

# Imprint, a Community Visualization of Printer Data: Designing for Open-ended Engagement on Sustainability

Zachary Pousman, Hafez Rouzati, John Stasko

School of Interactive Computing / GVU Center

Georgia Institute of Technology

Atlanta, GA 30332

{zach, rhythm, stasko}@cc.gatech.edu

## ABSTRACT

We introduce Imprint, a casual information visualization kiosk that displays data extracted from a printer queue. We designed the system to be open-ended, and to support a workgroup in reflection and conversation about the data. Imprint's visualizations depict environmental issues, such as energy consumption and paper consumption of the printers, as well as social information, such as popular concepts from the printed matter. Imprint is intended to spark reflection and conversation, and to bring data into discussions about paper usage and "waste." Our goal is not to explicitly reduce paper, energy, or toner consumption, but instead to open conversations by community members. Our work highlights a design approach for semi-public displays of personal data.

## Author Keywords

Information radiator, sustainability, casual information visualization, privacy, identity, persuasion

## ACM Classification Keywords

H.5.2 Graphical User Interfaces, I.7.m document and text processing.

## INTRODUCTION

The modern world is full of traces, ways in which human activity, either individual or aggregate, leaves behind some residue or trace. These traces are often invisible, fleeting and unexamined. As communication, commerce, work, and play become more and more technologically mediated, many areas of human life are subject to storage, tracking, and monitoring.

However, in many contexts, authority figures such as system administrators, accountants, and government personnel monitor people's activities as opposed to individuals being able to access their own data. Mann introduces a counter to *surveillance* (literally translated as "watching from above") that he calls *sousveillance*, a "watching from below", where a community watches itself

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW'08, November 8–12, 2008, San Diego, California, USA.

Copyright 2008 ACM 978-1-60558-007-4/08/11...\$5.00.

for its own ends and its own purposes [10]. Bowker [2], Star [13], and others have argued that studying the invisible aspects of infrastructure and making infrastructural data and processes visible can result in reflective thinking and new understandings. Following them, we hope to draw a community's attention to the ways that infrastructure shapes its understanding of the world and of itself, and to draw the community into reflection, contemplation, and conversation.

Our current focus is the trace data left by printer traffic. Workers at companies in the U.S. used 3.2 billion reams of paper in 2004 [11], much of it as they print and copy documents. Companies often set up group print facilities, such as a bank of laser printers shared by the workers on a given floor. Here, workgroup members print out emails, manuals, and personal documents. We argue that printed matter, when compared to other work-related trace data, such as web traffic or email, captures a lot of the important tasks and topics for many information workers. A worker might glance at many hundreds of emails, but she may choose to print only an important few. Sampling what workers print is one way to cull important topics, issues, and problems that a community is dealing with.

Imprint monitors a printer (or group of printers) and displays interactive visualizations of aggregated printed data on a large touch-screen kiosk in the printer area (see Figure 1). Imprint visualizations depict two broad classes of information: social information, information about which members of the community are printing, what concepts are popular, and what sub-groups and tensions are present in the community and second, environmental information, information about how much paper, toner, and energy is being used as individuals print their documents.

Our hypothesis, that we are currently testing in long term deployments, is that community members will use Imprint's visualizations to reflect on their habits and activities and on the community itself. In doing so, we expect that they might adopt printing practices to manage their social identities and perhaps to reduce their use of energy, paper, and toner. More important than member's actions are their motivations and thinking; we want to determine if Imprint can serve as the beginning of discussions, conversations, and explorations instead of a way to use the data to close or complete the discussion in a final way.

Our work on Imprint contributes the following design guidance for CSCW researchers. First, we show how casual information visualizations (see Pousman *et al.* [12]) can serve as a mirror for community members, displaying the often invisible parts of work life. Second, we describe how to display social data on a semi-public display while safeguarding individual workers' privacy. Lastly, we illustrate how an open-ended design, one that does not help workers complete proscribed task-based work, can still be valuable as a community resource for consensus building.

## **DESIGN GOALS**

### **Defamiliarizing printer data**

Imprint creates visualizations that remove print data from the everyday, following notions of defamiliarization [1]. Defamiliarization is a design strategy that appears in much of art, where the normal way of seeing or understanding some topic can be challenged by showing the object or topic from a novel perspective. Removing print data from its normal everyday "document" status helps to open new perspectives and to show this data in new ways.

We note prior examples of systems that display environmental impacts and traces to occupants of work spaces. Jeremijenko's *Stump* creates visual tree rings as coversheets instead of the usual coversheet metadata. The coversheets could be stacked to form a stump made from wasted paper [9]. Grasso and Meunier's *SmartCover* system also provides historical and social perspective on print activity [6]. However neither of these systems support interactive data exploration or hypothesis testing. Holmes' *7000 Oaks and Counting* is an ambient display and casual infovis system that reads building automation data to determine a real time carbon footprint of a building [7].

