A. Pardo\textsuperscript{1}, Giri K. Tayi\textsuperscript{4}, Jana Hrdinova\textsuperscript{1}, David F. Andersen\textsuperscript{5}, and Luis Felipe Luna-Reyes\textsuperscript{6}.

\textsuperscript{1}University at Albany, Center for Technology in Government, Albany - NY, USA, \{dsayogo;tpardo;jhrdinova\}@ctg.albany.edu
\textsuperscript{2}University of Muhammadiyah Malang, Department of Economics, Malang, Indonesia, dsayogo@umm.ac.id
\textsuperscript{3}Clark University, Graduate School of Management, Worcester - MA, USA, jizhang@clarku.edu
\textsuperscript{4}University at Albany, School of Business, Albany - NY, USA, gtayi@albany.edu
\textsuperscript{5}University at Albany, Rockefeller College, Albany - NY, USA, david.andersen@albany.edu
\textsuperscript{6}Universidad de las Americas Puebla, Business School, Cholula, Mexico, luisf.luna@udlap.mx

Received 26 August 2013; received in revised form 12 November 2013; accepted 22 November 2013

Abstract

Although the principle of transparency and openness is not new, recent years have brought increased attention to the need for greater government and private sector accountability driven mostly by the open government, smart disclosure and open data movements. But opening data in the government and private sectors brings about a set of challenges that need to be mitigated if we are to achieve greater information access. This paper aims to contribute to a better understanding of the challenges and motivations for data disclosure in the private sector. We used the sustainable coffee supply chain as a case study, gathering data through a workshop and a series of interviews with coffee supply chain participants and other stakeholders. We identified five challenges and six motivating factors facing data producers along the supply chain. These challenges and motivating factors can be further classified into four areas: market dynamics, information policies, data challenges, and technological capability. We believe that these findings can be generalized to inform discussion and policy design in other market areas. The paper concludes with a proposal for possible future steps to promote openness and innovation in the private sector.

Keywords: Open data, Sustainability, Supply chain, Ethical consumption, Private sector transparency.
1 Introduction

Over the last few years we have witnessed increasing emphasis on greater accountability and transparency in both the public and private spheres. Governments around the world have undertaken efforts to make their operations and decisions more transparent and open [10], [44], [54]; at the same time, private organizations are responding to pressures to become more accountable, transparent and ethical. [8], [17]. These trends are not new; for instance, the interest in government openness and transparency is closely linked to the constitutional right to know [46]. Similarly, increased emphasis on private sector transparency can be traced to the more inclusive definition of a stakeholder, broadening the previous focus on shareholders’ interest as the singular factor driving corporate decision making [8]. While the principles are not new, the intensity of these efforts has increased in the last decade driven by grassroots movements addressing perceived failures by states to protect workers [17]; and through increased consumer interest in sustainably produced goods and services [47]. Tactics used by these movements have also changed, moving from the confrontational stakeholder tactics used until the 80s, to partnerships and collaborative approaches where technology can play a key role [8]. The increased emphasis on the role of information and data seen in recent years stems from the technological and policy developments that make access to useable and accurate data possible.

Open data is defined as “non-privacy-restricted and non-confidential data, which is produced with public money and is made available without any restrictions on its usage and distribution” [30], p. 258. Despite the emphasis on publicly produced data, it is widely recognized that creating value through open data can and often does require using data produced by private organizations. Thus part of the US open government initiative highlights the need to change regulatory approaches to smart disclosure, providing consumers with information to help them make better choices [60]. The basic premise of smart disclosure is that increased access to public and private data will improve the operation of consumer markets. Additionally, providing such data in machine-readable formats has been identified as a key factor in increasing the data’s value for use and innovation [5], [9], [32], [67]. For example, providing technology companies with easy access to machine readable data is expected to encourage creation of innovative tools to process the information and create value for consumers [28], [60], [61].

While the potential benefits of promoting increased access to public data are substantial, realizing this potential faces a number of barriers and impediments, some of which we are yet to fully understand [30], [68]. In the case of opening privately-owned data there are additional concerns related to competitive advantages stemming from information and the expected returns of transparency [58]. In this paper we seek to contribute to this conversation by exploring the challenges to and motivations for promoting greater openness of private data. Using the sustainable coffee supply chain as a case study, we gathered information through a workshop and semi-structured interviews with key stakeholders to explore the challenges, motivations and perceived value of data disclosure.

The paper is divided into five sections, including this introduction. The second section provides a review of the domain of ethical consumption and examples of current efforts to use private sector data to promote better consumer choice. The third part of the paper provides a brief overview of methods used in our research. The fourth section presents our findings about the challenges to and motivations for opening of privately held data. The concluding part of the paper summarizes the main conclusions and proposes possible future steps to realizing the vision of smart disclosure, promoting innovation and supporting ethical consumption.

2 Open Data in the Context of Ethical Consumption

Governments around the world are promoting innovations to influence behavior of consumers, businesses, and other interest groups by implementing information policies that take advantage of new technologies and the vast stores of data being created in both public and private sector organizations. In 2011 the Obama White House announced its smart disclosure initiative, an effort to promote timely release of government-held data in a standardized, machine readable format that would allow them to be merged with proprietary data to increase the efficiency and effectiveness of consumer markets [60]. The basic objective of smart disclosure is to make public and private data available in a way that will help consumers make better choices. We begin this section by discussing the open data and smart disclosure policies as instruments for addressing information barriers, empowering consumers, improving market conditions, and promoting innovation. We conclude the section by introducing specific concepts in the field of ethical consumption, and three innovative efforts in this field built on the premise of open data.

