

Working Paper

More than “just shopping:” personalization, privacy and the (ab)use of data

Personalization and Customization Group*

April 19, 2006 - version a

Abstract

Emerging technologies often produce unexpected consequences that existing institutions and policies are unable to deal with effectively. Because predicting the consequences of technological change is difficult, responses to emerging technologies tend to be reactive (if not passive), rather than proactive. Improved understanding of the potential consequences of a particular technology would enable policymakers and analysts to implement appropriate measures more quickly and perhaps even act prospectively. This paper proposes a general approach that can be used to identify potential sources of disruption from emerging technologies in order to enable proactive policy actions to limit the negative consequences of these disruptions.

New technologies are often characterized through the use of metaphors and/or comparisons to existing technologies. While such comparisons provide an easy way to generate understanding of a new technology they often also neglect important aspects of that technology. As a result, the use of metaphors and comparisons creates a disconnect between what the metaphor suggests is happening and what is actually taking place. The incompleteness of the metaphors leads to a disparity in the appreciation of the benefits, opportunities, and pitfalls of a new technology. This disparity allows certain aspects of the technology to be ignored and/or exploited, with potentially disruptive social consequences. An analysis of the mismatch between metaphorical characterizations and the actual attributes of a new technology can help identify otherwise overlooked issues and determine if existing institutions and policies can adequately respond.

This paper uses a study of personalization technologies by online retailers to demonstrate the potential for disruption caused by failures of metaphor to adequately describe new technologies. Online retailing technologies have equipped firms with tools that allow them to move closer to the “mass market of one” — satisfying the demands of a mass market through individually-targeted sales strategies (i.e., personalization). While the metaphors of “shopping” and “catalog” have been used to describe online retail “stores,” these metaphors fail to capture several key aspects of online retail technologies such as aggregation, replication, persistence, and analysis of the personal data easily collected by such businesses. As a result, the institutions that exist to protect consumers when dealing with traditional, physical stores may no longer be sufficient. Furthermore, the pervasiveness of the metaphor undermines the ability of consumers to understand or debate the negative consequences of personalization, especially in the areas of privacy and identity.

*Jason Black, Kieran Downes, Frank Field, Aleksandra Mozdzanowska

Introduction

In Western societies, technological change has been a vital engine of economic growth, as well as a source of considerable social upheaval. As humans have devised increasingly sophisticated “ways of doing,” the contours of our day-to-day life have shifted — often in unexpected ways. While our technological abilities seem to expand at an ever-increasing pace, institutions and cultural norms seem to evolve at a more measured, if not sluggish, rate. Our ability to employ technology to improve and expand “what” we are able to do seems to outpace our ability to understand “how” we are doing it, limiting the ability of institutions and norms to judge whether the “what” is worthwhile in the context of the “how.”¹

New technologies have rarely been suppressed as a result of unanticipated consequences, although such extreme responses are not unheard of. Instead, a sequence of adoption, disruption, and response by innovators, industry, and governments that eventually settles into a new balance among stakeholders has become the norm. Unfortunately, this new balance is rarely path-independent — depending on how the problem first manifests and the nature of early reactions, important stakeholders may arrive late (if at all) to the discussion of remedies. If these stakeholders are disenfranchised, they may only offer grudging participation in the resulting end state (e.g., absence of a performance right in compulsory licensing of the broadcasting of recorded music [Bainwol, 2005; Berman, 1995]) or, worse yet, vocal and active opposition to that state.

While prediction and forecasting are inherently imperfect, identifying likely sources of disruption early on can enable anticipatory action to mitigate potential adverse consequences arising from new technologies. Identification of possible sources of disruption can inform technology developers, who may mitigate some potential effects through design changes. In addition, regulatory institutions and other stakeholders can develop proactive policies to minimize potential adverse effects of these technologies. By anticipating the type of problems that might emerge, and understanding their connection to their technological and social underpinnings, policymakers can better frame and resolve the fundamental issues that these new technologies might present.

Anticipating the potential disruptive consequences of a new technology first requires developing a method

¹Not to be a wicked pain in the ass right off the bat, but we may want to have some attributions here – for example, the idea that tech change outpaces regulation seems to be the subject of a good deal of poli sci literature, and in terms of questions about whether our new ways of doing things are worthwhile, we could cite something like Langdon Winner’s *The Whale and the Reactor*, as this is one of the fundamental questions that animates that book

to identify which new technologies are likely to lead to upheaval. This paper describes an important element of such an identification methodology, based upon the following observation and hypothesis:

Observation: There is frequently a disconnect between (1) the metaphors used to describe **what** a new technology does and (2) **how** the new technology actually works.

Hypothesis: The nature of this disconnect is

- a key indicator that the technology may be socially disruptive;
- an insight into the source of the possible disruption(s); and
- a pointer toward the institutions that are central to managing and mitigating the consequences of the technology.

To explore the consequences of this observation and to illustrate the applicability of the hypothesis, this paper will examine software personalization technologies, particularly those applied to online shopping. More specifically, this paper will explore how the use of metaphor in the promotion and adoption of new technologies can lead to the very disconnects that can prove so problematic. The following sections of the paper are concerned with unraveling and explaining how the process of metaphor adoption occurs and what issues arise when emerging technologies are employed to meet conventional ends. This paper first defines metaphor, mental models, and commensuration, before moving into the example of personalization technologies. In the example of personalization in online shopping, two key mistakes of perception, or “disconnects,” arising from the use of metaphor are identified. The simpler one is the fact that “online shopping” is not merely “shopping,” but something else, and that the use of analogy and metaphor has blinded the user to the differences [Lakoff and Johnson, 2003, Chap. 3]. The larger and more complex problem is that the digitization of human actions, exacerbated by the application of personalization technologies, has led to a truly dangerous kind of commensuration, in the formal sense of the term — the conflation of human identity with a measured, analyzed and deduced “digital identity.”

By applying the metaphorical disconnect hypothesis to personalization technologies, the paper attempts to establish a partial method for the identification of sources of disruption in new technologies.

Metaphor and New Technologies

The language of scientific discovery and technological innovation is marked by the use of metaphors to acquaint the public with the unfamiliar. The “horseless carriage,” the “electronic brain,” the “genetic code” — these metaphors have become inextricably intertwined with the technologies and discoveries they were coined to describe: the automobile, the computer and DNA. They are retained as a part of the history and the language of these developments, despite their deficiencies as either scientifically or technologically accurate descriptors.

These metaphors are the consequence of a common desire among creators and discoverers to make the value of their developments accessible to and usable by the uninitiated. $E = mc^2$ is probably one of the few formal scientific statements that *have* become synonymous with a scientific breakthrough — despite the fact that only a minuscule fraction of those employing the phrase even know what a “rest mass” is, much less are concerned about how to calculate its energy. Technical breakthroughs or scientific discoveries are more often described using language that only approximates the formal characterization of the development. For example, while “electronic brain” is, not at all representative of what a computer does, the phrase captures something of the potential of the technology in a way that is accessible to those who will never write a line of computer code. These colloquial expressions, largely devoid of rigorous descriptive content, can become vital components of the process of popularizing an otherwise inaccessible breakthrough.

The tactics of the developer in the creation of these evocative phrases are clear: (a) identify a characteristic or operational feature of the technical or scientific innovation that (b) connects to something familiar in the experience of the masses, and (c) use that commonality to give the public a useful way to (d) speak of and integrate the novel idea or mechanism into daily experience and, more importantly, (e) build upon that understanding to make creative uses of the discovery or development. Once the new becomes familiar and usable, the developer can focus on dissemination and promotion, without having to depend upon the public to fully understand what really makes the new thing “tick.”

In some ways, these phrases retain a certain charm — the naïveté implicit in them not only makes the developments more accessible, but can also tend to defuse some of the objections that might otherwise accompany new technologies. Consider, for example, the renaming of the “nuclear magnetic resonance”

(NMR) imaging system to its current “magnetic resonance imaging” (MRI). Nothing of the fundamental principle of the device’s operation has changed, but its name has been reshaped to avoid the negative connotation of “nuclear.” Such naming of new technologies is a natural part of the popularization of a new development, and is a powerful mechanism in the framing of the context and meaning of these discoveries. However, because they are based upon a necessarily incomplete conception of the actual technologies that underlie them, it is important to note that there are significant consequences that can accompany a reliance upon such imperfect descriptors.

The disconnects between a rigorous description of a new technology and the metaphors employed to popularize these innovations are not surprising; after all, that is something of the nature of a metaphor.² However, on occasion, the choice of metaphor can have important consequences for the way in which society adapts to the presence of new technologies and discoveries. The case of “intellectual property”, for example, suggests that there are circumstances under which this disconnect can be problematic for policy development. In an era where the rate of scientific and technological advance continues at a rapid pace while the level of public literacy in these fields may not be keeping up, the requirement to rely upon these metaphorical constructs suggests that the number of these disconnects will continue to increase, with a potential negative influence upon society’s abilities to adapt to these changes.³

Metaphors and Mental Models

When a new technology is developed and deployed, the immediate problem for promoters of the technology is how to explain this novelty to individuals who are unfamiliar with it. For those in the field, the explanation can build upon the knowledge that is common to their specific community. But, when speaking to a broader

²We should have a citation for this if we’re talking about the “nature” of metaphor – from Lakoff, maybe?

³Unraveling how the process of metaphor adoption occurs when emerging technologies are employed to meet conventional ends, and the consequences of the imperfections of these metaphors for social adaptation to new technologies, is the objective of the next several sections. This paper suggests that metaphors arise according to this general sequence:

- [a] In response to the deployment of a novel technological mechanism to accomplish a familiar task,
- [b] An imperfect, yet serviceable, metaphor is employed to explain and promote the novel mechanism.
- [c] This metaphor becomes widely accepted despite its imperfections (and possibly because there are information asymmetries arising from the imperfections that give some actors incentives to promote or sustain the metaphor).
- [d] As the metaphor gains traction, a variety of quantitative metrics, framed by the metaphor, are proposed to describe how the performance of the new technology is superior to the conventional ways of doing.
- [e] As these metrics continue to be employed (apparently successfully), people tend to embrace the metric as not merely an indicator of performance, but of *defining* the performance expected of *all* “ways of doing” in this context,
- [f] Tending to focus on a subset of performance attributes (to the exclusion of all others) reinforcing the metaphor despite its defects

audience, the promoter necessarily depends upon other mechanisms to convey the core concepts of the application. The typical approach is to find comparable activities or experiences, and then build upon that commonality to articulate the implications and utility of something novel.⁴

Those who study learning suggest that “mental models” are a particularly potent framework for thinking about how individuals adapt to the availability of new ways of doing. A mental model can be explained as a small-scale model of the external reality [Craik, 1943]. It is the internal model of the world or surroundings that can be used to try out scenarios, understand processes, and decide upon the best actions in a given situation [Johnson-Laird, 1983]. An alternative way of conceptualizing mental models is to employ the notion of the “black box”. Users perceive an input, a black box, and an output. Somehow the input is converted to the output in the box. The mental model allows the user to change the inputs to outputs to conduct the kind of scenario tryouts that users employ to develop a working understanding of new technologies and applications — the mental model is the user’s explanation of what takes place in the “black box” to convert inputs to outputs [Rouse and Morris, 1986].

Mental models are constructed largely through observations and feedback. They are incomplete, abstract, and dynamic. The model is based on gathered information and updated as more information is available or acquired. If a system does not behave as expected, users can update their mental models to include this case [Moray, 1999]. Thus, on one hand, a mental model is a simplification of the actual process, and one that is continually evolving. This evolution takes place because the user devises this model to reflect reality, and model predictions that fail to capture actual behavior drive the user to refine the model to reconcile these outcomes.