### **Opening sustainability not closing it**

Our goal is not to modify the behavior of members of the community in a simplistic manner; we do not want to build a system that bullies, badgers, cajoles or shames community members into changing their behaviors (such as printing fewer pages). This kind of intervention, even if it were achievable, might not cause long-lasting behavior change, but instead might cause workers to print more at home, routing around a technology they see as onerous or judgmental. Instead, we seek to engage community members in a data-driven discussion and contemplation, so that they can come to their own conclusions about printing practices and changes they may want to individually or collectively implement.

### **Building a community artifact**

Imprint is a technology probe [8]. Technology probes are deployments of an early-stage technology that designers use to explore the design space, not just to refine the design and evaluate usability or usage. Probes are deployed to provide rich feedback about the social particulars of the community, as well as to inspire both designers and users to "think of new kinds of technology to support their needs and desires." [ibid.]

With Imprint, we are designing a community artifact, a place for community members to view and reflect on their activity. The system is a semi-public display, and therefore we need to carefully balance information with worker's privacy concerns. We do this by displaying visualizations

that are either about *people*, and only implicitly include document information, or visualizations that are about *documents*, and only implicitly refer to people, so Imprint does not connect a worker to his or her prints too closely.

### **Designing for open-ended Interaction**

We take inspiration from Gaver *et al.*'s *Home Health Horoscope* [5] and the general approach to designing for ludic engagement. Ludic engagement is a design strategy for engaging the playful parts of human life, as opposed to targeting the task-centric and analytically solvable parts of human life (which is the implicit focus of much technological intervention) [4]. From here we take inspiration not to "solve" wasted printer pages, and instead to engage a workgroup community in playful reflection.

We think of Imprint as an intervention that has some of the properties of a meta-design project. Meta-design, a design theory and process proposed by Fischer (and his colleagues) [3] goes beyond User-Centered Design and Participatory Design to include users even more deeply in the design process, by making them active participants in the evolution of the system. By intentionally and carefully "under-designing" systems, leaving them open to interpretation, exploration, and different kinds of appropriation, users become more than passive recipients. End users become creative designers in their own right during the deployment. The systems serve as the "seed" for the community to begin to explore the issues at hand. Imprint launches with a set of visualizations but community members are encouraged to suggest modifications and even new visualizations of the data.

### **Simple metrics for 'game-ability'**

Imprint uses print data to create visualizations. But the metrics we chose to use in the seed visualizations are intentionally simple. We want workers to be able to connect their printing practices to the visualizations, and for the metrics that the system uses to categorize and organize documents to be understandable. In this way, community members can experiment with the system and are presented with ways to publicly express aspects of their identity, if they so choose.

## **SYSTEM DESCRIPTION**

In this section, we detail how Imprint monitors printing traffic on the network, how we parse and describe print outs in a modular way, and how the visualizations are created. We describe the first set of "seeding" visualizations, and describe the ways that user can interact with the system.

### **Monitoring and storing print data**

Imprint consists of three main components: a modified CUPS print server, a custom document parser, and an end user application all of which are backed by a persistent data store. The print server stores job files and initiates parsing. Imprint then converts raw job file data (Postscript) to PDF and then parses the PDF files, inserting all of the text, images, and layout information into the persistent store. We supplement the standard printing infrastructure with plug-in components (PyKota and Tea4Cups) that are designed for traditional print-accounting and quota-related monitoring. We are not using this data for their intended purpose.

Instead, we use them to determine print job ownership and to calculate paper and toner usage so that it can be fodder for visualizations for the community's benefit.

Imprint's Cocoa front-end implements a modified Model-View-Controller (MVC) design pattern that has been specifically adapted to maintain and efficiently represent the temporal and contextual relationships that exist between the users and documents. We model the data abstractly so as to provide a decouple the data from each independent visualization. This "proxy" design pattern efficiently represents our entity relationships compactly while minimizing data overhead.

### Imprint Visualizations

Imprint's touch-sensitive LCD display consists of a looping slideshow of multiple visualizations. Each visualization focuses on a particular aspect of data surrounding printing activity. We describe five of the eight "seeding" visualizations here (see Figure 1).

#### Tagcloud

The tagcloud (see Figure 1A) collects commonly printed words and can serve as a gestalt summary of the concepts that are popular. More popular words are larger and less popular words are displayed in smaller fonts. Recently printed words are highlighted, so that the last few documents can be compared to the historical data.

