

# Episodic Memory Integration of Personal Data in YourDigitalSelf

Varvara Kalokyri, Alexander Borgida, Amelie Marian  
Department of Computer Science, Rutgers University  
v.kalokyri@cs.rutgers.edu, borgida@cs.rutgers.edu, amelie@cs.rutgers.edu

## Abstract

Human episodic memory consists of past personal experiences organized into episodes. It is a crucial part of our lives, influencing our identity and interpersonal relationships. A large variety of digital devices and apps capture personal digital traces (PDTs) of our daily experiences, including messages, time and/or location data, browsing and purchase history, etc. Our goal is to reconstruct meaningful episodes from PDTs, in order to improve memory recall and aid individuals with memory problems. In addition to acquiring PDTs, this process faces many problems, including: the diversity and very large number of PDTs; describing the episode types of interest; finding and combining the relevant but very sparse set of PDTs that provide partial evidence for a particular episode having occurred.

To address these challenges, the YourDigitalSelf project employs semantic modeling and integration techniques to associate PDTs with events, themselves organized into prototypical narratives/plans called scripts. It introduces a formalized conceptual modeling language, hierarchical script definitions, and an evidence-based bottom-up approach to reconstruct script instances. A first evaluation of our implemented methodology using a mobile application demonstrates relatively successful integration of diverse digital traces and memory enhancement.

## 1 Introduction

In today's world, digital devices are seamlessly integrated into our daily lives, transforming the way we engage with people and our surroundings. These devices generate and store various kinds of Personal Digital Traces (PDTs), such as e-mails and messaging, location data, calendar entries, check-ins, reviews, web searches, and purchase history. These PDTs, serve as a partial record of individuals' lives, but pose a challenge to users because the full collection is very large, highly heterogeneous, and distributed in space and time, making it difficult to access, search, and learn from.

Drawing inspiration from Vannevar Bush's visionary "memex" concept, which can be seen as a personal data collection system, our work aims to exploit the potential of organizing personal data in order to aid remembering past events. By connecting the dots between personal information items, users can effectively retrieve specific memories or information in a coherent and meaningful way.

The goal is to discover evidence for users' actions from the digital traces these actions leave behind, and organize them into coherent episodes that look like high-level complex and meaningful workflows, enhancing the user's "autobiographical memory". This could be especially beneficial for remembering routine experiences, which has been shown to be harder to recall [1]. For example, users could use connections in their data to not only quickly retrieve digital artifacts, such as pictures from a trip, minutes from a meeting, or pictures from a birthday party, but also to retrieve specific information about events, such as the name of a restaurant they went to, the title of a movie they watched, etc.

Our research takes a step forward in supporting human memory and recall through the concept of episodic narratives, rooted in psychology and cognitive science. It enables users to organize PDTs into cohesive episodes, creating a personal knowledge base that users can query to remember specific events and their by-products.

Additionally, it provides users with narratives to stimulate their memory, which can be particularly helpful for those with memory difficulties.

To achieve this, we need to acquire and integrate all the heterogeneous data from various sources, including files used by apps on our devices (e.g., search history), and through service APIs. We focus in this paper on the following aspects:

1. A conceptual model to represent entities, especially PDTs, which helps extract semantically similar aspects (e.g., who? when? where?), thus integrating the heterogeneous sources.
2. A conceptual model for atomic activities/events. Note that PDTs may provide weighted evidence signaling the occurrence of an activity instance.
3. A conceptual model of complex activities, called scripts. These contain stereotypical commonsense plans, whose instances can again be signaled by the presence of either PDTs or activities/sub-subscripts. A novel notion of inheritance and subsumption helps one to create and organize libraries of scripts.
4. An algorithm for recognizing instances of atomic and composite events/scripts. The bottom-up nature of the algorithm, which merges evidence for components, allows us to deal with missing parts or atypical instances of scripts; a scoring scheme accounts for the varied strength of evidence provided by PDTs or script steps.
5. An Android app which instantiates our ideas, asking users to provide access to a specific set of data sources, loading the data into the PDT concepts, and running the above algorithm. This application is also used to evaluate our approach, through user studies and surveys.

We immediately acknowledge the sensitive nature of the personal data we are dealing with, and the very important privacy issues that they raise. All information obtained to evaluate our work resides on individuals' own mobile phones without disclosing any personal information.

Our study results demonstrate the potential of our approach, showing its effectiveness in integrating digital traces from different sources into coherent episodes and enhancing users' memory of past actions. Our ultimate goal is to empower users to create their personal knowledge base and better manage their digital memories.

## 2 An episodic script model for personal digital traces

Our approach to data integration and memory recall draws on principles from Cognitive Science and Psychology, specifically in the study of how individuals reminisce about their past experiences. These psychological findings have yielded several key observations.

**The Natural Mnemonic Framework:** It is observed that when people reflect on their past, they often do so by responding to a set of fundamental questions, which we refer to as the 'w5h' questions: what, when, where, who, why, and how [2, 3]. For instance, when attempting to recollect a significant event like one's sixth birthday, individuals instinctively seek to reply to these contextual cues. We use this inherent cognitive framework as the foundation of our data integration, by aligning PDTs and conceptual model instances along these essential questions.