Incorrect predictions may have a negative or even dangerous impact in some cases, but they also lead to learning. If an unexpected outcome occurs it is a clear sign that the mental model is incorrect or insufficient. A more dangerous case may be when the mental model is wrong, but the prediction seems correct. More generally, the notion of the mental model can lead to an understanding of some types of logical errors that people make, starting from the basis that such models are incomplete and based almost entirely upon observation rather than upon a systematic exploration of behavior.

Helping a user to work with a new technology can be viewed as helping her to construct a working mental

⁴We could add a citation here to Mark Katz’s *Capturing Sound*, where he makes this sort of argument about the differences between LPs, CDs and MP3s

model of the new technology. Commonly, the approach is not to ask the individual to construct an entirely new mental model; rather, through an appeal to metaphors with conventional experiences, the individual is led to refine and extend an existing mental model to accommodate the technological novelty [St-Cyr, 2002]. Such an approach is attractive because it limits the necessary teaching. Rather than trying to exhaustively train the user in the specifics of the new technology, using metaphors to update or create a new mental model gives them a starting point for making effective use of the technology, while also arming the user with a set of guidelines for refining that learning as the new technology generates unexpected outcomes.

The advantage of this approach is that the newcomer is able to rapidly adapt to the new technology, placing it within an existing conceptual framework that dovetails with past experience — “it works like X, only better.” However, there are also significant potential disadvantages.

The use of metaphor to understand a new technology leads to equating something new to something familiar, thereby easing the process of learning. However, because the metaphor is a necessarily imperfect instrument of description, it highlights only some qualities of a new technology or concept, while blurring or obscuring others. This process leads to a skewed understanding of the world, with the intention of generating both insight and misconception⁵

By associating the new technology with an established “way of doing,” the assumptions and presumptions that underlie the established way of doing can now also be associated with the new technology. And, because a mental model is only as good as the observations that the user makes, effort on the part of developers of the new technology to mimic the established way of doing can serve to mask differences between the two that might otherwise merit attention. In effect, the reliance upon **what** the technology does rather than **how** the technology actually works (and what is going on “behind the curtain”) can be a direct consequence of a reliance upon this kind of learning and adaptation in the face of a new technology.

The need to trade on familiar concepts in order to describe a new technology can be evident in examples

⁵Thompson [2002] argues that the moment of insight that is gained through the use of metaphor comes not when the metaphor is first formed to equate A and B, but when the metaphor breaks apart. The breaking occurs because the differences between A and B are recognized and understood, rendering the metaphor inadequate. See <http://www.engl.niu.edu/rthompson/metaphertigo/> Birner [2004] describes the process as follows:

The metaphor has foregrounded some aspects of the concept while hiding others, but with the lexical stabilizing of the metaphor, the fact of this foregrounding becomes lost. The skewed understanding of the world is taken to be a literal understanding. The metaphoric process is one from which proceeds not only insight, but also, necessarily, misconception.

showing the consequences of its absence. For example, TiVo, the company that sells the eponymous digital video recorder, has been struggling since its formation to build a market for the product. While the name has managed to breakthrough, it is still difficult to explain what it does, and TiVo's fortunes continue to lag.⁶ To illustrate just how bad the explanations of what TiVo does can get, consider that Michael Powell, then-chairman of the Federal Communications Commission and, thus, one of the key bureaucrats charged with regulating telecommunications technology, was obliged to resort to gradiloquent gibberish — “God's machine” — to describe the TiVo that he received for Christmas [Krane, 2003].

This inarticulateness is not limited to any one field. But description through metaphor carries with it a certain risk implicit in the mental model of learning — can the metaphor begin to govern, rather than merely approximate, the public's notions of what the technology does?

Commensuration

In the social sciences, a routine process that can lead to this kind of confusion is called commensuration — the formation of a metric relationship between apparently dissimilar, but necessarily comparable, objects or phenomena.

Commensuration has been described by sociologists Wendy Espeland and Mitchell Stevens as the process of “transforming qualities into quantities,” or “difference into magnitude.”[Espeland and Stevens, 1998] The purpose of this process is to reduce the inherent complexity of fundamentally different objects or concepts by assigning a numerical value to each that can then be easily compared. For example, while apples are obviously different from oranges, when assigned a quantifiable metric — based, for example, on caloric, vitamin, fiber and water content — it becomes possible to compare them. But the process of commensuration is nuanced, with a variety of attendant consequences that belie the otherwise logical process of comparing two numerical values with one another.

Often, commensuration is used to assign values to things that would seem to have no inherent value, or be of inestimable value. Espeland and Stevens use the example of an economist attempting (and ultimately failing) to assign a “value” to recreational activities on a river that would be eliminated by building a pro-

⁶Is this actually true? Do we have any data that suggests Tivo or other digital TV recorders are hurting? What about discussions among broadcasters that Tivo and similar devices are undercutting their business model because they enable users to skip over advertisements?

posed dam. Unable to devise a metric to correspond to the activities in question, the economist eliminates them from the analysis of the dam proposal, thus rendering this aspect of life around the river “invisible”. [Espeland and Stevens, 1998] The ease with which one can compare numerical values can obscure the true complexity of the compared items and their relationship, and encourage an acceptance of the metric over a more nuanced understanding and appreciation of what is being compared. Sociologists have observed that this abstraction away from context commonly happens with successful metrics, leading, in the best cases, to a kind of confusion about what precisely is at issue (such as with standardized testing) and, in the worst cases, to the imposition of a Procrustean bed to compel a kind of comparison that is socially and intellectually offensive [Espeland and Stevens, 1998].

However, it has been observed that, in many cases, once a metric has been devised and widely accepted, the association of that metric with the appropriate context can weaken [Espeland and Stevens, 1998].⁷ At some point, the metric can take on a life of its own, becoming the defining basis for comparison, irrespective of context. Unsurprisingly, this can lead to a kind of conceptual vacuum, where there is considerable concern about the application of such a metric without sufficient language to describe what’s upsetting about the use of the metric. In effect, while commensuration is useful for making comparisons, it carries with it the hazard of conflating commensurability with equivalence — e.g., the idea that, since one can use dollars to measure the relative desirability of an increment of safety, then one can assert that dollars and lives lost are equivalent. When stated so nakedly, of course, such confusions can be easily avoided⁸, but the influence of such confusion upon discourse can be immense.⁹

In the following sections of the paper, we illustrate how metaphors, mental models and commensuration are applied to “online shopping” and investigate the potential implications of their use.¹⁰

⁷See also Gentner *et al.* [2001] which suggests that human cognition may also reshape the metaphor itself in a similar fashion.

⁸See, for example, Weeks [2005] which discusses the foolishness of Amazon’s “Fun Stats”

⁹See, for example, “The Chemistry of a 90+ Wine” [Darlington, 2005], discussing the confusion arising out of the notion that analytical chemistry and a proprietary model can/should predict a critic’s score of a wine vintage. By striving to produce a product whose “Parker score” is over 90, a widely recognized ranking scale that has substantial influence in the marketplace, one loses track that there just might be more to winemaking than making Robert Parker happy. Yet, as one reads this article, all one finds is a vague sense of disquiet about the notion of winemaking to chase a specific score.

¹⁰One easy way we might make better transitions between sections is just being really explicit about what we’re going to discuss next. That said, any descriptions we include like this should follow the basic trajectory we’ve established in the intro.

Descriptive Review of Online Shopping

Online retail has been one of the major growth areas in Internet use in the United States in the last decade [U.S. Dept. of Commerce, 2004]. The ability to search for a broad range of products across a large number of online retail outlets is seen as one of the key drivers of this growth, and has led to the development of new networking technologies as well as new network-accessible retail outlets [Alba *et al.*, 1997; Petrison *et al.*, 1997]. Government policies — most notably, federal preemption of Internet-specific state sales taxes — have also spurred this growth [Pub. L. 105-277, 1998].

The Internet has also become a major marketing channel, enabling the delivery of information about product features, prices and availability not only to consumers actively seeking this information, but also to other Internet users through a wide range of Internet-based marketing and advertising instruments. Furthermore, Internet-based tools allow direct online retail transactions, taking order information, collecting payment information and directing shipping. In the case of digital goods (e.g. software, music), the network can also deliver the actual product.

The expansion of marketing and retail onto the Internet is only the latest in a series of evolutions in that industry as it seeks to exploit the potential of new communications technologies. Newspapers, radio and television each offered new opportunities to reach audiences, and marketing approaches and methods were revised accordingly¹¹. The marketing and retail industries have discovered that it is possible to integrate marketing intimately into the retail experience through the various technologies of *personalization* — that is, the technologies for inferring customer preferences and matching them with available products and services. The of both marketing and retail online, however, is dependent on both the specific technologies themselves and broad efforts to communicate effectively to consumers the “what” and the “how” of those technologies through metaphors.

Personalization Technologies

The rise of the Internet as a marketing channel posed as many problems for marketers as it offered solutions. At the most basic level, the Internet’s immediate, two-way communication infrastructure and low barriers

¹¹Ed note: a marketing history cite needed - not Petrison, though

to entry have increased the complexity of marketing by creating opportunities unavailable in media such as print and television.¹² While the use of television, radio and print media for advertising was well understood by the 1990's — having already gone through their own periods of development and maturation — designing easily navigable websites and constructing good user interfaces was not immediately obvious. Most marketers realized early on that a broadcast model suitable for a medium such as television was not appropriate for the Internet. At the same time, the explosion of Internet-based businesses meant that retailers needed to develop effective models of communication with customers to maximize consumers' willingness to spend online.

Internet retailers were faced with a paradox. On one hand, the potential of tying together webserver functionality with databases of product information could provide customers quick access to precisely the product they desired. On the other hand, the vast range of choices that could be presented through the use of these technologies could easily push customers away by overwhelming them with options.

The effort to provide easily navigated and understood websites stemmed from assumptions among usability experts neatly summed up in the title of Steve Krug's popular book on the subject: "Don't make me think" [Krug, 2005]. This philosophy held that consumers were unwilling to spend much time and energy understanding how to navigate a site before they could make their purchase — if confronted with a confusing site, users were more likely than not to go elsewhere.

Many early sites were often structured like online "catalogs" with limited searching and sorting ability, difficult or unattractive interfaces, and a lack of sufficient and clear transactional paths. By the late 1990s designers had hit upon the notion of "personalizing" the online experience as a way to cope with the competing objectives of simplicity and completeness. In its most basic sense, personalization meant tailoring a user's experience at a site according to that user's behavior — not only over the course of a particular visit, but across subsequent visits. For example, while a customer might be obliged to indicate her preferred language the first time she arrived at a webpage, designers wanted to ensure that the user did not have to repeat that selection each time they returned to the site. This personalized experience was intended both to ease the navigability issues and to provide a more "welcoming" experience in the absence of the person-to-person interaction characteristic of in-store retail.

¹²[editorial note: it may be important to expand upon the other features of the Internet that distinguish it from radio, TV, etc – ease of anyone to participate, anonymity, low barriers to entry, others.]

Unfortunately, implementing personalization was challenging within the confines of the original technologies of the web. A fundamental design feature of the hyper-text transfer protocol (HTTP) lay at the heart of the problem. HTTP was and is a “stateless” protocol, in that the server is not expected to retain any information between any one communication with the client and the next. With no “memory” of what the client has already done, retailers’ webservers were unable to “remember” anything about a previous visit without putting the visitor through off-putting registration and login procedures. While there were many possible methods to resolve this operational difficulty, the combination of the then-widespread use of the Netscape web browser and that firm’s specification of the “cookie” became the basis for the standard methods, later enshrined in RFC2109, for retaining the “state” of the web client as the user navigates a website [Netscape Corporation, 1999; Kristol and Montulli, 1997, 2000].

With the deployment of “cookie” technology in webservers and clients, it became possible to store identifying information on client computers that could be tied to records maintained by retail companies (and accessible to their webservers). The ability to “remember” a user quickly became a basis for doing far more than merely simplifying website navigation. Information routinely collected during retail transactions and reused on subsequent visits — such as past purchases, mailing and shipping addresses, and credit information — gave Internet retailers further ways to ease the mechanics of the online retail experience in addition to “simplifying” the shopping experience. These refinements have included the ability to employ this past information to streamline online ordering [Hartman and Gehlen, 2005], suggest products based upon past purchase history [Bezos *et al.*, 2005] and even suggest products based on online behavior when at the retailer’s web site [Linden *et al.*, 2005].

