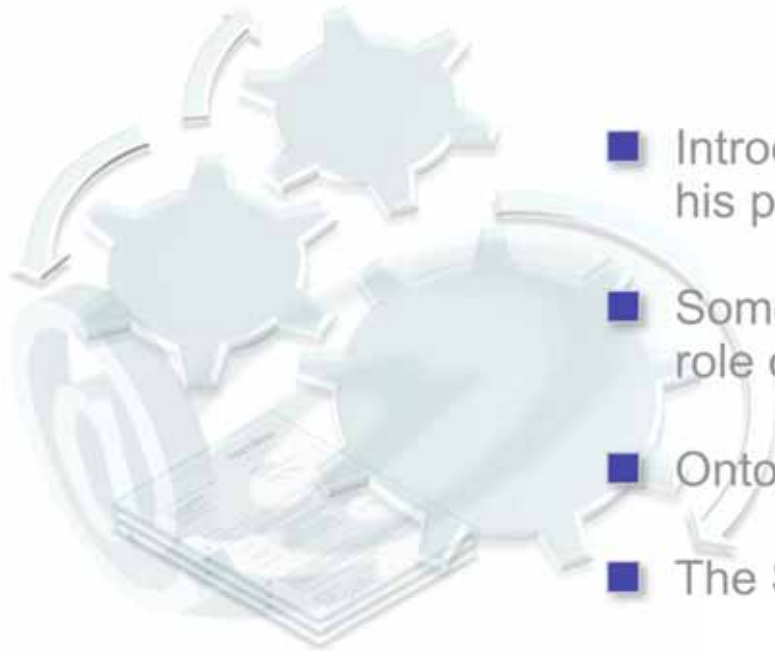# The Semantic Desktop as a Means for Personal Information Management

Andreas Dengel

Agenda

- Introduction to a typical knowledge worker and his problems in information management

- Some words about human perception and the role of ontologies

- Ontology creation and instantiation

- The Semantic Desktop

- Document Understanding via the Semantic Desktop

- Context-aware semantic services for the knowledge worker

- Dr. Thomas Mustermann is head of Customer Relationship Management
  at the German Research Center for Artificial Intelligence (DFKI) in Kaiserslautern, Germany

- He is one of those guys who are tackled by Constant Multi-Tasking Craziness. Right now he works on about 15 tasks at the same time one of which is the organization of a training course on DFKI's recent development: **The Semantic Desktop**

- When Andreas Dengel came back from IC3K 2009 held on Madeira, he told Thomas to invite Vitor Pedrosa (another knowledge worker) to the training course

- To fulfill his role and to complete all of his tasks, including the one he just received, Thomas has to acquire, organize, maintain, retrieve and use a whole bunch of information items

... but

Thomas' problem, as for many knowledge workers, is the need to know more than he may remember

E-Mails

Training Course Semantic Desktop
tbd

Agenda

On October 21st we will host a training course on the Semantic Desktop in Kaiserslautern and I was told to invite Vitor

Meeting Presentations

NEPOMUK
The Social Semantic Desktop
Training Course

Time Schedule
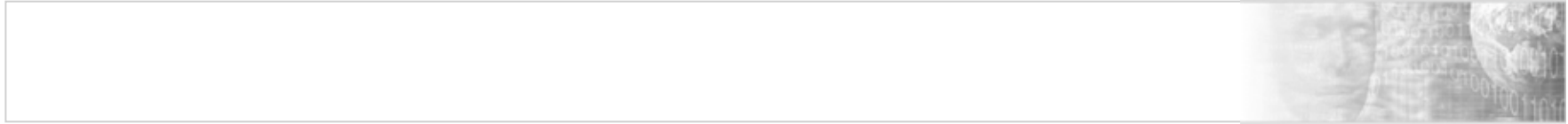
Training Course Semantic Desktop
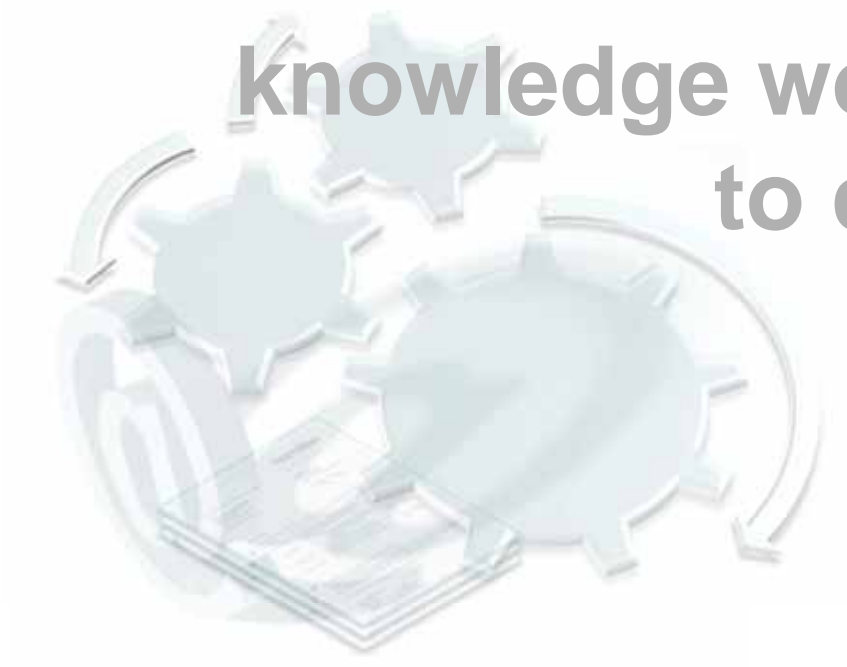
Contact Data

Petrosa, Vitor

Vitor Petrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

■ On Thomas' desktop, he stores around 12.000 files in about 2.300 folders in many applications and formats
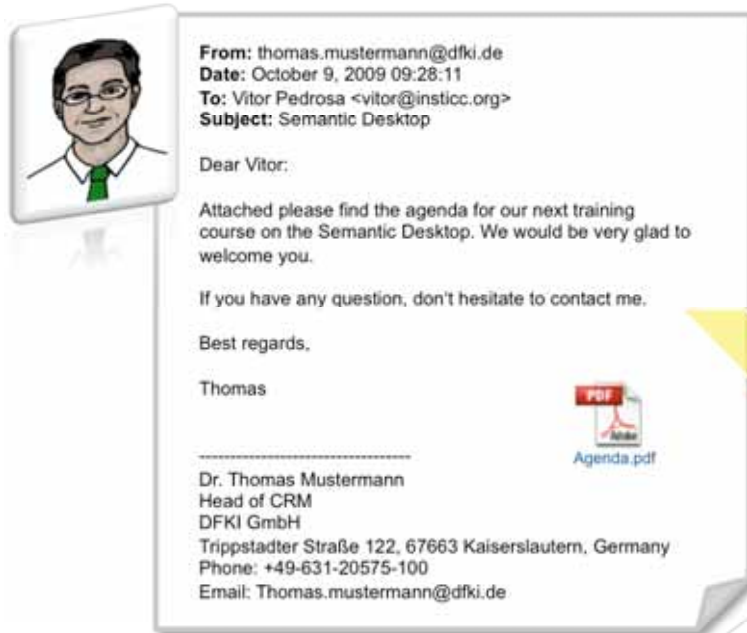
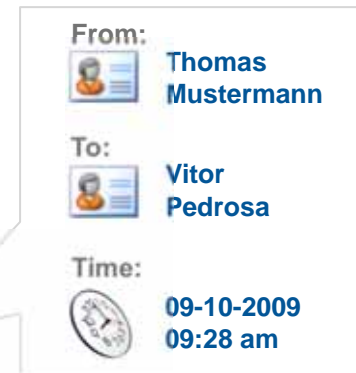# So what can we do in order to help knowledge workers like Thomas to do their job better?

