Diary Generation from Personal Information Models to Support Contextual Remembering and Reminiscence

Christian Jilek, Heiko Maus, Sven Schwarz, Andreas Dengel

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

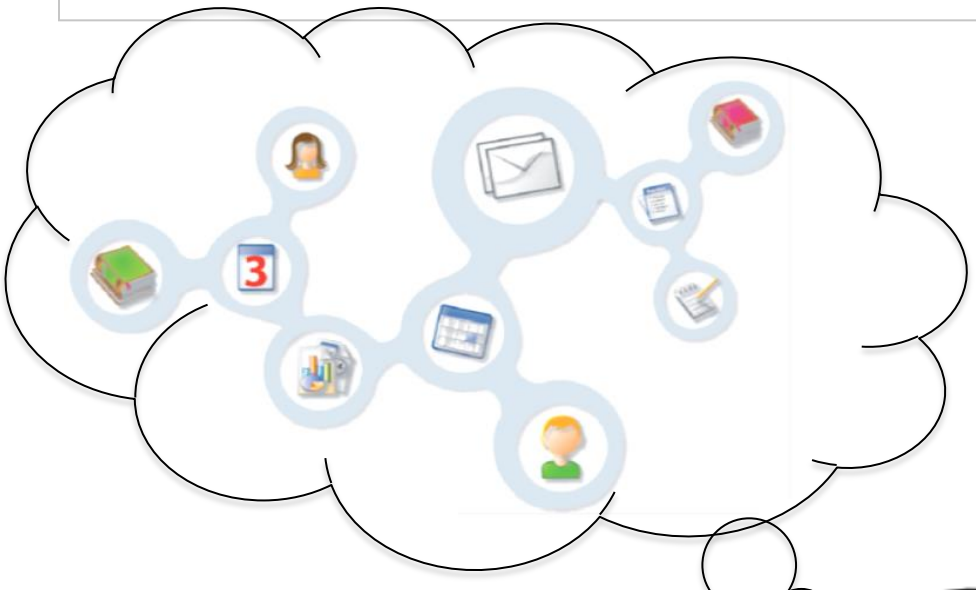
Contents

Can you name **five things**

you were **concerned with the most**

for an **arbitrarily chosen period of your life**,

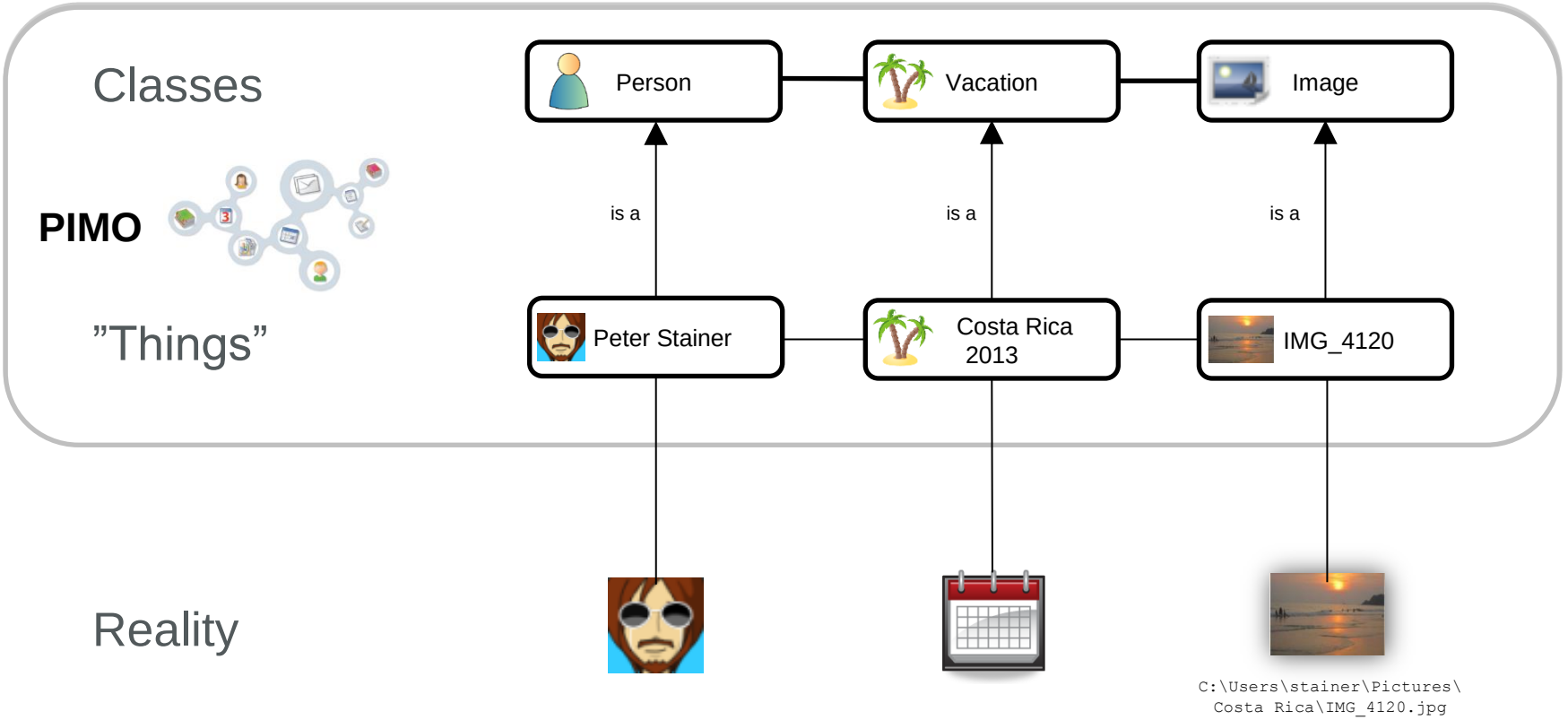
e.g. September 2008 or spring 2003 ?



