FACETED ID/ENTITY:

Managing representation in a digital world

by

danah boyd

A.B. Computer Science
Brown University
2001

Submitted to the
Program in Media Arts and Sciences
School of Architecture and Planning
in partial fulfillment of the requirements
for the degree of
Master of Science in Media Arts and Sciences
at the Massachusetts Institute of Technology
September 2002

Author
<div align="right">

Program in Media Arts and Sciences
August 9, 2002

</div>

Certified by
<div align="right">

Professor Judith Donath
Thesis Advisor
Program in Media Arts and Sciences

</div>

Certified by
<div align="right">

Dr. Andrew Lippman
Chair,
Departmental Committee on Graduate Students
Program in Media Arts and Sciences

</div>

FACETED ID/ENTITY:

Managing representation in a digital world

by

danah boyd

Submitted to the
Program in Media Arts and Sciences
School of Architecture and Planning
on August 9, 2002
in partial fulfillment of the requirements
for the degree of
Master of Science in Media Arts and Sciences

Abstract

In this thesis, i articulate a theory of how and why individuals use context to convey only a facet of their identity in social interactions. Through this lens, i discuss current issues in digital identity management. In this discussion, i focus on the role of design in affecting an individual's ability to maintain control of personal representation and identity information. I argue that the architecture of current digital environments has altered our notions of context, motivating users to develop new mechanisms for managing their presentation. I take the stance that users should have the ability to control their digital identity for the same reasons that they seek to control their physical identity, most notably to present themselves in an appropriate manner in relation to the current situation.

From this perspective, i argue for a design approach that will aid sociable designers in developing human-centered technologies that allow for individual control over personal identity. First, i argue the need for mechanisms of self-awareness and discuss what forms of awareness users should have. In doing so, i analyze current approaches to awareness and critique my own work on *Social Network Fragments*, a visualization tool for revealing the structure of one's digital social network. Alongside self-awareness, i present the need for identity management and critique my work on *SecureId*, a prototype intended to give users control over their digital presentation by offering security through identity-based knowledge.

This thesis argues for empowering users through awareness and control, so that they may provide the level of regulation that is desirable. In doing so, i offer a novel approach to context and identity management in digital social interaction.

Thesis Advisor: **Professor Judith Donath**
                             Program Media Arts and Sciences

FACETED ID/ENTITY:

Managing representation in a digital world

Thesis Committee

Thesis Reader
_____

Dr. Henry Jenkins
Professor of Literature and Comparative Media Studies
MIT Comparative Media Studies

Thesis Reader
_____

Dr. Genevieve Bell
Anthropologist
Intel Corporate Technology Group

for every lie i unlearn
i learn something new
i sing sometimes for the war that i fight
'cause every tool is a weapon
if you hold it right

(Ani DiFranco, *My I.Q.*)

THESIS READERS

Judith Donath is an Assistant Professor at the MIT Media Lab, where she directs the Sociable Media research group. Her work focuses on the social side of computing, synthesizing knowledge from fields such as graphic design, urban studies and cognitive science to build innovative interfaces for the online communities, virtual identities and computer-mediated collaborations that have emerged with the convergence of computing and communication.

Henry Jenkins is the Ann Feder Friedlaender Professor of Humanities and the Chair of the Comparative Media Studies Program at MIT. His research spans across the full range of contemporary and historical popular culture, but focuses primarily on the ways that everyday citizens draw on media to make sense of their lives, express their identities, and form communities. His column, "Digital Renaissance," runs monthly in Technology Review.

Genevieve Bell is a senior researcher within Intel's Corporate Technology Group. She is a member of Peoples and Practices Research, an interdisciplinary team of social scientists, designers and developers responsible for finding new users, new uses, and new user paradigms for technology. Her works focuses on the intersections of emerging technologies and social practice as they occur within domestic and public spaces. Bell also conducts ethnographic research in domestic spaces around the world. She holds a PhD in cultural anthropology from Stanford University.

ACKNOWLEDGEMENTS

Ideas only come as the product of interaction. I thrive on interacting with people, with communities and with society as a whole. By observing social behavior, i am motivated to think and analyze. The inspiration for this work comes from the most peculiar environments and unknowing people, not limited to my activities as a student. While i cannot even begin to thank all of those who have unknowingly served as my muse throughout the years, i would like to begin by acknowledging those who have actively provided support and sanity, and sometimes engaging drama. Without these people, the ideas contained in this thesis would never have come to light. In particular, i would like to thank:

Judith Donath, my advisor, for being my supporter over the last couple of years, for giving me the ability to investigate a wide variety of topics within the field of sociable media;

Henry Jenkins, for always challenging me to integrate my ideas in new fashions, for giving me the guidance to pursue this direction from a critical perspective;

Genevieve Bell, for showing me the value of social science, for encouraging me to follow my heart;

Colleen Capodilupo, Kelly Dobson, Andrew Fiore, Karrie Karahalios, Raffi Krikorian, Hyun-Yeul Lee, Matt Lee, Lisa Lieberson, Dana Spiegel, Derek Tang, Fernanda Viegas, my Sociable colleagues, for being other instigators of sociable ideas, for working with me to pursue these ideas from varied directions, for silly conversations and even sillier all-nighters;

Ronen Barzel, Ming-en Cho, Nelson Lai, Jeff Potter, Dan Ramage, for working with me over the last two years to implement some crazy ideas in meaningful ways;

Danyel Fisher, Dan Gould, Becky Hurwitz, Carlos Jensen, Scott Lederer, Mike Lewis, David Nguyen, Jeff Potter, for reading the large quantities of my written thoughts, listening to my babbles, providing feedback and being critical of my ideas;

Kathryn Boyd, Ryan Boyd, Rita Boyd, Richard Boyd, my family, for always being there to love me, for always reminding me that i am valuable, and for providing immeasurable support throughout my lifetime;

Andy van Dam, for continuing to be my father figure and my academic supporter in more ways that i can enumerate;

Ronen Barzel, Anindita Basu, Sascha Becker, Vered Birmaher, Gabriel Bos, Brian Bramlett, Tim Brennan, Tim Brooke, Ming-en Cho, Cassidy Curtis, Spot Draves, Ilana Frankel, Dan Gould, Becky Hurwitz, Wendy Ju, Steph Kawachi, Jess Lopez, Catie Lowder, Jon Moter, Vaughn Nuest, Karen Obel, Jeff Potter, Hannah Rohde, Andrew Shalit, Rose Shuman, Aaron Sorber, Jenn Steinfeld, Glenn Strauss, Ky Yokojima, Mike Zyda, my dear eclectic friends, for inspiring me during the last few years, for creating adventures and chaos, for supporting me when times were rough and entertaining me just cause it's fun;

Jess Collins, Sarah Davitt, Courtney Mallen, Emmet Quinn, my roommates, for putting up with all of the chaos and absurdity that a ridiculous grad student can provide;

Anitya/Changmian, and all of those in the psytrance/goa, cooperative living, queer/S/M, feminist, and activist scenes, for showing me the value of life, dance, spirit, and passion while providing me with an outlet for my energy and a home for my curiosity;

Intel, MIT Media Lab & Sponsors, for economic and structural support during these projects;

To all who have crossed my path during my lifetime, given me spiritual or intellectual guidance, appealed to my senses or my energy or my desire for adventures, or have otherwise challenged me to think in new ways, i thank you.

"guess there's something wrong with me / guess i don't fit in / no one wants to touch it / no one knows where to begin / i've got more than one membership / to more than one club / and i owe my life / to the people that i love" -- Ani DiFranco, *In or Out*

TABLE OF CONTENTS

Chapter 1:

INTRODUCTION

From its earliest days as a science fiction dream to its current commodified incarnation, the Internet has produced innumerable fantasies about a life free of physical and social constraints. Online society was to be utopian, prompting researchers and cyberanarchists alike to work towards this ideal. Unfortunately, as with all good dreams, we are reaching the moment of waking and becoming aware of the constraints of reality. Cyberspace is not our utopian fantasy; many of the social constraints that frame physical reality are quickly seeping into the digital realm.

Social interaction is a negotiation of identities between people in a given environment. One's identity is comprised of both a personal internal identity and a public social identity. As people engage socially, they project aspects of their internal identity into a social identity for others to perceive. Based on the situation, people only present a particular facet of their internal identity for consideration. Depending on their own need to self-monitor, an individual manages what is to be seen dependent on the environment, thereby creating a social performance where they offer different faces to convey different facets of their identity. The goal of such monitoring is to manage the impressions that others might perceive, to convey the appropriate information at the appropriate time.

In order to assess what is appropriate, people draw from situational and interpersonal contextual cues. By understanding the social implication of context cues and perceiving the reactions presented by others, an individual is given social feedback to adjust their behavior to fit the situation in the hopes of being perceived in the desired light. As people engage socially, they are continually drawing from their own experiences to

perceive others and the environment and presenting aspects of their identity that they deem appropriate to the situation. Yet, this negotiation occurs with little conscious effort.

Digital social interaction is not as simple. The underlying architecture of the digital environment does not provide the forms of feedback and context to which people have become accustomed. The lack of embodiment makes it difficult to present oneself and to perceive the presentation of others. As people operate through digital agents, they are forced to articulate their performance in new ways. Additionally, the contextual information that they draw from does not have the same implications online. Situational context can be collapsed with ease, thereby exposing an individual in an out-of-context manner. Unlike physical architecture, the digital equivalent is composed of bits, which have fundamentally different properties than atoms. The interface to the digital world is explicitly constructed and designed around a user's desires. As with any fundamental differences in architecture, there are resultant differences in paradigms of use, interpersonal expectations, and social norms. Performing online requires that people be aware of and adjust to these differences so as to achieve the same level of social proficiency that they have mastered offline.

In this thesis, i begin by expanding on these ideas – drawing on previous work to unpack the ways in which people negotiate social interaction, analyzing the underlying differences between the digital and physical architectures as they relate to sociability, and discussing what adjustments must be made to properly negotiate social interactions in a digital world. In this discussion, i bridge different theories of behavior and communication to offer a new approach for conceptualizing context and context management online.

Using these theoretical ideas as a foundation, i articulate what adjustments i feel are needed in order to provide users with a more sociable environment. In particular, i emphasize the need for self-awareness and identity management capabilities. By being aware of their behavior, individuals are able to monitor their own presentation. Likewise, by having the tools to control what aspects of their identity are presented, people can more appropriately organize their presentation. Awareness and control can provide some of the missing feedback that inhibits certain types of social interaction. My goal in this thesis is to reflect on the existing forms of social feedback and mechanisms by

which people engage socially, so as to offer suggestions for designers of digital system to more properly engage in human-centered development.

Assessing my perspective

I recognize that my biases frame the work in this thesis, particularly the perspective that i use to consider social interaction and regulation. As an academic, my notions of social interaction are entirely grounded in a Western, and primarily American perspective. As a researcher and system designer, i draw from a diverse set of disciplines, including sociology, psychology, cultural studies, queer theory, and computer graphics. While i am partially versed in all of these fields, i am by no means an expert in any of them. Yet, i come to this research as a technologist who is delving into the social sciences and as a long-time user of many of the sociable applications that are being discussed. In doing this work, my goal is to bridge the various disciplinary approaches as they relate to digital technology. As an activist in an American context, i value and seek to empower the individual, particularly those who are marginalized.

Issues of privacy and surveillance are embedded in my research. In handling these issues, i value the individual over corporations and governments and seek a privacy approach that makes data transparent to and controlled by their subject. In other words, i believe that an individual has complete rights to their own data and their presentation.

Although i take a performative approach to identity presentation, i do not believe that the individual is inherently fragmented. Instead, i see the modern individual as aware of and reacting to the diverse social climate that we are embedded within. In such an environment, i see the individual as managing multiple facets of their identity. I see social regulation as an operational force in social behavior, where the individual chooses how they react to such reactions. Although i do not believe that the digital world shall be the utopian space in which people can rid themselves of their prejudices, i do believe that it provides a novel social environment that allows people to interact in new ways. At the same time, i see the digital as limiting because of its architectural constraints. Thus, i am inclined to suggest structural and design adjustments to more adequately provide people with the level of expectations that they have developed in physical encounters.

I believe that the role of sociable designers should be to engage and empower users by developing human-centered applications. I believe in working with the needs and

expectations of all users, particularly those who are marginalized in physical interactions. Rather than relying on market or legal forces for regulation, i believe in constructing an architectural environment that provides users with the necessary information to regulate primarily through social norms. I believe that users should have the ability to manage and present themselves as they deem appropriate while simultaneously maintaining control over all of their digital expression and content.

Designing social applications requires a fundamental understanding of both social interaction as well as the underlying architecture of digital environments. To understand social interaction requires a deeper understanding of how people perceive themselves and others and what motivates them to interact in particular ways. The digital world is explicitly structured and constructed to meet the needs of its inhabitants. The architecture provides different resources for users, not all of which resemble physical possibilities. Users recognize the digital as a place for social interaction and thus seek to engage socially, often bringing their own assumptions about what the underlying structure provides. While some may argue that the digital architecture should not focus on engaging users socially, i believe that this is a very desirable and valuable application, as it provides a new form of sociability across time and space. It is with this desire in mind that i believe that sociable designers should not only understand what architectural possibilities exist, but have an understanding of how they impact social behavior. With that understanding, they can design a space explicitly intended for social interaction.

It is my belief that sociable designers should focus on developing this understanding in order to empower individuals by designing appropriate interfaces. I am aware that my approach in doing so comes from a Western academic approach and an American understanding of digital life. Thus, my design approach focuses on the needs and interests of the Western world, not because i believe that these issues are not applicable elsewhere, but because i cannot dutifully address them.

Goal and purpose

This thesis focuses on a particular aspect of underlying sociability issues, addressing the role of context, self-awareness, and identity management in social interaction. In doing so, i highlight how the architectural differences of the physical and digital realms affect perception and social behavior. With my biases known, my goal is to assess how

sociable designers can present awareness to create a digital environment that more adequately gives users control over their social interactions.

   This thesis offers three novel contributions to this area of research. After grounding the discussion in different notions of social interaction, i articulate a new theory of how **context** operates in the digital realm, focusing on how the underlying architectural differences require a new set of considerations. In particular, i tackle the problems that occur when situational contextual information is collapsed and how users reclaim this. As both contextual feedback and self-awareness are necessary for those seeking appropriate social presentation, i discuss the importance of **self-awareness** within the digital realm. In doing so, i discuss current approaches to self-awareness and offer a sample design approach for providing self-awareness. Contextual understanding and personal self-awareness are the building blocks that people use to properly control their identity and presentation during social interactions. Next, i focus on the relevance of **identity management** in giving people control over their social interactions. In this section of the thesis, i discuss current mechanisms for management and control suggest a theoretical framework for conceptualizing these issues. Finally, i introduce and critique two sample applications intended to test my theories. First, i discuss **Social Network Fragments**, a tool designed to reveal the social network structure that emerges in one's email interactions. Following this, i analyze **SecureId** as a prototype tool for identity management.

   The purpose of this thesis is to delineate important issues that sociable designers should consider when they develop structures intended to encourage social interaction. In doing so, i address both theoretical and computational contributions to this area of research. While i sketch a conceptual model for addressing these issues, the prototypes that are discussed reveal the challenges that we, as researchers and designers, must face. Rather than providing solutions, they expose the weaknesses in this area of research and suggest paths for future research. Explicitly structuring a system for social interaction requires overcoming many obstacles, as each new interface presents new confounding social effects. Attempts to mimic the physical world are flawed because the underlying structure is so different. Yet, to determine how to break from those assumptions and provide users with the necessary information requires far more than an understanding of social behavior. Thus, this thesis only provides the first level of information that is necessary to enrich the social atmosphere of the digital realm.

Motivation for aiding social interaction

From its inception, a primary use of the Internet has been to engage people in social interaction. From email to Usenet to instant messenger, some of the most popular applications have focused on building community and aiding in communication. Yet, while these systems are quite popular, the architecture also restricts the types of social interactions possible online. Online mediums are quite valuable for quickly sharing data, but they prove limited in providing the support necessary for building community. Community requires trust, yet building trust online requires understanding how trust is built and designing systems accordingly (Bos, et. al 2002; Rocco, et. al 2000). At the same time, notions of trust and privacy are not universal, which is problematic both for designers as well as participants. As people do not maintain the same notions of trust, the system must provide for negotiated ideas about trust and privacy.

Trust is complicated by the lack of consistent communication techniques and expectations. As Saville-Troike (1982) noted in reference to physical interactions, differing notions of communicative competence create misunderstandings; this also applies online, although the likelihood of differing communicative expectations is greater. Messages are often misinterpreted, resulting in flame wars or otherwise unnecessary arguments. Yet, assessing a situation for communicative norms requires the ability to determine interpersonal context; online, other people are difficult to see. Just as presence is difficult to ascertain, so are the cues that people tend to embody. Context takes on a different role in the digital realm, as does presentation of one's identity. Lack of embodiment and feedback about the situation and people make the digital world fundamentally different.

These differences should be embraced and appreciated. Although access is not universal, the digital era allows certain groups of people to connect across vast distances in unprecedented ways. Information can be accessed with ease and spread rapidly around the globe. Anyone online can publish their thoughts in a public space and connect with people who have similar thoughts. While the digital opportunities are invaluable, understanding and working with the differences to provide users with a more sociable space can only enhance the possibilities of the digital realm.

My goal is not to replicate physical social interactions, but to learn from them to ascertain what people need and want in social environments. Through experience with

the physical world, people have come to understand how they can operate their bodies to convey thoughts. The ease with which people present themselves comes through regular interactions. By understanding the fundamental structures that people use to engage naturally, sociable designers can build systems that provide different, but equally comfortable environments for social interaction. At the same time, while marginalized individuals are limited in what they can convey and how they may convey it offline, they have a new level of freedom online to present themselves without the implications of their bodies automatically associated with their presentation.

I believe that many people are hesitant to join digital communities because of problematic social norms and other fears of privacy. As social norms do not operate as a regulatory force, there are many incidents of people abusing the freedoms that the digital world provides. For example, as discussed in "A Rape In Cyberspace," one individual chose to use his account to harass others, resulting in collective aggravation without a real mechanism for stopping the behavior (Dibbell 1993). Without socially normative regulation or effective feedback channels, the digital environment makes people feel unsafe and powerless.

Rather than requiring that users accommodate for the current interface designs, i believe that designers should assess what people want in their social environments. Current interfaces only address a limited segment of the population and those designs make it difficult for people to maintain properly segmented lives online. In this thesis, i discuss what people are seeking when they go online and propose suggestions for designing such systems, focusing on empowering users through design.

Thesis structure

In order to develop the framework from which sociable designers, including myself, can operate, i begin this thesis by analyzing social interaction. Drawing from various social science approaches, Chapter 2 discusses topics such as the role of one's personality in social interactions. In particular, i take a Goffman-esque approach to discuss the mechanisms by which people perform and negotiate identity, relating this to the relevance of context in determining appropriate forms of interaction, and the ways in which we construct ourselves and others based on that information, our roles, and the facets of our identity. I address theories of self-monitoring, using this to discuss how people differ in the ways in which they negotiate social environments. In this chapter, i

focus on physical interactions, using theories that speak almost exclusively to face-to-face interaction. This chapter establishes the framework for this thesis, grounding social behavior in a theoretical discussion. By integrating a diverse set of concepts, the discussion gives the reader the basic background for considering the issues surrounding social interaction in the digital world.

While Chapter 2 is focused on social behavior in the physical world, Chapter 3 explores some of the ways social behavior is altered in the transition to digital interactions. After analyzing some underlying structural differences, this chapter examines how changes in the notion of context affect different aspects of social behavior, from performance to regulation. While situational and interpersonal contextual information is usually available during physical interactions, online it is often missing, misleading or collapsed. Because of its impact on the many arenas of social interaction, contextual differences are crucial for understanding social shortcomings of the digital environment. Thus, i provide an extended example of collapsed contexts and discuss how people attempt to manage contextual information locally. Chapter 3 also addresses the problems that emerge when bodies are no longer the agent through which people negotiate their interactions. Ultimately, Chapter 3 deconstructs the architectural differences in order to address what is missing when context and embodiment are altered.

In order to provide users with appropriate mechanisms for presenting themselves, i propose a two-tiered approach. First, users should have tools to be aware of themselves and others. Second, they should have tools to manage their identity and presentation.

Given this perspective, Chapter 4 introduces digital self-awareness tools. By presenting a selection of current tools, i discuss the motivations of designers in providing users with necessary feedback mechanisms, including visualization tools and data collection systems. From the perspective of empowerment, i critique this area of work and suggest desired improvements. After discussing what types of awareness people desire, i offer an example scenario and tool called *Digital Mirror* that is intended to provide interactive digital reflection.

Chapter 5 tackles the issues surrounding identity management, again presenting currently existing tools as well as discussing apparent needs. Here, i critique current management systems, most notably Microsoft's *Passport*. I also discuss why management is necessary for social regulation and articulate what is necessary for users to acquire

control over their identity presentation. I suggest a set of design standards that sociable designers should consider in order to more adequately meet the needs of multi-faceted individuals.

Following this, i switch modes to analyze the sample prototypes that i helped design to test these theoretical concepts. Chapter 6 introduces the applications section of this thesis, while Chapters 7 and 8 detail *Social Network Fragments* and *SecureId*, including the conceptual theory, the algorithms and the design approach.

Chapter 7 reflects on the design and concepts behind *Social Network Fragments*, a visualization tool that i built in collaboration with Jeff Potter. Beginning with an introductory background to social networks, i introduce the motivation behind this awareness tool, reflecting on the importance of social networks in understanding oneself as a multi-faceted individual. From here, i introduce the relevant algorithms and the design of the system so that the curious reader may understand the application. By analyzing the images produced from a sample dataset, i critique the application as a tool for awareness and discuss the issues unveiled in the process of developing the system. Specifically, Chapter 7 critiques the tool from a design perspective, analyzing the problems that arise when conveying highly dimensional data on a visual plane.

By focusing on the design issues that arose in developing *SecureId*, Chapter 8 analyzes the issues in developing a tool for identity management. First, i discuss conceptual aspects of the system, such as knowledge-based security of identity facets. Following this, i analyze *SecureId* through a series of images drawn from the prototype. In developing this prototype, i was able to reflect on the amount of work necessary to make the theoretical ideals of Chapter 5 a reality. Thus, the majority of this chapter exposes the problems that i encountered as i set out to design a tool for identity management.

Finally, in Chapter 9, i integrate these ideas, discussing the users' need to have appropriate cues for social interaction, the impact of the digital architecture, and offering an approach for designers that includes giving users self-awareness and management tools. My goal is to motivate designers to focus on designing systems that empower users, as this would only create more desirable sociable environments. While this document articulates much of the conceptual work that must be considered, the applications are embryonic. Thus, throughout this thesis, and most notably in Chapter 9, i argue for further research.

Chapter 2:

NEGOTIATING IDENTITY IN SOCIAL INTERACTIONS

During social interaction, people regularly present themselves while simultaneously reading the presentations of others. Depending on one's personality, an individual will adjust aspects of their presentation according to the reactions and presentations of those around them. Fundamentally, social interaction is a negotiation between individuals performing within a particular social context to convey aspects of their identity. This negotiation often occurs with little conscious thought; people comfortably interact with one another, revealing what is appropriate while assessing what information is being given. Although these interactions happen at an unconscious level, it is important to understand exactly what is happening, particularly since the goal of this research is to create digital systems that give equivalent social structures for sociable people.

In this chapter, i articulate some of the underlying motivations and actions that occur as people interact, focusing on face-to-face communication. In particular, i emphasize a multi-faceted approach to *identity*, Goffman's notions of *performance/perception*, and the importance of and mechanisms for *context awareness* and *regulation*. I have chosen to explicitly consider these four aspects of social interaction for their relevance to the design of digital spaces. These characteristics are affected by the underlying architecture; thus, i feel as though their subtleties require closer examination. In discussing them, i also relate psychological notions of *self-monitoring*, postmodern concepts of the *fragmentation of self* and the relevance of fashion. Although these concepts operate in tandem, by teasing them apart, i hope to more adequately prepare the reader for understanding the impact of digital architectures on social behavior.

Multiple notions of identity: the internal vs. the social

Self-awareness allows individuals to have a sense of who they are in relation to society and culture. By reflexively adjusting one's perception of self in reaction to society, people construct their individual identity. Approaches to identity abound, and they refer to many different ideas about the self, much of which is grounded in contemporary Western cultural values. Frequently, identity refers to at least two different aspects of the individual – that which is an internalized notion of the self, and that which is the projected version of one's internalized self. Researchers have constructed this distinction in various ways. Adam Smith (1976/1790) separates identity into the object versus acting self, while Mead (1934) refers to me versus I. Most controversially, Freud (1974/1923) distinguishes between a public ego, an internal selfish id, and an internal conscience or super-ego. While these approaches are vastly different, they all recognize that the self is complicated, in part because of a separation between internal notions and external ones. In other words, what people produce or convey to others is not necessarily the same as their internal perception of self. Lacan (1980/1968) presents an alternative to this approach, suggesting that there is no internal self, only an external one. As i disagree with this analysis, my approach will consider a duality of identity, where i collapse competing notions of the self into two categories – one's *internal identity* and one's *social identity*.

