Building a Benchmark for Task Progress in Digital Assistants

Jonathan Liono[∗]
RMIT University
Melbourne, Australia
jonathan.liono@rmit.edu.au

Johanne R. Trippas[∗]
RMIT University
Melbourne, Australia
johanne.trippas@rmit.edu.au

Damiano Spina
RMIT University
Melbourne, Australia
damiano.spina@rmit.edu.au

Mohammad S. Rahaman
RMIT University
Melbourne, Australia
saiedur.rahaman@rmit.edu.au

Yongli Ren
RMIT University
Melbourne, Australia
yongli.ren@rmit.edu.au

Flora D. Salim
RMIT University
Melbourne, Australia
flora.salim@rmit.edu.au

Mark Sanderson
mark.sanderson@rmit.edu.au
RMIT University
Melbourne, Australia

Falk Scholer
falk.scholer@rmit.edu.au
RMIT University
Melbourne, Australia

Ryen W. White
ryenw@microsoft.com
Microsoft
Redmond, USA

ABSTRACT

Understanding how tasks progress over time enables digital assistants to help with current activities and support future activities. Imbuing assistants with the ability to track task progress requires machine-learned models. In this paper, we describe an ongoing effort to collect signals from Cyber, Physical, and Social (CPS) activities, together with human assessments of task progression, to serve as a benchmark for training and testing such models. Collecting this data over time is inherently challenging in the daily sensing scenario. Consequently, lessons learned from our ongoing data collection are discussed to provide more insights for future research innovations in task intelligence.

ACM Reference Format:

Jonathan Liono, Johanne R. Trippas, Damiano Spina, Mohammad S. Rahaman, Yongli Ren, Flora D. Salim, Mark Sanderson, Falk Scholer, and Ryen W. White. 2019. Building a Benchmark for Task Progress in Digital Assistants. In *Proceedings of Task Intelligence Workshop at WSDM 2019 (TI@WSDM19).* ACM, New York, NY, USA, 6 pages.

1 INTRODUCTION

Intelligent digital assistants such as Google Assistant, Amazon Alexa, or Microsoft Cortana provide a useful way for people to manage their daily tasks. In particular, digital assistants have the potential to understand and predict the tasks the assistant's user will undertake.

We define a *task* as work that needs to be completed by a user. The task has a goal, progression, and completion status. It is composed of one or more activities and takes place in associated CPS contexts. To our knowledge, there is no rich dataset of user tasks, activities, and contexts to understand how users perform tasks and define the CPS signals associated with them.

[∗]Both authors contributed equally to the paper.

TI@WSDM19, February 15, 2019, Melbourne, Australia

Digital assistants can monitor a user's CPS activities to form a rich multi-dimensional CPS context. Ideally, this context can be used to effectively model and predict the personal tasks a user is undertaking. The assistant could also provide appropriate and aptly-delivered activity and information suggestions relevant to user tasks. Previous work [13] has shown the value of modeling a user's cyber (online) activities together with their physical location and trajectory. Automatically determining the current stage of a task could enable digital assistants to help users make task progress (through relevant recommendations), handle completion activities (such as summarizing work done), and so on. However, there are no benchmark resources to validate the assumption that CPS signals can be used to monitor the progress of the tasks performed by the users of a digital assistant.

The methodology described in this paper is currently being applied to generate a dataset that combines CPS signals with two different methods to collect user labels on task progression: the Experience Sampling Method (ESM) [5] and the Daily Reconstruction Methods (DRM) [6]. ESM provides a mechanism for conducting in situ sampling (e.g., using periodical and on-demand surveys through a mobile device) [10], whereas DRM uses a post hoc survey to ask participants to recollect and label a sample of their tasks, activities, and context. The combined dataset will enable the building and evaluation of models for task progress and recommendation.

2 RELATED WORK

Previous personal assistant research has investigated what users search for with assistants [3], how users interact with such systems [14], modeling user interests [20], recognizing intent from spoken and written utterances [7, 16, 19], tracking user intents [15], and helping users by providing information dynamically [1].