#### Popularity ('Hot or not')

As we sought to reflect community relationships and mores, we thought about the kinds of documents that are frequently printed in our community. In our home department at Georgia Tech, academic researchers print out many conference and journal articles, both for research endeavors and for coursework. These documents often contain extensive references sections, where other community members are mentioned. We designed simple a visualization to show how popular a given community member is (see Figure 1B). Imprint counts how often a worker's name is printed (we accept instances of "first name surname" and "first initial and surname"), and displays a bar graph for each user. We expect workers either to help us design more precise metrics for calculating popularity or perhaps they may enjoy the game-ability of this simple metric.

#### What's your neighborhood?

Imprint contains a visualization that clusters workers based on the documents they printed. We use Principal Components Analysis (PCA) to determine the distance, in document space, between each pair of workers. PCA reduces highly multi-dimensional data to a lower dimensional space. In our case, we project the dimensions of words and images into two dimensions, so that the relationships can be drawn on a 2D plane. We superimpose the 2D graph onto a world map (see Figure 1C). The map gives workers a well known context for discussing the groups and cliques amongst workers. They can, over time, begin to tease out what kinds of documents position them over which continents.

#### What are the printers doing?

Workgroup printers are often used frequently. However, the printers sit idle a good proportion of the time, going into a

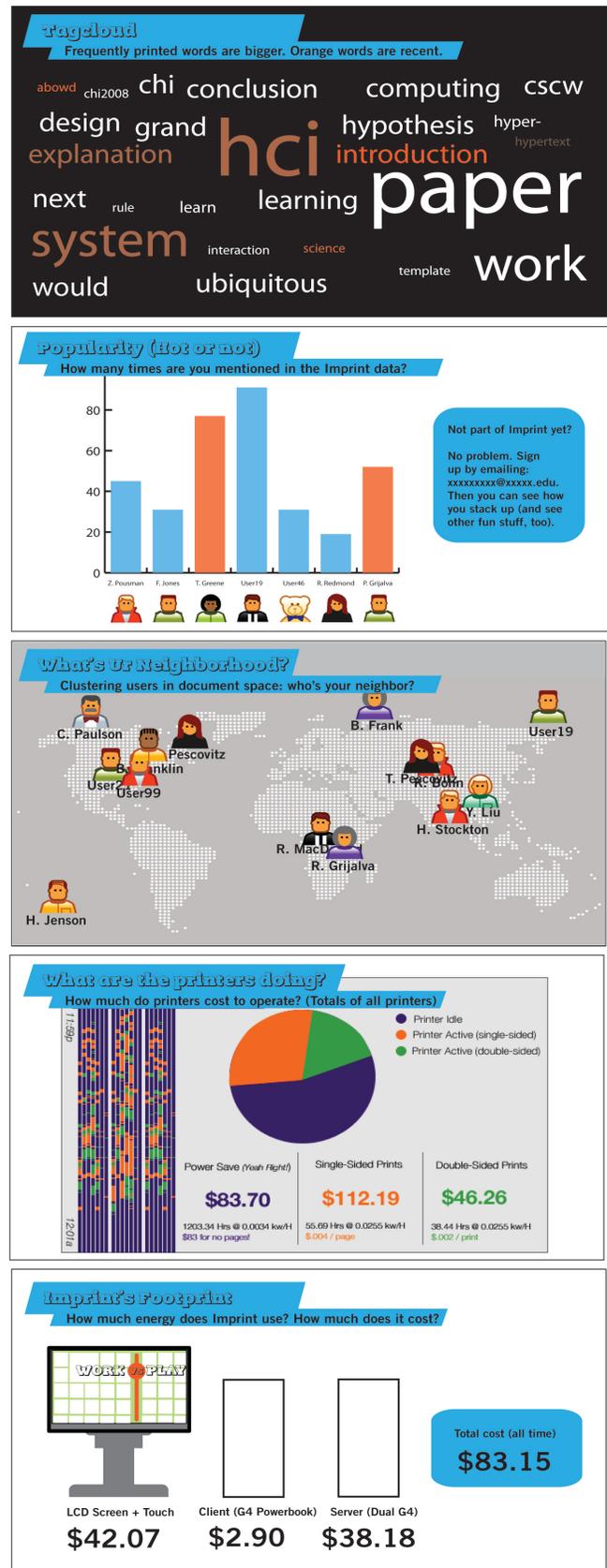


Figure 1: Imprint visualizations from top to bottom: 1A) Tagcloud, 1B) Popularity ('Hot or not'), 1C) What's ur neighborhood? 1D) What are the printers doing? 1E) Imprint's footprint.

“Powersave” mode. However, in our measurements with a Kill-a-Watt unit<sup>1</sup>, printers still use a good bit of electricity (in the hundreds of watts). Imprint displays a pie chart which aggregates the time that printers are working versus idle their times. We also display a strip chart for each day, where the colored bars indicate when the printer was busy or idle. We group the bars by week to show trends, such as busy weekdays with more idle weekends (see Figure 1D).