2.1 Open Data and Smart Disclosure

Since 2009, governments around the world have developed a number of policy initiatives to promote disclosure of information held by public and private entities. This concerted movement started when President Obama made a strong commitment to transparency and open government on his first day in office [45]. Through the release of the Open Government Directive, the administration made a commitment to promoting transparency in government operation, public participation in decision making, and collaboration involving a variety of stakeholders [35]. The directive instructed agencies to publish high value information online in an open and machine readable format. The
impact of the directive went beyond the US and led to an international effort to promote open government through the Open Government Partnership. As of 2013, 56 countries around the world have created or are in the process of creating such policies [26].

These efforts provide important context for smart disclosure policies. In its Open Government Partnership National Action Plan, the US government committed itself to promoting smart disclosure, to “ensure the timely release of complex information in standardized, machine-readable formats that enable consumers to make informed decisions in numerous domains” [60], p. 9. The directive recognized that government has information that can be leveraged to help consumers make better decisions and to aid scientific research. To promote better disclosure policies, the US National Science and Technology Council established the Task Force on Smart Disclosure.

According to the task force, smart disclosure refers to “the provision of data about consumer products and services, the companies that supply them, or about consumers themselves (personal data). Data about products, services, and companies can generally be made available publicly, while personal data is provided securely and privately to the individual who is the subject of that data or to recipients chosen by the consumer” [19], p. 7. The goal of smart disclosure is to foster creation of innovative products that help consumers make important marketplace decisions. Although the US government has historically disclosed consumer information, the emergence of Web 2.0 and internet technology further fueled the government’s interest in using this information to serve the consumer in areas such as health, education, energy, finance, and public safety. The task force laid out the benefits of smart disclosure as 1) enabling consumer decision making in complex market conditions; 2) empowering consumers with their own personal data, such as energy consumption patterns and healthcare coverage; 3) creating new data-driven products and services and spurring innovations; and 4) improving economy by enhancing market transparency and efficiency. In order to achieve these benefits, disclosure of information held by public entities is not enough. Vast amount of information that is relevant to consumer choices is held by private entities. To make such information useable and accessible, companies involved in production and distribution need to be encouraged to release data in machine readable formats. Government policy, especially in establishing data standards, can play a key role in promoting consumer access [19], [60].

It is still too early to assess the impact of smart disclosure policies. Open government initiatives overall have encountered many difficulties as opening data involves complex and ill-understood challenges [67]. Critics have pointed out that open government focuses on technology solutions while overlooking the need for adapting organizational practices, policy and culture. They further highlight lack of integration with existing legislation and regulations, lack of clear definitions and measurable goals, divergent and ambiguous goals, and uncertain sustainability [64]. In addition, the tradeoffs between transparency and national security, as well as the potential for economic return to private and public entities, have not been carefully studied and articulated [16]. Moreover, in some cases opening data has not taken into account data use and its context, which can result in the collection of meaningless data sets that will be of limited value to users [6], [21].

The tenet of providing the public with more data in a form that allows for direct manipulation is in accord with the mission of the open data movement [25], [44]. While the current definition of open data focuses on non-privacy-restricted and non-confidential data produced with public money [22], [30] this should be broadened to include private data. In some cases, such as the case of ethical consumption, private information is required to complement public data to create value. For example, combining energy consumption data held by private entities with information about government energy efficiency incentives provides better energy saving alternatives for consumers. Similarly, combination of the Toxic Release Inventory published by the US Environmental Protection Agency (site 5) with company’s sustainability report provides better information to consumers about environmental practices of a particular company. The open data movement believes that open data should share common principles such as openness, completeness, primacy, timeliness, ease of physical and electronic access, machine readability, non-discrimination, use of commonly owned standards, licensing, permanence and usage cost [5], [59]. Some of these principles govern the openness of data accessibility. Other principles such as ease of physical and electronic access, machine readability, use of common standards, and licensing and permanence, govern the manipulation and re-use capabilities of open data.

The principles of open data that encourage greater accessibility, data manipulation, and data re-use have been associated with various potential benefits [32], [67] including support of sustainable and value driven consumption. For example, using open data, Ecodesk is able to provide a web-based platform that enables organizations to search, analyze and compare potential suppliers based on the sustainability of their conduct. Any given organization can then work with the supplier to manage its emissions to enhance supply chain sustainability (Site 1). Another example is Neosis, a web-based platform that uses Green Button standards to help users track and analyze energy performance to improve energy efficiency [33]. But publishing data and information does not always lead to more use and value creation. Factors affecting data use range from user capabilities and data quality to social and technical barriers [27], [30], [68]. In the following paragraphs we briefly describe ethical consumption as a specific context where disclosing information can empower consumers to make better purchasing choices, and current innovations using privately held data to promote ethical consumption.
2.2 Ethical Consumption and the Opportunity for Open Data

Consumer markets have changed immensely since the times of Ricardo [49] and Adam Smith [57]. Instead of a meeting place where producers and consumers meet to exchange products or services in light of perfect information, modern consumer markets are now global networks where producers and consumers never meet or even know each other. Information about product sourcing is systematically stripped away in long supply chains and consumers are presented only with a final product and its final price. Information about how, when, where, and by whom products were produced is hidden in the nooks and crannies of the extended global supply chain.

Neoclassical microeconomic theory holds that commodity markets, such as the market for coffee, are instances of perfect competition and as such assumes the existence of perfect information (rational choice theory). Perfect information implies that both buyers and sellers know everything about every product at every point in time [42]. Given perfect information, rational agents will make decisions in the marketplace that reward the best producers and products [3]. In the area of ethical consumption, the assumption of the rational actor is also widely adopted [20]. Consumer decisions are considered to be guided by the combination of their moral sentiments and economic motives [7]. Carrigan & Attalla [14] found that consumers were dealing with three interacting factors in their purchasing decision: price, quality and convenience.