„a diary that writes itself“

on-demand diary generation from personal information models

PIMO represents the user‘s mental model as vocabulary for applications w/o confronting users with the formal knowledge representation

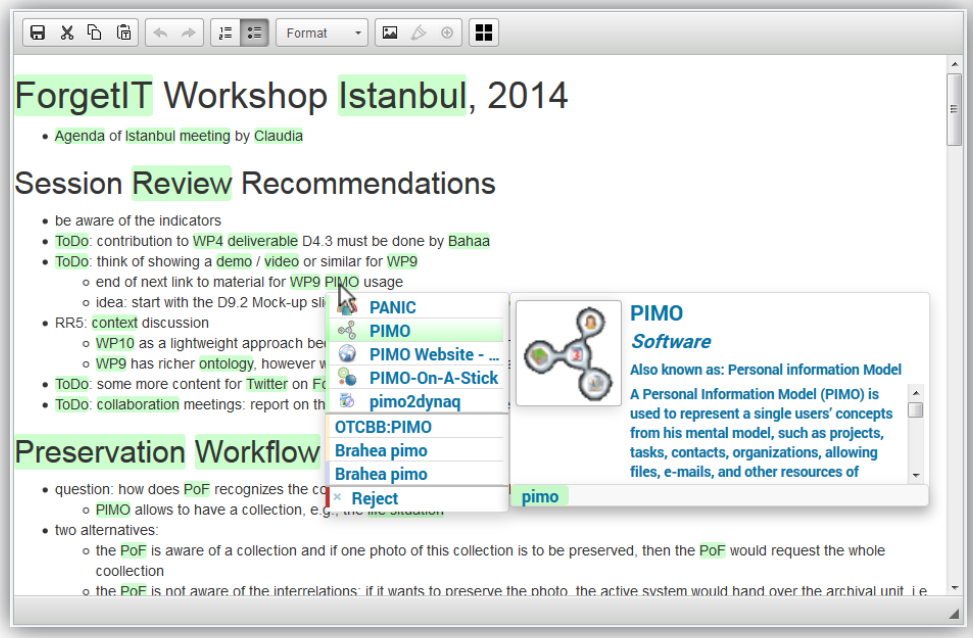


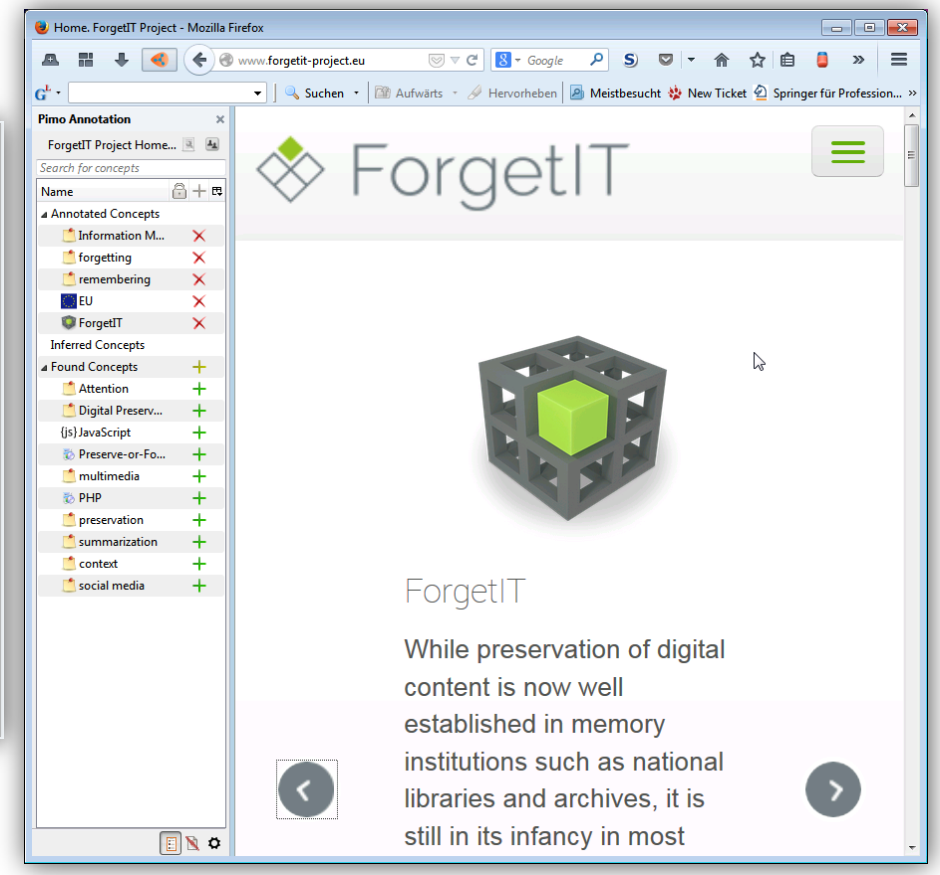
Ingredients of the Semantic Desktop Infrastructure

Dedicated PIMO Apps

Plug-in to (office) programs

Examples of Semantic Desktop Applications



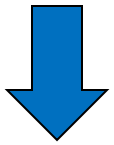


Semantic Editor (SEED)
[dedicated app]

FireTag for Mozilla Firefox
[plug-in]

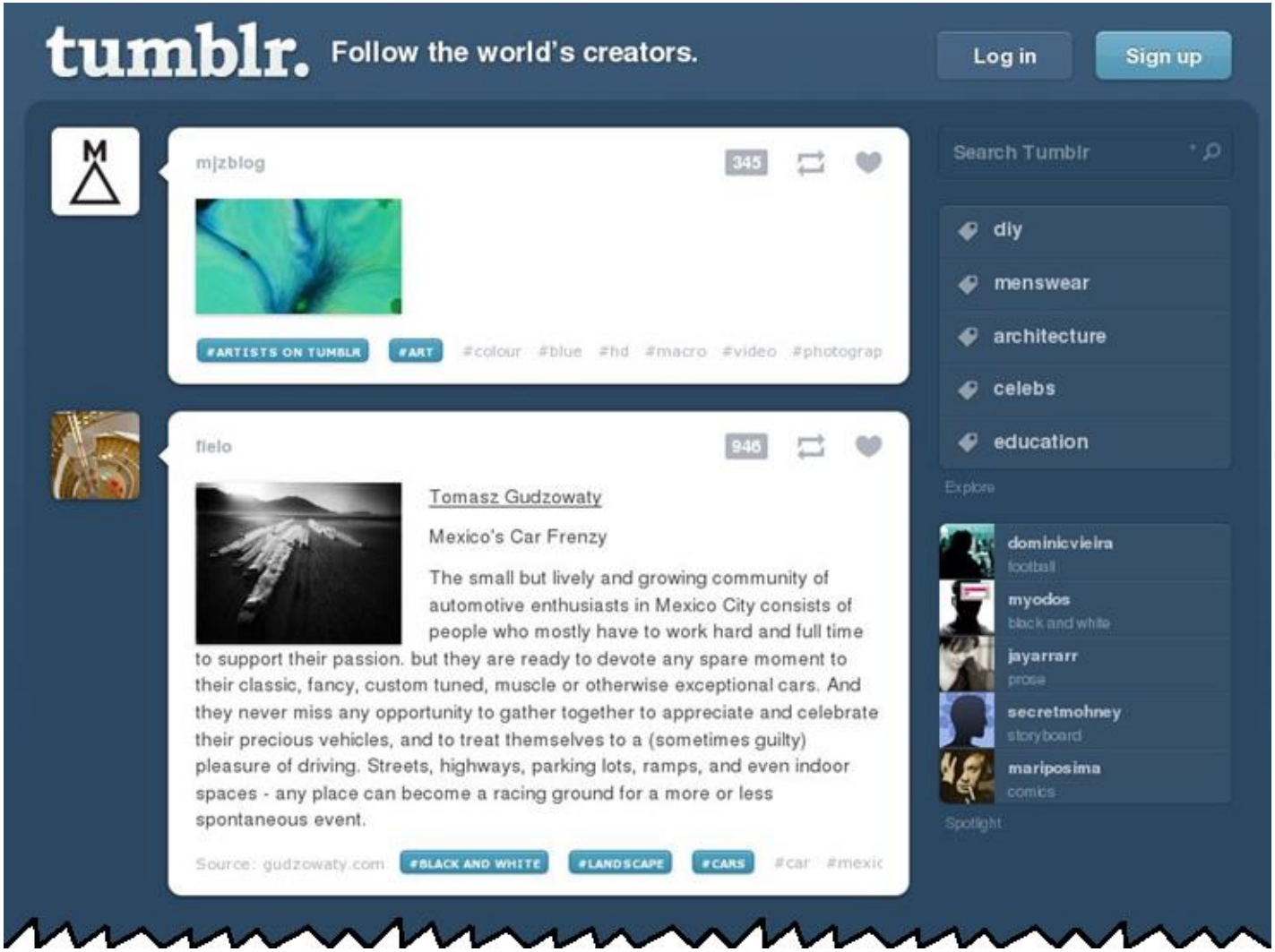
Vision (cont'd): Diversity within the Diary

