A Stage-Based Model of Personal Informatics Systems

Ian Li[1], Anind Dey[1], and Jodi Forlizzi[1,2]
[1]Human Computer Interaction Institute, [2]School of Design
Carnegie Mellon University, Pittsburgh, PA 15213
ianli@cmu.edu, {anind, forlizzi}@cs.cmu.edu

ABSTRACT
People strive to obtain self-knowledge. A class of systems called *personal informatics* is appearing that help people collect and reflect on personal information. However, there is no comprehensive list of problems that users experience using these systems, and no guidance for making these systems more effective. To address this, we conducted surveys and interviews with people who collect and reflect on personal information. We derived a stage-based model of personal informatics systems composed of five stages (*preparation*, *collection*, *integration*, *reflection*, and *action*) and identified barriers in each of the stages. These stages have four essential properties: barriers cascade to later stages; they are iterative; they are user-driven and/or system-driven; and they are uni-faceted or multi-faceted. From these properties, we recommend that personal informatics systems should 1) be designed in a holistic manner across the stages; 2) allow iteration between stages; 3) apply an appropriate balance of automated technology and user control within each stage to facilitate the user experience; and 4) explore support for associating multiple facets of people's lives to enrich the value of systems.

Author Keywords
Personal informatics, collection, reflection, model, barriers

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design, Human Factors

INTRODUCTION AND MOTIVATION
The importance of knowing oneself has been known since ancient times. Ancient Greeks who pilgrimaged to the Temple of Apollo at Delphi to find answers were greeted with the inscription "*Gnothi seauton*" or "Know thyself". To this day, people still strive to obtain self-knowledge. One way to obtain self-knowledge is to collect information about oneself—one's behaviors, habits, and thoughts—and reflect on them. Computers can facilitate this activity because of advances in sensor technologies, ubiquity of access to information brought by the Internet, and improvements in visualizations. A class of systems called *personal informatics* is appearing that help people collect and reflect on personal information (*e.g.*, Mint, http://mint.com, for finance and Nike+, http://nikeplus.com, for physical activity).

Personal informatics represents an interesting area of study in human-computer interaction. First, these systems help people better understand their behavior. While many technologies inform people about the world, personal informatics systems inform people about themselves. Second, people participate in both the collection of behavioral information as well as the exploration and understanding of the information. This poses demands on users that need to be explored. Finally, we do not know all the problems that people may experience with personal informatics systems. We know that people want to get information about themselves to reflect on, and that systems that support this activity need to be effective and simple to use. Identifying problems that people experience in collecting and making sense of personal information while using such systems is critical for designing and developing effective personal informatics.

To date, there is no comprehensive list of problems that users experience using these systems. Toward this end, we conducted surveys and interviews with people who collect and reflect on personal information. From this, we derived a model of personal informatics systems organized by *stages*, which emphasizes the interdependence of the different parts of personal informatics systems.

We provide three main contributions in this paper: 1) we identify problems across personal informatics tools, 2) we introduce and discuss a model that improves the diagnosis, assessment, and prediction of problems in personal informatics systems, and 3) we make recommendations about how to improve existing systems and build new and effective personal informatics systems.

In the next section, we provide a working definition of personal informatics and review related literature. We present the method and findings from our survey, and use them to introduce a stage-based model of personal informatics systems. We describe the barriers encountered in each stage and highlight opportunities for intervention within each stage. We also compare and analyze existing systems to demonstrate the use of the model for diagnosing

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.
CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
Copyright 2010 ACM 978-1-60558-929-9/10/04....$10.00.

and assessing problems. We conclude with a discussion of design guidelines for personal informatics systems and directions for future research.

A Definition of Personal Informatics

Collecting information about oneself and reflecting on the information has a long history. Benjamin Franklin tracked the days in which he accomplished one of his 13 virtues for 60 years [5]. Buckminster Fuller documented every 15 minutes of his life in a scrapbook he called the "Dymaxion Chronofile". These days, there are people like Nicholas Felton, who has been publishing annual reports about himself since 2005 (http://feltron.com), and Ellie Harrison, who created art projects on her personal behavior, such as food consumption and sneezes (http://ellieharrison.com). These are extreme examples, but regular people are tracking one or two types of personal information and using computing technology in the form of personal informatics systems to do so.

We define personal informatics systems as those that *help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge*. There are two core aspects to every personal informatics system: collection and reflection. Effective personal informatics systems help users collect the *necessary* personal information for *insightful* reflection. Personal informatics goes by other names, such as "living by numbers", "quantified self", "self-surveillance", "self-tracking", and "personal analytics" [21,22].

Personal informatics systems provide an advantage over simply trying to remember information about the self, because pure self-reflection is often flawed. This is because people have limited memory, cannot directly observe some behaviors (*e.g.*, sleep apnea), and may not have the time to constantly and consistently observe some behaviors (*e.g.*, manually counting steps throughout the day). Reflecting by using memory alone makes it difficult to see patterns and trends. People may also not have the expertise or knowledge to make the correct conclusions about their observations. Personal informatics systems help people by facilitating collection and storage of personal information, and by providing a means of exploring and reflecting on the information.

RELATED WORK

We discuss prior work related to the collection and reflection aspects of personal informatics, focusing on the collection of personal information, and the use of visualizations in non-work settings. We also discuss research projects and commercial products that combine collection and reflection.