The existence of information asymmetry, a condition in which companies in the supply chain have better information than consumers, restricts the ability of consumers to drive companies' commitment to ethical behavior through consumer demands [55]. Information has economic value because it enables consumers to make choices that yield greater expected value or utility than could be attained without the information [4], [43]. Increasing information disclosure represents one way to eliminate or significantly reduce information asymmetry by enabling verification of agent behavior [18]. By having information readily available, consumers could verify the company’s commitment to ethical and sustainable conduct; thus enabling them to make informed purchasing decisions. Consumer purchasing behavior in turn influences company conduct, ideally by increasing emphasis on ethical and sustainable models of production and operation.

Consumers seek information about products so that their social and environmental values can guide their ethical consumption behavior. Given that consumers tend to disregard company claims regarding their ethical and sustainable behaviors [15], companies use endorsements by external organizations to create consumer trust in their practices [14]. However, strategies based on attaching certifications and labels are becoming less reliable due to the rapid proliferation of certification and labels in a variety of markets. In fact, the number of third party labeling initiatives has grown rapidly since the 1990s with increasing consumer demand for ethical products [1], [29]. There are currently 444 eco labels worldwide in 25 industry sectors (Site 2). The large number of labels complicates consumer purchasing decision by making it increasingly difficult to discern whether standards employed by a given label are rigorous or represent another form of corporate “greenwashing” or “bluewashing” [23]. Given the lack of transparency, certifications and labeling are no longer seen as adequate for providing trusted product information [31]. This is especially problematic since a survey by TerraChoice in 2010 conducted in 24 stores in the US and Canada claimed that more than 95% of the 5,300 products committed at least one sin of greenwashing [38], [62].

The context described in these previous paragraphs constitutes an opportunity to innovate through open data, involving a diverse set of individuals and organizations, including producers, supply chain operators, certifiers, government agencies, non-governmental organizations (NGOs), and entrepreneurs who will aggregate and prepare data to be used by consumers. In the following paragraphs, we briefly describe some current examples of such innovations.

2.2.1 GoodGuide (Site 3)

GoodGuide was founded in 2007 by Dara O’Rourke, a professor specializing in global supply chains at the University of California at Berkeley. The mission of GoodGuide is to help consumers make value-based purchasing decisions. It provides expert judgment about the health, environment, and social performance of products and companies with a team of scientific and technology experts. GoodGuide provides four 1-10 numeric ratings, one general score and three sub-scores for health, environmental and social performance of the product or the company respectively. Data to build the scores are acquired from over 1,000 different sources, including scientific institutions, governmental agencies, commercial data aggregators, non-governmental organizations, media outlets and corporations.

GoodGuide offers a number of mechanisms for ensuring trust in its ratings. First of all, it strives to ensure a high level of data quality, which builds its trustworthiness. The comprehensiveness (1,000 sources), and timeliness (18-24 month update) of the data contributes to the relevance and validity of its scoring. Second, it uses a group of scientists to translate complex and descriptive information into a simple set of numerical values, making it easier for the less-informed consumers. This expert knowledge is further augmented by the social rating feature of the system. Users can recommend a product and leave comments. The last source of trust is the image of the company itself, being a for benefit organization, and being certified as a B Corporation™, a new kind of company which uses the power of business to solve social and environmental problems.
Despite their efforts, there are a number of data problems associated with GoodGuide ratings. First, the source data are not included on its site, thus limiting the drill down capability of the system. For example, users would not be able to see how a certain product received a score of 9 along the environmental performance dimension. Because of the lack of source data, the mechanism through which descriptive information is translated into the score is not transparent. Lack of transparency, according to GoodGuide, stems from the barriers to publishing original licensed data, use of secondary data already aggregated in a score, as well as the diversity of data sources. The second shortcoming of GoodGuide is that certain product information is difficult to obtain, thus the company level data is used as a proxy for product information. For example, environment and social scores are solely based on the company-level data for many products such as food. This is less desirable if the company produces a wide variety of food products in many regions of the world. A third limitation of GoodGuide is the issue of scalability. Currently, users can submit a request for products to be reviewed, which triggers labor-intensive data collection and processing. Promoting open data among supply chain participants can potentially solve some of these current limitations.

2.2.2 Green Button (Site 4)

Green Button embodies the idea that sharing information related to energy consumption with consumers will lead to behavior modification resulting in savings by those same consumers [28], [56]. Green Button is an application of the smart disclosure policy in the energy and utility industry in the US. It was designed to give consumers access to their own information on energy and electricity consumption to help them make more informed choices about energy spending [56], [65].

The development of Green Button’s capability can be seen in two stages: 1) the consumer-centric phase, and 2) the business-to-business phase [12]. The consumer-centric phase is customer-driven or human-interactive, meaning that the customer has to initiate the effort to download data from the company website. Once the data are downloaded, customers have the discretion to decide what to do with the data and to whom they would entrust the data should they decide to share. Ideally customers would share their data with a Green Button company, usually an information technology company, which would analyze the data and provide recommendations back to the customer. This customer-driven effort entails both benefits and risks. The benefit of customers downloading their own data is that it reduces privacy and security challenges for the energy and utility companies. On the other hand, the consumer-driven approach limits the scalability of the initiative because it depends on the stamina of customers to periodically download their data [51]. The second stage is the envisioned future capability of Green Button. This stage envisions business-to-business relationships where consumers will give consent to a (selected) Green Button company. Having the consent, the Green Button company will periodically download the data, analyze and provide periodic recommendations to consumers about their energy consumption and expenditure.

In understanding the success of Green Button implementation, Sayogo and Pardo [52] posit that successful implementation would not be possible without the combination of market and non-market drivers. The economic risks of the initiative were mitigated by non-market factors that become influential in the acceptance and adoption of public policy requiring for-profit entities to disclose their proprietary data. They also point to the necessary institutional environment enabled by the government and the strategic fit and alignment to the organization’s mission as the critical success factors.

