Information architecture for sustainable development

M. Silberman* and T. Heikkila†

Revision -1

June 3, 2007

Contents

1 Introduction **3**

2 The uses of information **3**
- 2.1 Strategic planning . . . 4
- 2.2 Work allocation and collaboration . . . 4
- 2.3 Performance assessment . . . 5
- 2.4 Financial management . . . 5
- 2.5 Relationship management . . . 5

3 Information architecture **6**
- 3.1 Whither information architecture? or, the parable of the oranges at the bazaar . . . 6
- 3.2 Data, information, and knowledge . . . 8
- 3.3 Quantity and efficiency; exploration and exploitation . . . 9
- 3.4 Flavors of data . . . 10
  - 3.4.1 Content and constraint . . . 10
  - 3.4.2 Unstructured data . . . 10
  - 3.4.3 Structured data . . . 10

4 Information architecture for organizations **10**
- 4.1 Strategic planning . . . 11
- 4.2 Work allocation and collaboration . . . 11
- 4.3 Performance assessment . . . 11
- 4.4 Financial management . . . 11

*Corresponding author. Arts Computation Engineering, University of California, Irvine. six@fiveplusone.net.

†Institute for Social and Economic Research and Policy, School of International Affairs, Columbia University.

4.5 Relationship management 11
4.5.1 Extra-agency collaboration 11
4.5.2 Stakeholder engagement 11
4.5.3 Accountability and transparency 11

5 Information applications **11**
5.1 Examples 11
5.1.1 Wikipedia 11
5.1.2 Facebook 11
5.1.3 Twitter 11
5.1.4 Craigslist 12
5.2 Challenges 12
5.2.1 Contextualization 12
5.2.2 Functionality 12
5.2.3 Usability 12
5.2.4 Extensibility 12
5.2.5 Customizability 12
5.2.6 Overload 12
5.3 Additional challenges for information applications for sustainable development 12
5.3.1 Internationalization and localization 12
5.3.2 Interface ruggedization 12
5.3.3 Maintenance 12
5.3.4 Infrastructure: electricity, computational power, bandwidth 12

6 Conclusions and future research directions **12**

1 Introduction

Our individual lives and the lives of our organizations are shaped by information. The unfolding of critical events often hinges on the timing and manner of its availability, and to whom it becomes available. In this paper we wish to explore ways in which organizations involved in sustainable development activities might make use of novel tools and techniques for structuring the physical, organizational, and temporal collection, storage, analysis, and distribution of information in order to more effectively achieve their aims.

Information—and especially asymmetrical information—plays a prominent role in the literature on political economy [C], and a substantial body of literature exists addressing the mechanics and methods of "learning" within an organizational context [C]. The role and practice of information architecture itself, especially as manifested in information technology applications, appears somewhat underrepresented in the academic discourse; discussion regarding specific opportunities and applications for organizations involved in sustainable development is, to our knowledge, nonexistent.

To establish a conceptual framework for exploring potential uses of information architecture in a sustainable development context, we first discuss in a nontechnical manner the ways in which the availability of information shapes the lives of individuals and organizations. We explore decision-making broadly speaking, as well as specific moments and processes in organizational life, including strategic planning; work allocation and collaboration; performance assessment; financial management; and intra- and extra-agency relationship management. We then define information architecture, and begin to build a vocabulary for thinking about information applications and the contexts in which they live. Next, we use this vocabulary to explore ways in which information applications can be made useful in an organizational context. Finally, we discuss some of the central challenges of information design, and some of the additional and unique challenges presented by the sustainable development context. We conclude by commenting on a few open theoretical questions of relevance.

It is our hope that this paper will serve to introduce formally into the sustainable development discourse concepts from the burgeoning field of information architecture, and especially from the nascent and endearingly rambunctious hodgepodge of theory and practice surrounding the theory object occasionally referred to as the 'semantic web.' We hope that these ideas will be of some use.

