

ForgetIT

Concise Preservation by Combining Managed Forgetting
and Contextualized Remembering

Grant Agreement No. 600826

Deliverable D10.4

Work-package	WP10: Preserve or Forget in Organizational Information Management
Deliverable	D10.4: Organizational Preservation Report
Deliverable Leader	Olivier Dobberkau
Quality Assessor	Heiko Maus
Dissemination level	PU
Delivery date in Annex I	31 January 2016
Actual delivery date	31 March 2016
Revisions	1
Status	Submitted
Keywords	TYPO3, CMS, Semantic Annotations, Organizational Preservation, Concept Detection, Learnings SME, Exploitation Ideas

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.

© 2015 Participants in the ForgetIT Project

Revision History

Date	Version	Major changes	Authors
1st Dec	0.1	Initial Draft	Dobberkau, Gsedl
21st Dec	0.2	Added Details	Dobberkau, Gsedl
7th Jan	0.3	Conversion to LaTeX	Goslar
25th Jan	0.4	Updated Structure	Goslar
Early February	0.5	Large Revisions and Additional Content	Goslar
Late February	0.6	Lessons Learned	Gsedl
Early March	0.7	Added Evaluation Chapter	Wolters
Mid March	0.8	Internal QA comments	Niederee
Mid March	0.9	Revised Evaluation Chapter	Wolters, Gsedl
End March	1.0	Quality assessment remarks	Dobberkau, Goslar, Maus

List of Authors

Partner Acronym	Authors
dkd	Olivier Dobberkau, Johannes Goslar, Ingo Gsedl
University of Edinburgh	Maria Wolters

Table of Contents

Executive Summary	6
1 Introduction	7
1.1 Take up of Project Results	7
1.2 Target Audience	9
1.3 Structure of the Deliverable	9
2 ForgetIT Pilot Application	11
2.1 Synergetic Preservation for Content Management	12
2.1.1 CMIS and Synergies	12
2.1.2 Preservation Support for Press Releases and News	15
2.2 Information Value in Content Management	15
2.2.1 Content Management and Value Creation	16
2.2.2 Advancing Digital Asset Management	16
2.3 Content in Context	19
2.3.1 Semantic Annotation	19
2.3.2 Semantic Annotation and TYPO3	19
2.3.3 Contextualization for Press Releases and News Management	23
2.3.4 Concept Detection in Pictures	23
2.4 Software Component Overview	24
3 Leveraging ForgetIT Results	28
4 Involvement of Free Software Communities	31
4.1 Previous Experience	31
4.2 Bringing ForgetIT to the TYPO3 Open Source Community	32
5 Evaluation	34
5.1 Participants	34

5.2	Usability Ratings	35
5.3	Summary of Observations	35
5.3.1	Technical Problems	36
5.3.2	Manual Annotation	37
5.3.3	MIMIR Search	37
5.4	Suggestions for Improvement	38
5.5	Next Steps	39
6	Lessons Learned	40
7	Exploitation Plans	42
7.1	Technology-related Exploitation	42
7.1.1	PHP & TYPO3	42
7.1.2	Docker Files	44
7.1.3	GATE	44
7.1.4	Possible software solutions based around the technologies and use cases elaborated in the ForgetIT project	45
7.1.5	Business Canvases	45
7.2	Consulting towards an organizational digital data strategy	48
7.3	Next Steps to be taken	48
8	Conclusions	50
8.1	Assessment of Performance Indicators	50
8.2	Vision for the Future	51
	List of Figures	53
A	Business Canvases	53

Executive summary

Organizational Preservation – Preserve or Forget in Organizational Information Management

This document recapitulates dkd's role in the ForgetIT project. As content management systems are our specialization, we will describe the specific extensions that we have built during our participation to prove the concepts and models as viable and valuable by integrating them with the TYPO3 content management system. We will revisit our pilot applications, main use-cases and their integration with the Preserve-or-Forget framework and its main ideas: Managed Forgetting, Synergetic Preservation and Contextualized Remembering.

To achieve this integration we worked on two use-cases identified in our content management domain: digital asset management and press release work. Alongside we introduced semantic annotation concepts, visual concept detection and a preservation workflow for TYPO3. We evaluated parts of the software components with users using a typical website called "Fake's famous fish-shop" as described in D10.3.

Further on we take a look back on lessons learned over recent years working within the ForgetIT project. Followed by a list of activities performed, TYPO3 extensions written, software created, open source projects undertaken, consultations offered regarding organizational strategies by dkd.

We also disclose possible exploitation ideas sketched in business model canvases. Finally we take a look at the future of ForgetIT and our conclusions.

1 Introduction

This document recapitulates the results of the work within the ForgetIT project with regard to organizational preservation. As industry partner the role of dkd, who is the main partner involved in WP10 and activities related to organizational preservation, was to contribute to the project by taking up central ForgetIT ideas and technologies and to integrate them into the TYPO3 content management system. In more detail, the main goals of WP10 are

- technical implementation of extended functionalities within the TYPO3 CMS in support of the ForgetIT approach and its respective connectors to the PoF Framework, thus creating application pilots for organizational preservation
- learning from the other work packages for extending the content management system and participating in the interdisciplinary work within the different areas of research and development
- researching the organizational endeavors in the context of organizational preservation
- business model creation from ForgetIT ideas

In this deliverable we summarize the results archived along the lines of those goals, we report on first evaluation results, and report on lessons learned. Last but not least we aim to explain on how dkd can continue the work with the gained insights and create a business model from the project's ideas for dkd and others.

This document is meant as a summary document. Thus, we avoid diving too deep into technical detail. We point to results and technical details presented in other project deliverables where needed.

1.1 Take up of Project Results

According to its description of work, the ForgetIT Project is the introduction of a radically more adoptable and sustainable approach to intelligent preservation management, which combines three novel concepts:

Managed Forgetting Inspired by the role of forgetting in the human brain we envision a concept of managed forgetting for systematically dealing with information that progressively ceases in importance and finally becomes obsolete, as well as for redundant information. This concept is expected to help in preservation decisions and to create direct benefits for active information use. We aim to replace such random forgetting processes with managed forgetting, where users are optimally supported in their explicit decisions about what to keep, and how what is kept to be organized and preserved. In particular, we envision an idea of gradual forgetting, where i.e.

complete digital deletion is just the extreme and a wide range of different levels of condensation for preservation is foreseen.

Synergetic Preservation For bridging the gap that still separates active information use from content preservation activities, the project envisions the concept of synergetic preservation, which couples information management and preservation management making the intelligent preservation integral to the content lifecycle in information management. This clearly supports easier adoption, and for by enabling a rich information flow from the information context to the preservation context for more intelligent and informed preservation decisions, e.g. for preservation selection and contextualization.

Contextualized Remembering To bring preserved information back into active use in a meaningful way, even if a long time has passed since their transition into the archive, we envision the concept of contextualized remembering, again inspired by processes in the human brain. The idea here is to already equip resources with rich context information when packaging them for preservation (thus preparing them for long-term interpretation) and to gradually evolve this context information, reflecting evolution in terminology, semantics and interpretation context, thus reaching a semantic level of preservation.

The ForgetIT project performed research in all three areas, the developed solutions are integrated into the PoF Framework.

Obviously, it is not trivial to take up solutions developed in a research project, which target leading edge research results into an operational system as TYPO3, which has to meet today's customer needs. An selective and adaptive process is required here. Within ForgetIT, the following guiding principles have been used in this process:

- **Anticipating TYPO3 customer needs:** Since dkd is a company, it is clearly crucial to select extensions for TYPO3 that will also be attractive for TYPO3 customers.
- **Aiming for immediate benefit** for the usage of an extended TYPO3 system: The benefits of preservation solution is rather long-term. For increasing the chances of adoption it, therefore, makes sense to try to create additional immediate benefit, e.g., by using synergies that result from adopting ForgetIT solutions.
- **Consideration of the trends and perception in the TYPO3 free software community:** Since TYPO3 is developed by a free software community it is crucial to involve and listen to this community in the process of selecting useful TYPO3 extensions based on the ForgetIT approach. This increases the chances that such extensions are taken up by the community creating additional resources for their further development.

Following those guidelines, take up opportunities were identified in all three of the core areas of the ForgetIT project, namely Managed Forgetting, Contextualized Remembering and Synergetic Preservation, in spite of the challenges mentioned above.

For the area of Managed Forgetting, the important conceptual idea of better understanding the value of data and content has been adopted by dkd. This resulted into concrete implementations of technical components such as the content dashboard for observing content value as well as to new ideas for business consulting related to the value of data and content in a company. As a concrete technical link to Managed Forgetting as implemented in the PoF Framework, the extended TYPO3 system has been linked to automated computation of Memory Buoyancy. Furthermore, a component for digital asset management has been implemented, which stresses the value of digital assets used in content management.

In the area of Synergetic Preservation, a CMIS-based connector to the PoF Framework has been developed. Following our guidelines, this connector does not only enable smooth transition to the preservation system. It also opens up opportunities of CMS interoperability due to the CMIS standard, thus, as an immediate benefit, enabling the transfer of content to other content management systems.

For the area of Contextualized Remembering, the benefits of contextualizing content with semantic annotations has been identified as a promising take up opportunity. Such annotation eases the re-use of content and improves its find-ability in the content management environment as well as once published, i.e. in the Web. Furthermore, such context information can be used in information value assessment (see above). For the purpose a semantic annotation tool has been integrated into the extended TYPO3 system.

1.2 Target Audience

The main target audience of this deliverable are other companies in the area of content management and information management in a wider sense, which are interested in investigating an approach of integrated information and preservation management in an organizational setting as well as in better understanding the value of content and data. Furthermore, the deliverable also targets developers working with the TYPO3 CMS and anybody interested in taking up ideas of the ForgetIT project such as Managed Forgetting.

1.3 Structure of the Deliverable

This deliverable consists of eight chapters. After giving a short introduction the first part of this deliverable is revisiting our pilot applications in "Organizational Preservation under ForgetIT" and summarizing its concepts and building blocks: uses cases, CMIS, semantic annotations and value creation in Content Management.

Next are sections discussing the results of all other work packages and their discovered relevance for our pilot application.

Afterwards our dissemination activity with free software communities is discussed.

Followed by the presentation of the formal user evaluation organized by project partner

University of Edinburgh.

The penultimate part of this deliverable is an outlook into the next endeavors exploiting project results.

The deliverable is concluded by an assessment of the key performance indicators set up in the beginning of the project and a future vision put in a nutshell.

2 ForgetIT Pilot Application

This chapter of the deliverable considers the Pilot application for organizational preservation in the light of take up of ForgetIT ideas and technologies. The chapter is therefore structured with the three main ForgetIT concepts a) Managed Forgetting and information value, b) Synergetic Preservation and c) Contextualized Remembering, here especially with focus on the contextualization of content.

The pilot application was developed based on two use cases extracted from customer interviews, community feedback, project interests and technical consideration (see D10.1):

1. the management of press release and/or news (creation, publication, and archivization);
2. digital asset management (data life cycle from configuring the system to archiving an asset (create, capture, organize, and pluralise).

See Figure 1 for a graphical representation of main cases and actors.