To clarify, my notion of *internal identity* refers to an individual's self-perception in relation to their experiences and the world. As it is reflective in nature, self-perception cannot be purely manifested internally. Without society and experience as a basis for reflexivity, there can be no internalized evaluation (Giddens 1991: 52-53). As such, history, experience and interaction provide the model by which individuals can give meaning to the physical, psychological, philosophical, and moral aspects of their identity. One's identity is not simply based on the characteristics that are written on the body or the circumstances in which one is born, but on how the individual reacts to and internalized these experiences.

Alternatively, when people interact with others, they convey aspects of themselves through a set of signals that others must learn to read and evaluate. As will be discussed in more detail in the next section, the negotiation between self-presentation and external evaluation can be viewed as a performance, which helps construct an individual's social identity. While internal identity is entirely constructed and maintained by the

individual, *social identity* is perceived externally, relying not on the intention, but the effective expression and perception of an individual's presentation. While one's social identity emerges from one's internal identity, its manifestation is r ead in light of body conveying it and the situation in which it is being conveyed. The environment plays a crucial role in the production and perception of one's social identity.

These two formations of the self do not operate alone; instead, the social identity and the internal identity are in constant interplay. The public version of one's self is impacted by the internalized version, which in turn evolves based on one's experiences. The more that an experience challenges an individual's notion of self in relation to society, the more it impacts their identity. People notice who they are in relation to the people around them, particularly noting that which is different. It is because of this that people are quite conscious of their position in relation to societal norms.

The social identity is what individuals use to interact with and relate to others. Yet, it is the internal identity that one is constantly comparing to others' in a social environment. In order to socialize, people take specific aspects of their internal identity, project it into their social identity and use this to construct a performance that will allow them to negotiate social situations.

Performing and monitoring one's social presentation

While interacting socially, people are aw are of and react to the feedback that they receive by the other people in an environment. They adjust their body posture, their facial expressions, and their general presentation. These adjustments are made not to be artificial but to convey appropriate social information for the situation. As articulated best by Goffman (1956), all social interactions can be seen as a series of interactive performances, where the actors are constantly altering their presentation based on their assumptions about what is acceptable in this situation and the reactions that they receive from others. People perform aspects of themselves in order to generate specific impressions, often so that others will perceive them in a positive light.

Furthermore, people not only perform their ideas, but all aspects of themselves. For example, while sex may be a biological trait, Butler (1990: 25) suggests that people perform their gender. Read in tandem with one's perceived sex, one's gender performance is used to create assumptions about their sexuality, their values and their

personality. For example, the notion of a butch woman is derived from a masculine performance coming from a female body. From this perspective, gender and other identity concepts are entirely constructed; normative ideas vary across cultures. What an individual presents is read in response to the cultural norms and reflects on the identity of the individual in a given context.

Drawing from Goffman's performance theory, there are three fundamental components to the passage of social information between individuals. When information is to be conveyed explicitly, it is *given*, but these messages are also impacted by the subtle, and perhaps unconscious messages that are *given off* by the actors, as well as the intention that the observer might *infer* (Goffman 1956: 2). Thus, any social message is not simply a set of factual data, but a negotiation in communication relying on both the signals presented by the actor as well as the signs perceived by the observer. The observer's impressions of a situation are based on inference, which results from mental models derived from previous interactions. As such, a viewer does not always perceive the intentions of an actor.

While interaction operates on impressions, people are often naturally (or neurotically) motivated to suppress their own desires in order to please others (Rank 1932; Moustakas 1972). In other words, they seek to create a good impression. Social conformity, or collective action, relies on this behavior. The internal need to conform and the fear of perceived social gatekeepers creates a mechanism for society to be regulated by social norms. Yet, while there is a general desire to follow the social order, individual personality characteristics determine how important and relevant conformity is.

In his theories of *self-monitoring*, Snyder (1974) suggests that personality determines the level at which people regulate their performance in relation to others' reactions. As such, an individual's reaction to socially normative pressures is dependent on where they are situated along an axis of self-monitoring. High self-monitors are highly attuned to the expectations and reactions of others, and are therefore extremely conscious of presenting themselves in a way that creates the desired impressions, either positive or negative. Conversely, low self-monitors fail to incorporate social feedback when constructing their presentation. Self-monitoring is important for considering how people negotiate their identity in social situations.

Additionally, people's previous experiences affect their perception. When reading an actor's performance, the observer is constantly integrating the portrayed information with all previous knowledge, experience, and relevant communicative situations (Saville-Troike 1982: 22). In evaluating an actor's presentation, people categorize and stereotype the interactions in order to position the actor within their mental model of human behavior (Simmel 1971: 9-10). While categorization provides an observer with a mechanism to quickly understand the information that they are being given, it also makes it difficult for an individual to overcome their initial impressions. Unfortunately, people are more likely to reinterpret future presentations to fit their early mental models then they are to adjust their initial classification of others (Aronson 1995). Recognizing this, people are motivated to make that first impression count.

When developing a presentation to create a desired impression, people assess what is appropriate and expected, while trying to determine how their presentation is going to be perceived. In other words, people constantly adjust for *context*.

The value of situational and interpersonal context

With little conscious effort, people assess the interrelated conditions of the environment in which they are presenting themselves. Contextual information provides performers with vital cues with which to determine what is appropriate behavior in a particular situation. Likewise, context provides readers with a model for evaluating one's behavior. In particular, two context cues provide the majority of the information that people actively integrate – *situational* and *interpersonal* context information.

Situational context refers to the aspects of the architecture and environment that suggest what activities normally take place *here* and *now*. Situational knowledge requires an understanding of the social qualities of the environment including the location, the time period, the particular occasion, and the general politics and values of the society. Based on previous experiences in a given context, people start developing mental models of these situations, just as they build mental models of people. These models allow people to associate particular architectural forms with functions and behaviors, allowing people to more rapidly process the situation. People have learned to understand particular design forms and they can quickly separate a fast food restaurant from a pub. Likewise, they understand the meaning of specific situations, thereby realizing that a solemn funeral is an inappropriate place to scream the latest football scores.

In addition to situational cues, people adjust for *interpersonal* context information. When an individual enters a room, they reflect on the others in that space. Even without conversing, people evaluate each other's performances, develop mental categories and get a sense of the people in relation to the space. In such situations, people recognize that they are being observed as well as observing and thus present themselves to be read. Interpersonal contextual information allows the observer to determine what are the appropriate roles in this environment, what types of social identities are acceptable and whether or not they will have anything in common with the other people. Not only does one evaluate the *type* of people around, but also each individual's presentation.

When assessing situational and interpersonal contexts, people also evaluate the level of porousness. In other words, what is the likelihood that the information presented in this situation to these people will reappear elsewhere? When unexpected recording devices or gossip replicates one's performance in an external context, there can b e significant social consequences. Thus, one must evaluate the likelihood that recording devices exist or the probability of information being spread by word of mouth. In some situations, this is perfectly acceptable, if not desired. Yet, even in public environments, porousness is not typically assumed. For example, when one presents oneself at a pub, most likely they do not expect that their presentation will reappear at work to be considered out-of-context. When evaluating for potential gossip, people also evaluate the trust of others. In environments where information is not to be spread, trust of those present is necessary.

When assessing contextual information, people rely on previous experiences and categorization. They compare the current environment to their mental model to determine what assumptions can be made. While these assumptions may be inaccurate, they provide the necessary framework for people to quickly determine how to best present themselves. By understanding the context of the environment, people know which aspects of their social identity to perform.

Reconsidering identity in relation to fragmentation, facets and faces

As previously discussed, there are two components to an individual's identity – the internal and the social. This social component is constantly being adjusted depending on the context of a particular environment. People present themselves differently in particular situations, not because they are hiding aspects of themselves, but because

some behaviors are more appropriate in one context than another. A working mother does not act like a mother in a boardroom meeting; the language that one uses at a pub is not appropriate for church; while leather skirts meant one thing in 1985, they mean something very different in 2002. Based on contextual cues, an individual determines what is acceptable behavior and what aspects of their identity they should perform.

   Because a variety of contexts affect individuals differently, one's social identity appears to regularly change in relation to the social situation. As such, an individual may appear to have many different and conflicting social identities. This realization appears to be philosophically contradictory to the humanist notions of a complete, manageable "Cartesian" self (Descartes 1641). Starting with Freud's divergent opinion (1974/1923), postmodern theorists began to think of the self as incorrigibly fragmented:

> We can no longer conceive of the 'individual' in terms of a whole, centered, stable and completed Ego or autonomous, rational 'self'. The 'self' is conceptualized as more fragmented and incomplete, composed of multiple 'selves' or identities in relation to the different social worlds we inhabit, something with a history, 'produced', in process. The 'subject' is differently placed or positioned by different discourses and practices. (Hall 1996: 226)

   Seeing the unconscious as a product of culture, not individuality, Lacan (1980/1968) suggests that the self is the product of imagination. Thus, their presentations reflect *multiple subject positions*, where people can be viewed an aspect of the *text* of a given situation; the subject is not separated from the situation. Given this take on the individual, it is not surprising that postmodern theorists view the modern individual as undergoing an identity crisis (Harvey 1990).

   Such an approach appropriately reacts to the needs of the contemporary individual to lead a *plurality of lifeworlds* (Giddens 1991: 83), where they must negotiate diverse social situations, each of which has its own norms and values. Yet, these theories fail to recognize the agency of the individual to separate their internal and social identities, fragmenting only the latter without creating a crisis for the former. Suggesting that an individual is inherently fragmented and undergoing an identity crises is problematic. In a society where people play many different roles and must constantly adjust for different social contexts, their presentation may appear to be fragmented, but this does not imply that they are. Instead, such adjustments suggest that the individual is maintaining and presenting multiple *facets* of their identity as appropriate.

In any given situation, an individual presents a *face* (Goffman 1972), which is the social presentation of one *facet* of their identity. I believe that an individual has a coherent sense of self, but in presenting only facets of their identity, they are perceived as fragmented. People maintain many different social facets and often associate particular facets, and therefore faces, with particular contexts. These multitudes of faces and facets do not indicate a collapse of the individual, but instead represent the control with which an individual manages their presentation in everyday life. With little consciousness, people quickly evaluate the context of a given situation, determine which facet of their identity they wish to convey, and construct a face from which to perform this identity.

Thus, in managing multiple facets, people are simply fragmenting their social identities. This form of fragmentation is not necessarily problematic, although it does require more flexibility in identity management. As Simmel (1971: xliii) recognizes, social fragmentation can be liberating because it allows for individuality, where people have the ability to portray a wide variety of the different aspects of themselves in different, yet appropriate situations. Maintaining multiple facets can offer relief and empowerment for marginalized individuals, as they can find acceptance and support in certain communities while being shunned by society as a whole.

As people negotiate multiple facets, they unconsciously associate different facets of their identity with particular contexts. For example, one may maintain a work-based facet that only appears when one enters the workspace. Such archetypes aid users in properly negotiating their presentation, knowing which facet to show given a situation.

Fashion as an example of the convergence of identity, context and facets

In post-industrial Western fashion, people are able to choose clothing to represent aspects of their identity and their relationship with the culture. The meaning of the fashion is contextually dependent such that the place, the time, and the viewers all determine the meaning of one's fashion presentation (Davis 1992: 5). Fashion indicates ones societal role and participation within various (sub)cultures. By evaluating fashion markers in relation to the context, the viewer makes assumptions about the performer's identity. Likewise, the performer often chooses clothing to represent the facet of their identity that is being performed. Thus, one might choose to wear different clothes to work and the pub. By understanding the situational context and the subtleties of fashion markers, performers can present subversive messages, such as retro. Without the subtle

nuances, the same articles of clothing are seen as outdated. Fashion is constantly changing and the values associated with it are different across different (sub)cultures. The ways in which people choose to adorn themselves indicates much about their values, their interests and their relationship to fashion.

Conflicting contextual cues; collapsed contexts

Situations that present conflicting, misleading or inaccurate contextual cues can be disconcerting. For example, it is embarrassing to arrive at a formal cocktail party in a risqué costume having understood the invite to be for a masquerade ball. Misunderstood contextual cues can lead individuals to present inappropriate faces, thereby giving off the wrong impression. When an individual wants to contextualize their presentation, such experiences can be perplexing. This is particularly true when segregated contexts are collapsed.

When an individual is placed into a social situation where they relate to different people through different roles, they must reassess what is an appropriate face to present. Situations where multiple contexts collide encourage individuals to react in one of two ways – either aim to present a face that is universally acceptable or risk the social consequences of conveying inappropriate information to some of those observing your presentation. While people seek to present themselves appropriately, they do not necessarily have control over what others reveal about their identity. When two worlds are bridged, information that may have been shared in one context can be shared in the other, potentially creating an awkward social situation. For example, introducing mom to all of one's friends can be a recipe for disaster.

In order to avoid such discomfort, when individuals maintain separate identity facets, they tend to segregate the associated contexts so that there is no collision of identity information. Individuals who present the same identity information across multiple contexts tend to be less concerned with explicitly separating their social contexts. Again, one's self-monitoring habits indicate the importance of maintaining separate contexts (Kilduff 1992). Given their intense focus on socially approval, high self-monitors might be more likely to separate facets along contextual lines and are probably more fearful of the social embarrassment of collapsed contexts.

Maintaining segregated social facets is advantageous for those whose identity strays from the norm, as it allows them to associate with other societal outcasts while still being able to maintain a public life. For example, many sexual minorities tend not to present their sexuality in every social occasion; thus, they are more inclined to separate contexts where this is shared from those where it is not. As society tends to assume normative viewpoints unless shown otherwise, there is a certain level of safety in socially driven "don't ask; don't tell" policies, yet such politics also weaken the power of marginalized individuals through obscurity.

In maintaining and adjusting their identity, people tend to be cognizant of their social surroundings. People control social presentations to meet their needs, including the desire for privacy, perceived social acceptability, fear of disgrace or harm, or perhaps an internal need to control different aspects of one's life through separation.

Crowd behavior and social regulation

Social regulation is effective when people feel the need to conform to social norms. Through fear of disapproval, social sanctions or other consequences, people will self-regulate their own behavior. While social pressure operates in almost any type of social interaction, its impact on crowd behavior takes on an entirely different form. The norms of a crowd are quite different than the average of the individuals' values, as they are fundamentally impacted by the opportunity for anonymity and deindividualization (Le Bon 1952/1985). At the same time, the collective pressure to conform in crowds is dramatically increased. By asserting one's individuality, one is no longer a part of the crowd. As the power of the crowd is quite effective, such nonconformity puts the individual at greater risk.

Just as the crowd alters the mental state of an individual, so does perceived authority. As Milgram (1974) showed in his seminal work, people will complete otherwise unthinkable tasks simply because of social pressure and fear of punishment from an authority. Both crowd behavior and obedience to authority indicate the magnitude of social regulatory forces. As people avoid social embarrassment, they are quite likely to behave according to the social norms laid out by the collective.

In order for these forces to function, certain social structures must be operational. First, people must be able to observe or otherwise understand the socially acceptable

behaviors. Second, others must be able to observe when an individual is acting out of line and signal their disapproval or suggest punitive possibilities. Third, there must be a mechanism by which people can publicly admonish an individual in an environment where people dread the effects of the potential punishment. Minsky argues that people need to have a sense of the other individual, of their existence, since "without the concept of an individual, we could have no sense of responsibility" (1985: 51). Thus, in anonymous situations, people's lack of fear of retribution or sense of other people undermines the effectiveness of social regulation.

In the crowd environment, it is not a sense of the individual that matters, but the sense of the group as a substitute for the individual. As such, it is more apparent to an individual that they will be punished for acting against the crowd than for acting with the crowd against a broader social norm. Thus, they are more likely to go along with the crowd, as individuality is what is punished in such an environment.

Social regulation helps create the norms that people use when they are determining how to properly act. By creating a set of social standards, regulation helps people properly assess the context of a situation. Social regulation also acts as a motivating force for people to perform their identity in a meaningful manner. Without the social pressures of inappropriateness, it is difficult for people to evaluate others and adjust their performance according to the social values, context and perception of others.

Concluding thoughts

While social interaction requires little conscious effort, there are complex processes continually at play. People must process a situation, read the contextual cues, present their internal sense of self in a meaningful way, adjust their presentation depending on others' reactions, and constantly negotiate what is socially acceptable. In all interactions, identity, performance, context and regulation are constantly operating and interacting.

Although understanding these behaviors may appear to be a futile academic exercise, it is necessary for designing digital environments. While these processes occur unconsciously in the physical world, the underlying structure that motivates them is drastically altered by the digital architecture. Such structural changes result in subtle but significant differences in social interaction.

Chapter 3:

RECONSIDERING SOCIAL INTERACTION FOR THE DIGITAL REALM

The intricate processes that comprise all social interaction are embedded in the underlying assumptions that can be made about the environment in which the interaction occurs. People learn to read and make use of the contextual information presented to them in the physical world. Yet, when they go online they inaccurately assume that experiences can be translated.

The architecture of the digital realm fundamentally conditions potential social interactions. Although designers and theoreticians have emphasized the metaphors that translate the physical to the digital, these metaphors are often inaccurate, if not deceptive. Architectural and spatial metaphors span the writings on cyberspace, suggesting that most aspects of the digital landscape can be compared directly to a physical replica. This metaphor is taken up in the spatial language that we use to discuss digital environments - chat *rooms*, web *sites*, message *boards*, and *portals* all exist in the realm of cyber *space*. Even the words that researchers use to separate the physical from the digital imply space: *world*, *landscape*, and *environment*. Yet, while these notions are sold for ease of comparison, they imply a set of architectural assumptions that are not applicable online. Thus, they mislead people into believing that they should act in a comparable manner and will receive the appropriate feedback.

Metaphors are one of the more effective means for people to build new conceptual models (Lakoff & Johnson 1980). This linguistic tool allows people to translate their mental assumptions from an understood concept to a new idea. Metaphors make the

new concepts seem intuitive by relying on previously understood ones. Of course, this is only successful when the assumptions can be accurately translated. In the case of the digital realm, translating physical expectations to the digital world is problematic. In physical rooms, people expect a certain level of privacy and control over their words because their experiences have indicated that social interactions are ephemeral and the average interaction remains in the context in which it was presented. Online, information is archived by default; thus, what is said in one room might not be as fleeting and immobile as the speaker believes. This immediately creates a tension between the expectation that an individual has and the reality of the architecture.

Although experienced users understand that the metaphors do not map directly, the architecture gives off an entirely different impression. Harrison & Dourish (1996) argue that the difference is that of space versus place. Thus, when the architecture implies that the virtual place is located in a spatial metaphor resembling the physical one, the architecture is deceptive. Thus, metaphors do not necessarily need to be retired, but they must be supplemented with mechanisms for architectural awareness.

Without this awareness, being taken out of context can be quite disconcerting. In order to address this, designers should convey the social norms through the architecture. They should simultaneously inform users of the underlying differences while providing the tools for people to more comfortably interact online.

This chapter presents some of the underlying differences between the physical and the digital, focusing on those that impact social behavior. I focus on two main architectural differences that impact social interaction – the power of architecture and the lack of embodiment. In looking at these, i am interested in the ways in which they impact one's ability to derive context and the other social cues necessary for communication.

Underlying differences in architecture

The architecture of the Internet is code (Lessig 1999), which is comprised of digital bits. Over seven years ago, Negroponte (1995) proselytized the notion that bits were not the same as atoms and thus must be treated differently. Shortly following, William Mitchell (1995) constructed an early critique of how the architectural differences would impact social interaction. Yet, even with such awareness, designers failed to inform users of this.

In the world of bits, many tasks are trivial when compared to their physical counterpart. Copying data is a core function of code; transporting bits over wires takes moments; altering data, images and text requires little effort and leaves few traces. Digital information is easily stored, manipulated, sorted and copied. Thus, most data that has passed through the Internet exists in many different forms on all of the systems through which it passed. While a typical conversation leaves nothing more than impressions in people's minds, online conversations are often recorded because of the nature of their passage. Whether they exist in email or on Usenet, this data is frequently archived, sorted and searchable.

Although it may seem advantageous to have historical archives of social interactions, these archives take the interactions out of the situational context in which they were located. For example, by using a search engine to access Usenet, people are able to glimpse at messages removed from the conversational thread. Even with the complete archive, one is reading a historical document of a conversation without being aware of the temporal aspect of the situation. As such, archived data presents a different image to a viewer who is accessing it out of the context in which it was created.

Digital archives allow for situational context to collapse with ease. Just as people can access the information without the full context, they can search for information which, when presented, suggests that two different bits of information are related. For example, by searching for an individual's name, a user can acquire a glimpse at the individual's digital presentation across many different situations without seeing any of this in context. In effect, digital tools place massive details at one's fingerprint, thereby enabling anyone to have immediate access to all libraries, public records and other such data. While advantageous for those seeking information, this provides new challenges for those producing sociable data. Although the web is inherently public, people have a notion that they are only performing to a given context at a given time. Additionally, they are accustomed to having control over the data that they provide to strangers. Thus, people must learn to adjust their presentation with the understanding that search engines can collapse any data at any period of time.

In the physical world, the public space still has boundaries; people are not performing for the entire world, across all time. They are performing in a particular environment and draw from the contextual cues of that environment. Online, when an individual

performs for a particular chatroom, they make certain assumptions about who has access to their presentation. When these interactions are recorded, the conversation can be repositioned into a different context. Although recording is an inherent attribute of shared bits, the digital design does not inform the users as they have come to expect offline. Thus, people are still startled when public presentations reappear elsewhere. The history of Usenet provides a clear example of the social impact of collapsed contexts.

Usenet: an example of destroyed context

In the 1980s, most people who had access to the Internet were either associated with universities or corporations. Many of these people regularly participated in conversations on Usenet, an asynchronous threaded messaging system that was available to everyone. Usenet was divided into topical groups, which represented many of the interests of these people and thus spanned an extensive range of topics. Yet, while there was diversity of interest, there was still an assumed homogeneity to the participants; it was not until 1992 that an AOL user posted to Usenet (Google 2001). Posters often knew each other and were equally familiar with the digital terrain.

Posters knew that they were posting to public forums and that anyone who had access could read their posts. Perhaps a little bit of hindsight makes it seem obvious that the Internet could one day be comprised of most people and that those posts would be permanently archived and reassembled with search engines. And perhaps those posters should have had that foresight, but many of them did not. People posted messages with a particular thread and group in mind, having a full understanding of who tended to post to that forum. They generally assumed that most readers had some vague interest in the topic at hand, but that their message was always read with the other messages and the thread for context. People often expected that their messages would last for a few months, as they routinely saw old messages fade away from their server. Posters had a sense of interpersonal and situational contexts, derived in part by assuming that it was like any group meeting space, where some people were vocal and others remained anonymous in the background.

Yet, as time marched on, the masses jumped on the digital bandwagon and started to participate in all of its forums. Usenet grew rapidly; new groups were added; old inhabitants left; and the culture of the groups changed over time. In 1995, DejaNews was

introduced as a searchable archive-based interface to Usenet; in 2001, Google acquired DejaNews and expanded the archive to 20 years worth of posts.

Suddenly, with a few keystrokes, millions of grouped postings could be condensed into those that pertained to a given keyword. Perhaps ideal for searching for answers to questions, this interface quickly removes any of the original context in which the post was created. While the date and links to the thread are included beneath an excerpt of the message, the interface allows you to automatically browse these messages out of temporal or group order. Although messages were created within a particular context, it is not necessary to know anything about that space to browse the messages. Nothing distinguishes the posts of one group from that of another, one time from another, or one individual from another.

Without knowing the context and history of a given newsgroup or individual, or the social norms of a given time period, messages can be easily misinterpreted. If a search for an individual shows postings from rec.pets.cats and alt.flame, and the searcher is not aware that angry postings are expected in the latter, the poster might easily be perceived in a negative light. Without knowing the context of the space, people do not know how to assess the specific social norms separate from a general view of normative behavior. Even a date-based search for my advisor, Judith Donath, suggests that the two most related groups to her are rec.arts.books and rec.autos.antiques; without knowing the group or the information being discussed, one might easily misinterpret what these "related groups" mean.

DOCTOR FUN

24 January 96

Copyright © 1996 David Farley, d-farley@tezcat.com
Distributed by United Feature Syndicate
http://www.unitedmedia.com
This cartoon is made available on the Internet for personal viewing only.

Suddenly, just as Paul was about to clinch the job interview, he received a visit from the Ghost of Usenet Postings Past.