The notion of Cyber-Physical Systems was described by Lee [9] as systems "with feedback loops where physical processes affect computations and vice versa". Although user context is commonly exploited in personal assistant systems, particular aspects such as time, location, and demographics have mostly been studied separately [2, 4, 8, 11].

To the best of our knowledge, there is no research where heterogeneous contexts that describe cyber, physical, and social behaviors have been integrated together to enhance personal assistants. Although mobile-based experience sampling in integration with sensor data logging has been studied for psychological and behavioral studies such as emotion sensing [12], we are unaware of prior work where such sampling has been integrated with task understanding, a focus of the study described in this paper.

3 METHODOLOGY

The proposed methodology to build a task progress dataset comprises several CPS signals, captured by different mobile and desktop apps, and associated to manually collected annotations of tasks.[1]

In the following sections we describe the procedure to collect the data (Section 3.1), the CPS signals collected (Section 3.2), and the annotations provided by the participants via both ESM and DRM (Section 3.3).

3.1 Procedure

The data is collected by asking a set of participants to install a set of apps on their devices and perform annotations in situ and at the end of the day. This procedure consists of five phases (recruitment of participants, preparation of accounts, intake, weekly meetings, and end of the study) which are detailed below.

Recruitment of Participants. Participants are recruited via an Online Recruitment System for Economic Experiments (ORSEE)[2] managed by RMIT University. This system has exclusive access to a pool of more than 1,000 potential participants. The user study was also advertised via online classified advertisements and community websites. Potential participants completed an Expression of Interest form, which acts as a 'screener' to identify suitable candidates. The requirements for participation in the study are: willingness to actively participate for four consecutive weeks, and ability to come to the intake and weekly meetings; being employed part- or full-time; owning an Android smartphone with version $\geq$ `5.0`,[3] and not having an anti-virus app installed or being willing to uninstall such apps. Participants that expressed interest and met the requirements were invited to an intake session.

Accounts for Participants. For each of the invited participants, the researchers create individual e-mail accounts that will be used by the participants to log into apps for collecting the data.

Intake. The intake session is guided by the researchers who explain the aim of the study, ensure the plain language participant information sheet is read and signed, assist participants to install the apps on their devices, and answer any questions participants may have.

Individual Weekly Meetings. Individual weekly meetings are held between a research investigator and each of the participants. These meetings have the following purpose: (i) verify that the mobile apps used to collect signals (as described in Section 3.2) are running appropriately (i.e., not being restricted for background processing on the phone); (ii) ensure a level of consistency in the way participants annotate tasks (e.g., ensure they understand the notion of *tasks* in a consistent way); (iii) remind the participants of the importance of the annotations and they are going to be used, e.g., to link between different annotated tasks, to learn/predict/measure progress of tasks); and (iv) correct or extract additional information that might not be captured during the ESM and DRM surveys. Researchers created weekly dashboards of the survey answers, in the form of printed spreadsheets of daily (from DRM-based survey) and hourly (from ESM-based survey) task annotations. These dashboards were then used in the review sessions with each participant to allow retrospective recall of the annotations and their associated contexts in the past week(s). Hence, the result of such recall provides the opportunities for both researchers and participants to identify activities which needed clarification, or unexpected gaps in their timeline which needed explanations. At the end of the weekly meeting, each participant receives a gift card of $50 for the contributions made.

End of the Study. In the last meeting (corresponding to the fourth weekly meeting), in addition to the steps described above, researchers pull anonymized summary information from the participant's calendar for the period associated with the study. The participant is then given the usual $50 gift card, plus a bonus gift card of $100 after completing the four weeks of the experiment.

Figure 1 gives an overview of all the apps/platforms and the instruments for collecting sensing signals (i.e., sensor data) and task annotations from the participants.

