Personal Bits: Mining Interaction Traces for Personalized Task Intelligence

Fanglin Chen
fanglin@cmu.edu
Carnegie Mellon University
Pittsburgh, USA

Jason I. Hong
jasonh@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, USA

ABSTRACT

As we work, play, shop, and communicate in digital interfaces, we continuously generate traces of information. To turn such noisy sources of personal data into actual insight, my research introduces *Personal Bits*, a service that enables personalized task support in various kinds of information tasks such as instant message handling, information retrieval, and text entry. *Personal Bits* mines a user's interaction traces with web apps and native mobile apps and extracts task-centric entities. I present three example apps for *Personal Bits*: *Deja Wu*, *MessageOnTap*, and *ContextBoard*, to address inefficiencies presented in these information tasks. *Personal Bits* acts as the central nexus for intelligence between apps and interaction traces, making it easy for apps to acquire personally relevant task entities in fine granularity.

CCS CONCEPTS

• **Information systems** → **Users and interactive retrieval**; *Task models*.

KEYWORDS

Task Intelligence; Personal Data; Intelligent User Interface; Interaction Traces; Productivity; Messaging; Information Retrieval; Language Model

ACM Reference Format:

Fanglin Chen and Jason I. Hong. 2019. Personal Bits: Mining Interaction Traces for Personalized Task Intelligence. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*, September 9–13, 2019, London, United Kingdom. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3341162.3349310>

1 PROBLEM STATEMENT

Our digital life is composed of various kinds of small information tasks, such as looking for a document, creating a new contact, or replying to an incoming text message. In my empirical research on such short-term informational tasks, we found people spend a great deal of time handling tasks prompted by instant messages, can be easily at lost in looking for something seen before, and

sometimes experience inefficiencies in communicating personal data in text entry. Everyday user interfaces such as messaging apps, browser search bars, and keyboards are frequently visited to complete various tasks, yet are still rather rigid in supporting users.

Personal data is the key to personalize these short-term informational tasks. Computers accumulate a massive amount of personal data that can tell where we go (from location-based apps), whom we know (from contact apps) and talk to (from phone dialer), what we are interested in (from browsers). However, much as it manifests, personal data is still siloed in different web and mobile apps. This hurts interoperability and therefore results in fragmented user experience in completing daily tasks. For instance, if a user, who has just booked a flight using one app, wishes to reserve a cab from the airport, she would have to do so by manually re-entering the location, date, and time from the flight app into the cab app. To enable users to retake control of their siloed data, better interoperability between apps, and more personalized task support, my thesis explores *interaction mining* techniques to produce structured and semantically meaningful information out of user interaction traces. These task-oriented entities are “personal bits” that can characterize tasks and provide feedback back to the user interface to generate personalized task support.

This proposal starts by discussing prior work related to personal data, task intelligence, and interaction trace mining. Next, I present three empirical studies on instant messaging, information seeking, and mobile text entry, which helps to understand how to address task inefficiencies. Through the collected empirical data, we also seek the common taxonomic structure of entities related to these informational tasks. I then present methods to extract personal bits through a continuous interaction mining service running in web browsers and mobile phones, which keeps synchronizing information to a private cloud of the user. Finally, as shown in Figure 1, I introduce three personalized task supports powered by personal bits. 1) *MessageOnTap* provides users quick access to relevant external apps and task shortcuts to get tasks done without leaving the conversation, based on the personal bits extracted from the conversation (e.g., date time, restaurant name, personal contact name, etc.). 2) *Deja Wu* helps users to reach their previously visited web contents faster, by matching user query to associative memories (e.g., type of content to look for, author of the post, or original query used to land on the content) they might have on their prior browsed content. All the associative memories are based on the personal bits collected through prior user web traces. 3) *ContextBoard* unifies all extracted personal bits into a language model, presenting users with input suggestions while they are typing in the keyboard.

