ForgetIT

Concise Preservation by Combining Managed Forgetting and Contextualized Remembering

Grant Agreement No. 600826

Deliverable D9.5

Work-package	WP9: Personal Preservation
Deliverable	D9.5: Personal Preservation Report
Deliverable Leader	Heiko Maus, DFKI
Quality Assessor	Doron Chen, IBM
Dissemination level	PU
Delivery date in Annex I	M36
Actual delivery date	February 2016
Revision	1.0
Status	final
Keywords	Personal Preservation, Semantic Desktop, PIMO, Personal Information Management, Preservation Selection, Evaluation

Disclaimer

This document contains material, which is under copyright of individual or several ForgetIT consortium parties, and no copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the ForgetIT consortium as a whole, nor individual parties of the ForgetIT consortium warrant that the information contained in this document is suitable for use, nor that the use of the information is free from risk, and accepts no liability for loss or damage suffered by any person using this information.

This document reflects only the authors' view. The European Community is not liable for any use that may be made of the information contained herein.

Revision History

Date	Version	Major changes	Authors
November	0.1	initial version	DFKI
December	0.2	outline of sections; scenario 1	DFKI, LUH,
December	0.3	evaluation results	UEDIN
January	0.4	Scenario 1 application	TT
January	0.5	finalizing	DFKI
January	0.6	brushing; prepare for QA	DFKI, UEDIN, LUH, TT
February	0.7	final version for QA	DFKI
February	0.8	comments from QA	DFKI,L3S,TT
February	1.0	2nd QA; finalized	DFKI

List of Authors

Partner Acronym	Authors
DFKI	Heiko Maus, Sven Schwarz, Christian Jilek
UEDIN	Maria Wolters, Stephen Rhodes
LUH	Andrea Ceroni
TT	Gürkan Gür

Table of Contents

Glossary   **67**

Executive summary

This deliverable presents ForgetIT's *Personal Preservation Report* which is an outcome of the Personal Preservation application scenario of work package WP9. It contains the final reports on the evaluation results, the final presentation of the application for supporting users in selecting photos for preservation, the Personal Preservation Report, the discussion of the reviewer's recommendations, and concludes with the final assessment of WP9's success indicators.

The deliverable presents the final results of the student group PANIC derived from the usage statistics and final questionnaire. Moreover, the final evaluation of WP9 is presented where 10 participants used the Personal Preservation Pilot to organize their personal photo collections and could inspect the preservation decisions based on their Preservation Strategy.

The application 'Preservation of Personal Photo Collections' already presented in D9.3, supports users in selecting photos for preservation in their personal photo collections. Besides additional features using new ForgetIT technology, the application has been connected to the Preserve-or-Forget (PoF) Middleware allowing to immediately preserve photo collections. By introducing situation search, it allows a novel way for finding photos in the archive.

The Personal Preservation Report presents WP9's approach for Personal Preservation by embedding it in a user's Personal Information Management (PIM) activities. The obstacles for Personal Preservation identified in D9.1 are revisited and the contributions to overcome these obstacles with the presented approach are investigated. This report is complemented by lessons learned from the DFKI PIMO – which is in everyday use by the DFKI team – using ForgetIT technology and goals (Managed Forgetting, Contextual Remembering, and Synergetic Preservation).

Finally, the reviewers' recommendations from Year 1 & 2 reviews are discussed including their impact on the development in WP9.

A final assessment of WP9's success indicators shows that WP9 reached the intended goals of the work package.

1   Introduction

This deliverable presents the *Personal Preservation Report* which is an outcome of ForgetIT's application scenario work package WP9. The report builds and concludes the deliverables D9.1 [Maus et al., 2013] and D9.2 [Maus and Schwarz, 2014], and the Personal Preservation Pilots I & II presented in D9.3 [Maus et al., 2014] and D9.4 [Maus et al., 2014], respectively.

The deliverable presents WP9's evaluation results in Section 2 which are twofold. First, Section 2.1 details the final evaluation results of the student group PANIC derived from the usage statistics and final questionnaire. Second, the final evaluation of WP9 is presented in Section 2.2 where 10 participants used the Personal Preservation Pilot to organize their personal photo collections and then could inspect the preservation decisions based on their Preservation Strategy.

Section 3 presents further work on the application 'Preservation of Personal Photo Collections' (in the review referred to as "Scenario 1") which supports users in selecting photos for preservation in their personal photo collections. Besides additional features using new ForgetIT technology, the application has been connected to the PoF Middleware allowing to immediately preserve photo collections. By introducing situation search, it allows a novel way for finding photos in the archive.

The Personal Preservation Report in Section 4 presents WP9's approach for Personal Preservation by embedding it in a user's PIM activities (Section 4.1). Section 4.2 revisits the obstacles for Personal Preservation identified in D9.1 and investigates the contributions of WP9 to overcome these obstacles. Section 4.3 complements this report by lessons learned from the DFKI PIMO – which is in everyday use by the DFKI team – using ForgetIT technology and the goals Managed Forgetting, Contextual Remembering, and Synergetic Preservation.

Section 5 presents and discusses the reviewers' recommendations from Year 1 & 2 reviews including their impact on the development in WP9.

Finally, Section 6 presents the final assessment of WP9's success indicators which shows that WP9 reached the intended goals of the work package. A glimpse on future visions closes this deliverable.

1.1   Target Audience

The target audience are all readers interested in how ForgetIT goals are realized and evaluated in the personal preservation application scenario.

To fully understand the presented results, the audience should be familiar with the Personal Preservation Pilots I & II D9.3 [Maus et al., 2014] and D9.4 [Maus et al., 2015], respectively, as the results build upon these pilots and the terminology explained there.

1.2   Structure of the Deliverable

The deliverable is structured as follows:

Section 2 starts with reporting the results of a formative evaluation with the Semantic Desktop with a user group in Section 2.1. Section 2.2 presents in detail the final evaluation of WP9 Personal Preservation scenario.

Section 3 describes the extension of the application to assist users in the preservation of Personal Photo Collections (also known as Scenario 1 for the reviews in 2015 & upcoming 2016).

Section 4 is dedicated to the final report on the Personal Preservation Application Scenario. This also includes lessons learned from the daily usage of the Semantic Desktop at the DFKI.

Section 5 presents our reactions to the ForgetIT review recommendations of year 1 & 2.

The Conclusions in Section 6 provide the final assessment of the success indicators and presents a vision for the future.

2   Evaluation

This Section presents the result of two evaluations conducted in WP9 in close cooperation with UEDIN. First, Section 2.1 concludes the evaluation of the student group PANIC where intermediate results were reported in D9.3. Second, Section 2.2 details WP9's final evaluation using the Personal Preservation Pilot.

2.1   Formative Evaluation with the PANIC User Group

Formative evaluations with end users are an important part of the software development cycle. Such evaluations provide feedback on intermediate stages of a system and highlight potential problems early so that they can be integrated into the task list and appropriately prioritised during development.

For the Personal Preservation Pilots, this data was gathered in a series of long-term case studies where four students used the Semantic Desktop for six months (May-October 2014) as part of their practical (a "research project" as part of their study), for which they received course credit. Thus, the situation is similar to one where the Semantic Desktop is deployed in a work environment where staff is incentivised to use it.

The acronym for the group, PANIC, is derived from the students' names. As described in D9.3 [Maus et al., 2014], the four students vary greatly in terms of their background, their attitude to preservation, and their technology literacy. All four were friends, took some classes together, and met frequently outside of university.

For the usage period, the PANIC group had access to a dedicated PIMO Server where only the students were registered as users. Each of the four students had their private instance of a Personal Information Model (PIMO) with the possibility to share with the group. To respect their privacy, together with DFKI's data protection official, a data protection agreement was developed which allowed anonymized analysis of collected data but no access to the PIMO anyone but the PANIC group, and the deletion of the PIMOs once PANIC stops using it. All information on PIMO contents were deliberately communicated by PANIC.

The PANIC group used the Semantic Desktop infrastructure and early prototypes of components such as the PIMO5 interface, Seed, the PIMO Photo Organisation app, and PIMO Diary to organise their personal, student, and professional lives. The students met with the DFKI project team regularly over the course of six months, during which PIMO was constantly developed, and system performance and stability was improved.

In this document, we summarise the overall approach to evaluation (Section 2.1.1), present final results on usage patterns and acceptance (Section 2.1.2), and discuss the implications of those results for the project (Section 2.1.3).

2.1.1   Design of the PANIC Evaluation

The PANIC study provided ongoing formative feedback to the DFKI team during a time when much of the basic functionality of the final system was being developed and implemented. The system was therefore not ready for unsupported, independent use, and a formal evaluation of the kind that was conducted for the final research prototype (see Section 2.2) was not appropriate. Instead, the evaluation focused on two questions: Are users able to appropriate the Semantic Desktop for their own needs in a meaningful way, and what are the main barriers of usage?

Identifying main barriers of usage allows the developers to prioritise functionality improvements, while appropriation indicates that people are likely to find the system useful once the barriers are eliminated or mitigated. Appropriation [Dix, 2007] is a technical term from Human-Computer Interaction that covers users' ability to use a system in unexpected, innovative ways while adapting it to their particular needs.

Appropriation and main barriers of usage were assessed through a questionnaire. In addition, we consulted system usage logs during the period. The questionnaire consisted of six open questions, summarised in Table 1, and was sent to the four students 10 months after the end of the period of intensive PIMO use. Thus, the students' assessment of PIMO reflected their strongest, most enduring impressions of the system, and allowed them to judge PIMO in light of other data management approaches that they had adopted afterwards.

While students agreed to fill in the final questionnaire as part of their course activities, the questionnaire was distributed and completed well after the students had received their course credits, and in at least two cases, after they had been awarded their final degree. Therefore, there was no incentive to paint the system in a favourable light for credit. In addition, the evaluation was conducted by the UEDIN team, who are completely independent from the developers at DFKI. Both aspects were intended to encourage honest answers.

2.1.2   Summary of Findings

As the overall usage statistics show, the students actively engaged with the system. They accessed it frequently (see Figure 1) and customised it thoroughly and shared it to ensure that the PIMO vocabulary used for annotation reflected their own needs and concerns (Figures 2 and 3). Therefore, we are confident that the students' answers are based on extensive practical knowledge of the Semantic Desktop.

Positive Aspects of the Semantic Desktop

First of all, students enjoyed working with a system that was still being developed, because it gave them unique insights into the process. They also liked to experiment and

Table 1: Final Student Questionnaire

1. What was your best experience of the PIMO system?
   Additional prompts: time saving, fun, particular added value

2. When and why did you stop using PIMO?
   Additional prompts: no longer worth it, end of practical, potential incentives to keep using PIMO

3. What PIMO functionality do you miss the most?
   Additional prompts: work/home, functionality you would like to see in other services

4. How do you organise your data now?
   Additional prompts: work/home, documents/photos/e-mails, operating system specific solutions

5. How do you coordinate projects and events with other people now?
   Additional prompts: work/home, projects with many documents versus travel

6. What are the three key aspects where PIMO should be improved?
   Additional prompts: desktop vs. mobile, new additional functions, collaboration

Figure 1: Access patterns of the PANIC group over time (Start in May and official end in the 2nd week of October. 2 weeks after that the group broke up.)
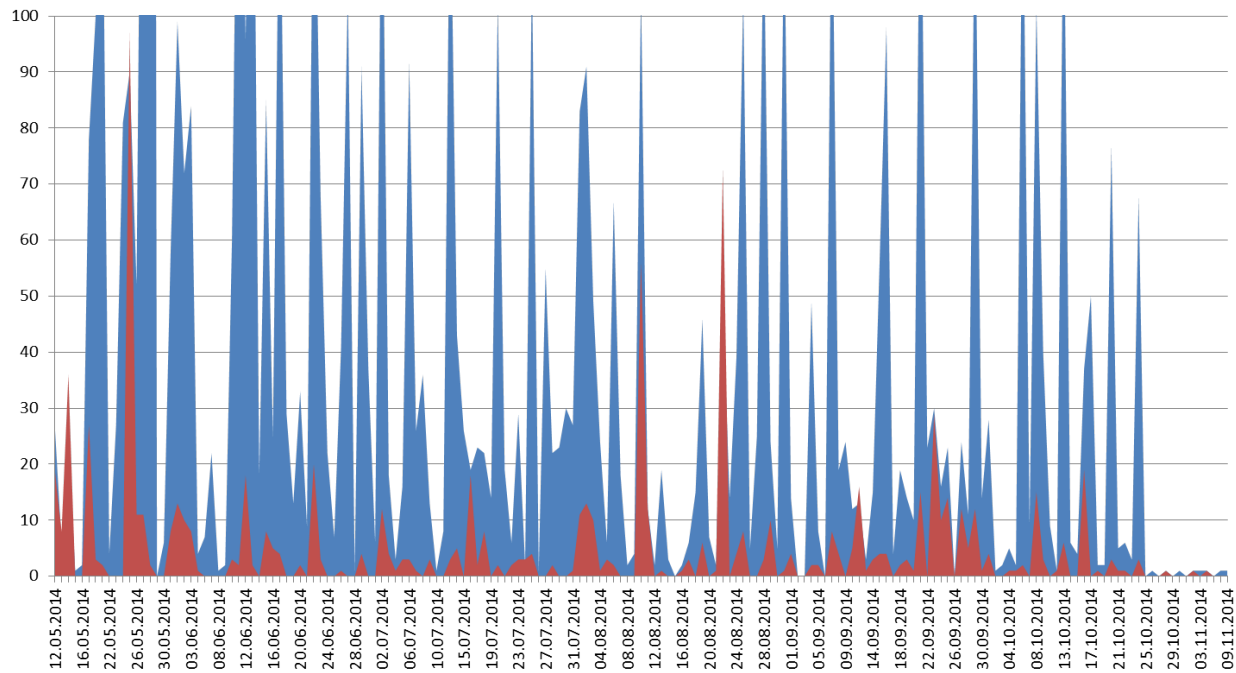
Figure 2: Create (blue) & Share (red) Actions for PANIC group.

7052 Things		
	6286	Media (Docs, Webpages, Images)
	167	Topics
	153	Events
	144	Agents (Persons, Groups, Orga.)
	105	Jobs (Tasks, Projects)
	41	Locations
	32	LifeSituations
	12	Tools
	112	Other (Thing as type)

Figure 3: Number and Types of Things created over time

discover new functions, and they were able to suggest functions that were implemented by the team, such as the PIMO5 Home Screen which showed all relevant new activity.

They highlighted that the Semantic Desktop was a very comprehensive system that could be used for multiple domains (work / private life / study). It provided a single home for all documents, which was accessible from any device that connects to the Semantic Desktop. Thus, there was no longer need for multiple cloud services, and important documents, such as task lists, were always at the users' fingertips.

The Semantic Desktop infrastructure was a decisive advantage for PIMO. All students mentioned that they missed this functionality once they had stopped using the system. Annotating documents, web pages, and e-mails with topics was straightforward. It afforded far more flexibility than a hierarchical folder structure. The topics were easy to maintain and useful for searching. One of the students, whose digital data had been relatively disorganised, found it easy enough to organise her digital life using the Semantic Desktop tagging and annotation functions. Student A demonstrates neatly how the Semantic Desktop makes data easier to manage. She noted that when using PIMO, she could easily assign a Topic to each e-mail. Without PIMO, she is forced to use hierarchical folder structures, which she only creates for clearly demarcated, recurring groups (online orders, work e-mail). This leaves a substantial amount of e-mails uncategorised.

While the initial effort to create relevant topics was substantial, it was mitigated by the ability to share annotations. The students also appreciated the Semantic Editor Seed, which had become more stable and reliable compared to the intermediate interviews reported in D9.3. Concepts were detected automatically, linked, and stored in the PIMO, which made it easier to manage and file notes.

Students also highlighted the potential of the rich, automated analytics afforded by PIMO, as exemplified by the diary function. The PIMO Diary allows users to easily see what happened in a given time period, and thus makes it easy to reminisce.

The final important advantage of PIMO was its strong support for collaboration. Student P noted that "[i]t became like a useful version of Facebook." (Student P). Two other students noted that PIMO could be usefully integrated with Facebook, or replace it altogether, with a large enough user base.

The students used PIMO for collaborations ranging from planning joint trips to preparing for common exams. Sharing document annotations was an easy way to let others benefit from one's own research. Overall, the collaboration functions were regarded as a key time saver, and were frequently mentioned as the main reason for using PIMO as frequently as they did.

Figure 2 shows the distribution of creating things and sharing actions of the group. As can be seen by the distribution of creation actions (in blue), that over the period things were continuously created (Figure 3 shows the distribution among the types). To consider here is that in the PIMO, things created are private by default and need to be shared explicitly. Therefore, the sharing actions (in red) show a significant amount of sharing continuously took place among the group members. However, there is still a difference as only 9% of

all things were shared. Considering personal preservation, this is a first indication that is seems feasible that private things are covered using the PIMO (as opposed, e.g., to the nature of social media such as facebook, where users intentionally post most material to be seen in public).