Usenet highlights the contextual problems associated with digital data. Although users post-1995 were not told about DejaNews, many were aware that it existed. By being aware, users were able to adjust their presentations to accommodate for the potential collapsing of contexts due to the change in architecture. Prior to that, many users lacked the assumed foresight; they did not anticipate these conditions. The architecture made archiving possible,

but posters did not predict that their messages would continue to persist and impact their interactions years later. Although almost everyone concedes that posts were public, the notion of *public* in the physical realm does not mean persistent across all space and all time. When a twelve-year-old states an opinion to a group of strangers in a public park, it is not assumed that this will be quoted out of context in a job interview fifteen years later. Likewise, it is not certain that society should require that level of accountability for past statements; even the credit bureau forgives an individual after seven years.

As massive quantities of Usenet data are aggregated, it is not surprising that researchers analyze it. While most of the analysis results in academic papers, Microsoft's Netscan (Smith 2001) provides a tool for users to see the resulting statistics about a given newsgroup, a given conversation, or a given person. While this data helps users gain perspective about the various groups and people, it can also be socially problematic. The statistics about groups are not put in a given context. If a group has 50 active members, is it more like 50 people in a football stadium or 50 people in a bedroom? Without having to know anything about the context in which posts originated, one can explore statistics on anyone's posting habits. What does it mean when someone posts messages to which there are no responses? Does this mean that the person is quite knowledgeable and is answering a question or that everyone would prefer to ignore this individual? Usenet comprises lots of different types of social norms. As discussed in the next section, presenting statistical data can be problematic.

Digital architecture provides different social cues

While Google provides a mechanism for collapsing contexts in Usenet, it also provides a means for people to instantly access extensive information about others throughout the world. This tool has both advantages and disadvantages. On one hand, having access to data about others informs the curious individual, as is noted by those who scour the Internet for personal information on potential dates (Schoeneman 2001; Rosen 2000: 199). Search engines allow people to sift through data to get a glimpse at someone of interest in order to evaluate potential connections. At the same time, this information can be misleading or inaccurate, thereby misinforming the individual. Perhaps the data is from an untrustworthy source or does not represent the individual in the current situation. Or perhaps the data reveals information about someone with an identical name. When people acquire information online, they are not aware of the validity of their sources.

Even if extensive, accurate information about an individual were to exist, users are not likely to read it all. With a limited sample, impressions may be inappropriately distorted.

Not only is the reader disadvantaged by not having the tools to properly evaluate the information, but also the subject lacks the ability to control their representation. As information is archived, it is also difficult for a subject to correct inaccuracies, let alone adjust potential impressions. With such data available, it is difficult to resolve old issues and one must be prepared to justify their past continuously. Such records are problematic, as they require people to "live their lives knowing that the details might be captured by a big magnifying glass in the sky" (Lee 2002).

In the archives of the digital world, the records of heated flame wars and other digital mistakes remain persistent. For some, this is a source of anxiety, shame and embarrassment. In the midst of my research, i received an email from someone who wanted to know if i had any solutions for purging old data:

> I had a rather bad public battle and due to being outnumbered by a bunch of jerks, I was made to look VERY bad many years ago and these same individuals feel the need each year to rehash the past and keep this wound open and painful to me, and I have no way of getting these "crap" purged. (Anonymous 2002)

Past posts are consistently part of a user's digital present in ways not comparable to the physical domain. Slander and gossip are archived, but the subject has no recourse for adjusting this data. In such incidents, people feel misrepresented and powerless.

Not only must one accommodate for their historical presentations, but they must also be prepared to deal with the quantitative data that is produced to represent them. For example, sites such as eBay can tell you about a users' reputation through a set of numbers. This simplification might make sense when you are evaluating a users reputation is one particular context (i.e. as a capable seller), but if reputation scores are calculated across different behavioral contexts, as is proposed by Microsoft Research (Smith & Fiore 2001), this could have tremendous social consequences. Using author profiles to evaluate someone's reputation, as one number based on 21 years of Usenet data, can be quite confounding. How do these reputations accommodate for context? Is one's post to alt.flame acceptable if it is not a flame? What value does verbosity have in evaluating an individual's worth? Do people have any say in how these statistics are used to represent them?

When a system presents reputation data, it alters the social structure. Thus, the design of such systems must be handled delicately. Researchers at AT&T accidentally discovered this problem when they equipped a chat bot with the ability to tell people various social statistics (Isbel, et. al. 2000). Although intended as a friendly feature for people to understand their own statistical behavior in relating to others, people quickly used it as a method for seeing how valuable they were in their friends' social network, which developed tremendous social tension. Numerical representations rarely convey the nuanced details of a situation, leaving room for abuse and misrepresentation that is more destructive than helpful.

Although the intention is to provide meaningful feedback, this is only helpful when it is representative and accurate. Inaccuracies come not only from mistakes but also from those who abuse the system. As these systems impact those that they represent, it is important that methods of recourse exist whereby users can challenge the results. The United States government recognized this need and drafted the *Code of Fair Information Practices*, which mandates transparency of governmental data with an explicit recourse protocol (Garfinkel 2000: 7). Limiting an individual's ability to control their representation is problematic. With identity theft on the rise, systems that emphasize scores for privileges, but provide no accountability, are open to harmful abuse.

As these problems are inherently architectural, users have two choices: either learn to manage with these systems, or demand designers to adjust the systems to meet the needs of the user. Although i argue primarily for the latter, it is also essential for users to be aware of the current structure and act accordingly. In order to encourage awareness, system designers should provide behavioral and systematic feedback that conveys the norms of the environment. This is important, as system interfaces not only affect a user's ability to derive context, but also to present one's identity.

The value of embodiment

Communicative performance typically utilizes the subtle nature of one's body. People know how to utilize their bodies to convey nuanced details and attitudes, and to otherwise affect the tone of any verbal cue. Through experience and mental models, people know how to read those subtle cues and evaluate people's bodies. Yet, online people must operate through a different medium. They project their ideas through the computer interface and perceive the output that the computer provides. Social

interactions are limited by what people can convey and perceive in the mediated space. In current systems, both the performer and viewer have limited channels for expression and perception. Thus, much is lost in digital conversations; attempts to convey intention can be frustrating.

   The spatial qualities of digital environments are devoid of meaning or functionality. If there is any decoration in the space, it is in the form of digital wallpaper or images that are supposed to mimic physical objects, such as graphical beer glasses. These items have no use in the digital environment; people cannot actually drink from a glass online. Additionally, the decorations are not tied to any fundamental aspects of the space, regulated by market forces or usage. A digital Van Gogh has no value. Digital decorations represent what the space wants to convey, not necessarily what it is. While these decorations are not particularly helpful, most online spaces lack even that level of spatial cues, relying simply on the digital equivalent of a piece of paper as the interface for interaction. None of these environments are affected by previous usage; history is told in logs, not through the effects on the space. Yet, in the physical world, the marks on the floor, the scratches on the table and the aging of the wallpaper convey subtle details that people evaluate in assessing a space. Online, everything always appears untouched. Unlike the physical realm, digital environments show no information about temporality, do not change over time through interaction, and do not communicate their history.

   Online, we are unable to see much of the interpersonal context cues – how many people, common characteristics of the people, fashion statements, gender, age, activities, etc. Yet, by quickly glancing at a physical crowd, one can easily determine these as well as what the social norms are, and how many people are abiding by various sets of rules. Crowds online are invisible. Body language cues and facial expressions are missing. What remains is a set of textual descriptions and expressions, with perhaps a graphical representation of oneself or a collage-like homepage that indicates manually articulated aspects of one's presentation. In order to detect crowds, people try to make meaning out of the download speed of websites or the tickers on websites to indicate visitors (Xiong & Brittain 1999). Online, people are given limited signals, and those are often inaccurate or inadequate for people to properly develop their mental models.

Recognizing conventional and assessment signals

People rely on the signals that others exude in order to assess their identity presentation. Yet, for those signals to be meaningful, they must have the ability to determine the validity or relevance of a given signal. *Assessment signals*, which are costly to possess, are quite reliable (i.e. a person with large muscles can be reliably perceived to be strong). Conversely, people can present *conventional signals* with little effort, but they are far less reliable (i.e. owning a T-shirt that says oneself is strong is far less meaningful than possessing large muscles) (Zahavi 1997). The signals that people present must be evaluated for both their meaning as well as their reliability, for if someone is to challenge a signal, it is important to understand how reliable that information is.

Because of their reliability, assessment signals are far more desirable for the presenter and the reader. Yet, they are far more costly to possess and maintain. Online, people present themselves primarily through text. Physical signals, such as one's strength, must be converted to textual statements, thereby converting assessment signals into conventional ones. Yet, just as the reliability of the signal is decreased, so is the likelihood of harm when challenging the signal. Different forms of assessment signals evolve online, such as an email address at a prestigious domain or certain types of public archives. Both online and off, assessment signals require time and complexity to develop and present. Yet, online, conventional signals typically evolve from the documentation of assessment signals being challenged, rather than just existing as an end result. Because of the amount of time necessary to evolve assessment signals online, people constantly interact with conventional signals, which must be challenged or accepted despite the low level of reliability. As a result, deception runs rampant, as people are too likely to trust the signals that they are given, particularly those that refer to the body (sex, age, race, etc.), which are rarely challenged offline (Donath 1999).

While text does provide some information about one's identity, it is not nearly as rich as the detailed information that one conveys through body and fashion. Online, minimal information can often be harmful, as coarse data requires that people interpolate from missing information in order to build their mental models (boyd 2001). This approach is particularly problematic because people are not likely to reevaluate their initial impressions (Aronson 1995). When engaging with another in social environments, people construct a mental image of that other person, even if the only information that they may have is data such as 21 years old, white, female with blonde hair. Should their

mental image resemble Britney Spears, they are most likely going to be wrong, resulting in an uncomfortable dilemma for both parties. As people read one's performance in relation to their mental image, conversation subtleties may be inaccurately perceived. Such is the case when people inaccurately assume someone's sex (O'Brien 1999).

Embodiment provides both social cues as well as a mechanism for people to properly present themselves; by not providing this information, the digital world fails individuals. This results in a slew of peculiar interactions, fundamentally due to a failure to properly communicate.

Regaining context through account maintenance

Inadvertently, users have formulated new behaviors for managing context online. As data is primarily collapsed through one's name or email address, people create multiple accounts and associate particular accounts with particular contexts. The most obvious example of this is the separation between work and personal email addresses. By managing multiple accounts, people are able to regain some control and privacy. In doing so, they are also formulating a new paradigm for conceptualizing context – localization.

Maintaining multiple personas online satisfies many goals for the digital individual. In the early days of MUDs and MOOs, people regularly explored their identity by playing with different online personas. Because people chose to fragment their social identity, digital researchers such as Sherry Turkle (1995) and Sandy Stone (1998) saw this play as indicative of a postmodern, fragmented self. Yet, the play in which people engaged simply gave them the ability to reflect on, experiment with, and process their own identity. Fragmented social presentations online provides even greater flexibility for the multi-faceted individual, as it allows them to walk through common spaces presenting different aspects of themselves rather than being required to maintain one persona per space, as is necessary offline (Reid 1998: 37). While role-playing is a fascinating, it is only one of the motivations behind maintaining multiple accounts.

Seeking privacy or segregation of lives, people maintain multiple accounts that represent different facets of their internal identity. In the realm of Usenet, this allows the user to use one account to discuss topics related to programming and one to talk about recreational interests. As an alternative to anonymity, this allows people to build

reputations and friendships while only revealing particular aspects of their identity. So long as people maintain a strict boundary between accounts (i.e. not providing one's name or other identifying information), this provides a barrier when archives aggregate across or allow access to data through individual identification.

By maintaining multiple accounts, users associate context locally. In other words, rather than adjusting one's presentation according to the situation or current population, one can maintain an account that represents a specific facet and present oneself through that. In doing so, people take their internal facets and create external representations for them. Thus, faces function directly from externalized facets, or accounts, rather than through the individual themselves. When reading for situational and interpersonal context information, people assess which facet should be associated with that interaction and use it exclusively. As one moves from one ephemeral context to another, one simply switches accounts or facets. Thus, when one logs into one's work email, one knows that one is presenting the work face uniformly through this account.

In doing so, people have started a new paradigm of social interaction online. Although this may initially appear peculiar, multiple email addresses/handles fill a desired void of the digital realm – the ability to manage the given context. They have minimized the collapsed contexts by maintaining the contexts locally; thus, what is aggregated is done so across a particular facet instead of a particular individual. Of course, people maintain a varying number of accounts and they differ as to how strictly they segregate their different facets. People's consciousness of this behavior is often dependent on how much they feel it is necessary to maintain segregated facets.

While such control mechanisms work as a substitute for the failure of digital context, they are only temporary bandages for a larger problem. It will be collapsed in the near future, accidentally or maliciously by those who want to reveal people online, through new technological advancements, or systematically by initiatives such as Microsoft's *Passport*. Managing separate facets is neither convenient nor intuitive; thus, only those with the greatest need put forth the effort to segregate their facets.

As is discussed in more detail in Chapter 5, *Passport* encourages users to maintain only one account. It is in the market's best interest that a user be unable to present facets, for marketing purposes as well as control. Thus, people's motivation to start regaining context in a unique way suggests the importance of such behavior in digital interactions.

Concluding thoughts

With information fundamentally missing, people are trying to find new ways to make sense of their interactions and regain awareness and control over their presentation. As people will inevitably adjust to the architectures that they are given, the goal is not to eliminate the possibilities that are afforded by the underlying potential of digital environments. Instead, designers must recognize what users are trying to do and provide them with the tools that will make it easier.

First, people need *self-awareness*. They need to understand their representation and role in digital interactions. While others see their presentations and have immense data about them, people are not often aware of the traces that they leave behind. Without this awareness, control seems impossible. Thus, people react to the problems without having an idea of how to stop them from occurring.

Awareness is necessary at both an individual and group level. People must be aware of the group as a whole, what the norms are and how other people are behaving. They must be aware of reactions as well as presentation, people as well as the virtual space. They must be aware of the contextual information that surrounds them. Without this awareness, people act in a *disinhibited* way, suggesting that increased awareness will result in increased self-regulation (Joinson 1998: 51).

Awareness is the first step for people to be able to manage their presentation and identity online. In addition, they need *management* tools to properly organize and present themselves. As they are not able to present their bodies, they need tools that will allow them to represent their digital equivalent, often facets of their identity as opposed to their whole being. By managing their facets, these tools should allow users to present faces as they see fit.

Awareness and management provide feedback that makes an environment more socially comfortable, as they provide information that people use to present themselves. Such information also provides users with some of what they need to self-regulate. By enhancing digital environments with desired channels for feedback and control, designers can empower users and create the environment for more fluid social interactions. Thus, the remainder of this thesis focuses on what is necessary to provide such awareness and identity management.

Chapter 4:

SELF-AWARENESS IN SOCIAL INTERACTIONS

Awareness empowers individuals, as it gives them the ability to understand their position in a given system and use that knowledge to operate more effectively. In social interactions, people want to be aware of their own presentation, of what is appropriate in the given context, and how others perceive them. In the physical world, this awareness comes relatively easily, as people know how to derive meaning from the information conveyed by their bodies and those around them. In daily interactions, people are aware of their presentation: they know what they are wearing, they have a sense of their facial expressions, and they can easily comprehend the reactions presented by others. Yet, online people produce immense quantities of data about their identity and behavior without an awareness of what that data is, let alone what it represents. People do not have the tools to be aware of their presentation online. Likewise, they are unable to gain access to the implicit data produced by others. Yet, these two components are essential for interpersonal contextual awareness.

Context awareness is a fundamental concern of the ubiquitous computing community, as awareness is necessary for interaction. Yet, much of the research in this area focuses on revealing environmental factors that the system can sense, including functional qualities of the space and quantitative interpersonal information such as presence. As exemplified by Anind Dey's work (2000), context-awareness in ubiquitous computing focuses on revealing external activities to the user. Although environmental awareness is essential, it is also necessary to have self-awareness. Users must not only be aware of the environment, but also of themselves within it.

In this chapter, i begin by discussing the data that people produce online and then highlight current approaches to social awareness by addressing various systems. After critiquing different approaches to awareness, i prescribe a tool that attempts to provide awareness as a user interacts with various systems.

Considering the data that individuals produce

As i have already discussed, digital data is inherently archiveable. This means that systems are able to track anyone's digital habits, including what websites they visit, who they email or instant message (IM) and when, what forms they fill out, and when they are online. Any data sent over the network, whether intentionally or unintentionally, can be archived and used to help represent an individual's behavior. Some data that individuals produce is done so intentionally, such as the messages that someone writes in an IM window. Other data is archived by servers without explicit consent from users, such as the footprints that one leaves when exploring the websites. While messages are freeform in their structure, people also relay structured data such as the profiling information required by many websites. Whenever the go online, people produce immense amounts of data about themselves without even realizing it (Behr 2002):

> External logs (web, IM): login time, duration, files accessed, referring website, connecting
>   to what people
> Personal ISP logs: time/date/duration, tracked web contact access, email messages
> Profile data: age, sex, address, email, occupation, industry, income
> Affiliations: website/email domain
> Personal website content: links, interests, bio, photos
> External websites references to an individual
> Message content: email, chat, IM, SMS, Usenet/bboard posts, journals/blogs
> Sharable data: MP3s and other files
> Social networks: IM, email, chat, Usenet/bboard
> Presence data: IM buddy lists, Outlook calendars
> Shopping habits, browsing habits, recommendations
> Reputation as buyer, seller, advisor
> Archives of data over time: conversations, websites

Unlike the data that one typically produces in the physical world, all of this data is stored and can be accessed with relative ease. Currently, this data is not centrally located; each server logs an individual's behavior on that machine alone. Unless a user only uses one machine, a complete set of data is not maintained locally either. Since most

users are not aware of the unintentional data that they produce, they are unlikely to store the aggregated data.

The aggregation of this data is quite powerful in helping construct a complete image of an individual, as marketing companies and corporations such as Microsoft (*Passport*) have already recognized. While external systems are working to reconstruct people through their data output, individuals are not even aware of this data, let alone how it could be perceived and used. Although i do not condone most of the corporate goals with regard to this data, i believe that the mechanism for empowering individuals starts with giving them access to this data in a meaningful way, entirely for their personal use.

For data to be comprehensible, users need more than access to the raw data. While one's browser history is quite interesting, logs are not intuitive. Telling a user that they spent over half of this week's web time at online bookstores is far more meaningful. With the quantity of information available, it needs to be distilled and encapsulated in order to be comprehensible. Doing so requires tools that are intended for this purpose.

Tools for creating awareness

Awareness can either be provided post-facto or integrated into the application. While the former provides reflection, the latter is more desirable as it allows the user to immediately respond to the information. Yet, most research focused on self-awareness deals with post-facto data both for simplicity and because of a lack of access to application source. This work focuses on revealing underlying patterns to the user, quite often through social visualizations.

In order to give the reader a sense of the different approaches, i have selected a sample of awareness tools and offer a brief analysis of their strengths and weaknesses. These tools focus on making social information available to the users, yet the information is not always simply about them; often it is about their relationship to others and to groups. The pieces that i have chosen as examples either emphasize making the raw data accessible or use the data to convey more generalized notions of the people in their space. Much of the work that i address comes from Sociable Media, my own research group, as we continue to be the dominant group working on social visualization.

Making data accessible: Netscan & Blogdex

Although Usenet data is public by nature, it is difficult to ascertain what the trends are within these environments. *Netscan* (Smith 2001) captures and processes Usenet data and makes it available to the public in the form of statistics – how many people, how regularly do they post, how often do people reply, what groups does an individual participate in, etc. This data provides a digital portrait of users and groups through their statistical habits.

Blogs are quickly emerging as a trendy way to share information with others on the web, as they let people post links to interesting sites and comment on others' posts. While one person's blog is quite interesting, the phenomenon as a whole is even more fascinating, particularly looking at what is fashionable to post, who links to who, and how rapidly the trends change. By analyzing as many blogs as possible, *Blogdex* (Marlow, 2001) provides a tool for people to see what the trends in blogging are, how their blog relates to the habits as a whole, and what their relation to other bloggers is.

Both *Blogdex* and *Netscan* provide a mechanism for aggregating data and conveying information that is often obfuscated, so that users can see their habits within the larger system. Yet, they do not pull out the meaningful trends or convey what the statistics might mean to the user. For example, while *Netscan* lets a user see that a particular group has 50 active members, this data is most likely meaningless to the user. Even when compared to other groups, the user cannot easily determine if 50 people in a space suggests that the room is more like an empty football stadium or a living room. Additionally, trends are relative. When *Blogdex* was reported in the BBC, over 200 Farsi blog owners added their blog to *Blogdex*. As a result, the rankings quickly changed due to the increased Farsi traffic; thus, it was clear that the rankings are only appropriate for the types of blogs who have added themselves to *Blogdex*. At the same time, it is difficult to get a sense of what types of blogs have been added, and which have not. In both systems, determining the trends or the meaning of the data can be challenging.

While these tools fail to make the leap between the data and their value, they are particularly noteworthy because they take the first step in making otherwise uncollected data accessible in unique ways. By using them, the curious and thorough user can develop their own intuition about the environment by scouring the statistical data for meaning. While these tools currently provide the equivalent of a well-structured system

log, they present the most salient statistics for environments that are otherwise not accessible.

Visualizing statistical data: PostHistory & Live Web

Focused on revealing an individual's behavior, *PostHistory* (Viégas 2001) was developed to give users a sense of their email habits. Drawing from one's email archive, the system analyzes the data to understand who converses with whom, when, and how often. *PostHistory* conveys this statistical data in an elegant and compelling visualization where users are able to easily see information such as which people write to them the most, what the relation between time and people is, and how often they receive personal messages versus group messages. While the current implementation is graphically compelling and legible, it only provides the essential data and makes no attempts to evaluate it for the user. This is both a strength and a weakness, as users are encouraged to reflect on their own behavior yet they are unable to delve into the data to understand its contextual relevance, partially because *PostHistory* does not allow users to access the underlying message data. For example, just because an individual sends the largest quantity of messages does not mean that they are that valuable to the user; each message may only be comprised of a few words or might be solely associated with a listserv. Without being able to see why the individual was rated so high, the information may be misleading.

When people surf websites, they are sharing the space with others, yet this aspect of social awareness is difficult to perceive other than recognizing that a site is slow. *Live Web* (Xiong & Brittain 1999) visualizes the data traces that each server maintains about visitors. Thus, users can see who else has recently visited a website and what path they took as they followed various links. This type of a system makes the social aspect of system logs accessible to the public, letting them get a sense of interpersonal context. While people are able to observe one another within one site, they are not able to gather more information about the people or follow them outside of the particular site. *Live Web* does not provide enough information for anyone to be more than simply an intimate stranger, as it does not provide the motivation or detail for people to communicate with one another.

Developed in Sociable Media as tools for visualizing inaccessible data, both *PostHistory* and *Live Web* focus on revealing the underlying logs of the data. They do little to imply

information about the user in relation to others or the community itself. Yet, by making the designs so compelling, they reveal data in a meaningful way, thereby offering the first step in providing users with the knowledge necessary to understand the social behavior around them.

Impression-driven visualizations: Loom2, Visual Who & Social Network Fragments

In *Loom2* (boyd, et. al. 2002), Hyun-Yeul Lee and i began exploring how a visual language could be developed to convey the socially salient features articulated by Whittaker, et. al. (1998) that *Netscan* (Smith 2001) exposed. By trying to understand the relevance of social data to the user, we created a series of artistic and computational sketches that allowed people to interactively explore different aspects of Usenet environments. The *Loom2* project focused on a series of sketches and designs that explored different aspects of information presentation, including some that were too complex to fully integrate into current systems. For example, we recognized the power of text in serving as both a functional mechanism for gathering meaning about the message as well as a beautiful form that could convey underlying intentions. Thus, one aspect of *Loom2* was to explore how glyphs could be animated with motion to personify the textual individual. Although this was only done through a handful of interactive prototypes, we recognized the power of conveying impressions as well as meaning.

Loom2 started to reveal the importance of giving people multi-layered data, such that visual information could help them create quick meaningful impressions, but also provide them with the detail necessary to explore the actual raw data at a lower level. The value of an interactive visualization system is that it draws both on the power of visual cues as well as layered information, or what Ben Shneiderman (1987) refers to as the Visual Information-Seeking Mantra: *overview first, zoom and filter, then details on demand*. *Loom2* recognizes that people need more than just simple access to data – they need to understand how data relates to them and how they relate to others. By approaching this issue through design, we began to develop a visual language that focused on providing data awareness by relying on cues that people understand, including aspects of motion, color and graphical layout.