**Episodic and Semantic Memory:** Psychological research [4, 5] distinguishes between two primary forms of memory. Semantic memory pertains to the accumulation of general world or cultural knowledge that individuals collect over their lifetimes (e.g., it encompasses the understanding that payment is required when making a purchase) and episodic memory involves the capacity to re-live specific past events or episodes, (such as the memory of celebrating one's 40th birthday by dining out). Our approach combines the representation of both semantic memory (scripts) and episodic memory (script instances). This fusion allows us to assist users in recalling information related to past events. To achieve this, it is imperative to connect individual atomic events and organize them into broader activities, even when these events occurred at different times. This linkage facilitates the retrieval of high-level activity-based memories, enhancing the user's ability to reconstruct their personal history effectively.

## 2.1 Conceptual modeling of personal digital traces

As mentioned, our objective is to develop a framework for modeling digital traces that unifies information based on the ‘w5h’ questions. To achieve this, we start with a standard-centered conceptual modeling language, as first presented in [6]. This language employs entity classes, such as ‘Documents’ and ‘Persons,’ which contain individual instances. These instances possess atomic valued features, such as the size of a document, and are interconnected through properties, such as those associated with an Email, including ‘from,’ ‘to,’ ‘date,’ ‘subject,’ and ‘body.’

Within this modeling framework, classes can be specialized into subclasses. For instance, ‘PERSON\_WITH\_EMAIL’ is a specialization of the ‘PERSON’ class, and new properties, such as ‘email\_address’, can be introduced, or existing properties may be further constrained. Notably, properties can also undergo specialization, where, for example, ‘from’ and ‘to’ are sub-properties of ‘who’. A large subset of the properties of PDTs can be viewed as specializations of w5h, when these are viewed as properties themselves, resulting in the integration of heterogeneous object schemas.

Note however that there is no sensible way to ask questions like “when?” or “who?” of many objects. However, by associating an object with a primitive action, it becomes possible to derive the ‘when’ aspect from the action. For instance, we may not directly ask ‘who email’, but we can ascertain ‘who sent an email’. This requires attaching documents to primitive actions.

## 2.2 Conceptual modeling of atomic and complex events

A script-based model is adopted for the purpose organizing atomic actions into coherent episodes. This describes event flows, encompassing steps that humans intuitively understand. For example, for going out to eat, these include steps like organizing the outing, making reservations, taking transportation, and making payment. These steps often generate PDTs, serving as evidence for the activities. For example, we then want to group emails concerning a particular dinner, a reservation that was made at a restaurant, a Facebook check-in with photos, and a credit card payment. The atomic actions that generated these documents are organized as part of the narrative for going out to eat. Essentially, each data trace presents evidence of various strengths for the occurrence of atomic event instances.

To achieve this, we need a set of higher-level plans that the users and their community frequently engage in, which describe the connections between all those events. The idea of scripts is inspired by the work of Schank and Abelson in AI [7], which are defined as plans describing stereotyped sequences of actions that describe some well-known situation.

For describing such scripts we use the standard notation used in business process management (BPM) [8]. As such, scripts are characterized by their goals, participant information, w5h aspects, component sub-scripts, and allowed sequencing and timing of sub-scripts/atomic actions - the latter captured in BPM.

However, in order to make a fully functional system, we need script definitions for many variants of everyday activities. These plan descriptions serve as prototypes due to the vast variability inherent in such situations, making it impractical to predetermine every possible variation. This will be facilitated by placing the script definitions in an ontology and then using specialization and inheritance, so that more specific processes, such as GoingOutToEat, GoingOutToConcert, GoingToOpera, etc. are obtained by indicating only differences from a common superclass, GoingOutForEntertainment. In turn, GoingOutToEat can be specialized to EatingOutAtRestaurant, EatingFastFood, etc. By using specialization hierarchies we can organize large knowledge bases, support reuse and change propagation, as well as prevent repetition errors.

Description logics (DLs) are ideal for the creation of specialization of such hierarchies via reasoning. We have introduced a new extended regular expression-based Description Logic [9] able to represent the control flow semantics of structured BPM. For example, specific concepts A and B can be specified, and then we can assert that a class A is subsumed by B, or discover this relationship by reasoning, meaning that all instances of A are

instances of B. We can also detect inconsistencies, (e.g. we might give a description of going out to a rock concert which contradicts going out to a concert. This will provide evidence that this is not a proper specialization).

Instance recognition for DLs by the use of inference could be used for our purposes for recognizing activities. However, this approach cannot be used because: not all steps occur in every instance, their order may vary, and some steps may leave no digital trace. Instead, a bottom-up approach is proposed where partially instantiated scripts are constructed from individual PDTs, and compatible instances are merged, accumulating information and strengthening evidence for them. This approach is flexible and accommodates various types of scripts, making it suitable for different scenarios. In the next section, the algorithm for creating episodes from PDTs is described.