Appropriation

We already saw strong evidence of appropriation in the intermediate interviews, where the successful use cases were searching annotated guitar tabs, planning a weekend trip, planning a trip to a festival, and organising course material. The students showcase some of this range of functionality in the video they created.[1]

The final questionnaire data reinforces this finding. Perhaps the best example for appropriation is the shared recipe collection described by student N. All four students shared annotated recipes with each other, there was no longer a need to search different web sites, because the relevant web pages with recipes had been annotated and added to the PIMO, and the group acted as an additional quality control filter.

An important theme across all key use cases identified in the final questionnaire was collaboration. The PIMO5 Home Screen offered a well curated view of what was happening in the other students' personal and professional lives. Students also shared information about deadlines and appointments. In one instance, Student C only remembered an appointment because others had filed it in PIMO. Sharing annotated resources saved time, because students were able to divide the work of annotating and curating documents.

Barriers to Use

Despite many positive aspects, and clear evidence for appropriation, all four students stopped using PIMO shortly after as the practical ended (see Figure 1). There were two main reasons for this, disintegration of the group and issues with the PIMO implementation.

While the group of students had been working together relatively closely during the time of the practical, immediately afterwards, their study paths diverged as they began writing their Master's Thesis (as also can be seen in the steep drop of usage in November in Figure 1). Since PIMO was only open to them, but not to others, they could no longer use it to collaborate across both their student and their private life. Thus, it became yet another channel, in competition with widely used apps such as Facebook, Doodle, and WhatsApp, and one of the unique selling points was lost. It also fell out of daily use.

The other unique selling point, the Semantic Desktop, could have been especially advantageous now that students embarked on their final thesis, which had to be completed in three months. However, parts of the Semantic Desktop were difficult to install on new

---

[1]See the Pilot documentation at `https://pimo.opendfki.de/wp9-pilot/pimopanic.html`

computers[2], slow, resource intensive, required for the students a substantial amount of disk space[3], and still had stability issues. This meant that it was not judged to be a suitable tool for a very important, time-critical piece of work. In this context, the rich Semantic Desktop features turned into a disadvantage. While puzzling out functionality was interesting and fun during a practical, it is an unwanted distraction when one needed to focus on a major piece of work.

The improvements suggested by the students fall into four categories:

New functionality – Chat: An integrated chat function would allow users to talk and co-ordinate more efficiently, instead of having to use an external service such as WhatsApp, Facebook Messenger, or e-mail. The function should also allow users to send each other data and files.

Easy on-ramp: If the Semantic Desktop were easier to install and had a more intuitive user interface design, this would make it easier to recruit new users

Broader support of 3rd party applications: The implementation used by the students only supported Firefox and Thunderbird. Support for other browsers and e-mail clients, seamless interfaces with other apps such as Facebook, and easy file sharing would make it easier to integrate the PIMO into existing workflows.

New modes: Students suggested replacing the existing HTML5 browser-based implementation (PIMO5) with a native PIMO mobile app. They also agreed on the need for an off-line mode, because they often wanted to use PIMO in situations without Internet connectivity, such as on train rides.

2.1.3   Implications for the Project

While the PANIC study was small, using a case study design with only four students, we are reasonably confident that the questionnaire data covers the main barriers to use and additional functions that need to be addressed in order to deploy the Semantic Desktop with groups of people, such as family, friends, or coworkers. For both aspects, we reached saturation, i.e., after analysing 2–3 questionnaires, the fourth did not yield any new data. In qualitative research, this is regarded as an indicator of comprehensive coverage.

The four students clearly saw the value of both PIMO as a unified system for organising and preserving one's personal and professional life, and of the Semantic Desktop approach. Even students who regarded themselves as relatively disorganised succeeded in creating useful structures, and two students noticed that their exposure to PIMO changed

---

[2]e.g., Mac OS is not supported by the SemanticFileExplorer and handling Java there is challenging for users.

[3]The disk space required for that kind of installation amounts to approx. 50MB with SemanticFileExplorer 10 MB, FireTag add-on twice (for Thunderbird and Firefox) 65KB = 130KB, and the Java GUI 28 MB. PIMO5 – on the way to fully replace the Java interface – simply uses the browser cache and can be neglected.

Table 2: Success Indicators and Metrics

Indicator	Metrics
Ease of use of the personal Preservation Strategy	task success, number of steps required for task, workload
Low additional effort for personal preservation	number of changes required after personal preservation settings have been applied
User satisfaction with the personal Preservation Strategy	System Usability Survey, interviews

how they managed their data. For example, student A previously only used folders for filing documents, but after PIMO, she also uses folder structures for e-mail and web browser bookmarks.

What makes PIMO more beneficial than another knowledge management systems is the strong support for collaboration between group members. It was one of the main benefits highlighted by the students and featured in all of the successful use cases discussed in the final questionnaire.

The main barrier to successful deployment and commercialisation appears to be the lack of a dedicated core development team for product-like development. Almost all of the areas for improvement highlighted by this study concern the implementation of the Semantic Desktop infrastructure, and require far more resources than those available to a single research team. Once this is in place, PIMO will be an excellent environment for personal preservation, because it will be simple and straightforward to enrich material to be archived with sufficient context and detail.

2.2   Personal Preservation Use Case: Results

In this section, the evaluation of the personal preservation use case is presented. The goal was to show the contribution of the Personal Preservation Pilot to the success indicators of WP9, which are summarised in Table 2.

2.2.1   Evaluation Setting

The evaluation setting included 10 participants hired by the University of Edinburgh using parts of the Personal Preservation Pilot to organize their own photos and finally, to judge the preservation proposal of the pilot. Please refer to deliverable D9.4 [Maus et al., 2015], Section 2, for an explanation of the user interfaces and services used in this evaluation.

To help users who are new to a personal preservation service, it would be useful to have a predefined set of policies and rules, so that unexperienced users would not need to care about defining a rather complex Preservation Strategy. Therefore, the Personal Preservation Pilot allows to select one out of four personas as individual Preservation Strategy

used in the Semantic Desktop. As described in detail in D9.4, these personas represent attitudes towards personal Preservation Strategies and were derived from the ForgetIT survey on personal preservation of photos conducted in WP2 (see [Wolters et al., 2015], D2.4 [Logie et al., 2016]). The four personas identified are Safe Curators, Safe Filer, File & Forget, and File & Forget Curator.

The evaluation used hired participants which were previously unknown to us. To identify the respective persona, a participant answered in the first session several questions used in the WP2 survey allowing to specify their persona. This persona is then set as Preservation Strategy in the PIMO5 options by the experimenter.

Once the Preservation Strategy is set, the so-called *time capsule* can be invoked. The selected Preservation Strategy is used for a new calculation of the Preservation Value and the time capsule view is opened.

The time capsule gives an overview of the photo collections and the respective set of photos to be preserved and those that will not be preserved. Participants were able to inspect these automated decisions and change the decision manually. Changes done by the users were logged and used for evaluating the persona-based Preservation Value (PV) Assessment.

2.2.2 Summary of Evaluation Method

The ten participants used the PIMO5 user interface to the Semantic Desktop to work with their own photographs. Participants brought 40–80 images from a holiday to the first session, which were organised into specific events (as photo collections). Table 3 presents the number of photographs brought by each participant and the number of collections made, as well as the number of duplicates detected by the Semantic Desktop.

In the second session, participants performed a series of tasks with their two largest collections. The five tasks were as follows:

1. to review and change preservation preferences,
2. to apply one of the detected visual concepts as a filter,
3. to add a note describing the collection as a whole complete with key concepts,
4. to annotate individual photographs using concepts to make them more searchable, and finally
5. to search the PIMO for a user-defined concept.

In the third and final session, participants reviewed the time capsule generated by the Semantic Desktop based on evidences gathered in the first two sessions.

	N collections	N photos	N duplicates
eval01	4	40	2
eval02	5	57	14
eval03	3	89	12
eval04	6	49	0
eval05	3	68	44
eval06	7	73	11
eval07	4	54	17
eval08	4	60	3
eval09	3	76	2
eval10	6	78	50
Mean			
SD			

Table 3: Number of collections and photographs for each evaluation participant

Figure 4: Concept icon in PIMO5: clicking on the label will open the thing, clicking on the magnifying glass will start a search with the thing as filter.

2.2.3   Task Performance

Participants completed the five tasks, outlined above, for their two largest photo collections. For tasks 1, 3, and 4 all participants eventually succeeded, however, for tasks requiring participants to apply concepts as filters to search the PIMO (tasks 2 and 5) errors were frequent. With the first collection, 4 out of 10 participants successfully applied a visual concept as a filter (Task 2) and 6 were able to correctly search for a user defined concept (Task 5). Success rates were similar for the second collection with 3 out of 10 participants succeeding at Task 2 and 7 for Task 5. The error made by participants when searching through the PIMO was to click on the concept name, thus revealing details of the concept itself, rather than clicking on the magnifying glass icon to apply said concept as a filter (see Figure 4).

Unsurprisingly Task 4—which required participants to annotate individual photographs—took, on average, longer to complete (see Figure 5) and required more clicks (see Figure 6). Although accuracy was very high, in general, observation revealed additional issues encountered by participants while working with the PIMO. These issues primarily centred around annotating collections and photographs (tasks 3 and 4, respectively) with Seed. Often participants would repeatedly add a new PIMO 'thing' (e.g., a friend's name) as there was a slight time-lag between adding a concept and the relevant text becoming highlighted (signifying that the text was successfully linked to the newly created thing). This duplication of PIMO things occasionally caused problems in Task 5 when

	SUS	NASA-TLX					
		1	2	3	4	5	6
eval01	30	7	1	11	11	8	14
eval02	62.5	7	2	3	7	8	9
eval03	80	9	4	7	5	7	6
eval04	70	18	4	11	9	17	2
eval05	77.5	8	1	1	11	8	2
eval06	60	13	1	3	4	12	12
eval07	82.5	2	3	8	4	3	1
eval08	87.5	6	1	7	5	10	5
eval09	82.5	1	1	7	11	1	1
eval10	5	13	3	12	3	13	12
Mean	63.75	8.4	2.1	7	7	8.7	6.4
SD	26.6	5.2	1.3	3.7	3.2	4.7	5

Table 4: Results of the SUS and NASA-TLX for each evaluation participant. The SUS is scored from 0–100 with larger scores indicating better usability. The NASA-TLX probes 6 dimensions of task difficulty; 1) mental demand, 2) physical demand, 3) temporal demand, 4) performance (how successful were you in performing the task), 5) overall effort, 6) frustration. Each is scored on a 1-21 scale.

participants came to search for user-defined concepts; a number of identically named things were returned by the search and it was unclear which were successfully linked to photographs.

2.2.4 Participant Feedback

Once participants had completed the five tasks across two collections they were asked to complete both the SUS and the NASA-TLX, the results of which are presented in Table 4. In general the system was rated quite highly, as seen in the overall mean SUS rating and also in the fact that 6 participants gave a rating of 70 or higher. However, two participants (eval01 and eval10) gave particularly low ratings (30 and 5, respectively); the results of the NASA-TLX suggest some reasons for this. It is interesting to note that the only dimension of the NASA-TLX to correlate significantly with the SUS, even with only 10 participants, was the item on frustration during the task ($r = -0.78$, $p < 0.01$) and participants 1 and 10 gave particularly high scores on this item.
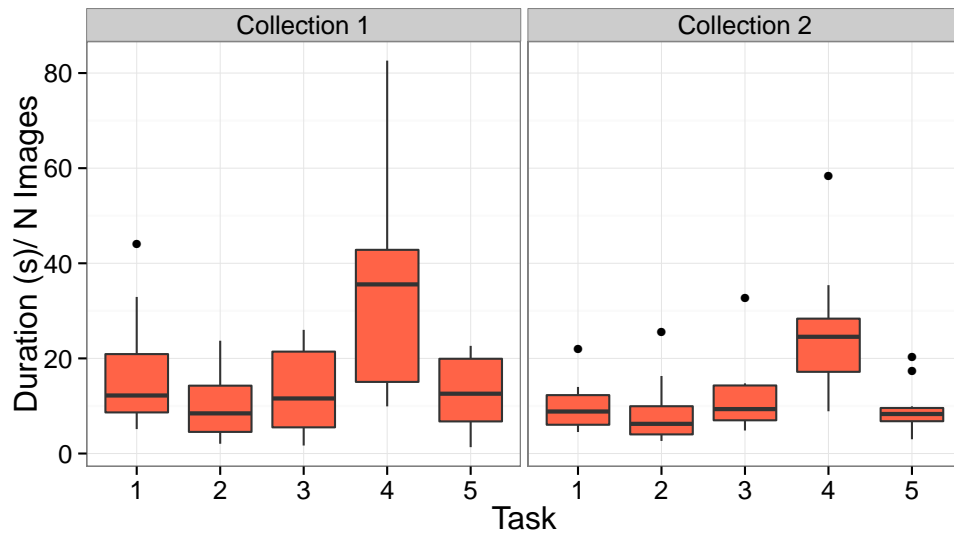
Figure 5: Time taken to complete each task for each collection worked with, normalised by collection size.
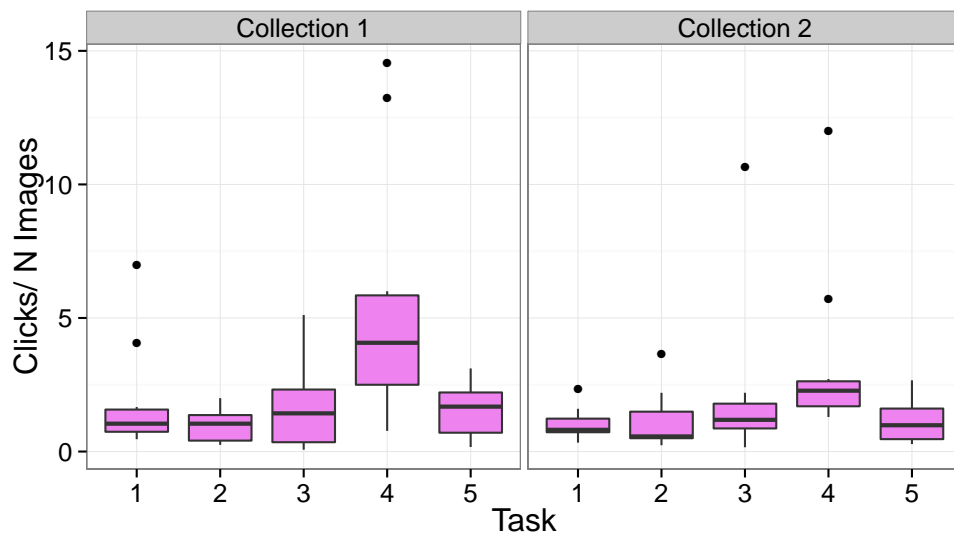

Figure 6: Number of clicks to complete each task for each collection worked with, normalised by collection size.

Further detail regarding each user's impression of the system was gained via a semi-structured interview following completion of the tasks in Session 2. This interview covered:

1. the user's overall impression of the system,
2. their impression of adding annotations to collections and photographs,
3. their impression of searching for concepts, and
4. any improvements they would suggest for any other issues identified.

In what follows, we provide a summary of the key themes emerging from this interview.

Overall impression of the system

In line with the findings from the SUS (see Table 4), participants generally reported liking the system overall. In particular, a common statement was that participants felt it would be useful for those with a large collection of photographs to organise and search through. The ability to add informative annotations and create new concepts (PIMO things) was mentioned most often (by 4 participants) as a useful aspect of the system. Further, for some participants (in particular eval09) the ability to hide—rather than completely delete—an image was an appealing aspect that is not shared by other photo organising software. Several participants, however, commented that they felt they would need more practice with the PIMO in order to use in properly. Particular problems that users noted tended to relate to navigating the PIMO. Many of the users mentioned that the back button along the top bar of the interface did not take them where they expected and that it often led them out of the collection they were working with (this was a particular issue for eval10 that may have contributed to frustration during the tasks).

Annotation

As outlined above participants generally liked the ability to create new concepts that could be searched for later. Some felt, however, that the number of categories into which new PIMO things could fall (Location, Person, Organization, Event, Media, Topic) was too restrictive and they would like more options (although none were suggested). Several users commented that they found adding annotations easy, especially for straightforward tags like well known locations that were easily found in databases (i.e., freebase and Wikipedia in this evaluation setting). Some users (evals 6, 7, and 10) encountered problems creating new PIMO things where it was not clear that the new thing had been linked to the text. This resulted in duplicated instances of PIMO things and this was mentioned as a particularly confusing aspect by these users.

Regarding the automatically generated visual concepts (by CERTH components), the views were mixed. Participants generally found that they were accurate at detecting aspects of their images. However, some of the users questioned the potential usefulness of vague concepts, such as 'outdoor' or 'vegetation', when coming to filter through a large collection. Eval06, for example, mentioned that she would like to be able to reject auto-

matically generate concepts (the procedure for doing this was not demonstrated to users).