2 The uses of information

Before going on to talk about what information architecture is, and how we might be able to use it to aid our organizations in their work, it seems reasonable to at least attempt to outline the problem domain under inspection. Although

information availability affects so many decisions in the lives of individuals and organizations, an enumeration of organizational processes in which information plays a central role is instructive: it allows us to understand each process as a possible locus for design intervention.

2.1 Strategic planning

Generally speaking, organizations are brought into existence to fulfill a particular objective or **set of objectives**. These objectives can be hierarchical, temporally specific, and self-modifying. A for-profit corporation, for example, has as its primary objective the production of shareholder value. Contingent on the expertise and resources assembled within the corporation, this primary objective may generate secondary objectives, such as 'produce and sell n vacuum cleaners per month.' The use of incomplete or inaccurate information in deciding objectives can lead to adoption of impossible or conflicting objectives which are understood as such only later, after much time and resources have been wasted.

A set of objectives taken in conjunction with an understanding or **model of the system** under consideration—the demographic to which the vacuum cleaners are to be sold, for example—yields a **strategy** or **plan** for achieving the articulated objectives [1], [2]. Here again the organization relies on accurate information about the context in which it is operating: if the model employed to formulate the strategy is based on mistaken assumptions, incomplete or inaccurate data, or faulty reasoning, the model may not provide an accurate understanding of the systems and processes under study. Implementation of the strategy thus derived may yield outcomes different from the expected outcomes.

2.2 Work allocation and collaboration

The strong effect of technological change on labor organization is well documented in literature on firms in both Europe and the United States, but the economic literature focuses primarily on the impact of declining communication costs while abstracting the structural details of the communication itself [3], [4], [5], [others?]. It is well understood that intra-organizational comparative advantages among departments and individuals lead to an optimal allocation of labor given a particular set of tasks [6], [others?]. Further, given a time-dependent series of tasks and a set of workers with a nonidentical but not necessarily unique skills and finite time-constrained output capacities, there exists possibilities for useful collaboration [C?]. Suboptimal labor allocation schemes and overlooked opportunities for collaboration represent distributional inefficiencies which can be corrected given complete and timely information about worker capabilities and availabilities.

2.3 Performance assessment

Development of effective performance assessment criteria and mechanisms is a notoriously challenging task, and historically many organizations have made use of a hodgepodge of subjective and objective criteria in evaluating worker performance [C]. The comprehensiveness of these criteria are constrained by construction: the cost of retaining or firing a worker based on incomplete or inaccurate information is often difficult to estimate [C], so firms and organizations tend to err on the side of inaction [C? I don't know if this is actually true]. We assume that proof beyond reasonable doubt of contract noncompliance must be accumulated before a worker can be justifiably let go, but if the time rate of information collection regarding the worker's activity (i.e., their net benefit or cost to the organization) is low, this caution may ultimately be costly. Higher rates of information collection allows for more rapid but equally informed decision making [C?]. Under some information architectures, however, raising the sampling rate of information collection may be costly, or processing costs may be prohibitive. In some situations, the one-time cost of overhauling an organization's information architecture may be less than the cost of raising the sampling rate under the existing architecture, and less the accumulated cost of persistently inadequate or inappropriate distribution of information. Because the latter cost is generally not explicitly accounted for, however—and of course, an accurate understanding of the costs of information underavailability or misdistribution is prevented by the lack of adequate information—strong incentives may exist against dramatic change or adoption of new tools or methods.

2.4 Financial management

All organizations, both for-profit and not-for-profit, must acquire and manage capital resources. The acquisition of capital by any organization is largely a process of action based on information acquisition and management: for-profit organizations must identify, analyze, and develop business opportunities, while not-for-profit organizations must identify potential funding agencies and partners, cultivate donor relationships, and demonstrate financial and operational responsibility.