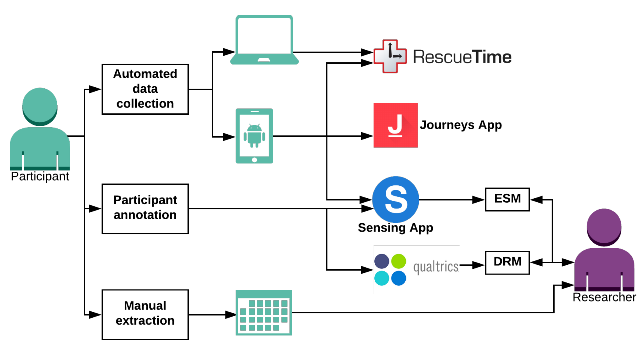

Figure 1: Overview of logging apps and their instruments.

3.2 Collecting Cyber, Physical and Social Signals

To capture context-rich activities and behaviors of our participants in a comprehensive manner, we devised a logging procedure which facilitates the smartphone-based collection of cyber, physical, and social aspects (i.e., signals) associated with different tasks. At the beginning of the logging process, participants were asked to install

[1] The methodology was reviewed and approved by the Human Research Ethics Committee at RMIT University.
[2] http://www.orsee.org/web/
[3] In our pilot experiments, we have identified issues with running the data collection apps in different Android versions, and an app developed in-house is only available for Android. We plan to develop an iOS version in the future.

three apps on their smartphones: an in-house developed app (hereafter referred as the *Sensing* app), *RescueTime*[4], and *Journeys*[5]. In addition, events from participants' calendars were extracted.

Sensing App. The Sensing app logs the following sensor data from participants' Android smartphones:

- Accelerometer sensor data (e.g., measurement of smartphone motion).
- Gyroscope sensor data (e.g., measurement of orientation and velocity of the smartphone).
- Magnetometer sensor data (e.g., measurement of magnetism around the phone).
- GPS (location) data.
- Smartphone barometer sensor data (e.g., measurement of atmospheric pressure).
- Device state (e.g., battery level and screen-lock status).
- Representative information of raw audio data (sampled from the device's microphone, such as noise level). This information is extracted in the audio recording module of the sensing app. It should be noted that only representative audio data (e.g., noise level and power spectrum information) are stored during data collection, with no general recording of ambient audio, to preserve the privacy of participants as well as other people who may be within the audio sensing range of the device.
- Call logs (e.g., information of incoming/outgoing calls/SMS, with anonymized phone numbers).
- WiFi scan data (surrounding WiFi networks of the smartphone, including their signal strengths).
- Bluetooth scan data (e.g., surrounding Bluetooth devices of the smartphone, including their signal strengths).
- Cell tower data (e.g., surrounding cell tower information of the smartphone, including their signal strengths).
- App usage statistics (e.g., information about mobile apps used throughout the day).

The *Sensing* app also facilitates the in situ task annotations performed via ESM (see Section 3.3).

RescueTime App. This app logs the online (cyber) behavior of participants. Specifically, it captures the time participants spend interacting with different applications and websites at an hourly granularity. These interactions are also grouped into various categories.

Journeys App. This app logs sensor data from smartphone users to detect the surrounding contexts in real-time (e.g., transport modes, location clusters, health scores). It also provides user profiling based on historical data.

Calendar Events. Events corresponding to the data collection period are extracted using the OAuth 2.0 protocol for authentication and authorization to read their online calendar (e.g., Google Calendar). The extracted events are obfuscated with only the following information being recorded: start/end time of an event, number of participants, and whether a particular location was indicated. Table 1 shows an example of the information extracted from a calendar after obfuscation.

[4]https://www.rescuetime.com/
[5]https://play.google.com/store/apps/details?id=com.sentiance.journeys&hl=en

3.3 Acquisition of Task Annotations

Task annotations were acquired by utilizing two methods: ESM [5] and DRM [6]. The data collection was conducted on weekdays (i.e., Monday–Friday). Figure 2 shows a general overview of surveys that are triggered through various methods: **mobile app notification**, **self-driven annotation**, and **email notification**. The ESM-based annotation acquisition is achieved through in situ surveys (triggered by notifications through an app, or initiated by the user at a convenient time during the active mode of data collection by the *Sensing* app), while DRM-based annotation acquisition is conducted at the end of the day (through email notification, which redirects the user to an online survey form to retrospectively reconstruct the tasks they engaged in throughout the day).