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. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6869-8/19/09...\$15.00

<https://doi.org/10.1145/3341162.3349310>

Figure 1: Personal Bits underpins personalized task support in the browser search bar, keyboard, and messaging app. 1. Tina has been looking into getting her real ID at Pittsburgh. After clicking through several web pages, she figured out the right DMV to go to. Her search queries, clicked items and concrete nouns in her viewed regions of interest were all used to construct a meaningful “personal bit” (i.e., related information of the DMV office) in the background. 2. Earlier today, she planned to go to the DMV office but forgot about which one to go to (i.e., Bridgeville). She can still find out the opening hour of this DMV office using Deja Wu. 3. After getting all the required materials ready, she ordered a Uber to get to the Bridgeville office. ContextBoard suggests her with the matching address after Tina typed Bridgeville. 4. In a conversation with her friend, ContextBoard suggests her recently mentioned location while she types. Once settled with a time to meet later today, MessageOnTap suggests her a shortcut to create a new calendar item (i.e., new event: Hidden Harbor, 6 pm).

2 RELATED WORK

Personalized task intelligence has its root in personal data. However, due to the siloed ecosystem [19], apps often lack a holistic user understanding. In related work, I start by revisiting the “siloed app” problem and prior attempts to address this problem to facilitate tasks. Then, I will move on to how to break app silos through *interaction mining* (i.e., information extraction from user interfaces).

2.1 Personal Data for Task Intelligence

To personalize task support, task interface needs to access the context of the ongoing task. Context, as any information that describes the user’s current interaction with the interface [9], can even include prior user interactions in completing the tasks [3]. Therefore, personalized task intelligence entails a much richer picture of personal activities than just the current time zone and current location, which are commonly used in conversational agents such as Siri, Alexa, and Cortana.

It is very challenging for individual apps to collect a wide range of personal data they might need to model ongoing task context, due to the complexity of interfacing different forms of data through

separate APIs [17]. Most apps and services also do not have APIs to make their siloed personal data available. On smartphones, information platforms such as Google and Apple now start to allow apps to index their content [10, 13]. This device-wide centralization of personal content can help users address their ultimate information need faster, without juggling between different apps to search around. However, app indexing requires significant developer efforts, and the coverage of the indexed in-app contents is quite low.

There is also a lack of understanding of what kinds of data constitute informational tasks. Despite the richness of different kinds of short-term informational tasks, the personally relevant data that flows in these tasks is often not as complex. Take the scenario in Figure 1 as an example. Tina’s mental model centers around one key concept only - “Applying for the Real ID at DMV”. This concept is reflected in personal inputs in different interfaces (i.e., Chrome, the Uber app, and Facebook Messenger). Similar to Twitter’s interest graph [18] which organizes interest, Facebook’s social graph [11] which organizes social connection, and Google’s knowledge graph [8] which organizes public-domain factual knowledge, one goal of this dissertation is to seek the common taxonomic structure in the realm of short-term informational tasks. Existing methods

on unifying personal data touch on limited ontology categories, most of which are not task-oriented. For example, Epistynet [5] and Google App Indexing [14] define Message and Person as ontology types, but there are further task-relevant concepts that can be unpacked from Message and Person. An incoming Message “*check out Kanye West’s latest ig post!*” contains two critical bits of information related to the upcoming task the user is likely to do. “ig” (i.e., the Instagram app) specifies the app of interest, and “Kanye West” (otherwise classified as Person) specifies a celebrity music artist the user is likely to look up next.

Personal Bits sits on a task-centric taxonomic structure of in-app contents. To derive this structure, I curate a task-related corpus based on the field data from three empirical studies on different informational tasks. These tasks are common tasks that users engage in, including information seeking, mobile text entry and tasks prompted by instant messages. This structure acts as a layer of abstraction between the unstructured interaction traces and intelligent task support used in different kinds of interfaces.

2.2 Mining Task Semantics in UIs

Personal Bits explores the use of interaction traces users have on mobile and web apps as the resource to unify personal data and break app silos, minimizing developer effort in integrating and specifying the kind of data to be indexed to the user. Given the limited screen real estate, app developers typically have proper description and hierarchy in their UIs, in order to help users navigate their tasks. A continuous collection of interaction traces allows us to achieve holistic user profiling, making it easy to understand user intents in navigating different interfaces to complete tasks.