Visual Who (Donath 1995) is an interactive visualization of mailing list and other group/member data. By interacting with the system, users are encouraged to comprehend highly dimensional data about their relationship to groups based on the

stereotypes of the members of those groups. Thus, *Visual Who* offers users a tool for comparing themselves to the group, where the groups' value is based on the external activities of all its participants. Closeness does not suggest that an individual is interested in the associated group; merely, it suggests that the individual has much in common with the members of that group. For example, the system strongly associated a Media Lab professor with skateboarding; he was not even remotely interested in skateboarding. As many of his students were skateboarders and he had a lot in common with them, he became associated with that group. One of the problems with this piece is that users can easily mistake the feedback they are receiving as indicative of their relationship with other people. When the system positions two people nearby, it simply suggests that the individuals have the same relative pull to the groups present. Thus, the only thing that they have in common is the same tie ratio.

In order to provide users with an awareness of the structure of their social networks via email interactions, Jeff Potter and i developed *Social Network Fragments (SNF),* which is detailed in Chapter 7. By analyzing email behavior, we associated a value for different types of email relationships based on how much they indicated an awareness or knowledge of others on a similar message. For example, when a user sends a message to two different people and blind carbon copies another, what can we say about the various ties in terms of how well the people know one another? By assigning a value to each of these ties, we developed a language for quantifying the weight of two people's relationships. Using this, *Social Network Fragments* visualizes the complete graph of people's relations with one another, focusing on conveying the structure of the social network.

All three research pieces focus on providing impressions by constructing a legible social landscape, as described by Donath (1996). By developing a language for relating people and information, both *Social Network Fragments* and *Visual Who* offer users an interactive interface in which to explore the social information that the system derives from the data that they produce. As the information that they convey is impression-driven, these systems are bound to be misleading at times. In *Visual Who*, users often mistake the graphical distance between users to be meaningful, while the only meaningful relationships are between the people and the various groups. Likewise, in *SNF*, the clustering algorithm can collapse dimensions in a way that places unrelated people near each other on the two dimensional surface.

As was recognized by the *Loom2* project, conveying impressions is a delicate process and the mistakes extend beyond just readability. Not only must researchers concern themselves with how the data is analyzed, but they must also take these qualitative values and convey them as impressions on the screen. Thus, there are bound to be errors in both steps. Yet, this approach is also important, as it is impressions that people want, not simply a vast quantity of unanalyzed data. Even the imperfect impressions that are conveyed by *Visual Who* and *Social Network Fragments* are quite compelling, because they are providing insight that is otherwise inaccessible to the users.

By giving people access to both data and the possible connotations that can be drawn, people are able to see a different perspective on their behavior. This awareness provides cues that may not be fully accurate, but neither are most impressions in the physical world. Awareness comes not simply from understanding the statistics that one produces, but by understanding the possible impressions that this makes in relation to the individual. Thus, while it can be perceived as a weakness that these system imply potentially inaccurate information, it can also be seen as a virtue, because it is precisely these impressions that users need to be aware of when they are engaging in social interaction.

Application-driven awareness

While the aforementioned awareness research systems provide users wit h post-facto awareness, systems have also been built to integrate social transparency into the system. For example, updated versions of *ChatCircles* (Viégas 1999) share a user's historical movements by leaving traces on the background of the chatroom while Erickson, et. al. (1999) have integrated awareness mechanisms of presence and participation into *Babble*. Both of these systems provide feedback to the users, including: who is there, who can see them, who is participating and with what level of activity. They provide a record of interactions, allowing users to see more than just the current data. The feedback mechanisms in these systems are intended to improve users' experience by making them more aware.

Likewise, many non-research tools incorporate feedback so that users can use the systems in a more effective manner. In particular, these systems reveal some of the data that the user has provided the site. Although this information is available, it is often obfuscated, as it is primarily intended for sporadic review. That which is readily

available is intended to help the user browse. Most often, that which provides the best awareness is not intended for such; yet, it can be co-opted by users to reflect on their own behaviors.

For example, some webboards, such as *ezboard*, give users tools to see the history of their posts, to see and edit a public profile and to track the responses to their messages. *Amazon* lets the user know that they are being observed by welcoming them by name in a manner that allows the user to see and edit much of the data that that they have stored about them. *Yahoogroups* lists all of a member's associated groups for their direct access. *Ebay* provides a user website that lets users modify their preferences as well as respond to the feedback about them and see how their reputation has been affected by others. Presence information can be seen through most instant messaging programs. Most e-commerce and communications sites provide some aspect of data awareness, whether it is the history of one's interactions or a profile that one is presenting to the company and other users.

By providing users with a centralized location where all of their membership data is located, these sites give users an opportunity to observe how the site and others see them. Unfortunately, what is typically provided to the users is not complete transparency; people still do not know how they are given a particular recommendation or why they receive a particular advertisement. Additionally, the structure of most sites does not indicate the level of observation that is occurring. By welcoming the user, *Amazon* provides a counter-example; the hello informs the user that the system is watching, thereby providing architectural feedback. Yet, for the most part, sites have no motivation to provide awareness as their data collection is usually for advertising purposes. What awareness they do easily provide is usually about other users, such as reputation scores.

Bridging research applications and web feedback

The aforementioned research systems provide direct feedback and make hidden data accessible, while the typical website provides feedback incidentally. Yet, both approaches are advantageous. While sometimes obfuscated, web feedback is incorporated into the system and changes over time. It is limited because it typically provides minimal feedback. On the other hand, the research applications convey rich data; most of them are more effective as portraits than ongoing awareness tools.

Unfortunately, when t hese systems run off of live data, they sit as separate applications, not directly integrated with the application being used. The feedback that they provide is often about the system as a whole, not simply about the user's role within the system. Thus, they may be considered separately.

Recognizing that most self-awareness applications are focused on giving users an overriding image of self and not one that is integrated with the current context, i started to consider what would be a better way of providing awareness. In doing so, i imagined a tool that interactively and continually provided awareness about the user as they operated in the digital world.

Digital Mirror: A tool for reflective self-awareness

A mirror provides an image in which we can see ourselves, our identity and postulate what others see. From Lacan's perspective (1977/1966), the *mirror stage* in development is when children first get a notion of themselves as unique individuals. This mirror reflection provides a source of feedback t hat allows us to adjust our presentation in order to convey what we want to project.

The mirror is an interesting metaphor for consideration, as people do not operate with such awareness in the physical world. In fact, performing in front of a mirror takes on an entirely different aura than performing without one. Yet, in our embodied selves, we have a decent sense of what we are projecting. Online, we lack the body with which to project ourselves and thus we project our ideas into a digital representation that serves as our online agent. By operating our agent, we assume we are able to perceive ourselves, as we can access our profiles, manipulate our location, and create textual messages. Yet, this presentation is deceptive as it is not what others can see.

Those who see us are also seeing much of our past. For example, when one logs into a website, the website does not just see the current set of actions, but aggregates them with all previous interactions. As we interact, the information that can be accessed about us is potentially great. Although it is inconvenient to log all conversations, this is potentially available to others. As discussed earlier in this chapter, given any application, much data can be stored and accessed.

Given this, one approach to empowering users through awareness is to give them access to all that could potentially be seen about them. By presenting this data in an accessible manner, an individual could determine what is meaningful. Just as with the mirror, the user not only sees what they believe they are presenting, but with the image that others can see. By revealing what can be seen given the facet that one is presenting, the system could provide the user with a different level of interpersonal contextual information. Certainly, this does not mimic the physical world, nor will the resulting behavior. Yet, by providing such feedback, people can understand how their facets operate online and have the ability to adjust them. Users do not see everything that they may have shared, but everything that is accessible in this given context, with this given facet. By integrating the tool into the interactions that one has and presenting the feedback explicitly, such a mirror system encourages the use of awareness to adjust one's behavior.

Privacy Mirror

As i was contemplating the interface for a digital mirror, i stumbled on Nguyen & Mynatt's (2002) concurrent work in constructing a *Privacy Mirror* for people's online interactions. By recognizing the power of accountability and awareness in inciting change, Nguyen presents a set of ideas that most closely resemble my own thinking. Yet, while i was imagining a tool directed at users for considering their own output in a multi-faceted contextually collapsed world, their system focuses on creating large-scale transparency in public environments such as websites. With such a system, people would be aware of all data logs, not only their own; they would be able to see the history of the people's interactions at a given site. *Privacy Mirror* would provide detailed transparency, eliminate the "secrets" behind access logs, and otherwise let users know what detailed data is being logged during their interactions.

While i agree with Nguyen & Mynatt that awareness is essential for giving people control, i do not agree with the approach of making all logged data universally public. By doing so, the system would allow for an even greater amount of contextual information to be collapsed. Although many advertising agencies have this information, consider the impact of *Privacy Mirror* if a boss discovered their employee's off-hours interest in a controversial topic. Such a system does not provide individual privacy, but transparency. Although Brin (1998) argues that such transparency is crucial for addressing the issues of privacy online, i feel as though universal awareness can only

bring about harm, as it would provide further drive towards a heterogeneous society where all people are performing for the universal public. As this is not the society that i am interested in helping develop, i decided to consider these weaknesses and imagine the interface to an improved mirror.

Digital Mirror: Example scenario

Imagine a tool, shall we call it one's *Digital Mirror* (or *Mirror* for short) that is a hovering presence on a user's system[1]. In its window, the *Mirror* shows an image of the user that changes as the user interacts with various applications. This image is constructed for the given user and is not accessible to anyone else.

Perhaps this image is abstract, showing iconic information to represent different information. Or perhaps the image is of a person who is caricatured based on the information provided. Both of these approaches have their weaknesses. On one hand, presenting a caricature shows more detail than is truly representative and thus creates an impression that has the same confounding issues as a profile by relying on minimal data to present an entire picture. Yet, at the same time, this is precisely what other people do; perhaps such a representation would make the user consider the data that they are providing. Having bits of data shown through visual iconic bits provides the data explicitly, but without the level of impression that often impacts social interaction. It is uncertain as to which approach is more valuable, and thus a fine example of needed future research. In either case, imagine that we are observing "Sarah" as she interacts with different applications using *Digital Mirror*.

---

[1] *Digital Mirror* is only a conceptual piece; there is no prototype.

Sarah logs into an IM client as *zephoria*. The people on her buddy list see that *zephoria* has logged in; they see her profile, which lists her as male and located in Boston. Many of the people on her list know her as Sarah, mostly from offline interactions; those who only know her online know her as Zephyr.

Sarah's *Mirror* now indicates her relationship with the IM client and her buddy list, indicating the profile information that is hidden to her when she opens her client. Perhaps the male identity is shown through a ♂ symbol, or perhaps the caricature is given male features. Her location could also be shown through a representative icon, perhaps a state map. Data that is accessible to all those on her buddy list is also integrated into this representation, perhaps the public Google-able data about *zephoria*.

Seeing one of her friends online, Sarah opens up an IM conversation with *Bob123*; they have talked many times before.

Although she is still presenting a facet of her identity, the context is narrowed by this direct link; thus, her *Mirror* changes again, to reflect the facet that she is in direct contact with *Bob123*. As they've shared long chats, images of conversations scatter the background of Sarah's *Mirror*. By selecting the conversations on her screen, Sarah can access these previous interactions. Recognizing that the IM character *Bob123* is identical to the email character *bob@bob.com*, the system includes their email interactions as well, as these pertain to the image that Sarah is presenting to *Bob123*. Drawing on the ideas from *Conversation Maps* (Sack 2000), *Mirror* provides users with unique words and expressions that stand out during their conversations, springing this information from the icons containing it. Thus, it is not surprising that personal qualities that Sarah has revealed in chatting litter the representation, indicating her love of music and Italo Calvino. A small graph appears, indicating the parts of Sarah's social network that *Bob123* knows about, using mechanisms derived from *Social Network Fragments*.

Bob123 asks about the well being of Taylor. Not remembering what she has shared about Taylor, Sarah turns to her *Mirror*, focusing in on the social network graph. In this graph, she can see all of the people that they've spoken with together online, some through IM, some through email. She can also see all of the people that she has mentioned to *Bob123*, including Taylor. Focusing in, she is referred to two emails and a chat log where they have discussed Taylor. Realizing that *Bob123* is referring to Taylor's health, she returns to the conversation and responds accordingly, noting not to tell him about Taylor's newfound love.

The *Mirror* lets the user delve into the data to understand from where the representations come. In this way, the system integrates previous work, where the initial *Mirror* image is a fingerprint, indicating the general information. Simultaneously, it is an interface, allowing Sarah to delve into the fingerprint to see its components; thus, it is the gateway for Sarah to access her facets.

> Switching to surf one of her favorite newsmags, Sarah's representation quickly changes to present the other facet's data. Logged in from her work machine, the website quickly notes her IP address and the website from which she came. The advert on the site is being pulled from DoubleClick, an advertising company, which is also aware of her IP address and all of the sites that she has surfed using this IP.

Unlike the more social environment of the IM world, the website is interested in data about her, often to provide her with targeted advertisements. The *Mirror* reveals this data by showing her representation through her digits, showing the site where she came from and giving her a timeline of her interactions at this newsmag.

> Sarah clicks on the timeline to remind herself of the last time she has visited. She's fascinated by the patterns, noting that she seems to come twice every day– once in the morning during her usual check-in routine and once, for a far more extended period of time, when she is anxiously awaiting the end of the work day.

As DoubleClick has also received her data, the *Mirror* links this current interaction to DoubleClick. By recalling the previous sites that she has given them, Sarah's *Mirror* connects this site to all of those other sites, producing a highly dense graph of the network of Sarah's websurfing.

> Intrigued by the suddenly large graph in her *Mirror*, Sarah decides to navigate the data in order to understand what it means. As DoubleClick has detailed and connected logs of her websurfing habits, she finds interesting tidbits about herself. For example, she always seems to go directly from CNN's website to the New York Times, and both the NYTimes and DoubleClick are aware of this incoming link. In zooming into more details about the specifics of her presence at the NYTimes, she is able to see the profile information that she has provided them, along with her history of articles read. She chuckles as she sees that the NYTimes recognizes her as a low-income male working in the financial district while living on Pennsylvania Avenue in D.C and reading all articles related to queer culture and military abuses in Afghanistan (without once looking at a stock price). As she zooms out of this profile information, returning to the parent link of DoubleClick, she is able to see a more general profile, which suggests that she is 83% male. The graph also provides her

with a view of what facets of her online presentation have been collapsed with others, mostly notably by sites who started to note which IP addresses that she has logged in to them from, using the same account.

Most people are unaware of the amount of information collected about them online, let alone how easily it is collapsed. By conveying this information in Sarah's *Mirror*, she can quickly see what is being revealed about her behavior. Plus, as *Mirror* attempts to highlight the most salient characteristics, Sarah can see the most obvious patterns in her behavior – her timing trends, the generalized categories of the sites she visits, the profiling information that has been collected about her, etc. The information should be provided in a highly dense visualization, a technique that Tufte values as being a design that gives viewers control over the data by allowing them "to select, to narrate, to recast and personalize data for their own uses" (1990: 50). Thus, dense visualizations provide good tools for reflection and awareness.

With awareness tools, people want to know how they are perceived. They are not able to see another's face so they must resort to understanding the data that others use to evaluate them. Yet, it is not only the data that matters; the situation in which the data is created drastically affects the impressions that others gather. Understanding how systems perceive a person is much easier, as that observation is usually calculated using out-of-context and numerical data. Thus, the impressions that Sarah can derive from her *Digital Mirror* when interacting with the web are far more meaningful than those she can derive when seeing her conversational history with an IM friend. At the same time, Sarah probably has a more intuitive sense of how her friend perceives her than how a computer system does, as their conversation inevitably provides feedback in the way that a data-hungry system does not.

This scenario articulates some of the feedback that i imagine would be useful to users, so long as it is distilled in a meaningful way. What is provided goes above and beyond the magnitude of information that people have offline, yet the environment is also quite different. Simply put, "information is power and currency in the virtual world" (Billy Idol, "Cyberpunks"). As people have the ability to access massive amounts of data about others, it becomes useful for users to be aware of what is out there about them. While i believe the widespread transparency is problematic for marginalized individuals, i do feel that personal data should be transparent to their subjects, as they are essential for self-awareness and identity management.

Desired qualities for self-awareness tools

By contemplating *Digital Mirror* and a potential usage scenario, i highlighted what i believe to be essential characteristics of user-focused awareness tools:

1) The tool should be contextually dependent. Thus, it must be integrated with the actual applications, collecting data from them and presenting it back to the user in an accessible manner as they are using that application. When one's facet bridges multiple applications, the presentation should include this data. Such an application would provide situational contextual information by allowing users to see what applications their facets transcend.

2) Awareness tools should only provide the data that might also be available to the system or person with whom the user is interacting, not just everything that the user presents. In doing so, the tools take into consideration the types of faceting that a user has developed and return meaningful information for the interaction at hand. As links between data are shown, the visual aspect of the interface should indicate how likely that link is to be made. For example, a set of email interactions between two people years ago should not have the same weight as ones made more recently.

3) The representation should provide both raw data and impressions, such that the user can quickly ascertain the value of the information or understand any of the impressions that are offered. By utilizing the value of an interactive multi-scaled approach, users can delve into the high-level fingerprints in order to understand how they are being constructed. Through interactive interfaces, the user should be able to get to the raw data from the higher-level impressions. The interface should be compelling and attempt to convey information as legibly as possible.

While i can proselytize such ideas, i have only hypothesized what such a design might look like. The hardest task in bringing these ideas to fruition is that of creating a comprehensible design. Such a system should draw from the developments that have been made by researchers who design visualizations that represent people's behavior post-facto. Simultaneously, further research is necessary to determine what data is appropriate to convey and what better mechanisms exist for making it accessible.

To do so requires strides in two directions – analyzing the appropriate data and representing it in a meaningful way. The efforts made in *Loom2* make it quite clear that this work has hardly just begun. Presenting statistical data in a key-based accessible manner is not that difficult; people can learn to read keys and evaluate the numbers. But numbers do not complete the picture; people need to derive meaning from their environments beyond what is easily computed. Yet, evaluating qualitative data requires such delicacy in order to pull out the desired impressions. Then, once the data is available, conveying those impressions requires yet challenge, as design is not a systematic art. For example, what is more helpful – abstract statistical representations or potentially inaccurate caricatures? How much abstraction is meaningful? What is an acceptable margin of error for conveying impressions? The goal is to provide a visual tool that requires little more than a glance to get a meaningful impression, but that also offers an interface for extended detail.

Concluding thoughts

Awareness online need not resemble its offline counterpart, as the available data is not comparable. By providing awareness online, the goal is not to mimic offline knowledge, but to supplement the dearth of available digital feedback. In doing so, people can feel more settled by understanding how they are seen, even if they cannot determine the reaction of others. Even this level of awareness increases one's ability to appropriately self-monitor.

Self-awareness allows users to understand who they are in a particular environment, how facets of their identity are manifested and aggregated, how other people and sites can see them. Such awareness places an individual within the society at large, in relation to other people. While awareness allows users to begin controlling their presence, it is only the first step. Awareness alone is not effective in giving individuals control; they must also have the ability to instigate change of how they are perceived by having the tools to manage their presentation directly.

Chapter 5:

DIGITAL IDENTITY MANAGEMENT

When we present ourselves to others, we want both awareness and control over what and how we are presenting ourselves. Without the ability to manage the aspects of the self that we perform, awareness is simply a reflective exercise. To truly empower people's interactions online, they must have the ability to manage the impressions that they construct, the information that they provide. Yet, giving people these abilities is a challenging design task.

Offline, interactions have an ephemeral quality. While initial impressions certainly impact all future negotiations, data is not persistent. Thus, future interactions are only impacted by memory-driven impressions, not by the constant reemergence of previous interactions. Online, social data is quite persistent. Thus, it is not only the initial impressions that matter, but also how well the data from previous interactions persists in an archived, out-of-context manner. While the ability to research someone online is quite valuable, finding a young professional's angst-ridden tirades from early teenage years is not necessarily valuable or appropriate in deriving an impression. Yet, with that data archived, the young professional has no way to eliminate that decade-old data and must always confront the impressions that it renders. Persistent interactions create immense challenges for identity and impression management.

I refer to both identity and impression management, because they are quite tangled conceptually, yet they cover separate ideas that must be considered. In social interactions, a viewer perceives both the identity information that one is conveying as well as more underlying information that strikes impressions on the viewer. While the former can include things such as one's occupation and political leanings, the latter is

much more difficult to tack down. People leave impressions on others simply in the way in which they smile. While people aim to leave specific impressions, they must rely on and react to the other's perceptions. *Impression management*, as detailed by Goffman (1959), is the negotiation of leaving and receiving impressions.

On the other hand, *identity management* is more concerned with the underlying structure of what's presented about the individual when making impressions. Identity management is the controlling systems behind impression management, as it is the facets of one's identity that one controls during presentations. Identity management is highly affected by the impressions that one leaves, but one manages one's identity regardless of those impressions. Impression management is completely tied to the reactions of others; without those reactions, there are no impressions.

While impression management is certainly crucial for identity management and for the construction of oneself online, it requires a level of awareness of others' reactions that is not currently possible online and is outside of the scope of this thesis. Thus, for the remainder of this chapter, i focus solely on identity management and address the resultant impressions when appropriate. The goal of this chapter is to discuss what, beyond self-awareness, is needed for one to properly manage their digital identity. I start by discussing why control is necessary online, introduce some of the current systems for digital identity management and then propose some of my own thoughts on this matter. This chapter prepares the reader for considering the issues raised by *SecureId*, an identity management tool for users that is addressed in detail in Chapter 8.

Why control? Why management?

As i discussed in Chapters 2 and 3, identity management empowers people to regulate their social behavior and engage in more meaningful social interactions. Between the persistence of data, the collapsing of contexts, and the marketability of their identity, people have very little say in how their identity is represented online. For this reason, people desire the ability to manage and control their presentation.

By lacking even basic control over the system's abilities, many people feel immediately disempowered. The market encourages both surveillance and profiling. Online, people cannot access many services without submitting to the profiling requests of corporations; if they do not agree to the terms of service, they have no mechanism to dissent and still

utilize the systems. The data that they provide to one service can be bought and sold, where the terms of service are changed and implemented with no form of recourse by the user; they do not own the data that goes through other people's servers.

Yet, users choose to use these systems because they provide a service that people see as valuable. Without realizing how valuable their data is people are willing to sell it in return for what appear to be free services. Yet, this reduction in privacy awareness and automatic protection of data is precisely what worries many privacy experts (Rosen 2000; Garfinkel 2000; Lessig 1999; EPIC 1994). While corporations are more rigorously requesting profiling data, and privacy experts attempt to educate the public, online participants are working within the systems to provide what they believe to be anonymous or falsified information. The current environment encourages anonymity and deception by users who seek out privacy and have no other method of access.

People lack control because the architecture makes it easy for the market to seize access to such people's presentations, as assets owned by the companies. People lack control because they do not realize how valuable their information is, what they are giving away, or how corporations use the data to profit at the expense of individuals. People lack control because they are not aware of their own presentation, let alone understand what it would mean to have the tools for control. Yet, the people's naïveté is not an excuse for the abuse of their privacy. While the law curbs the most egregious abuses of data control, it will not provide the level of protection that users need to develop a rich social environment. Thus, it becomes the responsibility of designers to consider the needs and interests of the users and construct environments that provide them with the ability to control their data and barter it at will, not on demand.

The value of identity control is not simply autonomy and freedom, but it is the underlying structure necessary for people to develop rich social environments. Lacking the ability to manage one's presentation in a faceted and contextual manner, anonymity will remain the only option for those who seek control. Even the mechanisms by which people create context discussed in Chapter 3 only provide a temporary bandage over the growing wound in individual control.

Regulation through federated identity

When online users attempt to regain control of their identity online, they do so through anonymity or multiple accounts. Although these mechanisms create control for some, they also provide an environment in which fraudulent behavior, harassment, hate speech, abusive deception, and other less desirable qualities of society can flourish. Not surprisingly, corporations are seeking accountability, if for no other reason than to eliminate the fraudulent abuses that are costing them economically.

Seeing these abuses as intimately tied to the ability for users to have anonymous and multiple digital personas, there has been a recent push for genuine authentication combined with the elimination of multiple logins. With proposals such as Microsoft's *Passport*[2] and Sun's *Liberty Alliance*[3], corporations are drumming up support for single login systems as a mechanism to end abuses and ease the hassle that users experience by maintaining numerous accounts. While many of the intentions of these systems are admirable, they not only ignore privacy issues and put users at notable risk, but they also fail to accommodate the need that users have for controlling their own data and representation. Without serious design reconsiderations, such systems run the risk of providing the ideal digital Panopticon, where an authority figure is able to observe every action of all individuals without them knowing what is being observed, when or for what purpose. As Foucault (1995/1975) recognized, such structure provides external discipline

---

[2] Various information on *Passport* can be found at:
http://www.passport.com
http://www.microsoft.com/netservices/passport/
http://www.epic.org/privacy/consumer/microsoft/
 (All sites live as of August 1, 2002)

[3] The *Liberty Alliance* website can be found at:
http://www.projectliberty.org/

and control out of fear. Such an environment is not advantageous to social interaction, particularly for marginalized individuals.