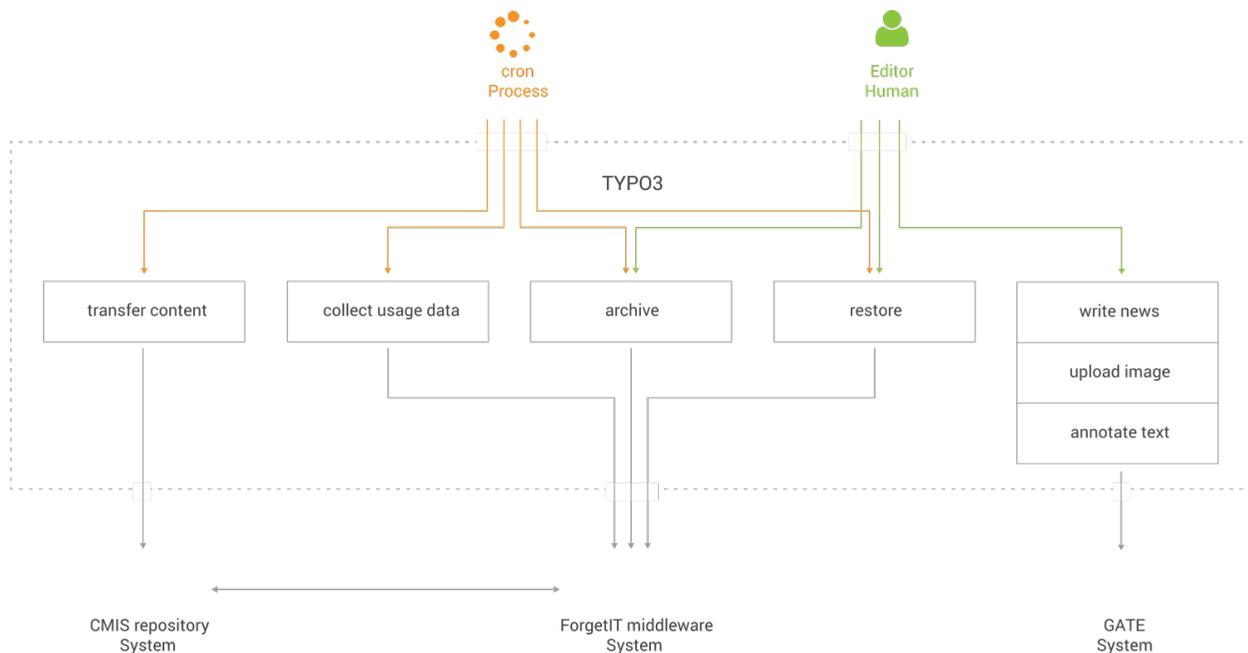


Figure 1: Use Case Overview

For a full step-by-step description (single scenarios, involved personae and machinae) and story cards please refer to D10.1.

2.1 Synergetic Preservation for Content Management

In our initial survey on websites in organizations we identified a need for preservation of web-content. Looking into the preservation activities of the participants we found that preservation was confused with doing regular backups of the HTML, the database and the assets of the website.

Synergetic preservation as defined by the ForgetIT project offers the ability to organizations to close the gap between active systems such as content management systems and preservation systems with the promise of operational synergies and cost reductions.

In the next sections of we will explain how the pilot application enables organizations to activate this advantage.

2.1.1 CMIS and Synergies

Content management systems, such as TYPO3, have become central to the communication needs of organizations over the years.

The nature of heterogenous systems in place in organizations led to the need of a standard for the exchange of content between different content management systems.

A joint group of vendors in the content management sector under the lead of the OA-SIS consortium developed the 'Content Management Interoperability Services' (CMIS) an open standard for inter-operation among different content management systems over the internet.

During early architectural discussions within the ForgetIT Project dkd proposed the use of CMIS as we saw benefits in adopting a standard promising interoperability between system such as our active system and our preservation middleware in order to implement synergetic preservation.

By adopting this approach content cannot only be transferred into the preservation system, but also prepared for being exchanged between different content management systems, thus making investment into content management more sustainable.

It is noteworthy to mention that TYPO3 CMS had almost no CMIS support at the beginning of the project or that the integration was missing large aspects such as the ability to connect to CMIS 1.1 compliant repositories.

In our aim to use CMIS as our main standard we also found out that the required libraries to connect our PHP based CMS with other systems were outdated or did not represent the actual state of the art. We therefore decided to implement a new PHP library following the reference library of the Apache Chemistry project.¹

¹<https://chemistry.apache.org/>

This newly developed PHP CMIS library² created interest in the CMIS community. At the moment we are discussing code contributions into the Apache Chemistry project.³

While developing the content analysis and transformation components described in D10.3 in subsection 3.3.6 we found out, that though the CMIS standard is providing the required abstract functionalities to interchange simple formatted content with simple structures and relationships easily, most advanced content management systems are likely to have heavier data-structures, which are not easily projected into CMIS formats and transformed back.

We decided to add some required type definitions for the CMIS repository used for the current TYPO3 content < – > CMIS content exchange, see ⁴ for details.

If a TYPO3 installation is only using the new CMIS file abstraction layer extension these types are not required, as only files are accessed from and transferred to the configured CMIS repository and no special knowledge of the content structure is needed.

The currently enabled types and the queue status are visible in the newly developed CMIS Manager which we added in year 3, refer to Figure 2 and Figure 3 for a visual impression.

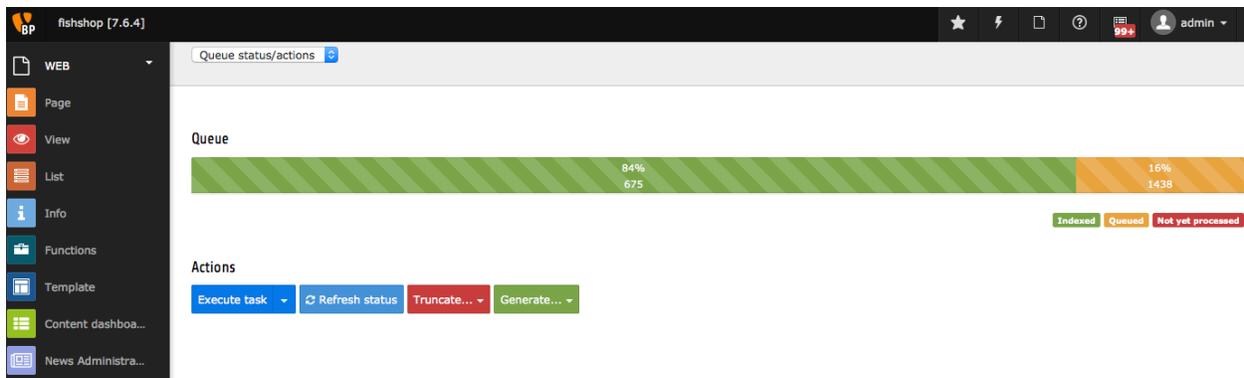


Figure 2: CMIS Manager: Queue

²<https://github.com/dkd/php-cmis-client>

³http://mail-archives.apache.org/mod_mbox/chemistry-dev/201510.mbox/browser

⁴<https://github.com/dkd/docker-alfresco-typo3>

The screenshot displays the CMIS Manager interface with a top navigation bar and a main content area. The top bar includes a user profile 'admin', a search icon, a help icon, a star icon, and a lightning bolt icon. The main content area is divided into three sections, each representing a different table: 'tt_content', 'pages', and 'CMIS'. Each section contains a table with columns for 'Setting' and 'Value'. The 'tt_content' section shows settings for 'Enabled?' (YES), 'Primary CMIS type' (D::dkd:typo3:sys_file), 'Secondary CMIS type(s)' (P:cm:titled, P:dkd:typo3:aspect:forgettable, P:dkd:typo3:aspect:assessable), and 'Property mapping' (None). The 'pages' section shows settings for 'Enabled?' (YES), 'Primary CMIS type' (F:dkd:typo3:pages), 'Secondary CMIS type(s)' (P:cm:titled, P:cm:localizable, P:cm:ownable, P:dkd:typo3:aspect:forgettable), and 'Property mapping' (TYPO3 title, subtitle). The 'CMIS' section shows settings for 'Enabled?' (YES) and 'Property mapping' (CMIS dkd:typo3:general:originalContent).

Setting	Value
Enabled?	YES
Primary CMIS type	D::dkd:typo3:sys_file
Secondary CMIS type(s)	P:cm:titled P:dkd:typo3:aspect:forgettable P:dkd:typo3:aspect:assessable
Property mapping	None

Setting	Value
Enabled?	YES
Primary CMIS type	D:dkd:typo3:tt_content
Secondary CMIS type(s)	P:cm:titled P:cm:localizable P:cm:ownable P:dkd:typo3:aspect:forgettable
Property mapping	TYPO3 bodytext

Setting	Value
Enabled?	YES
Primary CMIS type	F:dkd:typo3:pages
Secondary CMIS type(s)	P:cm:titled P:cm:localizable P:cm:ownable P:dkd:typo3:aspect:forgettable
Property mapping	TYPO3 title subtitle

Figure 3: CMIS Manager: Tables

We also followed the ongoing CMIS4DAM discussion, joined the technical committee as contributors under the TYPO3 Associations Membership in the OASIS⁵ and plan to update our components respectively to solve above mentioned challenges and make use easier.

Happily we can report that we have extended TYPO3 CMS to be able to profit from the use of CMIS and its interoperability with other systems. With this and the adaption of CMIS with the ForgetIT Framework synergetic preservation has become an option to organizations using TYPO3 and other CMIS compliant systems.

2.1.2 Preservation Support for Press Releases and News

We had chosen the press release use case because of the high use by organizations and their requirement to publish press releases and news on their websites. Therefore it was very natural for us to offer preservation support for such content types.

The preservation support for press releases and news is using the CMIS to provide two distinct preservable CMIS structures.

In case of a press release it is possibly a full TYPO3 page, which should be preserved, thus the whole page and all content elements within it are present within the CMIS repository. The PoF collector then can retrieve this page and all elements referred by it.

In case of a news item, a single CMIS element is preserved initially, storing attributes like author and news text. In both cases the contextualizer is able to pick up additional entities by analyzing the annotations added by TYPO3. Restoration happens via CMIS as described above.

2.2 Information Value in Content Management

Applying the concepts of Managed Forgetting of the ForgetIT project to content management systems offers a new approach to the way organizations deal with their assessment of the information value of their content. The collection of data on the usage of content made possible by our components of the pilot application will enable new insights on where values are created and which effects occur with the active interaction with such new information.

In the following two sections we will describe our assumptions regarding the state of content and digital asset management and how our content dashboard component offers a new unified way to handle content and files.

We present the main findings and are confident that we have laid a good foundation for upcoming challenges in content and digital asset management area.

⁵<https://www.oasis-open.org/committees/cmisis4dam/>

2.2.1 Content Management and Value Creation

Looking back at the history of content management systems it can be said that their evolution went from being a publishing tool for few technical staff members to non-technical more communication oriented users. Those users seek nowadays insights on how their communication is performing and this means not only quantitative aspects, such as web-page impressions offered by tools like google analytics or piwik. The CMS users are interested to find possibilities to segment their website users in many ways such as:

- identify: who is the customer?
- track: what are they doing?
- page: what web page are they on?
- screen: what app screen are they on?
- group: what account or organization are they part of?
- alias: what was their past identity?
- client history: Which products do they own? Which do they desire?
- campaigns: Which campaign did reach them?

Current CMS are neither offering such insights nor include a way to store this information near the content produced by them. At the present state most of the installed CMS based websites are making the migration to mobile and/or app based content distribution approach. The value collection as mentioned above will then become in our opinion the next step to be taken by the stakeholders for digital content in the organizations.