time	diary entry		
CW 42	I was working at John Deere.		
CW 43	I was working at John Deere.	I attended the GSE lecture.	I attended the MMKI lecture.
CW 44	I attended the GSE lecture.	I attended the MMKI lecture.	
CW 45	I was working at John Deere.	I attended the GSE lecture.	I attended the MMKI lecture.
CW 46	I attended the GSE lecture.	I attended the MMKI lecture.	I attended the GRA lecture.
CW 47	I attended the GSE lecture.	I attended the MMKI lecture	I attended the GRA lecture.
CW 48	I attended the GSE lecture.	I attended the MMKI lecture.	
CW 49	I attended the GSE lecture.	I attended the MMKI lecture.	
CW 50	I attended the GSE lecture.	I attended the MMKI lecture.	I attended the GRA lecture.
CW 51	I attended the GSE lecture.		



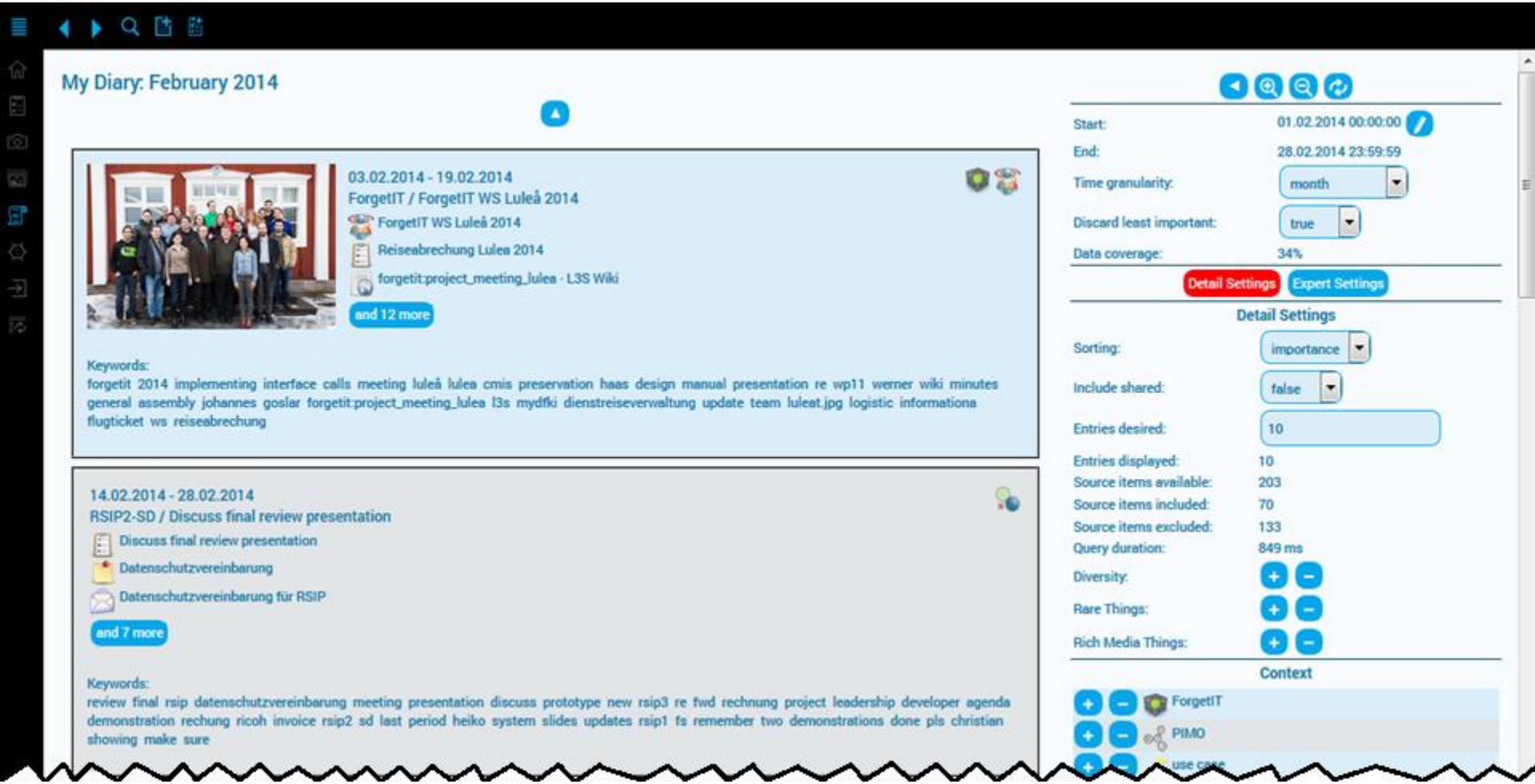
time	diary entry
CW 42-45	I was working at John Deere.
CW 43-50	I attended the MMKI lecture.
CW 43-51	I attended the GSE lecture.
CW 46-50	I attended the GRA lecture.