Considering Microsoft's Passport

In order to reflect on the design issues of these systems, consider Microsoft's *Passport*. As the name implies, this system is designed to provide a singular access point to many sites on the Internet. Yet, as is poorly indicated through such a metaphor, Microsoft maintains the information in one's *Passport*. When the user creates a *Passport*, they are asked to provide traditional corporate profiling information: name, email, sex, occupation, income, postal code, etc. In order to gain access to the federated sites that have integrated *Passport*, users must provide the site with their *Passport*. When authenticating the user's login, the site can also access the profile information that Microsoft has collected about the user. The site may then link this information with its own database of information and provide the content to its advertisers. As it appears from the technical notes on *Passport*, Microsoft does not currently receive any of the information that other companies collect about the user. In addition to the profile data that Microsoft maintains, *Wallet*, which is a component of *Passport*, maintains encrypted credit card information about the users for their ease of access. The metaphor of this is also noted, as one does not hold one's *Wallet*; Microsoft maintains it for the user.

While any site can pay to join the *Passport* authentication system federation, many of the sites that require *Passport* are Microsoft's, and not just those that focus on e-commerce. Microsoft's *Communities* portal, which provides users with *Hotmail* email access, chatrooms, message boards and instant messenger requires users to authenticate with *Passport*. As these technologies are the basis for many people's digital experience, Microsoft can easily associate one's profile data with one's social network, IP address, login habits, and other data. Therefore, regardless of its connections with other sites, Microsoft maintains most of the valuable data about one's digital presentation.

As users can only be logged in to one *Passport* at a time, it is not simple to maintain separate *Passports* for separate application contexts. This is magnified if users want to regularly access their instant messenger or email, applications which users tend to leave running throughout the duration of their connection. With the latest version of Windows, users are limited in what applications or information they may get if they are not logged into their *Passport*; thus, upon their initialization of Windows XP, they are

actively encouraged to create an account. As this information is integrated in both social environments as well as commerce ones, certain information cannot be hidden, as the user is unlikely to purchase a book and have it shipped to a false address because of their desire to maintain privacy. Thus, by requiring the user to provide certain accurate and authenticated information in the commerce environment, they are bound to convey the same information in the social environment, regardless of its potential impact. Such a scheme provides Microsoft and their collaborators with a system that practically requires users to provide authentic data.

   Just as a person can maintain passports for each nationality, a digital individual may currently control multiple *Passports*. While multiple logins provide users with the ability to present the proper form of identification in the proper scenario, they raise some of the same questions as their physical counterpart. When is it appropriate to provide which passport? Once you enter a country with one passport, you must use that one throughout the duration of your stay. What happens when aspects of that passport are considered socially unacceptable? Why can you not travel on multiple passports at once?

   Such hassles limit the number of *Passports* that users are motivated to maintain, as it is quite inconvenient to have to log off of IM in order to check an email account that is associated with a separate facet. Therefore, only the highest self-monitors are likely to maintain these distinctions, just as some of the few cell phone users who maintain separate SIM cell phone cards are gay men (Green, et. al. 2001) and business men who work in both Hong Kong and China (Bell, 2001). Those with the greatest risk recognize the social and personal consequences.

   Although managing the separate *Passports* is a nuisance, it does provide a strict boundary between two different facets of one's identity. At any given time, an individual can only be presenting one facet. Such separation allows for the strict separation that employers desire, so as to limit their employees from surfing and checking personal email at work. While this separation more accurately mimics physical life behaviors and an employer's ideal situation, it is not in synch with the typical user's behaviors, as most users are frequently managing unrelated interactions simultaneously.

   While it is possible to maintain and manage multiple *Passports*, this is not encouraged behavior. With security-driven screams for a national ID, both in the United States and abroad, and an increased desire for authentication, it is it is quite reasonable to assume

that it will not always be possible to separate one's identity online. The designers at Microsoft certainly recognize that a system such as *Passport* is a valuable way to curb unacceptable online behavior, yet they fail to acknowledge that they are also upsetting certain types of beneficial social behavior. With a uniform *Passport*, a sociable user is required to choose one of two values for "gender" – male or female. By default, this marker is accessible to anyone with whom the individual interacts, regardless of the social setting. As i discussed in "Sexing the Internet," this alters the social realm by sexualizing the environment and creating unnecessary expectations, built on poorly constructed mental models drawn from coarse data (boyd 2001). While intended for aggregate use only, even the Federal Trade Commission (2000) recognizes that online profiling must be addressed. The limitations of profile data, particularly static and uniform profile data are one of the weaknesses of a system such as *Passport*.

Perhaps the most problematic impact of *Passport* is that it eliminates the user's context replacement without providing a reasonable alternative. Although users can create multiple accounts if they feel the pressure to separate their facets, this system magnifies the difficulty in doing so and does not help provide the contextual information and separation that users are seeking to recover. By creating a uniform login across multiple sites, *Passport* furthers the collapsing of contexts. Prior to *Passport*, advertisers might have guessed when users from different sites represent the same user, often through IP address matching or connected email addresses. With *Passport*, Microsoft does not even need to collect all of the data for it to be collapsed outside of the user's control. *Passport* requires that users have the same login name for all of the different sites. Thus, any information recorded in the cookies for a given login is guaranteed to be the same individual; collection of mass data becomes quite a bit simpler.

Corporate control of personal data

In 1965, worried about potential unethical abuses of a national databank, the United States Congress decided to not pursue a National Data Center until individual privacies could be guaranteed (Garfinkel 2000: 14-15). With that decision, and the privacy regulations that unfolded in the 1970s, the United States made it difficult for the government to collect and maintain integrated records on its citizens. Yet, there are no restrictions on what the private sector can collect. While government agencies and credit bureaus are required to publicize their algorithms for computing scores and provide users with a mechanism for disputing the data kept on the individual, the private sector

has no such regulations. Corporations do not need to make available the data that they have collected, nor the methods by which they evaluate their users. They do not need to address users' disagreements nor do they need to change inaccurate information. As long as the fine print reminds users that their accounts can be terminated at any time, for any reason, corporations can deny service without even offering an explanation (Scheeres 2002). Since users have no alternative to these contracts, they are bound to a set of unregulated restrictions that rely on a set of values that are at the whim of the site.

   Users lack the recourse options for dissenting to contracts or challenging the data about their behavior that has been collected. They also lack the ownership of their own data. When Google purchased the Usenet archives owned by Deja, they also purchased all of the Usenet content collected[4]. The content is a collection of public statements made by individuals, yet those words were bought and sold without the permission of the users. Not surprisingly, users did not appreciate the commodification of their knowledge (Hauben 2002). In order to have their words removed from the archive, users must contact Google directly, either using the address from which the posts were made or otherwise proving their identity. Google promises to do their best at removing the data, yet they make no guarantees. Additionally, had an individual's statement been directly quoted by another user in the same thread, the individual has no recourse for removing that aspect of their content.

   Any site that collects data on users can sell that data without the permission of the subjects and the purchaser does not have to abide by the contracts that the user agreed to when they gave the original site permission to use their data. At will, sites my change

---

[4] Google Press Release:
http://www.google.com/press/pressrel/pressrelease48.html

the contract, sell the data, and deny service without informing the user. For example, when eGroups was purchased by Yahoo!, users were surprised to find that they were locked from their data unless they provided Yahoo! with a complete profile and agreed to a new terms-of-service agreement. Had the user declined to do so, Yahoo! still owned their data and the archives of their correspondences. In October 2001, various listserv owners were stunned when all of their archives and data were deleted; they were given no explanation nor any form of recourse; all attempts at contacting Yahoo! resulted in a lack of response. Even reflecting on the terms of service offered no explanation, as most of those affected could see no conflict there. It was not until a Washington Post article (Cha 2001) was published that these owners even knew why their data had been deleted – Yahoo! had declared them terrorists.

Perhaps the reader is thinking that they might have been terrorist organizations, and perhaps many of them were. Instead, i am inclined to believe that many victims of this abuse of data ownership resembled my own situation during this time. Out of the 20+ listservs that i moderated and 50+ listservs that i receive messages from, two of the most heavily trafficked listservs that i moderated on Yahoo! disappeared without notice in early October. Their topical content was identical, as they were both listservs intended for college and worldwide organizers of V-Day productions. A non-profit aimed at raising money for organizations working to end violence against women, V-Day and its associated listservs had two offending qualities: they conversed about helping women in Afghanistan and they used "pornographic" terms, as they raise money through productions of "The Vagina Monologues." Throughout October, my attempts to get an explanation were ignored. After the Washington Post article was published, i contacted Yahoo! again, offering an explanation as to what the organization was, what we did and why we were not terrorists. Although i received no response, most of my archives were reinstated within the week. While i was relieved to understand why my listservs had suddenly disappeared, i was horrified to realize how little control i had over the content that i managed. Not only could my access be taken away at a moment's notice, but also Yahoo! continued to own my data after they deleted my access, such that the data could be recovered when it interested them.

While the tech-savvy user has the ability to avoid using corporate services to host their data, no one is free from the impact that this control has. When a user sends an email message to a *Hotmail* account, Microsoft now owns that data on their server. When an

archiving system records webpages or Usenet posts, that system owns the data. Lack of control is about privacy as well as control and it affects everyone online.

Approaches to identity management

Lawrence Lessig argues that there are four mechanisms by which behavior can be controlled: the *law*, the *market*, the *architecture* and *social norms* (Lessig 1999). In Chapter 3, i dissected some of the underlying forces of the digital architecture and explained why the underlying architecture does not provide the means for people to enact socially normative regulation. In the last section, i introduced some ways in which the market regulates social behavior and personal identity. And while the law is only beginning to address issues of cyberspace, it is still entrenched in the metaphors between the physical and the digital, offering legislation and decisions that fail to acknowledge how the digital architecture is constructing a very different social environment.

In Lessig's model, regulation works best when the various forces are all operating effectively, yet this is not the case online. With the architecture dramatically affecting what is possible, social norms are often ineffective and the market is capitalizing on these changes while the legal community is not acting as though this space must be regulated differently. Although the law is already starting to impact what is acceptable usage (i.e. Intel vs. Hamidi[5]) and acceptable architecture (i.e. Napster), its approach to architectural change has focused on protecting corporate interests and copyright, ignoring individual interests and the underlying architecture. For example, when Napster was declared

---

[5] Intel charges Hamidi with trespass to chattels by sending distributed email to employees encouraging them to file complaints against Intel. Legal documents can be found at: http://www.eff.org/Cases/Intel_v_Hamidi/

illegal, it was forced to shut down because its architecture promoted the exchange of copyright materials. As a result, peer-to-peer networks were built such that no one could be held responsible. Thus, ISPs began to regulate their traffic and most recently, new technologies are being considered to eliminate the ability to copy music and other data. Additionally, bills in Congress (such as the CBDTPA[6]) are attempting to legislate architecture without an understanding of the architectural confounds. Thus, the legal impact has mostly been an impetus for system designers to work around the barriers that the law has created.

In any case, the legal approach will only handle the most egregious of incidents; it is up to designers to adjust the architecture to give people control. In particular, architects have the opportunity to create environments that promote self-regulation instead of relying on the market and law to develop or require such construction. In order for people to properly self-regulate, they must be able to manage their representations. Thus, designers must develop systems for identity management that authenticate users in a manner that does not also degrade their ability to control their presentation in a meaningful way.

1) In order to empower users, an identity management system should give the individual ownership over their data, its use and its distribution. In effect, people must own the rights to their words, thoughts and data. Copyright and intellectual property (IP) are not simply about the protection of registered artists and their managers, but the publication of the thoughts of all people.

---

[6] Consumer Broadband and Digital Television Promotion Act
http://www.eff.org/IP/SSSCA_CBDTPA/
http://www.politechbot.com/docs/cbdtpa/

2) The system should allow users to choose when and to whom what types of information should be revealed. Individuals should be able to develop and maintain the facets of their identity and have control over the contexts in which those facets are presented. Users should be aware of what can be seen about them and have the ability to adjust that information.

3) Users should have the ability to present the level of information that they perceive is appropriate. Systems should not require users to share personal data in order to gain access, as this allows for discrimination.

4) Users should have control over the redistribution of their data. If personal data is worthy enough for companies to trade it in return for free services, users should have the right to acquire those services at a price comparable to the value of their data and users should be compensated for the profits made from their data. No system should aggregate or distribute a user's data without their permission.

Certainly, these ideas are utopian in the current digital era. The architecture does not support such control; data can be easily transferred and copied such that having control over the data is near impossible. Yet, as companies develop technology intended to protect copyrighted material, these efforts should be appropriated to afford users the same level of protection as artists. Much of what is needed requires cooperation from the companies that so actively seek to profit from their sole control over a user's data. Thus, changes must come from the architectural level, with social and legal support.

In order for the architectural changes to be effective, they must be implemented at one of two levels. Either the foundation of the digital environment must be fundamentally altered to allow control over bits, or mechanisms must be placed on top of the current environment to regain control. Although the former is ideal, the latter can be implemented without the cooperation of most corporations. It is with that in mind that i designed *SecureId* as a prototype to consider the issues in building an identity management tool. As is discussed in detail in Chapter 8, the process of developing *SecureId* revealed the immense challenges that lie ahead in order to properly give users identity management tools. Although i stand behind the theoretical approach that i have outlined above, i realize that it is only embryonic, as much work is necessary both conceptually and functionally to provide users with the proper information.

Chapter 6:

EXAMPLE APPLICATIONS

Based on theoretical considerations, i have articulated some of what i feel is necessary to design and construct applications intended to empower individuals, most notably through awareness and identity management. Yet, my approach is predominantly theoretical in nature, based on observation and experience. In order to test these ideas and reveal the problems that they unveil, i have helped design and develop two different prototypes.

Social Network Fragments is a visualization tool that reveals underlying social patterns, most notably the social networks that evolve as people interact with others online through email. By providing users with a visualization of their habits, *Social Network Fragments* offers a unique view of otherwise obfuscated data. In this way, the system offers a level of awareness that is not typically available. In an attempt to provide users with a tool for identity management, i designed and implemented a prototype of *SecureId*. This system attempts to provide users with a way of controlling and managing their presentation online, through the management of facets and the information one might provide through such facets of their identity.

In the following two chapters, i switch from my theoretical discussion to focus on the issues that arose in the process of designing and implementing these systems. In doing so, i critique my own theoretical approach by recognizing why these problems are far more complicated than i initially suspected. For each application, i discuss the theoretical ideas that i intended to tackle in addition to the overriding goal of providing awareness and identity management tools. Using this, i highlight the most crucial algorithms that the reader needs in order to understand what the application provides. I provide usage

examples for each system through a set of screenshots and mockups intended to convey the output and interaction schema. Finally, i analyze the systems in reaction to their intended goal, providing critical responses to the actual results.

Both systems are simply prototypes, intended to explore these ideas and ground my theoretical ideas through practical experience. Thus, they are not provided as examples of ideal systems, but rather systems for in-depth critique and consideration. In effect, they are the critique of my theoretical ideas. In my struggles to design such systems, i convey why this problem must be more fully considered and why the theoretical notions conveyed in earlier chapters are only a framework for contemplation as sociable designers begin the process towards empowering users. Rather than conveying solutions, they provide fuel for future research.

Chapter 7:

SOCIAL NETWORK FRAGMENTS: A SELF-AWARENESS APPLICATION

Since its conception, email has been the most popular use of the Internet and those online regularly engage in sending messages between one another (PEW Foundation 2001; Harlan 2001). It is through this forum that people keep in touch with loved ones, coworkers, and digital strangers. These interactions reveal characteristics about the individual, including their social networks. Yet, this data is often obfuscated by the system, making it difficult for people to easily grasp the patterns and social interactions that they engage in daily.

Motivated by the depth of information that email provides, *Social Network Fragments* is interested in explicitly revealing the social networks patterns that emerge in email, emphasizing the structural forms of one's network and providing an interactive tool for people to reflect on their own habits. Understanding one's social network is quite important for awareness and empowerment. People manage their social network as one aspect of managing the context of their lives. Thus, awareness of one's digital network allows the individual to begin managing it online.

The purpose of this chapter is to detail the theoretical and practical components of *Social Network Fragments*, providing information as to its value in relation to self-awareness tools. As one of the applications chapters, it is intended to provide a detailed example of the process that we considered in developing a reflective awareness tool, as discussed in Chapter 4. By providing this detail, i intend to convey some of the challenges with which designers are faced.

As this project relies on the theories behind social networks, i begin by providing the relevant background material. Following this, i discuss the structural components of *Social Network Fragments*, including the system architecture and interface design. While discussing this tool, i critique our decisions and discuss the value of *SNF* as an awareness tool.

Background to social network analysis

Although personally constructed, one's identity is impacted by one's interaction with others. Many people have a variety of roles in an individual's life and therefore they provide a variety of impact, ranging from the stranger on the bus to one's best friend. Not only does the strength of an individual's connection to others play a role, but also the context, the value and a wide variety of uncontrolled events. These people help comprise an individual's *social network*, or the collection of people that the individual relies on for a variety of purposes. Although these people surround the individual, they may not all know, or even be aware of, one another. These holes in awareness or knowledge can be described as *structural holes* within an individual's network, where the only relationship that one person has to another is through the ego whose network is being considered.

The structure of one's social network conveys a great deal about an individual. How often does the individual maintain distinct relationships between groups of people? Do they have a few close friends or a large collection of less regular interactions? Are clusters within one's network separated based on roles (i.e. work vs. family)? While most people manage their social network with minimal effort, it is difficult to gauge the structure of one's digital network as the ego lacks the visceral experiences of shared space.

Considering related social network theory

Social network analysis is the study of the connections between people. These connections are valuable, because they are how people gather the different types of support that they need – emotional, economical, functional, etc. The types of connections – or *ties* – that an individual maintains varies, but they often include family, friends, colleagues, and lovers. In addition to a difference in type, ties vary in value or *strength*. Most commonly, social network theorists refer to two levels of ties – *strong ties* and *weak ties*, where a strong tie is able to offer a much greater magnitude of support than a weak tie. Although it may seem as though weak ties are not particularly valuable, Granovetter

(1973) shows that there are distinct advantages to having weak ties, including increased information flow and social mobility. Since weak ties require less effort to maintain, it is in an individual's best interest to maximize their weak ties, if they should want increased access to information.

   In most cases, an individual has great control over the structure of their social network. Although there are times and places when societies are so small or so tightly integrated that everyone knew everyone else, for many people this is not their experience. More likely, individuals will develop associations with people who are not even aware of most of the people in an individual's collection of acquaintances. These ties have a variety of purposes, and with each purpose, they have a difference in strength or importance. In some cases, a new tie might be neatly integrated in one's previously formed cliques. In others, that tie will be kept completely separate or only introduced to a limited number of one's ties. By controlling who knows who, an individual is able to explicitly manage their social network, providing connections as they see fit. When clusters of one's network are kept separate, a series of holes in the network develop, such that the individual becomes the bridge between the clusters; this is known as *structural holes.*

   Burt's structural holes argument (1993) builds on Granovetter's weak ties argument (1973). Burt argues that the advantages of weak ties are magnified for an anchor who is connected to different social clusters which have no other bridging connection. In other words, an individual who is the only person connecting one clique to another is advantaged. Not only does the individual gain from having access to a different set of information, they have the power to control what aspects of this information can be shared with the different social clusters to which they belong. Burt's discussion of structural holes is heavily motivated by the flow of social capital and the competitive advantages of controlling information flow. In his scenario, maximizing and controlling the flow of information is essential and empowering, such that an individual seeks to acquire as much information as possible. Therefore, by being the bridge between multiple social clusters, an individual maximizes their ability to acquire and control information. Although Granovetter argues that all bridges must be weak ties (1973), Burt rejects the relevancy of tie strength, but emphasizes that weak ties in bridges are more advantageous.

While Burt suggests that being a bridge is purely advantageous, Krackhardt (1999) argues that it is also constraining for the individual who acts as the bridge. In his analysis of cliques, Krackhardt develops the idea of Simmelian ties, where an individual is Simmelian tied to another if they 1) have a strong tie to one another and 2) share at a strong tie to at least one other person in common (i.e., they are part of a clique). Individuals who are members of a clique are constrained by the social norms of that clique such that Simmelian triadic ties are more constraining than simple, dyadic ties. Since each clique has a series of social norms by which its members are expected to follow, Krackhardt concludes that an individual who is a member of two separate cliques is constrained by the social norms of both groups, thereby needing to find the intersection of those norms in developing a socially acceptable face. Rather than seeing the bridging role as empowering, Krackhardt views it as a restrictive position, except in the case of private behaviors. In private scenarios, where only the particular clique and ego know about the behavior, the ego is advantaged by being the bridge, because they can act differently in different groups. Thus, if an individual seeks to maintain different social behaviors in different contexts, they become motivated to control social situations such that two cliques cannot converge, thereby guaranteeing private scenarios.

It is precisely these private scenarios that an individual desires when they want to maintain a multi-faceted individual identity. The individual produces their own identity information; therefore its initial flow comes from its creator and they control its initial recipients. Although trust and motivation plays a significant role in the passage of personal data, connections are also important. Regardless of trust and motivation, if information is passed to an individual with minimal ties, it is unlikely that the information will spread far. For this reason, one's social network is a considered factor when valuable private information is being shared.

From an individual's perspective, personal information is exceptionally valuable and therefore the individual wants to control its spread and content. The more valuable the information, the more closely the individual wants control. Should valuable information spread, it becomes gossip. Although individuals who are far removed are less likely to care to continue to spread the gossip, they are also less motivated to suppress its spread, as trust is less likely to override one's desire to spread information. In order to account for the potential of gossip, it is in an individual's best interest to minimize the ways in which it can spread. The most obvious mechanism is to only share information with ties

who are close, trusted, and have no motivation to share the information. Another effective approach is to minimize access by developing and maintaining structural holes. Structural holes provide security by 1) limiting the percentage of your social network that can learn any bit of information from other members; 2) increasing the number of degrees necessary for information to jump from one clique of associations to another. Although Milgram (1967) shows that few degrees are necessary to connect any one individual to another, by assuming that non-ties are less likely to continue the spread of gossip, increasing the degrees of separation effectively limits the passage of information.

Since flow of identity information can be more easily managed in a network with holes, it is in an individual's interest to maintain structural holes whenever possible, particularly when different cliques have different social norms. By being the only bridge between a set of work colleagues and a set of friends, an individual can portray two distinct social identities. Yet, once this faceting is started, it becomes more crucial that the structural holes are maintained. From Burt's information flow perspective, an added bridge simply weakens the power of the original bridge. When segmented identity information is involved, an added bridge can be considerably destructive for the ego, depending on the potential impact of revealing unknown identity information. In both cases, an individual is empowered by being able to act as a sole bridge between two different social clusters, although for slightly different reasons.

Some individuals instinctively separate many of their social clusters, if for no other reason than to minimize restrictions and maximize privacy. Just as Krackhardt noticed, when social clusters are bridged and Simmelian ties are built, an individual's behavior becomes constrained because they must follow the social norms of both communities simultaneously. Likewise, when an individual interacts with two cliques simultaneously, their behavior is effectively public, requiring a participation that will be appropriate for both forums. Although aggregated conformity might be expected for some individuals, Kilduff's (1992) earlier work implies that the impact of such convergence might be highly dependent on one's personal qualities, in particular their self-monitoring style. Because high self-monitors are quite likely to be influenced by their social surroundings, it follows that these individuals will be constrained when presented with combined social cliques with different norms. Conversely, as low self-monitors are less likely to adjust to social expectations, converging social clusters might not be so problematic.

It is important to note that converging social circles not only increase potential information flow and restrict acceptable behavior, but they also automatically increase identity information knowledge by making each cluster more aware of the individual's network. Should a clique be associated with particular activities or interests, others are likely to assume participation or interest. Depending on the difference in values and interests between the two groups, this may not be problematic. For most people, *homophily* alleviates this concern, such that any clusters that an individual might have are likely to be very similar to the individual, and therefore likely to be similar to one another (McPherson, et. al. 2001). Problems are most likely to occur when an individual maintains a cluster of people whose similarities to them do not overlap with the similarities they have to another cluster. For example, converging one's "anti-corporate/Marxist/activist" friends with one's corporate colleagues not only constrains appropriate behavior, but makes each group aware of the individual's involvement in the other.

Considering digital social networks

While most social networks literature is concerned with the physical world, Wellman, et. al. (1996) maintain that the same concepts are equally valid for those networks built and/or maintained in the digital world. While the theories remain the same, the ways in which people can manage and control their social networks are inherently affected by the strengths and weaknesses of the interaction paradigms possible when using digital tools. Additionally, the logged nature of one's digital interactions provides a more complete record of one's social network than is usually possible in the physical world. As such, researchers have shown that email offers great insight into an individual's social network (Garton, et. al. 1999; Wellman & Hampton 1999; Rice 1994; Sproull & Kiesler 1991).