Interaction traces have been studied in the information retrieval community to improve search relevance [1], primarily in the context of desktop web browsing. Even user semantics can be extracted by open Accessibility APIs from mobile app pages, it is not until recently that interest of interaction mining began to resurface for mobile. Deka *et. al* [6, 7] collected a large-scale dataset of mobile app interaction traces to discover patterns of mobile app design. The most relevant effort to task semantics mining involves screen-based search. Search engines (Bing snapshots [15], and Google screen search [12]) can suggest users with relevant public information and mobile deep links related to what being presented on the smartphone screen. The limitation of screen-based search is that it only supports information needs that are easily identifiable using public knowledge graph. For example, it detects person name and links it to a public Wiki, but cannot tell whether a person is a personal contact or not. Since the extraction method only applies to each individual text field, it also cannot make much sense of the unstructured texts such as emails and instant messages. Furthermore, as tasks usually span across a series of screens, the semantic mining service should have the capability of making sense of task interaction through traces rather than screens separately.

Personal Bits has the following advantages in information extraction: 1) Platform-agnostic. The extraction service works on both web and mobile apps. 2) Continuous. *Personal Bits* operates on the notion of tasks rather than single screens and can make sense of user goals based on consecutive screens in the task. 3) Personal. Given the rich user-generated contents collected through *Personal*

Bits, recognizing textual contents can not longer solely rely on public knowledge. 4) Task-centric. The ontology types defined in *Personal Bits* capture relevant semantics in informational tasks and are directly indicative of how users complete tasks.

3 METHODOLOGY

The key idea of *Personal Bits* is to transform interaction traces into structured task-oriented entities, in order to help apps to design support for many kinds of informational tasks. To outline a taxonomy of these task-related entities, I conduct empirical studies that investigate information seeking, mobile text entry, and handling instant messages. This allows us to systematically analyze what kinds of personal data are being exchanged in instant messages, how people describe a piece of information to look for, and what types of data are being entered on a mobile phone.

The empirical studies also help us to identify costs in completing these three informational tasks, which informs us how to design elements of personalized task support in the browser search box, messaging app, and keyboard. After describing empirical studies, I introduce how to extract personal bits from interaction traces, and how do the extracted personal bits enable the intelligent task support in these informational tasks.

3.1 Empirical Studies

3.1.1 Study 1: Handling Mobile Instant Messages. We recruited 22 students (mean age 24.8, 11 females) through flyers posted on campus and participated in our experiment for 2-4 weeks. Our in-situ data loggers collected 19015 exchanged messages. The interaction traces in the external apps after users switch out of the conversations were annotated as to whether are message-related tasks or not, given the relevance between the conversation and the texts shown in the external apps.

3.1.2 Study 2: Breakdowns in Past Information Retrieval. To identify opportunities in personalizing task supports for information access, we have run an observational study (N=21) to learn user strategies in looking for a previously seen content and the breakdowns along this process. Participants were recruited in a lab and instructed to demonstrate their recent revisitation to a piece of information. While demonstrating, they were asked to describe what they look for and why. The demonstrations were recorded in videos for interpretation.

3.1.3 Study 3: Mobile Text Entry. We conducted a mixed-method study including surveys, interviews, diaries, and in-situ logging to characterize the users’ mobile text entry behaviors across all apps. In our logging study, we recorded 401,875 were text-editing events, representing 476,322 characters entered into 155,300 app usages of 364 unique apps over 2 weeks (N=17). We also gathered reflections on the text entry through interviews and diaries.

3.2 Extracting Personal Bits

To derive the taxonomic structure of task-centric entities, we use the collected field data to generate individual task-relevant texts, which will be used for classification in a later card-sorting session. Example incidents include a destination location to enter in a ride-sharing app (from Study 1), a date-time specified for an upcoming

meeting (from Study 2), and a meeting room number mentioned in an email (from Study 3).