Searching

Given the number of errors made in the tasks requiring users to apply concepts as filters to search the PIMO, it is unsurprising that many participants commented during the interview that the search results were confusing and incomplete. Several users commented that they did not find applying concepts as filters intuitive and, when explicitly asked, eval10 mentioned that the separate buttons for opening concepts and applying them as filters (see Figure 4) were not clear. He suggested giving separate background colours as a potential way of distinguishing the two functions. Given that when participants made an error during Tasks 2 or 5 it was due to clicking on the concept name rather than the magnifying glass icon, we can be fairly confident that users did not understood the separate functions, despite this being covered in the Session 1 tutorial. However, in principle users thought the ability to add searchable tags to find images in a large collection was a good idea, in particular the ability to go beyond the people in the image (as is standard in social media applications like facebook) to add other informative concepts.

Suggestions for improvement

The main issue raised by users, in addition to their comments on annotation and searching, was the PIMO's handling of duplicates. Currently images deemed duplicates are linked to a single 'master' image and decisions made on this image affect all duplicates. Users with many duplicates in their collections (e.g. eval02, eval08 and eval10) commented that they would like control over duplicates and to be able to make decision on each individually. This issue was also raised in Session 3 when making time capsule decisions (see below). Another aspect mentioned less frequently concerned minor aesthetic details of the system. While some liked the lack of clutter (eval07) others (in particular eval04, eval06 and eval08) felt the large amounts of blank space could be better utilised. This may be partly due to the evaluation being conducted on a desktop computer with a 26" screen.[4] Specific suggestions were to possibly be able to choose an image from the collections to act as the background and to enlarge the image icons in collection view.

Considering the problems of the participants to distinguish between browsing to a thing and starting a search with the thing as filter, PIMO5 has now a clearer distinction between the thing's label and the magnifying glass button (see Figure 7).

---

[4]The same computer was used for the Festival 2013 and 2014 studies, the WP9 and the WP10 evaluations. The screen size had been selected with the Festival Studies in mind, and was required by the design of the interface of the WP10 solution.
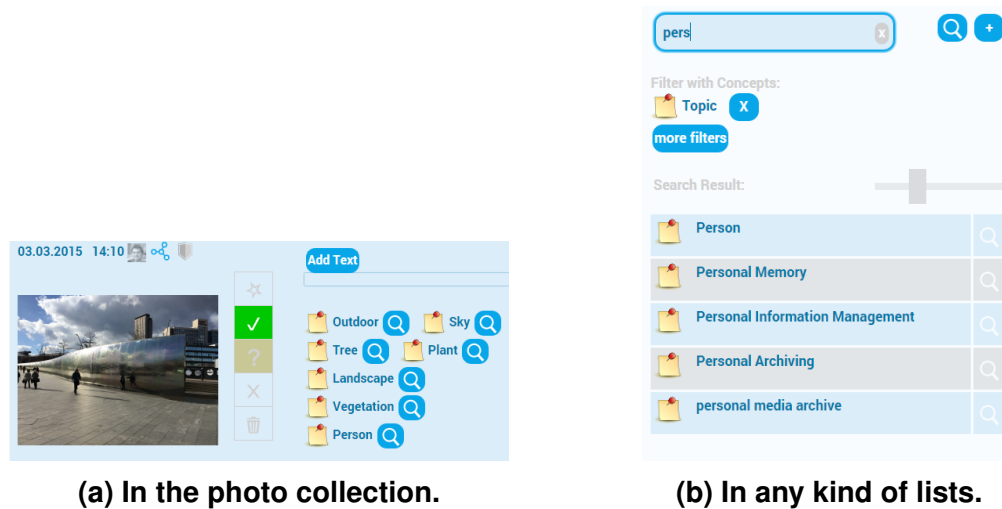
(a) In the photo collection.   **(b) In any kind of lists.**

Figure 7: Taking the recommendation from the evaluation into account, now a clearer distinction btw. a thing and the button magnifying glass was introduced in PIMO5.

2.2.5   Favourite/Show/Hide/Trash Decisions

Considering the 20 collections that were fully modified by participants during the course of Session 2 (2 collections each participant) we find that, of the 348 Favourite/ Show/ Hide/ Trash (FSHT) decision made by the Semantic Desktop at upload, 167 were changed by the user. The majority of these values were changed once (151) although some were changed 2 (13) or 3 (3) times. Figure 8 depicts the changes that users made to FSHT values. It is clear that when an image was hidden by the PIMO at upload participants generally agreed with this decision (only 8 hide decisions were modified) and as Figure 9 shows 33 hide decision were left unchanged. Participants generally agreed that they would not want poor quality images in their collections and liked the ability to hide, rather than completely delete, photographs. The reasons given by participants when hiding images were that the image was of poor quality, that it captured something that was relevant at the time but no-longer meaningful, or that it was a duplicate or captured redundant information (e.g., the same object from a different view).

As is clear from Figures 8 and 9 the most utilised decision was 'show' with 190 images falling in this category, this was followed by 'hide' with 67 images hidden. The 'favourite' and 'undecided' categories received a similar number of images, with 44 and 46, respectively.

2.2.6   Time Capsule Decisions

Table 5 presents results from the third session in which users reviewed their time capsule. The exact parameters used to produce the capsule depended on the user's persona (see column 2,) which was determined via responses to the ForgetIT survey (as explained in Section 2.2.1). The third column of Table 5 shows the variability in the percentage of
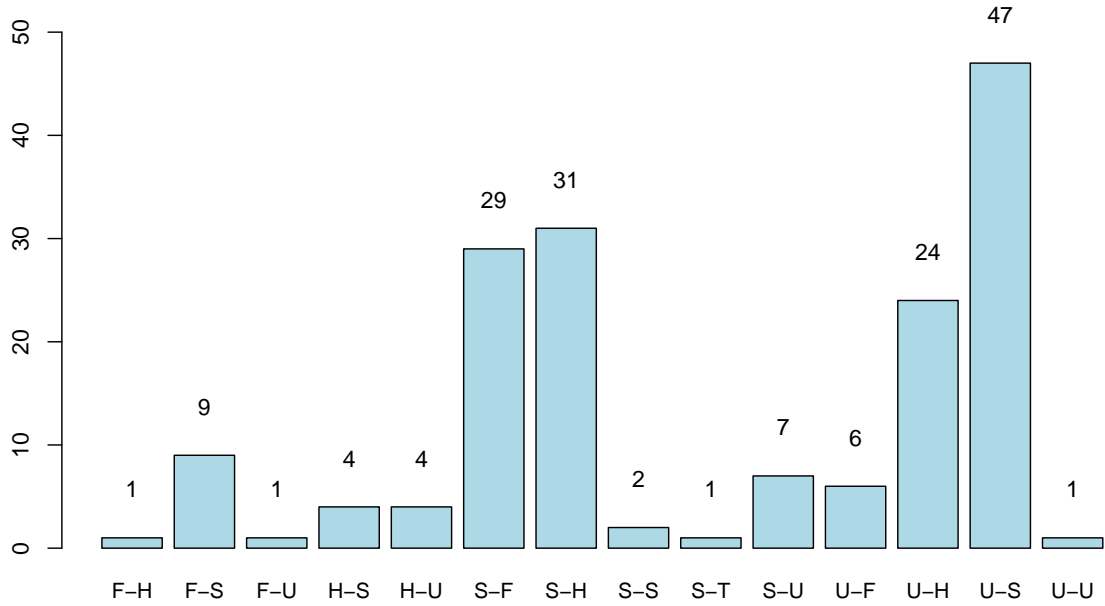
Figure 8: Changes made by users to Favourite/Show/Hide/Delete values. F = favourite, S = show, U = undecided, H = hide, T = Trash. Labels describe the transition from the initial upload decision by PIMO and the final decision made by the user.
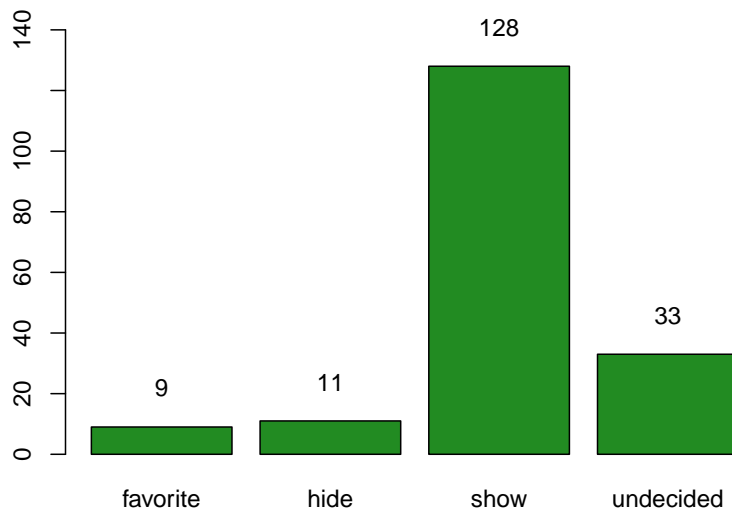

Figure 9: Values left unchanged by users from the initial upload decision made by PIMO.

	Preservation Strategy/Persona	% archived	Participant decision		
			N archived	N removed	% agree
eval01	File & Forget	76	4	2	76
eval02	Safe Curator	71	1	1	93
eval03	Basic Curator	83	4	0	94
eval04	Basic Curator	61	0	2	93
eval05	File & Forget	73	1	0	93
eval06	Safe Curator	83	1	2	87
eval07	Basic Curator	76	0	2	92
eval08	Safe Curator	55	1	1	96
eval09	File & Forget Curator	56	1	2	96
eval10	File & Forget	60	0	0	100

Table 5: Time capsule statistics for each participant. Personas for the Preservation Strategy were determined by the ForgetIT survey.

images from the two modified collections that made it into the capsule. However, column 6 is of primary interest as it presents the level of agreement between the user and the decision made by the Semantic Desktop. That is, it presents the proportion of the time capsule decisions left unchanged by each user. When a change was made, columns 4 and 5 present the number of images that were moved to or from the capsule, respectively. The reasons given for removing an image *from* the capsule were invariably because the image captured redundant information that was captured well by a single image. The reasons given for moving images *into* the capsule were similar, in that a specific event was omitted from the capsule therefore an image was added to represent that event.

One important thing to note is that many participants had a large number of duplicates in their collections (see Table 3) that they were unable to control. Many evaluation participants commented that they would want to retain specific duplicates that captured a slightly different aspect of a place or event. Thus the level of agreement may be slightly overestimated given that participants lost control over a number of their images.

2.2.7 Conclusion

Overall, it appears that the research prototype tested in this formative evaluation fulfills the three success indicators, ease of use, low additional personal preservation effort, and acceptable user satisfaction (see Table 2). Although participants struggled with the search functionality, this is a user interface design issue that can be addressed relatively easily, and the number of steps required was in line with the overall complexity of the task.

Using the standard threshold score of 68, 6 out of 10 participants judged the usability of the Preservation Strategy to be above average, scores from two participants are slightly below threshold, and only two participants, eval01 and eval10, gave the system poor scores. This indicates that the research prototype is very close to a stage of development where it can be used meaningfully by beta testers outside of DFKI for long-term field stud-

ies that are not as closely supervised as the PANIC study. While the PANIC team tested the system during development, and the participants of the 2014 Festival study interacted with an early version of the mobile interface, the version of the Preservation Strategy tested in the final formative evaluation clearly benefited from the feedback provided in earlier iterations.

The effort required for personal preservation was low. As Table 5 shows, participants mostly agreed with the preservation decisions that had been made based on their overall Preservation Strategy. The main improvement that participants requested was the ability to review the duplicate analysis, which automatically excluded some photos that should have been preserved. The current classification based on the ForgetIT survey requires users to answer less than 20 questions as part of the initial set up. Some of the items can also be determined through analysis of user behaviour (e.g., how often people delete photos or how often they use keywords), and the number of items could conceivably be reduced further by selecting the items with the highest predictive capacity. Once the existing research prototype has been revised and the user interface issues outlined above has been addressed, it should be ready for a 12–24 month study where user preservation behaviour can be observed and analysed more meaningfully than in a brief end-of-project assessment.

3   Preservation of Personal Photo Collections

The sustainable and efficient management of personal digital assets is a challenging topic as described meticulously in ForgetIT's literature. Therefore, we kept investigating the problem of supporting users in managing their personal photo collections. In the Year 2 review, we demonstrated the results of this application as "Scenario 1". We chose to dedicate further effort to this scenario since photo taking is getting more and more effortless, accepted nearly everywhere, and it causes people easily to end up with a very large collection of photos taken during various occasions, such as events at different locations and times. Simply dumping photos on some cheap large-capacity storage device is not only prone to the risk of "digital forgetting", but it also often results in having "dark archives" of photo collections, which are rarely accessed and enjoyed again (due to large effort and time required to revisit, annotate and sort photos). Therefore, users need support for browsing, storing, and retrieving their collections in an efficient and maintainable way.

Scenario 1 is instrumental for illustrating how the PoF Middleware can provide Preservation as a Service. Basically, it provides an example for extending stand-alone applications with preservation functionality by connecting to the PoF Middleware.

We do not assume any manually provided input (e.g., tags, textual descriptions, semantic annotations) beyond the mere image content to perform automatic selections, with the goal of keeping the user effort in managing large personal collections as low as possible. In principle, users could directly import their collections from cameras into the application and get an initial selection, without having to revise or annotate the collections. This makes the scenario different and, to some extent, complementary to the approach to personal preservation performed within PIMO, which embeds preservation as an additional service to the users PIM activities as well as exploits semantic annotations and relationships among resources. We do, however, allow users to annotate collections for indexing and retrieval purposes, as will be described in Sections 3.1.2 and 3.1.4.

In D9.3 [Maus et al., 2014] we have developed an application where users can exploit different methods to semi-automatically select valuable photos for subjecting them to special preservation activities. Building on extending the available application, we introduced the following new functionalities both in the front-end and in the back-end of our application:

- The back-end selection model has been improved.

- Videos are considered along with images.

- The user provides collection information that is exploited for indexing and retrieving.

- Collections are archived in the ForgetIT archive.

- The user can search over archived collections by making queries and fetch relevant content.
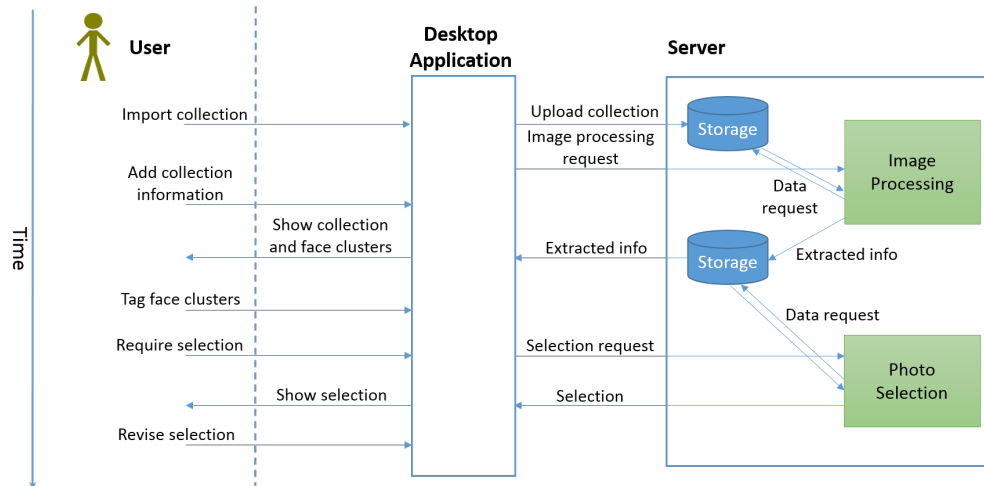
Figure 10: Workflow of the selection process.

These features and related project work are described in the rest of this section. In particular, Section 3.1 gives a conceptual description of the functionalities, while Section 3.2 presents how they have been developed within the user application.

This extended application will also be presented as "Scenario 1" in ForgetIT's final review.

3.1 Supported Functionalities

3.1.1 Improved Selection Method

The photo selection model developed within WP3, identifying important photos as a first suggestion that the user can modify at his/her convenience, has been expanded by extracting additional and more comprehensive information consisting of image aesthetics, low-level content information, sentiments, and face clustering. Experiments conducted within WP3 showed that the exploitation of such information improved the precision of the generated selections, which became closer to what user would have selected. The information extraction process has been part of WP4 and WP3, therefore detailed descriptions can be found in Deliverables D4.3 [Solachidis et al., 2015], D4.4 [Solachidis et al., 2016], and D3.4 [Zhu et al., 2016]. Also, please refer to D3.4 for an extensive description and evaluation of the new selection model applied in this scenario.

