



# Managed Forgetting, Data Condensation & Preservation in Application

Christian Jilek, Heiko Maus, Sven Schwarz, Andreas Dengel

Knowledge Management Department  
German Research Center for Artificial Intelligence (DFKI) GmbH

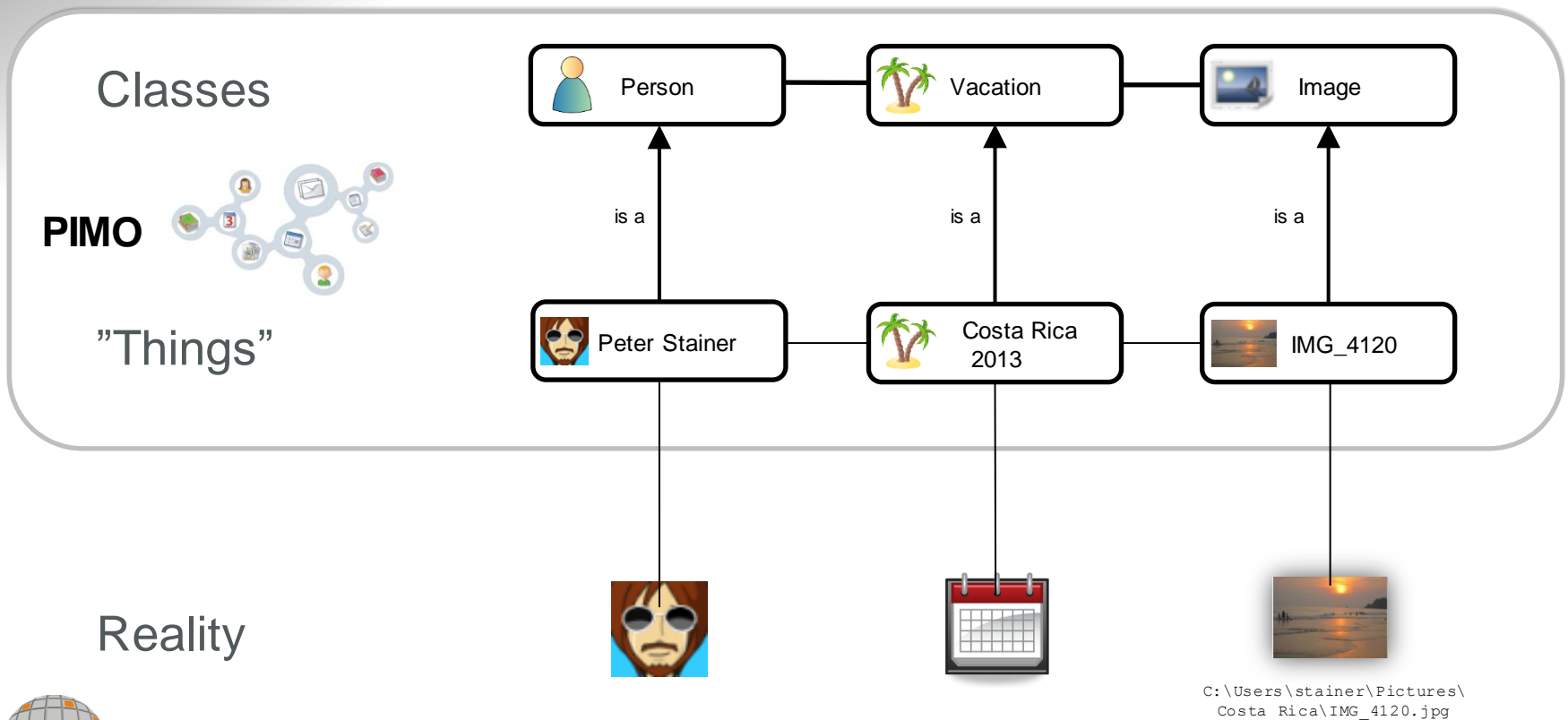


# The Semantic Desktop & PIMO

## Basics



PIMO represents the **user's mental model** as **vocabulary for applications** without confronting users with formal knowledge representation



# The Semantic Desktop & PIMO

## Dedicated Apps & Plug-ins



ForgetIT Workshop Istanbul, 2014

- Agenda of Istanbul meeting by Claudia

### Session Review Recommendations

- be aware of the indicators
- ToDo: contribution to WP4 deliverable D4.3 must be done by Bahaa
- ToDo: think of showing a demo / video or similar for WP9
  - end of next link to material for WP9 PIMO usage
  - idea: start with the D9.2 Mock-up slide
- RR5: context discussion
  - WP10 as a lightweight approach better
  - WP9 has richer ontology, however v
- ToDo: some more content for Twitter on Facebook
- ToDo: collaboration meetings: report on the

**PIMO Software**

Also known as: Personal information Model

A Personal Information Model (PIMO) is used to represent a single users' concepts from his mental model, such as projects, tasks, contacts, organizations, allowing files, e-mails, and other resources of

Preservation Workflow

- question: how does PoF recognizes the collection
  - PIMO allows to have a collection, e.g.
- two alternatives:
  - the PoF is aware of a collection and if one photo of this collection is to be preserved, then the PoF would request the whole collection
  - the PoF is not aware of the interrelations: if it wants to preserve the photo...the active system would hand over the archival unit...i.e.