Vision (cont'd): Modern Look & Feel



screenshot by *www.web-rater.com*, 2014

PIMO Diary: User Interface



Structure of a Diary Entry

date and headline

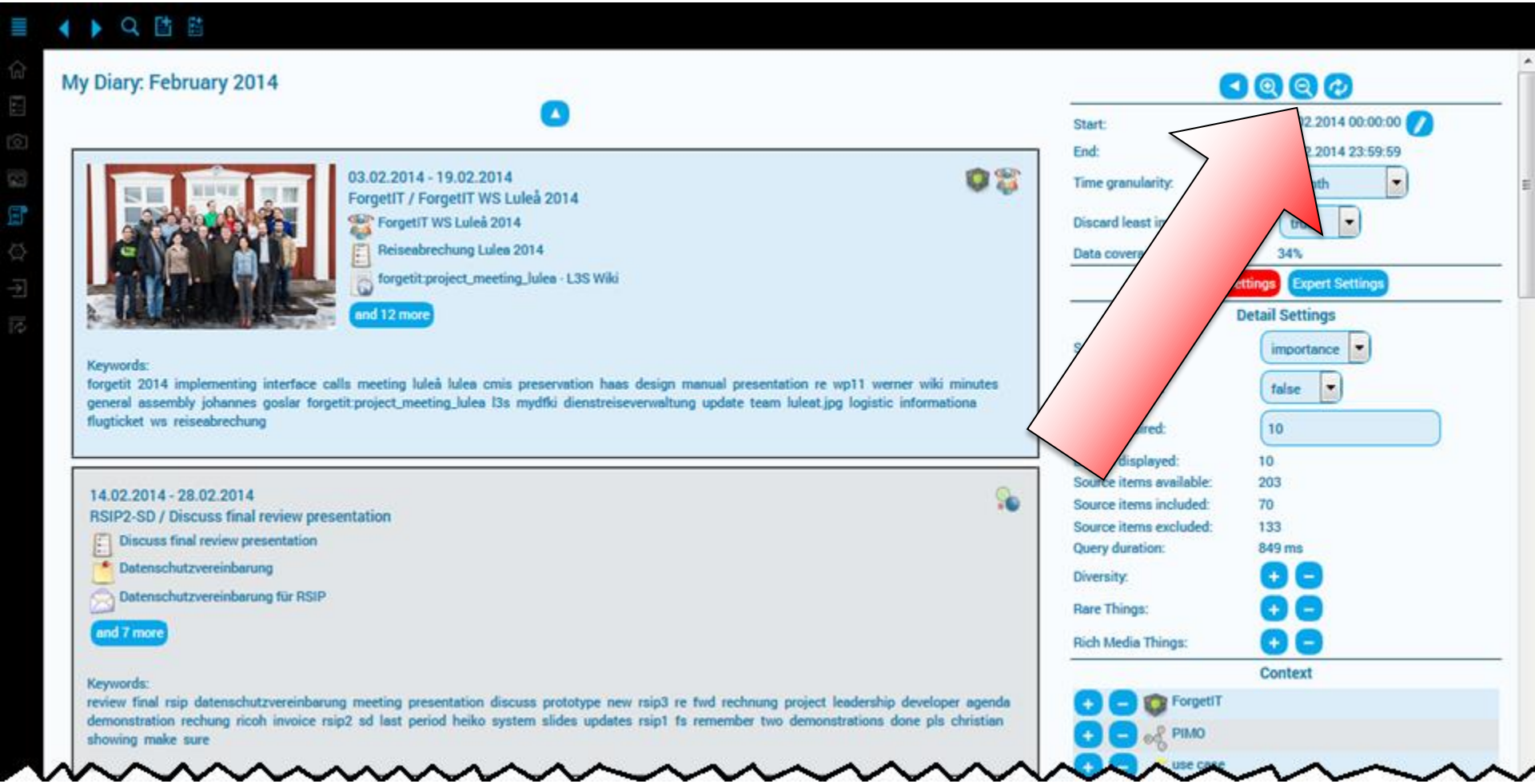
associated photos

most prominent things

most prominent annotations

most prominent keywords

Zoom out



Zoom out

My Diary: 2014

06.01.2014 - 22.12.2014
ForgetIT
ForgetIT Description of Work
ForgetIT WS Luleå 2014
ForgetIT Workshop Turin 2014
Review ForgetIT 2014 in Kaiserslautern
Istanbul 2014
[ForgetIT-all] Istanbul Meeting and the HOTEL
ForgetIT-Team-Luleat.jpg
QRPR Q1/14
and 722 more things

331 more images

Keywords:
jpg forgetit 2014 review meeting dfki wpleader wp9 pimo y2 edinburgh re report istanbul preservation q3 diary jilek reports 14 workshop survey quarterly wp4 update advisory l3s board kaiserslautern wp3 q1 telco information heiko mb 2015 evaluation diplomarbeit pptx forgetting

15.01.2014 - 10.11.2014
Schüler-Praktikum 2014
Schüler-Praktikum 2014
AirScript
HugoBallInternship2014 – kmintern
and 18 more things

4 more images

Keywords:
2014 dfki praktikum schülerpraktikum schüler jpg intern ausgabe 14 zusage 24.07.2014 hugo ball 3.01 airscript august september fwd topics 30 möglichkeit docx internship 32 praktikanten 31 november kmintern anhang bescheinigung bestätigungen praktikumszeugnisse tester img_0001 img_0005 img_3854 infopage hugoballinternship2014 teilnehmen teil