3.1.2 Acquiring User Input

The first part of the envisioned scenario, namely the phase where the user imports a collection and performs selections helped by the available method, has been enriched with more user interaction. The selection process is summarized in Figure 10. While a given collection is being imported in the system, the user is asked to provide different informa-
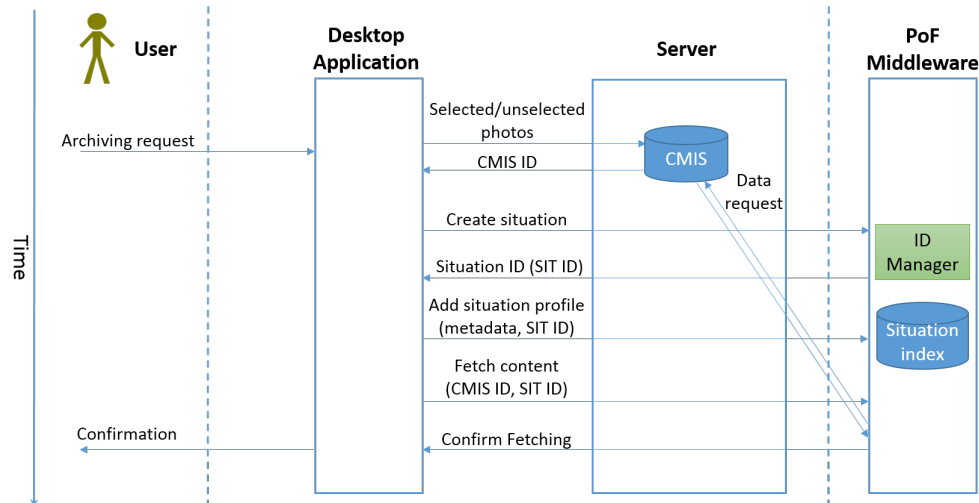
Figure 11: Workflow of the collection archiving process.

tion about the collection, namely a summarizing title, the time span and location (when missing from the EXIF metadata), the main topic (e.g., business trip, vacation), and any possible *memory cue* that the user might want to add. Memory cues represent aspects or facts that the user might want to remember about a collection. This set of manually provided information is meant to complement the one automatically extractable via image processing techniques. Another important information about a collection consists in the people occurring in it. To acquire this, the system shows the identified face clusters to the user, asking him/her to provide a name for each one. Thanks to the face cluster, the user does not have to spend time going through the whole collections looking for people. Moreover, simply asking the user to specify names without showing face clusters would fail when the user does not remember the collection content well.

3.1.3   Archiving Collections

Once the user has created and revised the set of most important photos within the imported collection, he/she has the possibility of storing the data within the ForgetIT archive by connecting to the PoF Middleware. The archiving process is shown in Figure 11. The collection is first exposed in a public Content Management Interoperability Service (CMIS) server, separating between sets of selected and not selected photos. Then, the available collection information, both acquired from the user or automatically extracted via image processing, is aggregated in a set of collection metadata. They are used to add a situation profile, representative of the collection, into the situation index. Situations are events or life experiences, as described in detail within deliverable D6.4 [Greenwood et al., 2016], and situation profiles are sets of descriptive metadata about the situation. Situations profiles are indexed so that, in case the active system does not longer exist, situations can form the unit of access to the archive. Note that the indexed situation profile only contains lightweight metadata, while the actual content of the situation (collection in our scenario) is stored in the ForgetIT archive. Please refer to D6.4 for further details on situations.

After having stored the collection metadata within the situation index for future retrieval, a fetch request to the PoF Middleware is performed to store the actual content of the collection from the CMIS server to the ForgetIT archive. Although the entire collection content is archived, keeping the distinction between selected and not selected photos has a twofold importance. First, different preservation policies can be adopted depending on the set so that, for instance, selected photos are archived at a higher degree of safety and redundancy. Second, when retrieving a collection, only the previously selected photos could be visualized to give a concise view of the most important content (still the user could explicitly ask to visualize the entire collection if he/she wants to).

3.1.4 Fetching Collections

The last phase envisioned in our scenario is the retrieval of collections given user queries and is shown in Figure 12. Thanks to the different fields of the situation index containing the extracted situation (collection in our case) metadata, the user can perform a faceted search over the desired index fields. The search can then support queries like: "Find collections taken in London between 2013 and 2015", "Find collections containing animals but not buildings", "Find collections with my wife".

The set of situation profiles matching the query are acquired in the application and shown to the user. Note that no actual content has been sent to the application so far, only collection metadata. Given the potential high transfer time that the fetch of the collection content might cause, the idea is to give to the user a first overview of what collections in the archive satisfy his/her query via their metadata, which already can provide useful summaries to somehow recall what a collection is about. The user can browse the retrieved situation profiles and possibly select those for which he/she wants to retrieve the actual content. The content of the selected profiles will be finally fetched from the ForgetIT archive via the PoF Middleware and visualized in the application.

3.2 User Application

The user application for Scenario 1 is a GUI-based tool which allows a user to manage, filter and archive collections for content ingestion while he can also contextualize and search pre-archived collections based on numerous parameters such as memory cues and situations. The main functionalities of the tool developed via prior work, such as photo selection and basic tagging, have been explained in D9.3 [Maus et al., 2014]. Therefore, in this deliverable we focus on the additional functionalities and developments integrated during the project work in Year 3 for the sake of brevity.

The main improvements in this year are related to the functionalities described in 3.1. The presented application will be part of the Scenario 1 demo in the final review. The presented screenshots are partly mock-ups based on the current state of the application showing the intended functionality for the review.
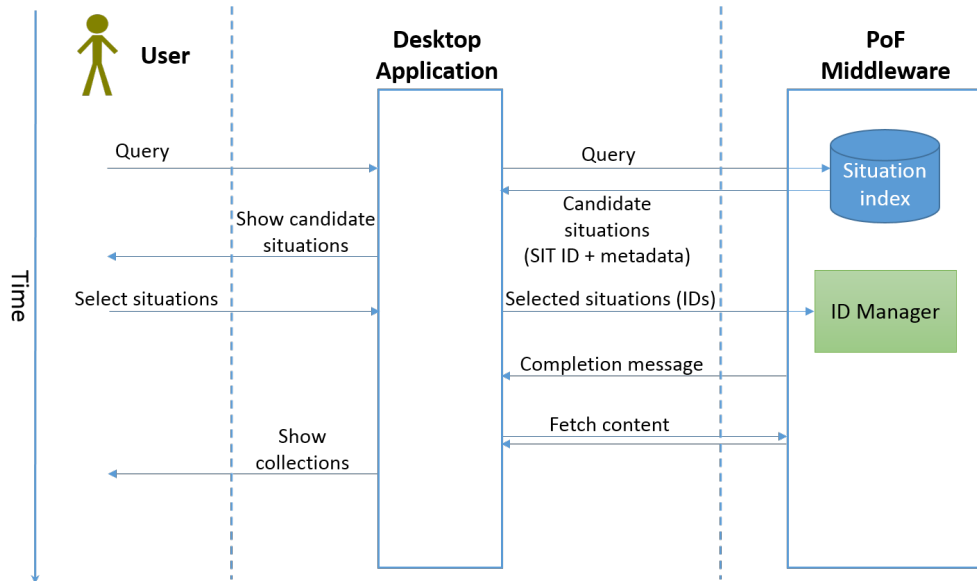
Figure 12: Workflow of the collection retrieval process.

3.2.1   Search and Fetch Operation for Collections

The user can search archives and fetch desired content from the Digital Preservation System through the PoF Middleware. As shown in Figure 13, the user can provide some input data such as location, people involved, date range and initiate a search in the archive by using the situation index in the PoF Middleware.

The retrieval of collections/situations is split into two parts: (i) performing the query and getting from the situation index only the metadata of situations satisfying the query (the retrieved metadata are shown to the user); (ii) letting the user pick one or more collections she is interested in, and retrieving their actual content. These two options depend on the actual external factors such as bandwidth availability towards the server-side. For low-quality connections, the user can be provided with metadata only, where with high-quality broadband connections visual data for content may be preferred. The default case is the former as shown in Figure 14. In that specific instance, the user has selected Collection 1 (marked with a tick) using metadata presented by the UI to retrieve that content from the system. Please note that the visual content in the upper part would be available for the latter case. The user may prefer rather a more streamlined metadata-based listing than a collection of images.

If "Enrich" is clicked in the user interface, an intermediate step takes place where additional images from external resources are presented to the user using the WP6 image contextualization service (see D6.4 [Greenwood et al., 2016]). This allows the user to contextualize her images with material from external sources and shows the capabilities of this service.

Finally, pressing "Recall" retrieves the content selected in the previous step from the preservation system and downloads it to the desktop application.
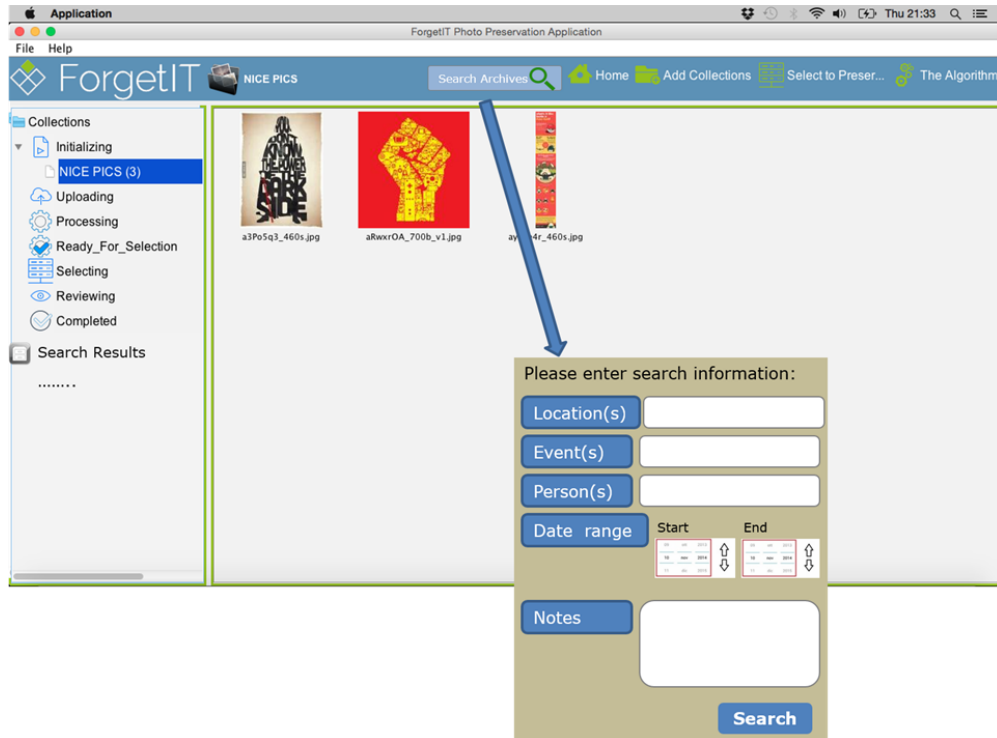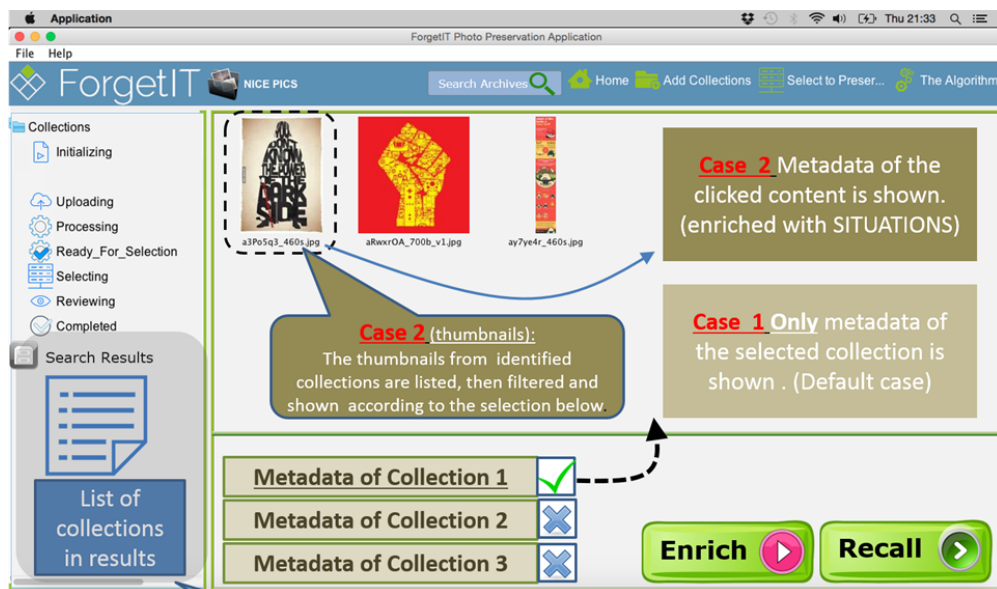
Figure 13: User interaction for searching archive in the application: user input.

- If the thumbnails are not available (**default case-Case 1**) due to system design or practical issues such as inability to retrieve file due to size, only the metadata are shown.
- All metadata are in summary format and expandable.

Figure 14: User interaction for searching the archive in the application: presentation of unguided search results (here, of Collection 1).
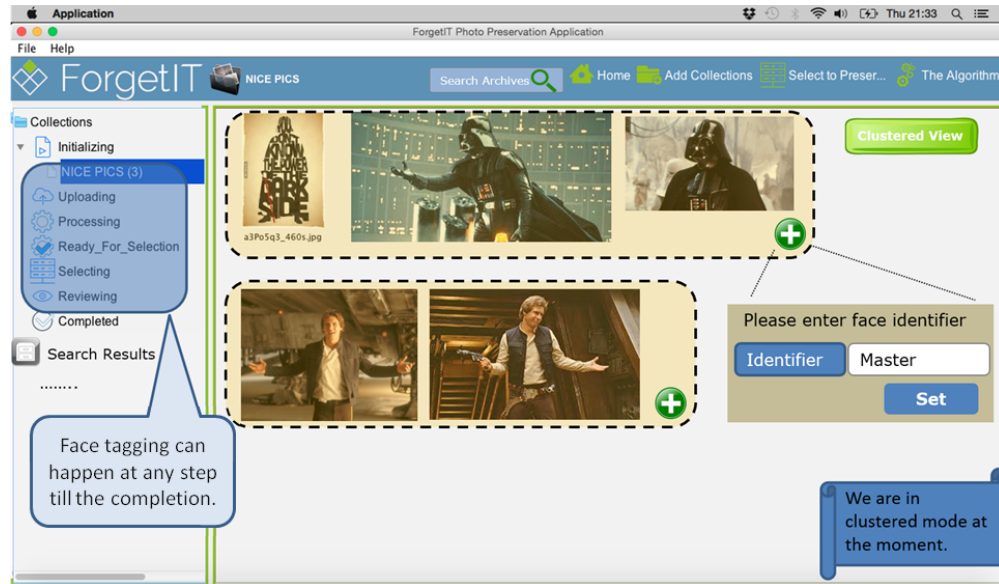
Figure 15: Face clustering during content preservation.

3.2.2 Improved Selection Capability via Face Clustering

During preservation, the photos are processed by a service in the back-end which also identifies faces using the face detection approach developed in WP4 (see D4.4 [Solachidis et al., 2016]). The photos are grouped according to the face clusters they belong to. Each cluster gets a tag by the component which serves as an identifier and also a cue for the content as shown in Figure 15. The user can also manually tag pictures with faces which were not identified by the component. It is up to the user to select which pictures to tag manually.

The application also provides some usability properties. For instance, the tags attached to face clusters by the user are permanently stored in the application, so that they can be used as suggestions for tagging face clusters in new collections. Moreover, when the user puts the cursor on the textbox for tagging (or while he starts typing), the list of known tags appear and assist the user during data entry. The user can also select photos and mark them as "having face(s)" with identifiers if they are not identified by the back-end service. Multiple selection and tag-at-once for identified faces is possible.