### 3 Algorithm for Episode Recognition

Our algorithm, starts with a specific script (S) and a comprehensive set of PDTs (P). The primary objective is to generate potential script episodes, which are instances of (S), and establish connections with the PDTs.

**Step 1. Script Specification:** A script specification has the following components: a top-level (outer) script (e.g. 'EatingOut' script), which we want to instantiate; several component sub-scripts and atomic actions; as well as sequencing relationships among them. In addition, every script contains its w5h properties.

For example, 'EatingOut' comprises events like inviting people, making reservations, going to a restaurant, paying the bill, and sharing photos. Some of these events (e.g., making a payment) offer strong evidence that the person indeed participated in this activity/script. On the other hand, receiving an email mentioning "dinner" or "lunch" provides weaker evidence for planning to dine out, which, in turn, offers weaker evidence of actually going out since the plan may remain incomplete or canceled. Our algorithm employs the concept of strong and weak evidence to prioritize potential script instances.

We gather information about the strength of evidence in a similar manner to acquiring common-sense knowledge about events within a given context. For that purpose, we have developed a self-evaluating multiplayer web-based game for collecting digital trace descriptions for every step in a scripted scenario and potential strength of evidence [10]. It is essential to note that there are limited "objective" methods for scoring evidence because scripts represent imprecise, culturally-shared, commonsense knowledge describing typical human activities.

**Step 2. Retrieving Document Set for Script Instantiation:** Once a script is parsed, the subsequent task involves identifying the set of documents, denoted as (D), that serve as evidence for the occurrence of its instances. These documents correspond to PDTs and act as "noisy sensors," indicating potential instances of corresponding scripts or atomic actions for which strong evidence exists.

To locate documents that correspond to a specific piece of evidence, it is crucial to identify specific cues within the documents. These cues can take the form of verbs to search for (e.g., "eat" in an email to identify an "Initiate Going Out" event) or attributes and metadata that a document may possess (e.g., categorizing a payment as "Restaurant" or "Supermarket" to find evidence for "Making a Payment at a restaurant or supermarket respectively). To ensure the replicability of this process for various scripts, the w5h components of the script or atomic action are obtained from its FrameNet frames[11]. Standard sources of synonyms and hyponyms such as WordNet and ConceptNet5 [12, 13] are then utilized to identify additional words for searching. The generated lists of words and phrases from this process are stored in a text file and used to retrieve all potentially relevant documents.

Finally, the set of documents (D) is pre-processed by: (i) expanding information (e.g., terms like "tomorrow", "on Friday", are made absolute dates); (ii) performing entity resolution for people and places (who and where dimension) using Stanford's Entity Resolution Framework [14]; (iii) grouping certain kinds of documents (e.g., related email threads, or related sequences of tweets) into a single individuals d in (D); (iv) finding the places/venues that the user has visited from the geo-location coordinate history (gps) [15].

**Step 3. Creating Initial Script Instances:** Every document in the set (D) serves to instantiate atomic actions or sub-scripts, creating candidate instances of the outer script (S) in a bottom-up fashion. Similarly, the w5h

properties are propagated from the document into the atomic actions and the script hierarchy.

**Step 4. Merging Script Instances:** Merging script instances is a distinctive feature of our system, using multiple sources of evidence for the same script instance. Each script is assigned “keys” that rate the importance of w5h (sub)properties in identifying instances. For ‘EatingOut’, the keys are ‘whereEatingOccurred’, ‘whenEatingOccurred’, and, to a lesser extent, ‘who’. The ‘what’ and ‘how’ properties of this script are not important because they would often lead to incorrect merging (e.g., two instances of eating pasta (‘what’) need not be merged). Since every script can have different keys, this has to be explicitly mentioned in the algorithm. When two instances share similar keys, they are considered candidates for merging. The w5h property fillers are combined, and a score is computed for the merged instance based on Hooper’s rule [16] for combining probabilistic evidence.

The key point here is that the merging process cannot be completed in a single step. After the script instance is established with a certain degree of certainty, additional PDTs may be collected as part of the instance when examining the script definition.

## 4 A mobile app for YourDigitalSelf

So far, we have given in Section 2 an ontology of entities (especially for PDTs), and of atomic and composite events/scripts, which involve these entities. In Section 3, we discussed how specific PDT concepts provide evidence of varying strength for specific scripts, and an algorithm which, given a script (episode type), uses this information to instantiate very small episode fragments and then merge them into larger ones, with associated strengths.

To build our YourDigitalSelf Android app (see [17] for details, including screen shots), we first need to have users provide access to the data sources from which the PDT instances will be created, and then interactively choose a specific script for which they wish to see rank-ordered instances retrieved. The types of PDTs along with the data sources currently supported by the app are the following:

- Messaging: Messenger, phone text messages
- Social Media: Facebook, Instagram
- Email: Gmail
- Calendar: Google Calendar
- Financial Data: Plaid API and directly downloaded .csv files from bank institutions
- Location Data: Google Maps location history, GPS data
- Photos: Google Photos

The system then gathers the raw data and uses it to populate the corresponding PDT concepts. In general, each API has a standard format, but in the not-infrequent case where APIs change, this mapping must be modified. Recall from Section 2.1 that properties for the ‘w5h’ questions and their sub-properties are the key to semantically “integrating” the extremely heterogeneous forms of PDTs.