24.02.2014 - 19.12.2014

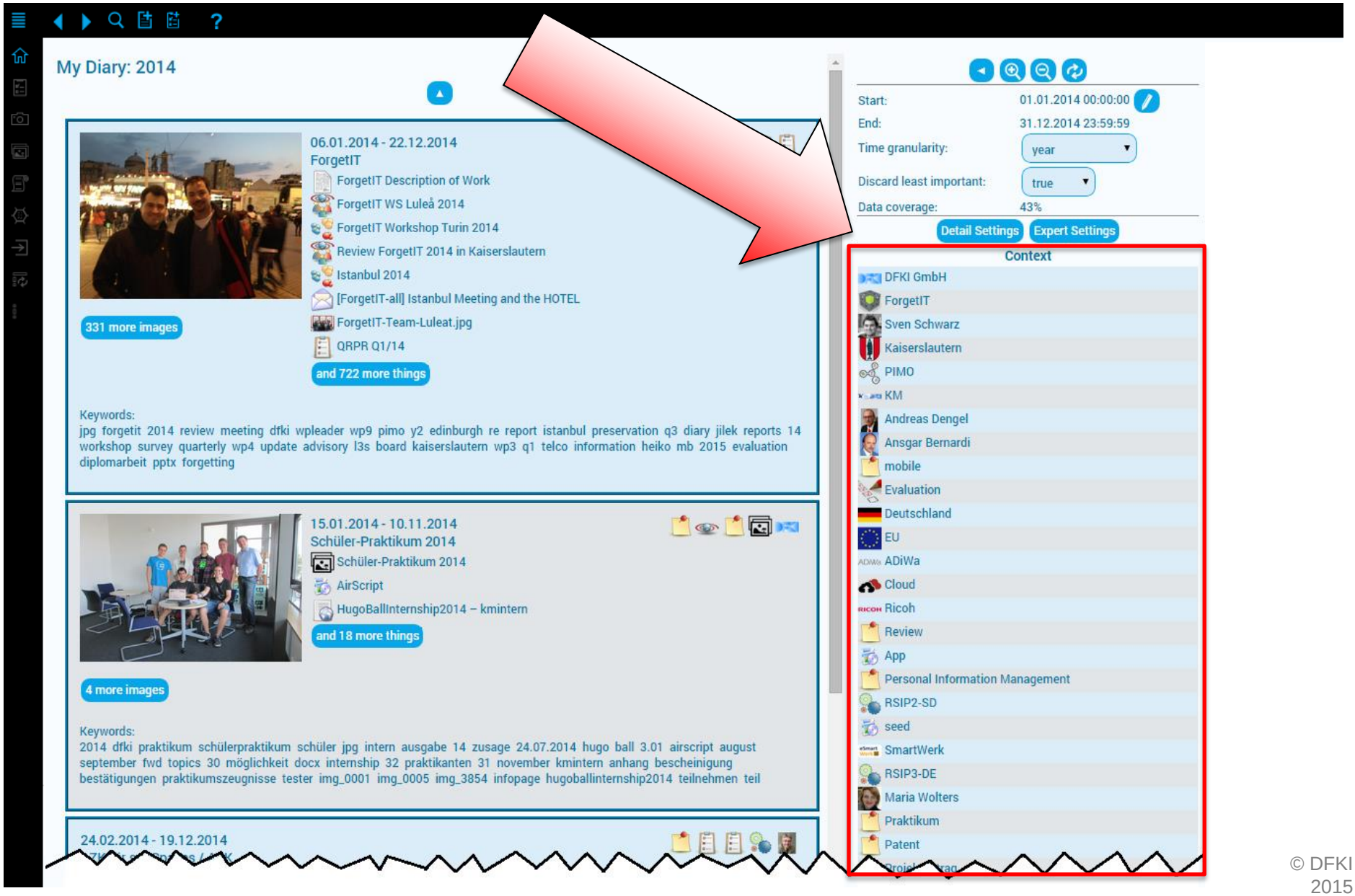
Start: 01.01.2014 00:00:00
End: 31.12.2014 23:59:59
Time granularity: year
Discard least important: true
Data coverage: 43%

Detail Settings | Expert Settings

Context

- DFKI GmbH
- ForgetIT
- Sven Schwarz
- Kaiserslautern
- PIMO
- KM
- Andreas Dengel
- Ansgar Bernardi
- mobile
- Evaluation
- Deutschland
- EU
- ADiWa
- Cloud
- Ricoh
- Review
- App
- Personal Information Management
- RSIP2-SD
- seed
- SmartWerk
- RSIP3-DE
- Maria Wolters
- Praktikum
- Patent

Overall Context



Overall Context

2012

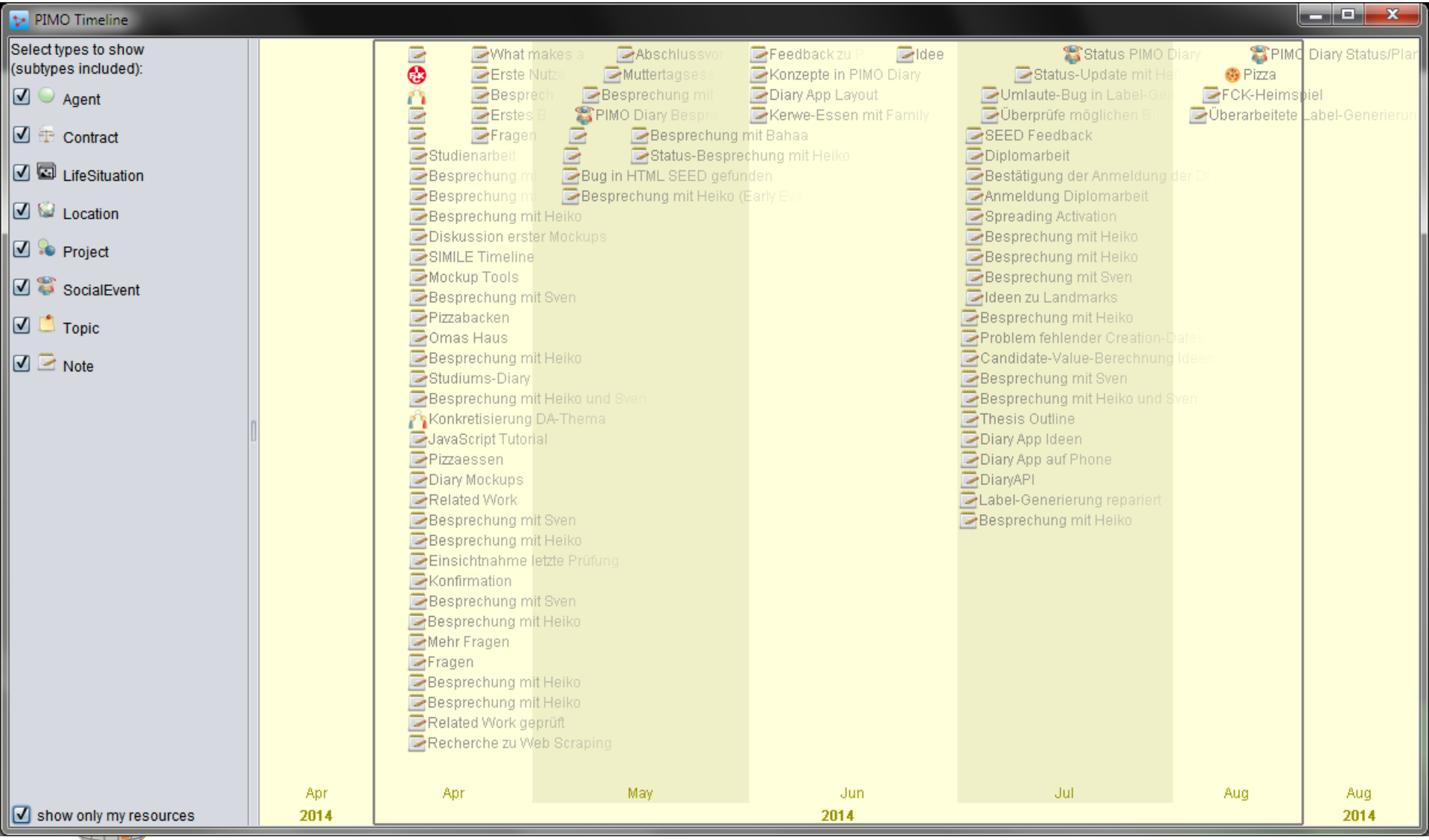
- ADiWa
- Deutschland
- BMBF
- SmartOffice
- Köln
- ConTask
- ARIS
- Google Inc.
- Apple
- Innovation
- Semantic File Explorer
- Wissensmanagement
- Düsseldorf
- IEEE
- skipforward
- Infos zu gnwosis und elsevier Challenge für internen Newsletter
- ADiWa G4 Deliverable Evaluation
- CoPLANT