3.2.3 Enriched Archiving via Improved User Input

While importing a collection, the set of automatically extracted metadata (the concepts, in particular) is shown in the application interface and made editable by the user. This is facilitated by the visual concept detection service in the backend which processes the visual content(s) and provides a list of detected concepts in them. In this way, the user has the last word on what goes into the situation index used as identifier during search. Moreover, the user is asked to provide additional information that will be stored in the
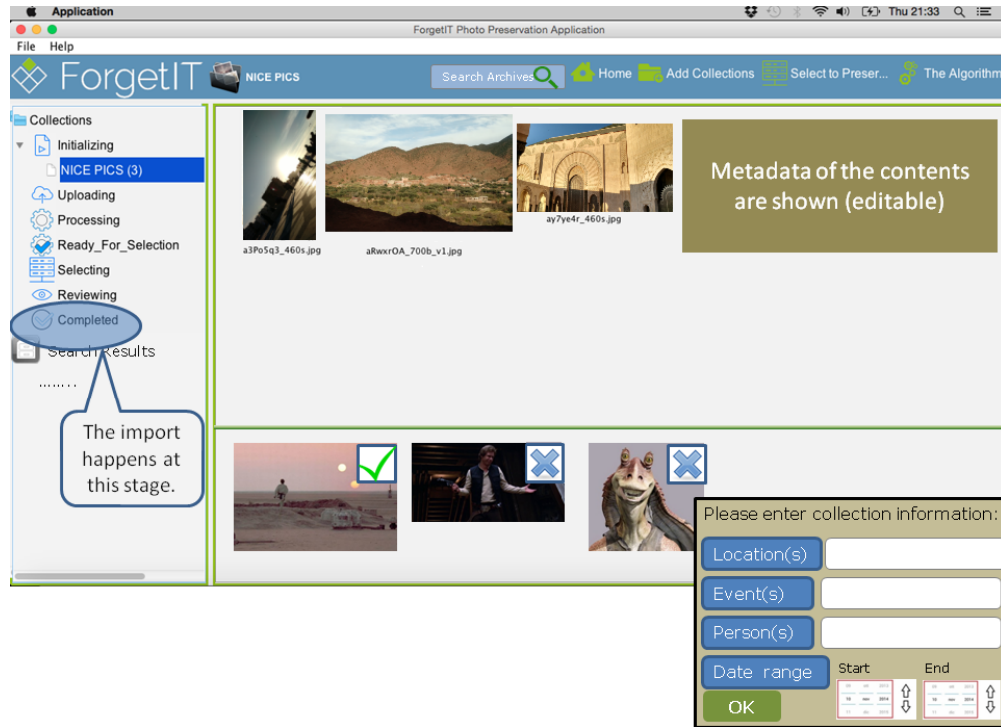
Figure 16: User data support during import of a collection.

situation index along with some information automatically extracted via analysis services as shown in Figure 16.

The application also provides some usability properties. For instance, the tags attached to face clusters by the user are permanently stored in a common file, so that they can be used as suggestions for tagging face clusters in new collections. Moreover, when the user puts the cursor on the textbox for tagging (or while he starts typing), the list of known tags appear and assist to the user for data entry. The user can also select photos and mark them as "having face(s)" with identifiers if they are not identified by the face clustering service in the PoF Middleware. Although it is not shown in the figure, multiple selection of images and tag-at-once for identified faces is possible. This feature allows the user to tag multiple images having the same face with one action.

4   Towards Using Personal Information Management to Support Personal Preservation

This section presents our approach for realising Personal Preservation using PIM in Section 4.1, addressing how WP9 overcomes the obstacles of Personal Preservation which individual users face in Section 4.2, and finally, discusses lessons learned with the approach in the DFKI PIMO installation in Section 4.3.

4.1   Approach for Personal Preservation

This section provides an overview of the steps to be taken to enable Personal Preservation for Personal Information Management (PIM) activities.

In WP9 we enabled personal preservation for users by overcoming the main obstacles for preserving (see Section 4.2). This is achieved by integrating and embracing daily PIM activities of users using the Semantic Desktop approach.

In the following, the steps to realize personal preservation are outlined for the case of using the Semantic Desktop approach. As shown with Scenario 1, dedicated applications extended with preservation capabilities are also possible. The extent of their support of personal preservation and the required effort depends on the functionality offered by the tool (such as covering only photo collections with Scenario 1 and focusing on the preservation purpose only).

The respective deliverables showing these results are D9.3 [Maus et al., 2014], D9.4 [Maus et al., 2015] (as well as the Personal Preservation Pilot online documentation[5]). Deliverables from other work packages also deal with some of these steps such as D5.4 [Nilsson et al., 2016], D8.6 [Gallo et al., 2016], and D11.4 [Akşener et al., 2015].

A prerequisite of this approach is the availability of a Semantic Desktop infrastructure providing a PIMO for the user. This could be offered by a service provider as pointed out in D11.4. As an outlook, the team at DFKI is working on a device for using the PIMO at home offering a PIMO Server and required infrastructure. The goal is to have a device with close to zero configuration for ease of use at home. The idea is similar to a home router of a telecom provider offering further services in the home network such as network storage or video streaming.

The approach for Personal Preservation has three phases: **Preparation**, **Usage**, and **Access to Preserved Material**. The steps in these phases are as follows:

---

[5]https://pimo.opendfki.de/wp-pilot/

Preparation

- Creating an account at a Semantic Desktop service: a service providing a PIMO and applications to access it (D9.3).

- If required, install apps and plug-ins for desired services, assistance, and sources for PIM (D9.3).

- Registering with a Preservation Service Provider (D9.4, D11.4)

- Arrange the Preservation Service Contract

  - selection of options which fit the user's preferences (also considering the dimensions features vs. costs) (D5.4)
  - enter the person(s) to contact in case of the user's death (D9.4, D5.4)

- Selection of Preservation Strategy & checking the consequences (D9.4)

Usage

- Usage of the Semantic Desktop infrastructure for PIM (D9.3)

- Preservation Activity

  - The Semantic Desktop will take care of computing and assigning Preservation Value Categories to things in the PIMO based on the selected Preservation Strategy (D9.4).
  - The PoF Middleware will take care to automatically preserve material according to the Preservation Service Contract (D9.4, D8.6, D5.4).
  - If desired by the user: additional manual preservation in Semantic Desktop applications (D9.3).
  - The PoF Middleware (more precise, the Context-Aware Preservation Manager of WP5) will take care for long-term maintenance such as changes in formats (D5.4).
  - During long-term usage it could be advisable to check if the chosen Preservation Strategy is still valid, e.g., if changes in life occurred.
  - Revisit and adapt Preservation Strategy: if enhancements of Semantic Desktop-enabled applications or plug-ins are installed by the user which lead to new rules or policies in the Preservation Strategies, e.g., support of quantified self hardware or applications (D9.4).

Access to Preserved Material

- The Semantic Desktop provides possibilities to access preserved material, inspect it, and reintroduce it into active usage (D9.3).

- The PoF Middleware allows to access preserved material without requiring the Semantic Desktop. Thus, users are able to access their material only using the web interfaces offered by the preservation service (D8.6).

- The Preservation Contract considers an alternative contact person which should be contacted in case the user died[6]. This person can either create an account on the PIMO Server and get access to the preserved material (D9.4) or use the PoF Middleware web interfaces of the preservation service (D8.6).

The presented approach is covered by the current Personal Preservation Pilots and the PoF Framework. However, improvements need to be undertaken for realising this as a product.

As the Scenario 1 application of Preservation of Photo Collections shows, manual preservation activities can also be supported by specialized applications, not part of the Semantic Desktop infrastructure.

4.2   Overcoming Obstacles of Personal Preservation

In deliverable D9.1 [Maus et al., 2013], we assessed the current state of Personal Preservation and identified five main obstacles which we wanted to address in WP9's approach.

In the following, those obstacles are revisited (and stated in the beginning of the respective section in gray boxes) and we discuss how these obstacles are now overcome by the solution provided by ForgetIT and realised in WP9.

4.2.1   Awareness

Users are **not aware of personal preservation** of digital content. There is a huge gap between current practices, such as backup by copying material to a different hard disk, and a proper preservation strategy.

The Personal Preservation Pilots as well as the Scenario 1 application (see Section 3)

---

[6]Detecting the death of a user was not in the scope of WP9. However, possibilities include to detect inactivity over a longer period and no response to messages by the service provider. Or, by information from the alternative contact person, or indications by other users. For instance, facebook allows friends of a user to indicate the death of a user, leading to an inactive state of the account to remember the user.

---

showed how a preservation service can be realized by using the PoF Framework. Although both use the PoF Middleware, the respective approach is different.

First, the Scenario 1 photo preservation application is a single application extended with preservation abilities by connecting to the PoF Middleware. Thus, it shows how "normal" applications can be extended with a preservation functionality and benefit from services provided by the PoF Middleware.

Second, the Personal Preservation Pilots show how a complete infrastructure such as the Semantic Desktop as well as various applications can be extended with preservation functionality and by their usage, provide valuable resources and context for preservation. Moreover, they show how an application ecosystem can contribute to realize Synergetic Preservation.

Apart from implementation approaches, the pilots also show the benefits of embedding preservation in users' PIM. The PANIC group evaluation (see Section 2.1) showed the capabilities and potential of the Semantic Desktop for reaching out to personal resources. Furthermore, the Personal Preservation final evaluation in Section 2.2 showed that users actually like the extended capabilities for organizing their photo collections.

After choosing a Preservation Strategy, the Synergetic Preservation is started with no additional effort for the user apart from occasional checks if the resource selection is sufficient (which would be optimized for product-like services).

The pilots also show the ease of setting up a preservation service (see also Section 4.1). However, the prerequisite of using a Semantic Desktop-like infrastructure still holds. Applications presented in WP9 such as the SemanticFileExplorer, the Photo Organization app and Scenario 1, show that user resources can be covered to some extent with less effort and could be realized as stand-alone applications.

Nevertheless, the Semantic Desktop infrastructure is also a step towards embedding Synergetic Preservation into an operating system. Helpful capabilities such as PIM, photo collections, or tagging, are already prominently introduced, e.g., in Apple's Mac OS X operating system.

4.2.2   High up-front costs

When starting with personal preservation, the user faces **high up-front costs** in terms of time, effort, and resources, and there are **very few tools** to help users prepare material for preservation and interact with an archiving service.

The Personal Preservation Pilots show the ease of setting up a personal preservation service (see also Section 4.1).

The high up-front costs in making personal resources suitable for preservation are reduced by embedding the required functionality for preservation into users' daily PIM ac-

tivities or at least providing applications supporting selection and connection to a preservation service as shown by Scenario 1. The evaluation with the PIMO Photo Organization app showed that early contextualization is well accepted by users if the benefits are immediate (such as the textual explanation of the context of a photo or automated search capabilities).

Furthermore, D11.4 [Akşener et al., 2015] shows, that a telecom provider sees a potential business model in offering a personal preservation service to their customers. If that would be accomplished, the high up-front costs would be reduced for finding and connecting to a preservation service for end-users.

Concerning the availability of tools, the two application scenarios show that extending an Active System with preservation capabilities is possible and the connection to the PoF Framework by implementing the interfaces is manageable. Thus, both possibilities – embedding preservation services in an infrastructure (such as the Semantic Desktop) or extending a single application with preservation capabilities (such as Scenario 1 and WP10's TYPO3) – are now open to enhance tools with preservation functionality which could foster the availability in future.

4.2.3   No Personal Preservation Services

There is no **personal preservation service** for the majority of end users which supports the whole preservation process. Cloud storage alone is not preservation.

As mentioned before, D11.4 presented a service for personal preservation which could be offered by a service provider. The business domain of a telecom provider nowadays covers lots of activities (e.g., mobile internet & access at home) and resources of users (e.g., photos, e-mails, cloud storage) with potential to extend it even further (calendar, services for friends & family, see also D11.4). In this light, extending their services with preservation capability fits well in the existing business models.

The digitalisation of various daily activities, resources held in internet platforms such as social media, lead to challenges of how to access, inherit, and preserve the material after a users death. Recent lawsuits concerning access to material in internet platforms (such as material in facebook or resources connected to the AppleID) shed a light on the increasing need for services embracing these sources. However, this will include legal issues to be covered when extending the preservation service. Furthermore, we can predict that the public awareness will increase for such services also including all digital resources. This would be a good market entry for a personal preservation service.

Finally, apart from telecom providers, ForgetIT shows that with the PoF Framework various flavours of Digital Preservation Systems can be supported. Therefore, we see the possibility also for a take-up by specialized service providers offering PoF Middleware as a service.

4.2.4   Increasing Amount of Digital Content

> The **vast increase in digital content with relevance to a person's life** poses challenges to Personal Information Management as well as preservation.

The variety of applications in the Semantic Desktop infrastructure and its openness to integrate further applications, plug-ins, and information sources allows to include various digital content in the user's PIM. WP9 showed by using technologies from ForgetIT work packages for analysing digital media that applications could support users in handling the large amount of digital data. For instance, visual concept detection allows to search large photo collections without any effort for manual annotation. Research presented in WP4 and WP6 show further technologies to manage large media content such as photo and video analysis and classification, face detection, contextualization of text, images, and video (see deliverables D4.4 [Solachidis et al., 2016] and D6.4 [Greenwood et al., 2016], respectively). WP3 presented with the RememberMe application a step towards reaching out to social media (facebook) and harvesting digital content (see D3.3 [Kanhabua et al., 2015]). Connecting to the PoF Middleware would then allow for preservation of the social media content.

4.2.5   Understanding for Future Generations

> Designing and organising an archive so that its **structure can be understood in a century from now** is cognitively challenging for users.

The PIMO – as a semantic representation of the mental model of a user – allows to structure existing content of the user's PIM activities and digital content. Further extensions such as the life situations (*pimo:LifeSituation*) can cover more specialized situations and involved material.

Applications in the Semantic Desktop infrastructure take care on structuring this information for the user (using the PIMO) without requiring her to model this by herself. The PIMO Photo Organization app is an example for this approach. Users organize their photo collections, add texts and maybe also annotations. The app stores this in a machine understandable semantic structure which covers the photo collections. It includes information from the user as well as extracted metadata (e.g., from EXIF data) and enriched metadata from components such as visual concept detection.

These structures allow a machine understandable access to the digital content as well as a human understandable presentation. They are used in the Semantic Desktop for providing added value such as semantic search or reusing content for other services such as the PIMO Diary.

Moreover, the combination of personal knowledge in the PIMO, early contextualization with the PIMO services as well as world knowledge (e.g., entity recognition and annotation

in texts using PIMO as well as external knowledge sources such as dbpedia in Seed) provides an even richer context for the digital content.

This rich semantic structure as well as the enriched digital content builds the basis for an understanding of the structure for future generations. The structure is transferred to the PoF Middleware in the *Local Context* for each and every resource covering semantic information of the resource itself as well as relations to other resources and concepts.

Technically, the PIMO is represented using RDF/S, standardized by the W3C. The *Local Context* handed over to the PoF Middleware is a formalized textual representation (see D9.4). This ensures readability and accessibility for long-term preservation.

Finally, the ForgetIT PoF Framework retrieves the content and adds its own contextualization for interpretation by future generations.

To summarize, the Semantic Desktop approach and the PIMO are the means to reduce the effort for users to generate structures. The combination of the *Local Context* – generated from the semantic information in the PIMO – and the additional contextualization in the PoF Middleware provides the means for the understanding of the content for future generations.

4.3   Lessons Learned from the DFKI PIMO

The following sections address lessons learned from experiences using ForgetIT technology with the DFKI PIMO. Before starting, let us have a closer look at the DFKI PIMO to get a better impression and this PIMO.

The DFKI PIMO

Starting with a core team of 3 users in March 2012, the DFKI PIMO steadily increased its number of users to over 20 (including temporary students and colleagues already left DFKI) with a permanent user group of 8. By using the Semantic Desktop in daily work, the DFKI team follows the paradigm of "eat-your-own-dogfood". Therefore, the DFKI PIMO and its Semantic Desktop infrastructure is now in everyday use for over 4 years. This allowed to deploy, test, and finally, use functionalities and services developed in ForgetIT with the DFKI PIMO in practical use. We continued this by deploying the PoF Middleware at DFKI premises and preparing to continuously preserve resources from the DFKI PIMO.

Figure 17 shows the total number of things and their respective types available in the DFKI PIMO. Several research and industry projects are covered by the DFKI PIMO. Several screenshots in the WP9 deliverables and videos in the online pilot documentation show real data from the DFKI PIMO.

16.843 Things	
11092	Media (Docs, Webpages, Images)
554	Topics
2091	Events
750	Agents (Persons, Groups, Orga.)
1541	Jobs (Tasks, Projects)
227	Locations
85	LifeSituations
270	Tools & Software
228	Other

Figure 17: Number and Types of Things available in the DFKI PIMO (as of February 2016)

4.3.1 Using GATE for PIMO text services

In cooperation with USFD, a PIMO Server uses GATE (General Architecture for Text Engineering) as a module for its Ontology-based Information Extraction (OBIE) service, exploiting the PIMO as vocabulary for entity recognition. It therefore utilizes GATE's *Gazetteer* functionality which has been slightly modified for the domain and integrated in a way that iterative updates are passed to its underlying finite state machine as soon as changes occur in the PIMO (i.e., creation, adding of an alternative label, label changes, merging, and deletion).

The OBIE service is used throughout the Semantic Desktop infrastructure for getting proposals of PIMO concepts for textual content. This is used in the FireTag extensions for Mozilla Firefox (for web pages), Mozilla Thunderbird (for e-mails), and the SemanticFileExplorer (for files on the file system). For web pages, we use a GATE function to extract only the relevant content of a web page[7].

These extensions now benefit from a more robust, more precise, and faster recognition. Besides the usual one-time view of a text document and getting proposals, the PIMO Server now also offers bulk upload of objects which are then analysed using the OBIE service.

Thus, we see the benefits of the GATE infrastructure and already started to use it for other purposes.