Personalization and Online Shopping¹³

While personalization may have started out as a shortcut around good website design [Nielson, 1999], the marketing community believed that personalization would be a vital dimension/element of their goal of creating a “mass market of one” [Keenan *et al.*, 2002]. The effective application of information and personalization technology had the potential to deliver, on a cost effective basis, “mass market scale” sales using advertising messages designed at the scale of the individual (see Petrison *et al.* [1997] for a brief history of the evolution of these marketing techniques). A tutorial from 1999 is very explicit about the importance of information-mediated personalization in online marketing:

Business-to-consumer e-commerce led to a rebirth of the concept of personalized or one-to-one marketing, this time on a mass scale. One-to-one marketing is the use of information about an individual to market specific products to that individual that are assumed or projected to be of interest to him or her. It is not a new concept. It is what storekeepers did with their regular customers for as long as there were small town or neighborhood stores. But the advent of automobiles, suburbia, department stores, superstores, shopping malls, and the like, largely turned personalized marketing into a historical relic. It is the thesis of this tutorial that recent advances in information technology, specifically the Internet, the World Wide Web, practical, large-scale database management, techniques for effectively processing large-scale databases, and, indeed, much faster processors, created an environment in which one-to-one marketing was not only reborn, but in which it can be practiced on a mass scale for the first time. [Gillenson *et al.*, 1999, pp. 4-5]

Descriptions of personalized shopping routinely use similar language to bemoan how the evolution of retail from the small shopkeeper to the mega-mall has meant an increasingly impersonal retail experience for the customer [Ling and Yen, 2001]. The introduction of personalization technologies, however, has revised the corporate business models that led retailers away from an individualized retail experience. Careful analysis of data that an individual generates during online shopping (via browsing history, inquiries and sales) enables mass market retailers and partnering financial and credit institutions to mimic attributes of the small-scale retailer, mainly selecting items from a vast inventory to satisfy the specific preferences of any

¹³Ed. note: An issue raised by several commenters has been that we seem specifically focused upon the shopping metaphor without discussing other operating metaphors, most notably the “catalog” metaphor that we spent a certain amount of time discussing ourselves. What might be interesting to think about is to compare the disruptive nature of the “catalog” metaphor with the “shopping” metaphor. In our constellation of harms, is the catalog metaphor more, less or equally disruptive? On the one hand, it makes it possible for the retailer to watch you read the catalog and try to make inferences from that. On the other hand, a catalog, if used for shopping, requires that you give up a certain amount of information in order to make the purchase, but there’s little of the kind of relationship-building that we seem to see in the shopping metaphor sites. Thoughts? — FRF [I added a comment or two about this to one of the paragraphs in the previous sections. May not be enough, but it’s a start.](#)

individual customer. This process of personalization directly offsets the impersonalization that large scale operations typically impose while maintaining operating efficiencies.¹⁴

As the costs of hardware and software used for personalization have decreased, personalization itself has become increasingly cost effective, enabling its application across the retail spectrum. [Petrisson *et al.*, 1997; Tedeschi, 2005] Moreover, the business of retail has increasingly become one of continuous information collection, aggregation, and analysis of customer behavior to develop better targeted marketing and retailing messages. The retail and marketing industry has been able to provide targeted, if not always individualized, attention to the consumer experience while cost effectively increasing the scale and scope of its operations [Ling and Yen, 2001; Wehmayer, 2005; Winer, 2001], through careful application of these tools (e.g., database marketing (DBM) [Lewington *et al.*, 1996] and consumer relationship management (CRM) [Winer, 2001]). Marketers are striving to create a customized retail experience — the “mass market of one” — by collecting individuals’ information and then analyzing that information to extract or infer consumer preferences. Marketers utilize the familiar metaphors of shopping to build the online experience according to existing customer expectations.

While personalization technologies have gotten cheaper, the specific benefits extracted from them continue to be difficult to measure. Though a variety of customer data can be captured and stored during an online retail experience, the value of that data beyond basic demographics and purchase history remains unclear in most marketing contexts.¹⁵

MOVE¹⁶. Others believe that the privacy and nuisance concerns raised by abuse of email are best allayed by strategic applications of personalization technologies that can be used to demonstrate that customer preferences are being noticed and adhered to by marketers.¹⁷

¹⁴Ed. note: are we conflating “shopping” and “marketing” too much here? It might be worth adding a comment or two about how shopkeepers of yore were, essentially, the only marketers anyone ever encountered for a while — marketing as a separate industry evolved alongside the formation of a consumer identity in the United States at the end of the 19th and early 20th century (then went apeshit again in the postwar period). I don’t think we’re causing ourselves a lot of trouble, but it might be worth a comment. I’ll think of something.

¹⁵cite needed: How would we cite an informal interview? My “Pocket Style Guide” has several possible ways - the MLA cite for a “Personal Interview” is “Subject Last Name, Subject First Name. Personal interview. Interview date.” The Chicago Manual of Style refers to a “Personal communication” and would be comparable, possibly amplifying upon who conducted the interview, e.g. “Personal interview with Kieran Downes.”

¹⁶email is not a personalization tech — does not fit here: Additionally, the issues associated with fraud, spam and phishing have, as one marketer put it, “ruined email” as an effective communication channel between businesses and customers, much less an effective marketing channel

¹⁷DELETE: Again, the rate of adoption of particular technologies and the rate of proper application of these technologies are different, which is the source of a good deal of concern among consumer-advocate and privacy groups

In the following section of the paper, we explore the areas where the application of shopping metaphors to online retailing fails to capture the nuances of how personalization technologies operate.

Unpacking Online Shopping

Personalization technologies can be considered as an antidote to increasingly impersonal experiences, shopping or otherwise, that are often the consequence of large scale operations. Increasing the scale and scope of any operation increases the size of the audience (and potential market). In the age of mass marketing, which has existed for some time, marketers have sought methods to expand the size of and profit from their potential markets while cultivating customer loyalty. Repeat customers have long been the cornerstone of many retail ventures, and this fact remains unchanged in the digital era. Retaining the characteristics that encourage repeat business — loyalty to brands, positive customer experiences, personal service, delivery of value proposition — in the face of new retail technologies, however, presents marketers with another paradox: how to feel like the corner store while attracting the audience of the supermarket.

It is instructive to explore the degree to which notions of “shopping” have changed since its migration into the online world. In particular, while personalization has been effectively used to offset the “one size fits all” consequences that typically accompany the development of a mass market retail operation, important changes have occurred in the relationships between retailers and their customers — changes whose implications have been felt, yet treated as the inevitable price of progress.

It might be best to start by reexamining the “ideal” notion of personalized shopping in the conventional sense, as cited in marketing literature. A frequent customer of a retail establishment will develop a personal relationship with the employees and often the main proprietor of that establishment. To a certain extent, these employees will come to tailor their interactions with the frequent customer, for example by identifying specific products whose attributes match up well with the customer’s revealed preferences. A favorite waiter might know how to tailor a food item for a specific customer, or to offer insights into specific menu items according to the customer’s tastes.

Such interactions are beneficial to both the customer and the retail establishment. The customer gains access to market and product information that would otherwise be difficult and expensive to obtain (the

scope of products available, the degree to which those products' attributes match up with the customer's preferences, insights into style, trends, etc.). At the same time, the firm gains through more efficient targeting of its sales message to the customer, as well as fostering loyalty that will help promote future sales.

As a consequence, many retail operations have tried to organize themselves around processes that help to facilitate the development of this type of seller-customer relationship. Salespersons have their "book" of customers, restaurant patrons have their favorite waiters, tables, etc., and there are a host of promotional instruments that have been developed around forming, nurturing and maintaining these relationships.¹⁸

In the online retail operation, technologies have been developed to attempt to provide a comparable experience. Through the agency of cookies and other identification techniques, these sites strive to match up a customer's identity with sales histories (*sales "books" parallel*), store or catalog browsing behavior (*traditional salesperson interaction - suggestion, rejection, refined preferences*), comparable purchases by other customers (*trends/fashion parallel*) and other information to meet the customer's individual preferences¹⁹.

More importantly, the salesperson of the "brick and mortar" retail outlet is replaced with a "digital salesperson" comprising web content delivery technologies, databases, and complex algorithmic processes used to provide personalized recommendations for products /footnoteCUT-determined by an aggregation of the user's purchase history and the purchase histories of other customers [Murthi and Sarkar, 2002; Karypis, 2001]. The online retailer replaces the traditional salesperson with this combination, recommending products and services for sale, conducting the sales transactions and organizing the delivery of the purchased product. The "digital salesperson" knows what is for sale, the exact state of the seller's inventory and delivery infrastructure, a specific customer's purchase history, the purchase histories of virtually every other customer, and other intimate "back office" information that virtually no conventional salespersons could know — and certainly not at the instant the customer enters the store. This "digital salesperson" is engineered to deliver a combination of services at a pace a conventional salesperson could never match.

Thus, the "digital salesperson" can provide clear benefits to the customer seeking a convenient, fast and easy online retail experience. The technologies that enable these benefits, however, merit more careful examination. While personalization technologies have been deployed to recapture the "look and feel" of an

¹⁸Ed note: find cite for loyalty cards; there's one already in the bibliography.

¹⁹Ed note: may need to work on this list

idealized, small town shopping experience, there are significant differences in the “back-end” of personalization technologies which can be quite arresting, if not upsetting. In fact, these differences suggest that, rather than speaking of a “digital salesperson,” it is more apt instead to speak of a “digital familiar” [Barrett, 1999; Black *et al.*, 2005; West *et al.*, 1999] (see appendix B).

Digital Familiars

In the world of literature, from Shakespeare to the Harry Potter novels, a “familiar” is often portrayed as a medium between one world and another. It exists to support the actions and needs of an individual, but its true allegiance is mysterious.²⁰ While it appears to serve the individual, familiars also serve their own ends, and the line between the two is blurry. Familiars possess the ability to know all, but choose to reveal only snippets of information at times of their choosing. The individual who is reliant on a familiar for aid or direction never knows if what the familiar reveals is designed to help the individual achieve her own goals, or to help the familiar at the expense of the individual, or some combination. The individual, without the gift of sight possessed by the familiar, is left with little choice but to trust that the familiar’s intentions are good and the information is reliable. While not necessarily “evil,” familiars are not necessarily trustworthy, either [Wilby, 2000].

It could thus be said that a salesperson at a bricks-and-mortar shop is a familiar — the salesperson might recognize a customer, greet her by name, and direct her to a newly arrived product the salesperson has determined may interest the customer through some calculus that is not necessarily obvious. The salesperson additionally acts as a familiar in that her allegiance is divided between the shop itself, which desires to sell products and make a profit, and the customer, with whom the salesperson may have formed a personal relationship. Such a relationship is, itself, complex — does it exist because the salesperson genuinely likes the customer, or because the salesperson hopes to charm the customer into returning to the shop to buy more products on a regular basis? While sophisticated customers are aware of this complexity and attempt to shape the relationship accordingly, even the most jaded are often influenced by “good customer service”.

A user visiting²¹ an online retail establishment encounters a similar situation. The webpage may greet the

²⁰cite passage re: Ariel and the thing from Harry Potter about not trusting something if you can see where it keeps its brain

²¹metaphor alert: Do you actually “visit an online site??

customer by name, list specific recommendations based on her last visit, or notably alter itself (particularly in terms of navigation) to correspond to settings or actions the user has influenced in some way. The decisions or data points that inform these personalized attributes, again, are often not clear to the user²². A user visiting a personalized site may not, in fact, even realize that the content is being personalized unless they are able to observe another customer accessing the same site and receiving different content.