To extract task-centric entities, *Personal Bits* continuously analyzes the render-time properties of the nested elements in the view hierarchy, and the user actions performed on a UI state resulting in transitions to other UI states within the application. Since implicit user behaviors, such as copy-paste and highlighting can tell a lot of user preference towards the viewed contents [16], they will be also considered while extracting personal bits. Extracted personal bits will be associated with one entity category defined in the taxonomic structure, based on a hybrid algorithm consists of existing personal bits matching, public knowledge graph retrieval, and natural language processing. The extraction service contains page-level analytics on each screen in the trace, to identify metadata for contents such as emails (e.g., email title, etc.), instant messaging conversations (e.g., sender name, etc.), and web page (e.g., date of the post, etc.). Finally, there is a session tracker to segment different informational tasks according to both the temporal and semantic patterns of the consecutive screens.

3.3 Supporting Personalized Task Intelligence

3.3.1 MessageOnTap. In Study 1, we found that when prompted by instant messages for tasks, users can spend a significant amount of time looking for the right app. Texts in the conversation are indicative to the actual task completed in the external app, adding more unnecessary costs in text entry. To mitigate such costs, MessageOnTap [4] suggests users with apps and even shortcuts for them to complete the task (see 4 in Figure 1).

One limitation of the current MessageOnTap prototype is its scalability. It only supports five common apps and the extracted entities are limited to generic types such as location, person. We are developing information extraction method that can support entity recognition in conversation in finer granularity (e.g., restaurant name, personal contact name, etc.) to support template matching of tasks in much more apps.

3.3.2 Deja Wu. Modern browsers do use users' web histories to enhance the search result in the browser search box, but they only do so by matching user queries to the web titles of the previous web histories. However, through our interview results in Study 2, we find that users can easily lose track of the web titles, for instance, during when they were sent with a link from others or were overwhelmed with a lot of contents when they saw the content for the first time. For example, a participant described she had trouble finding a video she saw earlier. All she can remember is that the video was originally explored from an answer posted on Quora about job interviews.

Figure 2: Deja Wu provides direct answers in the search box by mining user web traces. It matches user query to all the relevant personal bits (e.g., preceding website, data type, etc.) around a web content (i.e., a video).

After spent several minutes in going through the web history, she realized that the video title does not include the keyword “job”.

To circumvent such breakdown, Deja Wu (Figure 2) matches personally relevant descriptions to retrieve content visited before. Users can issue intuitive queries according to their associative memories to the content, specifying how they originally landed on the content (e.g., “Quora” in Figure 2), the type of the content (e.g., “hours” in Figure 1-1), or other contextually relevant information. All of these associative memory cues are captured in personal bits. Explanations that justify the suggestions [2] are also provided to signify the search box’s capability of learning past interactions.

3.3.3 ContextBoard. In Study 3, we found structured data such as URLs and email addresses are rarely typed, copy-and-paste is rarely used, and that sessions of smartphone usage with text entry involve more apps and last longer. To streamline the text entry, we propose a language model that is grounded on the personal contexts, infusing personal bits into the input suggestions that are available on the user’s keyboard. Training such a language model typically would require a large-scale dataset of individual personal data and their corresponding text entries. To minimize privacy concerns, we introduce a synthetic dataset, based on the crowd-sourced text entries grounded on activity logs collected from the American Time Use Survey (ATUS), which keeps the naturalistic human constraints (e.g., number of daily activity, diversity of activity types, and statistical validity). The goal of ContextBoard is to help users to autocomplete texts related to their personal bits, as shown in Figure 1 2&3.

4 EVALUATION

To evaluate the information extraction method, and the usefulness and the usability of the task support powered by personal bits, we will deploy Deja Wu for 2 weeks, conduct a large scale evaluation on MessageOnTap’s recommendation algorithm on a task-related instant message dataset, and conduct human evaluation on the fluency (syntactically sound), informativeness (enough personal bits) and faithfulness (right personal bits) the input suggestions generated by ContextBoard. We will also design APIs to make personal bits accessible to third-party apps. To evaluate the usability of the APIs, we will recruit developers to use Personal Bits APIs to design intelligent task supports in think-aloud protocols.