Many research areas focus on collecting personal information. Lifelogging research explores the use of sensors to collect various types of information about people's daily lives. MyLifeBits [7] envisions a future when daily activities of people, such as computing, web-browsing activity, electronic communication, and media usage, are recorded and archived. SenseCam, a wearable digital camera, takes photographs throughout the day while worn by the user [10]. The device also contains other sensors, such as light sensors, an infrared detector, and an accelerometer. GPS and microphones can be added to collect additional information. Many issues of collecting personal information throughout a lifetime have been discussed in several CARPE (Continuous Archival and Retrieval of Personal Experiences) workshops [8].

The experience sampling method or ESM is used in studies to collect personal information *in situ* (in the actual situation) over a long period of time. Researchers have developed many techniques to motivate people to participate in these time-intensive studies, such as improved questions and mobile devices that facilitate data input [20]. Context-aware devices alleviate interruptions by alerting the participant at more opportune times [12]. However, reflection on the data collected by ESM is for the researchers conducting a study and not the study participants. Some experience sampling projects have been developed that allow reflection on collected information: Track Your Happiness (http://trackyourhappiness.org) and Hsieh and colleagues' system [11].

Personal Information Management (PIM) focuses on how people manage their information so they can perform their tasks more efficiently [13]. PIM also explores how people can retrieve their information, but the focus is less on self-reflection and more on staying organized.

On the reflection side, Casual Information Visualization [19] and Slow Technology [9] help people reflect on everyday patterns. Casual Information Visualization aims to expand the definition of information visualization beyond work-related and analytical tasks to include non-experts. Slow Technology is a design agenda aimed at encouraging the development of systems that foster users to slow down to reflect, rather than speeding up performance. These areas of research focus primarily on reflection and less on collection of personal information. While these systems discuss personal reflection, it is not their primary focus.

The areas mentioned above have examined collection and reflection separately. Personal informatics distinguishes itself by considering the parts of collection and reflection as a whole process. Since the data must be about the person and the person must reflect on that data, the user is involved in both collection of and reflection on the data.

There have been a number of research projects that have combined collection and reflection on personal information. There are research physical activity systems, such as Fish'n'Steps [15], Shakra [16], and UbiFit [3]. Fish'n'Steps used a pedometer to count steps and a public display to visualize step counts between people. Shakra and UbiFit used mobile phones to collect and visualize physical activity information. Research systems in sustainability have also used technology for collection and reflection. StepGreen is a web site where people can report their

sustainable actions and see visualizations of their progress [18]. UbiGreen is a mobile phone system that tracks and visualizes green transportation habits [4]. Mycrocosm is a visual micro-blogging site that allows users to collect and reflect on various types of personal information [2]. Many commercial personal informatics systems have leveraged the ubiquity of access to information afforded by the Internet and mobile devices to help people in various domains such as finance, health, physical activity, and productivity (*e.g.*, Mint: http://mint.com, CureTogether: http://curetogether.com, DailyBurn: http://dailyburn.com, and Slife: http://slifelabs.com, respectively). There are also systems that allow collection of various types of personal information (*e.g.*, Daytum: http://daytum.com and your.flowingdata: http://your.flowingdata.com).

More systems are being created today, but there is no comprehensive list of problems that users encounter when they collect and reflect on personal information. There is also no common vocabulary to compare and contrast these systems. This paper identifies problems in existing systems and defines a model of personal informatics to help designers and developers create more effective systems.

METHOD

To better understand personal informatics systems and their users, we conducted a survey of people who collect and reflect on their personal information.

Survey

The survey asked participants to list the types of personal information they collect and reflect on. From their list, participants selected one that was the most interesting and relevant to them. The rest of the survey focused on the participant's selection and had three sections. In the first two sections, participants answered questions about collection and reflection: what tools they used, when and how often, their motivation for use, problems they encountered, and suggestions for improvement. In the last section, the survey asked what patterns, trends, and surprises participants found from reflecting on their information. The survey ended with demographics questions (*e.g.*, gender, age range, marital status, employment, education, and technology use). The following are example questions from the survey[1]:

- *How difficult is it to collect this personal information?*
- *What was your initial motivation to reflect on this collected personal information?*
- *What patterns (repeating events) have you found when exploring this collected personal information?*

Participants

We recruited participants from a blog dedicated to personal informatics (http://quantifiedself.com), a blog about general information visualization (http://flowingdata.com), and forums at two personal informatics web sites (http://slifelabs.com and http://moodjam.org). We chose these web sites because their readers and users were more likely to have used one or more personal informatics systems. Survey participants were entered into a raffle for a $25 Amazon gift certificate. We interviewed a subset of these participants using instant messenger to collect additional details about their responses. Interviewees received an additional $10 Amazon gift certificate.

We had 68 people complete the surveys, and 11 agreed to participate in the follow-up interviews. 37 participants were male. Ages ranged from 18 to 64 with a median age range of 26 to 30. About half had graduate degrees and another half were in college. More than half worked full-time. Participants were technologically savvy. 90% of participants used email or instant messenger daily to communicate. 60% of them used social networking sites daily. Most read news websites and blogs, ordered products online, and managed bills and bank accounts online.

RESULTS

We created affinity diagrams to analyze the survey and interviews. We did not determine a coding scheme beforehand; instead we identified themes from the data as we processed the responses. During the analysis, we identified the types of personal information collected and reflected on by participants, motivations for collection and reflection, and problems experienced. Next, we analyzed the survey and interviews for experiences with tools and barriers that users encountered, identifying a model of personal informatics.