While the interaction with a human or digital familiar may produce similar results in terms of customer experience, a digital familiar differs from a human sales person in a number of important ways. More specifically a digital familiar is characterized by the following technological and infrastructural features:

- the ability to replicate data;
- data persistence;
- and data analysis and integration.

The following sections analyze these features and their potential implications.

Replication

“Replication” refers to both the potential and ease of copying and transmitting digitally stored data. Information technologies, databases in particular, enable data replication for little or no marginal cost. Digital technology can easily make information collected by a single source available throughout an entire organization. In a retail context, although human salespeople can (and do) share information about customers with the business and with one another, there are significant costs for each instance/transaction, and operational limits to the scale of this sharing. Furthermore, cultural and social influences shape the human salesperson’s willingness to share everything. Factors like competition with other salespeople, trust relationships, and basic logistics tend to mitigate the human salesperson’s ability (and willingness) to share information.

However, digitized personal information can be far more readily shared at little-to-no-cost, either economic or social. The rapid decline in the costs of data storage, coupled with the increasing prevalence of

²²While relating to a real salesperson, a customer can always inquire about the rationale underlying a particular suggestion, such an option is rarely afforded in the online case.

standardized methods of transferring and storing data, have contributed to the overall shrinking of costs associated with maintaining digitized personal information for retail establishments. Digital familiars, which collect detailed information during the course of customer interactions, can therefore pass along any information gained to unlimited numbers of others (with or without the customer's knowledge). These others can include salespeople or telemarketers within the company or outside entities, with few significant technical limitations [O'Harrow, 2006; Solove, 2004].

While replication can have benefits for consumers in terms of easing interactions with large organizations, in other ways it can pose risks to consumers. Ease of duplication facilitates the ability of unauthorized individuals to "steal" the data, leaving the original in place. Without data auditing mechanisms in place, there can be little or no evidence that the information was ever stolen, much less forensic information to facilitate investigation of the crime. The ease of data replication also makes it difficult to determine how many copies of a particular data set are in existence. Thus, for example, it may be difficult for a company to purge its records concerning a particular customer or transaction in response to the consumer's request or even a court order, since simply tracking the number of copies of a particular record is difficult. Morgan Stanley, for example, was cited by a judge during a case in which it continually discovered new email records after claiming they had turned over all such records to the court [Anthes, 2005]. Detecting and assigning responsibility for breaches of data security becomes very problematic under these circumstances, especially when companies/entities are sharing data from a variety of sources.

In addition to draw backs there are also benefits to data replication, such as when a hotel or restaurant chain instantaneously knows the preferences of a customer no matter which location the customer visits and tailors its service accordingly. A hotel may know that a customer is allergic to feathers and therefore replace its standard feather pillows in the customer's room prior to arrival. Such benefits entice consumers into participating in data collection programs, especially when combined with price reductions or other incentives [Schoenbachler *et al.*, 1997; Winer, 2001; Wehmayer, 2005].

Persistence

When dealing with a digital familiar, expectations based on the vagaries of human memory are rendered irrelevant. This is a radical departure from traditional shopping because not only do human familiars forget,

but they also quit, are fired, retire, and so on. Given these facts, customers can expect that at least some of their information has a kind of half-life within the system. For digital retailing, this expectation is no longer reasonable, since every aspect of every transaction can be catalogued and available for recall. As with data replication, this difference offers both benefits and costs.

The benefit is that an individual's preferences will not be forgotten. When a human agent is replaced, much of her knowledge of her customers goes with her. Digital familiars can remember an individual's preferences "forever," cataloging them against future applications. The customer will not have to deal with a new trainee taking over for her favorite salesperson and struggling to meet her preferences. The digital familiar can present a predictable experience every time in perpetuity.

The negative aspect of data persistence is indistinguishable from the statement of the benefits — a customer's actions will be remembered forever. This persistence can be used to exploit weaknesses — for example, a former smoker continually receiving offers for free cigarettes. Additionally, it may create situations where one is continually haunted by one's past — for example, a youthful indiscretion or an injudicious association with unsavory characters or organizations. Just as a criminal record can follow someone for her entire life, one's shopping record may potentially have a similar effect. Consider, for example, the consequences should one's bad credit record be associated with the shopping record. The CRM literature discusses, quite matter-of-factly, the notion that a retailer may elect to "fire" an unappealing customer [Winer, 2001].

Analysis - Integration, Aggregation and Data Mining

Integration includes the ability to combine data from various sources/vendors — Amazon, WalMart, grocery stores, credit card companies, etc. This aspect of digital shopping is mainly a change in scale and scope, since credit card companies have had the ability to track historical purchasing records for many years. Integrating information from personal ads and dating services, chat servers, email, and online purchasing can give rise to a broader range of information, but is not obviously different than the information collection conducted already (e.g. credit bureaus). Particularly notable has been the increase in the use of government-collected data [McMillen, 2003; O'Harrow, 2006; Solove, 2004]

What is different is the extent to which such integration is taking place, and the increasing sophistication of the tools being deployed to exploit it. Data mining is a set of analytical technologies that can be used

to attempt to correlate disparate information about a person to predict future behavior (e.g. purchasing preferences) and has given uneven results to date²³.

It is possible that the dramatic increase in information collected will overwhelm the ability to analyze data, creating information overload, and that data mining will not provide a significant advantage over current data analysis tools. It may be that a few key pieces of data will remain most relevant to marketers/manipulators and the additional data collected will be as valuable as noise. Certainly, the current questioning of the reliability and utility of the methods of customer relationship management techniques suggests that there may be something to this concern.[Russo Dos Santos and Gros, 2003] On the other hand, many believe that database marketing remains the more effective and reliable technique, suggesting that the problems lie not with the notion of data collection and analysis, but in the specification of the objectives of the analysis. [Russo Dos Santos and Gros, 2003]

Fundamental Disconnects

In all of these and in other ways, online retail experiences are significantly different from traditional retail experiences. Notions of "shopping" that are brought from an offline to an online environment clearly fail to highlight these differences, and thus should, conceivably, raise serious questions on the part of consumers, retailers, and regulators alike. But one would be hard-pressed to find much in the way of formal institutional responses to those differences. Aside from specific elements of tax policy, in fact, there has been little attention given at the governmental or agency level to the novelties that personalized online shopping technologies have introduced.

Additionally, this is not simply a retail issue. These technologies underlie a series of profound changes not only in what it means to "shop," but also to the very meaning of identity — and not merely in the digital domain. Given that these changes are redefining key notions of fundamental social interactions in the marketplace, they are leading to changes whose pervasiveness raise important questions that *require* attention.

²³Ed note: On the other hand, the failures to do well have done little to limit the research into doing it better. Moreover, the fact that it works "well enough" for some applications has led to a broader application of the technology — applications within organizations that are less fastidious about the consequences of type I and type II errors in their classification of individuals: Total Information Awareness and its inheritors, for example [O'Harrow, 2006].

(I have a bit of an issue with us saying that the questions are coming. I think to a large extent they are here and have been for a while. If anything we are already a bit late in getting a debate going. So I took out that sentence.)

24

While the vagaries of the political process or the nature of the news media might be able to explain the vicissitudes of any particular incident, it also is very clear that there are substantial obstacles to the development of a useful discourse addressing personalization. We argue that part of the difficulty stems from the continued application of certain metaphors that fail to convey or, worse, obscure the subtle, yet critical, differences between “traditional” shopping and online shopping. Most online retail sites utilize metaphors from the bricks-and-mortar world to describe what is taking place during a user’s transaction: the “shopping cart” icon and ability to retain things the user has selected during their visit to the site; “checking out” when they are finished shopping; and so on. The shopping cart metaphor is understandable to anyone who has ever visited a grocery store. It is a useful way of explaining that, if a user elects to add an item to the online shopping cart, the item will be “remembered” over the course of the user’s session (if not longer), and will still be listed at the time the user wants to purchase the item. However, in a strict sense, there is no shopping cart during an online transaction — there are no “magic gnomes” filling carts in some far-away warehouse at the direction of the web-based shopper. Rather, there is only a comparable functionality, implemented in software and engineered to mimic certain elements of the behavior of real-world shopping carts.

The digital shopping cart, a largely benign example, illustrates our point — the metaphor that is used to explain what transpires during an online transaction and what “actually” transpires are fundamentally different from one another in ways that can obscure important intrinsic attributes of the online experience that are distinguishable from the traditional shopping experience: persistence, replication, and analysis of customer data. Ordinarily, such inconsistencies should lead to a revision of the metaphor — an “update” of

²⁴Before turning to preparations, it is important to examine why there has been such calm acceptance of these profound changes. Organizations (c.f., *EPIC*, the Electronic Privacy Information Center and, more narrowly, *CASPIAN*, Consumers Against Supermarket Privacy Invasion and Numbering) have attempted to bring about a debate, with little success. Sporadic attention is given to these problems from time to time, usually in response to a particularly sensational revelation (online price discrimination [Turow *et al.*, 2005], monitoring of online behavior²⁵, or theft of computerized records²⁶). Yet a full-scale policy or substantive political debate about the underlying causes and/or remedies to these issues never seems to emerge²⁷ (I have to admit that I’ve never really liked this paragraph. I think that we might want to relegate its specifics to a footnote, and just work a transition from the preceding paragraph to the next.).

the mental model, as previously discussed — yet such revisions are not taking place. In fact, an opposite effect can be observed. Increasingly, as the metrics of online shopping performance are being employed to evaluate other retailing approaches, retailing is being reshaped to conform to their dictates (see, for example [Cha, 2006], as well as the following section). In effect, the acceptance of metrics, which demonstrably evaluate performance imperfectly, nevertheless interferes with the revision of the mental model.

“Clicks and Mortar” to Bricks and Mortar

More generally, as users of the new technology become habituated to thinking of the applications the technology in terms of apparently workable metaphors, models, and metrics, they also assume that these new applications occupy the same institutional, legal and social milieu as does the metaphorical basis from which their appreciation of the new technology arises (again, see [Turow *et al.*, 2005]). Under these circumstances, users may operate under assumptions that are appropriate for the conventional ways of doing, but are fundamentally incompatible with the ways in which the new technology operates. For example, despite the name “email,” text messages transmitted over the Internet via the simple mail transfer protocol (SMTP) are, in terms of content privacy, less like “mail” and more like “postcards.” In online shopping, these kinds of disconnects have been well documented [Turow *et al.*, 2005]. *I think more importantly email is not governed by the same body of laws as snail mail. It’s not a federal crime to open and read an email, but it is a crime to read someone’s mail. The only problem with this example is that i’m not sure people actually think of the two as being governed by the same rules*

This disconnect can disrupt, if not defeat, the realization that the social, legal and institutional protections covering a conventional “way of doing” may not address or protect the new “way of doing.” Moreover, adherence to the metaphor despite indications of its failure — which can be reinforced if apparently-objective commensuration metrics become the dominant way of discussing the technology and its applications — can defeat institutional efforts to remediate the negative consequences of these new technologies.²⁸

Some of the consequences of this disconnect in the adaptation of personalization technologies amply demonstrate the problem.

²⁸I was just thinking to myself that we haven’t carried the commensuration stuff through much — if it’s going to be useful, we need to at least refer to it throughout, as we should with metaphors and mental models (which I think we do more often)

(should we add in our 2 areas of disconnect here?? identity)

Opting Out is Hard to Do

The implementation of personalization technologies, while striving to create an experience online that mirrors that of real world shopping, rests upon a series of “behind the scenes” operational details that have substantially different implications for the user. Unlike real world shopping, there is only one “salesperson.” Moreover, this “salesperson” is simultaneously “aware” of what every other user has ever purchased. Further, this “salesperson” also knows (or, at least, has access to) every thing that each user has ever scrutinized (the so-called “clickstream” information). This abstracted “salesperson” is wholly alienated from the conventional social contexts of shopping, so that the customer has no assurance of privacy in her transactions or shopping history, nor any sense of the communication cues that are otherwise available in face-to-face interactions with a salesperson that might engender the kind of trust relationships that are a part of real world personalized retail experiences.