When sending out the invitation, State-of-the-Art technologies may extract and store metadata about a document
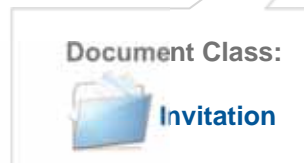
Example:

„Thomas sends invitation"

From: thomas.mustermann@dfki.de
Date: October 9, 2009 09:28:11
To: Vitor Pedrosa <vitor@insticc.org>
Subject: Semantic Desktop

Dear Vitor:

Attached please find the agenda for our next training course on the Semantic Desktop. We would be very glad to welcome you.

If you have any question, don't hesitate to contact me.

Best regards,

Thomas

Dr. Thomas Mustermann
Head of CRM
DFKI GmbH
Trippstadter Straße 122, 67663 Kaiserslautern, Germany
Phone: +49-631-20575-100
Email: Thomas.mustermann@dfki.de

Agenda.pdf

■ Metadata may be extracted from the Email

From:
Thomas Mustermann

To:
Vitor Pedrosa

Time:
09-10-2009 09:28 am

■ Email content may be indexed

Index:

■ Email content may be classified

Document Class:
Invitation

... but is this enough?

*We should no longer ask
if a single information item we can avail of is
sufficient.*

*We should rather ask
how to establish a vivid, interactive Information
Butler helping us to utilize the bits of knowledge
captured within our directories.*

The knowledge of humans is associative and perspective in which computers are lacking!

- Human beings may relate the content of a document into context because they have appropriate background knowledge

- Understanding a message leads to concepts and ideas being part of mental models generated while reading

- Computers may read the contents of documents but are not able to understand

The limits of today's desktop information management lead to a cut between mental models and document contents

From: thomas.mustermann@dfki.de
Date: October 9, 2009 09:28:11
To: Vitor Pedrosa <vitor@insticc.org>
Subject: Semantic Desktop

Dear Vitor:

Attached please find the agenda for our next training course on the Semantic Desktop. We would be very glad to welcome you.

If you have any question, don't hesitate to contact me.

Best regards,

Thomas

Dr. Thomas Mustermann
...ad of CRM
...GmbH
...ter Straße 122, 67663 Kaiserslaute... ...any
...-631-20575-100
E... ...mann@dfki.de

PDF Agenda.pdf

- Different applications manage different data

- Emails is filed in Email folder

- Attachments are stored in file folder

- Sender of an Email is stored in an independent address repository

- Related Websites are are disregarded

Emails
- ...
- Training Courses
- Seminars
- ...

File System
- Abstracts
- Advertisements
- Agendas
- ...

Contacts
- ...
- Filipe
- Petrosa
- ...

Bookmarks
- ...
- IC3K2009
- INSTICC
- ...

Mental Concepts

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda. I am glad to accept and confirm my participation in the Semantic Desktop course on October 21,

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

„Vitor confirms invitation"

- A document is like a key, which while reading opens a system of links to other documents, to events, locations, persons, or tasks

- As part of a process, a document contains many textual relations to existing knowledge that is complemented by new facts and relationships captured in the document

- Thus, the unit of a message is variable and relative, depending on who reads it at what time and in which context
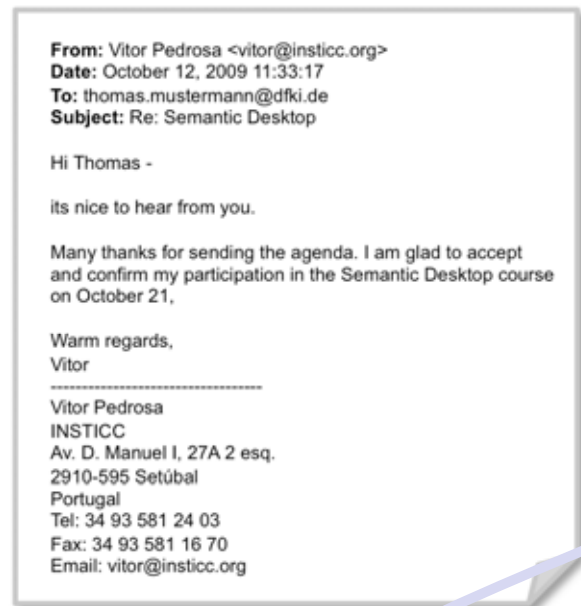
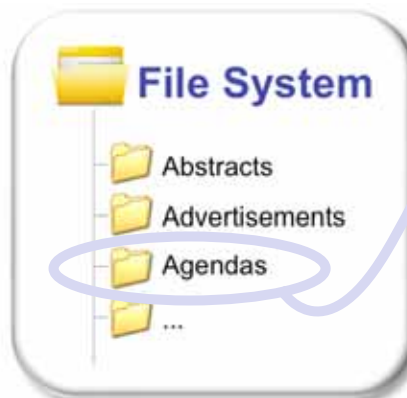# Imaginations without terms are blind and terms without imaginations are empty*

Imagination

evokes

is related to

"Portugal"

represents

Symbol

Reality

Semiotic Triangle

* I. Kant (1724 – 1804)
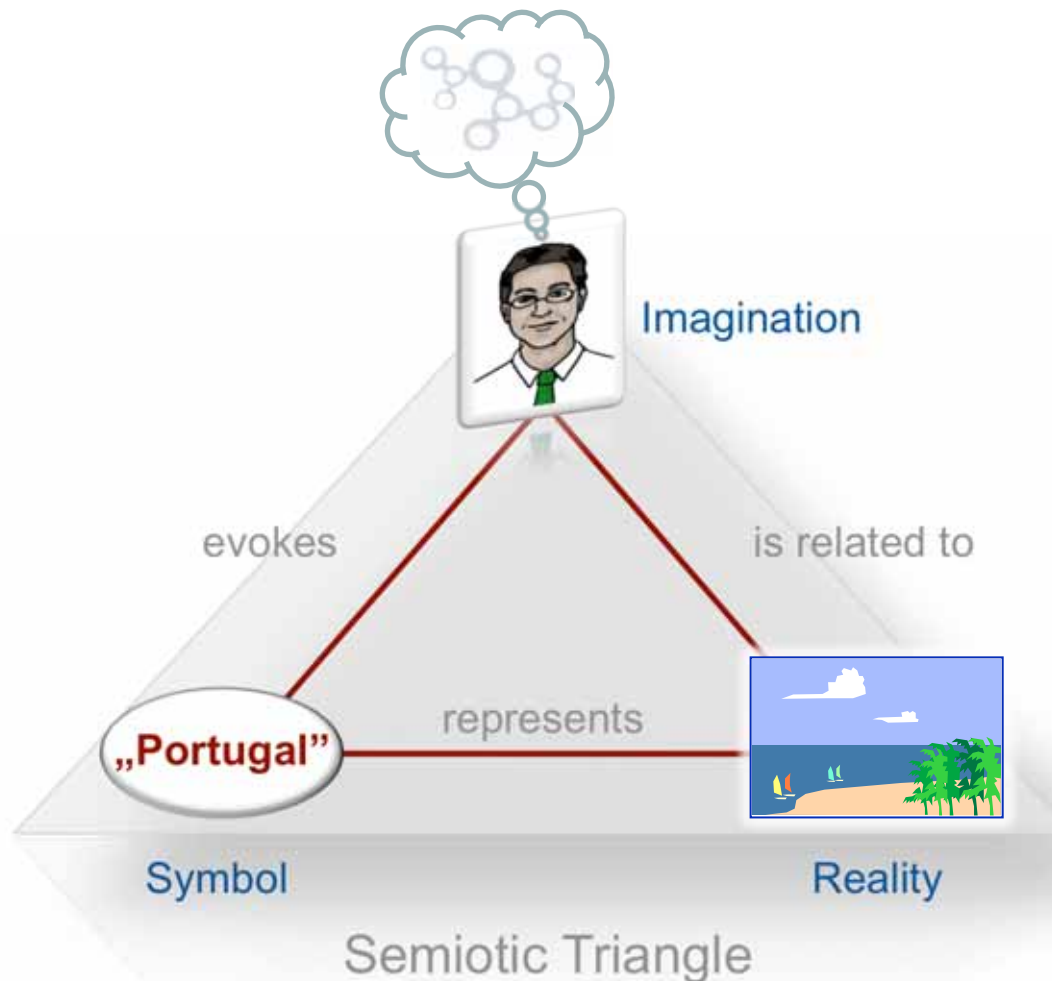
- Our environment consists of items, facts and events that are „real" and determine our lives („what is going on")