---

[7]By using GATE's `gate.creole.boilerpipe.BoilerPipe`; this allows to filter out decoration of the main content such as header, footer, and menus.

4.3.2   CERTH media technology

In the course of realising Pilot I with the PIMO Photo Organization app, we started to use the Photo Organization app with the visual concept detection from CERTH also in the DFKI PIMO. For reasons of privacy and performance, CERTH deployed their services on a dedicated virtual machine at DFKI premises. Starting with the visual concept detection, the virtual machine was extended over time with new services such as near duplicate detection.

The initial visual concepts from CERTH were filtered for the use cases of the DFKI PIMO which has business (pictures of whiteboards) as well as recreational (sightseeing) photo collections (which, finally, were also used in the WP9 evaluation). In addition, CERTH did a dedicated training of those visual concepts we have chosen for the DFKI PIMO which are now part of the service.

The use of these automatically detected and annotated visual concepts showed us the potential of including photo collections as a use case in the PIMO and the applicability of the visual concepts detected by CERTH services. We are planning to extend the development for photo collections in future work.

4.3.3   Seed as text editor in PIMO5

The iterative development of Seed in WP4 and constant deployment and tests in the Semantic Desktop allowed us to include Seed as an editor for text in the PIMO and early after introduction, already use it as the only editor in PIMO5.

Besides the really useful possibility to use the PIMO as rich and simple textbook (with notes, todo-lists, protocols, research, texts with colleagues, . . .), the most helpful functionality of Seed is the automated annotation of already existing PIMO concepts while writing. Thus, finally, allowing to easily find things later by using facetted search.

Furthermore, the entity recognition using also external sources such as freebase and db-pedia, allows for an explanation of entities in text. This hints at the potential for proactive information delivery for knowledge workers using external or more important – organisational sources such as a product catalogue. We will continue this research thread which we call "Semantic Writing" (see also [Eldesouky et al., 2015]).

4.3.4   The main ForgetIT goals

By deploying and using functionalities developed in ForgetIT also in the live DFKI PIMO, we gained insights on the applicability of the main ForgetIT goals.

Managed Forgetting

By extending the DFKI PIMO Server with a Memory Buoyancy (MB) calculation based on the work in WP3 (see D3.4 [Zhu et al., 2016] and our joint paper [Tuan et al., 2016]) and the forgetting extensions in the PIMO5 user interface, we already use part of the Managed Forgetting functionality in daily work (as of now, over 2 years).

We can see benefits for coping with the information overload originating from daily usage of the PIMO as lots of things are created with decreasing relevance on the long run such as tasks, events, web pages, or notes. Such once relevant things are now forgotten over time and are hidden from direct view during browsing and search. Furthermore, due to the MB's semantic graph, although not accessed directly, things do not drop that much in MB due to their relation to high buoyant concepts preventing them from being forgotten.

While we still need more fine tuning of the algorithm and experience with the MB for the DFKI PIMO, we see that Managed Forgetting already provides benefits for the knowledge work and still offers more potential.

On the other hand, we also see further research required considering user behavior and their expectations. For instance, how to predict and cope with the circumstances in the user interface that sometimes users want to access exactly those things long forgotten.

Our first attempt to cope with this added a slider in the PIMO5 search interface to manipulate the MB threshold used in search (current default 0.3). Things below this threshold are not shown in the result list (unless it is an exact label match). Now, an often observed behavior is to use the slider to set the threshold to zero (i.e., show everything) if the results are not satisfying, instead of modifying the query; whereas with lots of results, the slider is usually ignored. That means, there is still not enough trust in the algorithm if the results are not as expected.

The PIMO allows us to consider further research using the information from the semantic graph. For instance, although we hide usually things being forgotten, this could be changed for certain types using specific heuristics for a domain. For instance, partial matches in labels could be shown if the thing is of type "Topic" but ignored if it is of type "WebPage". Moreover, further research is required to additionally take into account the current context of the user for deciding to include forgotten things in search results (e.g., of a task).

We see lots of potential in the Managed Forgetting approach to tackle challenges of information overload in information systems especially when dealing with an evolutionary approach as supported by the Semantic Desktop. Therefore, we did a joint project proposal in the German Research Council (DFG) program on "Intentional Forgetting in Organizations"[8] together with LUH where we hope to continue developing the Managed Forgetting approach in the Semantic Desktop.

---

[8]Details (in German): `http://www.aow.ruhr-uni-bochum.de/spp1921/index.html.de`

Contextual Remembering

The PIMO already allowed to see resources in their past context by inspecting explicit relations to other things – usually entered by the user.

With the introduction of the PIMO Diary in Pilot I, we could show the potential for Contextual Remembering by using the rich material in the PIMO including semantic representation as well as content (textual or visual).

Personal insights from the diary give an interesting timeline. We found that the PIMO Diary is also enjoyable due to its capabilities to identify and present clustered activities of the user's "electronic footprint" in the PIMO.

Although focused on the personal application scenario, by "playing around" with it in the DFKI PIMO we see lots of potential for usage in Knowledge Management applications. For instance, it can serve as a kind of reporting tool giving an overview of one's own activities in a period of time or it can be used as a starting point for explorative search, especially for new colleagues to get an overview of a group's activities.

Synergetic Preservation

In Pilot II (D9.4), we introduced a Preservation Strategy initially as a predefined set of rules and policies.

Taking the advantage of actually having a live PIMO as a real data set for an organisational application scenario, we introduced a more fine grained Preservation Strategy with several heuristics and rules which could be applied to the DFKI scenario (see also the explanations in D3.4 [Zhu et al., 2016]).

Especially the possibility of changing the strategy, re-calculating, and inspecting the results provides interesting insight into the data set. The numbers in Figure 17 show lots of material accumulated over the years for this group PIMO. Besides the photo collections, particularly interesting from organisational point of view are other items proposed for preservation such as documents, presentations, notes, or topics. Notwithstanding, there are still resources assumed not to be worthwhile to be preserved.

Enhancing the Preservation Pilots with group functionality (in the personal use case these are family & friends) provides a group or a team at work also with a means for organizational preservation on a group level. Considering the sharing activity in the DFKI PIMO, we see that nearly approx. 47% of the things are public for the group. Thus, the PIMO contains private material as well as material considered relevant for the group. Thus, at least the shared resources would be subject to a preservation for the group.

With lots of professional businesses based on knowledge work such as product development, research, or consulting, companies get the opportunity to acquire and preserve valuable knowledge items from team interaction without requiring effort to document and preserve manually.

We can see this as contribution from the Semantic Desktop approach to an Organizational Application Scenario. By using the Semantic Desktop approach with groups having their individual PIMO and a group PIMO (so-called "GIMO"), we will have the opportunity to include preservation as a service for Knowledge Management projects.

We see in this Preservation Value Assessment also a first step towards Information Value Assessment for Knowledge Management purposes using new evidences for calculating content value provided by the Semantic Desktop infrastructure and the PIMO. This will be a research path we will follow.

Finally, we are looking forward to gaining experience from the PoF Middleware installed at DFKI premises including DSpace as a Digital Preservation System.

5 Review Recommendations

In the following, we want to present our reactions to the reviewer's recommendations concerning WP9 following the two project reviews.

5.1 First Year

In ForgetIT's review report of the first year, the reviewers listed as part of the "recommendations concerning future work" the following recommendation:

> 11. The PIMO approach and the use of a personal semantic infrastructure appear to be very promising, in particular with respect to a concretisation of a context implementation. The crucial point in this respect is the involvement of users in annotation, which might be supplemented by automated annotations to realise a semi-automatic handling. This requires the development of mechanisms to increase the user acceptance. The project is encouraged to address these relevant questions and evaluate the success of user involvement.

Furthermore, this point was stressed again in the reviewers' WP9 assessment:

> In contrast to the organisational context the personal use of the framework raises the question to what degree users can be motivated to provide the necessary annotation or to what extent annotation can be automated to ensure effective preservation?

To motivate users and to reduce the user (modeling) effort is an inherent challenge of the Semantic Desktop approach. Therefore, the following three functionalities are an integral part of the approach:

First, the Semantic Desktop approach strives to integrate those information sources which are already available for users such as calendars or address books. This reduces the user's need for duplicating information and provides insights on the user's context for the PIMO.

Furthermore, it is not expected that (most) users will change their long-term standard applications, e.g., their favourite e-mail tool, in exchange to a novel (maybe also with reduced functionality) Semantic Desktop application. Here, providing plug-ins providing PIMO functionality for such standard applications addresses this issue.

Finally, from our experience from knowledge management research (see, e.g., [Lampasona et al., 2012]), we also see for the personal application scenario the requirement to embed the approach into the user's daily life or activities. Moreover, added-value services providing immediate benefits for the users are required.

The PIMO Photo Organization (and PIMORE in D9.2) is such an application which provides immediate benefits to the user when organizing photos, such as neat condensed views of life situations, visual concept detection, and semantically enriched texts for a better photo organization as well as an enhanced search experience.

In this sense, the following functionalities also support users and reduce the effort of manual annotation. The functionalities were either brought into the project or developed as enhancements during the project.

- Supporting PIM activities: notes, tasks, (cloud) files, reminders, calendar, contacts.

- Embracing existing sources: calendar, address book with standard protocols (Cal-DAV, CardDAV) to access, analyse, provide, and reuse this information. This also allows to be connected to PIM applications such as Mozilla Thunderbird calendar and address book. Furthermore, these protocols also allow to connect mobile devices. For instance, on the iPhone, this PIMO information is embedded in the standard iOS apps Reminders, Calendar, Contacts, and Tasks[9].

- Special-purpose applications: PIMO Photo Organization, PIMO Diary, Task Management developed for the personal preservation scenario.

- PIMOCloud Service proving cloud storage embedded in the PIMO allowing to access (desktop) files on any device.

- Use of geo-location and metadata from photos in the PIMO.

Finally, the following services were implemented for reducing annotation effort:

- Automated tagging while writing texts with Seed using entity recognition from PIMO and proposals of new concepts from external sources (currently, freebase, dbpedia, and wikidata).

- Annotation proposals with FireTag and SemanticFileExplorer using ForgetIT text services, i.e., General Architecture for Text Engineering (GATE).

- Visual concept detection & near duplicate recognition of photos using ForgetIT services.

The following extensions were added to the PIMO Ontology. These properties support to reduce the annotation effort. In the following, these properties are listed with their respective inverse:

---

[9]see for these apps `http://www.apple.com/ios/what-is/`

pimo:hasSuggestedTopic/pimo:isSuggestedTopicOf Various plug-ins already allowed to annotate documents, e-mails, web pages, or files. They made also initial proposals if entities were found in the text. Once a user accepted a proposed thing to be annotated, the corresponding thing was explicitly added using `pimo:hasTopic`. However, those proposals which were not manually accepted by the user were lost. For reducing the annotation effort, we decided to change this: now all proposals by the entity recognition are added using `pimo:hasSuggestedTopic` – if the thing to be annotated is just created or already in the PIMO.

The main achievement is that users don't need to explicitly annotate everything for a large document, e-mail, or web page which could be annoying at some point, e.g., for large documents. The downside of this is that entities might get annotated although the user wouldn't have annotated them in the first place (due to wrong meaning or not relevant for the overall text). Therefore, we introduced this additional property to indicate that it is an automated suggestion and the user did not approve the suggestion yet (i.e., did not approve the occurrence in the text). The enriched context of resources comes in handy in semantic search, e.g., when searching for documents and having the opportunity to also filter with suggested things.

The optimization of this problem will be future work.

pimo:hasAggregatedTopic/pimo:isAggregatedTopicOf In a photo collection, some photos are worthwhile to be annotated. However, it is cumbersome to repeat all the annotations to describe the whole collection (i.e., also additional effort). Therefore, this property is used for a `pimo:LifeSituation` which allows to inherit the things annotated to the images in its collection.

pimo:hasInferredTopic/pimo:isInferredTopicOf Previously this property was only used for inheriting things down the hierarchy of a task-tree. Extending this to `pimo:LifeSituation` now allows to inherit its topics to the images in its collection. Thus, images get a richer context by the life situation they are embedded in.

pimo:wikiTextMentions/pimo:isMentionedInWikiText Used for representing entities detected and annotated while writing a text using Seed. This is used for early contextualization as well as a means to use semantic search also for written texts.

A final thought on the effort required for having a context for preservation purposes. The ForgetIT approach uses context information from the Active System as Local Context. However, the Contextualization realized in WP6 does not rely on getting context from the Active System. On the contrary, it adds World Context information independently and automatically when it comes to preservation. Thus, the Contextualizer allows to enrich resources with World Context even if poor or no context information from the Active System would be available.

5.2   Second Year

In ForgetIT's second year review report, the reviewers gave the following recommendations wrt. WP9 (again, listed in gray boxes).

5.2.1   Social Capacity

> It would be good to follow the guidance arising from the evaluation of WP9 done by the PANIC cohort and add "social capacity" to the Semantic Desktop.

In the course of the last years, the PIMO has developed from an individual knowledge base to a group knowledge base. Additionally, requirements arising from daily work in the DFKI PIMO as well as insights from the PANIC formative evaluation resulted in further functionality supporting "social capacity":

- Notifications on other users' activities on the PIMO5 Home Screen (e.g., sharing, someone contributes to my shared or subscribed things).

- Explicitly sharing with others and notifying them on the PIMO5 Home Screen as well as directly in the Windows system tray on the desktop.

- Augment a user's diary with resources from the group: as an option, the PIMO Diary can include shared material in diary generation (see D9.3[10]) but still keeping the focus on the user.

- Early chat functionality: talk about things.

- Connecting 3rd party instant messenger allowing to send things & comments to a group channel.

In the following, we present those functionalities which were not yet presented in other deliverables.

Activity of the Group

The PANIC group lacked information on the group's activity as well as the possibility of notifying others on their finished results. The entry page in the beginning of the PANIC formative evaluation was the search page listing those things recently touched by the user. Things shared by others were not shown unless explicitly searched or navigated to. Therefore, we extended PIMO5 with a Home Screen listing important things such as stacks for tasks and things. To foster the visibility of the group's activities, the Home Screen got three extensions for raising social interaction (compare with Figure 18):

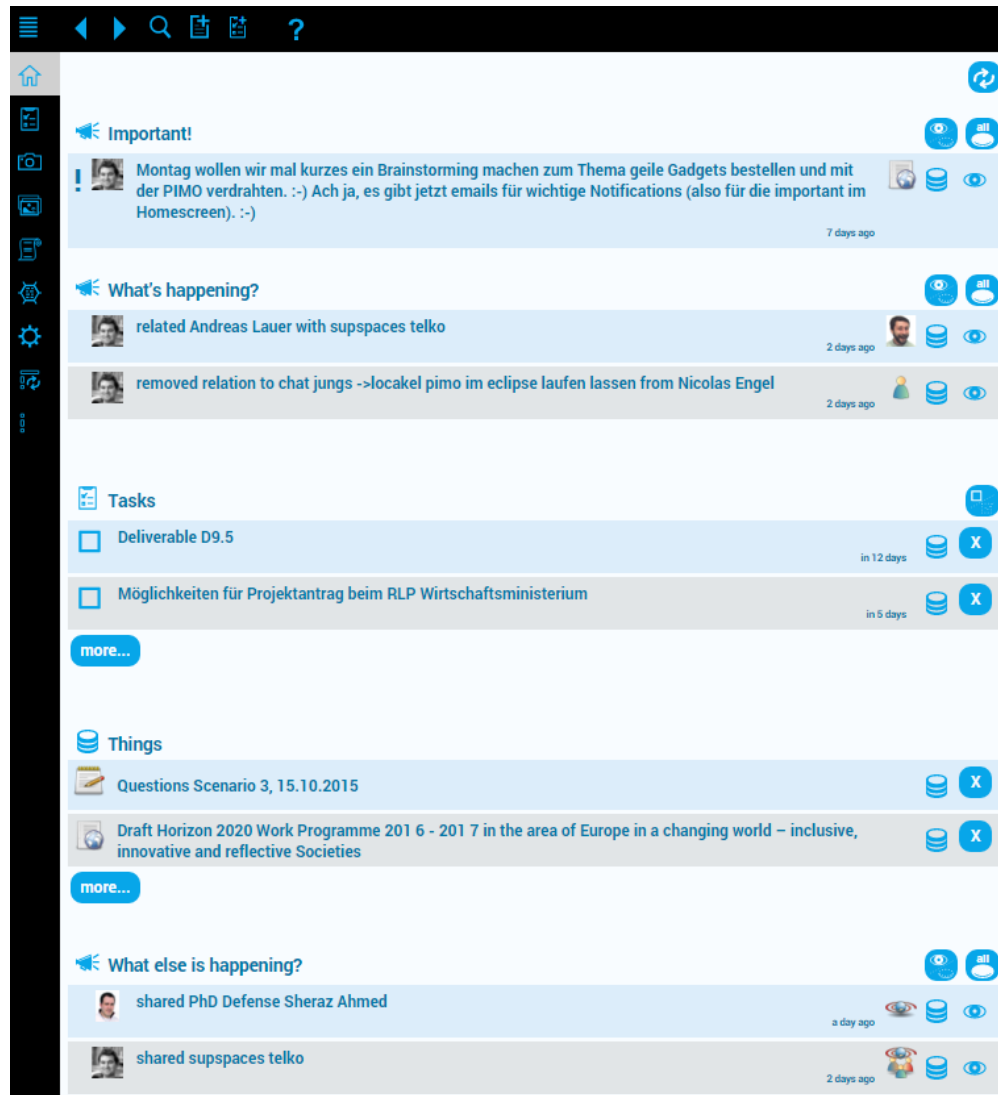[10]or the online pilot documentation **PD.8.1.7.** Group diary

Figure 18: PIMO5 Home Screen with notifications and important stacks.