2014

- DFKI GmbH
- ForgetIT
- Sven Schwarz
- Kaiserslautern
- PIMO
- KM
- Andreas Dengel
- Ansgar Bernardi
- mobile
- Evaluation
- Deutschland
- EU
- ADiWa
- Cloud
- Ricoh
- Review
- App
- Personal Information Management

© DFKI 2015

My Timeline of 2014



My Diary of 2014

25.02.2014 - 13.08.2014
Diplomarbeit Christian Jilek / diary
- PIMO Diary Besprechung
- Status PIMO Diary
- erste mock-up Session Jilek
- treffen mit jilek
- Jilek Konkretisierung Thema Diary
- Besprechung mit Bahaa

and 62 more things

Keywords:
besprechung diary heiko sven pimo app evtl 1 diplomarbeit status fragen idee seed jilek mockups todo work 2 ideen related beispiel recherche z.b text landmarks meeting konzepte 5 feedback evaluation web timeline thesis treffen update datenverdichtung anmeldung 3 erstes abstraktion

29.03.2014 - 09.08.2014
Pizza / Pizzabacken
- Pizza
- Pizzabacken
- Pizzaessen

Keywords:
pizzaessen pizzabacken pizzabacken.jpg pizza 29.03.2014 lecker family 05.04.2014 friends

02.07.2014 - 04.08.2014
Notes / Bug
- Umlaute-Bug in Label-Generierung von Notes repariert
- Überprüfe möglichen Bugfix für Umlaute
- Überarbeitete Label-Generierung für Notes

and 1 more things

Keywords:
label generierung umlaute notes repariert bug überprüfe bugfix überarbeitete bahaa evtl behoben absprache nochmal verbessert hoffentlich seed text bekommen abstimmung möglichkeit newlines

27.06.2014
Spreading Activation

Keywords:
spreading activation z.b topic idee zumindest person relationen bestimmte beachten unterschiedliche gewichtungen einsetzen ismemberof manages ismanagerof projekt haspart ispartof hassubtopic hassupertopic

01.06.2014
Kerwe-Essen mit Family

Keywords:
essen kerwe family lecker 01.06.2014 hähnchen zeug schwarzwaldständchen

21.02.2014 - 21.05.2014
Diskussion Fragen zur Evaluation / Themenfindung Christian Jilek
- Diskussion Fragen zur Evaluation
- Themenfindung Christian Jilek
- Themenabstimmung Jilek

Keywords:
jilek evaluation themenabstimmung fragen themenfindung christian diskussion

18.03.2014 - 13.05.2014
Abschlussvortrag Studienarbeit Wirtschaftsinformatik
- Abschlussvortrag Studienarbeit Wirtschaftsinformatik
- Studienarbeit
- Einsichtnahme letzte Prüfung

and 1 more things

Keywords:
studienarbeit letzte wirtschaftsinformatik einsichtnahme diplomprüfung abschlussvortrag prüfung 24.03.2014 vergessen dbaw 11 15 uhr ag prof härder 36 termin insb 18.03.2014 yeah bestanden datenbankanwendung en kapitel 13.05.2014 letzter echter uni juhu 03.04.2014 feedbackgespräch kleinere tipps abgeholt hinsichtlich formulierenung kapiteleinteilungen merke lang

11.05.2014
Muttertagsessen

Keywords:
muttertagsessen anlässlich essen 11.05.2014 muttertags großteil familie salat pizza pasta

17.03.2014 - 14.04.2014
FCK-Heimspiel / Kaiserslautern
- FCK-Heimspiel
- 1. FC Kaiserslautern

Keywords:
1 fck heimspiel fc kaiserslautern fckticket.jpg endlich auf'm betze köln

04.04.2014
Omas Haus
- Omas Haus

Keywords:
omas haus omashaus.jpg fertig 04.04.2014 ausgeräumt

23.03.2014
Konfirmation
- Konfirmation

Keywords:
konfirmation aussicht.jpg 23.03.2014 anlässlich freunden bekannten essen coole aussicht

27.02.2014
Proof of Concept

Keywords:
proof concept 27.02.2014 meilenstein prinzipielle durchführbarkeit vorhabens belegt quelle wikipedia

Group Information MOdel (GIMO) Diary

Heiko‘s initial entry

12.03.2014 - 31.03.2014
Related work DA Jilek identifizieren / Diplomarbeit Christian Jilek

- Related work DA Jilek identifizieren 1
- Vorarbeiten des DFKI mit Relevanz zu Diary 2
- Jilek Konkretisierung Thema Diary 3
- Jilek Kick-off 4
- ChurchCherubiniOliver14: A large-scale study of daily information needs captured in situ 5
- Diplomarbeit ab nächster Woche 6

hide less important things

Keywords:
jilek diary thema kick dfki related work woche daily large information diplomarbeit konkretisierung nächster study identifizieren vorarbeiten relevanz churchcherubinioliver14 scale captured situ sven bahaa seed app alzayat firetasks rahman pimostation pimomeeting michael kraus pimocal nops mohamed badreldin

include shared information

now including Christian‘s shared data

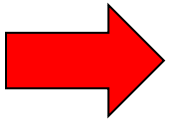
12.03.2014 - 31.03.2014
Diplomarbeit Christian Jilek / diary

- Related work DA Jilek identifizieren 1
- Vorarbeiten des DFKI mit Relevanz zu Diary 2
- Jilek Konkretisierung Thema Diary 3
- Jilek Kick-off 4
- Besprechung mit Heiko
- Besprechung mit Heiko
- Besprechung mit Heiko und Sven
- Besprechung mit Sven
- ChurchCherubiniOliver14: A large-scale study of daily information needs captured in situ 5
- Related Work geprüft
- Related Work
- Besprechung mit Heiko
- Besprechung mit Heiko
- Diary Mockups
- Besprechung mit Sven
- Mehr Fragen
- Fragen
- JavaScript Tutorial
- pimo5 – smartoffice
- Diplomarbeit ab nächster Woche 6

hide less important things

Keywords:
besprechung diary heiko sven work related jilek fragen mockups app pimo dfki pimo5 seed tutorial diplomarbeit thema geprüft woche daily large information javascript kick 5 smartoffice buoyancy preservation konkretisierung nächster study identifizieren vorarbeiten relevanz churchcherubinioliver14 scale captured situ z.b 4