When asked, the app then essentially runs the algorithm in Section 3 to gather episodes (script instances) of the script asked for, and then displays them via a GUI. Note that each episode has attached a variety of PDTs that provided evidence for it; this is a second way in which related data is “integrated”.

Instances of the concepts in conceptual model cannot be efficiently stored directly on the mobile phone, and must instead be mapped to a relational database natively managed by Android. The mapping of the different concept types to tables is declaratively described using pairs of conjunctive queries.

After the mapping to our model, the next step involves designing an indexing structure to facilitate data queries. We leverage a combination of full-text and column indexing techniques provided by the Android platform. The indexing occurs whenever the user decides to include data in the app, either manually or through automated periodic updates.

In general, the project has made a significant effort to support extensibility and modification by declarative

representation of other aspects as well, such as evidence to look for in scripts and clues to search in PDTs, as well as making scripts parametric/generic (e.g, Book<T>, T="accommodation", "concert",...).

## 4.1 Experimental Setup

To assess the effectiveness of our approach, we conducted experiments using real user data with the primary objective of detecting instances where individuals engaged in dining out at restaurants. We opted for this specific script example because it offers a rich variety of PDTs and shares commonalities with other scripts related to entertainment, such as attending theaters or concerts.

We employed the YourDigitalSelf Android application for this purpose, which allows users to install it locally on their devices, and it is designed to serve future research purposes as well. It is paramount to clarify that our approach did not entail direct access to users' personal data. Instead, users provided responses solely in the form of Yes/No answers, without revealing any personal information. The application is configured to collect PDTs from a diverse range of services, including messaging, social media, email, calendar, financial data, location data, and photos.

We ran an in-depth study<sup>1</sup> with 16 users. Before the experiment, we asked participants to try to remember the occasions of them going out to eat at restaurants and then carefully go over their past month's digital information, writing down their outings, including name of the restaurant - where, date they went - when, with whom they went - who. We used this information as a proxy for recall. Then during the experiment participants were shown all candidate script instances of them going out to eat at restaurants, and had to indicate Yes/No for each of the instances first, and then for each of the w5h properties.

For a comprehensive description of the entire study and its results, we refer interested readers to the detailed study in [18].

## 4.2 Experimental Evaluation

The evaluation of our approach focuses on the assessment of our matching algorithm and scoring function which encompasses two key aspects: (1) the effectiveness of our approach in identifying script instances and (2) its proficiency in recognizing and abstracting 'When,' 'Where,' and 'Who' information from sub-scripts and atomic actions into the overarching script instance.

For both facets of evaluation, we employ two distinct dimensions of relevance:

### 1. Binary Relevance

A script instance is deemed relevant if the user actually engaged in dining out at a restaurant, even if the 'w5h' information was only partially correct whereas each 'w5h' information is considered relevant only if it is precisely correct (neither a subset nor a superset).

### 2. Graded Relevance

A proposed script instance falls into one of the following categories:

Exactly relevant: When the user indeed dined out at the specified restaurant.

Relevant but too broad: When the identification of a restaurant outing is accurate, but the user's purpose for the visit did not involve dining (e.g., for dancing or socializing with a friend).

Relevant but too narrow: When a restaurant outing is correctly identified, but the user did not stay at the establishment to eat (e.g., it was a takeout order).

Partially relevant: When a planned outing was correctly identified, but the user did not follow through with it.

Not relevant: When the identified instance does not pertain to a restaurant outing.

---

<sup>1</sup>Our study has received approval from the Rutgers University Institutional Review Board (IRB) committee.



	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16
Our approach	13	15	10	5	<b>14</b>	<b>24</b>	13	6	19	17	19	7	11	5	17	15
User memory	7	8	8	5	8	10	5	5	9	9	13	5	6	3	8	9
User data	13	14	11	5	<b>7</b>	<b>14</b>	11	6	20	14	15	5	11	5	16	14
<b>Recall</b>	1	1	<b>0.91</b>	1	1	1	1	1	<b>0.95</b>	1	1	1	1	1	1	1

Table 1: Number of identified EatingOut actions by users vs number of correct events our approach retrieved per user as a proxy for Recall