2.2.3 I-Choose

Our current research aims to create a set of data standards in the form of a formal ontology of the certification and inspection process that would allow certifiers and inspectors to post their data in a standard way and allow consumer advocates to gain structured access to the full information package that rests behind each certificate. Consumer advocates would know who did what, when, where and under whose authority to create a given certificate. Such data architecture constitutes a third example of an innovation that uses open data to support ethical consumption. We focus on coffee grown in Mexico and sold in Canada and the US, which gives our initial prototype a relatively tight research boundary with a product with advanced product labeling and inspection standards. As illustrated in Figure 1, the I-Choose supported system would have three basic operational components [36]. The first component is a hand-held shopping support device connected to the Internet, such as a smart phone. The device will be used to scan product identification information, which can be done through a barcode, an RFID tag or a QR code.
The second component is the data architecture described above. The project has developed and successfully tested a prototype using an ontology-based data standard to share product and certification information. The project team is currently working on increasing its robustness through application to real-world data. If fully developed, it would allow multiple stakeholders to seamlessly, reliably, and easily exchange data and to assemble in real-time an information package describing details of any particular coffee.

The third component is a user interface inspired by existing mobile apps such as GoodGuide. This component involves collaboration among consumer advocates and developers who would use I-Choose architecture to create and support innovative services to make available value-based consumer information. This service would overcome current data challenges of systems such as GoodGuide by drawing and interpreting published data using ontology-based data architecture. These three components represent a scalable, trusted information system linking consumers to producers and their supply chains. The success of our prototype, however, is dependent on the willingness of private entities such as producers, roasters and certifiers to make their data widely available and accessible. The challenges and motivations surrounding such information disclosure are the focus of the present paper as described in the remaining sections.

3 Methods

This paper explores the challenges and motivation for greater disclosure of privately held data among the members of the sustainable coffee supply chain in the NAFTA region. We used two primary methods to collect our data: a workshop and in-depth interviews with representatives from the certified sustainable coffee industry (see Table 1).

Table 1: Summary of data collection methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>Timeline</th>
<th>Participants / Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workshop</td>
<td>August 2011</td>
<td>Experts in the certified sustainable coffee supply chain</td>
</tr>
<tr>
<td>Interviews</td>
<td>• May – June 2012</td>
<td>• 9 producers and exporters</td>
</tr>
<tr>
<td></td>
<td>• Nov 2012 – May 2013</td>
<td>• 5 mission driven roasters and importers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 6 representatives of third party certification:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 2 with Institute for Marketecology (IMO) Fair for Life</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1 with Fairtrade USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1 with Commercio Justo Mexico</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1 with Control Union</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1 with internal control of UTZ certified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 2 NGOs in the US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 2 consumers advocates</td>
</tr>
</tbody>
</table>

The objective of the workshop was to identify main stakeholders and how they would be affected, both in a positive and negative way, by greater data disclosure. The two-day workshop involved brainstorming and clustering tasks, as well as small and large group discussions. The workshop was followed by a series of interviews designed to gain more in-depth understanding of the challenges and motivations faced by stakeholders along the certified sustainable coffee supply chain.
coffee supply chain in respect to providing greater access to their privately held data. Each interview lasted between 45 and 80 minutes and was recorded and transcribed. All interviews were semi-structured to provide the flexibility to follow up on new information presented in the context of an interview, and to explore new findings in depth. The interview protocol included the following topics: the mission and objectives of the organization represented by the interviewee, an exploration of information that is currently shared by their organization as well as information that should be shared, the level of trust in the information provided, barriers and incentives to sharing data, and collaboration, governance and certification processes.

In addition to collecting primary data, we also used openly available secondary data in the form of process documentation from seven major coffee certification initiatives, namely: Fairtrade Labeling Organization (FLO), UTZ, 4C, Rainforest Alliance (RAN), C.A.F.E Practices, Organic, and Nespresso AAA (for more information refer to Site 6). We reviewed two types of document: a) the annual report and/or sustainability report (if available) of the seven major coffee certification initiatives for the years 2009 to 2011, and b) standards and compliance criteria documents from FLO, UTZ and 4C. The annual report and sustainability report were important for understanding the financial capability, cost of disclosure, intensity of competition, and self-declared sustainability practices. The standards documents were needed to verify the information collected from the interviews in terms of information chain and data challenges.

4 Findings

This section of the paper presents our main findings drawn from the primary data collected through the workshop and interviews, as well as analysis of the secondary data. The findings are presented in four parts: identification of primary data producers, data gaps in the currently publicly available data, challenges to increased data disclosure, and motivating factors for greater information openness.

4.1 The Primary Data Producers in Certified Sustainable Coffee Supply Chain

Based on the data collected in the workshop in August 2011 we identified seven primary stakeholders in the sustainable coffee supply chain: consumers, producers, certifiers, retailers/roasters, distributors, cooperatives, and consumer associations. The identification and roles of these stakeholders were discussed in detail in Sayogo et al. [53] and Zhang et al. [66]. From these primary stakeholders, we further identified primary data producers of certification and inspection data by identifying the most direct trading route for certified sustainable coffee. These are data producers that generate, store or act as stewards of data related to the sustainable certification of coffee. The interview findings and secondary data analysis indicate that the shortest path of trade consists of producers, exporters, importers, roasters, and third party certifiers (Figure 2). Among these five actors, three are essential data producers of certification and inspection data (producers, roasters and third party certifiers), and two are mediators (exporters and importers).

![Diagram of the sustainable certified coffee supply chain]

Producers in the certified sustainable coffee supply chain generally refer to an organization of small farmers (cooperative), association of cooperatives, or large plantation owners depending on the certification schemes. Producers are the initial source of information and the owners of certification data that are stored in a certification database. Roasters transform raw coffee into finished coffee products, and hold information connecting the finished product with the origins of the coffee, including information about the producers of the coffee. Third party certifier
retains and stores all information and data collected from both the producers and roasters. Third party certifier acts as the steward of the data and information owned by producers and roasters.