Important! Here, direct user messages and important notifications of the system (e.g., preservation status, finish of image analysis of a photo collection) are listed.

What's happening? Changes or contributions of other users to things which the user either shared, subscribed to, or is involved in (as in an event).

What else is happening? Here, usually sharing actions from other users are listed. In order not to overwhelm users with notifications on a heavily used PIMO, only few are listed and if ignored will vanish after a week from the notification stack. Furthermore, this section is listed on the bottom of the Home Screen.

Clicking on a notification leads to the corresponding thing (which was shared, modified, or the user talked about).

In the following, we give some examples of the message possibilities newly introduced.
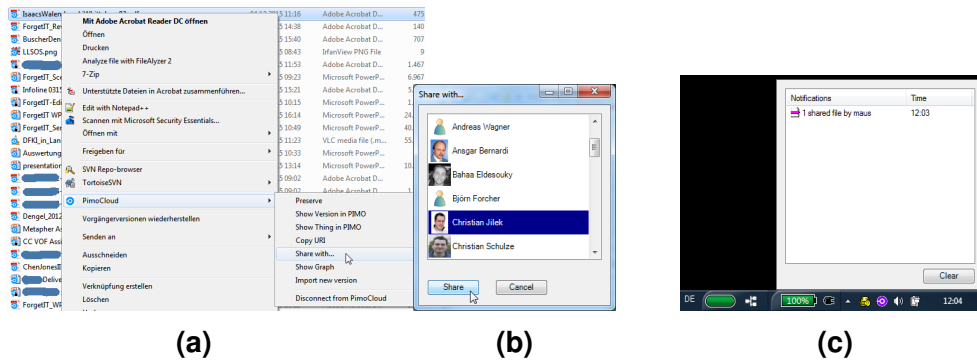
(a)   **(b)**   **(c)**

Figure 19: A file's context menu allows to share with a PIMO user (a) to be selected from a
            list (b); here, with Christian. The PIMOCloud Tray Icon will change its (usually
            blue) color and get a balloon message. Furthermore, users can inspect their pre-
            vious share notifications (c). These notifications are also shown on the PIMO5
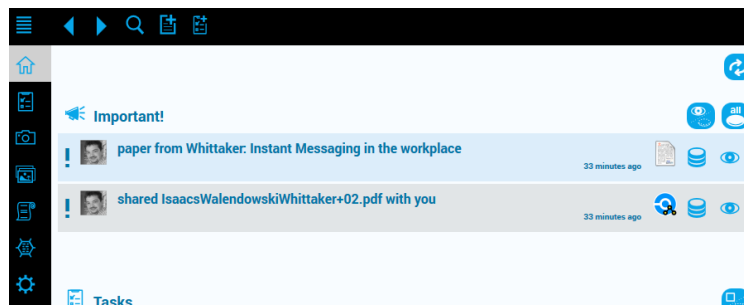   Home Screen (see Figure 20).

Figure 20: Two notifications on the PIMO5 Home Screen (of the user Christian): The lower
            one is the notification about the file sharing in Figure 19. The upper one, is a
            direct message from another user about a document which was issued in Fig-
            ure 22. Clicking on a notification opens the thing in PIMO5 directly.

Sharing of Files

As reaction to the PANIC group's desire for more immediate possibilities to notify others
while cooperating with them, we introduced the "share with" functionality in the PIMO-
Cloud context menu: a file can be shared with one or more users who will get notifica-
tions about that activity. Figure 19 shows this using the PIMOCloud Service and also a
notification via the tray icon, allowing to download and access the file immediately.

Chatting about things

Further possibilities were introduced to foster communication in the Semantic Desktop.
Now it is possible to chat about every thing in the PIMO as shown in Figures 22 and 23.
Those chats are stored and listed in a thing's chat tab for later reference. These chats
also give more information on things or their usage. These chats will also be subject to
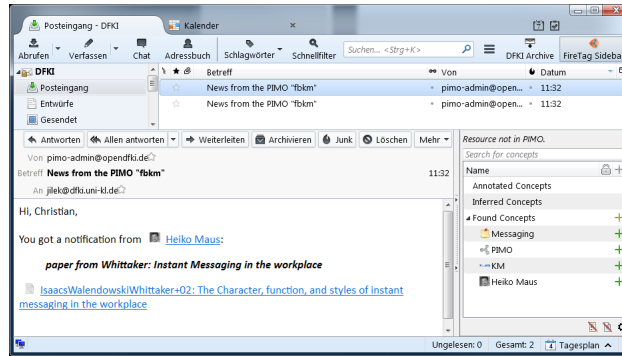preservation as part of a thing's context.

Figure 21: The Semantic Desktop also sends direct messages between users as e-mails.

Finally, the Semantic Desktop also sends direct messages to a user as e-mail as shown in Figure 21.

Instant messenger integration

As a first step towards realizing a full-fledged PIMO-embedded chat client, we integrated the popular group messenger slack[11] using their slack bot API[12] for external access. Therefore, we allow a user to push messages to a slack installation (by selecting the "slack integration", see Figure 22) which lists the notifications in a dedicated group channel (see Figure 24). By clicking on things listed in the messenger, they can be opened in PIMO5, or – if the URL is external such as a web page – directly accessed.

However, the slack integration is limited as slack is a closed source system. Therefore, we are working on a tighter integration with the open source group chat client RocketChat[14] which will allow to keep track of all chats and use them in the PIMO.

---

[11]https://www.slack.com/

[12]https://api.slack.com/bot-users

[14]https://rocket.chat/

[14]If the resource would be a web page, then its URL is listed which allows open the web page directly.
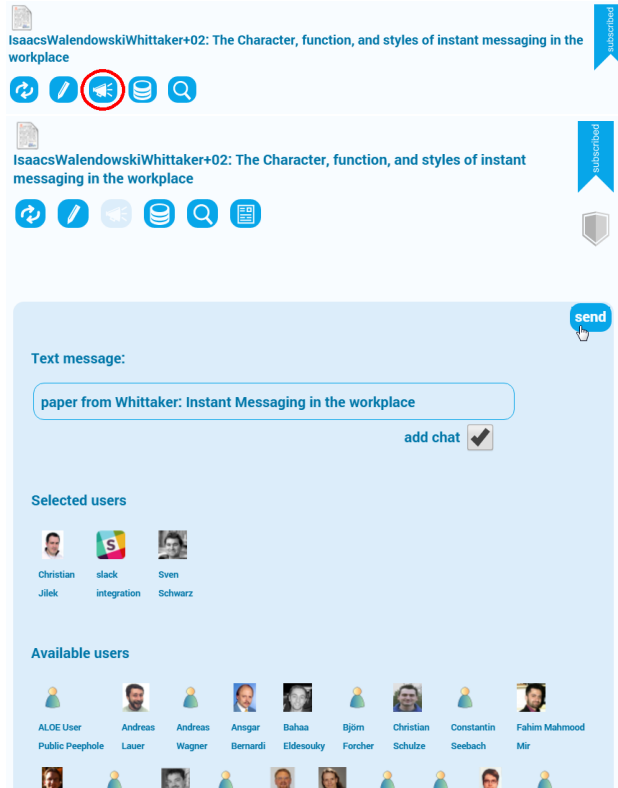
Figure 22: Each thing in PIMO5 has a chat button (red circle) allowing to send messages about this thing to other users. The users can be selected by clicking on their icons. Selecting the slack Instant Messenger icon sends this message to the PIMO channel in the slack installation of the team as shown in Figure 24. If the "add chat" option is enabled (which is enabled as default), the text is added and listed as chat entry in the thing's chat tab (see Figure 23).
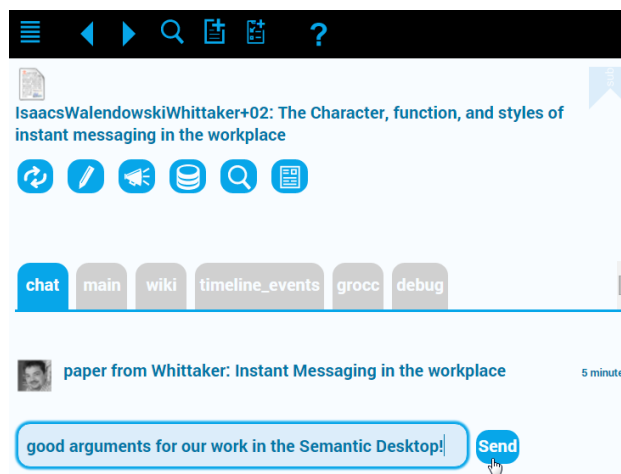

Figure 23: The notification is also listed in the chat tab. Users have the possibility to start chatting about a thing. These chats are stored and listed with a thing.
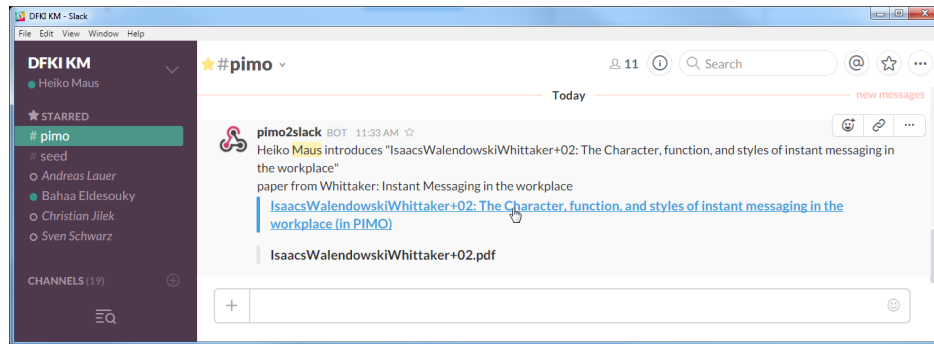
Figure 24: The message from Figure 22 is also sent to the slack instant messenger via the pimo2slack-bot. The corresponding thing is listed as hyperlink allowing the user to directly open it in PIMO5 (here, the document as pdf[13]).

5.2.2 Preservation Strategy

> While the semantic handling is well supported by the Semantic Desktop it is not so obvious how users handle the preservation strategy or change their preservation policy, respectively, since it has been mentioned that users organise their information objects according to their personal style. This could be considered in more detail in the final year of the project.

Pilot II introduced the Preservation Service Contract allowing users to set and change their Preservation Strategy including preservation levels for various Preservation Value Categories (D9.4). They can use a pre-defined Preservation Strategy (based on four personas) and, if desired, change the pre-set policies and rules to match their own preferences. The evaluation showed that the four personas were a good starting point as the participants overall agreed with the resources selected for preservation.

This is a start, further or different rules and policies can be introduced, the handling, selection, and usage in the PV Assessment will remain the same.

5.2.3 Evaluation addressed in D9.3 assessment

> [. . .] However, to make a full evaluation of the research and development work completed under D9.3 it is essential to test it, much as the PANIC Student group did in the Formative evaluation phase.

WP9's final evaluation presented in section 2.2 used to a large extent the Pilot documented in deliverables D9.3 and D9.4. Focus was on the PIMO Photo Organization application and the automated preservation decisions based on personas and Preservation Value Assessment of user interactions with the PIMO.

However, evaluating the full range of the presented Pilot I & II functionalities would have

required participants using it over a long period of time including a similar engagement as the PANIC group showed. The remaining time frame did not permit such as long-term study at the end of the project. However, in Section 4.3 we present the insights we got at the DFKI during the usage of Pilot features over the three years of ForgetIT project runtime.

> We do not get any indication as to how intensively the students used the system during this period and whether their impressions were influenced as much by their information practices and their attitude towards preservation more generally as by their experiences using the system itself.

The evaluation report on the results of the PANIC group now addresses these indications (see Section 2.1).

6   Conclusions

This last deliverable concludes the work on the Personal Preservation application scenario done in WP9.

The work package's first deliverable D9.1 [Maus et al., 2013], identified the obstacles for personal preservation, defined the application scenario and derived requirements for adequate uses cases covering the intended results from the research done in ForgetIT.

Taking these requirements, deliverable D9.2 [Maus and Schwarz, 2014] then developed those use cases and presented mock-ups and early prototypes for the personal preservation scenario addressing the use cases. In this Year 1 deliverable, we could already show an integration with the first version of the PoF Middleware allowing manual preservation. Thus, the application scenario was also a driver in the project towards a coherent view on preservation, requirements towards ForgetIT technology, and system integration.

In Year 2 with D9.3 [Maus et al., 2014], the first Personal Preservation Pilot showed how the Semantic Desktop approach for PIM contributes to the ForgetIT goals Managed Forgetting, Contextual Remembering, and – first steps towards – Synergetic Preservation. The Pilot also kept up with the progress of the project, integrating technologies from work packages, deepening the PoF Framework integration, and showing a fully integrated preservation process.

Then D9.4 [Maus et al., 2015] – by presenting the second Personal Preservation Pilot – focused on preservation by showing how the Semantic Desktop infrastructure and the PIMO allow for an embedded Preservation Value Assessment for personal resources. By introducing the Preservation Strategy including personas as well as more fine grained rules and policies, Pilot II showed how the claim of WP9 – supporting Personal Preservation by embedding it into PIM activities of the user – can be realized. Moreover, ForgetIT and the Pilots successfully addressed nearly all use cases listed in D9.2 (see final assessment in D9.4).

The evaluation reported in this deliverable D9.5, now shows that we are on the right track. Embedding preservation technology into daily activities is feasible and – provided that the products mature – we can realize a Synergetic Preservation overcoming the obstacles identified for personal preservation.

To reach this, WP9 interacted with all work packages, gave requirements for research as well as implemented some directly in WP9 such as Managed Forgetting, and successfully integrated several technologies to realise a coherent view on personal preservation.

To finally show the success of WP9, the success indicators for WP9 are finally assessed in the following.

6.1   Assessment of WP9's Success Indicators

In the ForgetIT's Description of Work [ForgetIT, 2012], the following expected outcomes for WP9 were listed together with success/progress indicators:

Expected Outcomes

1. *A Preservation Strategy suitable for Personal Information Management using the Semantic Desktop approach*

2. *Prototypes showing the realization of the personal preservation*

3. *Evaluated best practice of preservation-oriented knowledge modelling and personal preservation*

In the following, the success of these outcomes are discussed along the success indicators. The first two were already finally assessed in D9.4 on the basis of the pilots. The third one was missing due to the required evaluation for this. For completeness, success indicators 1 & 2 are repeated here. The addition in this deliverable is to assess the evaluation results for the third success indicator.

A Preservation Strategy suitable for PIM using the Semantic Desktop approach

- *Availability of several documented use cases for personal preservation, with appropriate presentation materials*

  Deliverable D9.2 [Maus and Schwarz, 2014] listed in Section 2 relevant use cases which are addressed by the pilot. Their fulfillment was assessed in deliverable D9.3 [Maus et al., 2015] and reassessed in deliverable D9.4 [Maus et al., 2015]. We have shown that most of the use cases are successfully covered in both pilots. Furthermore, the pilot documentation demonstrates the realization of several use cases.

- *Documented Preservation Strategy suitable for users applying PIM*

  The initial Preservation Strategy realized and documented by the Personal Preservation Pilot I is as follows:

  The pilot shows how PIM is supported by the Semantic Desktop infrastructure. In doing so, various evidences are collected for the PoF to make preservation decisions. Furthermore, all resources known to the Semantic Desktop are candidates to be preserved automatically without user intervention. In addition, manually triggering preservation on resources is also possible even if the resources are not initially known to the PIMO.

  Making resources known to the Semantic Desktop is simply done by either annotating a resource, defining file folders where files are automatically introduced into the

PIMO, or using a specialized app from the Semantic Desktop infrastructure such as photo organization, calendar, task management, or note taking, thus, being part of user's PIM.

The preservation is then realized by connecting the Semantic Desktop with the PoF Middleware. This is documented in the web-based documentation of the Personal Preservation Pilot I.

Thus, users using at least some of the components offered by the Semantic Desktop infrastructure for their PIM are supported by synergetic preservation of PoF. As pointed out in this deliverable, even if a user just uses one app from the infrastructure, e.g., photo organization for their photo collections, their resources are connected to the PoF.