Analyzing email spools can provide a great deal of information about the ego. At the most basic level, one can derive to whom the ego speaks, how often, how much, and including which other people in the conversation. At a deeper level, one could derive what types of content are shared, what the differences in sending and receiving are, where the people are located and when references to real life events are made. These patterns are quite rich and can be used to say a lot about their authors. At the same time, they are not perfect – not all conversations happen online and some of the most frequent conversationalists may not be the closest ties in one's network. While social network analysts should not want to use email as the sole source for understanding an

individual's behavior, this information is quite useful to the actual ego as they are quite able to separate out why some people are more prominent than others. As a result, the data still stands to convey rich information to the ego.

The social network of most people is quite large; manual studies of social networks have found that people average approximately 1500 ties of all different strengths (Killworth, et. al. 1990). Because of the ephemeral nature of people's connections, there are even more ties documented in email; many digital connections are so tangential that offline researchers would not even consider them. The quantity of ties impacts the dimensionality of one's network, because rarely do people maintain social networks where all members of their network are unaware of all others. Instead, there are many different types of ties between the different members of one's network. By simply trying to imagine what the graph of such a system would be, it is easy to realize that this largely dimensional dataset is quite hard to comprehend. In response, *Social Network Fragments* seeks to make this information accessible through an interactive visualization.

Building *Social Network Fragments*

After considering the theoretical concepts introduced above, i recognized the value in making digital social networks accessible to people, for their awareness as well as management. So that they may consider the impact of their network on their identity, i wanted to create a system that would reveal the structure of their social network. To do so, i began collaborating with fellow Brown University alumnus, Jeff Potter. We both recognized the power of visualizing largely dimensional structural data through spring systems, as i had previously worked on Judith Donath's *Visual Who* (Donath 1995) and he had worked a spring-based visualization tool emerging from the *Memex* project at Brown (Large Scale Design GISP 1998; Simpson 1995).

Using much of his original code, Jeff reworked the spring system to provide a layout algorithm for this largely dimensional dataset of email connections, based on a set of weighting systems that we determined were appropriate to numerically describe the relationship between any two individuals. Built on top of Jeff's layout algorithm, *Social Network Fragments* consists of a visualization tool that allows users to interactively explore their data, accessing different clusters and see the data over time.

In this section, i begin by explaining the technology behind *Social Network Fragments*, including the input that users provide, the relational weighting we devised, the spring system that Jeff created, and the larger interface that we used to give users access to their data. Following the technological overview, i discuss the results by considering an example dataset provided by "Drew."

Data Input

When a subject offers data for us to visualize, they must first provide the information necessary for Jeff to analyze the messages for time/date/sender/receiver, evaluate the relations between people based on the relationships discussed in the next section, and pre-compute the layout. Ideally, such a system would recognize when two email addresses belong to the same individual or who belongs in which facets of the individual's identity, but such is not currently the case. Thus, participants are asked to provide us with a set of four files that indicate the potential contexts and their colors, which email addresses should be associated with the subject, which email addresses are actually listservs, and a collapsing of all email addresses associated with any given individual. In the latter three files, subjects are encouraged to associate particular email addresses/listservs/people with particular contexts, which will affect the coloring in the system. The information that the subject provides helps the system more accurately determine the relationship between people, not just between email addresses.

Relational Dataset

At a fundamental level, our first priority was to determine the strength of the relationships between the different people in one's network. Traditionally, people talk about strong ties and weak ties, but we are attempting to analyze one's ties computationally, without any feedback from the user. Thus, we set about to categorize the different types of ties that exist in an email spool.

Knowledge ties. We assume that if A sends a message to B that A 'knows' B. (We do not assume that B knows A; we also do not assume that A knows B if the message went through a listserv.)

Awareness ties. We assume that if B receives a message from A that B is 'aware' of A.

Weak awareness ties. If B and C both receive a message from A, we assume that B and C are 'weakly aware' of each other.

List awareness ties. If B receives a message from A through a listserv, we assume that B is 'listserv aware' of A.

Trusted ties. If *A* sends a message to *B* and blind carbon copies (*BCC*'s) *D*, we assume that *A* 'knows' and 'trusts' *D*. We assume this because *D* has the ability to respond and reveal that *A* included people without *B*'s awareness.

We assume that most senders do not distinguish between the *To* and *CC* fields so we treat them identically (referred to as the *To* field from this point forward). We also assume that if no one is in the *To* field and everyone is *BCC*'ed that privacy is assumed and that there are no trusted ties. As this system only analyzes messages sent to the subject, we only know about the people that the subject *BCC*s and the people who BCC the subject. We do not know of anyone who might also have been *BCC*'ed on a message.

Example. Consider the following message:

    From: Drew
    To: Mike, Taylor
    BCC: Morgan, Kerry

This produces a set of ties as follows:

    Drew knows Mike; Drew knows Taylor; Drew knows & trusts Morgan; Drew knows & trusts
        Kerry
    Mike is aware of Drew; Mike is loosely aware of Taylor
    Taylor is aware of Drew; Taylor is loosely aware of Mike
    Morgan is aware of Drew; Morgan is loosely aware of Mike and Taylor
    Kerry is aware of Drew; Kerry is loosely aware of Mike and Taylor

We maintain bi-directional links in order to understand the strength of ties. Using this information, we construct a matrix of ties, including the quantity and type. Each type of tie is given a level of importance, such that trusted ties are more valuable than knowledge ties, which are more valuable than awareness ties, etc. Each value is additive, such that the more that two people converse with one another, the greater the value of the tie that connects them. While we found a set of values that seem to apply to many datasets, these constants can be altered depending on how much they make sense for a given subject. For example, if one consistently *BCC*s people for a reason other than privacy, it is foolish to overvalue trusted ties.

The relationship between any two people is given one numeric value, ranging from 0 (no connection) to 1 (most strongly connected). Time is divided into two-week intervals. For each time slice, the relational value is determined based on the additive value of each type of tie in relationship to its importance. It is done by time slice so that the subject can see as people begin connecting to one another. The weights are scaled across all people (except the ego/subject) and over all time such that each type of weight has a separate scaling. As the ego exists in a large percentage of the ties, the ego/subject's weights are scaled separately just over time; otherwise, this would devalue all other weights to near zero. These weights also affect the visual properties of individuals. Color is determined based on the context in which the individual knows both the subject and the other people in the network. Personal contexts override listserv contexts, which override email address contexts. For example, consider the following message, where Drew is the subject and is writing from his WORK email address:

From: Drew
To: Mike, Taylor, Morgan

Assume that the following people have been overridden with particular contexts:

Mike: COLLEGE
Morgan: FAMILY

In such an example, the following weights would be used to determine the coloring:

Drew: 1 COLLEGE context from an awareness tie; 1 FAMILY context from an awareness tie; 1 WORK context from an awareness tie (because Taylor is not overridden)
Mike: 1 COLLEGE context from an awareness tie; 1 FAMILY context from a weak-aware tie; 1 WORK context from a weak-aware tie
Taylor: 1 WORK context from an awareness tie; 1 FAMILY context from a weak-aware tie; 1 COLLEGE context from a weak-aware tie
Morgan: 1 FAMILY context from an awareness tie; 1 COLLEGE context from a weak-aware tie; 1 WORK context from a weak-aware tie

Spring System

In order to determine how an individual is geographically positioned with respect to the others in the system, Jeff implemented a simple spring system that reacts to a combination of forces that pull and repel different anchors from one another. First, there is a gravity force that pulls all nodes towards the center of the graphical world. Without the gravity force, island nodes would expand infinitely. There is also a repulsion force

that makes all nodes repel from nodes that are within a certain distance, which also means that things do not actually expand infinitely. These two forces result in a system that only considers repulsion and gravity forces, such that without any other forces, the system would settle into an evenly spaced circular ball of data. The constants that determine these two forces are dependent on the size of the dataset, and are tweaked to magnify the structural features.

   In addition to the default repulsion, each pair of nodes/people is assigned an attraction spring based on the weight of the ties between the two people. When two people are not tied, the repulsion spring is the only force operating; when people are strongly tied, the attraction force outweighs the repulsion force. Although the attraction forces are directly mapped to the strength of the tie, the repulsion constant is altered to be appropriate for the given dataset. As clusters of nodes form, their aggregated repulsion force further repels unrelated nodes.

   At the start of Jeff's pre-compute system, all nodes are randomly positioned in the geographical world. The system begins stepping through a series of iterations in order to find a layout in which most nodes are relatively settled. For each iteration, the nodes assess the system of forces that are impacting them (gravity, local repulsion, connected node attraction) and determine what an ideal geographic position would be given those forces. As moving to that ideal position assumes that all other nodes would be staying in the same position, each node only moves a fraction of the distance in the direction of its ideal position. This procedure is then repeated. Over time, nodes settle into a position where the distance that they must move is so minimal that they begin to visually shake, because they are moving back and forth between two pixels, trying to find a position between the two. At this point, the system has reached a settled phase.
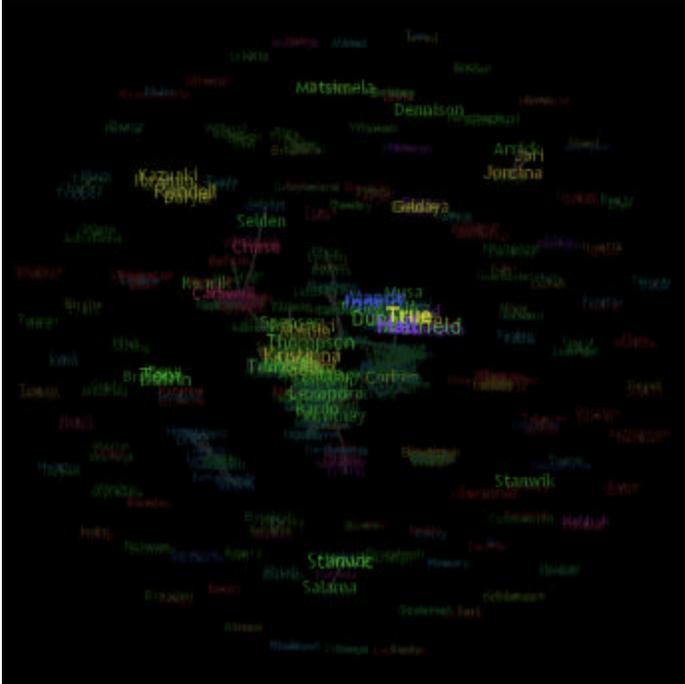
Figure 7-1. This image is derived using our example data provided by Drew. In the image, we see all of Drew's nodes in a position where they've reached a settled layout. The result resembles a galaxy, with various solar systems. Due to the initial random layout and the general centrifugal gravity, this algorithm frequently pulls the most highly connected clusters towards the center.

While the settling algorithm results in beautiful clustering of related nodes, it is still compressing n-dimensional data into a 2.5[7] dimensional space. As a result, some nodes are geographically close to one another, even though they are unrelated. Such scenarios occur due to a fundamental restriction in graphing highly dimensional data on a plane.

---

[7] 2.5D uses both X and Y coordinates as well as applying layers. Thus, the information appears to be laid out along a third dimension that cannot be navigated.
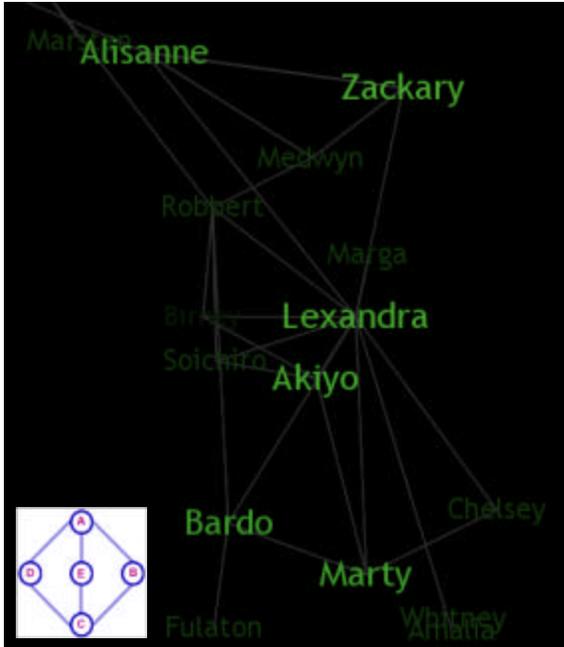
Figure 7-2. By focusing on one cluster in Drew's network, we can see the graph layout problem. In order to explain the problem, consider the inset.

While *E* is completely unrelated to *D* and *C*, they are just as geographically close as the two nodes with which *E* has strong ties. This occurs because both *A* and *C* are also strongly tied to *D*, *B*, and *E*. Additionally, the length of ties fails to accurate represent distance, as the length of the tie between *E* and *A* is much shorter than between *A* and *D* even though they represent the same strength.

In Drew's network, the problem is even more convoluted, as the dimensionality of connected nodes is greater. Such complication makes it difficult to determine who is really closer to whom.

This graph theory problem is inevitably a weakness of a system such as ours, which collapses n-dimensional data into a 2.5D space. While we considered doing a 3D version of this, we both agreed that the confusion that 3D adds does not outweigh the advantages of an extra graphical dimension, particularly since we are dealing with >500 dimensions of data.

Another dimensional weakness of our current spring system implementation is that it does not take into account time in determining layout. While we present the data to the user over slices of time, the layout algorithm assumes that all items will be shown simultaneously. By using two dimensions of space and one dimension of time in laying out the graphs, we would dramatically improve the visibility of the data.

User interface

The interface for *Social Network Fragments* is comprised of two frames, the *network frame* and the *history frame*. The temporal length of email interactions is divided into two-week periods such that every slice of the animation shows two weeks worth of data. When the system is initialized, the network frame shows the data from the first time slice and the history frame highlights the time slice that is currently being observed. The subject is not

shown in the network frame, but their overriding data is the default view for the history frame. Consider Figure 7-3 for a view of the entire interface.
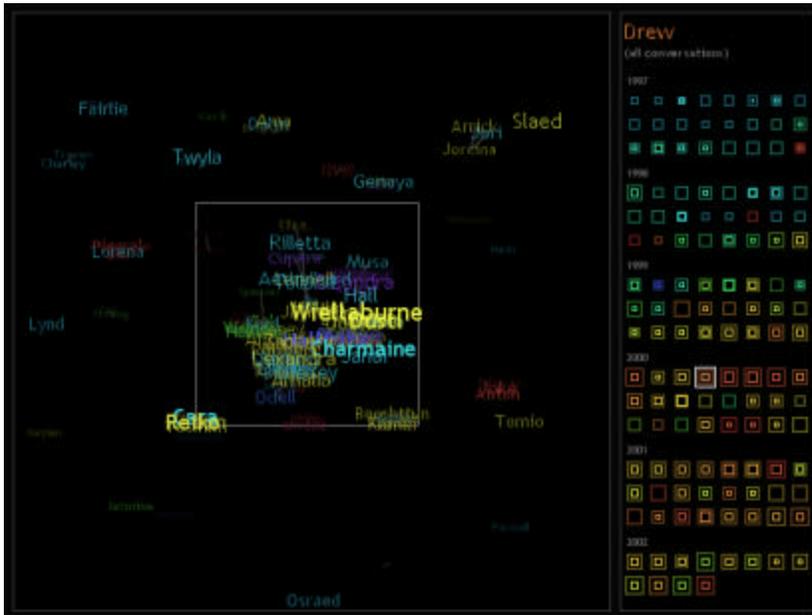

Figure 7-3. While earlier figures showed segments of the network view, this image shows the complete interface. Drew's data is obfuscated, but the structure remains in tact.

The left frame is the network frame while the right is the history frame. The history frame has the current time slice highlighted and the box in the network frame shows the region in which the user is currently zooming.

The colors represent the contexts in which Drew knows the various people, where yellow refers to her social friends, teal to her college mates, and red to her research colleagues.

Positioning. The aforementioned layout springs helps determine the positioning of the nodes within the network frame. Because integrating the settling into the system is too computationally intensive, we pre-calculate the position of each node. As i mentioned before, this presents a problem, such that people are not laid out with each time slice in consideration. This creates the appearance of unrelated people near each other without connections because their strong connection appears during a different time slice. Thus, a preferred mechanism would be to consider laying out each time slice separately. Yet, this also poses an interface problem, as it would mean that individual nodes would jump from one section of the screen to another, making it difficult for the user to see the continuity. To deal with this, one would want to show the traces of movement, so that the user appears to be moving from one cluster to another over time. For the layout to be more meaningful, time needs to be directly integrated into the geographical positioning.

Connections. Given a minimum weight for a time slice, a line is drawn between two people to recognize that they are related to one another. This allows the viewer to know if two people are truly connected or nearby due to a function of the layout mechanism. These connections are more clearly visible in Figure 7-2.

Coloring. The size and brightness of an individual relates to the magnitude of the person's weighted connections at a given time slice. Color is predetermined based on the context and weight of all relationships at that time. The ego's color is also a general weighting of the various connections for each time slice.

Zooming. By selecting a region of the network with a mouse, the system slowly zooms into that area, holds, and zooms out. Further zooming or a mouse down will stop the system from zooming out. A square is drawn around the desired region so that the user can see what they are aiming towards. Our zooming mechanism can be confusing for users because there is no easy way to pull out of a given zoom or shift over without waiting for the system to start zooming out.

Time. As time marches progresses, people emerge and become connected with other people. In the current system, time marches at a speed of one day per second. There is a minimum weight for an individual to be shown in a given time slice, although they are phased in during the time slice before and phased out in the one following the slice when their weight drops below the minimum. The history panel serves as both a system clock as well as a temporal overview of the entire dataset. Each square region represents a time slice of two weeks, the default time period, where the oldest areas are at the top of the screen. By clicking on a square, the current time is adjusted to the starting time of that time slice.

Time overviews. Inside each time slice in the history panel are potentially three smaller squares. The outer square represents the weight for all awareness ties; the middle square represents the weight for all knowledge ties; and the inner square represents the weight for all trust ties. While these squares are relative to the scaled weights, the meaningful value is in their comparison; no square is drawn if there is a zero weight.

History view. The history panel shows an overview of the entire system. By default, the ego is shown, where all of the time overviews are the ego's relation to others. Yet, by

clicking on the name of someone in the network panel, the history panel reflects the historical interactions between the ego and the selected user, as shown in Figure 7-4.
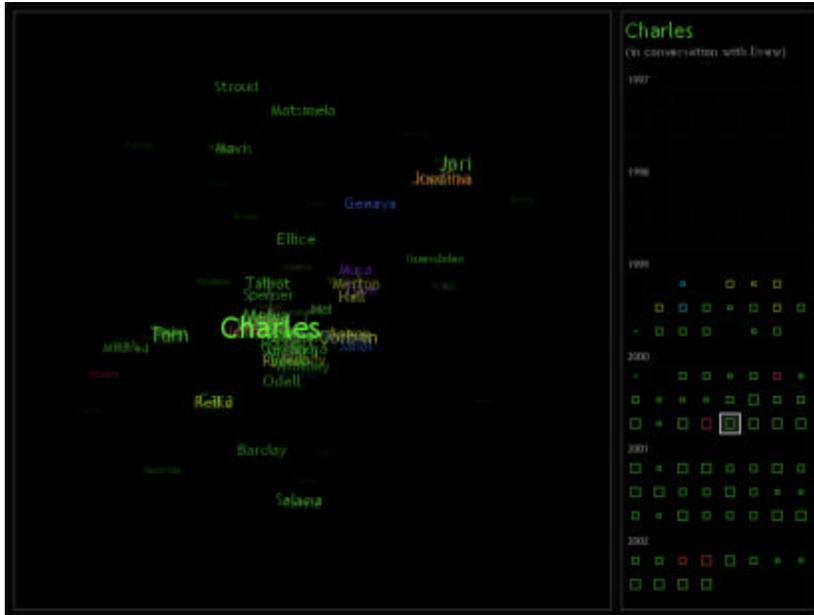

Figure 7-4. When a name is selected in the network panel, the person's name is highlighted and the history panel changes to reflect the relationship between the person and the ego. Thus, in this image, we see all of the connections between Charles and Drew, noting that they only starting conversely a few months into 1999 and that their interactions were pretty consistent over time.

The interaction paradigm that we used in designing this required only a mouse, as it was designed for exhibition. Such a paradigm limited the types of interactions people could have and our zooming mechanism was not intuitive. For personal use, we also added additional features such as a find box and key commands for immediate zoom and filters. Such interactions would be useful for those engaging with this interface.

Discussion

The awareness that people using *Social Network Fragments* seek regards their social network structure. Thus, the most compelling aspect of *SNF* is viewing the clusters that form and trying to understand the meaning that they have to the user. Clusters develop because of common ties; by looking at the clusters, the subject is able to have an image by which they can tell their story and the story of various encounters. Just like a photo album snapshot, these images are often more meaningful to the subject than the researcher. Consider one of the clusters that appear in Drew's network.
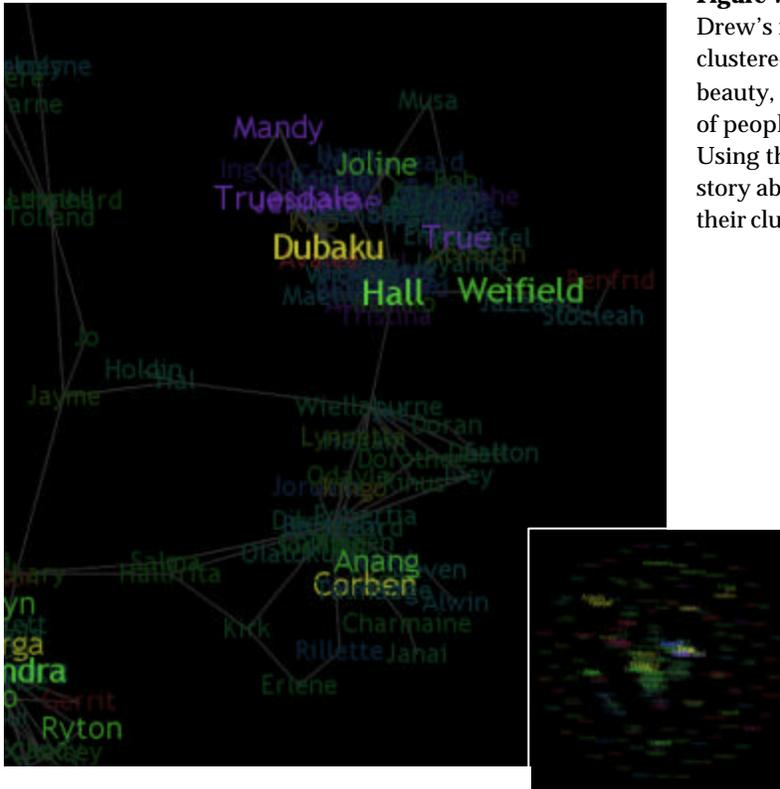
Figure 7-5. By zooming into a region in Drew's network, we can see a collection of clustered people. Aside from its structural beauty, this cluster represents a collection of people that are important to Drew. Using this image, Drew is able to tell a story about who these people are and why their clusters are meaningful.

Drew first noticed the colors. She had assigned purple to represent her activist friends, green to be associated with her collegiate context, and yellow as a default color for friends. Although it was not specifically assigned, blues result from people who Drews knows from both the collegiate and activist contexts of her life.

Drew is actively involved in both the collegiate and national organizations associated with her particular activist community. The cluster containing Dubaku consists of people who are associated with the national organization, some of whom know Drew simply through the organizational context, while others also know her through one of her collegiate contexts. Hall, the primary outreach coordinator is also in regular communication with one of the local productions that Drew directs at her college, primarily through the Drew's co-organizer, Wiellaburne. Although Drew is in charge of the collegiate organization, her primarily relationship with the people in this cluster is through the college. Some of these people are also connected to other college mates of Drew's, and the unseen network branches on the left lead to other facets of Drew's

college life. Another noticeable feature of this data is that, at this time slice, Drew is more heavily involved with the national organizers than with the local members.

While this anecdote may seem meaningless to the reader, the recognition of these relationships were quite powerful for Drew. The self-portrait provided a visual mechanism for her to recollect historical events and activities and to notice aspects of her communication that she had not previously realized. When i showed Drew the bridge between Hall and Wiellaburne, her initial response was shock, because she did not realize that she and Wiellaburne were the only people who communicated with Hall. After pausing, she explained that this actually made a lot of sense, as she could not recall anyone else who knew Hall. This type of reflection is one of the ways in which the ego can use these images to be aware of their own connections and that of those around them.

We created six personal portraits and each person who saw their images were able to share stories about particular clusters or connections. The data was not surpr ising, as much as it was revealing. By being forced to explicitly consider and reflect on the relationships that are taken for granted, people recognized that their interactions could be graphed and that reading this graph is meaningful.