Collected Personal Information

Participants reported a wide variety of information that they collected and reflected on (see Table 1). Many participants reflected on automatically collected information such as bank statements and email history, since they are readily available. Automatically collected information was split

automatic	#	manual	#
bank statements	54	calendar events	27
email history	52	status updates	22
credit card bills	38	work activities	22
phone call history	26	blog posts	21
SMS history	25	weight	21
IM history	25	exercise	20
financial software	23	browser bookmarks	20
electricity bill	23	time at work	18
browsing history	23	social bookmarks	18
search history	20	mood	17

Other automatically collected: heating bill (12), travel (2)
Other manually collected: journal/diary (16), pictures taken (14), sleeping habits (12), food consumption (12), productivity (10), health (9), medication intake (7), caloric intake (5), symptoms (5), miles ran (4), sports activities (4), blood pressure (4), blood sugar level (2), dream journal (2), step counts (2), relationship status (2), books read (1), habits of newborn baby (1), transportation (1)

Table 1. Top 10 types of personal information, automatically and manually collected by the participants.

[1] Visit http://personalinformatics.org/lab/survey/ for a complete transcript of the survey.

between those recorded by industry infrastructures, such as financial transactions from banks and energy consumption from utility companies, and those recorded by computing applications and services, such as search history and email history. Since manually collecting information requires more time, fewer participants reflected on them. However, there is a greater variety of manually collected information because many types of information cannot be recorded automatically. The most popular was calendar events, since it is critical for time management. Status updates were also popular because of the rise of micro-blogging services (*e.g.*, Twitter) that facilitate input using various media.

The remainder of the discussion will focus on the personal information that participants selected as *the most relevant and interesting to them*. The four categories of information most relevant and interesting to participants were finance, journaling, exercise, and general health. Finance is the prevalent information type since there is a strong incentive to keep track of where one's money goes, and there is a reliable infrastructure for tracking the information. Journaling is a common activity for recording one's thoughts and experiences, and people are using new tools such as blogs and microblogs (*e.g.*, Twitter, Facebook status updates). Exercise is also popular because it is an activity for which people want to track their progress, and many devices and web sites exist for collecting such information. General health information, such as food consumption, weight, symptoms, medication, amount of sleep, and alcohol/caffeine intake is also of strong interest. As for the outliers, there are people who collected information on productivity, status of relationships, computer usage, transportation, the habits of a newborn baby, and books read.

Tools Used

People used a variety of tools to collect information. Some used pen and paper because of their flexibility and ease of use; people can take them anywhere and they are easily accessible when a note needs to be written. This group faced a problem in having to transcribe their data to an electronic format in order to visualize their data. Some used Excel spreadsheets for graphing. Many used existing personal informatics web sites for collecting and reflecting on various information, such as finance, food consumption, mood, and physical activity. Some used physical devices such as pedometers, the WiiFit, and a continuous positive airway pressure (CPAP) machine for tracking sleep apnea. Some activities had an infrastructure that automatically records information such as financial transactions, search history, and communication tools (*e.g.*, email, instant messenger, and IRC), which people later explored for personal information. Some with programming backgrounds devised their own way of exploring data (*e.g.*, statistical packages for analysis, and programming languages to organize and cull information).

Reasons

Participants gave several reasons for collecting and reflecting on personal information: natural curiosity, interest in data, discovery of new tools, suggestion from another person, and trigger events.

Some people cited curiosity about themselves prompted them to collect personal information. P22 said, "Curiosity: Q: how much would I walk if I didn't ride my bike? A: kind of a lot, but not as much as you'd think." Some people identified themselves by their interest in quantitative data. Some participants used the terms "data nerd", "a student of information visualization", and "geekiness" to describe themselves. There is also the added value that the data is about them. P40 said, "I'm an engineer, so numbers and trends and stuff just interest me in general. Plus this data is about ME (her emphasis)."

Finding personal informatics tools also encouraged people to start collecting information. P48 said, "I've been following Nick Felton's annual reports so when he started Daytum, I joined to start tracking which restaurants I ate at." Others cited trigger events, such as problems in relationships (P60, P1), sleep patterns (P44), and weight (P21, P23, P37). Sometimes, the trigger event is combined with an extra push from another person. P49 started collecting blood sugar level and blood pressure information because of "a doctor's recommendation (new medical issue, new medications)."

STAGES AND BARRIERS

We will now introduce the stage-based model of personal informatics that we derived from analysis of the survey and interview data. The model is composed of a series of five stages (Figure 1): *Preparation*, *Collection*, *Integration*, *Reflection*, and *Action*. We define each stage and, for each one, describe the barriers that participants experienced.

Preparation Stage

The *Preparation* stage occurs before people start collecting personal information. This stage concerns itself with people's motivation to collect personal information, how they determine what information they will record, and how they will record it.

Barriers in the Preparation stage are related to determining what information to collect and what collection tool to use. As noted, some people stumble upon tools, which drive them to start collecting. However, this becomes a problem when the tool does not satisfy their information needs. This causes them to switch to another tool, which has two negative consequences: 1) they abandon their previous data because most systems do not support data exporting, and 2) if they can export data, the formats between the applications may not be the same. For example, P48 used Google spreadsheets to record food and drink consumption then she switched to Daytum where she recorded restaurant information instead. When she discovered your.flowingdata, she returned to recording actual food items. Better preparation in tool selection would have