	Sources	Precision
User #1	Social Media, Calendar, Financial Data, Location Data, Google Photos	0.87
User #2	Social Media, Location Data, Financial Data	<b>0.94</b>
User #3	Email/Messaging, Social Media, Calendar, Financial Data	0.66
User #4	Email/Messaging, Social Media, Calendar, Financial Data, Location Data, Google Photos	0.66
User #5	Email/Messaging, Social Media, Calendar, Financial Data, Location Data	0.74
User #6	Social Media, Calendar, Financial Data, Location Data, Google Photos	0.89
User #7	Email/Messaging, Social Media, Calendar, Financial Data, Location Data, Google Photos	0.76
User #8	Email/Messaging, Social Media, Calendar	<b>0.6</b>
User #9	Email/Messaging, Social Media, Calendar, Financial Data, Google Photos	0.74
User #10	Email/Messaging, Social Media, Calendar, Financial Data, Location Data	0.81
User #11	Email/Messaging, Social Media, Calendar, Financial Data, Location Data, Google Photos	0.76
User #12	Social Media, Calendar, Financial Data, Location Data, Google Photos	0.86
User #13	Email/Messaging, Social Media, Calendar, Financial Data, Google Photos	0.73
User #14	Email/Messaging, Social Media, Calendar, Financial Data, Location Data, Google Photos	0.83
User #15	Email/Messaging, Social Media, Calendar, Financial Data, Google Photos	0.77
User #16	Email/Messaging, Social Media, Calendar, Financial Data, Location Data, Google Photos	0.79
<b>Total</b>		<b>0.78</b>

Table 2: Overall precision for each user

We establish an analogous framework for assessing the relevance of ‘When,’ ‘Where,’ and ‘Who’ (‘w5h’) information, classifying them into relevant categories based on the same criteria.

To measure the performance of our approach, we rely on the following metrics, considering both binary and graded relevance:

- **Percentage of instances retrieved:** This metric indicates the percentage of all ‘EatingOut’ events identified by users that our scripts successfully retrieved. It serves as a proxy for recall.
- **Mean Average Precision @ k (MAP@k):** Utilizing MAP as a binary relevance assessment, we determine the percentage of the top-k identified script instances that correspond to actual ‘EatingOut’ events.
- **Normalized discounted cumulative gain (nDCG):** This metric is employed to assess the ranked results while considering graded relevance, as described earlier.

### 4.3 Experimental Results

Table 1 shows the number of correct EatingOut instances retrieved by our approach compared with the number identified by users from memory, and by searching their PDTs. A first observation is that the results clearly indicate how hard it is for users to recall their outings, either from memory or even when asked to go through their digital information. Our approach identified more correct instances than the users were able to recall and find in all cases but two (users 3 and 9 in bold). In addition, users found it hard to search through their data, since most of the applications have keyword-based search. Users 5 and 6 in bold, who were able to retrieve only half of their outings, clearly show this issue.

Table 2 shows the overall precision of the identified script instances along with the sources each user incorporated in the application. Our approach achieves a total precision of 78% for all the users. User#2 achieved

	when	where	who
MAP	0.85	0.81	0.21

Table 3: MAP for when, where, who dimensions for all users

the highest precision of all, since the sources they chose to include in the study contained bank transactions, google maps location history, Instagram and Facebook, sources that tend to be of high quality, whereas User#8 achieves the lowest precision of all since they included their private phone text messages without any high-quality source.

However, retrieval systems typically return results in a ranked order, and users expect the first few results to be the most relevant. We now look at the quality of the returned answers by evaluating the mean average Precision@k metric.

Figure 2 shows the Mean Average Precision@k for all the identified script instances for all the users. As shown, our approach achieves a really good precision even for low values of k. The reason for that is that our approach does include many different kinds of sources and is able to account for all the different kinds of user behavior.

Figure 3 shows the normalized discounted cumulative gain (nDCG) for the ranked results when taking into consideration the graded relevance. The nDCG was computed by normalizing the DCG@k with the ideal DCG value or IDCG@k. It is clear that our ranking quality is high, and our approach is able to recognize and distinguish highly relevant PDTs in favor of irrelevant PDTs.

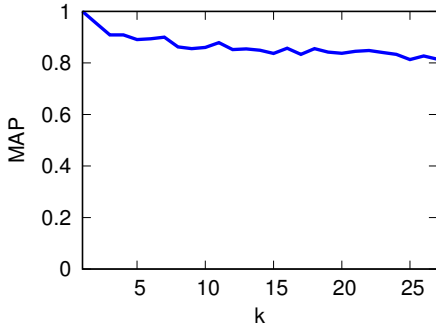


Figure 2: MAP@k for the recognized instances

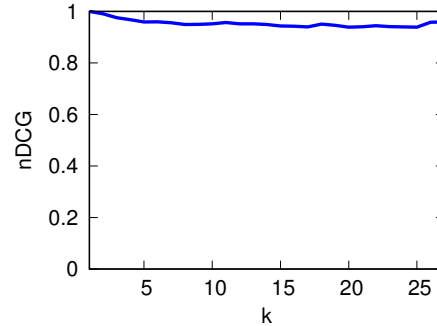


Figure 3: nDCG@k for the recognized instances

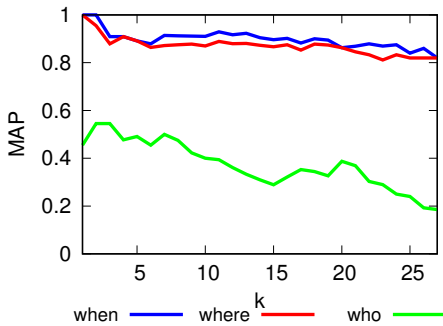


Figure 4: MAP@k for when, where, who dimensions

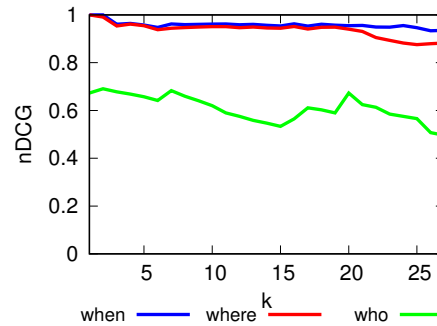


Figure 5: nDCG@k for when, where, who dimensions

We then report the same metrics (MAP, MAP@k, and nDCG@k) on the when, where, and who dimensions.