4.2 Data Gaps: An Unresolved Issue in Current Efforts to Support Sustainable Consumption

The three examples presented in section two of this paper reveal a currently unresolved problem that restricts innovations supporting sustainable consumption such as GoodGuide, Green Button and I-Choose. The unresolved problem relates to the existence of data gaps in the supply chain, which restrict entities such as GoodGuide from being innovative in offering trusted information services and product recommendations to their customers.

Data gaps create omitted variables bias situations in which data that would be relevant for thorough assessment are not available, and the exclusion of the omitted data makes for potentially biased assessment results. GoodGuide, for example, is unable to acquire detailed information on third-party certifications. As a result, the GoodGuide system assigns the maximum sustainability points to any product that has a sustainability third-party label on it. As a result, a company that decides to donate 3 cents per pound of coffee sold gets the same grade as a company following a full set of environmental and fair trade standards that are inspected by a third party certifier. This challenge is one of the key motivations behind the I-Choose architecture that aims to distinguish between robust and non-robust certification systems [53].

Based on the workshop results and secondary data, there are two possible sources of data gaps in the supply chain: the integrity of data sources and the reluctance of stakeholders in the supply chain to open their data. The methods used by GoodGuide for generating its ratings point to the problems associated with lack of access to data and its website acknowledges this limitation: In many cases, data that would be relevant for a thorough assessment of an important attribute is unavailable for either a product or company. This may be because we have not yet identified a credible data source for a given issue or topic. Or it may be that the data are not publicly available because companies have not disclosed critical information (Site 3).

The results from the workshop also point to the issue of assuring accuracy of information and the ownership of data. Consequently, bridging data gaps is a challenge for open data initiatives such as GoodGuide and I-Choose. An understanding of the challenges and motivation for each data producer in the supply chain to open its data is necessary as a starting point to find solutions to data gaps. This understanding could facilitate identification of what is needed to fill the data gaps, including incentives, policies and governance mechanisms. Identifying incentives for individuals and organizations in the supply chain to participate in open data initiatives is crucial for formulating effective information and data-sharing policies [51], [52].

4.3 Challenges to Opening Data to Support Sustainable Consumption

This section outlines five main challenges facing producers, roasters and third party certifiers as data producers, data owners and/or data stewards to opening their data to the public. A more detailed discussion of challenges to each data producer was presented by Sayogo [51].

4.3.1 Data Challenges: Collection, Accuracy, and Credibility

The first step to making information more widely available is ensuring that the needed data is collected and is accurate and credible. Our interview findings indicate that producers and roasters face a number of challenges when attempting to collect accurate and credible data.

Producers. Certification and inspection data for certified sustainable coffee consists of many pieces of data collected from small farmers by the dry millers, usually a cooperative or a middle-man. We refer to these first middle men as producers. Farmers represent the smallest unit of data source. Ensuring continuous data supply from small farmers is thus crucial for the whole information supply chain for certified coffee. The interviews indicate that requesting farmers to maintain consistent documentation of data is the most difficult challenge for two reasons. First, farmers do not understand the value of maintaining documentation other than obtaining a price premium. Thus, without a premium, farmers saw documentation as a waste of their time. Second, farmers are reluctant to record information about their products if the documentation process is complicated. As a consequence, producers frequently have to assume additional costs to ensure data is collected. Producers often distribute pre-defined and easy-to-fill forms to farmers and/or assign staff, usually the internal control, to solicit data from farmers through interviews.

Additionally, producers have to contend with ensuring accuracy and credibility of data in situations where small farmers motivated by geographical location and financial issues sell their coffee to local intermediaries who in turn sell it to the producers. Our interviewees stated that these local intermediaries occasionally fabricate information about the certification of the product in order to quickly fulfill orders. Most often the intermediaries record and report non-certified products as certified. This issue is significantly lessened when the producer is a local cooperative whose information can generally be more trusted due to well established internal controls.
Roasters. Roasters’ primary challenge is ensuring access and credibility of data due to the interruption in understanding of data provenance. Roasters generally procure their coffee from traders or importers. Our interview findings indicate that some importers do not want to reveal their information or source of information for two reasons. First, our interviewees indicated that some traders **cheapen the spirit of fair trading** by manipulating the producers in the negotiation. These traders do not want to disclose their source of information. Second, hoarding or hiding information is seen by some traders as a competitive advantage and strategy to mitigate threat of substitution. As a result, roasters are often faced with incomplete and inaccurate data regarding the origin of the product they purchase, which in turn affects their ability to provide accurate and complete data for their roasted product.

### 4.3.2 Technology Capability: Technical Expertise, Hardware and Communication

For producers, roasters and third-party certification bodies, opening data is also challenging due to three aspects of limited technological capability: limited technical expertise, limited access to technology (hardware), and issues with communication technology. First, interviewees from roasters and third-party certification asserted that they are sometimes limited by their lack of technical expertise despite their willingness to disclose their data. For instance, several of our interviewees abandoned their efforts to publish coffee contract documents online due to their limited technical expertise. Second, our interviews of producers point to the challenge of accessing technology due to the geographical remoteness of their location. For instance, when trying to sell organic coffee such as US Department of Agriculture (USDA) organic, they are challenged with accessing the certifier information system to publish the required organic certificate. Finally, limited access to technology also affects communication between producers and traders (roasters/importers). For producers, limited access to communication technology lengthens the communication points with traders. Very often these producers have to assign additional staff in a location with access to technology to help with the communication problem.

Technology capabilities could also help during the data collection process at the small farmer level. However, technical expertise and access to technology is a major challenge at that point in the supply chain.