Critique

In providing subjects with a tool to grasp the structure of their relationships, *SNF* provides a level of awareness about one's social network that is not normally available. Although this data is at the fingertips of all digital social beings, people rarely consider it. While pieces such as Viégas' *PostHistory* (2001) provide a compelling look at the statistical data of one's email interaction, *SNF* provides a much more qualitative perspective. Yet, this approach has significant weaknesses.

Evaluating ties. Although we have given serious consideration to the mechanisms by which we evaluate the value of a relationship, the impact of our numerical representation must be considered, as it only provides one perspective on the relationships described. Nowhere is this more obvious than in the evaluation of the BCC ties. While we have assigned them to indicate trust, that is not a universal use of BCC. Likewise, just because someone receives an immense quantity of messages from another does not mean that they are closer ties, yet our system assumes so.

Actual awareness. Even when discussing the notion of awareness, we assume that the person's browser reveals all of the *CC*ed people, yet this is not true; many people do not even know that the message that they received including many other recipients. Thus, assuming loose awareness can be inaccurate.

Layout. While the clustering is quite stunning and appealing, the design portrays misleading information as an artifact of the layout algorithm. As i discussed earlier, the collapsing of highly dimensional data into a 2.5D space presents a visual image that is quite misleading; many geographically close people are not actually strongly connected; it is an artifact of the algorithm.

The major weakness in systems such as *SNF* stem from our attempts to convey qualitative data in a manner that gives resounding impressions. By using computational evaluations to produce qualitative ideas, we are faced with both the problems of evaluating the data and conveying the impressions. In this way, the problems that we face in visualizing data in *SNF* resemble many of the problems that Hyun and i only began to address in *Loom2* (boyd, et. al. 2002). As impressions are so crucial for giving people awareness, the weaknesses of *SNF* indicate why this problem is so challenging.

Yet, while the weaknesses are many, the images still provide a valuable insight into an individual's social network structure. Seeing the structure of one's interactions for the first time is quite thought provoking, as it provides a level of insight and awareness to one's interactions that are normally inaccessible. This level of awareness not only provides a system for telling stories, but also for reflecting on one's behaviors and intentions through a compelling interface. By addressing this need, *Social Network Fragments* would be quite a valuable tool if we redesigned certain aspects of it to address the underlying flaws.

Chapter 8:

SECUREID: AN IDENTITY MANAGEMENT APPLICATION

In order to address the issues involved in identity management, i designed and constructed *SecureId*, a prototype for considering how people would manage their digital presence. My intention was to design the type of tool an individual would need in order to properly present themselves by controlling facets of their identity, associating data with those facets and controlling the access to them. Although a redesign of underlying architecture would be ideal for providing such tools, i focused on what would be possible on top of existing architectures. In the process of designing this system, i started exposing the challenges of digital identity management.

In this chapter, i introduce the project and discuss the ideas of the prototype through a series of screenshots and mockups. Although this project allowed me to explore what it means to have digital identity management, it also revealed the weaknesses of my system and the problem as a whole. For example, the lack of embodiment in digital interactions requires a level of explicit management that is far more complicated than the natural management that people take for granted in the physical world. In order to treat the concerns that *SecureId* raises, i integrate the problems that i encountered into this discussion. Finally, i discuss what is needed in order to properly address identity management.

Conceptual background

In order to effectively manage one's identity, users must be aware of their presentation, the contexts in which they want to share information and have the ability to control which people can gain access to their personal data. In order to give users this control,

SecureId focuses on three primary conceptual ideas – *identity awareness*, *facet control* and *knowledge-based security*. Combined, these three represent some of the crucial components that people use when managing identity on a daily basis. Thus, before discussing what *SecureId* does, i refresh the reader of these relevant ideas.

Identity awareness

As was discussed in Chapter 5, self-awareness of one's identity and presentation is crucial for managing oneself in a public space. In the physical world, individuals embody the agent that presents them to the public and thus they are both able to convey information comfortably and present a wide range of identity features. Not only does the body convey biological characteristics such as age, but people are also able adorn it with fashion articles that convey other aspects of their identity. In addition to what is written on the body, people are able to maneuver comfortably, presenting nuanced details about themselves through their facial expressions and body language. Online, this information must be explicitly articulated, yet most people are terrible at doing so.

Based on the memory of previous interactions, the individual also has a sense of what previous knowledge has been shared. Online, information travels in different paths and an individual is not always certain of what data is available to the other person during their interaction. Previous interactions have produced logs of data that are far more accessible than the ephemeral conversations of the past.

Identity management requires the awareness of both what the individual is presenting and what previous information has been shared. In other words, people must have a sense of what they are presenting to others. Personal awareness is one of the founding needs for contextualizing social interaction.

Facet control

People negotiate their presentation based on different facets of their personality. These facets are often associated with different roles or contexts in which people engage with others. For example, *graduate student* is one facet of my identity. Given this facet, i interact with some people based on this role and give away certain data about myself to anyone who knows this facet of me. In my daily life, i use this facet of my identity whenever i enter my laboratory or whenever i present myself at a conference. Yet, in the

digital arena, contextualizing this facet of my identity is not as simple. I may have my email address and website associated with my role as a graduate student, but they can be easily combined with the other facets of my identity. As was discussed in Chapter 3, facets can be collapsed online and thus people must negotiate new mechanisms for contextualizing the facets that they present.

  Identity management requires the ability to properly understand the immediate context and harness the appropriate facet to present an acceptable face for a given situation. Thus, people need the ability to manage their facets as a way of managing their identity.

Knowledge-based management

  Although some social signals are assumed to be universal, people also present coded signals that are only intended for those who understand their underlying meaning. By having such knowledge, the viewer gains more information about the presenter. Coded signals that are only intended for limited audiences are particularly common amongst subcultures (Hebdige 1991/1979), but they are a powerful way in which people manage their identity information in public.

  Fashion presents the most frequent place where such coded signals can be found. For example, someone may wear a T-shirt with a symbol associated with a particular musician, such as a Grateful Dead bear. If the viewer does not know what that symbol represents, it becomes meaningless and is just seen as another T-shirt. Yet, if the viewer can associate the bear symbol with the Grateful Dead culture, they can make assumptions about the music and subculture interests of the T-shirt owner.

  Such signaling is particularly common for people who fear potential risks for revealing their participation in particular subcultures, such as those who are considered sexual deviants. Throughout history, a wide variety of symbols have been used in to indicate one's sexuality and sexual desires, including pinky rings, earrings in the right ear, green carnations, pink triangles, and rainbows (Pink Zone 2002). In addition to these simple markers, an extensive set of meanings has developed around the use of handkerchiefs (de Moor 1997). Known as hanky codes, the color and placement of simple handkerchiefs are used by members of the BDSM community to indicate the type of sexual play that is desired. While these various markers are easily recognizable by other

queer individuals, the majority of the population is not aware of the coded meanings. Thus, the symbols provide a perfect set of knowledge-based identity markers.

In the digital world, fashion markers are much more challenging. Although fashion can be seen across homepages (Chiou 2000) and in the profiles that people present, those presentations do not offer the fluidity of clothing. Unlike their physical counterparts, digital fashion markers are focused on presentation, not sharing. While someone might see your Grateful Dead homepage, you cannot tell that the person observing your website is also the owner of a Grateful Dead homepage. Without that shared knowledge, the markers do not begin a shared experience unless the observer chooses to initiate a conversation.

Identity management capitalizes on shared knowledge. By utilizing shared knowledge, people can put forward a facet that will only be seen by those who relate to it. As such, people have to do less management because the markers manage themselves.

SecureId scenario

Based on the conceptual issues discussed above, i designed *SecureId* so that users could explore what identity management means in a digital environment. Although an ideal interface would allow users to ubiquitously manage their digital identity in the same fashion as their physical one, this is not currently possible. With the lack of bodies comes the challenge of managing otherwise natural presentations in a space that requires explicit behaviors. Thus, i chose to give users an interface that gave them the control while simultaneously forcing them to consider how unnatural such management is in the current digital realm. The prototype of *SecureId* was built so that users could feel the difficulty in explicitly managing the identity data that they take for granted everyday.

When designing *SecureId*, i focused on the three conceptual components discussed above. While it would be preferable for the data to be gathered as people interact, the prototype asks people to produce their own data. When using *SecureId*, users can add any type of data about themselves and they can control the level of knowledge necessary for someone to access it. The interface also provides feedback to the users about who can gain access, what facets exist and what information they contain.

What follows is a series of screenshots and mock-ups produced as our example character *Gaia* uses the system. In order to introduce the interface, i discuss the images and the interactions that Gaia experiences when using the system. The system is comprised of three main sections: creating one's profile, creating one's facets, and interacting with the data and others.

Profile Creation

Upon initializing *SecureId*, Gaia is asked to login and create a profile. The purpose of the *Profile Set-up* is to create a set of public and comparable information about an individual. It is this information that is used by the system to see Gaia and relate her to others.

Figure 8-1. In the Profile Set-Up, Gaia is given a list of potential profile fields to create. The only mandatory one is the *Public Name*, which is how others see her in the system. In this shot, Gaia has selected to make a new *Email Address*. She fills in the address and then chooses one of the images on the right based on the level of privacy that she seeks.

Gaia can choose to fill out as much or as little in her profile as she wishes. For each item in her profile, she must choose one of three different types of privacy: 1) public; 2) searchable & comparable; 3) comparable. Profile data that is public can be seen by anyone who sees Gaia. Data that is searchable can be found if someone searches for this information. Thus, if Gaia makes this searchable, anyone who looks for this email address can find Gaia's profile. Finally, something is comparable if is can be seen by the

comparison system. Thus, if Gaia chooses comparable for this email address, anyone who accepts people who are also part of brown.edu will see her. Data that is searchable is also comparable. All data put into the profile system must be at least comparable.

By creating comparable data, she can position herself in relation to others and start constructing the facets of her identity that fit. While people share the digital space with many others, it is hard to ascertain who is out there and who might have something in common. Just as two people with Grateful Dead T-shirts might recognize one another and initiate a conversation, having data that can be compared to others within the system gives users a point of interest on which to potentially connect. Likewise, it allows for people who have something in common to be initially linked. For example, when Gaia lists her email association with Brown University, she can be linked to Brown and thus everyone who is also associated with the university. It is in Gaia's best interest to put as much information here as possible, as others use this data to pass their comparison tests.
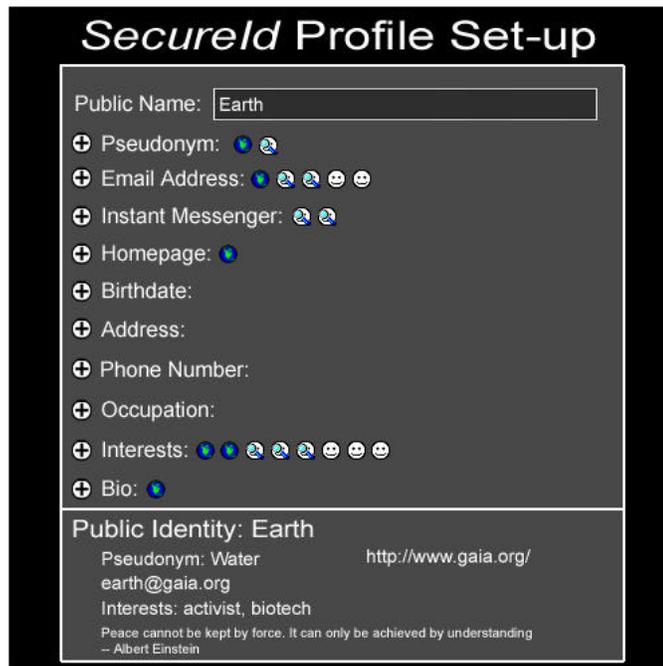

Figure 8-2. When Gaia's profile is complete, it appears as a set of icons that represent the different information she has given to the system. At the bottom, there is a panel that indicates all of the public data, for her awareness. It is this data that anyone who finds her or has access to her via comparable databits can see. Thus, the *Profile Set-up* reminds her that this data is accessible.

In the *Interests* section, Gaia was able to choose predefined interests as well as choose her own. Likewise, her *Bio* is simply a selection of statements that Gaia assocites with. In her case, she chose an Albert Einstein quote, but she could have chosen anything.

Most of the initial profile data only reflects one aspect of a person's identity – specifically, who they are as a unique individual and how they can be located. This data is not personal, but simply one mechanism to systematically differentiate people within an organized society. These are the types of data that one typically finds on a business

card. Had i added income and sex, this would look like a standard profile collected by marketing companies. Certainly, they say something about an individual, but what they say is not a complete picture, and certainly not the personal identity that an individual tends to self-construct. Yet, what is that picture?

The system also allows the user to add interests and a bio, but that is also quite artificial and difficult to ascertain. Given a person's website, we might be able to systematically derive more meaning about the person, but the majority of people either do not have homepages or use their homepages to present a professional image. Perhaps a picture might present more of the subtle details about a person, but what is an appropriate image for the world?

The *Profile Set-up* made me realize that ascertaining information about an individual is quite difficult. When someone walks through a room, they do not need to state their identities at the door; people perceive them immediately. Having to do so online is quite disconcerting, yet it is difficult to start managing one's identity digitally without any notion of who an individual is. Ideally, this section would not require a systematic approach by the individual. Instead, the *Profile Set-up* should be derived from all of the information that an individual does present online with this section being appropriate for editing. For example, one's email and instant messaging addresses can be found on an individual's system. Perhaps a more appropriate set-up should attempt to learn from the user's system and present them with what it finds, allowing them to alter the level of privacy that any data has.

Facet Creation

Based on what Gaia constructs in her profile, she is offered a series of potential facets to create. In *SecureId*, the *facets* are holders of information and data about the individual. Just as people maintain facets of their identity in their head, they are asked to articulate those facets in this digital environment. As facets relate to certain roles or associations that the individual maintains, they also operate as the context from which the individual presents aspects of themselves.

Figure 8-3. Based on the interests and associates that Gaia presented as she was setting up her profile, she is asked to create a set of potential facets. Each comparable databit – email addresss, area code, zipcode, occupation, interests, bio keywords, etc. – is considered by the system for potential comparisons to others. This list is determined based on the ones that the system recognizes as comparable.

She does not need to choose any of these facets as she can always create facets later; this selection list is just to give her a sense of what she can create immediately.
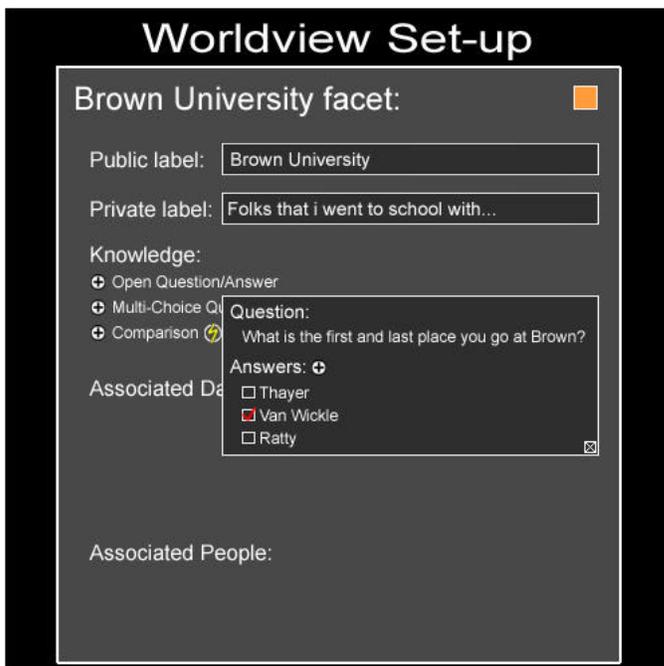

Figure 8-4. When Gaia chooses to create a Brown University facet, she is given a simple interface to edit associated information. The color determines how the facet appears in her world. The public label indicates how the facet is seen by others who might be trying to gain access while the private label is for her own consideration.

By default, a *Comparison* knowledge item is created, where the comparison is *@*brown.edu. Thus, anyone with a brown.edu email address passes the comparison knowledge. In this image, Gaia is making the facet even more secure by adding an additional *Multi-Choice Question/Answer* knowledge bit.

Associated Data / People are filled in as people interact with this facet; data appears here if Gaia chooses to edit this facet after use.

The facet creation stage allows Gaia to initiate facets based on the comparable data. When she chooses to create a facet, she is given initial information based on what the system has derived. For example, the *Public Label* is automatically created, as is a *Comparison* knowledge bit. Gaia can accept these defaults or make her own. She may delete the *Comparison* knowledge bit if she does not want anyone to gain access by comparison. If she chooses to add additional knowledge bits, the individual must be able to answer all of them to gain access. There are three different types of knowledge in this system: comparison, open question/answer, and multiple-choice.

Comparison Knowledge is done through regular expression matching. Thus, the system suggests a comparison to make. As noted above, a sample comparison might be *email=*@*brown.edu*. The prototype assumes that if someone adds an email address, it is a confirmed email address and thus anyone with a Brown address should be able to gain access. Comparisons are based on any of the information in the profiles.

Open Question/Answer Knowledge is also done through regular expression matching. The user can create a question and when someone answers it, the answer is compared against the answer(s) that are associated with the knowledge. When someone creates the question, they can put multiple possible answers so as to make answering easier.

Multiple-Choice Knowledge is simply done through a set of checkmarks. When someone creates a multiple-choice question, they choose which is the correct answer. When someone tries to get past that knowledge bit, they must also choose the same answer.

These three types of knowledge protect the facets. As an individual is creating the facets, they may choose to have as many questions as they wish. Thus, when someone is trying to get access to the facet, they must be able to answer all of the questions and have a profile that matches whatever comparisons are called upon.

Knowledge is an interesting way of protecting information, yet to do so online is cumbersome. While the comparison mechanism is quite useful, it requires the profile data to be accessible and complete. The two question/answer mechanisms force the user to explicitly state what they know and for the owner of the facets to figure out appropriate questions to guard the information. Digital knowledge is not nearly as unconsciously shared online as its offline equivalent; thus, it fails to provide the same level of value for assessing people.

The worldview

After creating a profile and initializing different facets, Gaia enters the *SecureId* worldview. Within the worldview, Gaia is able to see the landscape of her facets, who has access to different facets, and what information she shares within the different facets. From this space, she can also manage her world and explore other people's shared information. As such, this space acts as both a mirror, reflecting Gaia's self back to her, as well as a portal into other people's shared data.
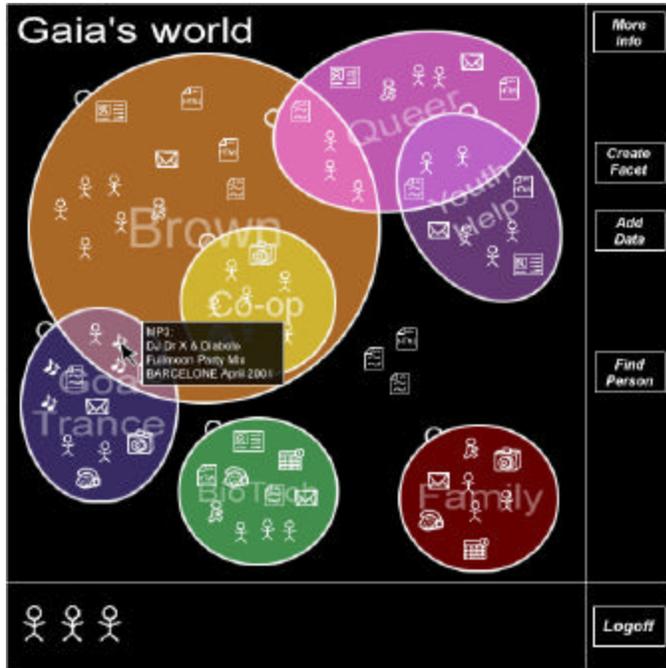

Figure 8-5. Over time, Gaia's world reveals lots of shared information and people; this is the view of her world that she sees. In this image, Gaia is maintaining seven different facets, where some facets have shared people and information. The Co-Op facet requires access to the Brown facet for its existence to even be visible.

Data that exists outside of a facet is public data while data in a facet is only accessible to people who have passed the knowledge requirements posed to gain access.

When Gaia highlights an icon, the information about that icon is displayed. The icons represent different types of data; the people represent those who have gained access to Gaia's facets. People who see Gaia or that she sees who do not have access to particular facets are shown in the bottom pane.

The worldview is a place for Gaia to adjust the presentation that she wants to give. The style for this interface draws from Viégas' work on *Collections* (1997), which is interested in designing an interface for people to manage collections of their information. In her work, Viégas was interested in defining different clusters of access to data and giving the user an interface to maintain access to these clusters. Facets operate as different groups where their knowledge structures their access lists.

Gaia can create new facets and define the data that exists in them. When Gaia creates a new facet, she is given an interface similar to that in **Figure A2**-4. After she creates the facet, it is placed on her worldview for her manipulation. She can enlarge it and move it to be placed where she sees fit.
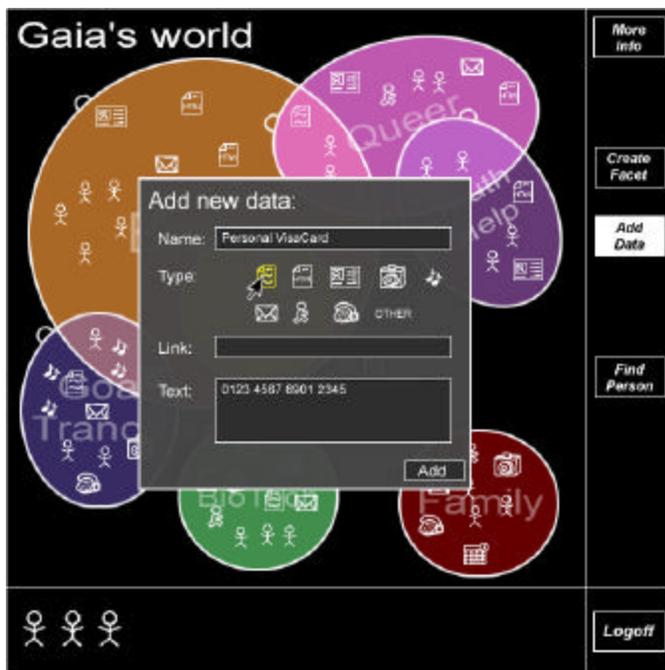

Figure 8-6. In *SecureId*, the user can choose what the different data represents. It can be pointers to information, or information itself. In the prototype, data is entirely text, but images and other media could be added. An icon can be chosen to represent the information.

Similar to facets, when Gaia creates new data it is placed on the screen for her to move or alter. Gaia can place new databits in any of the facets or in the public region. By double clicking on a databit, she can adjust its properties.

Data and people associated with a given facet stay associated with it upon movement or manipulation. Data can be removed from a given facet but the only way to remove people is to alter the knowledge locks that function as the guards for a given facet. When the user changes the knowledge bits, all people are dispelled from the facet.

Not only does the worldview function as the interface for Gaia to see her own data and who has access to it, but it is also from here that she can seek others. By clicking on a person that has gained access to her data, she can peek into their data. Likewise, she can search for a person based on known information. Both create a mechanism for her to gain access to others' facets.
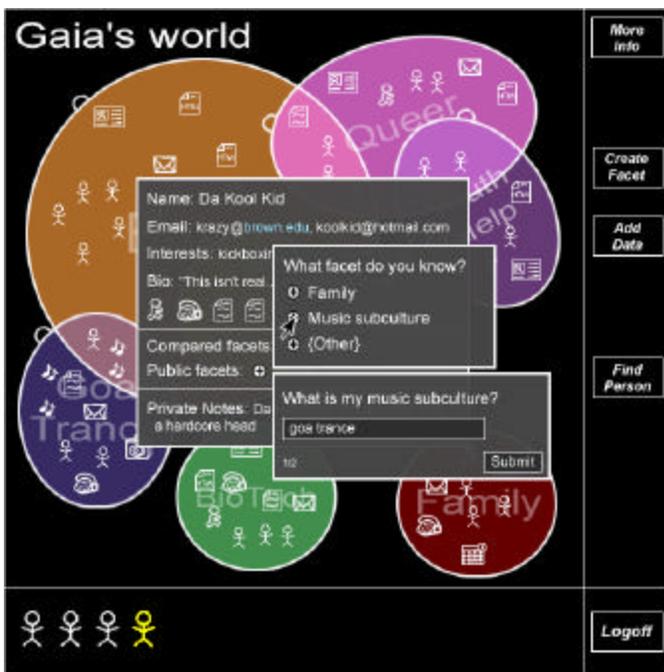

Figure 8-7. After searching for krazy@brown.edu, Gaia is given access to Da Kool Kid's public data and facets. As she knows this person as Damien, she is able to add a private note to remind herself of what she knows.