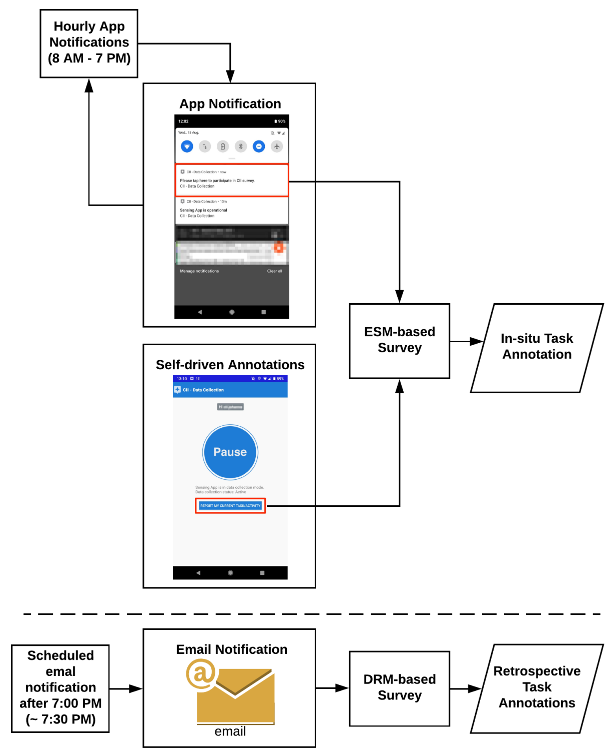

Figure 2: Triggers and acquisitions of in-situ and retrospective task annotations.

ESM-based Annotation. These annotations are based on a short survey; in order to minimize interruption to daily activities and tasks, the questions in relation to the performed tasks should not be too long and must be straightforward. ESM aims to minimize human cognitive bias while reducing the reliance on participants' ability to accurately recall earlier experiences [18]. Hence the ESM process typically does not include questions which ask the mobile user about the actual task start time. In this study, the annotations acquired from the ESM process are defined as in situ annotations.

Figure 3 details the workflow of the *Sensing* app, triggered when a user accepts the notification prompt from the app to perform an in situ annotation (ESM-based survey). Alternatively, this ESM-based survey can also be initiated by the user whenever they are