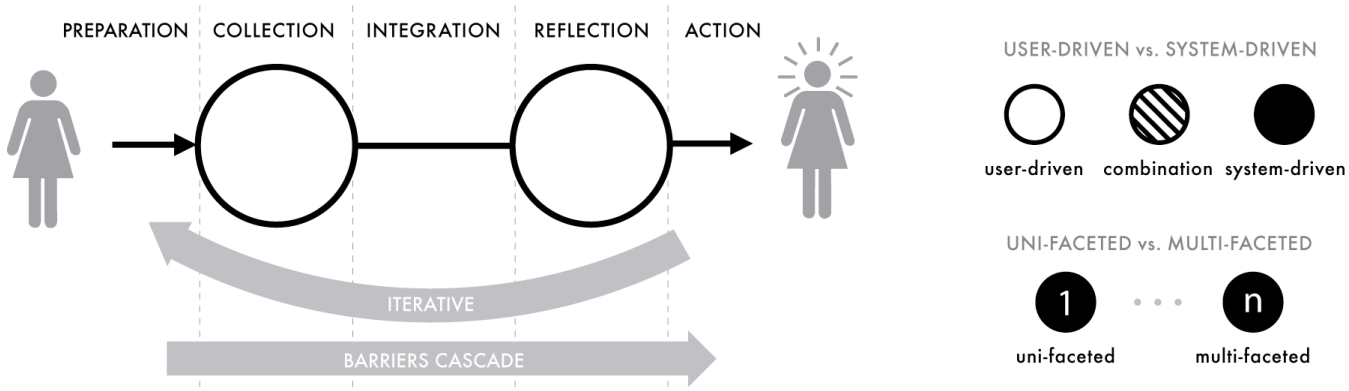

Figure 1. The Stage-Based Model of Personal Informatics Systems and its four properties: 1) barriers in a stage cascade to later stages; 2) stages are iterative; 3) stages are user- and/or system-driven, and 4) uni- or multi-faceted. The visuals for 3) and 4) can be used to show these properties for a particular system.

helped her avoid a gap in her data. Another example is P23 who used spreadsheets to record jogging and biking times before switching to DailyBurn, She did not transfer data from the spreadsheets to DailyBurn, because she would have had to manually transfer the data, which takes time. This is a lost opportunity for more longitudinal reflection about her physical activity.

Collection Stage

The *Collection* stage is the time when people collect information about themselves. During this stage, people observe different personal information, such as their inner thoughts, their behavior, their interactions with people, and their immediate environment. Participants reported different frequencies of collection: several times a day (*e.g.*, food consumption), once a day (*e.g.*, amount of sleep), several times a week (*e.g.*, exercise), or a few times a month (*e.g.*, symptoms, books read).

People encountered several barriers in the collection stage (Table 2). Many of the problems are because of the tool used for collecting information. Some problems occurred because of the user, either because they lacked time, lacked motivation, or did not remember to collect information. Other problems are data-related: 1) data may rely on subjective estimation (*e.g.*, how many calories were expended when lifting weights? P23); 2) data may rely on subjective ratings with no standard for entering data (*e.g.*, P1 wanted to rate his relationship satisfaction, but noticed that his ratings were not consistent); 3) data may be hard to find (*e.g.*, P54 said, "Sometimes life isn't interesting enough to make me want to write it down, other times I can't find any worthy writing material.")

Collection Barriers	Example Quote
Tool (13/68)	"not having ready access to a computer at the time symptoms happen" P6
Remembering (12/68)	"Forgetting to record it. Because I am often not at my personal computer." P57
Lack of time (11/68)	"not difficult, time consuming at times." P16
Finding data (7/68)	"Sometimes life isn't interesting enough to make me want to write it down, other times I can't find any worthy writing material." P54
Accuracy (6/68)	"Guestimating mass of food matching homemade or restaurant foods against database entries" P5
Motivation (5/68)	"keeping up the motivation to do so, finding payback for the investment of time and effort." P4

Table 2. Collection barriers.

Integration Stage

Integration is the stage that lies between the Collection and Reflection stages, where the information collected are prepared, combined, and transformed for the user to reflect on. In Figure 1, the Integration stage is represented as the distance between the Collection and Reflection stages. This distance is determined by how much effort the user has to put into preparing the collected data for the reflection stage. The Integration stage can be *long*, meaning that the user has to do many things to prepare the collected data for the reflection stage. An example of this is when the data to visualize is collected on paper. In such a system, the user has to gather all of his notes and transcribe the data into a graphing application such as Excel. The Integration stage can also be *short*, meaning that the user bears little responsibility in preparing the collected data for reflection. An example of this is Mint, which automatically integrates financial data from bank accounts and credit card companies. Another example is Nike+, which automatically synchronizes runs between an iPod and the Nike+ website.

Integration barriers prevent users from transitioning from collection to reflection of data (Table 3). Users encountered these problems when collected data comes from multiple inputs, reflection of data happens in multiple outputs, and the format of collected data is different from the format necessary for reflection.

Integration Barriers	Example Quote
Transcribing data (10/68)	"It'd be neat if I could graph it straight from the website instead of manually typing in the data to a spreadsheet" P41
Organization (8/68)	"Collecting is simple. Organizing it takes some time." P29
Scattered visualizations (4/68)	"A bit cumbersome going to so many different sites [for visualizations]" P6
Multiple inputs (3/68)	"Difficult to keep organized because sometimes data are kept in separate places" P31

Table 3. Integration barriers.

Reflection Stage

The *Reflection* stage is when the user reflects on their personal information. This stage may involve looking at lists of collected personal information or exploring or interacting with information visualizations. Users may reflect on their information immediately after recording the information (*short-term*) or after several days or weeks involving extensive self-reflection (*long-term*). Short-term reflection is valuable because it makes the user aware of their current status. For example, pedometers show a current aggregate count of steps. In contrast, the BodyMedia SenseWear armband (http://bodymedia.com) does not have a display, so the user is not aware of the amount of calories they have expended until they synchronize with the desktop software. Long-term reflection is valuable because it allows users to compare personal information between different times and it reveals trends and patterns.