Table 3 shows the MAP for the three dimensions for all users. We observe that the when and where dimensions are easier to extract than the who dimension due to the metadata that the PDTs have. On the other hand, the who dimension is more difficult to extract. For example, we observe in figure 4 how the precision for the who dimension drops significantly as  $k$  increases. This happens due to the fact that for low values of  $k$ , the score of the instances is low, which means there are not many digital traces to account for these instances, therefore either there is no information, or our approach either lacks or recognizes more people in an outing. This is actually demonstrated in figure 5, which shows the normalized discounted cumulative gain for the ranked results where we can now observe how much better the accuracy is for the ‘who’ dimension.

## 5 Literature Review

In this section, we provide a brief overview of various related research areas, highlighting the interdisciplinary nature of our work.

**Personal Information Management:** Personal Information Management (PIM) research, dating back to the 1980s, focuses on assisting users in managing and organizing digital data. Researchers have suggested PIM interfaces for web activities [19–21], email [22, 23], and local files [24, 25]. Existing systems often utilize domain-specific ontologies to identify relevant objects and their relationships [26, 27]. In contrast, our approach is dynamic, emphasizing the integration of PDTs through narrative connections, shifting the focus from static information to dynamic data integration based on the ‘who, what, when, where, why, and how’ (w5h) properties.

**Process Representations:** Our interest lies in representing autobiographic events and their instances. This area intersects with various formal process representation languages, graphical notations, and complex event recognition [28],[8]. While traditional formalisms concentrate on enacting processes, we focus on descriptive formalisms that enable script instance recognition, underlining the challenges of recognizing multiple, concurrent, and personalized script instances. This leads us to several relatively closely related areas such as Activities of Daily Living, Ambient Intelligence, Pervasive Computing, and LifeLogging. The following are some papers that survey an entire field (e.g. LifeLogging [29]) or the use of ontologies and inference [30–32].

**Activities of Daily Living, Pervasive Computing, LifeLogging, and Memory Tools:** Memory aids play a crucial role in rehabilitation, particularly in improving prospective memory. Existing tools like Sensecam [33], and Kalnikaite’s browser [34] are akin to our approach, as they aim to trigger autobiographical memory by recording images and linking them to daily activities. Other tools include the MemoClip, the Cyberminder, and Memory Glasses [35–37]. However, we stand out by using pre-existing digital traces rather than capturing new data. In addition, the area of life-logging is quite similar, and is surveyed in [38, 39]. Bell has pioneered this field with MyLifeBits [40] for which he digitally captured all aspects of his life. A recent paper [41] focuses on the creation of lifelogs, positioning them as a critical resource for personal assistants to provide tailored advice within specific contexts. Another relevant paper is that of Meditskos et al [42] which used the technique of multi-sensory data analysis along with egocentric video recording from a bracelet to aid the dementia patients recognize their daily living activities.

**Planning:** Plan recognition, a prominent field in AI, focuses on recognizing plans from actions [43, 44]. Some similarities exist with our approach, particularly in recognizing multiple, concurrent, and interleaved script instances. We emphasize data integration, while traditional planning approaches are concerned with action sequences and transitions.

**Personal Digital Assistants:** Personal Digital Assistant (PDA) systems, exemplified by Amazon Alexa, Siri, and Google Now, share commonalities with our work. These systems are engineered to provide users with event reminders based on their personal data, often driven by commercial objectives. However, their functionality is confined to utilizing data within the vendors’ proprietary ecosystems. It’s noteworthy that these PDA systems primarily emphasize future tasks, offering reminders based on time or event triggers [45]. In contrast, our current application scenarios center around retrospective tasks, particularly the organization of past memories.

Nonetheless, our data integration approach can complement prospective methods for tailoring activity recognition.

Our approach emphasizes the dynamic integration of Personal Digital Traces (PDTs) through narrative connections and the recognition of multiple, concurrent, and personalized script instances, setting it apart from conventional methods in these related areas.

## 6 Conclusion

In summary, our research takes steps forward in advancing autobiographical memory by linking diverse PDTs based on their shared objectives, organizing them into cohesive episodic narratives. We introduced a conceptual model encompassing entities, PDTs, and the actions that create them. Scripts, which describe complex activities and are amenable to reasoning via description logic, played a pivotal role. Our merging algorithm identifies script instances, providing a valuable resource for users. Through the YourDigitalSelf app, we evaluated and underscored the potential and promise of our approach.