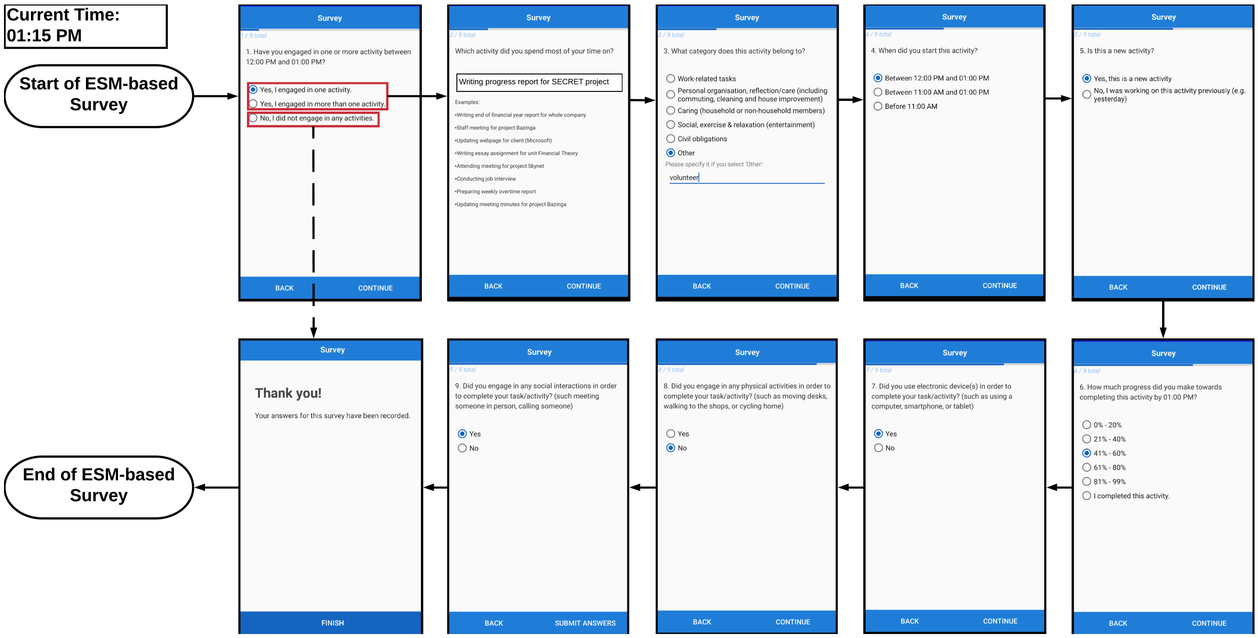

Figure 3: Workflow of ESM-based survey in the *Sensing* app.

engaged in a task (within the duration of active data collection of the Sensing app). Therefore, the flexibility of the in situ annotation process can be guaranteed by providing these two trigger types (app notification and self-driven annotations). In our study, the active data collection from the *Sensing* app covers the logging of mobile CPS signals from 6 AM to 7 PM.

The ESM-based survey is designed to cover the recent tasks that participants were engaged in during the previous one-hour time block, including the contextual aspects of the task, as depicted in the ESM survey workflow in Figure 3. Suppose a participant is asked to fill in the survey at 01:00 PM, but is only able to respond to such a notification fifteen minutes later. At the time the participant initiates the survey at 01:15 PM, the questionnaire is focused on tasks within the time block of 12:00–01:00 PM. The contextual aspects associated with the reported task include the categorization of tasks [17], the occurrence of other activities that may overlap with the reported task, perceived time spent, a binary indication of whether this task is newly initiated (not a continuation of previously reported tasks), estimated progress by the end of the time block, and the CPS presences associated with the engaged task. Ideally, the user should report a recent task that could be time-consuming or complex for its completion. However, this ESM-based survey is presented in an open-ended form, in order to support flexibility, and learn how participants perceive tasks according to their daily routines. The *Sensing* app was configured to trigger notifications for ESM surveys on an hourly basis, from 08:00 AM to 07:00 PM. If a participant reports that no particular task was performed within the previous time block, the ESM-based survey will be concluded immediately.

DRM-based Annotation. At the end of a day, each participant was asked to list all tasks, including estimated start and end times, that were performed during the day. The process of asking the mobile user to recall their tasks within one day allows us to derive the most significant (and possibly time consuming) tasks which they made progress on or completed. The main advantage of the DRM method is to inform the design of future applications of assistive technologies, by focusing on acquiring annotations for important tasks as perceived by users. For example, imagine a scenario where time-consuming tasks can be supported by the seamless recommendation of activities by an intelligent assistant, in order to speed up the completion of these tasks. Inherently, the annotations should include the actual start and end time of the tasks, as perceived

Table 1: Example of events extracted from a participant's calendar (artificial data).

Event	*Start*	*End*	*Has Location?*	*# Participants*
Busy	2018-11-01T12:30:00+11:00	2018-11-01T14:00:00+11:00	TRUE	13
Busy	2018-11-03T15:00:00+11:00	2018-11-03T16:00:00+11:00	FALSE	0
...	...	...	...	...
Busy	2018-11-22T14:00:00+11:00	2018-11-22T14:30:00+11:00	TRUE	2
Busy	2018-11-23T11:00:00+11:00	2018-11-23T12:00:00+11:00	FALSE	0

by the participants. In order to maximize the advantage of this technique, the participants are encouraged to view the DRM as a diary study, by recalling the activities/tasks since they woke up as a string of episodes. The products of this DRM-based survey process are denoted as retrospective task annotations (as depicted in Figure 2). In our study, the trigger of this process is achieved through daily email notifications sent at the end of the day (07:30 PM). Appendix A shows the workflow of the DRM-based survey questions. For the implementation of the DRM-based survey, we used Qualtrics[6] as the online platform to record the retrospective task annotations.