### 4.3.3 Challenges Related to Data Ownership and Conflict of Disclosure Policy

The third-party certification body is the steward that collects and stores all data related to certification and inspection as well as to sustainable trading. Thus, opening the data available in the third party certification database would provide substantial benefits for the development of innovations to support sustainable consumption. However, third party certifiers face two interrelated challenges to publishing their data: data ownership and conflict with applicant’s disclosure policy. First, certification and inspection data is owned by the applicant, usually the producer and roaster, and not the third party certifier. Releasing, accessing and/or publishing these data require consent from the applicant as the owner. Interviewees from Fair for Life certification indicate that the requirement for data owners to publish certification results is voluntary, and there is an option for the applicants to opt out from this requirement. Consequently, certifiers are limited in the amount of data they can publish despite their inclination to publish all information online. Second, the disclosure policy of the certification body often conflicts with the disclosure policy of the applicants with regard to publishing certification data. Applicants have their own information policy governing the release of their data. Applicant takes into consideration their own disclosure policy before deciding to publish their certification data; they are especially concerned about the impact to their brand and reputation.

### 4.3.4 Information Policy: Confidentiality and Economic Value of Information

The interviews identified challenges related to lack of information policies that would address the types and amount of data/information to be disclosed, and restrictions to disclosure in terms of privacy, confidentiality and security of information. The interviewees, especially from mission-driven entities, point out that a major barrier to opening their data is making a decision about what kind of information and how much information to disclose that adds value without violating disclosure restrictions. The interviewees showed willingness to open their data if doing so adds value to the organization and takes into consideration restrictions related to the confidentiality and economic value of information. Some information is closely related to competitive advantage and disclosing it potentially endangers the organization’s market share. There is also the issue of ensuring data privacy. As alluded to in the paragraph above, disclosing certification results might compromise the reputation and competitive advantage of the applicants.

### 4.3.5 Cost of Disclosing Data/Information

Disclosing data and information to support open data for sustainable consumption can be costly for the data producers. The interviewees indicate that indirect costs of certification can be very expensive even if mechanisms exist to mitigate direct costs. For instance, third-party certification generally requires producers to maintain records and documentation to support the certification and/or traceability efforts. Maintaining records and documentation represents a major cost for producers due to the challenges discussed in sections above. Publishing their data online is costly for both roasters and third-party certification bodies as well. For roasters, additional work is required to transfer information from an offline format to online or to public domain. The time spent by their staff on publishing information online is time spent away from their regular job duties, which increases expenses. For third-party certifications bodies, the information can be extensive; for example, audit results might consist of 30 pages with 10 or 20 control points for each category, making open data costly for them.
4.4 Motivation for Opening Data to Support Sustainable Consumption

Despite the challenges presented in Section 4.3 above, there are factors that motivate data producers to support innovation in sustainable consumption through open data. The motivations presented in this paper mostly reflect the perspective of mission-driven traders. Mission-driven traders endorse and advocate for environmental sustainability and social equity by embedding these beliefs into their mission and operating with sustainability and equity in mind [48].

4.4.1 Intensified Market Competition and the Use of Transparency as Competitive Strategy

The interviewees from the first companies to enter the organic and fair trade market in North America highlighted the increased competition in the niche market of sustainable certified coffee as one major motivation for opening data. They cite three major factors driving the increased competition: entrance of mainstream players, declining supply of sought-after coffee strains, and changes in the landscape of third party certification.

First, giant food corporations such as Nestle, Sara Lee, Tchibo, Procter & Gamble, and Kraft, which dominate the mainstream conventional coffee market [63], are increasing their presence in the specialty-coffee market through direct branding and acquisitions. For instance, Nestle plans to invest approximately $5.3 million in its coffee project by 2020 [40] and Sara Lee plans to triple the amount of UTZ-certified coffee purchased in 2011 [50]. Second, production of popular coffee strains such as Colombian coffee has been affected by a leaf rust infection, which has significantly lowered production levels in South America. This has resulted in a market in which large and small scale companies are competing for a dwindling supply of desirable coffee variety. Finally, the fair trade market was significantly affected by the split of Fairtrade USA from FLO in 2011. Fairtrade USA reduced its requirement from 8:2 to 9:1 composition of non-certified to certified coffee for the brand to carry the Fairtrade USA label. In this scheme, a brand is allowed to attach a Fairtrade USA label if it contains at least 10% of certified fair trade coffee in one bag of coffee [41], as compared to 20% under the original requirement. This benefits large plantations with economies of scale and undermines small farmer cooperatives in accessing the market. As a result, small scale producers need to find new ways to differentiate themselves from plantation giants in order to maintain their share of the certified coffee market.

Increased competition creates strong incentives for mission-driven traders as well as small farmer cooperatives to use information disclosure as one such differentiation strategy. According to the interviewees, transparency allows mission-driven companies to compete against large players who have the advantage of economies of scale, greater financial resources, wider brand recognition and an established consumer base. The interviewees from mission-driven companies welcomed the idea of pursuing a disclosure strategy based on open data and interoperable technology such as I-Choose. They also posit that the same goes for small farmer cooperatives, as stated by one of the interviewees, So small farmers who are going through the trouble of creating functioning coops and selling it to Fairtrade, they want to be able to find tools to distinguish themselves from plantation owners … so they are already motivated.

4.4.2 The Value of Opening Data

According to interviewees, companies have an incentive to disclose information if two conditions are met: disclosure of information creates value for the organization and disclosure only requires certain types of information. The interviewees from mission-driven importers/roasters emphasized that disclosure of any information is probable provided that disclosure adds value. For mission-driven companies, their foremost value is in strengthening their mission of environmental and social justice to increase or maintain the organization’s credibility to their consumers. As explained by the interviewees, mission-driven companies embrace open data and technology-based smart disclosure under the perception that it allows them to publicize their exceptional conduct related to economic, social and environmental sustainability. However, most companies have limitations on what information they are willing to disclose, mostly driven by privacy, confidentiality, and economic considerations. For example, companies are willing to publish certification and audit results to showcase their environmental and social conduct, but they will be unwilling to disclose information related to production processes as this information is seen as a strategic asset.