Barriers in the Reflection stage prevent users from exploring and understanding information about themselves. These problems occurred because of lack of time or difficulties retrieving, exploring, and understanding information (Table 4).

Reflection Barriers	Example Quote
Lack of time (10/68)	"Having time to go through everything, but that is also one of my biggest pleasures is finding that time." P19
Visualization (6/68)	"It's hard to get a holistic view of the data since the time filters are at most one month and I'd like to look at several months at once." P48
Self-criticism (5/68)	"It's extremely difficult (psychologically) to look back on my earliest journals. Much of that information is very emotional and innocent." P12
Interpretation (5/68)	"Sometimes its very difficult to interpret the media" P54
Search (4/68)	"not too tough. sometimes have to wait while search occurs... but it's a couple minutes at most" P14
No context (3/68)	"Not having an overlay of changes in circumstance" P11
Sparse data (3/68)	"Not enough; My collection of data has been intermittent enough that I don't get good time series." P44
Data is not useful (3/68)	"it's really not very useful and it's kind of annoying. I mean, I walk a lot. What else do I really want to know?" P22

Table 4. Reflection barriers.

Action Stage

The *Action* stage is the stage when people choose what they are going to do with their newfound understanding of themselves. Some people reflect on the information to track their progress towards goals. From the understanding of their information, people may tailor their behaviors to match their goals. Some systems alert the user to take actions. For example, Mint alerts users when their bank account reaches a minimum amount. The WiiFit shows an avatar that acts like a personal trainer; P37 said, "The WiiFit avatar gets excited (or crestfallen) at my progress, which is kind of cool, since s/he acts as a personal trainer." Some systems provide incentives to motivate users to take action. For example, Slife introduced Slife Rewards, which uses donations to selected charities to encourage users to achieve their productivity goals.

Some people reflect on the information to inform them on what actions to take. Most systems do not have specific suggestions on what to do next, which is a barrier to applying understanding of personal information. Different systems have resolved this in multiple ways. Some systems are more explicit such as Mint providing suggestions about bank and credit card accounts with lower fees. Some systems involve input from others. For example, some people collect data and share them with their doctor for advice on next steps (P1, P40, P49). Some systems with sharing capabilities allow feedback from one's social network, *e.g.*, SparkPeople (http://sparkpeople.com), DailyBurn (http://dailyburn.com). P23 said about DailyBurn, "they have forums, which is good to get advice from others."

The stage-based model extends our view of personal informatics systems beyond a pairing of collection and reflection tools. We have just defined the specific stages of personal informatics systems. We will now describe the properties of the model as a whole.

PROPERTIES OF THE STAGES

Here we talk about the holistic properties of the stages and their implications on the development and design of personal informatics systems. We will discuss four properties of the stages: cascading barriers, iteration, user-driven vs. system-driven, and uni-faceted vs. multi-faceted information. We describe the implications of these properties on the design of personal informatics systems.

Barriers Cascade

An important property of personal informatics systems that the stage-based model reveals is that barriers cascade, *i.e.*, problems in earlier stages affect the later stages. Problems in the Preparation stage, such as not using the right tool or not collecting the right data, lead users to change tools, which wastes their time. Since users cannot go back in time and collect the right kind of information, their collected data are rendered somewhat, if not totally, useless. Problems in the Collection stage may lead to sparse datasets that are insufficient for reflection. P44 lacked time and motivation during the Collection stage because he did not have enough data to get a good time-series visualization. When thinking about what he learned from his data, he said "I wish I could report successes on this front, but my lack of regular collection has made this difficult." Problems in the Integration stage, such as scattered visualizations and difficult organization, make reflection harder. P6 noted that "each website has their own way of organizing [data]" and that it was "a bit cumbersome going to so many different sites." Problems in the Reflection stage, such as having trouble using visualizations effectively, prevent users from transitioning to the Action stage. P64 said, "I waste too much time exploring the variations of presentation rather than considering personal changes that might be necessary."

Cascading barriers suggest that a holistic approach to the design of personal informatics systems is critical. Focusing only on one stage ignores the whole experience of the user with the system. While we can take inspiration from different fields to resolve barriers within each stage (*e.g.*, visualization techniques from information visualization research), creating an effective personal informatics system requires the consideration of *all* of the system's parts.

Stages are Iterative

Another property of personal informatics systems is that the stages are iterative, *i.e.*, users incorporate new data, tools, and/or processes as they progress through the stages. We asked our interviewees whether they knew what personal information to collect when they *first* decided to collect information. Many said they knew, but some changed their minds when they found a different collection method. For example, P37 switched between multitudes of tools (whiteboard, spreadsheet, WiiFit) to collect physical activity information. As mentioned before, P48 changed tools as she found new ones (Google spreadsheets, Daytum, your.flowingdata). The iterative process was not supported well in either case. For example, P48 did not transfer her spreadsheet data because Daytum did not support importing data (at the time she used the system). When she switched to your.flowingdata, she could not import her Daytum data.

Some participants changed what personal information they needed while going through the stages, which resulted in tool changes. P23 changed collection tools because of changes in the types of physical activity she performed: 1) spreadsheet for biking and running several years ago; 2) she got a WiiFit last Christmas for physical activity games; 3) she uses DailyBurn now for her gym attendance. P44, who programmed his own visualizations to learn about his sleeping habits, continually added new information (*e.g.*, caffeine consumption, meal times, weight) into the reports depending on his hypotheses (*e.g.*, what is the effect of caffeine consumption on my sleep patterns?).