Home. ForgetIT Project - Mozilla Firefox

www.forgetit-project.eu

Suchen | Aufwärts | Hervorheben | Meistbesucht | New Ticket | Springer für Profession...

**Pimo Annotation**


ForgetIT Project Home...

Search for concepts

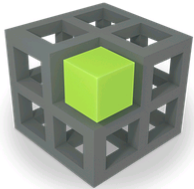
Name	
Information M...	X
forgetting	X
remembering	X
EU	X
ForgetIT	X

**Inferred Concepts**  
**Found Concepts**

Attention	+
Digital Preserv...	+
(js)JavaScript	+
Preserve-or-Fo...	+
multimedia	+
PHP	+
preservation	+
summarization	+
context	+
social media	+



# ForgetIT



ForgetIT

While preservation of digital content is now well established in memory institutions such as national libraries and archives, it is still in its infancy in most

## Semantic Editor (SEED)

[dedicated app]

## FireTag for Mozilla Firefox

[plug-in]



# Managed Forgetting

## Memory Buoyancy → item's short-term value



during workshop

... after 8 months

... after 2 years

ForgetIT WS Luleå 2014

agents

- ForgetIT Advisory Board
- Elaine Niven
- Berker Loğoğlu
- Parvaneh Afrasiabi Rad
- Jørgen Nilsson
- Göran Lindqvist
- Sven Schwarz
- Mari Wolters
- Heiko Maus
- Viktor Mayer-Schönberger

jobs

- ForgetIT
- ForgetIT Meeting in Luleå 2014
- Reiseabrechnung Luleå 2014
- WP9 Presentation for Advisory Board
- Reiseantrag Luleå 2014
- WP9 in Luleå: Status, next steps, review, mock

locs

- Luleå
- Elite Stads hotell i Luleå

media

- forgetit/project\_meeting\_lulea\_L35 Wiki
- Luleå-WP9-AdvisoryBoard-2014.pptx
- Luleå Flugticket 2014
- Minutes of ForgetIT General Assembly Meeting
- notizen in lulea
- [ForgetIT-sl@ WP11 Presentation
- Logistic informations for the ForgetIT Meeting
- [Forgetit-wp9leader] Revised Agenda
- [Forgetit-research] ForgetIT General Assembly

ForgetIT WS Luleå 2014

agents

- ForgetIT Advisory Board
- Elaine Niven
- Berker Loğoğlu
- Parvaneh Afrasiabi Rad
- Jørgen Nilsson
- Göran Lindqvist
- Sven Schwarz
- Heiko Maus
- Viktor Mayer-Schönberger

media

- Luleå-WP9-AdvisoryBoard-2014.pptx

ForgetIT Workshop Luleå 2014

persons

- Søren Schaffstein
- Elaine Niven
- Sven Schwarz
- Heiko Maus
- Olivier Dobberkau
- Mark Greenwood
- Maria Wolters
- Vasilis Mezaris

events

- Fotos ForgetIT Workshop Lulea, Schweden

jobs

- ForgetIT

show forgotten ↑ MB 0.5 ↓

ForgetIT Workshop Luleå 2014

persons

- Søren Schaffstein
- Elaine Niven
- Sven Schwarz
- Heiko Maus
- Olivier Dobberkau
- Mark Greenwood
- Maria Wolters
- Vasilis Mezaris

events

- Fotos ForgetIT Workshop Lulea, Schweden

jobs

- ForgetIT

show forgotten ↑ MB 0.5 ↓

access to forgotten information.