4.4.3 Compelling Information Policies Motivate Sharing

One of the factors contributing to open data innovation in the energy and utility industry is the pre-existence of information policies governing the use and distribution of data, especially the clear demarcation of privacy requirements [52]. Similarly, the three exceptions to disclosure—privacy, confidentiality and security—become one of the major challenges for organizations in certified sustainable coffee supply chains. The interviewees from mission-driven companies and third-party certifiers indicated that the presence of compelling information policies will motivate them to disclose and open their data. As stated by the interviewees, mission-driven companies recognize the value of transparency and existence of robust information policy will help them ensure secure and appropriate disclosure of information.
### 4.4.4 Government Enactment of Disclosure and Environmentally Friendly Procurement Policies

Government regulatory actions can also motivate private sector entities to disclose information and/or to adopt disclosure technology. For instance, the US federal and state governments are beginning to create regulations and policies that might indirectly affect the certified coffee market. We highlight three such government regulations and policies: the US Food and Drug Administration (FDA) Food Safety Modernization Act (FSMA), the California Supply Chain Disclosure policy, and increased emphasis on environmental practices in government procurement.

The Food Safety Modernization Act (FSMA) was signed into law in January 2011 with the objective to improve the safety of the US food supply by granting federal regulators the authority to prevent contamination and health safety risks. FSMA includes three requirements to which the coffee industry should pay attention: food safety plans, supply chain management and record maintenance, and access policies [34]. One of the changes brought by the FSMA is the requirement for product tracing. The FDA description of product tracing systems involves a requirement for mechanisms to enable tracing a product back to the source or forward through the distribution channel. The tracing requirement will affect coffee stakeholders in ensuring records maintenance and access policies to support the product-tracing mechanism.

The California Supply Chain Transparency Act of 2010 is designed to support ethical production of goods and services through greater supply chain transparency. This act specifically requires … retail sellers and manufacturers doing business in the state to disclose their efforts to eradicate slavery and human trafficking from their direct supply chains for tangible goods offered for sale…. Although the law affects only the state of California, its impact can be far greater as companies who want to sell their products in California are required to show that their products were ethically produced at the point of product origination. This in turn provides motivation for producers in international markets to document and share information about their production practices, thus increasing the transparency of the system.

Finally, as pointed out by the interviewees, government can indirectly influence the market by acting as a sustainable institutional buyer to motivate companies to be more sustainable. Efforts by government to support sustainable production have already begun, focusing initially on greenhouse gas emissions. For instance, the US General Service Administration spearheaded the promotion of sustainable procurement, using their buying power to help their vendors to become more sustainable [63]. This stemmed from Executive Order (EO) 13514 issued in October 2009 that calls on federal agencies to establish sustainability-related targets throughout their operations [13], [24]. Considering the magnitude of procurement by the government, the effort to promote sustainability along the supply chain could create strong incentives for greater transparency of production methods along a variety of supply chains.

### 4.4.5 Interoperability and Data Standardization Motivate Sharing

As previously argued, the high cost of disclosure caused by limited technological capability presents a major challenge for organizations in certified sustainable coffee supply chains that are willing to open their data. The interviewees indicated that if the technological costs can be kept to a minimum, mission-driven companies are motivated to take part in open data. One of the interviewees from an importer organization stated, *If we have the map of how to do this better, we’ll be jumping to be more fully transparent and I think there are a lot of companies like us…. There’s nothing we wouldn’t put out there*. They further emphasize that technological costs encompass not only financial resources, but also time, complexity and learning curves. As shown in the case of Green Button, interoperability and data standardization reduce such costs for organizations. Thus data standardization and data interoperability efforts in the domain of certified coffee such as I-Choose can serve as motivating factors for mission-driven companies to pursue technology-based disclosure strategies.

### 4.4.6 Social Pressure: Social Media and Consumers Activism

The interviewees indicated that social pressure serves as a motivation for organizations to pursue technology-based disclosure initiatives. Creating noise and interest in the market is a good way to get industry support behind the initiative. Companies that yield to public demands exerted by NGOs and mission-driven organizations fear reputation damage and resulting threats to their market positions. There is an inverse relationship between brand recognition and the cost of reputation damage. Companies with large brand reputation are more susceptible and very sensitive to safeguarding their reputations because damage to brand reputation is often very costly and can be hard to remove [2]. The interviewees indicated that using social media as a tool to provoke consumer activism is desirable for two reasons: low cost and larger impact. Social media are desirable because the interactivity and network effect magnify the impact of shared information. Intensive use of social media in the *Raise the Bar, Hershey!* campaign created significant attention that eventually forced Hershey to conform to the public demands to pursue more sustainable practices for their procurement and supply chain [11]. On the other hand, our interviews also revealed that producers and roasters in countries where consumers attach no value to sustainable practices show little or no interest in any certification or compliance process.
5 Conclusions

The key to successful implementation of data-driven tools designed to support sustainable and ethical consumption is increased public access to private sector data. Unlike government actors, private entities view information as a strategic asset and significant economic tool. Decisions to make data publicly available are driven by a variety of factors some of which support increased openness while others present challenges to greater information sharing. Understanding the interplay of such factors is crucial to designing policies and incentives to support smart disclosure of private data. The aim of this paper was to provide deeper understanding of the challenges faced by data producers along the sustainable coffee supply chain as well as motivating factors for greater openness. Although the research supporting our findings focused narrowly on certified sustainable coffee produced and sold in the NAFTA region, we believe that our conclusions are relevant for other industries as well as to other geographic areas.