The iterative property of personal informatics systems suggests that systems should be flexible to support users' changing information needs. Systems could be flexible by 1) supporting importing data from other systems, as well as, exporting to other systems; and 2) supporting different kinds of information. Since tool changes may render previously collected data useless, systems should support rapid iteration, so that users can quickly hone in on the questions they want to answer and select the appropriate tools to collect the necessary information.

User-driven vs. System-driven

Each stage can be classified as *user-driven, system-driven*, or a combination of both. When a stage is user-driven, the user is responsible for the activity in the stage. In the Preparation stage, the user decides what type of information to collect and what tools to use. User-driven collection is when users record information on paper or enter information into a spreadsheet, an application, or a website form. User-driven integration is when users have to transcribe collected information, so that they can reflect on it. In the Reflection stage, the user reflects on the information without the aid of visualizations or other tools for exploring the data. In the Action stage, the user is responsible for deciding on actions to take depending on their conclusions from the data without aid from the system.

User-driven stages demand time and attention from the user. This is problematic when the demand becomes too much (*e.g.*, high frequency of collection). There are ways to facilitate user-driven stages. For example, systems that use user-driven collection can motivate the user to collect the information. Hsieh and colleagues [11] have explored this in the context of experience sampling; they encouraged participation in an ESM study by showing reflective visualizations. An alternative way to dealing with the burden of user-driven stages is to transfer the responsibility completely to the system, *i.e.*, making them system-driven.

When a stage is system-driven, the system takes on the responsibility of performing the tasks in the stage. In the Preparation stage, the system can suggest to users the appropriate tools and information that will help them answer their questions about themselves. In the Collection stage, the system can collect personal information by using sensors (*e.g.*, pedometers), or keeping a record of transactions (*e.g.*, bank statements, search history). System-driven collection can also involve other people, *e.g.*, a nurse writes the data for the patient into a web site. In the Integration stage, the system can aggregate and prepare the information for reflection, *e.g.*, Mint aggregates data from different banks and credit card companies. In the Reflection stage, the system can help users reflect on and explore their information using visualizations. In the Action stage, systems can alert the user to take action, *e.g.*, Wakoopa (http://wakoopa.com) suggests new applications and websites depending on the person's computer usage.

This property of the stages suggests that there are opportunities within each stage of a system to use a system-driven approach to alleviate the demands on the user. However, designers should consider the tradeoffs between a system-driven approach (*e.g.*, inaccuracies of automated analysis, and loss of user control) and a user-driven approach (*e.g.*, burden and complexity). Developers and designers can select which stage could be facilitated by the system to benefit the user the most.

Facets

People's lives are composed of different facets. For example, people have their home life, their work life, their daily interactions with other people, their physical activity, *etc*. Personal informatics systems can be *uni-faceted* or *multi-faceted*. Currently, most systems are uni-faceted, showing only one facet of a person's life (*e.g.*, Mint for financial matters, Nike+ for physical activity, and Slife for productivity). Facets are not necessarily correlated with the number of pieces of information collected by a system. For example, the BodyMedia SenseWear armband is uni-

faceted because all the data it collects (*e.g.*, galvanic skin response, ambient temperature, skin temperature, acceleration) represent only one facet of life, *i.e.*, physical activity. An example of a multi-faceted system is MyLifeBits [7], which collects information about computing activity, web-browsing activity, communication (voice, email, and IM), and media usage (radio and television). Research projects in diabetes management have shown the value of associating multiple facets (*e.g.*, blood sugar level and food consumption) in patients' health [6,17].

Uni-faceted systems simplify the collection and integration of personal information because there is less data to manage. However, uni-faceted systems limit the type of information that people can understand about their life. For example, the WiiFit shows people progress toward their physical activity goals, but there is no awareness of the effect of food consumption, mood, sleeping patterns, and work on physical activity. Systems with multi-faceted collection stages such as MyLifeBits and Daytum allow collection of multiple types of information, which makes collection of data harder, but offers greater potential for becoming aware about different facets of life. Unfortunately, such systems usually present information about multiple facets in separate visualizations. Many participants expressed their desire to see associations between multiple facets of their lives. P26 described his motivation to collect multiple types of information as "to understand trends in symptoms, behaviors, and circumstances." There is an opportunity here for personal informatics systems to provide visualizations (Reflection Stage) that show people the relationships between different facets of their lives. However, the benefit of multiple facets has its cost. P49, talking about tracking medication intake, said, "I suppose if it were easily collected, information on food intake, calories, fat, cholesterol, sodium, *etc*., would make an interesting starting point for analysis. However, if it is too difficult to collect, I have better things to do with my time."

This property of personal informatics systems opens several opportunities for research and applications. Many existing technologies are capable of collecting various types of personal information, and researchers in ubiquitous computing and lifelogging are developing new collection tools. How can personal informatics systems leverage these new technologies to inform people about different facets of their lives? What would visualizations that show multiple facets of people's lives look like? How should they be designed so that non-experts (most people) can gain insights about their lives?

Limitations of the Survey and Model

We recruited participants from a blog dedicated to personal informatics, a blog about general information visualization, and two personal informatics web sites, so most survey participants were familiar with personal informatics. Since even these interested users had plenty of problems with different systems, we suspect that the problems they encountered may be a subset of problems that common users may experience. We think it would be interesting to study users with little or no prior experience with personal informatics systems to find specific barriers that they may encounter.