4 LESSONS LEARNED SO FAR

So far, we have completed the data collection for a total of 13 participants, who performed the previously outlined procedures for four consecutive weeks. The data has been collected in two batches, consisting of five and eight participants, respectively. We are currently running a third batch with four participants.

A number of challenges were encountered during the data collection process, related to: the recruitment process; the underlying notion of what a *task* is; and limitations and issues with hardware.

Recruitment. We found that recruiting a large number of participants for a long-term laboratory user study is not straightforward. Although the pool of potential participants was large (more than 1000), it was more challenging than expected to identify participants with availability for four consecutive weeks, and with a compatible Android device. Moreover, the time of the year in which the study is performed (tail end of the year) may have a negative impact (e.g. some potential participants may plan to travel for the holiday season).

Notion of Task. Although the instructions during the intake gave an explicit definition of *task* as "the set of actions/steps/activities needed to reach a particular goal" and provided examples (e.g. "staff meeting for project X", "designing webpage for client Y"), weekly meetings were extremely important to ensure that participants annotated their tasks in a consistent way. In particular, researchers encouraged participants to annotate the tasks as instances ("writing report for project X") rather than more generic classes of tasks ("writing a report"), aiming to increase the quality of the labels.

Hardware/Software Limitations by Heterogeneous Manufacturers. Most of the CPS signals and the ESM surveys were collected via the *Sensing* app, software developed in-house. Developing, testing, and maintaining mobile apps for multiple devices entails substantial effort. The current version has been successfully used on devices such as Samsung S6+, Google Nexus, Google Pixel, and Motorola Moto X. However, Android devices that use the specific Operating System (OS) settings of some brands (e.g., Oppo and OnePlus) introduced new types of problems for the data collection of mobile sensor data, including:

- **Aggressive battery optimization.** Several operating systems (depending on phone model/brand) might be subject to aggressive optimization (by default, regardless of Android permission settings to not optimize the app). Under such conditions, there would be a lack of records for sensor data. The majority of identified missing sensor data (including in-situ task annotations) are related to the following issues:
  - Freezing the app immediately when it goes from foreground to background.
  - Preventing the app from being run the background.
  - Disallowing the auto-start app process (such as when app is killed manually by the user).
  - Automatically switching the apps (*Sensing*, *Journeys*, and *RescueTime* apps) from "not optimizing battery" to "optimized". This issue was reported by a participant using a OnePlus device.
- **Optimized notification settings.** Due to custom optimization by a particular phone brand (such as Oppo devices), the OS may automatically limit the number of app notifications for ESM-based surveys. During our data collection from an Oppo user, there were no hourly notifications that are noticeable from the weeks of data collection (some of them would disappear immediately after pinging the user). Similar behavior was noticeable for a participant with a OnePlus device.

[6] https://www.qualtrics.com/au/

5 SUMMARY

In this paper, we presented a methodology that is currently applied in an ongoing research project to create a benchmarking dataset for task progress for digital assistants. Imbuing digital assistants with the ability to determine task progress enables them to deliver many new types of support, both for the current task phase and future task phases. We also shed some light on the issues and challenges that arose during data collection, together with lessons learned.

Near-term future work includes estimating the robustness and reliability of the collected data to train and evaluate machine learning models for task progression. We also plan to study the scalability of the methodology by training research assistants to perform the procedure in other contexts (e.g., in different geographic locations or languages).

ACKNOWLEDGMENTS

This research was partially supported by Microsoft Research. Any opinions, findings and conclusions expressed in this paper are those of the authors and do not necessarily reflect those of the sponsor.