- In order to express their thoughts, people use signs, symbols, or characters that may be understood by others („what I couch or explicate")

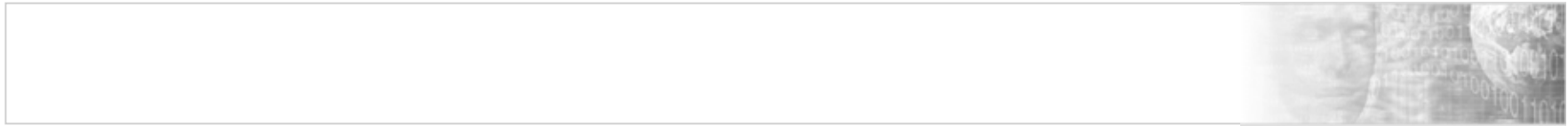- People reading texts put contents together and create their very individual imagination („what I mean")

When receiving new information, Thomas implicitly relates
the contents to his mental models he mirrored on his desktop

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda. I am glad to accept
and confirm my participation in the Semantic Desktop course
on October 21,

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

... but the linking only happens in his mind!

Emails
- ...
- Training Courses
- Seminars
- ...

File System
- Abstracts
- Advertisements
- Agendas
- ...

Contacts
- ...
- Filipe
- Petrosa
- ...

Bookmarks
- ...
- IC3K 2009
- INSTICC
- ...

Why don't we make these mental models explicit so that people like Thomas and Josep may better remember all they have to do?

One approach is to look on the ideas of the Semantic Web that builds on predication and ontology to formally represent semantics
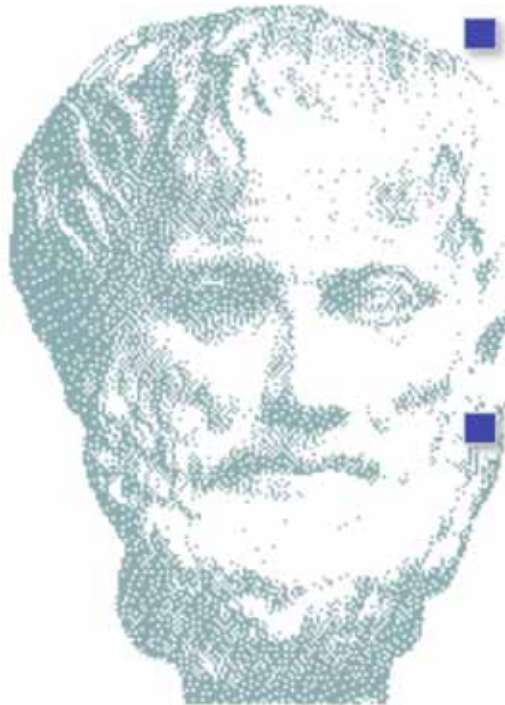
- A **Theory of Ontology** attempts to give answers to the question: **What is there?**

  (the Greek terms „ontos" and „logos" mean „to be" and „word")

  - Aristotle defined a system of ten categories, such as substance, quality, quantity, where, when, …

- A **Theory of Predication** tries to answer the question: **What is it to say something about something?**

  - A subject is what a statement is about

  - A predicate is what a statement says about its subject

- A common definition of **an Ontology** for Semantic Web researchers **is an explicit, formal specification of a conceptualization** *< Tom Gruber, 1993 >*

Subject   Predicate   Object

Vitor Pedrosa → 5082403

has-phone-no

```
<rdf:RDF>
 <rdf:Description
  rdf:about="http://www.insticc.org/Pedrosa">
<has-phone-no> 5082403</has-phone-no>
 <rdf:Description>
</rdf:RDF>
```
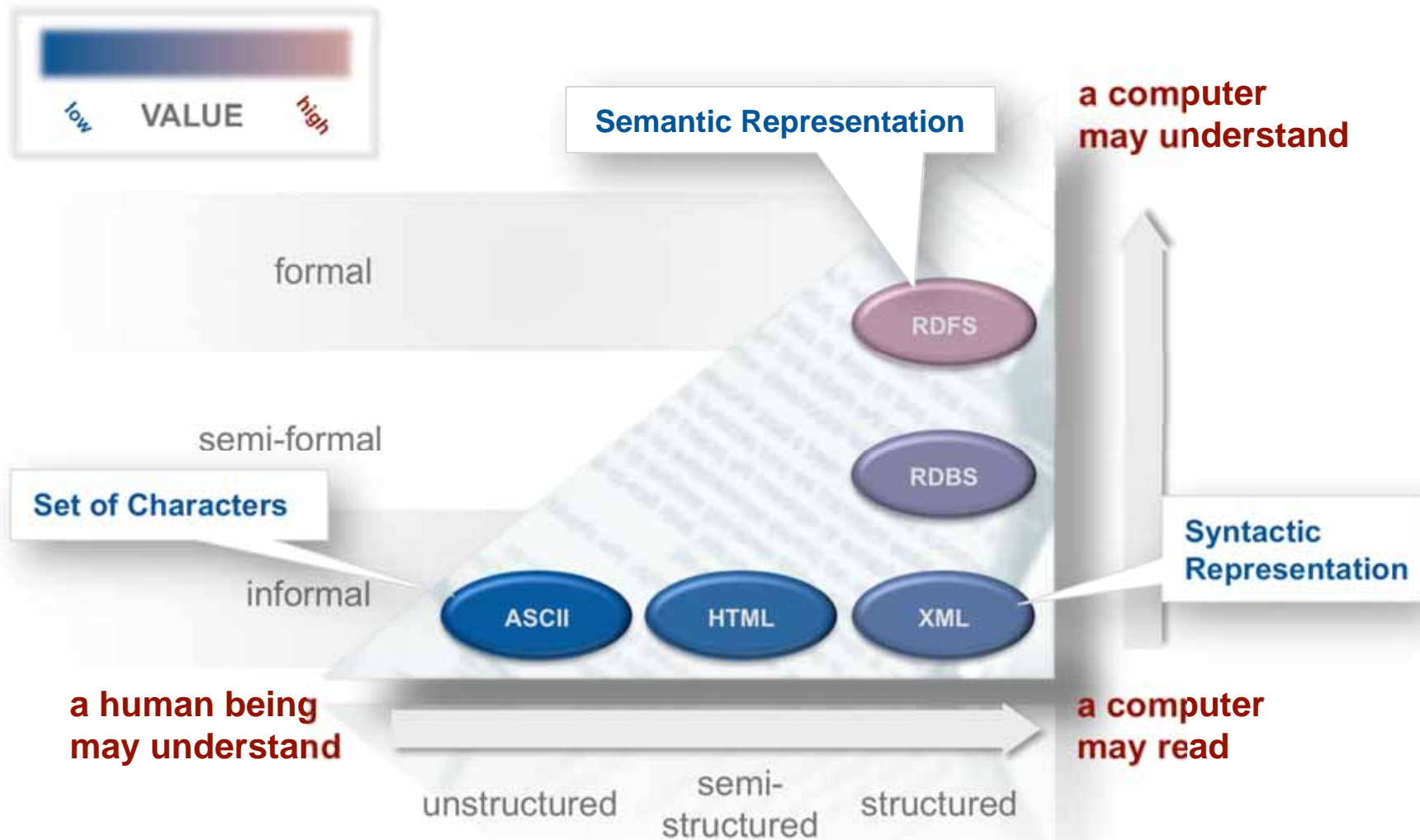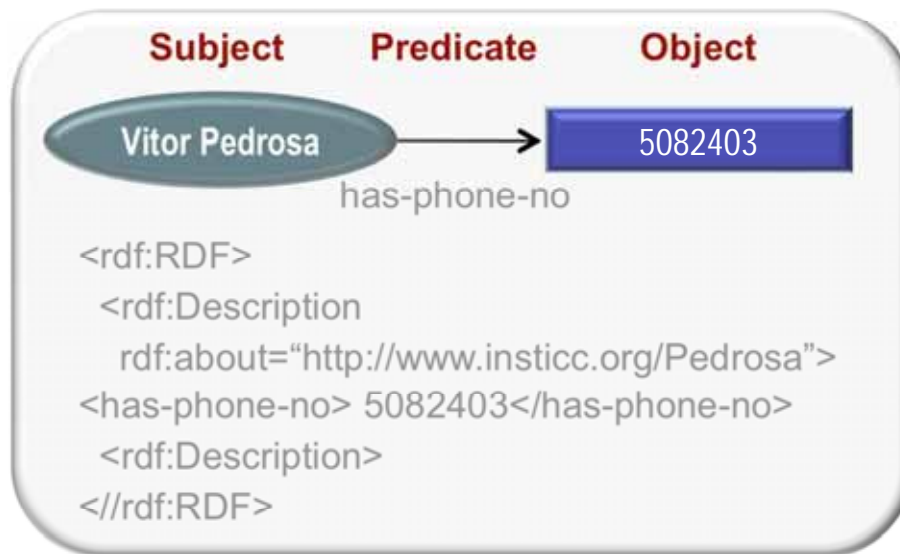
- A fact is expressed as a Subject-Predicate-Object triple