2.5 Relationship management

Organizations must also maintain relationships with both funders (or, in the case of for-profit firms, customers) and other stakeholders. Stakeholders may form a large a diverse network of private citizens, government officials, donors, agency functionaries, and private for-profit organizations. Serious challenges that fall under the broad rubric of "relationship management" and that confront almost every organization include **extra-agency collaboration** [7], **stakeholder engagement**, and **accountability and transparency**. These three are critically linked, and are fundamentally information collection, management, and distribution problems. Organizations dealing with the increasingly recomplicated and

interdisciplinary challenges posed by doing business (or not-for-profit work) in the digitized, globalized, hypernetworked 21st century are likely to find stakeholder responsiveness and transparency to be critical assets [8].

3 Information architecture

"Information architecture," as a practice and as a body of theory, is famously (and perhaps rather ironically) difficult to define satisfactorily. In some circles, it is assumed that the practice of information architecture is one limited mostly to the realm of the World Wide Web, and generally involves the arrangement of hyperlinked bits of content and the visual arrangement of the hyperlinks that connect them [10]. One popular articulation of the primary concern of information architecture is "the structural design of shared information environments" [11]. We wish here to take a slightly broader view, and to consider as well the problems generally studied under the moniker "interaction design." Although practical and somewhat compelling reasons for resisting this "discipline creep" have been advanced [12], the two cannot be separated easily when attempting to understand the role of information in decision making: the receipt of information by an individual must be understood as an *event* which leads to that individual making decisions different from the ones he or she would have made without that information. We explore the role of information in decision making in the following scenario.

3.1 Whither information architecture? or, the parable of the oranges at the bazaar

Alice arrives at the local open-air bazaar with the intent of purchasing as many oranges as possible for no more than $5.00. She first comes upon Bob, who sells oranges for $1.00 each. She's about to buy five oranges from Bob when she is accosted by Carol, who tells her that David is selling oranges for $0.80 each on the other side of the bazaar, and offers to show her the way to David's orange cart. Alice is now confronted with a decision: should she buy five oranges from Bob at $1.00 each, or should she follow Carol? Assuming all oranges are equal, that the opportunity cost to Alice of traveling across the bazaar to David's cart is negligible, and that Carol is telling the truth, Alice might be wise to follow Carol. But none of these assumptions are guaranteed to hold: David's oranges may be of a lower quality; the bazaar may take a great deal of time to traverse (and the opportunity cost of doing so may exceed the utility cost of the extra orange); and indeed Carol may be lying: David may not exist at all except as a character in Carol's nefarious scheme to divert Alice's attention while a shady accomplice relieves her of her $5.00.

What information will enable Alice to make the optimal (utility-maximizing) decision? At the very least, she needs to know if David exists at all; how far his cart is, how long it will take to get there, and the opportunity cost of making

the trip; and how good his oranges are compared to Bob's. There may be other considerations entirely unrelated to oranges that affect her decision: consider for example the possibility that the bazaar is split into two halves, with the half containing David's orange cart on the opposite side of a fiery chasm spanned only by a rickety rope bridge. If Alice has a paralyzing fear of rope bridges, she will reach the chasm only to find that she would have been better off buying Bob's oranges in the first place, and will forfeit the time opportunity cost of walking from Bob's orange cart to the fiery chasm and back again.

The scenario thus far has assumed that Alice arrives at the bazaar with no knowledge of the goods on offer; she merely assumes that she will be able to buy some nonzero number of oranges for less than $5.00. Let us maintain this assumption, but now conjecture that upon her arrival she encounters a directory of vendors, and learns that the entire bazaar contains three sellers of oranges and one seller of fruits. Even if she encounters Bob first, she may feel compelled to visit the other vendors to ascertain which will offer her the best price (and whether or not the fruit vendor sells oranges at all). Encountering the directory and obtaining the information that there are other vendors who might offer her a better price has changed Alice's behavior. The information has *intervened*; the directory itself constitutes a design intervention. Whatever nefarious character installed the directory can be said to have committed an act of information architecture.