REFERENCES

[1] Yun-Nung Chen, Dilek Hakkani-Tür, and Xiaodong He. 2015. Detecting Actionable Items in Meetings by Convolutional Deep structured Semantic Models. In *Automatic Speech Recognition and Understanding (ASRU), 2015 IEEE Workshop on.*

[2] Moustafa Elhamshary and Moustafa Youssef. 2014. CheckInside: A Fine-Grained Indoor Location-Based Social Network. In *UbiComp'14.*

[3] Adam Fourney and Susan T Dumais. 2016. Automatic Identification and Contextual Reformulation of Implicit System-Related Queries. In *SIGIR'16.*

[4] Dawud Gordon, Jürgen Czerny, and Michael Beigl. 2014. Activity Recognition for Creatures of Habit. *Personal and ubiquitous computing* 18, 1 (2014).

[5] Joel M Hektner, Jennifer A Schmidt, and Mihaly Csikszentmihalyi. 2007. *Experience Sampling Method: Measuring the Quality of Everyday Life.* Sage.

[6] Daniel Kahneman, Alan B Krueger, David A Schkade, Norbert Schwarz, and Arthur A Stone. 2004. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science* 306, 5702 (2004).

[7] Joo-Kyung Kim, Gokhan Tur, Asli Celikyilmaz, Bin Cao, and Ye-Yi Wang. 2016. Intent Detection Using Semantically Enriched Word Embeddings. In *Spoken Language Technology Workshop (SLT), 2016 IEEE.* IEEE.

[8] John Krumm, Dany Rouhana, and Ming-Wei Chang. 2015. Placer++: Semantic Place Labels Beyond the Visit. In *IEEE PerCom'15.*

[9] Edward A Lee. 2008. Cyber Physical Systems: Design Challenges. In *11th IEEE Symposium on Object Oriented Real-Time Distributed Computing (ISORC)*.

[10] Jonathan Liono, Flora D Salim, Niels van Berkel, Vassilis Kostakos, and A Kai Qin. 2019. Improving Experience Sampling with Multi-View User-Driven Annotation Prediction. In *IEEE PerCom'19*.

[11] Thuong Nguyen, Vu Nguyen, Flora D Salim, and Dinh Phung. 2016. SECC: Simultaneous Extraction of Context and Community from Pervasive Signals. In *IEEE PerCom'16*.

[12] Veljko Pejovic, Neal Lathia, Cecilia Mascolo, and Mirco Musolesi. 2016. Mobile-based Experience Sampling for Behaviour Research. In *Emotions and Personality in Personalized Services*. Springer.

[13] Yongli Ren, Martin Tomko, Flora D Salim, Jeffrey Chan, Charles LA Clarke, and Mark Sanderson. 2018. A Location-Query-Browse Graph for Contextual Recommendation. *IEEE Trans. Know. Data Eng.* 30, 2 (2018).

[14] Milad Shokouhi and Qi Guo. 2015. From Queries to Cards: Re-ranking Proactive Card Recommendations Based on Reactive Search History. In *SIGIR'15*.

[15] Yu Sun, Nicholas Jing Yuan, Yingzi Wang, Xing Xie, Kieran McDonald, and Rui Zhang. 2016. Contextual Intent Tracking for Personal Assistants. In *KDD'16*.

[16] Yik-Cheung Tam, Yangyang Shi, Hunk Chen, and Mei-Yuh Hwang. 2015. RNN-based Labeled Data Generation for Spoken Language Understanding. In *Sixteenth Annual Conference of the International Speech Communication Association*.

[17] Johanne R Trippas, Damiano Spina, Falk Scholer, Ahmed Hassan Awadallah, Peter Bailey, Paul N Bennett, Ryen W White, Jonathan Liono, Yongli Ren, Flora D Salim, and Mark Sanderson. 2019. Learning About Work Tasks to Inform Intelligent Assistant Design. In *CHIIR'19*.