Each of these features (as well as those cited in the preceding section) have a certain intrusive nature associated with them, one that individuals might find offensive or repugnant. As a consequence, customers might elect to take their business elsewhere, as one might expect would be the case in a free market. However, with the increasing prevalence of personalization, database marketing and technology-driven customer relationship management, the alternatives are becoming more and more difficult to find [Turow, 2005].

Additionally, the instruments of personalization, created to facilitate online transactions, have now migrated to physical retail as well. The techniques of “clicks and mortar” retail can now be found in “bricks and mortar” retail, largely in response to competitive pressures [Cha, 2006]. While retailers have focused on creating an engineered “mass market of one” through the application of personalization technologies, a world of continuous “dataveillance” has been created [Clarke, 1988]. With the creation of the world of “everyday surveillance” [Lyon, 2002], the customer has been urged to accept that a desire for privacy is something to “get over” [Sprenger, 1999]. While continually surveillance in a traditional retail context might otherwise be unacceptable, its tacit acceptance in the online world has eased its introduction.²⁹

²⁹We may need a smoother way of integrating this issue, but nothing’s springing to mind right at the moment.

Digital Identity

I think the Identity section should be a section and not a subsection, or at least one layer higher. This is sort of the consequences part 2. First we talk about the consequences in our specific case and then we go broader. I also think that we can bring in discussion of commensuration much more into this section as we talk about metrics and data coming to represent a person.

One of the inescapable consequences of data collection is that, once data has been accumulated, data analysts find new ways to employ it. Irrespective of the original intent behind its collection, once a data repository has been established, innovative uses wholly unanticipated at the outset are developed. Observers have noted that, in a world of increasing digitized data collection, storage and analysis, our day-to-day activities take place in a cloud of digital information that is collected and, though the application of personalization technologies, associated with an individual. With this data, firms have found that the same efficiencies in operation that lead to the “mass market of one” can also be generalized to a host of services other than marketing [O’Harrow, 2006; Solove, 2004].

For example, consider the credit market. When credit meant carrying a tab at the local grocery store, that service was generally only offered to customers personally known to the grocer. Someone unknown to the grocer would not have access to credit, and would have to develop a relationship with the local grocer before it would be offered. The transactions costs on offering credit were high, and the amount of credit offered was low. Today, such a notion of credit is quaint. Instead, credit is offered on a massive scale, facilitated by the availability of a set of standardized reports that have been found to reliably establish a borrower’s credit-worthiness without the need for the development of a personal relationship. In effect, the credit industry has “personalized” the market for credit through the establishment of a set of metrics, based on an analysis of customer data, that measures how good (or bad) a credit risk each customer is.

This story of the credit market is repeated in many other industries — insurance, health care, banking, employment, retailing. With the increasing availability of information tied to the individual, and commercially traded among firms, the costs of classifying individuals, inferring behavior and establishing identity are falling, and firms are devising a host of new services that make use of, and depend upon, that transactional efficiency. However, as more and more services become dependent upon the availability of datasets

of personalized information, these services also begin to institutionalize the very problem raised by commensuration — the conflation of the metric and its measures with the real. Additionally, the objects of the metrics — the people — are often totally in the dark about what metrics are assigned to them, what they mean, and how they are devised and implemented. Like Joseph K in Franz Kafka’s *The Trial*, consumers of credit, insurance or retail items can one day find themselves labeled as undesirable risks by firms with no sense of why or how this determination was made [Solove, 2004].

As Solove [2004, pp. 48-49] puts it in *The Digital Person*:

Increased reliance upon the easily quantifiable and classifiable information available from databases is having profound social effects. The nature and volume of information affects the way that people analyze, use, and react to information. Currently, we rely quite heavily on quantifiable data: statistics, polls, numbers, and figures. [...] The goal of this use of empirical data is to eliminate the ambiguity and incommensurability of many aspects of life and try to categorize them into neat, tidy categories. [...]

[T]he information in databases often fails to capture the texture of our lives. Rather than provide a nuanced portrait of our personalities, compilations of data capture the brute facts of what we do without reasons. [...] In short, we are reconstituted in databases as digital persons composed of data. The privacy problem stems paradoxically from the pervasiveness of this data—the fact that it encompasses much of our lives—as well as from its limitations—how it fails to capture us, how it distorts who we are.

The data stored about any individual, no matter how comprehensive, will always be limited and unable to capture the full identity of the person. Moreover, the algorithms employed to generate derived metrics or indicators of identity (e.g., good credit risk, potential fraud threat, likely terror suspect) will always be subject to type I and type II errors [Clarke, 1988]. Despite these failures, the incomplete digital identity will increasingly become the proxy for each person in the online world and will increasingly be employed in preference to “real identity,” if only because of the transactions costs. In fact, some would already argue that the distinction between real identity and digital identity is false [Zwick and Dholakia, 2004].

The use of the proxy by online entities will inevitably lead to misjudgments about the real person behind the proxy (see, for example, [Zeller, 2006] as well as the many examples in [O’Harrow, 2006]). At the same time, cunning individuals will strive to craft multiple online identities that are tailored to their particular online needs, an action about which authorities may be less than sanguine. The commensuration disconnect in terms of online identity will then work both ways in providing both benefits and drawbacks to both individuals and entities. Savvy online users will of course be more likely to accrue benefits, however, while

naïve users may have difficulty in manipulating their online identity and perceive more of the drawbacks of commensuration³⁰.

Compounding this issue is the degree to which this set of practices has migrated from the commercial world and has become a part of federal, state and local processes in the United States. There are reports that the US government has partnered with commercial data collection resources to devise data mining applications to identify security threats [O’Harrow, 2006]. The legal basis for these systems, which frequently mate commercial data sets with government records and criminal reports, is murky at best, and has been the source of violent reactions when fully appreciated by the public [American Civil Liberties Union, 2004; Associated Press, 2004]. And, when the government’s scrutiny is mated to databases whose contents and validity are the consequence of processes not fully appreciated (or, worse, kept purposely secret), the consequences of error for the individual can be far more dire than a mere rejection of credit (see, for example, “Database Tagged 120,000 as Possible Terrorist Suspects” [Associated Press, 2004]^{31, 32}

³⁰See, for example the ACLU’s “Pizza Shop of the Future,” available online at <http://www.aclu.org/pizza/>.

³¹Although it is beyond the scope of this paper, it is instructive to examine briefly how the defects in the understanding of identity informs one of the great debates of the digital age — the notion of a right to privacy. The unexamined reframing of context implicit in the increasing acceptance and power of digital identity helps to cripple the public debate that is increasingly necessary in the face of the problems and abuses that confront digitized societies. The legal notions of privacy in the US derive from a law review article written over 100 years ago in response to another set of emerging technologies. Warren and Brandeis [1890], surveying the rise of gossip and yellow journalism that accompanied the development of smaller (and, thus, no longer confined to a studio) cameras and improved photograph production and reproduction technologies, identified that something was amiss with the then-current legal structure of harm and redress, and strove to extract a set of principles implicit in the common law that suggested that privacy was a right to be defended. While the legal doctrines of privacy have evolved since then, many observers today indicate that there are fundamental problems at the heart of how one defines the notions of privacy and identity [Kang, 1998; Solove, 2004].

While the legal profession at least has a set of doctrines to fall back upon, the public debate of these topics is arguably completely bankrupt. The rise of the digital identity has made it possible for many to suggest that privacy, like the corner grocery store, is a bit of quaint nostalgia that one cannot afford to retain in the modern world [Sprenger, 1999]. The fact that these digital identities are largely the construct of commercial entities, who make a market in this information among one another, has led to the claim that these identities are property — and, moreover, that they are the property of the entities that created them, rather than those for whom they are stand-ins [Zwick and Dholakia, 2004]. The popular press is full of stories about the perils of these developments, yet the individual is largely left with a sense of helplessness in the face of them, not to mention next to nothing upon which to begin a meaningful debate about the nature of these practices and the uses to which these technologies have been put. The widespread acceptance of the digital identity as an appropriate substitute for human identity, and the reshaping of modern life around the use of that digital identity, has already made debate about the appropriateness of these practices almost impossible — not because there is no problem or because the change is inevitable; rather, it is because the use of metaphors, mental models and commensuration has already enabled the establishment of norms of digital identity without great scrutiny of the bases upon which they are built, or even the acknowledgment that these bases might be distinct from those of conventional identity. In the absence of this collective discussion, the two forms of identity have become interchangeable without the construction of the social and institutional forms that might remediate the negative consequences of the widespread supplanting of conventional identity with a digital identity.

³²I think we need something here or somewhere nearby that acknowledges there is more to this than the dangers of loss of privacy. For example: In the case of online shopping, it is important to acknowledge that the use of metaphors of traditional shopping by online retailers and marketers have encouraged individual users to see the benefits of their online experiences, and those benefits are not insubstantial. To a certain extent, retailers such as Amazon.com clearly provide consumers with products and services they want — otherwise, no one would spend their money there. The key problem is not that these business aren’t offering valuable and desirable services. Rather, it is because, in taking advantage of those services, consumers are accepting potentially perilous compromises with regard to their general privacy, often without realizing it. Or something like that.

What to do About the Disconnect

The increasing pervasiveness of personalization technologies, like so many other technologies, has been a double-edged sword, yielding both benefits and perils. This paper argues that a substantial number of the perils associated with emerging technologies derives from the mechanisms individuals use to cope with novelty. Reasoning through metaphor to incorporate new technologies into their lives, users of a new technology may erroneously assume that, because the new technology yields a result that is similar to the conventional technology, the underlying operation of that new technology is subject to the same set of institutional, cultural and legal constraints that govern the use of the conventional approaches. The basis for making the comparison among the new and conventional approaches, if sufficiently compelling, may ultimately lead to an unexamined reframing of the acceptable context within which both the technologies are used.

This (structural?) resistance to adjustments in the metaphors employed to explain and promote a new technology may help to illuminate a notable cycle associated with the rise of a new technology. In many cases, there seems to be a period of early exuberance and excitement associated with the initial application(s) of a new technology, followed by a period of reflection and reassessment of the consequences of the new applications (*c.f.*, other PoET case studies). In some of these cases, reconciliation of the conflicts arising out of this period of reassessment relies upon a crisis, rather than a measured approach and assessment of the issues arising from the introduction of a new technology. The necessity of a crisis or other heightened sense of urgency to achieve adjustments tends to lead to more extreme reactions than a more careful evaluation might actually suggest, but the intensity of the social unrest arising during the crisis tends to demand that strong measures be implemented/taken^{33,34}

It is almost certainly impossible to break the habit of metaphor construction by those facing technological

³³Editor's note: we really need at least one example. Various nuclear applications? Note that there were many programs around the use of nuclear bombs as part of large scale civil engineering programs, instead of conventional explosives (Project Plowshare – see <http://www.bbc.co.uk/dna/h2g2/A685109>, a film at the Internet Archive <http://www.archive.org/details/Plowshare1961>). The early days of genetic engineering? Pending shutdowns of file sharing networks over copyright, crippling other applications of peer-to-peer networks?

³⁴(I'm not convinced the following paragraph is necessary.) While there are many possible explanations for this kind of extreme overreaction in the face of a crisis, they tend not to explain why a crisis of this sort was allowed to develop in the first place. (I would argue, then, that we need to find a place for this subsequent set of thoughts)— This paper suggests that one key factor that leads to these crises is a kind of widespread complacency, driven by the extent to which commensuration supports the retention of an imperfect metaphorical description of a new technology. If this is indeed the case, then we are in real trouble. This reliance upon metaphor is becoming more and more common in the face of increasingly rapid technological innovation and a decreasingly technologically literate society. Moreover, because of this pace of technological development and deployment, necessarily imperfect metaphors of explanation will be needed by more and more individuals, irrespective of their level of education or technical know-how

innovations. Rather than trying to change this aspect of human cognition, attention must be paid instead to striving to limit the potential for policy overreaction that might accompany the public uproar arising when the metaphors are finally pushed past their breaking point and the public comes to recognize the Faustian bargain that has been made without their consent (or knowledge).. There are two possible strategies that should be given careful consideration.