CASE STUDIES

Having identified the stages model, the properties of the stages, and the barriers experienced within the stages, we can now apply the model to analyze personal informatics systems. In this section, we evaluate existing personal informatics systems to demonstrate the use of the model. In consideration of space, we only highlight some aspects of each system. We describe barriers that users may encounter and suggest potential solutions to explore further.

Twitter-based systems

There are several personal informatics tools that use Twitter as a collection tool. Some collect specific types of information such as eating (http://tweetwhatyoueat.com) and smoking (http://qwitter.tobaccofreeflorida.com), while some collect multiple types of information (*e.g.*, Daytum, Grafitter, Mycrocosm, and your.flowingdata). Users manually enter data into Twitter, but the cost of collection is reduced because users can collect information using different applications and devices (*e.g.*, browser, instant messenger, mobile phone) in different contexts. The Integration stage is also simplified because all data are funneled through Twitter. There are two areas that could be further explored for improvements:

Scattered visualizations. The uni-faceted Twitter-based systems do not have the data to associate multiple facets together. However, the multi-faceted systems have multiple data collected, but do not visualize the data together. P27, who used your.flowingdata, suggested that "relating many datasets at once" would improve the system. This design suggestion could be as simple as associating two pieces of information together. For example, a person who collects information about expenditure and eating locations may see directly how her eating habits are affecting her finances.

Lack of time to collect data. These systems also suffer from the tedium of manually collecting information. P48 said, of her experience with your.flowingdata and Daytum, "I just find it hard enough to track food regularly, so I haven't tracked other things that I would like to." A design suggestion to address this is to automate some of the data collection. Not all collection can be automated, but some data can be. One inspiration is the concept of *tweetjects* [1], sensors and devices that post to Twitter. Some tweetjects post personal information, such as @gareth_laptop (nearby devices), @andy_house (electricity usage), and @kickbee (baby activities in the womb).

Mint (finance)

Mint is a personal finance management website. Mint is mostly system-driven. Mint provides some support in the Preparation stage by helping users select the banks, credit cards and investments that the system will integrate. The

Collection stage is completely system-driven, supported by the bank and credit card infrastructures that record financial transactions. Mint integrates your transaction records from multiple bank, credit card, and investment accounts. P9 noted that one of the flaws of Mint is that it "automatically categorizes [transactions], which works 95% of the time, but not always. I need it to be almost perfect to use it for more than curiosity." The system-driven integration is not perfect, so manual integration is still needed to provide category labels for unlabeled transactions and to fix mislabeled transactions. It supports reflection with visualizations of a user's spending between different categories. Mint is one of the few personal informatics systems that have system support for the Action stage; it provides suggestions on how a user can save money by finding banks with discounted services and credit card companies with low rates, and alerting users of low balances and unusual activity.

No support for multi-faceted reflection. Reflection within Mint is uni-faceted. P12 noted: "I now want to record all the minutiae of my personal life that aggregates into interesting data. I want to graph the people I see, the things I do, the hours I devote to every significant task, and the money I spend and *why* (our emphasis). I want to have yearly data that shows, for example, that I spent 1,000 hours on programming, but only 400 on reading, or that I spent twice as much in coffee shops as I did on groceries." She does not use Mint, but would have found it useful for comparing her expenditures at coffee shops and groceries; however, Mint's visualizations would not help her answer the "why" question. She speculated that there are associations between different facets of her life; allowing input about different life facets in Mint may help her. Since Mint is largely system-driven, adding extra information to associate with spending may be a small burden to users, even if manually collected.

IMPACT (physical activity)

IMPACT is a mobile phone and web site system for collecting and reflecting on physical activity, with which we had first-hand experience in deploying to users for several weeks [14]. We highlight two problems with IMPACT that could be further explored for solutions.

No support for Preparation stage. The IMPACT system did not have support for the Preparation stage. The IMPACT system collected four types of information: step counts and the context in which those steps were taken (location, type of activity, and whom the person was with). The system imposed these types of information without determining whether the user might find all the information useful. We learned that users were interested in the effects of mood and weather on physical activity, but IMPACT did not support these types of information. Problems in the Preparation stage affected the Collection stage leading to barriers such as too much information to collect. Collecting the three types of context unnecessarily burdened the user, when other types of information may have been more useful (*e.g.*, mood, weather). There were other collection barriers such as intrusiveness (experience sampling type alerts), and the use of separate devices for monitoring step counts (mobile phone) and location (GPS device).

User-driven Integration. An initial version of IMPACT required users to transcribe step counts, time, and contextual information from a notebook to a web site. We fixed this in the current version with a system-driven integration, where the system automatically transfers data collected on a mobile phone to a web site. We could have avoided the problem with the initial prototype if we had considered the Integration stage earlier.

Notice that the two problems described above could have been avoided if we had analyzed IMPACT using the stage-based model of personal informatics systems earlier in its development. Addressing these problems may improve users' experience with the system and yield results for the use of context in improving monitoring of physical activity.

DISCUSSION

We have defined a stage-based model of personal informatics systems and identified a comprehensive list of the problems that people experience in each of the stages. We also described the properties of the stages, which have implications in the design of personal informatics systems. To build effective personal informatics systems, developers and designers should consider the following:

- Since barriers cascade to later stages, designers and developers should consider the system as a whole. This holistic approach requires integrating innovations and applying lessons from different areas of research, such as lifelogging, ubiquitous computing, information visualization, and persuasive technologies.
- Since users iterate in the kinds of questions they ask and the tools they use, flexibility within a system and between systems is important. Systems should allow users to easily change what kind of data they collect dependent on their needs and to transfer data from one system to another.
- There are opportunities within each stage to use a system-driven approach to reduce the demand on users and to make the experience more enjoyable and useful. Similarly, a user-driven approach is appealing because it leaves control in the hands of the user. Insights from the field of human-computer interaction should play a big part in the development of mixed-initiative approaches for effective personal informatics systems.
- Currently, most personal informatics systems are uni-faceted. Participants expressed desire for associating different aspects of their lives together. Creating multi-faceted systems may be difficult because of the extra data that need to be collected, however the insights gained may be worth it. There are opportunities to leverage the increasing ability of ubiquitous computing and lifelogging technologies to collect various types of

information and show users multiple facets of their lives. These multiple facets must be displayed in ways that users can understand them and gain valuable insight into their lives.