She is automatically given access to their shared Brown facet, which includes all of the files and bio information that she can see in the first section. There are also a list of *Public Facets* to which Gaia has not already gained access. By selecting this, she is given a list of facets from which to choose. Since she knows Damien through their favorite musician, she decides to gain access to that facet. Here, she is given 2 different questions that she must answer in order to gain access.

Once she gains access, more data is added to the profile she sees of Da Kool Kid. By selecting these icons, she can gain access to their data.

Just as Gaia can gain access to others' data, they can gain access to hers. They also see the public facets that she puts forward. When someone searches for Gaia and finds her, they see the public names for the facets from which they are not automatically exempt due to incompatible comparisons. Facets that are protected by comparisons that the user does not match are not accessible even for sight. Thus, if BioTech and Youth Help are

protected by such comparisons, the user will see Gaia's possible facets as Goa Trance, Family, Queer and Brown. They will not see the Co-Op facet because they must first have access to the Brown facet before that is made visible. Those with such access would see that as a possible public facet. When Gaia creates the names of her public facets, she must do so with care. Because they are explicitly named, they are made public. Thus, if she calls the Queer facet by such a name, its existence in a public list will reveal her participation in such a culture, which may not be what she wants.

When users attempt to gain access to a particular facet, they are only given one opportunity. Without such a limitation, anyone could gain access to any multiple-choice facet with a few tries. When the owner changes the facet's protection, those denied may try again. Likewise, the owner can explicitly place someone in one of their facets, thereby automating the access.

Discussion and critique

The interface designed for *SecureId* is clumsy at best. Users have to articulate detailed information to even proceed to use it. Gaining access to the facets of their friends requires a level of explicit sophistication that is cumbersome and problematic. Designing appropriate knowledge questions is challenging and users have difficulty accurately answering others' questions. Knowledge-based security is a desirable alternative to explicit access lists, but negotiating it is not similar to its physical counterpart. Likewise, organizing one's data is never easy, even if it can be easily placed into a given facet.

The explicit manner in which people must present and their identity is neither ubiquitous for presenters nor fulfilling for observers. Aside from data about themselves, people do not know how to present their identity. The subtleties of presentation are lost in such an explicit system and thus people are resigned to exist simply as a product of their output. Verbally and systematically articulating one's identity is quite challenging. People do not know how to present themselves from an external perspective; they simply know how to perform themselves from within their bodies. In order to present oneself online, one must step outside of one's body and describe oneself in a meaningful manner; this is not something that most people are fluent at doing.

Articulating one's identity and facets is quite disconcerting, as it requires a level of consciousness about one's interactions that most people do not maintain. Additionally, it

restricts the types of identity information that people can present as most components of identity are subtle and not part of what someone would consciously record. Identity information is not simply the construction of an individual's notion of self, but the relationship between the individual and the viewer. When the subject distills their identity into language, the viewer is not given the depth of information necessary to draw their own conclusions. Thus, explicit identity presentation also limits the viewer as they are once again receiving coarse data. The conscious control of information is cumbersome and limiting.

While the explicit nature with which one must articulate one's identity is a fundamental weakness of *SecureId*, the system does reveal the mental processes with which one normally construct social interactions. People are aware of what they are presenting to others offline, yet this type of information is often obscured online. Although it is often unconscious, individuals do have a notion of associating people with particular facets of their identity and assigning particular bits of information with those facets. *SecureId* requires that the individual be conscious about these practices. This consciousness, while irritating, provides a level of awareness that is not normally available.

Although the interaction paradigm for *SecureId* is fundamentally problematic for identity management, the results of such a system provide for some interesting reflections. First, it provides a level of awareness about identity management that most people do not consider. By having to articulate oneself, one has to consider what it means to present oneself as an individual. By having to distill one's identity into language, awareness is encouraged; at the same time, identity management can simply resort to data management about personal data.

Just as Viégas recognized in *Collections*, management of data is a challenging problem that is worth pursuing, as people want to control their collections of data and present it differently at different times. As identity presentation online is done through data presentation, some of same complications and advantages apply. Most notably, the facets that people maintain are quite similar to the situations in which people share different collections. Consider using a similar system as the access point to someone's website. Rather than being given the public page, the user is given a page associated with the knowledge that the individual has based on the facets to which they have access. Thus,

family members are given a homepage that is filled with family photos while those who know the individual on a professional level are given a website filled with a resume of previous work. The different homepages share both data about the individual and present their identity as a whole. Identity presentation can be done through situational facet-based data presentation.

Such an example is the type of goal that we must strive towards. While a tool such as *SecureId* can provide a manual mechanism for conveying information about oneself, it is too cumbersome to ever be useful. In order to be effective, a non-invasive design must be developed whereby people can manage their information without having to articulate it. It should provide feedback to the users and allow them to navigate with the least amount of effort possible. Explicit questions are not a desirable approach, but neither are explicit lists of who can gain access to a given set of information. While a knowledge-based approach is interesting in concept, implementing it online requires deeper thought. The approach that i took in *SecureId* is problematic simply because it is so restraining for users, both in articulating the questions and answering them. Instead, the system should learn from the user's practice, perhaps using the clustering work developed in *Social Network Fragments* to determine what facets exist and who should gain access to them.

In order for an identity management tool to be valuable, it must ease the amount of effort that an individual must invest rather than increase it. The system must develop an awareness of the individual and those with whom they interact. Not only should it automatically generate the data for such a system, but it should also begin to learn which people should gain access to what based on how the individual interacts with them and in what apparent contexts. The system should make guesses that simply allow the user to alter the assumptions.

At the same time, this still restricts the user to presenting data to convey identity. Unfortunately, this is a current limitation of digital social interactions. Conversations and impressions happen through text. Thus, identity management is derived from that text. This limits both what can be conveyed and what can be perceived, which inevitably makes identity management much more difficult.

In developing the prototype of *SecureId*, i realized that identity management tools highlight the fundamental differences between physical and digital social interactions. That which is so natural offline requires explicit consideration online. Yet, to do so is

unnatural. Not only is the unnatural element cumbersome, but it also limits the channels that people can use to present themselves. Developing a proper identity management system not only requires a deep consideration of how people can interact with the data that they use to present themselves, but how the digital environment can aid people in conveying subtle information in a meaningful way.

   While the ideas in this chapter and those embedded in the construction of *SecureId* address some of what is needed and challenging about developing identity management tools, they are only embryonic. Much is needed before identity management can be comfortably done online. It requires a level of ubiquity that is not currently available, nor designed. Explicit management provides new complications that affect the ways in which people interact with one another, thereby impacting all forms of identity presentation. Thus, the explicit nature of *SecureId* introduced new challenges that obfuscated the intended goal. The byproducts of digital interaction are even more heavily highlighted in identity management tools, as i have learned from my mistakes with *SecureId*. Thus, this chapter serves to articulate some of the issues that must be addressed in developing a more appropriate tool, but it does not provide the complete framework that one needs. An appropriate system must not only provide awareness but also make management easier, even if it will never be as natural as in the physical world. Yet, to design an interface that allows people to manage their identity comfortably is a challenge for future research.

Chapter 9:

CONCLUSION

Social interaction is negotiated between people with structural rules supplied by the surrounding environment. Depending on one's personality, an individual may adjust their presentation to present what they deem to be most appropriate in the situation. The individual draws from the situational and interpersonal contextual cues surrounding them, as well as by perceiving the reactions of others. As a result of this negotiation, people only present a segment, or facet, of their identity. Through experience, individuals learn to associate particular facets of their identity with specific roles, environments or contexts. As such, one learns to present one's work facet in environments associated with work. In order to properly present oneself, an individual must be constantly aware of the environmental feedback that they are receiving and adjust accordingly.

Although overly simplified, people engage in such negotiations in everyday life. Through experience, people learn to manage different aspects of their identity, present themselves and read the presentations of others. They develop mental models for understanding the cues that exist in a social setting and learn how to utilize them to assess a situation. Awareness and control are integral aspects of negotiating social interactions.

Online, the rules change. Although people have developed nuanced structures for negotiating social situations, many of this must be altered as people move into a digital environment. The underlying architectural features of the digital world present new challenges for social interactions. Users must learn to present themselves through an agent rather than through their natural body. Additionally, while people are accustomed

to reading contextual information in order to present themselves, digital contextual cues are not what is typically expected. As digital social interactions are comprised of archiveable bits, this information can be aggregated with ease. Searchable archives collapse situational context information, leaving the users vulnerable, as they cannot properly present themselves for a specific context without risking the information being collapsed with other presentations.

In order to negotiate the digital environment, people must adjust their behavior to accommodate for the architecture. To gain control over the possibility of collapsed contexts, users began creating multiple accounts and associating each with a role or facet of their identity. Using this mechanism, they localize context to the account and present it comfortably in any situation. This provides users with a temporary solution for acquiring control over their presentation. Yet, in order to maintain the separation between different accounts, the user cannot present any data that would allow these to collapse, such as an identifying name or email address. The architecture does not easily support such management; more recently, changes to the architecture of various websites and applications makes this separation even more challenging.

Although corporations such as Microsoft believe that people want aggregated data, they fail to recognize that people seek out separate accounts in order to properly adjust their social identity to the digital environment. As a designer, i believe that we should create systems that enable people to properly present themselves as they see fit. Rather than overriding the desire for control over context and faceted identities, i believe that we should design systems that offer better interfaces for managing this information. Thus, in this thesis, i have proposed a two-tiered approach to address this problem.

First, users must be given appropriate mechanisms for being aware of their presentation and that of those around them. Recognizing one's social presentation is far more challenging online, as people draw from a wider range of information. Yet, people want to have a sense of what information others have access to when they are constructing an image of the individual.

Secondly, users must have the ability to control the information that they are presenting. Social negotiation is about impression management, yet to manage impressions requires control over what data is presented and how. Identity management tools help users by providing them with a desired control over their presentation.

By empowering users through awareness and management, designers can enhance online social environments. Such tools enable users to adjust their presentations to be in line with the social environment. In turn, this develops a mechanism for social regulation; increased social regulation helps strengthen communities through shared responsibility. Many online social disasters result from communication failures, which are usually due to difficulty in perceiving and presenting oneself.

In this thesis, i have articulated how underlying architectural differences affect social interactions, first by discussing how people engage in the physical world and then addressing the confounds that emerge as this behavior moves online. I have offered a new approach to considering context in digital environments and addressed how people attempt to localize contexts by managing multiple accounts online. Stemming from my belief that designers should work to empower users, i have hypothesized that what users need includes tools for awareness and identity management. Addressing each type of tool specifically, i have analyzed current approaches and discussed what i believe should be developed. I have designed and/or constructed example prototypes and critiqued these systems and addressed the issues that they reveal.

I believe that empowering people to engage in meaningful and manageable social interactions is a worthy goal. To do so, i believe that systems must be built which address the social needs of all users. Thus, this thesis analyzes the issues that must be addressed by sociable designers. Additionally, in this thesis, i offer a set of potential approaches and critiques that help frame what must be done in future research. The hardest challenge that researchers in this area face is determining how to design interfaces that provide people with the necessary information. Although i have addressed some of the weaknesses in this area, i have only scratched the surface of what must be done. Thus, this thesis also serves as the initial steps and framework that i plan to use as i continue to develop and design systems intended to empower individuals.

BIBLIOGRAPHY

Anonymous. 2002, June 26. (Private email conversation; name withheld for privacy reasons.)

Aronson, Elliot. 1995. *The Social Animal.* New York: WH Freeman.

Behr, Mary. 2002, July 26. "All eyes are on you." *Popular Science.*

Bell, Genevieve. 2001. (Personal communications.)

Bos, Nathan D.; Judy S. Olson; Darren Gergle; Gary M. Olson; ZachWright. 2002. "Effects of four computer-mediated channels on trust development." *Proceedings of CHI 2002.* New York: ACM Press.

boyd, danah. 2001. "Sexing the Internet: Reflections on the role of identification in online communities." *Sexualities, Medias and Technologies Conference* (University of Surrey, June 21-22).

boyd, danah; Hyun-Yeul Lee; Daniel Ramage; Judith Donath. 2002. "Developing Legible Visualizations for Online Social Spaces" danah boyd, Hyun-Yeul Lee." *Proceedings of the Hawai'i International Conference on System Sciences-35* (Persistent Conversations Track; Big Island, HI; January 7-10). Los Alamitos, CA: IEEE Computer Society.

Brin, David. 1998. *The Transparent Society*. New York: Addison-Wesley.

Burt, Ronald. 1993. "The Social Structure of Competition." *Explorations in Economic Sociology* (Richard Swedberg, ed.), 65-103. New York: Sage.

Butler, Judith. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. London: Routledge.

Cha, Ariana Eunjung. 2001, November 19. "Online Companies Draw Fire For Removing 'Offensive' Postings." *Washington Post.*

Chiou, Ta-gang. 2000. "Virtual Fashion: Tracking and analyzing cultural dispersion on the World-Wide Web." Cambridge, MA: MIT Media Lab Master's Thesis.

Davis, Fred. 1992. *Fashion, Culture and Identity*. Chicago: University of Chicago Press.

de Moor, Des (Dirk). 1997, May 19. "Codes and Symbols." *the deviant's dictionary*. http://www.queernet.org/deviant/sscodes.htm

Descartes, René. 1641. "Meditations on the First Philosophy." *Which the Existence of God, and the Real Distinction of Mind and Body, are Demonstrated* (John Veitch, trans. 1901).

Dey, Anind K. 2000, December. "Providing Architectural Support for Building Context-Aware Applications." Atlanta, GA: Georgia Institute of Technology Ph.D. Thesis.

Dibbell, Julian. 1993, December 23. "A Rape in Cyberspace or How an Evil Clown, a Haitian Trickster Spirit, Two Wizards, and a Cast of Dozens Turned a Database into a Society." *The Village Voice.*

Donath, Judith. 1999. "Identity and deception in the virtual community." *Communities in Cyberspace* (Marc Smith & Peter Kollock, eds). London: Routledge.

Donath, Judith. 1995. "Visual Who." *Proceedings of ACM Multimedia '95* (San Francisco, CA, November 5 -9). New York: ACM Press.

EPIC: Electronic Privacy Information Center. 1994. http://www.epic.org

Erickson, Thomas; David N. Smith; Wendy A. Kellogg; Mark Laff; John T. Richards; Erin Bradner. 1999. "Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of 'Babble.'" *Human Factors in Computing Systems: The Proceedings of CHI '99.* New York: ACM Press.

Federal Trade Commission. 2000, June. "Online Profiling: A Report to Congress."

Foucault, Michel. 1995/1975. *Discipline and Punish: The Birth of the Prison* (Alan Sheridan, trans.). New York: Vintage.

Freud, Sigmund. 1974/1923. "The Ego and the Id." *The Standard Edition of the Complete Psychological Works of Sigmund Freud (SE)*, vol. XIX (James Strachey, et al., eds.) London: Hogart Press & The Institute of Psycho-Analysis.

Garfinkel, Simson. 2000. *Database Nation: The Death of Privacy in the 21st Century*. Cambridge, MA: O'Reilly & Associates.

Garton, Laura; Caroline Haythornthwait; Barry Wellman. 1999. "Studying Online Social Networks." *Doing Internet Research* (Steve Jones, ed.), 75-105. Thousand Oaks, CA: Sage.

Gelder, Ken and Sarah Thornton. 1997. *The Subcultures Reader.* London: Routledge.

Giddens, Anthony. 1991. *Modernity and Self-Identity: Self and Society in the Late Modern Age.* Cambridge: Polity.

Goffman, Erving. 1956. *The Presentation of Self in Everyday Life.* New York: Doubleday.

Google. 2001, December 15. "20 Year Usenet Archive Now Available." http://www.google.com/googlegroups/archive_announce_20.html

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78: 1360-1380.

Green, Nicola; Hazel Lachoee; Nina Wakeford. 2001. "Rethinking Queer Communications: Mobile Phones and beyond." *Sexualities, Medias and Technologies Conference* (University of Surrey, June 21-22).

Hall, Stuart. 1996. "The Meaning of New Times." *Stuart Hall: Critical Dialogues in Cultural Studies* (David Morley & Kuan-Hsing Chen, eds.). New York: Routledge

Harrison, Steve; Paul Dourish. 1996. "Re-Place-ing Space: The Roles of Place and Space in Collaborative Systems." *Proceedings of ACM Conference on Computer-Supported Cooperative Work CSCW'96* (Boston, Massachusetts), 67-76. New York: ACM.

Harvey, David. 1990. *The Condition of Postmodernity*. Cambridge, MA: Blackwell.

Hauben, Ronda. 2002, January. "Commodifying Usenet and the Usenet Archive or Continuing the Online Cooperative Usenet Culture." *Science Studies* 15:1, 61-68.

Hebdige, Dick. 1991/1979. *Subculture, the meaning of style.* London: Routledge.

Isbell, Charles; Michael Kearns; Dave Kormann; Satinder Singh; Peter Stone. 2000. "Cobot in LambdaMOO: A Social Statistics Agent." *Proceedings of the Seventeenth National Conference on Artificial Intelligence* (Austin Texas; July 30-August 2).

Joinson, Adam. 1998. "Causes and Implications of Disinhibited Behavior on the Internet." *Psychology and the Internet: Intrapersonal, Interpersonal, and Transpersonal Implications (*Jayne Gackenbach, ed.). San Diego, CA: Academic Press.

Kilduff, Martin. 1992. "The friendship network as a decision -making resource: Dispositional moderators of social influences on organizational choice." *Journal of Personality and Social Psychology* 62: 168-80.

Killworth, Peter; Eugene Johnsen; H Russell Bernard; Gene Ann Shelley; Christopher McCarthy. 1990. "Estimating the Size of Personal Networks." *Social Networks* 12:289-312.

Krackhardt, David. 1999. "The Ties That Torture: Simmelian Tie Analysis in Organizations." *Research in the Sociology of Organizations* 16:183-210.

Lacan, Jacques. 1980/1968. "The language of the self: The function of language in psychoanalysis." *Metaphors We Live By* (George Lakoff & Mark Johnson, eds.; A. Wilden, trans.). Chicago: University of Chicago Press.

Lacan, Jacquires. 1977/1966. "The mirror stage as formative of the function of the I." *Écrits: A selection* (Alan Sheridan, trans.). New York: W.W. Norton.

Large Scale Design GISP. 1998. "Memex and Beyond." http://www.lsd.brown.edu/design/ & http://www.buddygraph.com/projects/buddygraph-v2000/

Le Bon, Gustave. 1952/1895. *The Crowd: A Study of the Popular Mind.* London: Ernest Benn.

Lebo, Harlan. 2001, November. "UCLA Internet Report 2001: Surveying the Digital Future, Year Two." Los Angeles: UCLA Center for Communication Policy.

Lee, Jennifer 8. 2002, July 25. "Net Users Try to Elude the Google Grasp." *The New York Times.*

Lessig, Lawrence. 1999. *Code and Other Laws of Cyberspace.* New York: Basic Books.

Marlow, Cameron. 2001. "Blogdex." http://blogdex.media.mit.edu/

McPherson, Miller; Lynn Smith-Lovin; James Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415-444.

Mead, George Herbert. 1934. *Mind, Self, and Society* (C.W. Morris, ed.). Chicago: University of Chicago.

Milgram, Stanley. 1967. "The Small-World Problem." *Psychology Today* 1:62-67.

Milgram, Stanley. 1974. *Obedience to authority*. New York: Harper & Row.

Minsky, Marvin. 1985. *The Society of Mind.* New York: Simon & Schuster.

Mitchell, William. 1995. *City of Bits: Space, Place, and the Infobahn.* Cambridge, MA: MIT Press.

Moustakas, Clarke. 1967. *Creativity and conformity*. Princeton, NJ: D. Van Nostrand.

Negroponte, Nicholas. 1995. *Being Digital.* New York: Vintage Books.

Nguyen, David H.; Elizabeth D. Mynatt. 2002, June. "Privacy Mirrors: Understanding and Shaping Socio-technical Ubiquitous Computing Systems." Atlanta, GA: Georgia Institute of Technology Technical Report (GIT-GVU-02-16).

O'Brien, Jodi. 1999. "Writing in the body: Gender (re)production in online interaction." *Communities in Cyberspace* (Marc Smith & Peter Kollock, eds.). London: Routledge.

PEW Foundation. 2001. "Daily Internet Activities." *The Pew Internet and American Life Project.* http://www.pewinternet.org/

Pink Zone. 2002. "Pride Symbols Guide." *sexuality.org: society for human sexuality* http://www.sexuality.org/l/lesbigay/pridesym.html

Rank, Otto. 1932. *Will therapy and truth and reality*. New York: A. A. Knopf.

Reid, Elizabeth M. 1998. "The Self and the Internet: Variations on the 'Illusion' of One Self." *Psychology and the Internet: Intrapersonal, Interpersonal, and Transpersonal Implications* (Jayne Gackenbach, ed.). San Diego, CA: Academic Press.

Rice, Ronald. 1994. "Network analysis and computer-mediated communication systems." *Advances in social network analysis: Research in the social and behavioral sciences* (Stanley Wasserman & Joseph Galaskiewicz, eds.), 167-203. Newbury Park, CA: Sage.

Rocco, Elena; Tom A. Finholt; Erik C. Hofer; Jim D. Herbsleb. 2000. "Designing as if Trust Mattered." Ann Arbor: University of Michigan CREW Technical Report.

Rosen, Jeffrey. 2000. *The Unwanted Gaze: The Destruction of Privacy in America*. New York: Vintage.

Sack, Warren. 2000. "Discourse Diagrams: Interface Design for Very Large Scale Conversations." *Proceedings of the Hawai'i International Conference on System Sciences-33* (Persistent Conversations Track; Maui, HI; January). Los Alamitos, CA: IEEE Computer Society.

Saville-Troike, Muriel. 1982. *The ethnograhy of communication. An introduction.* Baltimore, Maryland: University Park Press.

Scheeres, Julia. 2002, July 11. "Dear Member: You've Been Deleted." *Wired News.*

Schoeneman, Deborah. 2001, January 15. "Don't Be Shy, Ladies– Google Him! Check Out His Search Engine First." *New York Observer.*

Shneiderman, Benjamin. 1987. *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Reading, MA: Addison-Wesley.

Simpson, Rosemary. 1995. "Memex and Beyond." http://www.cs.brown.edu/memex/

Simmel, Georg. 1971. *On Individuality and Social Forms*. Chicago: University of Chicago Press.

Smith, Adam. 1976/1790. *The Theory of Moral Sentiments* (D.D. Raphael & A.L. Macfie, eds.). Oxford: Claredon Press.

Smith, Marc. 2001. "Netscan: A tool for measuring and mapping social cyberspaces." http://netscan.research.microsoft.com.

Smith, Marc; Andrew Fiore. 2001. "Visualization Components for Persistent Conversations." *ACM Proceedings of SIG CHI 2001* (Seattle, WA; March 31-April 5), 136-143. New York: ACM Press.

Snyder, Marc. 1974. "Self-Monitoring of Expressive Behavior." *Journal of Personality and Social Psychology* 30:526-537.

Sproull, Lee; Sara Kiesler. 1991. *Connections: New ways of working in the networked organization.* Cambridge, MA: MIT Press.

Stone, Allucquere Rosanne. 1998. *The War of Desire and Technology at the Close of the Mechanical Age.* Cambridge, MA: MIT Press.

Tufte, E. 1990. *Envisioning Information.* Chesire, CT: Graphics Press.

Turkle, Sherry. 1995. *Life on the Screen: Identity in the Age of the Internet.* New York: Simon & Schuster.

Viégas, Fernanda. 1997. "Collections: Adapting the Display of Personal Objects for Different Audiences." Cambridge, MA: MIT Media Lab Master's Thesis.

Viégas, Fernanda. 1999. "ChatCircles." http://chatcircles.media.mit.edu/

Viégas, Fernanda. 2001. "Post History." http://smg.media.mit.edu/projects/PostHistory/

Wellman, Barry; Janet Salaff; Dimintrina Dimitrova; Laura Garton; Milena Gulia; Caroline Haythornthwaite. 1996, February. "Computer Networks as Social Networks: Virtual Community, Computer Supported Cooperative Work and Telework." *Annual Review of Sociology* 22: 213-38.

Wellman, Barry; Keith Hampton. 1999. "Living Networked in a Wired World." *Contemporary Sociology* 28(6), 648-654.

Whittaker, Steven; Loren Terveen; Will Hill; Lynn Cherny. 1998. "The dynamics of mass communication," *The Proceedings of CSCW '98*.

Xiong, Rebecca; Eric Brittain. 1999. "LiveWeb: Visualizing Live User Activities on the Web." *Proceedings of SIGGRAPH 1999* (Sketch; Los Angeles; August 8-13). New York: ACM Press.

Zahavi, Amotz. 1997. *The handicap principle: A missing piece of Darwin's puzzle.* New York: Oxford University Press.