1. Striving to ensure that the advocates of new technologies are “kept honest” by scrutinizing the metaphors that they employ to promote new technological applications to the public; and
2. Anticipating what will be needed to achieve an efficient resolution of the crises that the breakdown in these metaphors will engender.

In the case of personalization technologies, the question of what is ethically correct with respect to Internet and computer technologies has been debated, but has not been settled. As a result, there is no consensus on what the goal of any action to remedy existing problems would be. X? has proposed both technical and legal solutions. Technical solutions for the most part deal with restructuring the Internet to eliminate anonymity. [However, as the preceding discussion suggests, the problems do not lie with an inability to personalize online information in ways that are beneficial to consumers; rather, it is the problem that institutions are unable to provide these services without also compromising the security and privacy of the consumers, and consumers are not aware of the potential implications of those compromises.](#)

Technological solutions, no matter how well designed, cannot completely eliminate error. Yet they can create an infrastructure that generates and stores even more information and, as a result, end up posing a greater threat to privacy. The alternative is the follow in the footsteps of Europe and set up regulations outlining ownership of information protecting the privacy of customers [Dash, 2005]. Companies clearly have a significant stake in fighting any such legislation.

The larger explanation, and the one that this paper is predominantly concerned with, is the fact that, because of the how adaptation to these technologies has taken place, few of the stakeholders have an effective language for engaging the problems. The widespread acceptance of an inadequate metaphor and the resulting defective mental model has corrupted the dialog so that no meaningful framing of the issue can take place.

Ultimately, resolving the problems will depend largely upon consumers “unlearning” what think they

know about how these technologies operate, and develop a greater level of technical sophistication to both appreciate the risks they are taking, and organize more effectively around regulations and restrictions that protect their interests [Clarke, 1999; Lyon, 2002; Solove, 2004]. Until then, meaningful policy discussion will be quite difficult. In the meantime, the potential for misuse and direct, obvious misuses continue, with both regulatory bodies, online businesses, and consumers relegated to a predominately reactive rather than proactive position.³⁵ Zwick and Dholakia [2004, p. 33] put it like this:

We argue that because of database-driven marketing techniques and CRM, the creation of digital consumer identities is no longer under the control of the real consumer. Indeed, consumer identities are assembled as digital representations in multiple databases without the consumer always being aware of it. Hence, we witness a *multiplication* of consumer identities, as *varied* consumer profiles of the *same* consumer accumulate in dispersed databases. Therefore, in the final part of this article, we propose that only if consumers have direct access to companies' customer databases can they regain control over their consumer identities and reclaim their autonomy in the networked marketplace of the twenty-first century.³⁶

Summary and Next Steps

This paper proposes an approach to identify both potentially disruptive technologies and the likely sources of disruption that will emerge as these technologies are adopted. The use of metaphors to characterize and promote a new technology typically represents an effort to find familiar associations between the new technology and conventional processes and activities, at the expense of a rigorous accounting of the attributes of

³⁵Ed note: One topic that we've talked about before that might merit at least a discussion — the proposed requirement that, like availability of credit reports, organizations that have constructed a digital identity of an individual should make that information available to the individual.

³⁶One of the problems with this, and perhaps this can be part of how we include it in our discussion, is the whole idea that consumer identities in any form were ever under the control of the consumer — I think we've already determined from our historical findings (and basic logic) that this is not and has never been the case. If anything, consumer identities in the past that were useful to businesses were things that were formed by both the consumer and retailer, but consumers have always had limited knowledge of how a retailer constructs an identity for them and then uses that to market products to them. Indeed, the linking together of identities is also something that has existed in the past — that's basically what businesses do when they sell their mailing lists. This is, I'm afraid, once again a question of scale, and ease-of-use: if the Internet and personalization technologies make it easier for retailers to do things we already don't like that they do (like selling our names and information), that's bad, but it's not new. One of the problems with this article is that it's using this new, scary Internet thing as an excuse for suddenly making available data to consumers that has existed for a long time, but before somehow it wasn't enough of a concern to merit such a radical action. Why now? They never convincingly say.

the new technology. While such metaphors can effectively promote the adaptation and use of a new technology, failures to recognize the imperfections of the metaphor as a formal descriptor can lead to problems. The existence of a disconnect between what a technology actually does and common assumptions arising out of widespread acceptance of these metaphors signals that a disruption might occur. Commensuration, for example, is a process wherein these imperfections are not only ignored, but actually can become the basis for reforming established ways of doing to conform to the new, enshrining the metaphor's imperfections in a set of formal, quantitative metrics that, in the extreme, subvert the processes by which institutions respond to social pressures. By identifying such disconnects, it may be possible to anticipate the source and nature of potential disruptions. An analysis of the attributes that are not captured by popular metaphors can help identify areas of potential disruption and determine if existing institutions and policies are prepared to respond to the resulting issues.

A case study of personalization by online retailers has demonstrated this methodology. The popular notions of online shopping fail to capture the attributes associated with data collection & storage and the creation of online identities based solely on such data. As the technologies of online retailing have migrated to traditional brick and mortar shopping, the problems associated with such database identity creation are quickly becoming pervasive. This is occurring with little public debate or *a priori* policy action. Worse, the public debate on these real issues is crippled by the shared, but defective, perceptions that engender an inability even to describe the *nature*, not to mention the source, of the problems.

However, certain concerned organizations are trying to overcome these hurdles. Both the ACLU and EPIC have made efforts to raise awareness of issues related to privacy and the use of personalized data on the Internet. Additionally, both organizations provide channels through which concerned citizens can contact their elected representatives to encourage discussion and possible regulation to protect privacy. The problem is that they tend to stop there. Given the nature of the disconnect and the diffusion of interests around trying to solve for it, neither organization has been able to accomplish much beyond instilling fear in users and consumers who are concerned about privacy³⁷. Reactions from the public and from regulators have thus far been muted by the complexities of the issue coupled with a fundamental absence of a platform for discussion that cuts to the heart of the issue. We have argued here that the basis for discussion must originate from a

³⁷See the ACLU pizza movie, available at <http://www.aclu.org/pizza>

deeper level — that of metaphors and the assumptions they engender — than has thus far been reached³⁸.

Without a meaningful basis for public discourse, policy responses to these problems — mostly likely brought on by crises — are virtually guaranteed to be reactive, if not convulsive. This may lead to a public backlash focused on the symptoms, rather than the fundamental issues that underlie them. In effect, the pressure to take immediate action may lead to rash political decisions, further delaying the necessary public debate and, potentially, crippling (if not killing off) the offending technology. Analyzing the source of the disconnects in order to identify the long term implications prior to the development of a backlash will allow for proactive policy making.

Text Scraps

While many of these phrases are completely benign, it is worth briefly noting a few that carry with them some impressive baggage.³⁹

³⁸ If we're still planning to hose the "What to do about the disconnect" section, which is fine, we should probably have something in this section vis-a-vis the ACLU and EPIC efforts to confront this problem. Feel free to edit the hell out of this — it's sort of the best I can come up with for right now.

³⁹Ed note: Here's an alternative formulation, albeit one that may be less useful — thoughts? Some of these metaphorical expressions exemplify the basic approach taken: to suggest that the new development or application works like, or can be thought of being like, *X*, except that it's now being accomplished through a novel technique that can be inferred from the modifier *Y*. For example, the "horseless carriage" — it works like (and can be thought of as) a "carriage," *X*, only this one is "horseless," *Y*. Or, "computer language" — it works like a language *X*, except this one is for a computer, *Y*. The "electronic brain" — it's able to do (some) things that a brain can do, only it's done so electronically. Once one puts one's mind to it, the ubiquity of this kind of linguistic construct to integrate new concepts and techniques into the mainstream can be surprising.

What is even more striking is the degree to which these phrases persist long after these new concepts and techniques have been wholly integrated into daily life. The automobile has been around for a century, yet everyone still acknowledges that a "horseless carriage" is an automobile — even though a vanishingly small fraction of these people have any direct experience of the use of a horse-drawn carriage.

Appendix A — Applying the approach to other emerging technologies

This methodology can be applied to other emerging technologies as well. By focusing on the disconnect between the popular metaphors for a technology and the actual attributes of that technology, policy makers can identify the likely areas for disruption as the technology is adopted. There are several possible cases that might help to refine and further explain this perspective on the dynamics of adaptation to emerging technologies:

- Automobile technologies (horseless carriage; getting rid of one form of pollution (noise as a second form?); commensuration — “it’s cleaner;” discovering the limitations of definition of clean)
- Synthetic biology — This is a field that is presently in the process of developing a metaphor — synthetic biology is just like computer programming; one codes up a set of base pairs, compiles that code into a some genetic material in a sequencer, and then executes the code by implanting the genetic material into a cell [Ferber, 2004]. Moreover, this programming metaphor is already being employed to draw parallel policies from the computer world into the world of synthetic biology, most notably the notions of open source.
- Cellular telephone technology (it’s a “mobile telephone;” interesting dichotomy of notions of location; of course, with landline phones, you know where you are; yet how many times have you been asked “Where are you?” when someone calls you on a landline phone? Loss of a notion of location, while, at the same time, the notion of location is now only knowable by the operator of the telephone; users are alienated from knowledge of location; which they have come to accept. Other breakdowns? Does GPS/E911 phone remediate this problem? Or just introduce another one – since now I can decide whether or not to reveal my location (maybe). Alternatively, the phone becomes a greater GPS – it can not only supply location, but it can also supply advertising/information that enhances the location information – closest pizza shop, closest hotel, closest bookstore with the lowest price)
- Radio frequency identification tags (RFID). Here, the developers of the technology insist that the RFID tag is nothing more or less than a barcode, just one that can be read from a distance. However, adopters of this technology are already demonstrating that the technology is distinct from barcodes in many respects (for example, the security concerns raised over the inclusion of RFIDs in passports, the decisions to embed these devices into living creatures (pets and humans) to make them “findable”) (Note that according to an email I got today from some telecomm folks, there is an expectation that RFID is going to be added to cellphones - expectation of 40+% with RFID by 2008 — see <http://www.itu.int/ITU-T/worksem/rfid/program.html>)

Automobiles and “Horseless Carriages”

One common example of a metaphor that effectively described a new technology’s characteristics based upon an old technology, but ultimately engendered confusion and negative assumptions, is the common reference to the new technology of the automobile as a “horseless carriage.” Upon its introduction in the United States at the end of the 19th century, automobiles entered cities such as New York that were, literally, covered in manure. It has been estimated that, prior to the advent of the automobile, horses in New York City produced 60,000 gallons of urine and 2.5 million pounds of manure every day. The bodies of dead horses (as many as 15,000 annually) often rotted in the streets before they were cleared away at city expense. Respiratory ailments were blamed on dried manure that blew about in the winds during the summer [Cowan, 1997].

Referring to an automobile as a horseless carriage served several purposes: it established a link with an existing, well understood technology; it established the automobile as a replacement for that technology; and it emphasized the absence of a component of the old technology that many identified with nuisance, filth and disease. Though automobile exhaust didn't smell or look clean, it was diffuse and did not linger, unlike the waste produced by horses. Especially in its early days, the automobile was thus viewed as a "clean" technology that could aid in solving many urban problems associated with horses, and its label as a horseless carriage was an effective way of communicating its potential.