With the technologies developed in this work package we have laid the foundations for those upcoming demands.

2.2.2 Advancing Digital Asset Management

While talking with selected customers we discovered their need for more dedicated approach to digital asset management than TYPO3 CMS was able to offer. Often organization have large image archives and their integration into their website leads to non-optimal setups where assets have to be synchronized from image databases to content management systems. This duplication of data leads to risks such as image licensing litigation or reputation loss due to the use of outdated images.

This situation brought up the need of our second use case in the workpackage.

For digital asset management we developed a new TYPO3 backend module named 'Content Dashboard' which is working on top of our new CMIS integration with TYPO3 (see

section 2.1.1). The content dashboard is the new user interface to handle and use all manual ForgetIT functionalities within TYPO3 and see visualizations of the internal relations and calculated values relating to managed forgetting.

For a better understanding of the functionality, please refer to the following screenshot for a visual overview of the new content dashboard. This content dashboard works on top of the CMIS representation of TYPO3 content and enables automatic and manual interaction with the PoF middleware and shows memory buoyancy and preservation value.

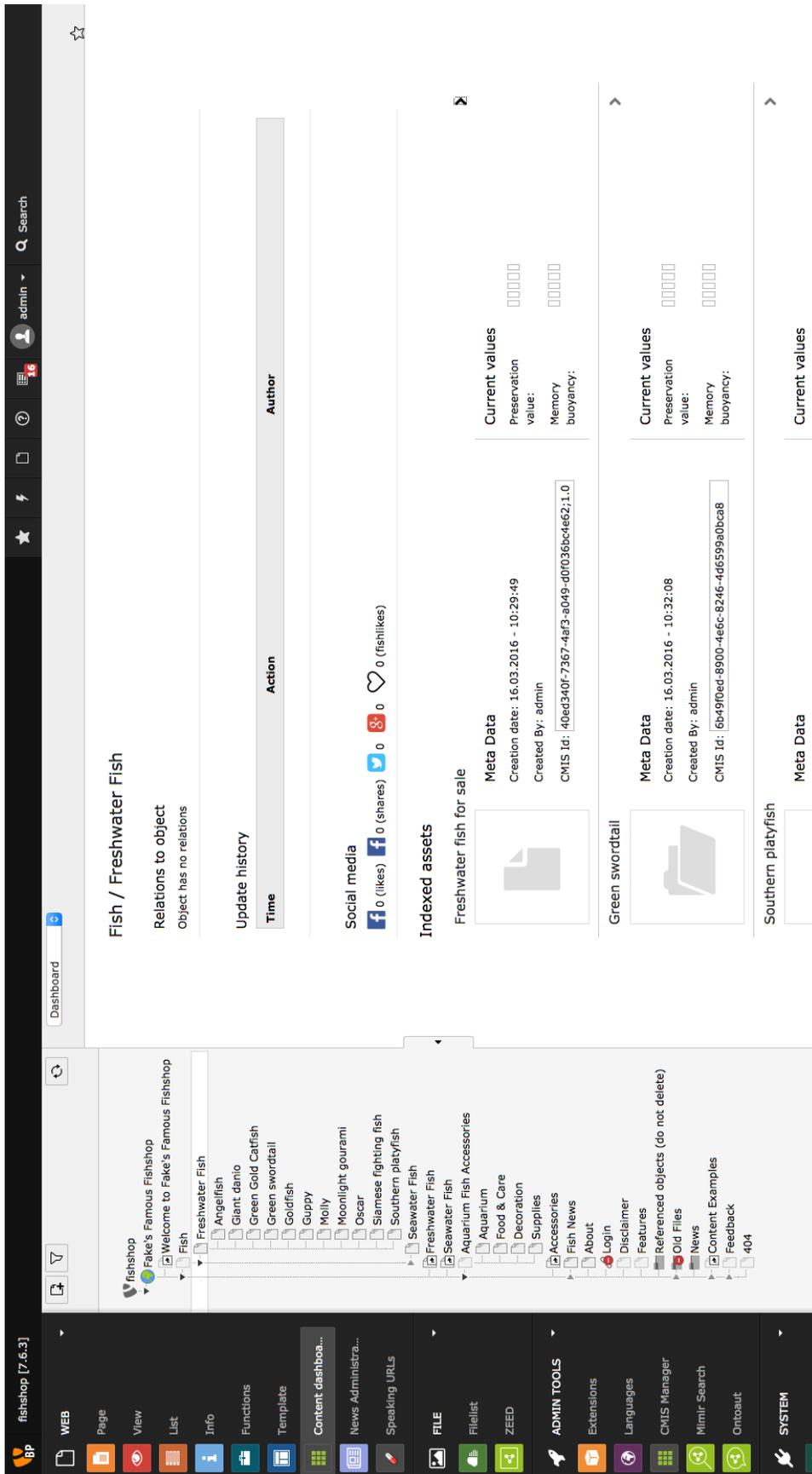


Figure 4: Content Dashboard

2.3 Content in Context

Content is not just characters and other plain data. A human is interpreting the content and working with it on a level of meaning and not raw data. To enable computer systems to work on a meaningful level will broadly expand their abilities in helping humans. A first step towards this is the examination of a content's meaning and its relationships, the following section describes our approach to integrate ForgetIT technology to bring TYPO3 nearer to the meaning behind the raw data.

2.3.1 Semantic Annotation

The current Web mostly consists of documents and media poorly or not annotated at all. The next step in developmental progress is the semantic web (SW). It adds a semantic layer of information relating to objects used in the Web. This layer makes the objects and documents in the Web searchable, weighable and interconnectable. SW adds meaning and context to documents in the Web, making it deeper and more dynamic and above all machine-understandable. This increased understanding of the content's internal meaning can than be used to facilitate better human-machine interaction.

Semantic Annotation facilitates contextualized remembering as well as value assessment for preservation, two main features of ForgetIT. It adds a value to data through context information. Foremost it helps making the decision, whether to preserve data or to forget it by encapsulating organization-specific knowledge (see D6.3 on the role of local context). Semantic Annotation makes it straightforward to search for information, to build synergistic relationships of data, and machine-driven summarization algorithms for texts, images, and multimodal objects.

2.3.2 Semantic Annotation and TYPO3

Semantic Annotations are currently not part of the actual TYPO3 CMS distribution. There had been some initiatives to introduce such functions in TYPO3 in the past, such as the Extensions "SCOTTY" or "semantic" but those extensions were discontinued or lacked the adoption by the TYPO3 users. Oliver Hader, member of the TYPO3 Core Team reports in his Slideshare Presentation about the past and future possibilities of semantic technologies in TYPO3. See ⁶ for details.

Our work in the ForgetIT Project allowed us to revisit the semantic technology proposition and we had support by other work package partners in implementing a promising technology. The following section explains some under the hood implementation of our EXT:annotate TYPO3 extension.

For a final version of the technical components involved in the semantic functions see Figure 5 and Figure 6 for a swimlane diagram, details can be found in D10.3. The user

⁶<http://de.slideshare.net/ohader/meaning-of-data-web-semantics-in-typo3>

interface is shown in Figure 7.

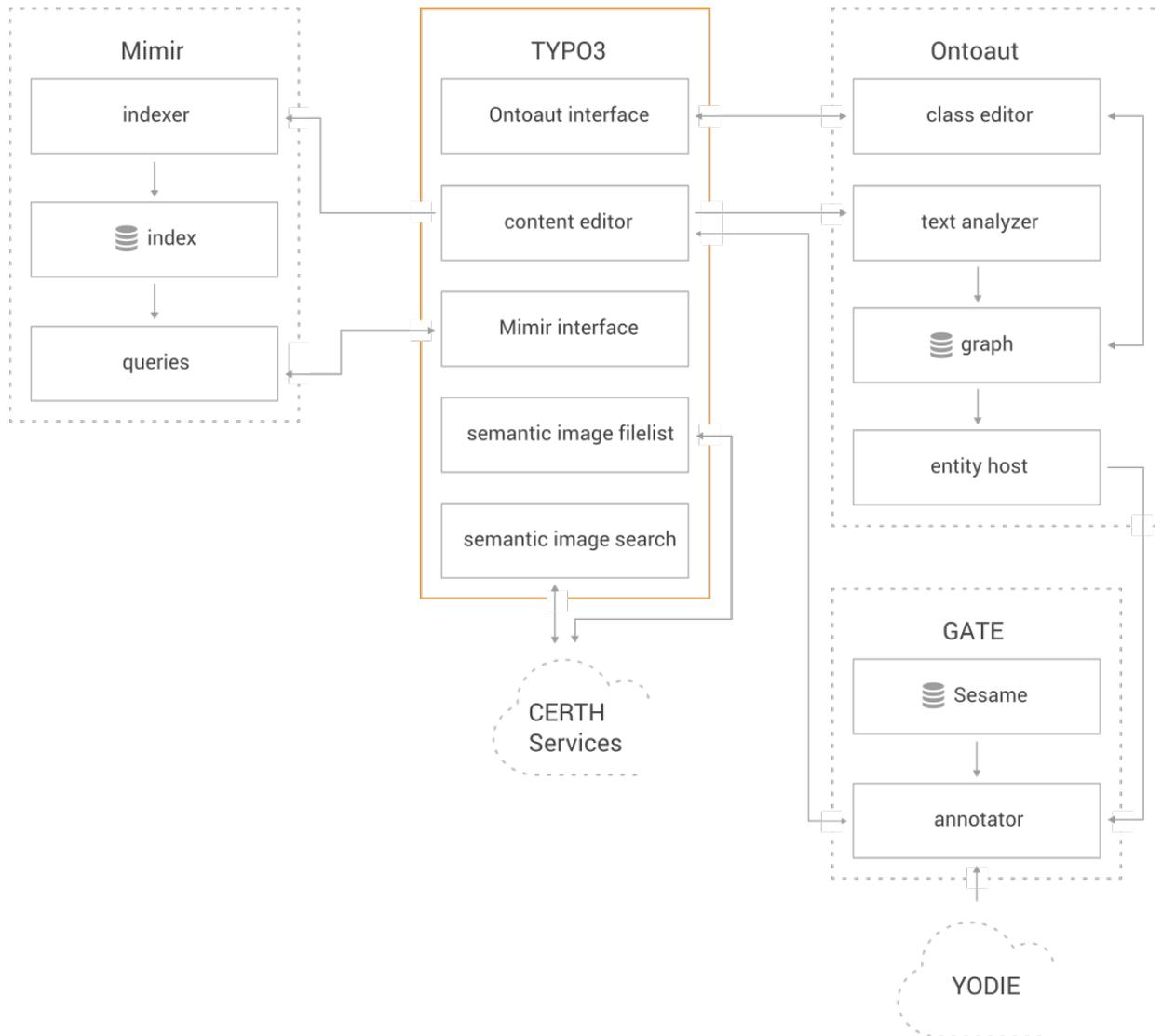


Figure 5: Annotation Component Overview



Figure 6: Swimlane depiction of steps in the annotation pipeline.