The Personal Preservation Pilot II extends the initial Preservation Strategy as follows:

Pilot II allows to set an individual Preservation Strategy by either selecting policies and rules along the 6 dimensions for Preservation Value Assessment resp. choosing among the four personas offered. The PIMO5 user interface allows for inspecting the consequences of the selected Preservation Strategy. Further details of a Preservation Service Contract are shown as proof-of-concept such as choosing different Preservation Levels which are send to the PoF Middleware as Preservation Broker Agreements (Extensions of this will be part of deliverable D5.4 [Nilsson et al., 2016]).

Based on the individual Preservation Strategy, the Preservation Value Assessment in the Semantic Desktop calculates the PV of each thing in the PIMO. The PoF Middleware via the Forgettor then is able decide on the set of resources to be preserved (using the selected PV Categories communicated in the Preservation Broker Agreement) and start the Preservation Preparation Workflow.

With the setting of the Preservation Strategy in the Semantic Desktop enables users to 'forget' about the preservation at all, because ForgetIT's approach of Synergetic Preservation will start to pull things of the PIMO into the archive if their PV reaches the thresholds resp. categories (gold, silver, bronze) set in the contract.

Additionally, D9.5 documents in Section 4.1 the steps required to realize the approach for Personal Preservation intended with WP9.

- *Level of contribution of Semantic Desktop for selection and preservation activities*

  Pilot I shows how user's are able to interact with resources in the Semantic Desktop. This allows to provide evidences for the PoF to decide for preservation. In deliverable D3.3 and upcoming D3.4 ([Kanhabua et al., 2015, Zhu et al., 2016]) the contributions for calculating the Preservation Value are explained. For the Memory Buoyancy, the calculation is based solely on interaction evidences and semantic model information from the PIMO. Pilot I shows several benefits which make use of the Memory Buoyancy to realize Managed Forgetting.

As mentioned above, Pilot II introduces Preservation Strategy and individual Preservation Value Assessment and thus realizes a contribution which shows the usefulness of the approach, of the infrastructure chosen, and also a proof-of-concept for adapting selection and preservation activities to the individual user's choice. Now users can continue with their PIM and resources will get preserved (according to the chosen contract) as well as manual preservation of material is still possible.

Furthermore, the PIMO provides rich context from personal knowledge for the contextualization in the PoF Framework which is usually not producible by world knowledge contextualization.

Prototypes showing the realization of the personal preservation

- *Existing prototype for a concise preserving personal desktop*

  Pilot I shows the implementation of the preservation workflow by connecting to the PoF Middleware. This enables the Semantic Desktop to preserve resources in the preservation system. Restoring preserved resources is also realized within the Semantic Desktop infrastructure.

  Several applications within the Pilot contribute to the ForgetIT goals of Managed Forgetting, Synergetic Preservation, and Contextualized Remembering as pointed out in the Pilot I documentation in D9.3.

  As mentioned above, Pilot II introduces Preservation Strategy and individual Preservation Value Assessment, therefore strengthening the achievement of Pilot I in the area of Preservation.

- *Existing prototype for a concise preserving mobile information assistant*

  In ForgetIT, the Semantic Desktop infrastructure was extended by an HTML5 interface (the so-called 'PIMO5') which is capable of running on mobile devices. Apart from the desktop clients for preserving on the desktop, PIMO5 is the main focus of development to realize the use cases for personal preservation. With this, the second pilot which focuses on mobile devices will embrace the functionality of PIMO5 in Pilot I.

  Pilot II realized with the Memory Buoyancy assessment and the Memory Buoyancy Sync Manager the synchronization of files and concepts from the PIMO which are of current relevance for the user. The benefit is to provide access to these files and concepts also if the user is off-line as well as reducing required mobile data bandwidth by saving the user the need to download the files first.

  The availability of native clients on mobile devices also enables to connect resources to the Semantic Desktop and thus, also to allow preservation by connection of the Semantic Desktop infrastructure with the SD/PoF Adapter to the PoF Framework.

  Depending on the mobile operating system, access to resources can be restricted.[15]

---

[15] For instance, on iOS access to files of other apps is not allowed, they need to be handed over explicitly by the user.

However, embracing these resources is a task for the future development of the Semantic Desktop infrastructure on mobile devices. For ForgetIT, the proof-of-concept in Pilot II shows the general realisability.

The following success indicators were not addressed in previous deliverables. Because now the evaluation results are available, finally these indicators can be assessed.

Evaluated best practice of preservation-oriented knowledge modelling and personal preservation

- *Ease of use of the personal Preservation Strategy, as shown in demonstration use cases*

  First of all, this deliverable presents in Section 4.1 the approach for enabling Synergetic Preservation for PIM activities given that the Semantic Desktop is used. The steps are few and easy to follow. Once set up, what is left is to do the usual PIM.

  The Semantic Desktop infrastructure covers activities such as task management, note-taking, and photo organization. Further activities can (and will) be supported. The formative evaluation with PANIC indicates that the Semantic Desktop has the potential to cover lots of personal resources of users and thus make them available for preservation. However, usability issues need to be addressed to realize also ease of use for the applications themselves.

  Based on this foundation, Pilot II introduces the Preservation Strategy and shows an easy set-up and selection of a personalized Preservation Strategy using two possibilities. First, simply selecting between one of four personas. These personas use predefined profiles with rules and policies adequate for the respective user behaviour. In this instance, the details are hidden from the user. The results of the WP9 evaluation show a high agreement on the resulting preservation decisions (see Section 2.2.7). Second, an extended persona profile shows the details and allows to easily switch on/off the predefined rules and policies to adapt the Preservation Strategy to personal preferences. Furthermore, the switches explain in natural language what the respective effect would be.

  Taking this into account, the Personal Preservation Pilots show the potential ease of use of a personal Preservation Strategy.

- *Low additional effort for personal preservation, as shown by experimental use of the prototypes*

  The results of the WP9 evaluation indicate a low effort for preservation (see Section 2.2.7). The evaluation was designed to let participants organize their photo collections with the help of an application (i.e., the Pilot) without preservation in mind. The additional opportunity to preserve the photo collections came as a "by-product". Therefore, for those people spending effort for organizing and further using their photo collections, the preservation would come as a complementary opportunity.

The effort left consists of the decision on the proposed material to be preserved. Here, the PIMO"time capsule" user interface provides a means to inspect the selected resources and the possibility to easily manipulate the decisions. This also helps in reducing the effort. In the evaluation, participants mostly agreed with the proposed preservation decisions based on their Preservation Strategy. This also supports an expected low effort in a final product. However, the effort spent for checking the proposed decisions will be lower compared to a fully manual selection for preservation.

The effort can even be lowered following the advise for personal preservation enabling Synergetic Preservation given in Section 4.1: setting the personal Preservation Strategy and then let the PoF Middleware take over to decide on material to be preserved. From then on, the PoF will continuously retrieve from the Semantic Desktop and ingest resources into the Digital Preservation System without user interaction. Nevertheless, a prerequisite for this is a good PV Assessment and trust in the resulting selection of material.

Finally, the inspection of the Preservation Strategy in the DFKI PIMO as pointed out in Section 4.3.4 also gives us an indication of an effortless possibility for preserving relevant material from the DFKI PIMO once connected to the PoF Framework deployed at DFKI premises.

- *User satisfaction with the identified personal Preservation Strategy, as assessed by user feedback during evaluation experiments*

  As pointed out in the conclusions of the WP9 evaluation (see Section 2.2.7) the overall satisfaction for the Pilot used in the evaluation was above average and the users were satisfied with the preservation decisions as they mostly agreed with them.

6.2   Vision for the Future

As outlined in deliverable D11.4 [Akşener et al., 2015], TT sees a potential market for a Personal Preservation Service as part of their services. The discussion on this will be continued to identify possibilities to realize such as a service. TT would also like to extend the Personal Preservation concept to other modalities (e.g., mobile client apps integrated into mobile devices provided with TT subscription contracts). Additionally, it may consider its large subscriber base for exploring a crowd-sourcing dimension regarding cooperative Preservation Value estimation which may extend the "value" concept of a content from an individual perspective to a community perspective (collective memory).

As pointed out in the lessons learned with the DFKI PIMO presented in Section 4.3, we can see that all three ForgetIT goals are beneficial to the Semantic Desktop approach, particularly, when used for evolutionary knowledge management. As the assessment of Memory Buoyancy and Preservation Value is already an integral part of a PIMO Server installation, the foundation is laid to use it in further research and industry projects.

We see lots of potential to continue research on these challenging topics in combination with the Semantic Desktop and PIMO. Particularly, the Preservation Value in combination with the Preservation Strategy allows new insights to valuable content of knowledge bases. We will continue this research towards information value assessment in evolutionary knowledge management as supported by the Semantic Desktop.

Furthermore, being able to offer preservation capabilities for knowledge management solutions is a novel service for DFKI and will be investigated further. Therefore, we will also continue to use the PoF Framework and stay in close contact with EURIX for further opportunities to continue work.

7　References

[Akşener et al., 2015] Akşener, F., Sözen, S., Niederée, C., Maus, H., and Nilsson, J. (2015). ForgetIT Deliverable D11.4: Personal Preservation as a Service.

[Dix, 2007] Dix, A. (2007). Designing for Appropriation. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...But Not As We Know It - Volume 2*, BCS-HCI '07, pages 27–30, Swinton, UK. British Computer Society.

[Eldesouky et al., 2015] Eldesouky, B., Maus, H., Schwarz, S., and Dengel, A. (2015). Seed, a Natural Language Interface to Knowledge Bases. In Yamamoto, S., editor, *Human Interface and the Management of Information. Information and Knowledge Design*, volume 9172 of *Lecture Notes in Computer Science*, pages 280–290. Springer International Publishing.

[ForgetIT, 2012] ForgetIT (2012). Annex I: Description of Work. Grant Agreement 600826, ForgetIT consortium.

[Gallo et al., 2016] Gallo, F., Ceroni, A., Tran, T., Chen, D., Andersson, I., Greenwood, M. A., Maus, H., Lauer, A., Schwarz, S., Solachidis, V., Papadopoulou, O., Apostolidis, E., Pournaras, A., Mezaris, V., and Goslar, J. (2016). ForgetIT Deliverable D8.6: The Preserve-or-Forget Framework – Final Release.

[Greenwood et al., 2016] Greenwood, M. A., Petrak, J., Gorrell, G., Solachidis, V., Papadopoulou, O., Apostolidis, K., Galanopoulos, D., Tastzoglou, D., Mezaris, V., Eldesouky, B., Tran, N. K., Hube, C., and Niederée, C. (2016). ForgetIT Deliverable D6.4: Contextualisation framework and evaluation.

[Jilek et al., 2015] Jilek, C., Maus, H., Schwarz, S., and Dengel, A. (2015). Diary Generation from Personal Information Models to Support Contextual Remembering and Reminiscence. In *Workshop on Human Memory-Inspired Multimedia Organization and Preservation (HMMP). IEEE International Conference on Multimedia and Expo (ICME-2015), June 29 - July 3, Torino, Italy*, pages 1–6. IEEE.

[Kanhabua et al., 2015] Kanhabua, N., Niederée, C., Ceroni, A., Naini, K. D., Kawase, R., Tran, T., Maus, H., and Schwarz, S. (2015). ForgetIT Deliverable D3.3: Strategies and Components for Managed Forgetting - Second Release.

[Lampasona et al., 2012] Lampasona, C., Rostanin, O., and Maus, H. (2012). Seamless Integration of Order Processing in MS Outlook using SmartOffice: an Empirical Evaluation. In *ESEM '12 Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement*. ACM Press.

[Logie et al., 2016] Logie, R., Alexandru, C., Niven, E., Rhodes, S., Rutter, J., and Mayer-Schönberger, M. W. V. (2016). ForgetIT Deliverable D2.4: Foundations of Forgetting and Remembering.

[Maus et al., 2013] Maus, H., Dobberkau, O., Wolters, M., and Niederée, C. (2013). ForgetIT Deliverable D9.1: Application Use Cases & Requirements Document.

[Maus and Schwarz, 2014] Maus, H. and Schwarz, S. (2014). ForgetIT Deliverable D9.2: Use Cases & Mock-up Development.

[Maus et al., 2014] Maus, H., Schwarz, S., Eldesouky, B., Jilek, C., Wolters, M., and Loğoğlu, B. (2014). ForgetIT Deliverable D9.3: Personal Preservation Pilot I: Concise preserving personal desktop.

[Maus et al., 2015] Maus, H., Schwarz, S., Jilek, C., and Gallo, F. (2015). ForgetIT Deliverable D9.4: Personal Preservation Pilot II: Concise preserving mobile information assistant.

[Nilsson et al., 2016] Nilsson, J., Andersson, I., and Gallo, F. (2016). ForgetIT Deliverable D5.4: Workflow model and prototype for transition between active system and AIS - Final release.

[Solachidis et al., 2016] Solachidis, V., Apostolidis, E., Markatopoulou, F., Galanopoulos, D., Tzelepis, C., Arestis-Chartampilas, S., Pournaras, A., Tastzoglou, D., Mezaris, V., Chen, D., Harnik, D., Khaitzin, E., Eldesouky, B., Maus, H., Greenwood, M., and Tan, A. S. (2016). ForgetIT Deliverable D4.4: Information analysis, consolidation and concentration techniques, and evaluation - Final release.

[Solachidis et al., 2015] Solachidis, V., Papadopoulou, O., Apostolidis, K., Ioannidou, A., Mezaris, V., Greenwood, M., Eldesouky, B., and Maus, H. (2015). ForgetIT Deliverable D4.3: Information analysis, consolidation and concentration techniques, and evaluation - Second release.

[Tuan et al., 2016] Tuan, T., Schwarz, S., Niederée, C., Maus, H., and Kanhabua, N. (2016). The Forgotten Needle in My Collections: Task-Aware Ranking of Documents in Semantic Information Space. In *Proceedings of the 1st ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR)*. ACM Press.

[Wolters et al., 2015] Wolters, M. K., Niven, E., Runardotter, M., Gallo, F., Maus, H., and Logie, R. H. (2015). Personal Photo Preservation for the Smartphone Generation. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, CHI 2015 Extended Abstracts, Republic of Korea, April 18 - 23, 2015*, pages 1549–1554.

[Zhu et al., 2016] Zhu, X., Niederée, C., Tran, T., Ceroni, A., Naini, K. D., Tran, N. K., Maus, H., and Jilek, C. (2016). ForgetIT Deliverable D3.4: Strategies and Components for Managed Forgetting - Final Release.

Glossary

CMIS Content Management Interoperability Service. 30, 31

FireTag is a PIMO-enabled sidebar plug-in for Mozilla Firefox web browser and Thunderbird e-mail client for annotating resources, i.e., web pages and e-mails. 16, 43, 49

GATE General Architecture for Text Engineering. 5, 43, 49

MB Memory Buoyancy. 45, 60, 61, 63

NASA-TLX NASA Task Load Index. 20

OBIE Ontology-based Information Extraction is a process that processes unstructured or semi-structured natural language text through a mechanism guided by ontologies to extract certain types of information and presents the output using ontologies. 43

PIM Personal Information Management. 7, 8, 28, 36, 37, 39, 41, 49, 58–62

PIMO Personal Information Model. 10–12, 14–19, 22–25, 28, 36–38, 41–56, 58–61, 63, 64

PIMO Diary is an application in the Semantic Desktop which uses the PIMO to generate a diary for various time periods w/o requiring the user to actually write one (see also [Jilek et al., 2015]). 10, 14, 41, 46, 49, 51

PIMO5 is an HTML5 client in the Semantic Desktop infrastructure allowing to browse the PIMO and containing several apps such as the PIMO Diary. PIMO5 is also available for browsers on mobile devices. 5, 10, 14–16, 18, 19, 23, 24, 44, 45, 51–56, 60, 61

PIMOCloud is a PIMO-embedded cloud service with synchronization abilities across devices, versioning, and preservation functionality. Extensions such as the SemanticFileExplorer allow to easily add files to the cloud. PIMO5 is able to access such cloud files anywhere. 49, 53

PoF Preserve-or-Forget. 7, 8, 28, 30–32, 35, 37–42, 47, 58–61, 63, 64

Preservation Strategy is a profile of rules and policies for a user which is used to calculate the Preservation Value. 7, 8, 17, 18, 26, 27, 37, 39, 46, 56, 58–64

PV Preservation Value. 18, 47, 56, 58, 60, 61, 63, 64

RDF/S is the Resource Description Framework and Schema, a knowledge representation language for the Semantic Web, standardized by the W3C. 42

Seed Semantic Editor. 5, 10, 14, 19, 42, 44, 49, 50

Semantic Desktop is an approach which introduces a knowledge representation layer (the PIMO) by semantic technologies on top of information sources on the desktop and mobile devices and opening up this to applications. 9–11, 14–18, 24, 26, 36–45, 47, 48, 51, 53, 54, 56, 58–64

SemanticFileExplorer is a PIMO-enabled plug-in for the MS Windows File Explorer. 16, 39, 43, 49

SUS System Usability Scale. 20, 22