Unfortunately, with the benefit of hindsight, it is clear that these associations with cleanliness were misguided. The popularity and accessibility of the automobile increased at rates no policy makers had anticipated, and automobiles became a fact of life in cities such as New York and elsewhere all over the nation before the magnitude of their impact — and distance from horses — was well understood. Framing the automobile as a horseless carriage may have thus encouraged a misguided way of conceptualizing the automobile as clean, preventing effective discussions of its own unique contributions to pollution at a time when they might have made a difference. The magnitude of the pollution issues associated with automobiles now is so vast that no policy solution that will make a positive difference in air quality and climate change could be enacted without serious economic and infrastructural suffering.

Appendix B — Digital Familiars OII White paper [Black *et al.*, 2005]

Personalized Digital Services: Power, Equity and Transparency in “Digital Familiars”

An exploration of the issues of personalization and customization

Jason Black, Kieran Downes, Frank Field, Aleksandra Mozdzanowska

(a working draft with links, etc. is at

<https://msl1.mit.edu/twiki/bin/view/Scratch/PersCustWorkingDraft3>)

ARIEL

All hail, great master! grave sir, hail! I come

To answer thy best pleasure; be't to fly,

To swim, to dive into the fire, to ride

On the curl'd clouds, to thy strong bidding task

Ariel and all his quality.

— *The Tempest*; Act I, Scene II; William Shakespeare

As computers and computerized services have become ubiquitous, there has been a concomitant increase in the mechanisms and modalities of personal interaction with these devices. However, the accessibility and understandability of the services being offered has continued to outstrip the public's grasp of the possibilities of these technologies.

One strategy that has been employed to ease the human-machine service interaction has been to shift the burden of understanding the operation and/or capabilities of a machine service away from the user and onto the programmer. By devising software interfaces that adaptively respond to signals from the user, the programmer can embed program facilities that can “learn,” identifying limitations in the user's appreciation of the operation or the features of the machine service being provided and attempting to anticipate the user's needs and wants without the user having to master the machine or system.

These efforts to create adaptive interfaces have gone through many iterations, and have seen varying degrees of success. As the computer has become more and more ubiquitous, these adaptive interfaces have become a part of the daily lives of the public.

One of the main streams of this kind of system design goes under the names of “personalization” and “customization,” largely distinguished from one another by the extent to which the user actively participates in shaping the performance and behavior of the software intermediary. While these tools are descendents of the original efforts to simplify the user's experience with complex software services, as their application has moved from the specialized to the mundane, a host of important issues has begun to emerge.

A brief historical exploration of a handful of the current exponents of this set of technologies can serve to illustrate the nature of these issues. Personalization's preeminent exemplar is Amazon.com, where a web-based software agent acts as the customer's personal shopper. Based upon the customer's purchasing history, and the purchasing history of all other Amazon.com shoppers, the software makes purchasing suggestions, leading the customer to what it expects are products that are likely to satisfy his wants. The user has the option to actively participate in “teaching” the software by ranking past purchases and commenting upon products purchased elsewhere, but the fundamental effort has been to mimic the behavior of an attentive salesman, who has a perfect memory and a firm grasp on the interests and buying habits of all of Amazon.com's customers.

At one level, there is nothing new about this kind of marketing, of course. While the field has gone through a host of incarnations, at its heart marketing has been about identifying what a customer has bought in the past, and drawing inferences about that to promote new purchasing options. Probably the key development in marketing has been the discovery that, sometimes, it is more effective to promote the seller/producer of the product than the product itself - leading to the rise of the notion of “brand” and its marketing. And the focus of that effort is the development of “brand trust” - the sense that, even if the customer does not know everything about the product offered, the fact that it is being offered by a company that the consumer has good feelings about is enough to close the sale.

Customization, where the user takes an active role in shaping the interaction with the machine service, has also moved into areas beyond software system interaction. For example, the delivery of news and other time-sensitive information is increasingly being offered through the agent of computer-driven site that sift the news data stream according to criteria set by the user. Whether delivered by pull (e.g., web-based) or push (e.g., email newsletters) technologies, the user can instruct the server, within a range of choices set by the programming, to classify and deliver information according to those classifications.

Again, in certain respects, nothing new here. News gathering and delivery strategies have evolved over hundreds of years, including coffee shops & taverns that specialized as collection points of certain kinds of information, diverse newspapers & news magazines focusing upon specific content or editorial positions, and various kinds of broadcast services whose content and delivery might commonly be shaped by the technologies employed and the strictures set by public policy.

The software industry has continued, of course, to work to refine the agents that act to improve the usability of their own products, seeking to increase the utility of (and, thus, the demand for) their products. Yet, it appears that there are limits to their abilities to accomplish this on the scale of the individual. The ubiquitous Microsoft Office family of tools is rife with instruments that actively seek to help the user. Auto-correction of typographic errors in Microsoft Word has been a boon to many - except when the software insists on making a change that the user does not want. Microsoft’s Office Assistant (“Clippy,” almost certainly an application of the ill-fated “Microsoft Bob” technology) has seen massive investment, only to be side-stepped owing to the virulent hatred that it has engendered in a sizable fraction of the population the tool was expected to service.

The potential of this family of technological developments is huge. As our tools for collecting, manipulating and acting upon information become more capable, they have also tended to become more complex, limiting their accessibility to those who might benefit most from their use. The introduction of the notion of a software-driven mediator between the user and the service that is programmed to adapt to the user’s strengths, limitations and revealed preferences is a clever approach to the problem, and proponents claim it has led to substantially more penetration of these services than might otherwise be expected. However, as these tools have migrated from the land of pure software and into the realm of information services more generally, there are reasons to inquire into how these “digital familiars” change the conventional into something less so.

Rather than “agent,” the notion of a “familiar” seems a more apt term in this context. As with Prospero’s Ariel, the familiar is an agent of the magician/user, and its talents in channeling magical forces helps the user to achieve his ends. However, the familiar does not slavishly hew to the user’s will, but is instead an independent agent whose motives are not necessarily aligned with the user’s.

For both of the non-software examples, sales/marketing and information/news gathering, the notion of “trust” emerges as a key element of the activity. In the case of sales, the goals of the sales agent is to develop a trust relationship with the consumer, so that she will be willing to make a purchase that will leave both the salesman and the consumer better off. In the case of news/information gathering, the issue of trust arises through the notion of the authority of the information source — can the information be “trusted?”

For both sales/marketing and news/information, society has seen the need to erect institutions to ensure equitable and transparent relationships in these transactions, particularly as the scale of the services has increased. Fair trade, truth in advertising, “fair and balanced” news — these are public policy goals that have emerged in the face of increasing concentrations of power on the side of the service providers, particularly as these interactions have moved from the personal to the corporate/industrial.

The “digital familiar” is presented to the consumer as an electronic servant. The attentive salesman, the easily-directed news clipping service and the host of complements being deployed today extend the abilities of the user/consumer in a host of exciting and innovative ways. But their introduction into daily life also raises a set of issues that, if addressed at all, is being handled without a complete consideration of their scope.

Fundamentally, the key issue is that, although the “digital familiar” poses as the servant of the consumer, the consumer is not the master of the familiar — either personalized or customized. While the familiar can mimic the development of a relationship, no such relationship is actually forming. Rather, the familiar is seeking to engender trust, without the reciprocal responsibilities that are a part of normal relationships.

As a consequence, the familiar is perfectly capable of sharing queries, information, analyses, and assumptions about the user that would be considered gross invasions and betrayals in the real world. Information divulged to a friend is constrained by the relationship, trust and social mores; the “digital familiar” may build upon those cultural assumptions, but it is not constrained by them. Moreover, the familiar, armed with an appreciation of the user that is developed through inhumanly attentive collection and analysis of user behavior, is positioned not merely to serve, but also to shape (if not control) the actions of a user whose whole relationship with the familiar founded upon his/her relative ignorance of the familiar’s true capabilities and inner workings.

Already the information collection, organization and reselling business has grown to a scale that has raised concerns among public interest groups and activists. As “digital familiars” become an ever-increasing part of the way in which users interact with the world around them, it is going to be vitally important that there be an exploration of the ways in which the gaps that these tools exploit can be filled, either through the expansion of current institutions of management and control, or the development of new ones. And, increasing attention will have to be paid to exposing the interests and enforcing the responsibilities of the men “behind the curtain” who are financing the programming of these familiars.

Otherwise, when the public finds out the extent to which the “servants” have been talking behind their backs to characters they deem unsavory, their reaction may lead to the crippling not only of this area of innovation, but many others in digital communications.

References

- Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Swayer and Stacy Wood. “Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces.” *Journal of Marketing*, vol. 61, pp. 38–53, July 1997.
- American Civil Liberties Union. “ACLU Unveils Disturbing New Revelations About MATRIX Surveillance Program.” Press release, May 20, 2004. Available online at <http://www.aclu.org//privacy/spying/15234prs20040520.html>. Last checked January 11, 2006.
- Anthes, Gary H. “Data: Lost, Stolen or Strayed.” *ComputerWorld*, August 1 2005. Available online at <http://www.computerworld.com/hardwaretopics/storage/story/0,10801,103541,00.html>. Last checked January 4, 2006.
- Associated Press. “Database Tagged 120,000 as Possible Terrorist Suspects.” 2004. Available online at <http://www.nytimes.com/2004/05/21/national/21database.html?ei=5007&en=041dfc5e74c6f507&ex=1400472000&adxnnl=1&partner=USERLAND&adxnnlx=1136986204-eM+wFhfxLP7tSP7dey0oDw>. Last checked January 11, 2006.
- Bainwol, Mitch. “Statement of Mitch Bainwol, Chairman and CEO, Recording Industry Association of America.” Testimony, Oversight Hearing, Content Protection in the Digital Age: The Broadcast Flag, High-Definition Radio, and the Analog Hole, November 3 2005. Available online at <http://judiciary.house.gov/oversight.aspx?ID=202>. Last checked February 17, 2006.
- Barrett, Neil. “Digital Familiars.” *BCS Review 1999 Online*, 1999. Available online at <http://www.bcs.org/BCS/review99/articles/emergingtechnology/digitalfamiliar.htm>. Last checked January 4, 2006.
- Berman, Jason S. “Testimony of Jason S. Berman, Chariman and Chief Executive Officer, Recording Industry Association of America.” Hearing on H. R. 1506 “The Digital Performance Right In Sound Recording Act of 1995”, June 21 1995. Available online at <http://judiciary.house.gov/legacy/489.htm>. Last checked February 21, 2006.

- Bezos, Jeffrey P., Maryam S. Mohit, Walker L. Lockhart, Ericka N. Lock, William W. Allocca and Suguna Subramaniam. “U.S. Patent #6,917,922: Contextual presentation of information about related orders during browsing of an electronic catalog.” USPTO Patent Full-Text and Image Database, 2005. July 12, 2005.
- Birner, Betty J. “Metaphor and the Reshaping of our Cognitive Fabric.” *Zygon Journal of Religion and Science*, vol. 30, no. 1, pp. 39–48, 2004.
- Black, Jason, Kieran Downes, Frank Field and Aleksandra Mozdzanowska. “Personalized Digital Services: Power, Equity and Transparency in ‘Digital Familiars’.”, April 2005. Unpublished white paper prepared for “New Approaches to Research on the Social Implications of Emerging Technologies;” Oxford Internet Institute, University of Oxford; April 15 and 16, 2005, Available online at <https://msl1.mit.edu/twiki/bin/view/Scratch/PersCustWorkingDraft3>. Last checked January 4, 2006.
- Cha, Ariana Eunjung. “In Retailing, Profiling for Profit.” *The Washington Post*, p. A01, August 16 2006. Available online at http://www.washingtonpost.com/wp-dyn/content/article/2005/08/16/AR2005081601906.html?nav=rss_business. Last checked January 19, 2006.
- Clarke, Roger A. “Information Technology and Dataveillance.” *Communications of the ACM*, vol. 31, no. 5, pp. 498–512, May 1988.
- Clarke, Roger A. “Internet Privacy Concerns Confirm the Case for Intervention.” *Communications of the ACM*, vol. 42, no. 2, pp. 60–67, February 1999.
- Cowan, Ruth Schwartz. *A Social History of American Technology*. Oxford University Press, New York, 1997.
- Craik, Kenneth. *The Nature of Explanation*. Cambridge University Press, Cambridge, UK, 1943.
- Darlington, David. “The Chemistry of a 90+ Wine.” *The New York Times Magazine*, August 7 2005. Available online at <http://www.nytimes.com/2005/08/07/magazine/07ENOLOGI.html?ex=1281067200&en=464de51d2ed0cd02&ei=5088&partner=rssnyt&emc=rss>. Last checked January 22, 2006.