Text:

The **mandarinfish** or **mandarin dragonet** (*Syngnchiropus splendidus*), is a small, brightly colored member of the **dragonet** family, which is popular in the saltwater aquarium trade. The **mandarinfish** is native to the Pacific, ranging approximately from the **Ryukyu Islands** south to **Australia**. The **mandarinfish** was first described as **Callionymus splendidus** in 1927 by **Albert William Herre**, an **American** ichthyologist working in the **Philippines**. It was later placed in genus **Syngnchiropus**. The generic name **Syngnchiropus** is from Ancient Greek *syn-*, meaning "together", and *-chiropus* meaning "hand-foot". The specific epithet **splendidus** is from Latin for splendid. The common name of the **mandarinfish** comes from its extremely vivid colouration, evoking the robes of an **Imperial Chinese mandarin**. Other common names include **mandarin goby**, **green mandarin**, **striped mandarinfish**, **striped dragonet**, **green dragonet** and sometimes **psychedelic mandarinfish**. The similarly named **mandarin fish** (*Siniperca chuatsi*), properly known as the **Chinese perch**, is only distantly related. The **mandarinfish** belongs to the perciform family **Callionymidae**, the **dragonets**, which counts 10 genera and more than 182 species. Genus **Syngnchiropus** counts 51 species, divided into 10 subgenera. The **mandarinfish** is in subgenus **Syngnchiropus** (**Pterosynchiropus**) along with the **Australian LSD-fish** (*S. occidentalis*) and the **LSD- or psychedelic fish** (*S. picturatus*). To date, *S. splendidus* is one of only two vertebrate species known to have blue colouring because of cellular pigment, the other being the closely related **psychedelic mandarin** (*S. picturatus*). The name "**cyanophore**" was proposed for the blue chromatophores, or pigment-containing and light-reflecting cells. In all other known cases, the colour blue comes from thin-film interference from piles of flat, thin and reflecting purine crystals. The **mandarinfish** has a body shape similar to a **goby**, but the resemblance ends there. The vivid coloration sports a bright blue background, with swirly orange stripes and a blue-greenish face

Path: body 578 words

Annotations

Annotate! Add annotation Index! Analyze!

> Mandarinfish x

> Dragonet x

vocab: http://dbpedia.org/ontology/

resource: http://dbpedia.org/resource/Dragonet

typeof: http://dbpedia.org/ontology/Fish

Add Property Around Selection Add Hidden Property

> Dragonet x

> Mandarinfish x

> Ryukyu Islands x

> Australia x

> Mandarinfish x

> Albert William_Herre x

Figure 7: Annotation editor and automatic annotations within TYPO3.

The primary usage of Semantic Annotation is to amplify searches, provide context and connect content created in CMS. Secondary usage maybe to make it machine understandable and to create a higher ranking of content with search engines. The secondary usage is not high in importance with the PoF Framework, but entices clients and commercial users to consider Semantic Annotations in their endeavors.

Semantic Annotations currently are experiencing an acceptance curve typical for new technologies.⁷

2.3.3 Contextualization for Press Releases and News Management

We expanded the content scope of the 'press release' use case to include generic news as well because we found that more often customers publish more informal news and not only formal press releases. Additionally we recognized that some of the features discussed in D10.1 were out of reasonable technical scope and distinction, so we slimmed the use case to the necessary functionality to publish semantic news and press releases and left the internal development and discussion of the to be published content out of the final implementation.

The functionality is based on the TYPO3 News module, and automatic semantic annotation and manual annotation editing has been implemented in a way that is fully consistent with the annotation mechanism for pages.

2.3.4 Concept Detection in Pictures

Semantic annotation within text-based documents is straightforward. Foremost it requires defined ontologies involved, on which to build annotations. Concept detection in pictures relies on software analyzing the material. Is it a landscape, or a face? Is it a photo, an artwork, or a copy? Based on the result the software may continue its inquiry, detecting whether the landscape picture depicts a sunset, mountains, or a cityscape, and so on. The software also detects metadata like picture quality and size, asks whether the picture is part of a sequence, or does it show something recognizable (e.g. the Eiffel Tower) or someone famous.

This kind of image analysis was one of the main functions that convinced our two company case study partners, ACO and Spielwarenmesse, to use the PoF system. Although both have extensive photo databases, searching for content is difficult, and it would be helpful

⁷ The technology adoption curve (TAC) is a theory about how individuals and organizations behave in implementing innovative technologies. A quick examination of the framework shows some similarities to the product life cycle curve discussed in business marketing courses. The theory is however more sophisticated than a life cycle or a diffusion model. The underlying model of technology adoption identifies 5 types of adopters of technology with very different interests and buying characteristics. The companies and individuals that are first to adopt a new technology are called innovators. The second type is known as the early adopters. The third type is called early majority, then the late majority adopters and, finally, the laggards. [Daniel J. Power, dssresources.com]

to have both generic annotations (people, landscape, object) and annotations that are specific to the company's needs, but not implemented yet (brick, drain, pipe). Due to the flexibility of CERTH's solutions, generating company-specific classifiers is feasible with sufficient pre-annotated data.

Most of the anticipated immediate benefit would be to the internal workflow of the companies, in particular the website management and updating workflow, which is then expected to result in a better user experience for the website users.

To achieve this functionality within TYPO3 we are using the technologies and services developed in work package 4 by adding a new search mode for the TYPO3 file browser and by adding context modules for each image, connecting also to the quality analyzer. This development is ongoing, results will be presented at the final review. Figure 8 shows the an example of the integrated concept detection and Figure 9 the quality analysis.

2.4 Software Component Overview

Following is a list of all developed or extended software components to support main functionality as described above.

- PHP CMIS library: Adaption of the Apache Chemistry CMIS library for PHP and TYPO3.⁸
- TCA to CMIS service extension: TYPO3 extension to convert CMIS objects to TYPO3 content and vice versa.⁹
- FAL CMIS driver: This TYPO3 extension enables the usage of CMIS objects in a CMIS repository as TYPO3 FAL (File Abstraction Layer).¹⁰
- Content Dashboard: TYPO3 extension that displays ForgetIT values for memory buoyancy and preservation value and allows manual archivization.¹¹
- Semantic Annotations: TYPO3 extension that expands the TYPO3 rich text editor so that automatic and manual semantic annotations are possible.¹²
- MIMIR Search: TYPO3 extension providing a semantic search and SPARQL queries for the annotations added to TYPO3 content.¹³
- Visual concept and quality detection: TYPO3 assets used in the TYPO3 CMS. At a later stage additional functionality might be added such as face recognition.¹⁴

⁸<https://github.com/dkd/php-cmis-client>

⁹<https://github.com/dkd/typo3-ext-cmis-service>

¹⁰<https://github.com/dkd/typo3-ext-cmis-fal>

¹¹<https://github.com/dkd/typo3-ext-contentdashboard>

¹²<https://github.com/dkd/typo3-ext-annotate>

¹³Contained in the annotate extension.

¹⁴<https://github.com/ksjogo/typo3-ext-semantic-images>

- **Ontoaut:** Independent software component that helps building an organizational knowledge source for the annotation pipeline. JavaScript library integrated as a backend module.¹⁵
- **REST/API:** TYPO3 service component that enables TYPO3 CMS to pass objects to the REST API of the different ForgetIT service endpoints.¹⁶
- **Aggregation:** TYPO3 service component to consolidate and store all kind of user interaction and website visitation logs as CMIS objects in order to calculate MB and PV of TYPO3 content.¹⁷
- **ZEED (Zero Example Event Detection):** TYPO3 extension that accesses the CERTH service enabling users to search for matches of text and images concepts.¹⁸
- **TYPO3 Scheduler Tasks:** a set of TYPO3 scheduler tasks that enable time based and asynchronous interactions with local and remote components.¹⁹
- **GATE plugins:** GATE (General Application for Text Engineering) functions such as dedup (de-duplication), RDFa Import und RDFa Export to enhance functions in the semantic pipeline.^{20 21 22}
- **Docker based infrastructure:** A set of Docker files to setup the active system and its required additional components such as: webserver, database, CMIS repository and many more.^{23 24 25}

¹⁵<https://github.com/ksjogo/ontoaut>

¹⁶Awaiting publication

¹⁷Awaiting publication.

¹⁸Contained in the semantic images extension.

¹⁹Contained within the cmis service extension

²⁰<https://github.com/dkd/forgetit-gate-anti-dup>

²¹<https://github.com/dkd/forgetit-gate-rdfa-import>

²²<https://github.com/dkd/forgetit-gate-rdfa-export>

²³<https://github.com/dkd/docker-typo3>

²⁴<https://github.com/dkd/docker-alfresco-typo3>

²⁵<https://github.com/dkd/docker-mimir>

Edit File Metadata "IMG_2291.jpg" on root level

General Access Metadata Camera Data Categories Semantic Image



IMG_2291.jpg (303 Kibytes)
Image (image/jpeg)
Location: fileadmin/ (auto-created) - /photos/IMG_2291.jpg

Semantic Image

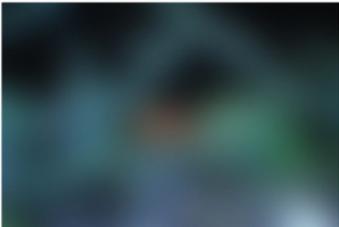
Quality Analysis:

Concepts:

- Outdoor 0.9980468
- Building 0.9320081
- Person 0.7757447
- Sky 0.6873762
- Urban_Scenes 0.6593029
- Daytime_Outdoor 0.6585699
- Windows 0.64571
- Cityscape 0.6162207
- Pavilions 0.5877316
- plaza 0.577307
- statue 0.5749765
- Suburban 0.5744335
- City 0.5449136
- house 0.4760518
- Religious_Building 0.4641779

Figure 8: Detected Concepts

General Access Metadata Camera Data Categories **Semantic Image**

 **nemo_blurred.jpg** (331 Kibytes)
Image (image/jpeg)
Location: fileadmin/ (auto-created) - /photos/nemo_blurred.jpg

Semantic Image

Quality Analysis:
Quality 0.764793
Blur 0.819307
Contrast 0.360624
Darkness 0.419159
Noise 0.439713

Figure 9: Low Image Quality

3 Leveraging ForgetIT Results

As an application work package in the project the idea is to build upon and leverage the ideas developed in the other WPs, especially the technical WPs.

Having established the necessary infrastructure in Years 1 and 2, we used Year 3 to focus on integrating technologies developed in other WPs. This also addresses one of the recommendations of the second year Review.

Working closely with WP8, we integrated components of the PoF Framework as they became available. We have now included core aspects of the framework such as the smooth transition to the archive, the retrieval from the archive, support for contextualization, and information value assessment methods for managed forgetting.

In the following we discuss which ideas, methods and technologies were used from the individual WPs.

WP2: Foundations of Forgetting and Remembering.

The results of WP2 highlighted fundamental differences between human and organisational forgetting, helped us understand the nature of organisational forgetting, and laid the foundations of a consulting model based on ForgetIT results. While human forgetting mechanisms are largely consistent, organisational memory is highly diverse and therefore requires careful study before we can tailor a ForgetIT solution to a particular organisation's needs. The two longitudinal case studies conducted in collaboration with WP2 as well as the three case studies of non-profits conducted by WP2 clearly show the degree of variability seen in practice.

WP3: Managed Forgetting Methods.

The main contribution of this work package was to establish a conceptual computational model of managed forgetting, the first pillar of the PoF Framework, that could then be applied to the TYPO3 context. We worked closely with L3S in order to develop a good operational understanding with the two key concepts for Memory Buoyancy and Preservation Value.

WP4: Information Consolidation and Concentration.