However, several directions for future research and development are worth considering. These include conducting a more extensive array of experiments encompassing a variety of scripts and involving a larger, more diverse population. Additionally, we recognize the necessity of incorporating Natural Language Processing (NLP) analysis to address instances where users mention activities indirectly. This could be potentially tackled with the use of today’s natural language capabilities (such as openAI’s chatGPT). Beyond its implications for personal data management and memory enhancement, our work offers opportunities for behavioral researchers. The combination of personal digital traces, AI, and self-reported surveys holds the potential to yield fresh insights into mental health assessment. Our tools are open source and publicly accessible on GitHub. This availability paves the way for educators and students to engage with their personal databases, fostering the development of novel, useful applications. Finally, our tools hold potential significance in the medical field, particularly for medical experts conducting studies related to memory rehabilitation in patients facing cognitive difficulties.

## References

- [1] N. M. Bradburn, L. J. Rips, and S. K. Shevell, “Answering autobiographical questions: The impact of memory and inference on surveys,” *Science*, vol. 236, no. 4798, pp. 157–161, 1987.
- [2] D. L. Schacter, *The seven sins of memory: How the mind forgets and remembers*. HMH, 2002.
- [3] W. Jones, “Personal information management,” *Annual review of information science and technology*, vol. 41, pp. 453–504, 2007.
- [4] E. Tulving, “10. episodic and semantic memory,” *Organization of memory/Eds E. Tulving, W. Donaldson, NY: Academic Press*, pp. 381–403, 1972.
- [5] M. A. Conway and D. C. Rubin, “The structure of autobiographical memory,” *Theories of memory*, vol. 103, p. 137, 1993.
- [6] V. Kalokyri, A. Borgida, A. Marian, and D. Vianna, “Semantic modeling and inference with episodic organization for managing personal digital traces - (short paper),” in *Proc. ODBASE 2017*, pp. 273–280, 2017.
- [7] R. Abelson and R. C. Schank, “Scripts, plans, goals and understanding,” *An inquiry into human knowledge structures New Jersey*, vol. 10, 1977.
- [8] S. Goedertier, J. Vanthienen, and F. Caron, “Declarative business process modelling: principles and modelling languages,” *Enterprise Information Systems*, vol. 9, no. 2, pp. 161–185, 2015.

- [9] A. Borgida, V. Kalokyri, and A. Marian, “Description logics and specialization for structured bpmn,” in Int. Conf. on Business Process Management, pp. 19–31, Springer, 2019.
- [10] V. Kalokyri, A. Borgida, and A. Marian, “One of us: a multiplayer web-based game for digital evidence acquisition of scripts through crowdsourcing,” in Proceedings of the 2023 Conference on Human Information Interaction and Retrieval, pp. 187–196, 2023.
- [11] C. J. Fillmore, C. R. Johnson, and M. R. Petruck, “Background to framenet,” International journal of lexicography, vol. 16, no. 3, pp. 235–250, 2003.
- [12] G. A. Miller, “Wordnet: a lexical database for english,” Communications of the ACM, vol. 38, no. 11, pp. 39–41, 1995.
- [13] H. Liu and P. Singh, “Conceptnet: a practical commonsense reasoning tool-kit,” BT technology journal, vol. 22, no. 4, 2004.
- [14] O. Benjelloun, H. Garcia-Molina, D. Menestrina, Q. Su, S. E. Whang, and J. Widom, “Swoosh: a generic approach to entity resolution,” The VLDB Journal, vol. 18, no. 1, pp. 255–276, 2009.
- [15] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma, “Mining user similarity based on location history,” in SIGSPATIAL Advances in geographic inf. systems, p. 34, 2008.
- [16] G. Shafer, “The combination of evidence,” International Journal of Intelligent Systems, vol. 1, no. 3, pp. 155–179, 1986.
- [17] V. Kalokyri, A. Borgida, and A. Marian, “Yourdigitalself: A personal digital trace integration tool,” in Proceedings of the 27th CIKM conference, pp. 1963–1966, ACM, 2018.
- [18] V. Kalokyri, A. Borgida, and A. Marian, “Supporting human memory by reconstructing personal episodic narratives from digital traces,” in Proceedings of the International AAAI Conference on Web and Social Media, vol. 16, pp. 453–464, 2022.
- [19] S. Dumais, E. Cutrell, J. J. Cadiz, G. Jancke, R. Sarin, and D. C. Robbins, “Stuff i’ve seen: A system for personal information retrieval and re-use,” in (SIGIR’03), 2003.
- [20] V. Kaptelinin, “Umea: translating interaction histories into project contexts,” in SIGCHI conference on Human factors in computing systems, pp. 353–360, 2003.
- [21] H. Murakami and K. Mitsuhashi, “A system for creating user’s knowledge space from various information usages to support human recollection,” International Journal of Advancements in Computing Technology, vol. 4, no. 22, 2012.
- [22] T. Ayodele, G. Akmayeva, and C. A. Shoniregun, “Machine learning approach towards email management,” in World Congress on Internet Security (WorldCIS), IEEE, 2012.
- [23] S. Whittaker, V. Bellotti, and J. Gwizdka, “Email in personal information management,” ACM, vol. 49, no. 1, pp. 68–73, 2006.
- [24] D. Barreau and B. A. Nardi, “Finding and reminding: file organization from the desktop,” ACM SigChi, vol. 27, pp. 39–43, 1995.
- [25] D. K. Barreau, “Context as a factor in personal information management systems,” Journal of the American Society for Information Science, vol. 46, no. 5, pp. 327–339, 1995.