- Subjects, predicates, and objects are given as names for entities, also called resources or nodes

  - Entities *represent something, a person, an appointment, a website, …*

- Names are URIs, which are global in scope, always referring to the same entity in any RDF document in which they appear

- The underlying structure of any knowledge can be viewed as a graph (of triples) consisting of nodes (subjects, objects) and labeled directed arcs (predicates) that link pairs of nodes
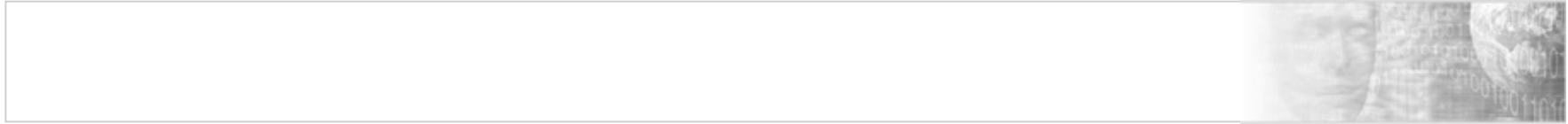
... but how can be provide
a shared vocabulary
that is understood by all users
of any semantic technology?

A Semantic Desktop is a device in which an individual stores all her digital information like documents, multimedia and message interpreted as Semantic Web resources. Each resource is identified by a Uniform Resource Identifier (URI) and all data is accessible and queryable as RDF graph.
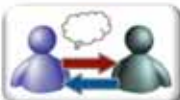
That way, resources from the web can be stored and authored content can be shared with others. Ontologies allow the user to express personal mental models and form the semantic glue interconnecting information and systems. Applications store, read and communicate via ontologies and Semantic Web protocols.

The Semantic Desktop is an enlarged supplement to the user's memory.

L. Sauermann, A. Bernardi and A. Dengel, *Overview and Outlook on the Semantic Desktop*,
Proceedings International Semantic Web Conference, Galway, Ireland (Nov. 2005), pp. 1-19.

Ontologies describe a particular vocabulary that can be used to describe aspects of real domains
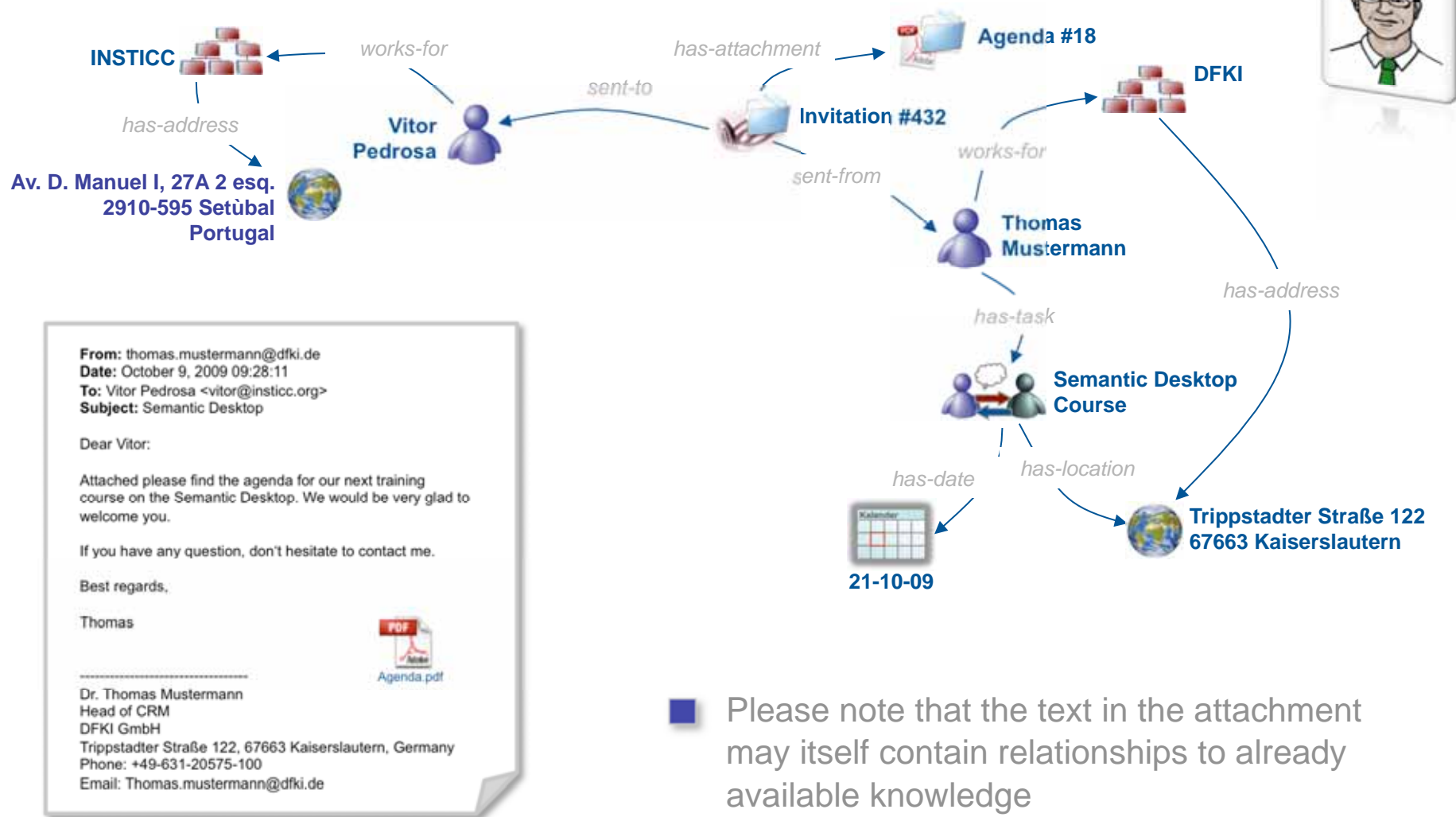
Document Classes