The crucial contribution of WP4 was to give us better tools for information assessment, consolidation and concentration, either analyzing multimedia data or text. For the moment we successfully integrated the visual concept detection and image quality assessment technologies developed by CERTH. In the long term, we want to integrate more developed functionality like face detection and duplicate detection into the CMS level, so

the user can be assisted at manual preservation tasks and check results from automatic preservation. The ZEED search will be integrated to automatically associate images with TYPO3 content elements based on their main text.

WP5: Joint Information and Preservation Management.

From this WP, we obtained essential information about the design and implementation of preservation systems, and a better understanding of the synergetic preservation process which is one of the three key aspects of the PoF Framework. The Information Continuum Model, which shows the typical life cycle of a piece of content from creation through storage to publication, was particularly useful.

WP6: Contextualization / Decontextualization

WP6 provided us with the tools for Contextualised Remembering, the third pillar of the ForgetIT approach. Members of the dkd team spent time at USFD to learn about the contextualisation approaches developed in the project, in particular semantic annotation, and we created a user interface that allowed end users to add and check textual annotations to all pages of a website, including news.

WP7: Computational Storage Services.

We learned in this work package about the OpenStack storage technology swift and its expansion by our project partner IBM called storlets. Throughout the project we have seen the storlet evolve from java application to docker containers, a technology that we also used in our development, testing and production environments. It is understood that storlets offer a good approach to solve the problem of manipulation of stored objects without the need to move the data to applications to do the computation. IBM also provided important lessons about how new technologies such as data containers are disseminated and taken up in large companies. A short Youtube Video summarizes the role of IBM technology well: <https://www.youtube.com/watch?v=3rXeNbps8wo>

WP8: Preserve-or-Forget Reference Model and Framework.

We pushed the CMIS standard into the ForgetIT project within the first year and were able to demonstrate the usefulness by interchanging content between TYPO3 and PIMO through CMIS at year 2 review. Our demonstration implementation with Alfresco as CMIS repository shows a way for other content management systems to follow, if they are interested in using ForgetIT technology. In the final year of ForgetIT we worked closely with EURIX, the WP8 lead, to bring the ForgetIT standards to the Open Archival Information System community, which fits in very well with the free software ethos of TYPO3.

WP9: Personal Preservation.

Although there are clear differences between personal and organisational preservation, DFKI's expertise was invaluable when integrating solutions which had been developed by researchers into a working system. Through several workshops with WP9, WP3, and WP2, we refined our own understanding of Managed Forgetting and discussed ways of implementing it in a working system. DFKI also provided us with invaluable insights into the semantic web and ontologies of concepts, which helped us integrate relevant WP6 technology.

4 Involvement of Free Software Communities

Part of the strategy of dkd with respect to development as well as exploitation was the involvement of the TYPO3 free software community. dkd is heavily involved in this community and has also previous experience with successful co-development within such communities. In this section we summarize the blueprint we used to ensure effective involvement (Section 4.1) and describe how this blueprint was adapted specifically for involving the TYPO3 community as early as possible in ForgetIT-related activities (Section 4.2), thus fostering later take up of project findings.

4.1 Previous Experience

When planning our open source product strategy for ForgetIT, we looked back at experiences gained developing a similar project in 2009, when we detected demand of a search solution for TYPO3 CMS.

Our inquiry into existing solutions produced Apache Solr as the best application. We identified essential software components, then estimated time for necessary programming labor. Cost for complete integration could not be covered by client assignments and in-house investment only.

We started a call-in to the TYPO3 community, asking for other companies interested in a better, faster search in TYPO3. Community members replied that they had detected the same demands for a better search and came to similar conclusions regarding realization of the project.

We proposed a solution, initializing an Early Access Program. This program had a financial component, giving access to first versions of its software and offering adopters to be included as competence spots for the search in this community.

Early partners quickly joined the financial program, allowing us to realize functions listed in a white paper. If a partner wanted to force development of a certain function, they would demand so and thus push it further up in the roadmap.

Feedback from the project enabled us to quickly identify fresh requirements and implement them.

Today we look at about 300 partners in the Apache Solr for TYPO3 program, each providing up to 4000 Euro in funding. Noteworthy is today's position of this TYPO3 extension: projects requiring complex interactive search use this solution, initialized by us. Hosting companies, specialized in TYPO3, offer Solr based search as a part of their product offering.

We see similar potential in integration of the CMIS technology today, although the use case is slightly more narrow than search provided by Solr. We are closely observing the current evolution of content management. Because demand of functionality with con-

tent management systems is expanding, the utilization of content repositories, as offered through CMIS, appears to be the natural solution.

4.2 Bringing ForgetIT to the TYPO3 Open Source Community

In order to ensure impact of ForgetIT related ideas, we had to closely involve the TYPO3 free software community. This meant that when developing for ForgetIT, we also had to keep in mind how new functionality would work in the wider context of the TYPO3 system, and whether it would be applicable to other solutions as well.

From a strategic point of view, if components that were originally developed for ForgetIT are widely adopted for content management, it is much easier to add the additional preservation functionalities later. Therefore, we promoted both the overall idea of preservation and managed forgetting, and the underlying reusable components, such as CMIS integration and automatic annotation facilities.

We reached out to the TYPO3 Community through regular presentations, discussion panels, and developer days.

Stations in chronological order:

TYPO3 Conference North America Mai 2013 ‘Digital Dark Age - Are we doing enough to preserve our website heritage?’ Introduction to the aspects of preservation presented to TYPO3 CMS users and integrators. The presentation showed how ForgetIT will help to preserve the value created through online content.

TYPO3 University Anney Mai 2013 We participated in a panel discussion about the future of content management systems.

TYPO3 Developer Days July 2013 Discussion in multiple open sessions how ForgetIT could improve content management.

TYPO3 Camp Mallorca September 2013 Discussion panel: ‘Outside the box CMS’

TYPO3 Conference October 2013 ‘Some store to remember, some store to forget’ - general overview over the topic of managed forgetting and main ForgetIT concepts.

Dutch TYPO3 Congress October 2013 Repetition of the previous talk for another audience.

TYPO3 East Europe November 2013 Session on digital preservation.

TYPO3 Camp Milano March 2014 Session on digital preservation.

TYPO3 Conference October 2014 ‘A love affair: TYPO3 and CMIS’ - Presentation about the newly developed CMIS functionality, gathered feedback from users wanting to easily connect existing systems to TYPO3

TYPO3 Developer Days July 2015 'Annotations' - introduction into the semantic web, how TYPO3 could profit from it and the current state of automatic annotations within it

TYPO3 Developer Days July 2015 'CMIS and TYPO3' - developer perspective on the CMIS implementation

TYPO3 Camp Mallorca September 15 'The ForgetIT project' - general overview on all achievements within and for the ForgetIT project, active discussions on different topics

TYPO3 Conference October 2015 'Annotations' - presentation on the semantic web and how TYPO3 might get there

TYPO3 East Europe 2015 'Annotations' - same presentation, small changes

These presentations, panel discussions and general community activities successfully planted initial seeds of ForgetIT ideas in the TYPO3 community shortly after the project started, and increased awareness for our ongoing outreach and open development process.

We discussed possibilities within TYPO3 and asked developers for user stories and ideas. How would they develop ForgetIT extensions to get a better overview on actual interests and needs? How did they adjust some of the tasks and goals to better fit these requirements? This in turn informed the development of our own use cases and the required functionality.

All together our community program can be considered a great success. We created a lot of interest and helped TYPO3 core team members gain interest in the semantic web and managed forgetting strategies.

Our CMIS library was integral to this success. We were able to engage not only with the TYPO3 community, but also with the CMIS and the broader php community. Thus, our preservation-enabled CMIS solutions are a perfect way of ensuring fast uptake of a more comprehensive ForgetIT solution.

At the time of writing, the `php-cmis` package created by dkd for the project, which was uploaded on the php package repository `packagist` on May 28, 2015, has been installed 557 times in 10 months, is being watched for updates by 17 developers, and was bookmarked (starred) by 21.²⁶

The package is also actively considered as a new standard CMIS library for php, which would bring ForgetIT technology also to other web content management systems such as Drupal.

²⁶Data retrieved from <https://packagist.org/packages/dkd/php-cmis>, March 14, 2016.

5 Evaluation

As this was the first time that user-facing components of the organisational ForgetIT integration had been tested with users outside of dkd and those members of the TYPO3 community who had access to the code, we used a strongly formative paradigm. Instead of assessing how successful participants were at completing tasks without further guidance, and measuring task completion times, we first familiarised ten participants with the TYPO3 CMS interface through a series of website editing tasks, and then asked them to use the annotation facility, providing them with extensive help throughout.

Observations were recorded using screen capturing tools and on tape. The experimenter debriefed each participant in depth about their experience and made detailed notes. The only formal assessment was the System Usability Survey (SUS), which participants completed both for the annotation mechanism and for TYPO3 itself. This allowed us to obtain summary judgements and determine the effect of users' lack of familiarity with TYPO3 on their assessment of the annotation tool. A detailed description of the design and methodology can be found in D2.4.

5.1 Participants

We recruited 10 participants aged between 21 and 50 from a variety of backgrounds.

5 participants were current PhD students in the University of Edinburgh. Three of them (P1, P6, P10) were experienced in web design and development, having worked on websites commercially. One of them (P6) had done some tagging (HTML metadata tags), but was not very familiar with search engine optimization (SEO). P8 had some experience as a web editor, while P9 had only worked on websites for university assignments. P9 had a strong design background, with a Masters in Design and Digital Media, and had been briefly involved in advising on the design of a commercial website in a public relations agency.

The remaining 5 participants were all employed in the IT sector. P2 and P3 were very experienced web editors. One of them (P3) was a Commercial Officer and had about 10 years' experience with websites, mostly as an editor, having occupied freelance, part time and full time positions. The other (P2) worked as an Illustrator, e-learning Designer and Developer, and had over 15 years' experience in the web. They both had experience with tagging and SEO in terms of using phrases which can bring people to the website, and naming objects carefully.

P4 and P7 were very experienced self-taught web designers and developers, having started learning about web technologies by themselves more than 10 years ago, and having worked both freelance and as employees.

P5 was an Oracle developer specialized in working with databases, who had a keen interest in web technologies, especially HTML5. He did not have any tagging or SEO

experience.

8 out of 10 participants were familiar with Content Management Systems (CMS) like TYPO3, although none had heard about TYPO3 before, because this CMS is not widely used in the UK. Most of the participants with CMS experience had used WordPress, and other content management systems mentioned were Drupal, Polopoly (used in the University of Edinburgh), Cubic, PHP-Nuke, and Joomla.

5.2 Usability Ratings

As Figure 10 shows, participant judgements vary widely. The mean score for TYPO3 is 69.5 (standard deviation: 19, range: 40–90), which is in the acceptable range according to existing norms for the SUS questionnaire. This validates the choice of TYPO3 as our main platform. Despite users being completely unfamiliar with the environment, they were able to navigate TYPO3 after a short introduction.

For the annotation tool, scores are worse, with a mean of 50 (standard deviation: 30, range: 0–85), and far larger variation in ratings. The participants who gave the annotation system the worst scores were P1, P2, P3, P4, and P7. The last four were more experienced web developers and editors, while the first was a PhD student with some coding and SEO experience.

Looking at the debriefing interviews, the scorers who were more critical were those who imagined using the manual annotation system in their daily web editing work, whereas the more positive responses came from participants who valued the ability to automatically annotate web pages and/or were sufficiently familiar with databases to be able to cope with the Mimir search syntax.