But is it a commendable act? Let us assume that the directory lists only the vendors and their locations within the bazaar. This is just enough information to be dangerous: no matter which vendor turns out to have the best price, Alice will have to visit *all* of them in order to find out which vendor has the best price. Consider the case in which her default action (absent any intervention) is to purchase oranges from Bob, who offers the second best price in the bazaar. What if the opportunity cost of obtaining the information about the vendor with the *best* price turns out to be greater than the benefit obtained from purchasing from the vendor with the best price rather than the second best price (i.e., from Bob)? In this scenario, the information provided by the directory has caused her to behave in such a way that her total utility is in fact *lower* than it would have been had she simply bought her oranges from Bob; here, a little learning is a dangerous thing. (The problem is even worse if Bob turns out to have the best price after all.) On the other hand, if her utility gain is less than the opportunity cost of obtaining the complete price list, the intervention is (on balance) good, at least for Alice.

This, of course, is a vastly simplified scenario: information problems in the lives of organizations—and even in the everyday lives of individuals—are orders of magnitude more complicated and may require a great many variables and possible state-action combinations to be modeled with useful fidelity. But the complexities of even this simple example serve to illustrate the difficulties confronting any would-be designer of information delivery systems: how much

is enough? Assuming a nonzero cost for design interventions: what is the optimal intervention for a given situation? Shedding that laughably unrealistic assumption: when is intervention wise, and when is it not worth its cost? Can information escape the law of diminishing marginal utility? How shall we organize and disseminate information so that optimal decisions are made? In the sections that follow, we will develop some theoretical machinery with which to approach these questions, and explore some architectures with which individuals and organizations have begun to address them.

3.2 Data, information, and knowledge

Up to this point, we have been using the term "information" rather loosely, to indicate all manner of things that some theorists and practitioners might perhaps prefer to call "data" or "knowledge" or indeed something else entirely. While it hardly seems appropriate to claim that there exists any consensus on formal definitions of these terms—"data," "information," and "knowledge"—we will claim here for convenience that they describe three (possibly overlapping, possibly disjoint) domains along what we will call a "continuum of meaning," or more precisely a *continuum of semantic context* (see Fig. 1). We will use this assertion to define these terms relationally, and claim specifically that "data" is associated with less semantic context than "information," which in turn is less contextualized than "knowledge."

Figure 1: A section from the continuum of semantic context.

Consider for example the statement

$$q = 50 \tag{1}$$

Claiming that this statement lacks much semantic context—lacks *meaning*, fails to impart much *knowledge* to the recipient, and most certainly fails to elicit any behavioral changes in said recipient—seems unlikely to provoke much objection. To claim that it lies firmly within the category we have described as "data" may be slightly more contentious, but we will claim this anyway.

Consider now, in contrast, the following statement:

$$\text{Bob has } q \text{ oranges, where } q = 50. \tag{2}$$

This statement builds upon statement (1) by adding semantic context: it changes statement (1) from an answer to an uninteresting, irrelevant question ("What does q equal?") to an answer to a question ("How many oranges does Bob

have?") of obvious relevance to at least one person (Alice). We will place statement (2) into the category "information," although with somewhat less confidence than that with which we labeled statement (1) a "datum."

Observe that statement (2) may or may not be *actionable*—that is, the receipt of statement (2)—by Bob or by anyone else—may or may not facilitate a *decision*. To know whether or not it facilitates a decision, we must add more statements; we must provide a broader semantic context. Consider for example an agent "Alice" who receives statement (2) along with the statement

$$\text{Bob sells oranges at price } p \text{ each, where } p = \$1.00. \tag{3}$$

This combination of statements can perhaps be considered "knowledge" if it fits into a preexisting theoretical framework; consider, for example, the possibility that Alice's utility for oranges is described by $U(q) = \sqrt{q}$, that Alice has an available capital stock of \$4.00, and that oranges provide Alice more utility than any other available goods, except pineapples, which are \$5.00 each. Receipt of statements (2) and (3) allows Alice to make a decision: buy 4 oranges from Bob!