Organizations

Groups

Persons

Events

Locations/Addresses

Appointments

Topics

Tasks

- The vocabulary may follow different "W-Dimensions" of knowledge (what, who, when, where, …)

- All workflow-relevant aspects of information can be described using a set of explicit categories

- The categories can be taken from other applications and formally represented using RDFS
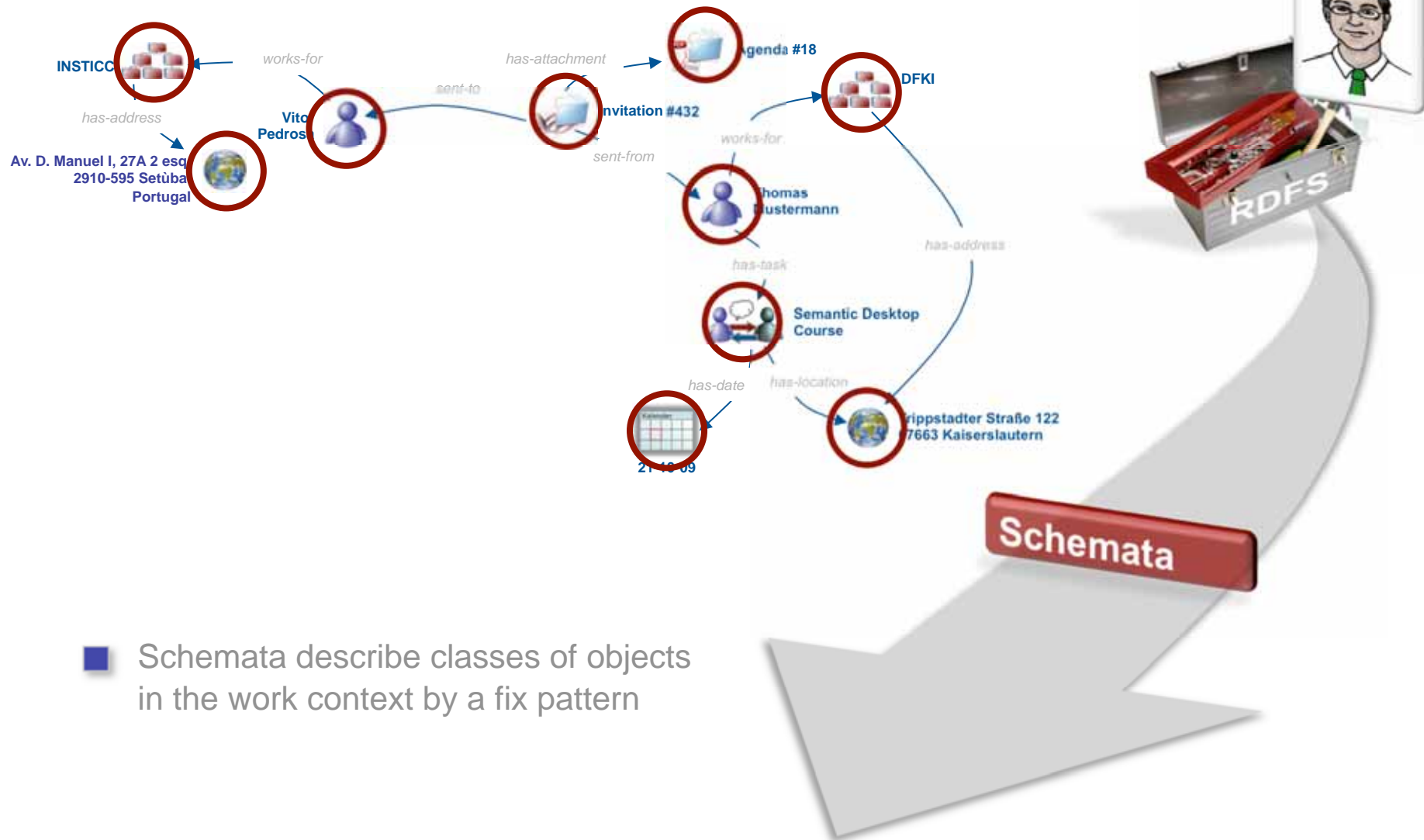
Exemplary categories for describing the work context (in RDFS they are called **schemata**)

Thomas already uses a Semantic Desktop on which he may create an application-independent **Personal Information Model**

INSTICC

works-for

has-attachment

Agenda #18

DFKI

sent-to

has-address

Vitor Pedrosa

Invitation #432

works-for

Av. D. Manuel I, 27A 2 esq. 2910-595 Setúbal Portugal

sent-from

Thomas Mustermann

has-address

has-task

Semantic Desktop Course

From: thomas.mustermann@dfki.de
Date: October 9, 2009 09:28:11
To: Vitor Pedrosa <vitor@insticc.org>
Subject: Semantic Desktop

Dear Vitor:

Attached please find the agenda for our next training course on the Semantic Desktop. We would be very glad to welcome you.

If you have any question, don't hesitate to contact me.

Best regards,

Thomas

PDF
Agenda.pdf

Dr. Thomas Mustermann
Head of CRM
DFKI GmbH
Trippstadter Straße 122, 67663 Kaiserslautern, Germany
Phone: +49-631-20575-100
Email: Thomas.mustermann@dfki.de

has-date

has-location

Kalender
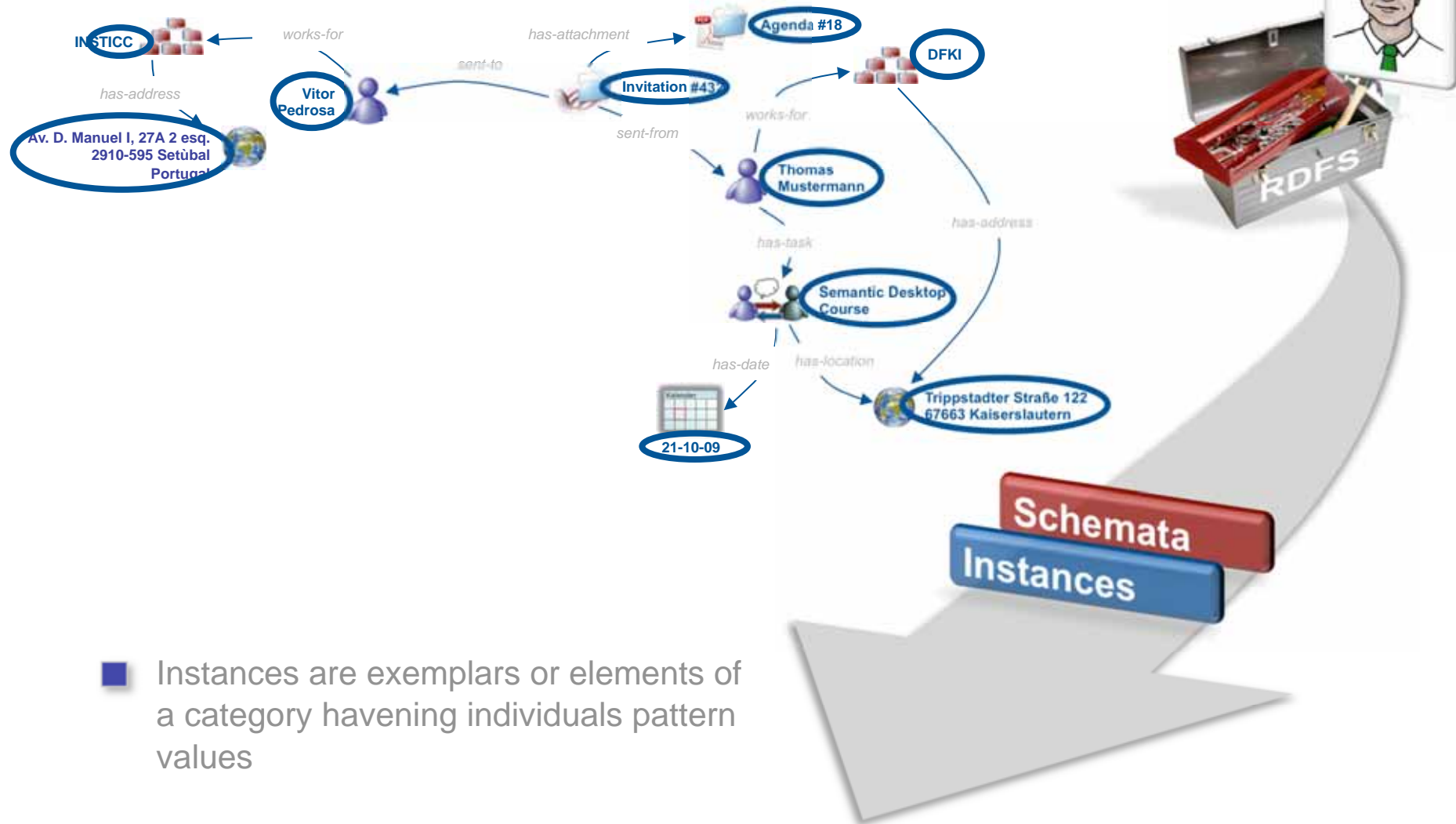
Trippstadter Straße 122 67663 Kaiserslautern