Basic, Detail & Expert Settings



Start: 01.01.2014 00:00:00
End: 31.12.2014 23:59:59
Time granularity: year
Discard least important: false
Data coverage: 100%

Detail Settings | Expert Settings

Detail Settings

Sorting: blog
Include shared: false
Entries desired: 10
Entries displayed: 12
Source items available: 94
Source items included: 94
Source items excluded: 0
Query duration: 8 ms
Diversity: + −
Rare Things: + −
Rich Media Things: + −

Expert Settings

Concepts in diary: 5772
Concept index size: 13770
Term index size: 25501
Max. sim. found (weighted): 0.90224
Better clusters in CC: 0
Remove image things: true
Use compact vectors: true
Show cluster comp.: false
Sim. threshold type: amountOfMax
Sim. threshold value: 0.25
Weight concepts: 0.5
Weight label terms: 0.25
Weight text terms: 0.5
Normalize sim. values: true
Max. weight related: 0.75
Centroid check (CC): auto
CC sim. thr. factor: 2
Find better clusters in CC: false
CC auto mode factor: 1.5
2nd Clust. Pass (2CP): 1rounded
2CP ann. limit factor: 0.75
2CP #incl. ann. things: 10
Ann. limit factor: 0.75
Rarity appl. factor: 10
Rarity bonus: 1
Rich media bonus: 1
Top elem. in context: 20
Label terms factor: 3
Keywords per entry: 40
Img. weight fav. 1
Img. weight keep/del. 0.3
Img. weight con. ann.: 0.1
Img. weight text ann.: 0.1
Img. weight quality: 0.5

- **Similarity Calculation**
  - using **term vectors** for **headlines** and **text bodies**
  - using **concept vectors** for **concept annotations**
  - **Spreading Activation** to find **indirect annotations** and
  - extend concept vectors

- **Clustering of similar entries**
  - also fosters **higher diversity** within the diary

- **Importance evaluation**
- **Headline Generation**
- **Text Summarization**

next slides

Importance evaluation of diary entries based on

- **annotation intensity**,
- presence of things having a high potential of being a **memory landmark**,
- **rarity** (idea: *rare persons/locations/etc. might be more memorable*),
- associations with **rich media**

09.08.2014 - 29.03.2014
Pizza

composite entry consisting of 4 single entries:
- Pizzaessen [pimo:thing#Note]
- Pizza [pimo:thing#Topic]
- Pizzabacken [pimo:thing#Note]
- pizzabacken.jpg [pimo:thing#Image]

importance evaluation:
- (annotations | highPrioThings | rarity | richMedia) = (1.104 | 0.699 | 0.000 | 1.000)
- importance = 2.803

27.06.2014
Spreading Activation

single entry:
- Spreading Activation [pimo:thing#Note]

importance evaluation:
- (annotations | highPrioThings | rarity | richMedia) = (0.588 | 0.699 | 1.000 | 0.000)
- importance = 2.287

Headline Generation based on

- labels of **annotations** **and** **information elements**
- intra-cluster importance evaluation (similar to previous slide)
- length of viewed time period
- possibly a split label
- example:

ForgetIT / ForgetIT WS Luleå 2014

03.02.2014 - 19.02.2014
ForgetIT / ForgetIT WS Luleå 2014
ForgetIT WS Luleå 2014
Reiseabrechung 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 reiseabrechung

Keywords as a summary of the entry

- weight(label terms) > weight(text body terms)

Early Evaluation: Setting

- *PANiC*:
  - group of 4 participants (50%♀, 50%♂)
  - „PANiC“: acronym of their first names
  - Industrial Engineering students
  - in their last year before earning a master‘s degree
  - **4 months** access to our **Semantic Desktop prototype**
  - then testing **PIMO Diary** for **3 weeks**

- Questionnaire:
  - items 1-12: **compact USE questionnaire** (Lund, 2001)
  - items 13-20: **specific questions** concerning our app‘s **core features**
  - text field to express any kind of **feedback or comments**

  - 7-point Likert scale, each item is phrased such that: 7: best value
    1: worst value

Early Evaluation: Summary of Results

- **Overall ratings (80 items in total):**
  40x 7, 30x 6, 9x 5, 1x 4, 0x 3, 0x 2, 0x 1 → overall average rating: 6.36

- **Item 19:**
  The app allows for an appropriate and **satisfactory retrospection** on those parts of my life that are reflected by my PIMO. → item‘s average rating: 6.75

- **Item 20:**
  The **overall context** provides a good impression, i.e. a quick overview, of those things (reflected by my PIMO) that concerned me the most in the chosen period. → item‘s average rating: 6.75

- **Comments / Feedback of the *PANiC* group:**
  - „This program is very ***innovative***, I havn‘t seen anything like this before.“
  - „I was surprised by its ***intuitive handling*** and the ***quality of the results***.“
  - „Using the app was ***fun***.“
  - „It‘s a nice add-on to the PIMO which ***helps in keeping an overview***.“