CONCLUSION

Personal informatics is a growing class of systems that can help people know themselves better. More people are building such systems and there is an opportunity to provide design recommendations for creating more effective systems.

In studying how people use personal informatics systems, we gained an understanding of their motivations and practices. We used this to define a model that describes personal informatics systems as a series of 5 stages. We compiled a comprehensive list of barriers that people encountered in each stage. We also described properties of the stages: cascading effect of barriers, iterative, system-driven vs. user-driven, and uni-faceted vs. multi-faceted. These properties have implications in the design of personal informatics systems: 1) they should be designed holistically; 2) they should improve support for iteration between stages; 3) an appropriate balance of automated technology and user control should be applied within each stage to facilitate the user experience; and 4) they should explore support for multiple facets of people's lives to enrich the value of systems.

We believe the model and its properties will be valuable for future research and development because it provides a common framework for describing, comparing, and evaluating personal informatics systems. Personal informatics systems pose new interesting HCI challenges. We believe that the stage-based model, the identification of the barriers within each stage, and the description of the properties of the stages will help HCI researchers and practitioners find solutions and explore new approaches in personal informatics more efficiently.

ACKNOWLEDGMENTS

We would like to thank our survey and interview participants and the websites that helped our recruitment.

REFERENCES

1. Alleyne, R. "The house that Twitters." *Telegraph.co.uk*, September 8, 2009. Accessed on September 16, 2009 at http://www.telegraph.co.uk/science/science-news/6156291/The-house-that-Twitters.html
2. Assogba, Y., and Donath, J. Mycrocosm: Visual Microblogging. *HICSS'09*, 2009, pp. 1-10.
3. Consolvo, S., McDonald, D.W., Toscos, T., *et al*. Activity Sensing in the Wild: A Field Trial of Ubifit Garden. *CHI'08*, 2008, pp. 1797-1806.
4. Froehlich, J., Dillahunt, T., Klasnja, P., *et al*. Ubigreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits. *CHI'09*, 2009, pp. 1043-1052.
5. Franklin, B. Autobiography of Benjamin Franklin. New York, 1916, pp. 146-155.
6. Frost, J. and Smith, B.K. Visualizing Health: imagery in diabetes education. *DUX '03*, 2003, pp. 1-14.
7. Gemmell, J., Bell, G., and Lueder, R. MyLifeBits: a personal database for everything. *Communications of the ACM*, 2006, pp. 88-95.
8. Gemmell, J. & Vemuri, S. "CARPE Research Area." *SIGMM*. Accessed on January 6, 2010 at http://www.sigmm.org/Members/jgemmell/CARPE
9. Hallnäs, L. and Redström, J. Slow Technology: Designing for Reflection. *Personal and Ubiquitous Computing,* 5(3), 2001.
10. Hodges, S., Williams, L., Berry, E., *et al*., K. SenseCam: a Retrospective Memory Aid. *Ubicomp'06*, 2006, pp. 177-193.
11. Hsieh, G., Li, I., Dey, A., Forlizzi, J., and Hudson, S.E. Using Visualizations to Increase Compliance in Experience Sampling. *Ubicomp'08*, 2008, pp. 164-167.
12. Intille, S.S., Tapia, E.M., Rondoni, J., *et al*. Tools for Studying Behavior and Technology in Natural Settings. *Ubicomp'03*, 2003, pp. 157-174.
13. Jones, W., and Teevan, J. *Personal Information Management*. UW Press, 2007.
14. Li, I., Dey, A., and Forlizzi, J. Using Contextual Information to Improve Awareness of Physical Activity. *Engaging Data Forum 2009*, Boston, MA, 2009.
15. Lin, J.J., Mamykina, L., Lindtner, S., *et al*. Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. *Ubicomp'06*, 2006, pp. 261-278.
16. Maitland, J., Sherwood, S., Barkhuus, L., *et al*. Increasing the Awareness of Moderate Exercise with Pervasive Computing. *IEEE Pervasive Health Conference*, 2006, pp. 1-9.
17. Mamykina, L., Mynatt, E.D., and Kaufman, D.R. Investigating health management practices of individuals with diabetes. *CHI'06*, 2006, pp. 927-936.
18. Mankoff, J., Kravets, R., and Blevis, E. Some Computer Science Issues in Creating a Sustainable World. *Computer*, 41(8), 2008, pp. 102-105.
19. Pousman, Z., Stasko, J.T., and Mateas, M. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics*, 2002, pp. 1145-1152.
20. Scollon, C., Kim-Prieto, C., and Diener, E. Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Studies*, 4, 2003, pp. 5-34.
21. Wolf, G. Know Thyself: Tracking Every Facet of Life, from Sleep to Mood to Pain, 24/7/365. *Wired*, 17.07, 2009, pp. 92-95.
22. Yau, N. and Schneider, J. Self-Surveillance. *Bulletin of ASIS&T*, June/July 2009, pp. 24-30.