- [26] D. R. Karger, K. Bakshi, D. Huynh, D. Quan, and V. Sinha, “Haystack: A general-purpose information management tool for end users based on semistructured data,” in CIDR’05, pp. 13–26, 2005.
- [27] V. Katifori, A. Poggi, M. Scannapieco, T. Catarci, and Y. E. Ioannidis, “Ontopim: how to rely on a personal ontology for personal information management,” in ISWC Wkshop. on The Semantic Desktop, pp. 258–262, 2005.
- [28] W. M. P. van der Aalst and A. H. M. ter Hofstede, “YAWL: yet another workflow language,” Inf. Syst., vol. 30, no. 4, pp. 245–275, 2005.
- [29] C. Gurrin, A. F. Smeaton, and A. R. Doherty, “Lifelogging: Personal big data,” Foundations and Trends in Information Retrieval, vol. 8, no. 1, pp. 1–125, 2014.
- [30] A. Bikakis, T. Patkos, G. Antoniou, and D. Plexousakis, “A survey of semantics-based approaches for context reasoning in ambient intelligence,” in European Conference on Ambient Intelligence, pp. 14–23, Springer, 2007.
- [31] N. D. Rodríguez, M. P. Cuéllar, J. Lilius, and M. D. Calvo-Flores, “A survey on ontologies for human behavior recognition,” ACM Computing Surveys, vol. 46, no. 4, p. 43, 2014.
- [32] E. Alevizos, A. Skarlatidis, A. Artikis, and G. Paliouras, “Probabilistic complex event recognition: A survey,” arXiv preprint arXiv:1702.06379, 2017.
- [33] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, “Sensecam: A retrospective memory aid,” in International conference on ubiquitous computing, pp. 177–193, Springer, 2006.
- [34] V. Kalnikaite, A. Sellen, S. Whittaker, and D. Kirk, “Now let me see where i was: understanding how lifelogs mediate memory,” in Human Factors in Computing, pp. 2045–2054, 2010.
- [35] M. Beigl, “Memoclip: A location-based remembrance appliance,” Personal Technologies, vol. 4, no. 4, pp. 230–233, 2000.
- [36] R. W. DeVaul, A. Pentland, and V. R. Corey, “The memory glasses: subliminal vs. overt memory support with imperfect information,” in Symp. Wearable Computers, 2003, 2003.
- [37] A. K. Dey and G. D. Abowd, “Cybrereminder: A context-aware system for supporting reminders,” in Int. Symposium on Handheld and Ubiquitous Computing, pp. 172–186, 2000.
- [38] C. Gurrin, A. F. Smeaton, A. R. Doherty, et al., “Lifelogging: Personal big data,” Foundations and Trends® in information retrieval, vol. 8, no. 1, pp. 1–125, 2014.
- [39] E. Van Den Hoven, C. Sas, and S. Whittaker, “Introduction to this special issue on designing for personal memories: past, present, and future,” HCI, vol. 27, no. 1-2, pp. 1–12, 2012.
- [40] J. Gemmell, G. Bell, and R. Lueder, “Mylifebits: a personal database for everything,” ACM, vol. 49, no. 1, pp. 88–95, 2006.
- [41] W. Tan, J. Dwivedi-Yu, Y. Li, L. Mathias, M. Saeidi, J. N. Yan, and A. Y. Halevy, “Timelineqa: A benchmark for question answering over timelines,” in Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023 (A. Rogers, J. L. Boyd-Graber, and N. Okazaki, eds.), pp. 77–91, Association for Computational Linguistics, 2023.

- [42] G. Meditskos, P.-M. Plans, T. G. Stavropoulos, J. Benois-Pineau, V. Buso, and I. Kompatsiaris, “Multi-modal activity recognition from egocentric vision, semantic enrichment and lifelogging applications for the care of dementia,” Visual Commun. and Image Representation, vol. 51, pp. 169–190, 2018.
- [43] R. P. Goldman, C. W. Geib, H. A. Kautz, and T. Asfour, “Plan recognition (dagstuhl seminar 11141).,” Dagstuhl Reports, vol. 1, no. 4, pp. 1–22, 2011.
- [44] G. Sukthankar, C. Geib, H. H. Bui, D. Pynadath, and R. P. Goldman, Plan, Activity, and Intent Recognition: Theory and Practice. Morgan Kaufmann, 2014.
- [45] D. L. Schacter, The seven sins of memory: How the mind forgets and remembers. HMH, 2002.