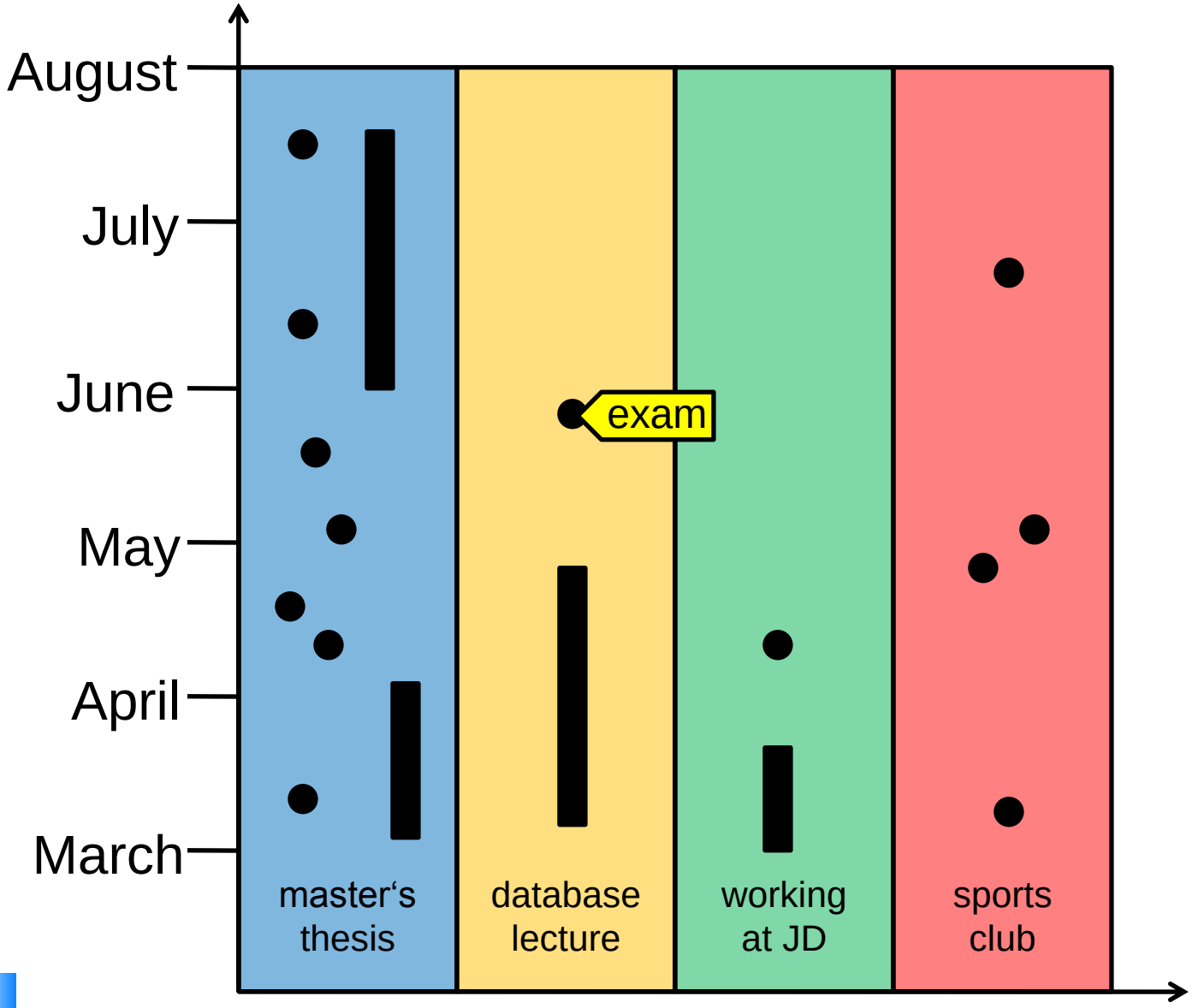
Conclusion:

- app enabling **early contextualization**
- **supporting** and especially **easing personal retrospection**
- innovative app:
  - **self-writing diary** with **blog look & feel**
  - **diversity** to make reading more exciting
  - **zoom** in and out of **time periods**
  - **manually shift emphases** (experimental)
  - **overall context**
- promising results in a first user experience evaluation

Open issues / outlook on possible future work:

- text **summarization using natural language** (sentences)
- things having a **time span**
- more **social media** capabilities (diary sharing)
- algorithm and parameter **tuning**
- **topic lanes**

Future Work: Topic Lanes



Thank you for your attention!

☺

Any questions?

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

References

- Sauermann et al., "***PIMO – a framework for representing personal information models,***" Proc. of I-Semantics, vol. 7, pp. 270–277, 2007.
- Sauermann et al., "***Overview and outlook on the semantic desktop.,***" in Proc. of the 1st Workshop on The Semantic Desktop at ISWC, 2005.
- Dengel, "***Knowledge technologies for the social semantic desktop,***" in Knowledge Science, Engineering and Management, pp. 2–9. Springer, 2007.
- Maus et al., "***Weaving personal knowledge spaces into office applications,***" in Integration of Practice-Oriented Knowledge Technology: Trends and Prospectives, M. Fathi, Ed., pp. 71–82. Springer, 2013.
- Eldesouky et al., "***Seed, a natural language interface to knowledge bases,***" in Proc. 17th Intl. Conf. on Human-Computer Interaction, Los Angeles, USA, 2015.
- Salton et al., "***A vector space model for automatic indexing,***" Communications of the ACM, vol. 18, no. 11, pp. 613–620, 1975.
- Liu et al., "***A simple and effective Concept Vector for WordNet semantic measurement,***" in Advanced Computer Theory and Engineering (ICACTE), 2010, vol. 2, pp. 342–345.
- Crestani, "***Application of spreading activation techniques in information retrieval,***" Artificial Intelligence Review, vol. 11, no. 6, pp. 453–482, 1997.
- Horvitz et al., "***Learning predictive models of memory landmarks,***" in Proc. CogSci 2004: 26th Annual Meeting of the Cognitive Science Society, 2004.
- Mavridaki & Mezaris, "***No-reference blur assessment in natural images using fourier transform and spatial pyramids,***" in IEEE Intl. Conf. on Image Processing (ICIP), France, 2014, pp. 566–570.
- Lund, "***Measuring usability with the USE questionnaire,***" Usability interface, vol. 8(2), pp. 3–6, 2001.

More References

- Logie et al., "***D2.2: Foundations of forgetting and remembering - preliminary report,***" Deliverable D2.2, ForgetIT consortium, March 2014.
- Seiffge-Krenke, "***'Dear Kitty, you asked me...': imaginary companions and real friends in adolescence,***" Praxis der Kinderpsychologie und Kinderpsychiatrie, vol. 50, no. 1, pp. 1–15, 2001.
- Sumi et al., "***Comic-Diary: Representing individual experiences in a comics style,***" in UbiComp 2002: Ubiquitous Computing, vol. 2498 of Lecture Notes in Computer Science, pp. 16–32. Springer, 2002.
- Cho at al., "***Ani-Diary: Daily cartoon-style diary exploits bayesian networks,***" Pervasive Computing, IEEE, vol. 6, no. 3, pp. 66–75, 2007.
- Liao et al., "***Smart Diary: A smartphone-based framework for sensing, inferring and logging users' daily life,***" IEEE Sensors Journal, vol. 15, no. 5, 2014.
- Plaisant et al., "***LifeLines: using visualization to enhance navigation and analysis of patient records.,***" in Proc. AMIA Symposium. American Medical Informatics Association, 1998, pp. 76–80.
- Ringel et al., "***Milestones in time: The value of landmarks in retrieving information from personal stores,***" in Proc. Interact, 2003, vol. 2003, pp. 184–191.
- MIT, "***SIMILE Widgets: Timeline,***" Application software, 2009, http://www.simile-widgets.org/timeline/.
- André et al., "***Continuum: Designing timelines for hierarchies, relationships and scale,***" in Proc. 20th Annual ACM Symposium on User Interface Software and Technology. 2007, UIST '07, pp. 101–110, ACM.
- Hailpern et al., "***YouPivot: Improving recall with contextual search,***" in SIGCHI Conf. on Human Factors in Computing Systems. 2011, CHI'96, pp. 1521–1530, ACM.
- Microsoft Corporation, "***Microsoft Research: Lifebrowser,***" Video, 2012, http://research.microsoft.com/apps/video/default.aspx?id=159531