Dash, Eric. “Europe Zips Lips, U.S. Sells ZIPs.” *The New York Times*, August 7 2005. Available online at <http://www.nytimes.com/2005/08/07/weekinreview/07dash.html?ex=1281067200&en=e502461cad24f6fe&ei=5088&partner=rssnyt&emc=rss>. Last checked January 27, 2006.

Espeland, Wendy and Mitchell L. Stevens. “Commensuration as a Social Process.” *American Review of Sociology*, vol. 24, 1998.

Ferber, Dan. “Microbes Made to Order.” *Science*, vol. 303, pp. 158–161, January 9 2004. Available online at <http://www.sciencemag.org/cgi/content/full/303/5655/158>. Last checked February 27, 2006.

Gentner, Dedre, Brian Bowdle, Phillip Wolff and Consuelo Boronat. “Metaphor Is Like Analogy.” In Dedre Gentner, Keith J. Holyoak and Boicho K. Kokinov, eds., *The Analogical Mind*, chap. 6, pp. 23–58. MIT Press, Cambridge, MA, March 2001.

Gillenson, Mark L., Daniel L. Sherrell and Lei-da Chen. “Information Technology as the Enabler of One-to-One Marketing.” *Communications of the Association for Information Systems*, vol. 2, September 1999. Available online at <http://portal.acm.org/citation.cfm?id=374468.374473>. Last checked December 16, 2005.

Hartman, Peri and John Gehlen. “Patent #6,907,315: Method and system for displaying and editing of information.” USPTO Patent Full-Text and Image Database, June 14 2005.

Johnson-Laird, Philip Nicholas. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press, Cambridge, MA, 1983.

Kang, Jerry. “Information Privacy in Cyberspace Transactions.” *Stanford Law Review*, vol. 50, p. 1193, 1998. Available online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=631723. Last checked January 27, 2006.

Karypis, George. “Evaluation of Item-Based Top-N Recommendation Algorithms.” In *10th Conference*

of Information and Knowledge Mangement (CIKM), pp. 247–254. 2001. Available online at <http://glaros.dtc.umn.edu/gkhome/node/124>. Last checked March 1, 2006.

Keenan, Faith, Stanley Holmes, Jay Greene and Roger O. Crockett. “A Mass Market of One.” *BusinessWeek Online*, December 2 2002. Available online at http://www.businessweek.com/magazine/content/02_48/b3810088.htm. Last checked December 16, 2005.

Krane, Jim. “FCC’s Powell declares TiVo ‘God’s Machine’.” Associated Press Breaking News, January 10 2003. Available online at <http://www.sfgate.com/cgi-bin/article.cgi?file=/news/archive/2003/01/10/financial11802EST0373.DTL>. Last checked January 25, 2006.

Kristol, D. and L. Montulli. “RFC 2109: HTTP State Management Mechanism.” Tech. rep., Internet Engineering Task Force, February 1997. Obsoleted by RFC 2965, Available online at <http://www.ietf.org/rfc/rfc2109.txt>. Last checked December 16, 2005.

Kristol, D. and L. Montulli. “RFC 2965: HTTP State Management Mechanism.” Tech. rep., Internet Engineering Task Force, October 2000. Available online at <http://www.ietf.org/rfc/rfc2965.txt>. Last checked December 16, 2005.

Krug, Steven. *Don’t Make Me Think: A Common Sense Approach to Web Usability*. New Riders Press, Berkeley, CA, 2 ed., 2005.

Lakoff, George and Mark Johnson. *Metaphors We Live By*. University of Chicago Press, Chicago, IL, 2 ed., April 15 2003.

Lewington, John, Leslie de Chernatony and Ann Brown. “Harnessing the Power of Database Marketing.” *Journal of Marketing Management*, vol. 12, pp. 329–346, 1996.

Linden, Gregory D., Brent R. Smith and Nida K. Zada. “U.S. Patent #6,912,505: Use of product viewing histories of users to identify related products.” USPTO Patent Full-Text and Image Database, June 28 2005.

Ling, Raymond and David C. Yen. “Customer Relationship Management: An Analysis Framework and

- Implementation Strategies.” *Journal of Computer Information Systems*, vol. 41, no. 3, pp. 82–97, Spring 2001.
- Lyon, David. “Everyday Surveillance: Personal data and social classifications.” *Information, Communication & Society*, vol. 5, no. 2, pp. 242–257, 2002.
- McMillen, David. “Privacy, Confidentiality, and Data Sharing.” A paper prepared for the Confidential Data Access for Research Purposes Workshop, October 16-17, 2003, October 2003. Available online at http://www7.nationalacademies.org/cnstat/Data_Access_Panel.html. Last checked February 28, 2006.
- Moray, N. “Mental models in theory and practice.” In D. Gopher and A. Koriat, eds., *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application. Attention and performance*, pp. 223–258. MIT Press, Cambridge, MA, 1999.
- Murthi, B. P. S. and Sumit Sarkar. “The Role of the Management Sciences in Research on Personalization.” Review of Marketing Science (ROMS) Working Paper, October 30 2002. Available at SSRN., Available online at <http://ssrn.com/abstract=407940>.
- Netscape Corporation. “Persistent Client State HTTP Cookies.” December 16 1999. Available online at http://wp.netscape.com/newsref/std/cookie_spec.html. Last checked December 16, 2005.
- Nielson, Jakob. *Designing Web Usability: The Practice of Simplicity*. New Riders Press, Indianapolis, IN, December 1999.
- O’Harrow, Jr., Robert. *No Place To Hide*. Free Press, New York, 2006.
- Peterson, Lisa A., Robert C. Blattberg and Wang Paul. “Database Marketing: Past, Present, and Future.” *Journal of Direct Marketing*, vol. 11, no. 4, pp. 109–125, Autumn 1997.
- Pub. L. 105-277. “Internet Tax Freedom Act.” Title XI (*Omnibus Appropriations Act of 1998*, originally S.442) (112 Stat. 2681-719), October 1998. Amended in Pub. L. 107-75; (115 Stat. 703) (HR 1552). Presently under consideration for permanent status, S.849, HR 1684, HR 1685.

- Rouse, W. B. and N. M. Morris. "On looking into the black box: Prospects and limits in the search for mental models." *Psychological Bulletin*, vol. 100, no. 3, pp. 349–363, 1986.
- Russo Dos Santos, C. and P. Gros. "Multiple Views in 3D Metaphoric Information Visualization." *YLEM Journal*, vol. 23, no. 10, pp. 4–7, September-October 2003. Available online at <http://www.ylem.org/Journal/2003Iss10vol123.pdf>. Last checked April 19, 2006.
- Schoenbachler, Denise D., Geoffrey L. Gordo, Dawn Foley and Linda Spellman. "Understanding consumer database marketing." *Journal of Consumer Marketing*, vol. 14, no. 1, pp. 5–19, 1997.
- Solove, Daniel J. *The Digital Person: Technology and Privacy in the Information Age*. New York University Press, New York, 2004.
- Sprenger, Polly. "Sun on Privacy: Get Over It." *Wired News Online*, January 26 1999. Available online at <http://www.wired.com/news/politics/0,1283,17538,00.html>. Last checked January 26, 2006.
- St-Cyr, Olivier. *An Ecological Approach to Mental Models: Towards Assessment Ecological Compatibility*. Master's of applied science thesis, University of Waterloo, Waterloo, Canada, 2002. Human Systems in Systems Design Engineering, Technical Report Series, HSSD-02-01.
- Tedeschi, Bob. "Small Internet Retailers Are Using Web Tools to Level the Selling Field." *The New York Times*, December 19 2005. Available online at <http://www.nytimes.com/2005/12/19/technology/19ecom.html?ex=1292648400&en=ceb94b2537c38cc5&ei=5090&partner=rssuserland&emc=rss>. Last checked February 10, 2006.
- Thompson, Richard Tony. "Metaphertigo, That Crazy Thing Metaphor Does to Language, Meaning, and the Human Equilibrium.", Spring 2002. Winner of the 2002 Arnold Fox Research Writing Award, Presented at the 2002 Arnold Fox Symposium, Northern Illinois University Department of Graduate Studies in English, Available online at <http://www.engl.niu.edu/rthompson/metaphertigo/>. Last checked March 6, 2006.

Turow, Joseph. “Have They Got a Deal For You: It’s Suspiciously Cozy In the Cybermarket.” *The Washington Post*, p. B01, June 19 2005. Available online at <http://www.washingtonpost.com/wp-dyn/content/article/2005/06/18/AR2005061800070.html>. Last checked January 26, 2006.

Turow, Joseph, Lauren Feldman and Kimberly Meltzer. “Open to Exploitation: American Shoppers Online and Offline.” Tech. rep., Annenberg Public Policy Center, University of Pennsylvania, Philadelphia, PA, 2005. Available online at http://www.annenbergpublicpolicycenter.org/04_info_society/Turow_APPC_Report_WEB_FINAL.pdf. Last checked December 16, 2005.

U.S. Dept. of Commerce. “A Nation Online: Entering the Broadband Age.” Tech. rep., U.S. Department of Commerce, September 2004. Available online at <http://www.ntia.doc.gov/reports/anol/index.html>. Last checked February 27, 2006.

Warren, Samuel D. and Louis D. Brandeis. “The Right to Privacy.” *Harvard Law Review*, vol. 4, no. 5, pp. 193–220, 1890. Available online at http://www.lawrence.edu/fast/boardmaw/Privacy-brand_warr2.html. Last checked January 27, 2006.

Weeks, Linton. “Amazon’s Vital Statistics Show How Books Stack Up.” *The Washington Post*, p. C01, August 30 2005. Available online at http://www.washingtonpost.com/wp-dyn/content/article/2005/08/29/AR2005082901873.html?nav=rss_technology. Last checked February 27, 2006.

Wehmayer, Kai. “Aligning IT and marketing — The impact of database marketing and CRM.” *Database Marketing & Consumer Strategy Management*, vol. 12, no. 3, pp. 243–256, April 2005.

West, Patricia M., Dan Ariely, Steve Bellman, Eric Bradlow, Joel Huber, Eric Johnson, John Kahn, Barbaraand Little and David Schkade. “Agents to the Rescue?” *Marketing Letters*, vol. 10, no. 3, pp. 285–300, 1999.

Wilby, Emma. “The Witch’s Familiar and the Fairy in Early Modern England and Scotland.” *Folklore*, vol. 111, no. 2, pp. 283–305, October 2000.

Winer, Russell C. “A Framework for Customer Relationship Management.” *California Management Review*, vol. 43, no. 4, pp. 89–105, Summer 2001.

Zeller, Jr., Tom. “Waking Up to Recurring ID Nightmares.” *The New York Times*, January 9 2006. Available online at <http://www.nytimes.com/2006/01/09/technology/09link.html?ex=1294462800&en=8c923f07c16f6d5c&ei=5090&partner=rssuserland&emc=rss>. Last checked January 27, 2006.

Zwick, Detlev and Nikhilesh Dholakia. “Whose Identity Is It Anyway? Consumer Representation in the Age of Database Marketing.” *Journal of Macromarketing*, vol. 24, no. 1, pp. 31–43, June 2004. Available online at <http://jmk.sagepub.com/cgi/content/refs/24/1/31>. Last checked January 30, 2006.