21-10-09

■ Please note that the text in the attachment may itself contain relationships to already available knowledge

For this purpose, he uses the rich RDFS tool box allowing
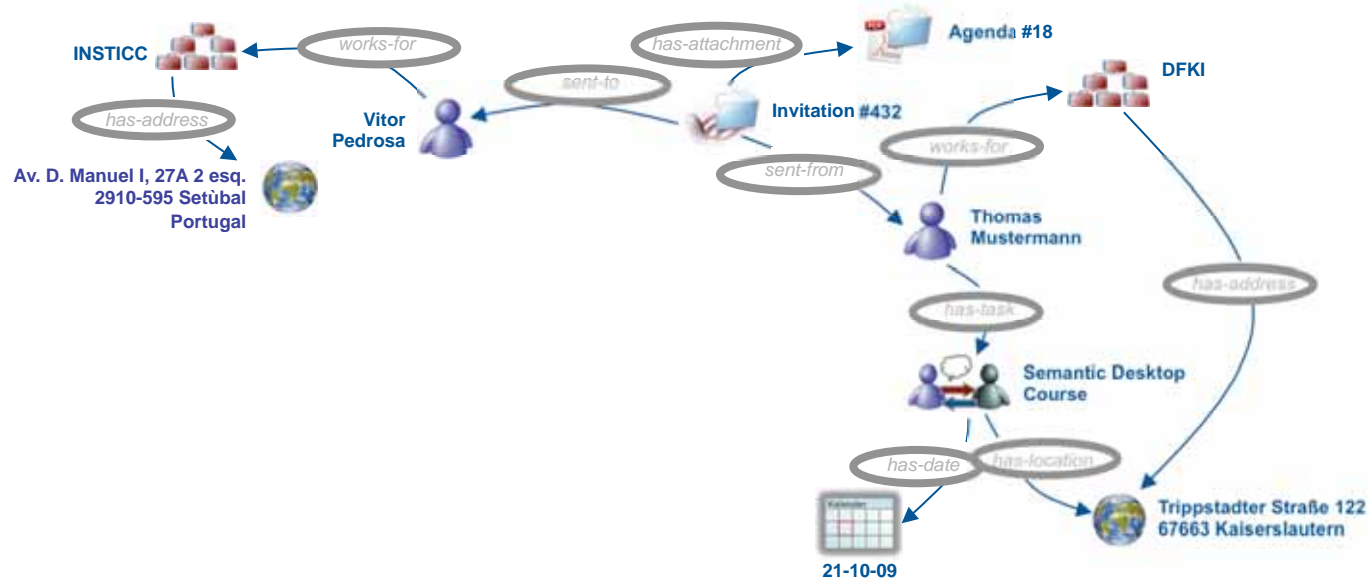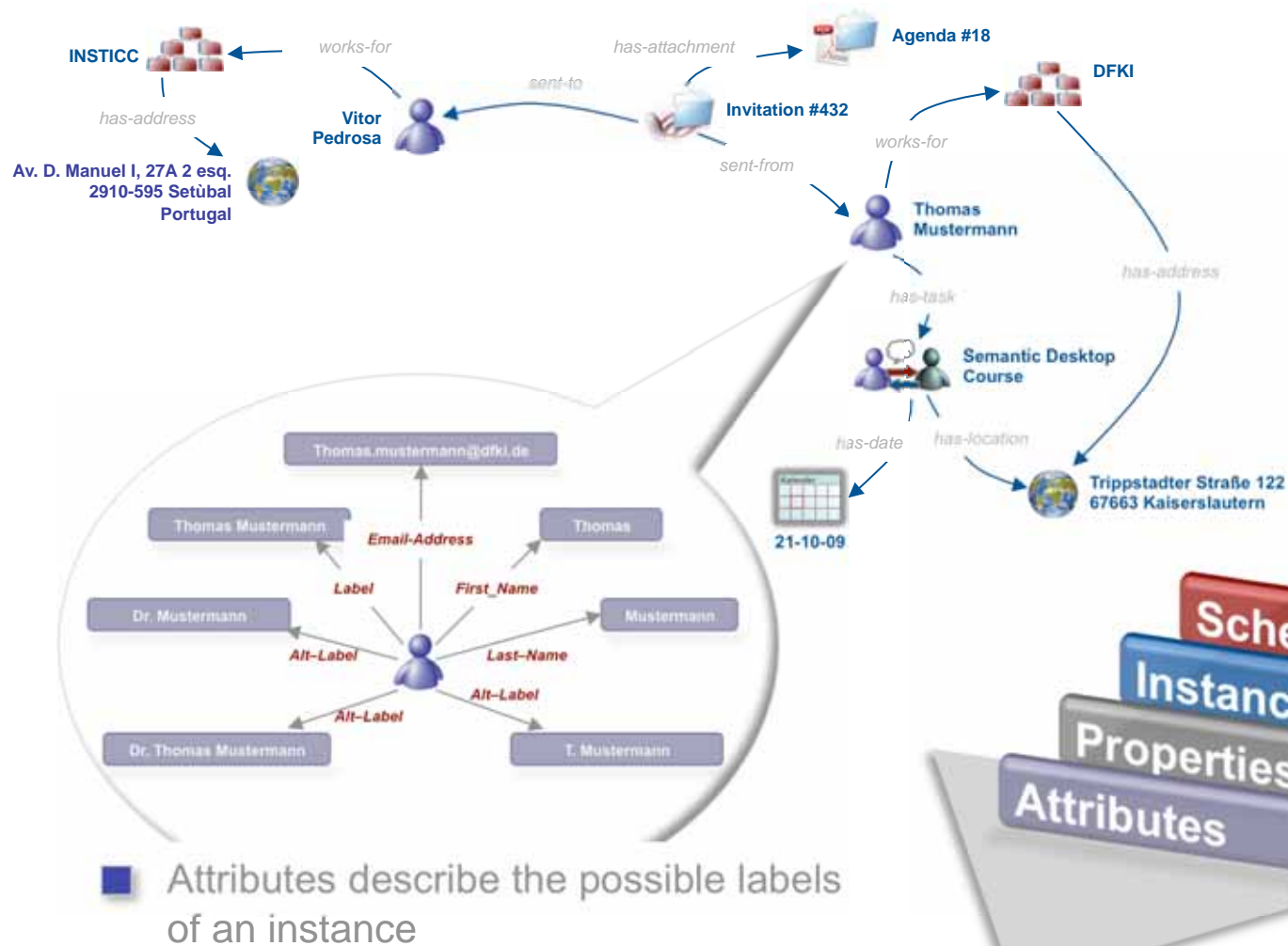him to formally represent all aspects of information he needs