Our research resulted in the identification of five challenges to and six motivating factors for greater data openness facing data producers in the certified sustainable coffee supply chain. These challenges and motivating factors can be further grouped into four general areas influencing smart disclosure: market dynamics, information policies, technological capabilities and data challenges. First, the growing intensity of market competition, increasing social pressures generated by mission-driven organizations and the public, and the economic value of information create market dynamics that motivate data producers to increase their transparency via open data as a competitive strategy. Second, information policies addressing issues of confidentiality, secrecy and privacy, are seen as one of the major future determinants of increased data access. Absence of such policies can hamper any future efforts within the sustainable supply chain due to the perceived unclear demarcation of legal boundaries. On the other hand, existence of compelling information policies addressing these issues can serve as a motivator for greater data openness. Third, insufficient technical capabilities and lack of adequate technical infrastructure create a major impediment for data creation as well as data sharing. At the same time, efforts aimed at increasing data interoperability through common data standards can reduce the cost of data production and collection, and thus can serve as a motivating factor. Finally, data challenges encompass issues connected to data integrity (accuracy and credibility), data ownership and conflicting disclosure policies. Finding mechanisms that can alleviate some of these challenges can increase the value of these data as well as the value in their disclosure.

Our research leads us to believe that some of these challenges can be addressed by the creation of a scalable socio-technical system to facilitate information sharing and interoperability among stakeholders in the supply chain [36]. Such a socio-technical system should at minimum consist of three different components: a set of data standards and data architecture to enable sharing of disparate information across the supply chain, a set of Application Programming Interface (API) standards to make it possible for developers and other interested groups to create specific applications to make this information usable by regular consumers, and a governance system charged with the creation and periodic revisions of these standards [37].

Data standards along with a well-developed data architecture have the potential to make a vast amount of product-related data available to consumers who can in turn use these data to make more informed purchasing decisions. However, providing consumers with a vast amount of data has limited value as only a few consumers can discern the meaning of the different data points of different certifications. For that reason, other actors such as consumer advocates, private organizations or NGOs are needed to create data-driven tools that would match a product to specific consumer values and thus help consumers use the information in a better way. We believe that opening data will benefit existing efforts such as GoodGuide by providing access to currently hard-to-obtain information. The second component of the socio technical system, namely a set of API standards, would facilitate easier development of innovative tools that take advantage of open data.

The two technical components need to meet the requirements of usability, utility, trust and openness to be widely adopted and used. To accomplish this objective, we believe that a multi-stakeholder governance structure is needed to collaborate on setting, maintaining and periodically adjusting the above described standards. The governance structure would need to involve a wide range of stakeholders including government regulators, industry associations, consumers, consumer associations, and producers. The governance structure would also need to have a well-developed set of internal policies and procedures to ensure fairness of the process to avoid situations where the standards are driven only by the large players in the supply chain.

Governments also have a role to play in this system. Governments already play an active role in providing information to markets by requiring labeling (for example FDA-mandated food labeling), by providing certification standards (USDA Organic certification), and by providing legal structures for consumer protection, especially in markets characterized by information asymmetries (automobile lemon laws). Certification and labeling schemes as private regulation mechanisms produce limited trusted information needed by consumers to make informed decisions. Consequently, regulatory policies focusing on transparency and disclosure play a role in helping consumers mitigate information asymmetry as well as possible unethical practices in the market such as greenwashing [21]. Such policies must be well designed to avoid information overload stemming from indiscriminate requirements for full disclosure. Such information overload can overwhelm consumers and consumer advocates, while at the same time obscuring important and relevant information that is needed by consumers to make informed decisions.
judgments [6]. Regulatory transparency policies must incorporate comprehensibility of information disclosure as a crucial element of information disclosure [21]. Consumers need tools as well as uncomplicated and relevant information to help them make purchasing decisions that are consistent with their values. Governments, particularly in the US and UK, have recently promoted policies on smart disclosure to give consumers more control over their own data to make informed choices for their purchases [60]. This policy encourages private entities to increase their disclosure practices to support consumers in making informed judgments and decisions. The release of such information is expected to encourage technology companies to create innovative tools based on information technology to manipulate information for the benefit of consumers [28], [60], [61].

Policies promoting smart disclosure are better strategies than those asking consumers to understand the intricacies of product information, which is especially true for complicated products such as financial investments. Consumer behavior theorists posit that consumers’ decisions are commonly affected by behavioral biases resulting from information overload and aversion to complexity [39] that consequently obfuscate critical information for making desirable consumer choices [61]. Smart disclosure policy encourages the development of simple tools that reduce complexity and information overload and facilitate consumers’ behavior in making good decisions. The combination of open data and smart disclosure policy will enable innovators to collect data pushed out by government, private entities and researchers, and create tools that can help consumers make choices that best reflect their values, including social and environmental values.

Acknowledgments

Research reported in this paper is supported in part by the US-NSF (Grant No. IIS-0540069), CONACYT-Mexico (Grant No. 133670), and COMEXUS-Fulbright Commission in Mexico. The views and conclusions expressed in this paper are those of the authors alone and do not necessarily reflect the views of COMEXUS, CONACYT or NSF. We also want to acknowledge help from Deborah Lines Andersen in the final preparation of this manuscript.

Websites List

Site 1: Ecodesk
http://www.ecodesk.com

Site 2: Ecolabel Index
http://www.ecolabelindex.com/

Site 3: Good Guide
http://www.goodguide.com/

Site 4: Green Button
www.greenbuttondata.org

Site 5: EPA Toxic Inventory Release
http://www2.epa.gov/toxics-release-inventory-tri-program

Site 6: Coffee Barometer 2012
http://issuu.com/spanhuysen/docs/ltc_coffeebarometer

References


[34] G. J. Kushner and E. B. Fawell, FDA Food safety modernization act: What it means to the coffee industry and how to prepare, presented at the Centennial Convention of the National Coffee Association, February 8, 2011.