[18] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *Comput. Surveys* 50, 6 (2017).

[19] Jason D Williams, Nobal B Niraula, Pradeep Dasigi, Aparna Lakshmiratan, Carlos Garcia Jurado Suarez, Mouni Reddy, and Geoff Zweig. 2015. Rapidly Scaling Dialog Systems with Interactive Learning. In *Natural Language Dialog Systems and Intelligent Assistants*. Springer.

[20] Liu Yang, Qi Guo, Yang Song, Sha Meng, Milad Shokouhi, Kieran McDonald, and W Bruce Croft. 2016. Modeling User Interests for Zero-Query Ranking. In *ECIR'16*.

A DAILY RECONSTRUCTION METHOD: SURVEY

Considering the following user task of **Writing WSDM workshop paper**, the questions in DRM-based survey would be presented as follows:

Q1. What time did you wake up today? (hh:mm)

Q2. How many hours did you spend for sleeping (in total)?

- More than 8 hours
- 8 hours
- 7 hours
- 6 hours
- 5 hours
- Less than 5 hours

Introduction to rest of survey: Thinking about today, we'd like you to reconstruct what your day was like, as if you were writing in your diary.

Think of your day as a continuous series of scenes or episodes in a film. Each episode is a task that you have performed or in progress towards the completion.

Each task should at least be performed in one-hour duration. In this study, we aim to understand how an intelligent assistant can help in recognizing and managing your daily tasks, to increase the overall productivity and your quality of life. **Next »**

Q3. Have you attempted/progressed on any tasks today?

- Yes **» Proceeds to Q4.**
- No **» Finishes the survey.**

Q4. Please enter the description of one task you attempted/progressed on today.

Q5. To which category does **Writing WSDM workshop paper** belong to?

- Work-related tasks
- Personal organization, reflection or care (includes commuting, cleaning and house improvement)
- Caring (household or non-household members)
- Social, exercise & relaxation (entertainment)
- Civil obligations
- Other: ____________________

Q6. To which of the activity/task-type does **Writing WSDM workshop paper** belong to?

- Communication
- Documentation
- Planning
- Admin and management
- Education
- IT
- Finance
- Physical
- Problem solving
- Low-level
- Project
- Customer care
- Meals and breaks
- Travel
- Other: ______________

Q7. What kind of trigger did you initiate **Writing WSDM workshop paper**?

- Deadline
- Reminder/alarm (e.g. through digital notification)
- Ad-hoc/spontaneously
- Needs for resources
- Other: ____________________

Q8. What is the approximate time when you started **Writing WSDM workshop paper** (hh:mm format)?

Q9. Approximate progress when you started **Writing WSDM workshop paper**:

- 0% – 19%
- 20% – 39%
- 40% – 59%
- 60% – 79%
- 80% – 99%
- 80% – 99%
- 100% (complete)

Q10. What is the approximate time when you stopped **Writing WSDM workshop paper** (hh:mm format)?

Q11. Approximate progress when you stopped **Writing WSDM workshop paper**:

- 0% – 19%
- 20% – 39%
- 40% – 59%
- 60% – 79%
- 80% – 99%
- 80% – 99%
- 100% (complete)

Q12. How satisfied are you with the progress of **Writing WSDM workshop paper**?

- Extremely satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Extremely dissatisfied

Q13. Thinking about the urgency of this task, what was your perceived priority of when you started **Writing WSDM workshop paper**:

- High
- Medium
- Low

Q14. Who were you directly interacting with in the progression of Writing WSDM workshop paper?

- None
- Spouse/significant other
- Household member(s)
- Friend(s)
- Co-worker(s)
- Boss(es)
- Others: ______________

Q15. Describe your activities and contexts involved for the progression of **Writing WSDM workshop paper**.

Q16. **Writing WSDM workshop paper**? Recalling today's tasks, is there any more task you attempted to progress on?

- Yes **» Loops back to Q4.**
- No **» Finishes the survey.**