INSTICC

works-for

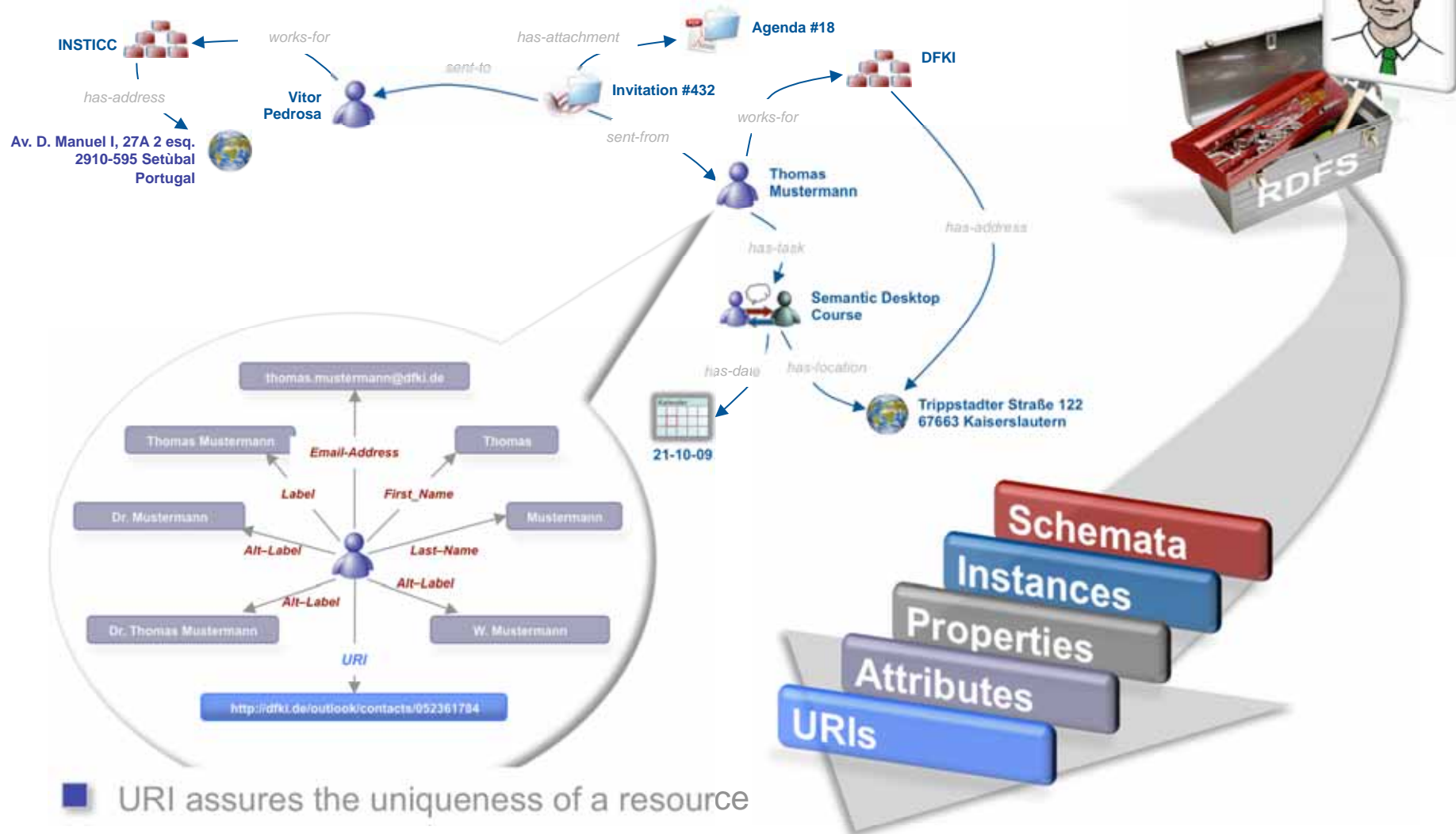*has-attachment*

Agenda #18

DFKI

sent-to

has-address

Vito
Pedros

Invitation #432

Av. D. Manuel I, 27A 2 esq
2910-595 Setùba
Portugal

sent-from

works-for

Thomas
Clustermann

has-address

has-task

Semantic Desktop
Course

has-date

has-location

rippstadter Straße 122
7663 Kaiserslautern

21-12-09

RDFS

Schemata

■ Schemata describe classes of objects
in the work context by a fix pattern

For this purpose, he uses the rich RDFS tool box allowing him to formally represent all aspects of information he needs

INSTICC

works-for

has-attachment

Agenda #18

sent-to

DFKI

has-address

Vitor Pedrosa

Invitation #43

works-for

Av. D. Manuel I, 27A 2 esq. 2910-595 Setùbal Portugal

sent-from

Thomas Mustermann

has-address

has-task

Semantic Desktop Course

has-date   *has-location*

Trippstadter Straße 122 67663 Kaiserslautern

21-10-09

Schemata

Instances

■ Instances are exemplars or elements of a category havening individuals pattern values

For this purpose, he uses the rich RDFS tool box allowing
him to formally represent all aspects of information he needs

INSTICC — *works-for* ← ● Vitor Pedrosa — *sent-to* — *has-attachment* → Agenda #18

has-address → Av. D. Manuel I, 27A 2 esq. 2910-595 Setùbal Portugal

Invitation #432 — *works-for* → DFKI

sent-from → Thomas Mustermann

has-task → Semantic Desktop Course

has-address → Trippstadter Straße 122 67663 Kaiserslautern

has-date → 21-10-09 *has-location*

Schemata
Instances
Properties

- Between the concepts of the ontology there are qualified relations called properties

- Each instance has a is-a-relationship to its class, i.e. it complies the defined pattern

For this purpose, he uses the rich RDFS tool box allowing him
to formally represent all aspects of information he needs

INSTICC

works-for

has-attachment

Agenda #18

DFKI

sent-to

has-address

Vitor
Pedrosa

Invitation #432

works-for

Av. D. Manuel I, 27A 2 esq.
2910-595 Setùbal
Portugal

sent-from

Thomas
Mustermann

has-address

has-task

Semantic Desktop
Course

has-date

has-location

Thomas.mustermann@dfki.de

Trippstadter Straße 122
67663 Kaiserslautern

Thomas Mustermann

Email-Address

Thomas

21-10-09

Label

First_Name

Dr. Mustermann

Mustermann

Alt-Label

Last-Name

Alt-Label

Alt-Label

Dr. Thomas Mustermann

T. Mustermann

Schemata

Instances

Properties

Attributes

Attributes describe the possible labels
of an instance

For this purpose, he uses the rich RDFS tool box allowing him to formally represent all aspects of information he needs

INSTICC

works-for

has-attachment

Agenda #18

DFKI

has-address

sent-to

Vitor
Pedrosa

Invitation #432

works-for

Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal

sent-from

Thomas
Mustermann

has-address

has-task

Semantic Desktop
Course

has-date

has-location

thomas.mustermann@dfki.de

Thomas Mustermann

Email-Address

Thomas

Label

First_Name

Dr. Mustermann

Mustermann

Alt-Label

Last-Name

Alt-Label

Alt-Label

Dr. Thomas Mustermann

W. Mustermann

URI

http://dfki.de/outlook/contacts/052361784

21-10-09

Trippstadter Straße 122
67663 Kaiserslautern

Schemata

Instances

Properties

Attributes

URIs

■ URI assures the uniqueness of a resource

Employing URIs, an application- and platform-independent unique representation for Thomas' resources is created

- Each information item is a semantic web resource whether it is file (folder or document), an email constituent (i.e. message, sender, recipient, attachment), an address (...), or a calendar entry, ...