Upon seeing this, one might be inclined to declaim further and more concretely something like: "data are merely numbers and facts; information is an *answer* to a *question*; and knowledge is information within a *theoretical framework*." Some authors have insisted that information becomes knowledge only when it is interpreted by a human being [**?**]. This distinction can be employed within the framework articulated above, but we will not insist on it. Indeed, we will not expend considerable effort distinguishing information from knowledge; rather, we will examine the ways in which the delivery and receipt of information changes behavior.

Let us assert finally that the design of an information architecture—to whom it delivers which data, how frequently, and with what context—reveals the theoretical framework from which it was designed. Certain decisions are made about which information will be useful to whom, how frequently it ought to be updated, and so on. Underlying these decisions are assumptions about the functioning of the actual system or environment the information describes.

3.3 Quantity and efficiency; exploration and exploitation

We have demonstrated how more information is not necessarily more useful. We wish now to explore this concept in further depth. To do this, we will employ two terms: **quantity** of information, measured either as a number of "information objects" or, perhaps more appropriately when considering digitally mediated information, as a number of bits, and **efficiency** η of an information interaction. We will define efficiency as

$$\eta = \frac{\partial c}{\partial k} \tag{4}$$

where ∂c is the fractional change in the **correctness** c of the outcome resulting from the decision, where we define correctness as a function of the normalized error ϵ, defined as the normalized difference between the expected outcome x^* of a decision and the actual outcome x, i.e, $\epsilon = \frac{|x^* - x|}{|x^*|}$; specifically,

$$c(\epsilon) = e^{-\epsilon^2} \tag{5}$$

and ∂k is the fractional change in the total cost k of acquiring information in preparation for making the decision in question. Here the expected and actual outcomes x^* and x of the decision lie in the same n-dimensional space, where n is the number of parameters that describe the objective to be achieved via the decision in question.

Let us now assume that there exists some smooth function $\epsilon(\cdots)$ that describes for a given architecture the error as a function of the [what?]

The opportunity cost of (a) gathering information and (b) turning information into knowledge

The critical importance of unplanned knowledge

3.4 Flavors of data

3.4.1 Content and constraint

The more constraints, the less possible information in a pure sense [13]. But structure adds accessibility (i.e., the capacity to turn information into knowledge) and machine readability may add it even more dramatically.

3.4.2 Unstructured data

Generic voice; text; video; audio; etc.

3.4.3 Structured data

Data that adheres to an *a priori* agreed-upon protocol, structure, or schema can allow for more efficient interpretation and contextualization. Additionally, machine-readable protocols like XML can allow for automated analysis and action based on predetermined criteria.

4 Information architecture for organizations

Good information architecture delivers information for producing and aggregating the knowledge necessary for competent decision-making, coordination, and strategic execution. Excellent information architecture does this so concisely and efficiently that stakeholders have time to invest in acquiring new knowledge

outside their domains of specialization.

"Wal-Mart has a highly designed system for the storage, indexing, tracking, and just-in-time delivery of MP3 players and Pez dispensers. Organizations for which knowledge is the lifeblood require analogous systems for the storage, indexing, and just-in-time delivery of information for the production of that knowledge."

Here we explore the ways in which a coherent, well-implemented information architecture can help or hinder the organizational processes and traits described above.

4.1 Strategic planning

4.2 Work allocation and collaboration

4.3 Performance assessment

Additional dimensions of "performance assessment" in the sustainable development context

4.4 Financial management

4.5 Relationship management

4.5.1 Extra-agency collaboration

4.5.2 Stakeholder engagement

4.5.3 Accountability and transparency

5 Information applications

Flavors and components of information applications: "Knowledge base" (infrequent updating) vs. "Keeping in touch" (frequent updating)

Monolithic applications vs. small pieces loosely joined ("glue applications"; f. ex. FacebookSync)

5.1 Examples

5.1.1 Wikipedia

5.1.2 Facebook

5.1.3 Twitter

vs. text messaging, AIM

5.1.4 Craigslist

5.2 Challenges

5.2.1 Contextualization

In the developed-economy context, many economic actors already make heavy use of information applications, which facilitate, streamline, and complicate the task of producing information products, delivering services, making logistical arrangements, and coordinating work between multiple individuals whose engagement may be separated in space and time. In this context, developers of novel applications aiming to simplify rather than complicate the management of information should seek to understand the already information- and application-dense environment in which their applications will be deployed.