5.3 Summary of Observations

In this section, we restrict ourselves to observations of the annotation tool, as the version of TYPO3 that participants used, 6.1, has since been superseded by Version 7.

Overall, the most useful aspect of the system shown to the participants were the automatic annotations. Take for example this summary from P7:

I can see this system being very very helpful for a big platform where you have a lot of data and lots of news, and in that moment a company can use in a much better way the code from annotations in order to provide more meaningful results in their search engines. (P7)

Since this was a development version of the annotation tool, there were no manuals or cheat sheets, but dkd produced two sets of short videos, one covering the use of TYPO3 and one covering the use of the annotation tool, which participants viewed before

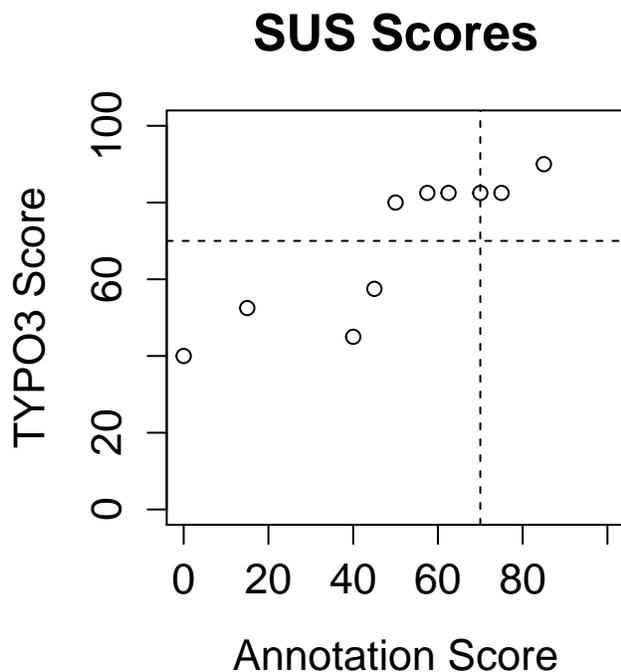


Figure 10: SUS scores for TYPO3 and the Annotation Tool

beginning their work. Although participants were allowed to take as many notes as they wished, and to return to the videos if they needed to, all participants would have benefited from a detailed manual and/or an online help function. In fact, some participants usually relied heavily on online help and web resources when trying out new software.

This lack of help affected especially the MIMIR search function, which required participants to type in queries in a context-sensitive query language. However, some participants, such as P5, the Oracle developer, picked it up quickly.

5.3.1 Technical Problems

Some problems that occurred during the evaluation were due to bugs or stability issues. For example, for three participants, the automatic annotation facility lost connection to some of the dbpedia servers. Sometimes, the annotation menu was not opened automatically, and in a few cases, annotations could not be deleted. One participant was unable to add properties.

Several participants were unable to nest annotations, i.e., to annotate a part of an already annotated phrase. The behaviour of automatic annotations was also not always consistent. Sometimes, manually added annotations remained after a new automatic annotation run, even after they have been deleted.

5.3.2 Manual Annotation

While most participants appreciated the automatic annotation process, the manual annotations proved more difficult. Some of this was due to the labelling of the buttons involved, as noted by P3:

[...] there are some funny things like with the annotation having variations of the word 'Annotate' 3 times in the page, without it being helpful what the different functions are (P3)

In addition to the naming issues, there were also no indications whether a button was active, should be used, or was applicable at all. The Index button confused several participants because it was grey, a color that is usually associated with inactive buttons.

The buttons also appeared in a place that was not consistent with user expectations, at the bottom of the text-box that contained the text to be annotated. Traditionally, such buttons are located at the top or at the side.

The design of the list of annotations also made it difficult to match an annotation with the phrase to which it referred. While automatic annotations were ordered according to their place in the text, new manual annotations appeared at the top of the list. The names given to the annotations were taken from the concepts that had been annotated, not from the text, which also added to the confusion.

But perhaps the most confusing aspect of the user interface was the fact that the list of annotations itself was positioned below the annotation buttons and the text itself. This further increased the difficulty of matching annotations to phrases.

The way in which the user interface had that list below, rather than just integrating with the existing... There's already a rich text editing box and everyone... it's like standard across all pieces of software, you know, Microsoft Word use the same thing, there's a ribbon bar at the top which has all the controls, it seems like you just put it in there. (P6)

P8 also suggests including the annotated phrase in the annotation list itself.

maybe when the label doesn't match the highlighted text, you could have in the header of the annotation mentioned, maybe in brackets, the text itself, so it's easier for the user to know, ok, that text has been annotated with this label. (P8)

5.3.3 MIMIR Search

The main issue with the MIMIR search was that it relied on a query language which was not well described or documented. While the IT experts among our participants who

knew SQL did not have many problems, those participants who had mainly worked as web editors or web developers were confused.

The results of the MIMIR search also contain both phrases which are highlighted in orange (for automatically added annotations), and phrases which are coloured orange (manually added annotations). No such distinction is made in the main annotation screen.

Finally, MIMIR search is implemented in TYPO3 as a generic view, such that when navigating the page tree, one stays in MIMIR search mode. This makes sense for experienced TYPO3 developers, but presented a problem for most of our participants, who were unfamiliar with TYPO3.

5.4 Suggestions for Improvement

When interpreting the results of the evaluation, it is important to bear in mind that the annotation facilities are only a small part of the whole solution. Most of the work done by dkd ensures the smooth functioning of the underlying system. The annotation facility evaluated here allows users to adjust some of those automatic calculations, namely annotation with underlying concepts.

Our debriefing interviews show clearly that the annotation and search facilities are promising, and many participants could see how they would be useful for managing and preserving large websites. These results are similar to the ones for WP9 as reported in D9.5, where users appreciated the ease of integrating new documents into an existing knowledge base with minimal overhead, once the main ontology had been built.

However, the current implementation still has some usability issues that would need to be resolved before deployment. A comprehensive online help, especially for the MIMIR search, together with a crib sheet for commonly used processes, will also greatly improve the user experience. Since this was a very early version of the user-facing system, such help was provided by the experimenter instead.

Following the design guidelines developed by Nielsen²⁷, the recommendations are:

Make system status visible: This concerns issues such as indicating whether there are unsaved annotations, or whether a new automatic annotation would delete existing manual ones. For example, the button colour or shading should be changed to show when it is active, and the Annotate! button should only be active when the annotation fields are not present.

Consistency: As far as possible, the interface should be consistent with cues and layout conventions that are widely used across operating systems, such as menu button position.

²⁷Nielsen, J., & Mack, R. L. (Eds.). (1994). Usability inspection methods. New York, NY: John Wiley and Sons.

Provide help and documentation: Users would greatly benefit from online help, in particular for the MIMIR search functionality. Since this was an early beta version, the evaluation was designed to allow for this help to be given by the experimenter instead.

5.5 Next Steps

In this study, we only examined some of the ForgetIT organizational preservation components. Most of the functionality that was implemented does not require manual intervention. Judging from our interviews with the two WP10 case study companies, this high level of automation will be a major selling point for the system in practice.

Therefore, the next step for the evaluation needs to be deployment on a small website that is not mission critical. This would allow us to evaluate all of the user interfaces provided, each of which covers small manual management and decision tasks, and to study how well the ForgetIT system with its substantial capacity for automation fits into the workflow of an organization.

Between the time of our evaluation and the time of writing the Deliverable, TYPO3 has also changed. The overall environment is now TYPO3 version 7. The rich text editor RTEHTMLArea which serves as the current cornerstone of the annotation interface has been a part of the TYPO3 interface for 15 years and is outdated. The community is considering the integration of an alternative, CKEditor (ckeditor.com) as result of our influence and dissemination of ideas of the ForgetIT Project. This editor has facilities that are perfect for annotation work, such as inline window widgets that can be displayed next to linked text. This alone would address one of the main usability problems seen in our evaluation, namely the difficulty of associating the annotation with the phrase to which it refers. It is noteworthy that CKEditor is the technical base for the SEED Editor also used in PIMO and we would expect synergies from this technical adoption.

6 Lessons Learned

As a small and medium enterprise participating in a European project, dkd had to answer a number of questions for shaping their contributions to the project:

- How do commercial organizations manage, preserve and publish digital information (content) and documents (digital assets)?
- How is web content and its context delivered to the PoF Framework and transferred into a future CMS?
- What role does TYPO3 play in the organizational preservation scenario? Which interfaces, extensions and standards need to be developed to ensure a managed preservation process?
- What does the final user experience look like?

DKD's expertise and focus is on creating data driven solutions for clients and customers. They are typically looking for solutions moving their business deeper into data-driven forms of enterprise. We support them with our deep understanding of content management systems like TYPO3. ForgetIT adds interesting new aspects to our view on content creation.

The ForgetIT project was the first scientific research project dkd Internet Service participated in. As such, we've learned a lot of different lessons from the project. They cover a multitude of aspects such as research methods, conceptual insights, technologies, organization of work in the sphere of European partnerships, research in scientific surroundings, etc. In the following we summarize our most important lessons learned:

Allow for start up time: In the beginning of the project, we spent quite some time to wrap our minds around what ForgetIT was, how we could comprehend what the project had already brought forward. Work packages, pace, and forms of collaboration are different from what is required working with clients.

We put dkd's work and activities in relationship to ForgetIT. We took in a lot of information transfers from our partners in ForgetIT. We phrased insights from this input. And searched for angles to let the project take off.

Understand and cope with differences in working style: Working as an agency, we are focused on accepting problems a client brings to us, then offer several solutions to solve the problem presented. In contrast, a research project such as ForgetIT is not so much driven by today's customer needs, but aims to envision future technologies. This gap had to be bridged in the first year of the project.

Choose the right team: We sensed how helpful it would have been, if we had a more stable team right from the start. Next time we will begin with a systems architect and

a project leader to supervise progress and at least one senior developer to keep the flow of progress steady.

Joint coding sessions are important: Moments of concentrated collaboration, when we actually conceptualized, coded, and collaborated with partners in ForgetIT where from our perspective among the highlights of this project. Next time we join a demanding project like ForgetIT, we would aim for more such meetings to propel the project towards practical solutions. In addition, regular developer talk about technological architecture and component communication are helpful.

Interdisciplinary conceptual discussions advance the project: All participants benefit from deeper and broader discussion of major findings and insights. That might generate better replies and contributions from all involved, a more comprehensive perspective for topics such as further sensitization toward data, ideas of information value, importance and functionality of annotations and metadata, as we experienced among ourselves.

Dealing with evolving Interfaces: The fact that the technologies and the need of the application partners to build upon such interfaces implies some coordination challenges. From the side of the applications, it would be helpful, if a clear description regarding component communication would be created early in the project. However, as this is a research project many of the methods and ideas are still under research, i.e. in flux, which also effects possible method interfaces.

Flexibility required: We learned that our approach of a lot of finely grained user stories has its problems. The scope regarding of our two use cases was changing during the project. Only with time a more robust ideas was developed, what is expected from the users and how TYPO3 would interact with them through the PoF Framework.

Change in attitude towards clients: Our new client attitude was influenced by the work with UEDIN and UOXF. Previously we looked at clients and asked “what they do” and tried to represent this with a website. Today we are more influencing, we are giving advice on “what they should do” – our role transformed from observer to adviser. With that came another change. It may be well worth our time and effort to develop technologies, even if they are not yet directly requested by customers – the future need is already foreseen by us and considered to be a future client request, giving us a tremendous head-start, when the actual requests comes in later.