# Data Preservation

## Preservation Value → item's long-term value



[Login](#) [Options](#) **Preservation**

---

[Contract](#) **Strategy**

Preservation Strategy Preset: [File & Forget Curator](#)

**INVESTMENT**

- number of annotations (things having a higher number of annotations are more likely to be preserved)
- wikitext length (the longer a thing's wikitext, the more likely it will be preserved)
- usage (the more frequently a resource is modified, the more likely it will be preserved)

**GRAVITY**

- connectivity (the more a thing is related to other things, the more likely it will be preserved)
- type-based heuristic (certain things such as contracts are more likely to be preserved)
- important projects (the higher the number of person involved in a project, the more likely it will be preserved)
- closeness to important things (things related to tasks or events are more likely to be preserved)

**SOCIAL GRAPH**

- important persons (the higher the number of projects a person is involved in, the more likely it will be preserved)
- PIMO user on photo (photos containing PIMO users are more likely to be preserved)

**POPULARITY**

- image rating (the higher an image's rating, the more likely it will be preserved)
- number of views (the more a thing is accessed/viewed, the more likely it will be preserved)

**COVERAGE**

- cover photo collections (at least one photo of each photo collection should be preserved)

**QUALITY**

- image quality (high quality images are more likely to be preserved)

[save settings & recalculate](#)

To check the results of these preservation settings: [show preservation overview](#)

**Maus07: Workflow-Kontext zur Realisierung prozessorientierter Assistenz in Organisational Memories**

- Maus07.pdf

**Interview Ergebnis Fragen für supSpaces**

- Interview-Ergebnis\_Fragen für supSpaces v3.0.docx

**WoltersNivenRunardotter+15: Personal Photo Preservation for the Smartphone Generation**

- wip362-woltersA-w.pdf

PV=0.900164008140564  
 Explanation:  
 connectivity=57 => 0.501  
 #annotations=12 => 0.339  
 usage: 5 months in 8 months => 0.303  
 wikitextLength=3131 => 0.829  
 typeBonus: Document => 0.1  
 closeToImportantThing=true => 0.5

**PV=0.900164008140564**  
**Explanation:**  
 connectivity=57 => 0.501  
 #annotations=12 => 0.339  
 usage: 5 months in 8 months => 0.303  
 wikitextLength=3131 => 0.829  
 typeBonus: Document => 0.1  
 closeToImportantThing=true => 0.5



# Data Condensation

## Overview of PIMO Diary



### My Diary: February 2014



03.02.2014 - 19.02.2014  
 ForgetIT / ForgetIT WS Luleå 2014  
 ForgetIT WS Luleå 2014  
 Reiseabrechnung Lulea 2014  
 forgetit:project\_meeting\_lulea · L3S Wiki  
 and 12 more

Keywords:  
 forgetit 2014 implementing interface calls meeting luleå lulea cmis preservation haas design manual presentation re wp11 werner wiki minutes general assembly johannes goslar forgetit:project\_meeting\_lulea l3s mydfki dienstreiseverwaltung update team luleat.jpg logistic informationa flugticket ws reiseabrechnung

### 14.02.2014 - 28.02.2014

#### RSIP2-SD / Discuss final review presentation

- Discuss final review presentation
- Datenschutzvereinbarung
- Datenschutzvereinbarung für RSIP

and 7 more

Keywords:  
 review final rsip datenschutzvereinbarung meeting presentation discuss prototype new rsip3 re fwd rechnung project leadership developer agenda demonstration rechnung ricoh invoice rsip2 sd last period heiko system slides updates rsip1 fs remember two demonstrations done pls christian showing make sure

Start: 01.02.2014 00:00:00  
 End: 28.02.2014 23:59:59  
 Time granularity: month  
 Discard least important: true  
 Data coverage: 34%

**Detail Settings** Expert Settings

#### Detail Settings

Sorting: importance  
 Include shared: false  
 Entries desired: 10  
 Entries displayed: 10  
 Source items available: 203  
 Source items included: 70  
 Source items excluded: 133  
 Query duration: 849 ms  
 Diversity: + -  
 Rare Things: + -  
 Rich Media Things: + -

#### Context

- ForgetIT
- PIMO
- use case



# Conclusion & Outlook



Some aspects investigated in the ForgetIT project:

- **Memory Buoyancy** → item's short-term value
- **Preservation Value** → item's long-term value
- **organize information accordingly** (i.e. rank/filter/hide/delete/sync)
- **data condensation** driven by contexts

- SD→LL:
- use SD technology for **sense making** of LL data
  - **better handle massive LL data** using automated condensation and forgetting technologies already realized in the SD for PIM

- LL→SD:
- SD could benefit from **more detailed sensory support**
  - SD could benefit from lessons learned when **dealing with massive data** (especially in form of a **continuous stream**)

Our recommendation for both, LL & SD:  
**focus on a more explicit handling of contexts**





# Thanks for your attention!



## Any questions?

The work presented was partially funded by the **European Commission** in the context of the **FP7 ICT project ForgetIT** (grant no. 600826)



and by the **German Federal Ministry for Education and Research** in the **project supSpaces** (grant no. 01IS15013B).