5.2.2 Functionality

5.2.3 Usability

5.2.4 Extensibility

5.2.5 Customizability

5.2.6 Overload

5.3 Additional challenges for information applications for sustainable development

5.3.1 Internationalization and localization

5.3.2 Interface ruggedization

5.3.3 Maintenance

5.3.4 Infrastructure: electricity, computational power, bandwidth

6 Conclusions and future research directions

Conclusions

The vocabulary we have developed here—"objectives," "model," "strategy," and so on—can, we suspect, be usefully refined and used to construct a quantitative framework for modeling the ways in which organizations acquire, manage, and use information to understand their environments, acquire and manage their resources, and formulate and realize their objectives. [MORE]

The problem of "information overload" is an increasingly nontrivial one in pervasively networked societies [14], and the very practical question of devising methods for obtaining relevant information ...

The ubiquitous accessibility of vast amounts of information brought about by the falling cost of communication is transforming the way in which individuals, organizations, and societies communicate, think, and live [15], [16]. We hope in future studies to continue developing a framework for exploring the ways in which information architectures do and might affect behavior at all scales: individual, organizational, and societal.

References

[1] Cohen, S. and W. Elmicke, 2002: *The Effective Public Manager: Achieving Success in a Changing Environment*. San Francisco: Jossey-Bass.

[2] Bryson, J. M., 2004: *Strategic Planning for Public and Non-profit Organizations*. San Francisco, J. Wiley & Sons.

[3] Borghans, L. and B. ter Weel, 2006: The division of labour, worker organisation, and technological change. *The Economic Journal* **116** (February 2006), F45-F72.

[4] Ichniowski, C., K. Shaw, and G. Prennushi, 1997: The effects of human resource management practices on productivity. *American Economic Review* **87** (3), 291-313.

[5] Ichniowski, C., K. Shaw, and J. Gant, 2003: Working smarter by working together: connective capital in the workplace. Mimeo, Stanford University.

[6] Lluis, S., 2005: The role of comparative advantage and learning in wage dynamics and intrafirm mobility: evidence from Germany. *Journal of Labor Economics* **24** (4), 725-767.

[7] Brinkerhoff, D. W. and J. M. Brinkerhoff, 2004: Partnerships between international donors and non-governmental development organizations: opportunities and constraints. *International Review of Administrative Sciences* **70** (2), 253-270.

[8] Levine, R., C. Locke, D. Searles, and D. Weinberger, 2001: *The Cluetrain Manifesto: The End of Business as Usual*. New York: Perseus. `cluetrain.com/book/index.html`.

[9] Kapor, M., 2006: Architecture is politics (and politics is architecture). *Mitch Kapor's Blog*. `blog.kapor.com/?p=29`.

[10] Subramanian, S., 2004: An introduction to information architecture. *Sitepoint*. `sitepoint.com/article/information-architecture`.

[11] Morville, P., 2006: Information architecture 3.0. *Semantics: Peter Morville's column about information architecture and findability*. `semanticstudios.com/publications/semantics/000149.php`.

[12] Garett, J. J., 2002: ia/recon. *jjg.net*. `jjg.net/ia/recon`.

[13] Shannon, C. E., 1948: A mathematical theory of communication. *Bell System Technical Journal* **27**, 379-423, 623-656.

[14] Spira, J. B. and D. M. Goldes, 2007: Information overload: we have met the enemy and he is us. Basex, Inc.

[15] Benkler, Y., 2005: *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale.

[16] Sunstein, C., 2006: *Infotopia: How Many Minds Produce Knowledge*. New York: Oxford.