Time of project ForgetIT comes to a close. “Preserve-or-Forget” has unleashed many new insights and developments. New ideas for usage and applications have sprung to life and we are looking forward to develop more integrations.

It is our goal to connect scientific insights found in ForgetIT with real-life demands and requests desired by our clients. Some demands were on the table for quite a while, others were nudged and introduced to clients through concepts created by Preserve-or-Forget such as value annotated to information through metadata.

7 Exploitation Plans

The following section is taken out from D11.5. Due to the nature of this delivery we thought it would be beneficial to include our exploitation ideas here.

dkd's current business model is centered around the open source content management system TYPO3. Most of the services offered are tailored to the customers seeking to use this CMS solution to drive their online and e-commerce activities. Therefore the technical outcome of the WP10 work in the ForgetIT project will be used to expand current customer solutions. This includes the offer of CMIS-based digital asset management solutions, a semantic annotation platform offering and a data preservation consultancy.

Being a small business can only gain market awareness and leave a footprint by inviting other companies to use our open sourced technologies. We gave out information about our developments during our dissemination activities throughout the project. It will continue to inform and educate interested third parties during the next year.

At dkd we communicated our insights to the internal project-teams. This empowered them to integrate findings in solutions offered. Also, we organize internal workshops with teams and management to brainstorm product and service ideas. We also plan to revisit a survey conducted in year one of the project. We will emend sections through more accurate questions.

Our exploitation ideas are centered around our expertise in the field of Content Management Systems. During the work within the ForgetIT Project we have gained farther insights into organizational requirements and encountered ever broader tasks. We developed and offered solutions for challenges, but at the same time learned our limitations due to our limited resources. We focused on what was feasible and accordingly offered adequate consulting and integration concepts. Our approach is three tiered:

- development and contribution to technical base components and libraries
- software solutions and integrations
- consulting on organizational digital data strategies

7.1 Technology-related Exploitation

The following table summarizes our achievements in terms of development and contributions to base components and libraries. This includes list of libraries and TYPO3 Extensions, their description, possible users, actions and next steps.

7.1.1 PHP & TYPO3

Component	Description	Audience	Action & Next Steps
PHP CMIS Client	PHP CMIS Client is a port of OpenCMIS (Java) to PHP. Interfaces are mostly the same so most OpenCMIS examples should be also usable for this PHP CMIS Library.	CMIS developers that need to integrate a CMIS Solution with their PHP based software.	Contacted the Apache Chemistry Community in order to donate the code base to the Community.
TYPO3 Extension CMIS Service	This TYPO3 Extension integrates the PHP CMIS Library into TYPO3 CMS so that other TYPO3 can use the CMIS functions.	TYPO3 Developers seeking integration of CMIS compliant solutions with TYPO3.	Availability on Github. Release on the TER (TYPO3 Extension Repository) planned.
TYPO3 Extension CMIS FAL	A TYPO3 CMS file abstraction layer driver for CMIS	Developers and Integrators that want to connect a CMIS repository to TYPO3 so that it displays digital assets stored via CMIS.	Availability on Github. Release on the TER planned.
TYPO3 Extension Annotate	A TYPO3 Extension that adds semantic annotation with RDFa using GATE project technology	Developers and Integrators that want to enrich their Content with semantic annotations	Availability on Github. Release on the TER planned.
TYPO3 Extension Content Dashboard	A TYPO3 Extension that displays content that was delivered to the PoF Framework.	Developers and Integrators seeking to use the PoF Framework in their TYPO3 Installation.	Demo the current state to a set of clients. Presentation at TYPO3 Events. Webinars & Video Screencasts to explain use and advantages.
TYPO3 Extension CMIS Client	TYPO3 Extension that offers CMIS Client functionality.	Developers and Integrators.	

7.1.2 Docker Files

Component	Description	Audience	Action & Next Steps
docker-alfresco-typo3	Docker files providing an alfresco instance configured to work with TYPO3 CMS.	TYPO3 Developers in need of a preconfigured development and production environment.	Maintenance and support of the provided docker image via Github and Dockerhub.
docker-typo3	Docker based TYPO3 Installation based on composer	TYPO3 Developers in need of a preconfigured development and production environment	Maintenance and Support of the provided docker image via Github and Dockerhub
docker-mimir	Docker Mimir Image for ForgetIT	TYPO3 Developers in need of a preconfigured development and production environment	Maintenance and Support of the provided docker image via Github and Dockerhub
forgetit-fish-example	This creates a docker setup so that we can test our ForgetIT functionality	TYPO3 Developers and ForgetIT project partners	Maintenance and further development

7.1.3 GATE

Component	Description	Audience	Action & Next Steps
forgetit-gate-anti-dup	duplicate remover for GATE	GATE technology users	Contribution to the GATE Project proposed
forgetit-gate-rdfa-import	RDFa importer for GATE	GATE technology users	Contribution to the GATE Project proposed
forgetit-gate-rdfa-export	RDFa Exporter for GATE	GATE technology users	Contribution to the GATE Project proposed

7.1.4 Possible software solutions based around the technologies and use cases elaborated in the ForgetIT project

We have identified the following software solutions while working with ForgetIT:

- (A) TYPO3 CMS preservation service for companies using TYPO3
- (B) TYPO3 CMS content platform and digital asset management via the FAL and CMIS extension
- (C) TYPO3 CMS semantic annotation service
- (D) TYPO3 CMS semantic enhanced search

7.1.5 Business Canvases

Common building blocks for the mentioned solutions mentioned above are following the Business Canvas model as shown in Figure 11. All filled out canvases can be found in Appendix A.

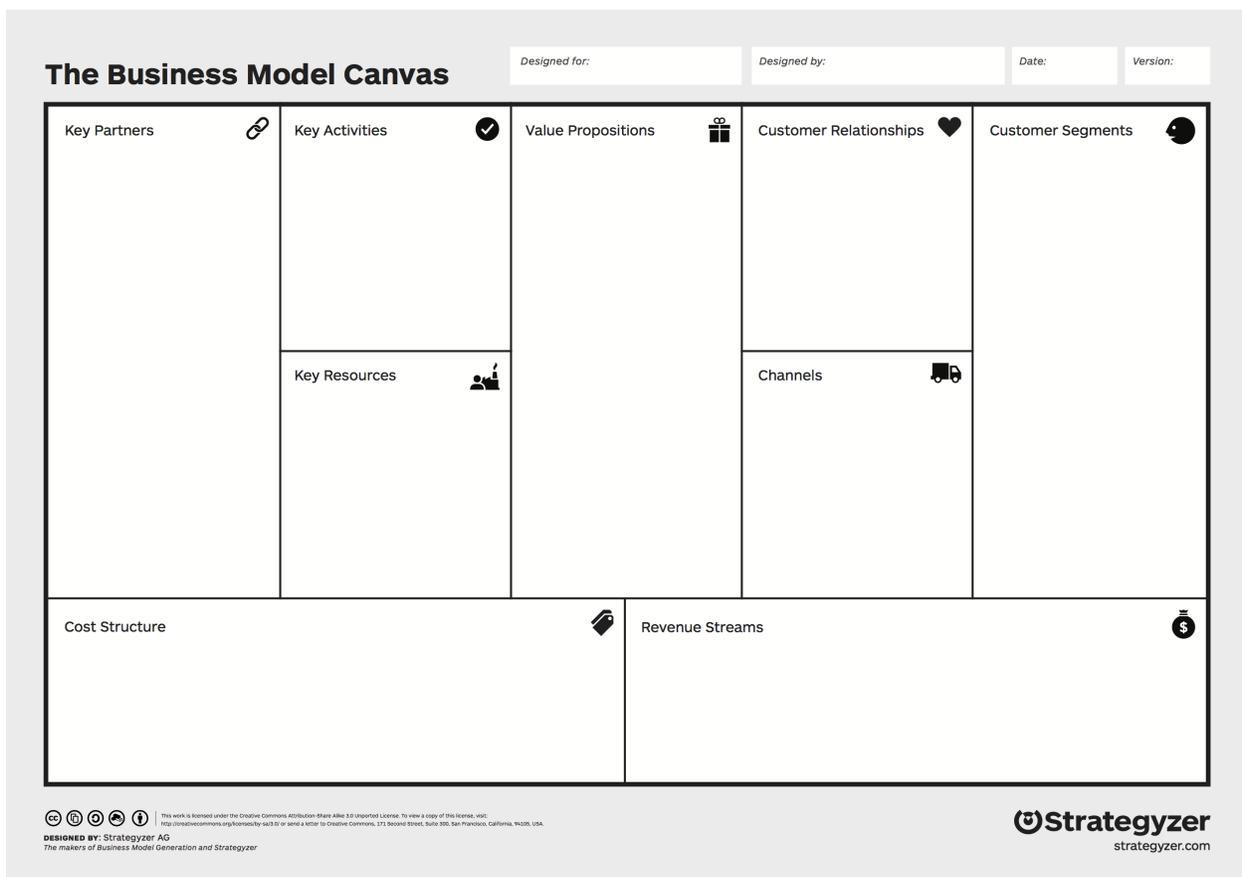


Figure 11: Business Canvas Model

Customer Segments

- organizations using TYPO3 CMS to build their website (end users) (A, B, C and D)
- organizations providing TYPO3 CMS integration and development services (technical agencies)(A, B, C and D)
- the market is a niche in the DACH Region

Value Proposition

- help organizations to preserve their TYPO3 CMS based website (A)
- help organizations to collect and distribute content and assets over several TYPO3 CMS and web applications using the CMIS standards (B)
- add semantic and concept annotation to the TYPO3 CMS based website of organization in order to gain more traffic from search engines and also enhance the daily work of the organization while creating new content (C)
- create new ways to present and navigate a website using semantic entities and concept detection (D)

Channels

- website informing on the offering (A, B, C and D)
- technology partners such as other TYPO3 Agencies (A, B, C and D)
- social media and online marketing (A, B, C and D)
- webinars (A, B, C and D)
- conference talks (A, B, C and D)
- white papers (A, B, C and D)
- press work (A, B, C and D)

Customer Relations

- technical support (A, B, C and D)
- sales and marketing support (A, B, C and D)
- end user support via ticket system such as Zendesk (A, B, C and D)
- newsletters and social media (A, B, C and D)

- self service platform such as an online shop (A, B, C and D)

Revenue Streams

- income based on monthly usage fees from a SaaS offering (A, B, C and D)
- service level agreements for end users and partners (A, B, C and D)
- volume or transaction based pricing (A, B, C and D)
- possible licensing or certification fees (A, B, C and D)

Key Resources

- cloud based infrastructure (A, B, C and D)
- application technology / ForgetIT components (A, B, C and D)
- software engineers and systems architects (A, B, C and D)
- support personnel (A, B, C and D)
- sales and marketing (A, B, C and D)
- management (A, B, C and D)
- possible external funding as a startup (A, B, C and D)

Key Activities

- production of the software components (A, B, C and D)
- platform operation (A, B, C and D)
- support and consulting (A, B, C and D)

Key Partnerships

- technology partners providing external components such as CERTH (A, B, C and D)
- cloud platform partners such as Amazon, Google or Microsoft (A, B, C and D)
- TYPO3 association

Cost Structure

- cloud based infrastructure
- development costs based on FTE and external workforce

7.2 Consulting towards an organizational digital data strategy

As described in D10.3 we have identified the need of organizations for a digital data strategy. Our approach is to create a consulting approach that is open and based on open source tools.