5 EXPECTED CONTRIBUTION

The underlying motivation for my thesis research is the unwavering belief that with the ubiquitous access of information, intelligent agents in many forms can adapt to facilitate informational tasks with user understanding in full-bandwidth. My dissertation contributes by breaking app silos through mapping interaction traces into the structured, task-centric personal bits, and demonstrate the effectiveness of personal bits in creating new kinds of intelligent task supports in messaging, information seeking and text entry. With the structured representation of task-centric entities, and the service which generates these task-centric entities, *Personal Bits* is the centerpiece that connects personal interaction traces to individual apps, which empowers different kinds of personalized task support with minimal effort in acquiring and processing personal data.

6 BIOGRAPHICAL SKETCH

In August 2019, Fanglin Chen will start his fifth year as a Ph.D. student at School of Computer Science in the Carnegie Mellon University. Fanglin's advisor is Dr. Jason I. Hong, Professor in the School of Computer Science. They are both affiliated with the Human-Computer Interaction Institute in Carnegie Mellon. Fanglin is expected to graduate at the August of 2020.

REFERENCES

- [1] Eugene Agichtein, Eric Brill, Susan Dumais, and Robert Ragno. 2006. Learning user interaction models for predicting web search result preferences. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 3–10.
- [2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for Human-AI Interaction. (2019).
- [3] Matthew Chalmers. 2004. A historical view of context. *Computer Supported Cooperative Work (CSCW)* 13, 3-4 (2004), 223–247.
- [4] Fanglin Chen, Kewei Xia, Karan Dhabalia, and Jason I Hong. 2019. MessageOnTap: A Suggestive Interface to Facilitate Messaging-related Tasks. (2019).
- [5] Sauvik Das, Jason Wiese, and Jason I Hong. 2016. Epistenet: facilitating programmatic access & processing of semantically related mobile personal data. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 244–253.
- [6] Biplab Deka, Zifeng Huang, Chad Franzen, Joshua Hibsman, Daniel Afergan, Yang Li, Jeffrey Nichols, and Ranjitha Kumar. 2017. Rico: A mobile app dataset for building data-driven design applications. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, 845–854.
- [7] Biplab Deka, Zifeng Huang, and Ranjitha Kumar. 2016. ERICA: Interaction mining mobile apps. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 767–776.
- [8] Google Developers. 2018. Google Knowledge Graph Search API. <https://developers.google.com/knowledge-graph/>.
- [9] Paul Dourish. 2004. What we talk about when we talk about context. *Personal and ubiquitous computing* 8, 1 (2004), 19–30.
- [10] Apple Inc. 2018. App Search Programming Guide: App Search Programming Guide. <https://developer.apple.com/library/archive/documentation/General/Conceptual/AppSearch/AppContent.html>.
- [11] Facebook Inc. 2018. Graph API. <https://developers.facebook.com/docs/graph-api/>.
- [12] Google Inc. 2018. Find info about what's on your screen. <https://support.google.com/assistant/answer/7393909?co=GENIE.Platform%3DAndroid&hl=en>.
- [13] Google Inc. 2018. Firebase App Indexing. <https://firebase.google.com/docs/app-indexing/>.
- [14] Google Inc. 2018. Firebase App Indexing. <https://firebase.google.com/docs/reference/android/com/google/firebase/appindexing/builders/Indexables>.
- [15] Jacob Kastrenakes. 2018. Find info about what's on your screen. <https://www.theverge.com/2015/8/20/9181889/bing-launches-android-google-now-on-tap-competitor>.
- [16] Diane Kelly and Jaime Teevan. [n. d.]. Implicit feedback for inferring user preference: a bibliography.
- [17] Yuanchun Li, Fanglin Chen, Toby Jia-Jun Li, Yao Guo, Gang Huang, Matthew Fredrikson, Yuvraj Agarwal, and Jason I Hong. 2017. Privacystreams: Enabling transparency in personal data processing for mobile apps. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 76.
- [18] Techcrunch. 2018. How Twitter Is Pairing Its Interest Graph With Ads. <https://techcrunch.com/2012/03/01/how-twitter-is-pairing-its-interest-graph-with-ads/>.
- [19] Jason Wiese, Sauvik Das, Jason I Hong, and John Zimmerman. 2017. Evolving the ecosystem of personal behavioral data. *Human-Computer Interaction* 32, 5-6 (2017), 447–510.