#### *Imprint's Footprint*

One of Imprint's visualization depicts the running total amount of energy used to power Imprint itself (see Figure 1E). Displaying this data lends credibility to the system and to the designers, since it is a disclosure of the environmental impact of installing the system. Disclosing this data can serve as another spark to conversation and reflection on the system and its usefulness.

#### **Interactivity: Supporting Exploration**

Imprint displays each visualization in a rotating slideshow, so that each time a worker passes by the system, he or she might see a different visualization. A worker can press large next and back buttons to move to different slides. Users may also explore the data of any visualization in time to see time-based patterns and historical views. Each visualization has a time slider on the bottom, which defaults to the last three days of printer data. The slider can be used to move this three day window back in time, to show the state of the data at that point. Beyond the default three days, users can select a 10-day window or view the visualization over “all time” (since install).

Community members can also print out visualizations that they find compelling or noteworthy. We expect that users may want to keep certain images for their own ends, to share a particular (and transient) view with coworkers.

#### **Displaying comments and eliciting suggestions**

Imprint also contains visualization slides that do not display data, but instead encourage users to give feedback, or to suggest new visualizations. We also plan to take over some of the wall space near the printers for a gallery for users to post (with any annotations, insights, or comments) visualizations that they find telling or insightful or funny. We also place a physical suggestion box in front of imprint to gather feedback and suggestions.

#### **CONCLUSION AND FUTURE WORK**

We have introduced Imprint, a resource for workgroup conversation and reflection on printed data. The system is open-ended; though we present visualizations of the printer trace data, we expect a community of coworkers will suggest new metrics, new visualizations, and new issues which can all be inserted into the system. Each deployment of Imprint will therefore diverge and evolve, including workers as more than just users, but as creative designers.

#### **Longitudinal Evaluation**

Our current work is to begin to test the system. We are beginning with a longitudinal deployment on our floor of an academic research building at Georgia Tech. Our plan is to measure the effects of the intervention, both in

qualitative and quantitative ways, using a mixed methods approach. Our aim is to determine the degree to which members of our workgroup interact with the system, use it as fodder for conversation, and, perhaps, modify their behaviors.

#### **ACKNOWLEDGEMENTS**

The authors wish to thank the College of Computing Academic Computing director, Randy Carpenter for assisting us in accessing printers and print stream data. We also thank A. Grimes, M. Chetty, B. Landry and E. Poole for their comments on early drafts of this paper.

#### **REFERENCES**

1. Bell, G., Blythe, M., and Sengers, P. Making by making strange: Defamiliarization and the design of domestic technologies. *ACM ToCHI*. 12:2(2005) pp. 149-173.
2. Bowker, G. Information Mythology: The World As/Of Information. In Bud-Frierman (ed), *Information Acumen: The Understanding and Use of Knowledge in Modern Business*. 1994. pp 231-247.
3. Fischer, G. and E. Giaccardi. Sustaining Social Creativity. *CACM*, 50:2(2007) pp 28-29.
4. Gaver, W., Bowers, J., Boucher, A., Gellerson, H., Pennington, S., Schmidt, A., Steed, A., Villars, N., and Walker, B. The Drift Table: Designing for Ludic Engagement. *Extended Abstracts of CHI 2004*.
5. Gaver, W. and Sengers, P. and Kerridge T., and Kaye, J. and Bowers, J. Enhancing Ubiquitous Computing with User Interpretation: Field Testing the Home Health Horoscope. *Proceedings of CHI 2007*, pp 537-546.
6. Grasso, A. and Meunier, J. 2002. Who can claim complete abstinence from peeking at print jobs? *Proceedings of CSCW 2002*, pp. 296-305.
7. Holmes, T. G. Eco-visualization: Combining Art and Technology to Reduce Energy Consumption. In *Proceedings of Creativity&Cognition 2007*, pp.153-162.
8. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., and Eiderback, B. Technology Probes: Inspiring Design for and with Families. 2003. *Proceedings of CHI 2003*, pp. 17-24.
9. Jeremijenko, N. Stump. 1999. Retrieved on March 29, 2008 from <http://onetrees.org/stump/index.html>
10. Mann, S., Nolan, J., and Wellman, B. Sousveillance: Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments. *Surveillance & Society*. 2003. 331-355.
11. Paulson, R. Green Procurement Requirements and the Use of 100% Post Consumer Fiber Paper. Organization: NADEP; Environmental Program Office, 2005.
12. Pousman, Z., Stasko, J., and Mateas, M. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics* 13, 6, 1145-1152.
13. Star, S. The Ethnography of Infrastructure. *American Behavioral Scientist*. 1999. pp 377-391.

<sup>1</sup> <http://www.p3international.com/products/p4460.html>