We have gathered the following key insights that describe our approach:

- All organisations are different. They can not be compared easily with each other.
- The amount of digital data in organisations differs, but we assume that it will increase overall.
- Dealing with digital data needs a dedicated strategy approach.
- Digital data is a fundamental factor for an organisation's success.
- We derive all our insights on the actual doing (the IS) of an organisation, not what a mission statement requires (the OUGHT).
- We will create a catalog of "action items", that an organization can put into action with a reasonable amount of time and effort.
- Our model helps us estimate an organization's digital data readiness.
- It will help the organization to identify its data creation and usage needs.
- Our toolset consists mainly of publicly available, standard-compliant technologies.

7.3 Next Steps to be taken

The ForgetIT team at dkd takes the following steps to exploit above mentioned ideas and possibilities:

- Collection of the status of the software components created by dkd and the ForgetIT partners based on the commercially and operational fitness.
- Presentation of the dkd created technologies to the management of dkd and technical key personnel in order to identify interests to maintain and develop those components.

- Management presentation of the products ideas described in the business canvas.
- Refining of the business canvas and first rough business plan.
- Seek internal or external financing.
- Creation of a dedicated business unit to start the business operations.

We can conclude our dkd chapter of exploitation ideas saying that there is a wide range of opportunities for us. As this is our first participation in such a project, we cannot predict which idea is the most likely to be successful.

We aim to lead and follow the technical development in our core market TYPO3-based websites. We see a high value in software-as-a-service offers. We compare efforts here with our Apache Solr based service “Hosted Solr”. As we see our main business change from web content management to digital based experience applications, we see an entrepreneurial chance in strategic consulting needed by organizations in the coming years.

8 Conclusions

Within the ForgetIT project, dkd developed three new services based on project results:

1. a Content Management System (CMS) content platform and digital asset management via the File Abstraction Layer (FAL) and Content Management Interoperability Services (CMIS) extensions
2. a prototype for a TYPO3 CMS service that adds additional annotations to text and photos based on technology developed by project partners USFD and CERTH, and provides several ways of searching the added concepts
3. a preservation service based on an interface between the newly developed content platform and the Preserve-or-Forget Framework.

Most of the project effort was dedicated to establishing stable, reliable internal infrastructure. Our long-term engagement with two very different companies, the large holding ACO and the smaller specialist fair operator Spielwarenmesse AG, shows that short-term benefits are key when introducing new functionality into companies. More efficient content management delivers those benefits and lays the necessary groundwork for packaging content and sending it to a PoF Framework.

A key selling point of the new CMIS library, `php-cmis` is the level of interoperability that it allows. Lack of interoperability is the bane of many content providers who need to publish to several different platforms, covering both platforms that are popular at the moment, such as Facebook, and platforms that may emerge in future. Through its Alfresco integration, `php-cmis` is already ahead of the competition, and provides easy access to dynamic content such as Google Documents.

Both the companies and the beta testers who used a preliminary version of the annotation interface saw the clear benefits of the extensive automatic multimodal annotation capabilities that are integrated into the current ForgetIT solution.

Once companies are using ForgetIT-enabled content management systems, the door is open to implementing archiving solutions that showcase the full benefit of managed forgetting. However, until then, there is a clear need to increase awareness of what preservation is, and how companies can benefit from it.

8.1 Assessment of Performance Indicators

As Table 3 shows, almost all of the main performance indicators were completed successfully, even though some milestones were reached later than anticipated.

The open source TYPO3 extension has been taken up both by the TYPO3 community and the wider free software / open source community. As the extension has not been

Table 3: Performance Indicators

Outcome	Indicator	
1) Documented organizational archive and presentation needs	persona models available	D10.1
	user stories available	D10.1/D10.3
2) TYPO3 pilot application	list of features implemented in pilot	D10.2-D10.4
	list of features from other work packages integrated	D10.2-D10.4
3) Open source TYPO3 extension	uptake by TYPO3 community	D10.4
4) evaluation	completed design number of users (formative) long-term consultations with companies insights into applying ForgetIT to organisational preservation	D2.4 10; D10.4 2; D10.3 D10.4

deployed in a full working system yet, it was impossible to perform an appropriate summative evaluation. However, given the community's interest, we expect that the extension will have a life well beyond the end of the project.

Since most of the development resulted in developer-side infrastructure, which is key for the further uptake of ForgetIT-based organisational preservation solutions, the user interface components were ready relatively late in the project. Therefore, the user evaluation was more informal than that conducted for WP9 and focused on providing formative feedback.

8.2 Vision for the Future

We have already detailed our exploitation plans in detail in Section 7.1. The main library, `php-cmis`, is well on the way to being accepted widely in the free software community. This provides us with a clear pathway for advancing the preservation agenda. We will continue to raise awareness of preservation in the free software community. Through our professional services, we will provide an integrated pathway for interested companies who wish to parlay existing, initial ForgetIT-derived infrastructure into a full preservation solution. Our approach, developed in conjunction with partner UOXF, combines a principled assessment of the IS versus the OUGHT with a straightforward implementation based on

ForgetIT concepts. This will ensure long-term impact.

This approach to professional services also extends the framework within which dkd operates. Projects typically consist of a plan-build-run cycle. Within this cycle, dkd has been specialising on the “build” and “run” aspects. Through the lessons learned from ForgetIT, the company is now well positioned to extend its services during the “plan” phase through innovative, evidence-based consulting.

List of Figures

1	Use Case Overview	11
2	CMIS Manager: Queue	13
3	CMIS Manager: Tables	14
4	Content Dashboard	18
5	Annotation Component Overview	20
6	Annotation Pipeline	21
7	Automatic Annotation	22
8	Detected Concepts	26
9	Low Image Quality	27
10	SUS scores for TYPO3 and the Annotation Tool	36
11	Business Canvas Model	45
12	Canvas A	54
13	Canvas B	55
14	Canvas C	56
15	Danvas D	57

A Business Canvases

The Business Model Canvas

Version: 1.0

Date: Jan 16

Designed by: dkd

Designed for: dkd

<p>Key Partners technology partners providing external components such as CERN cloud platform partners such as Amazon, Google or Microsoft TYPO3 association</p>	<p>Key Activities production of the software components platform operation support and consulting</p>	<p>Value Propositions help organizations to preserve their TYPO3 CMS based website (A)</p>	<p>Customer Relationships technical support sales and marketing support end user support via ticket systems newsletter and social media self service platform such as an online shop</p>	<p>Customer Segments organizations using TYPO3 CMS to build their website (end users) organizations providing TYPO3 CMS integration and development services (technical agencies) the market is a niche in the DACH Region</p>
<p>Key Resources cloud based infrastructure application technology / ForgetIT components software engineers and systems architects support personnel sales and marketing management possible external funding as a startup</p>	<p>Channels website informing on the offering technology partners such as other TYPO3 Agencies social media and online marketing webinars conference talks white papers</p>	<p>Revenue Streams income based on monthly usage fees from a SaaS offering service level agreements for end users and partners volume or transaction based pricing possible licensing or certification fees</p>	<p>Cost Structure cloud based infrastructure development costs based on FTE and external workforce</p>	



DESIGNED BY: Strategyzer AG
The makers of Business Model Generation and Strategyzer
This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit: <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Figure 12: Canvas A

The Business Model Canvas

Version: 1.0

Date: Jan 16

Designed by: dkd

Designed for: dkd

<p>Key Partners technology partners providing external components such as CEKTH cloud platform partners such as Amazon, Google or Microsoft TYPO3 association</p>	<p>Key Activities production of the software components platform operation support and consulting</p>	<p>Value Propositions help organizations to collect and distribute content and assets over several TYPO3 CMS and web applications using the CMS standards (B)</p>	<p>Customer Relationships technical support sales and marketing support end user support via ticket systems newsletter and social media self service platform such as an online shop</p>	<p>Customer Segments organizations using TYPO3 CMS to build their website (end users) organizations providing TYPO3 CMS integration and development services (technical agencies) the market is a niche in the DACH Region</p>
<p>Key Resources cloud based infrastructure application technology / ForgetIT components software engineers and systems architects support personnel sales and marketing management possible external funding as a startup</p>	<p>Channels website informing on the offering technology partners such as other TYPO3 Agencies social media and online marketing webinars conference talks white papers</p>	<p>Revenue Streams income based on monthly usage fees from a SaaS offering service level agreements for end users and partners volume or transaction based pricing possible licensing or certification fees</p>	<p>Cost Structure cloud based infrastructure development costs based on FTE and external workforce</p>	



This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit: <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

DESIGNED BY: Strategyzer AG
 The makers of Business Model Generation and Strategyzer

Figure 13: Canvas B

The Business Model Canvas

Version: 1.0

Date: Jan 16

Designed by: dkd

Designed for: dkd

<p>Key Partners technology partners providing external components such as CERN cloud platform partners such as Amazon, Google or Microsoft TYPO3 association</p>	<p>Key Activities production of the software components platform operation support and consulting</p>	<p>Value Propositions add semantic and concept annotation to the TYPO3 CMS based website of organization in order to gain more traffic from search engines and also enhance the daily work of the organization while creating new content (C)</p>	<p>Customer Relationships technical support sales and marketing support end user support via ticket systems newsletter and social media self service platform such as an online shop</p>	<p>Customer Segments organizations using TYPO3 CMS to build their website (end users) organizations providing TYPO3 CMS integration and development services (technical agencies) the market is a niche in the DACH Region</p>
<p>Key Resources cloud based infrastructure application technology / ForgetIT components software engineers and systems architects support personnel sales and marketing management possible external funding as a startup</p>	<p>Channels website informing on the offering technology partners such as other TYPO3 Agencies social media and online marketing webinars conference talks white papers</p>	<p>Revenue Streams income based on monthly usage fees from a SaaS offering service level agreements for end users and partners volume or transaction based pricing possible licensing or certification fees</p>	<p>Cost Structure cloud based infrastructure development costs based on FTE and external workforce</p>	



DESIGNED BY: Strategyzer AG
The makers of Business Model Generation and Strategyzer
This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit: <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Figure 14: Canvas C

The Business Model Canvas

Designed for: **dkd**

Designed by: **dkd**

Date: **Jan 16**

Version: **1.0**

<p>Key Partners technology partners providing external components such as CEKTH cloud platform partners such as Amazon, Google or Microsoft TYPO3 association</p>	<p>Key Activities production of the software components platform operation support and consulting</p>	<p>Value Propositions create new ways to present and navigate a website using semantic entities and concept detection (D)</p>	<p>Customer Relationships technical support sales and marketing support end user support via ticket systems newsletter and social media self service platform such as an online shop</p>	<p>Customer Segments organizations using TYPO3 CMS to build their website (end users) organizations providing TYPO3 CMS integration and development services (technical agencies) the market is a niche in the DACH Region</p>
<p>Key Resources cloud based infrastructure application technology / ForgetIT components software engineers and systems architects support personnel sales and marketing management possible external funding as a startup</p>	<p>Channels website informing on the offering technology partners such as other TYPO3 Agencies social media and online marketing webinars conference talks white papers</p>	<p>Revenue Streams income based on monthly usage fees from a SaaS offering service level agreements for end users and partners volume or transaction based pricing possible licensing or certification fees</p>	<p>Cost Structure cloud based infrastructure development costs based on FTE and external workforce</p>	



DESIGNED BY: Strategyzer AG
 The makers of Business Model Generation and Strategyzer
This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit: <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Figure 15: Danvas D