- All resources are identified by a URI (Uniform Ressource Identifier)
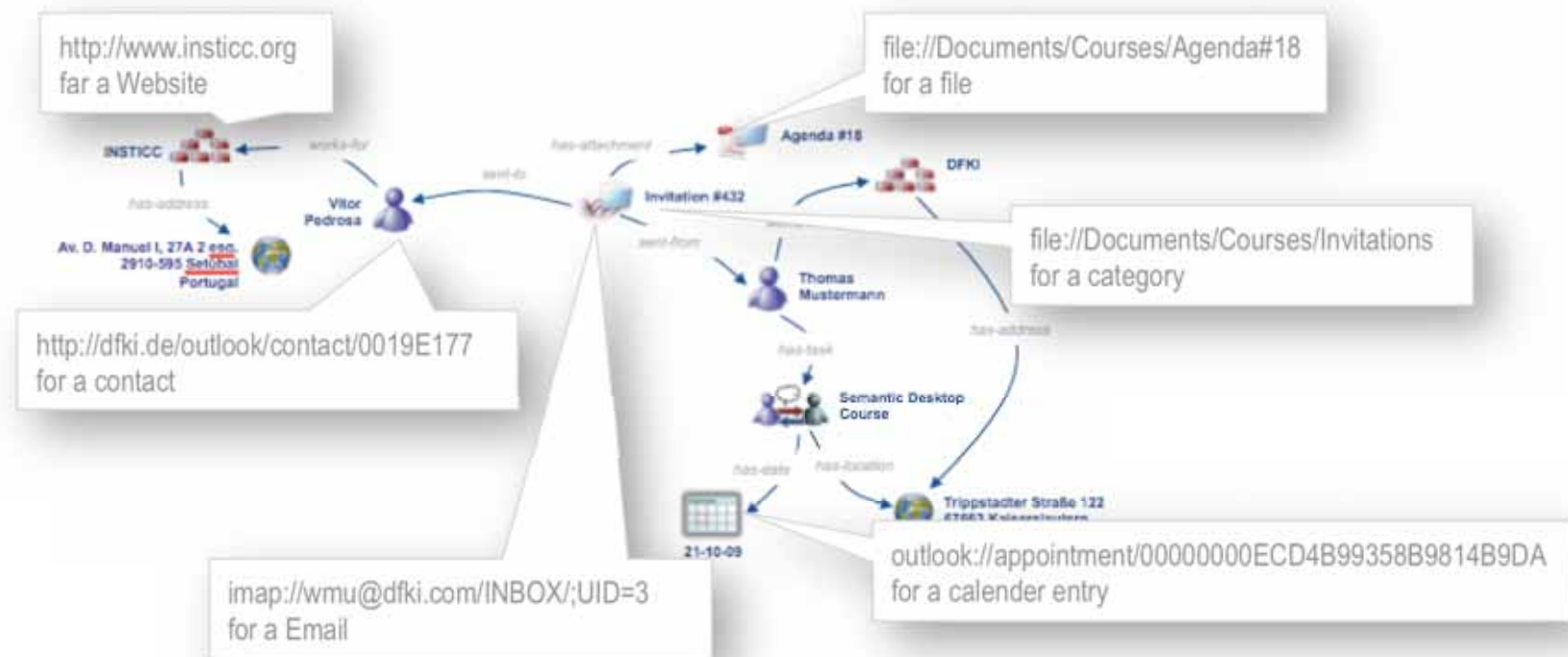
„ *The individual network of thoughts leads to a*

multi-dimensional and multi-perspective organization of content

and for this reason necessarily to a "dematerialization"

of traditional archiving concepts "

Over time Thomas' continuous work has transformed the initial schema system into a comprehensive domain vocabulary
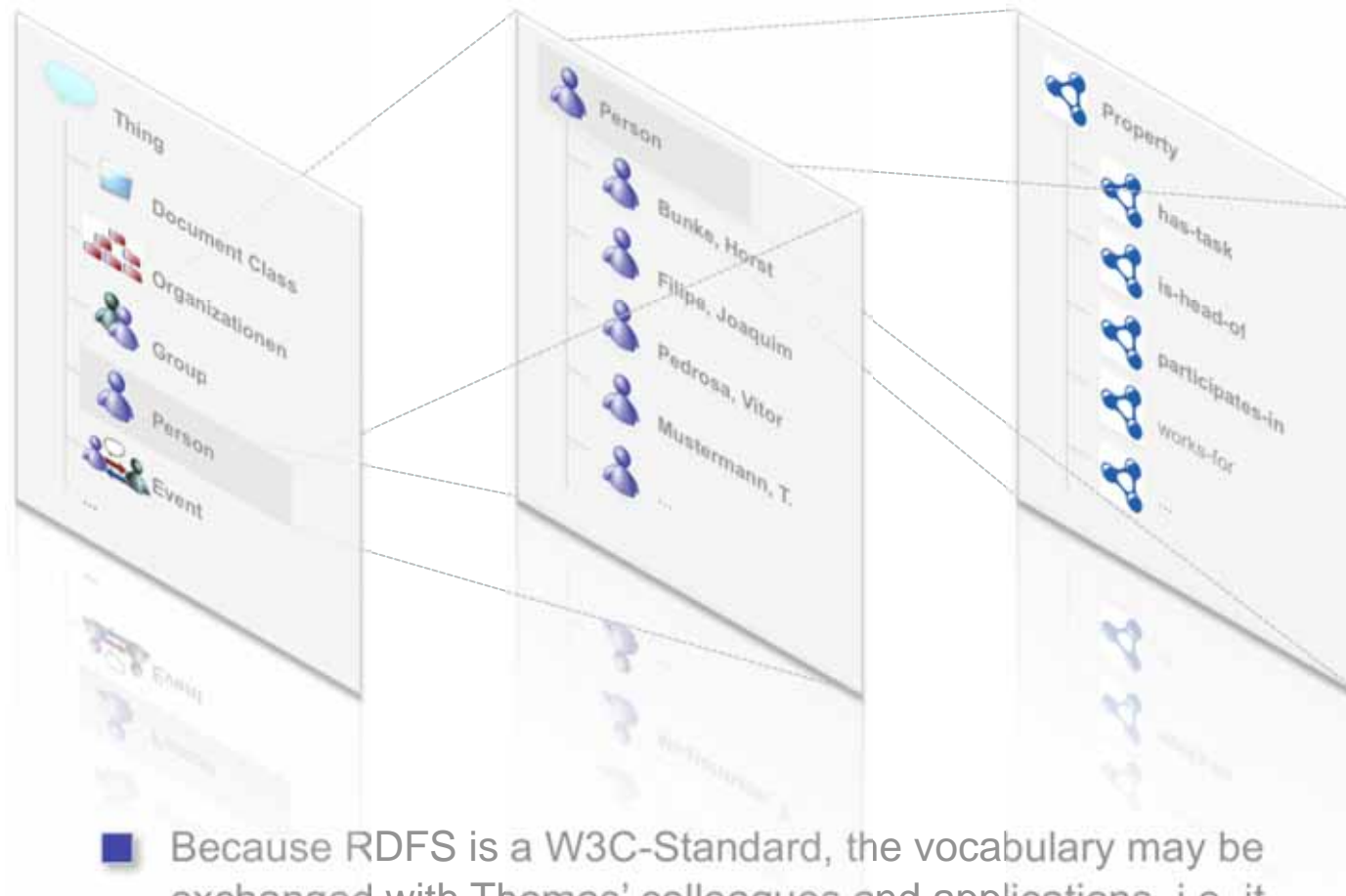
■ Because RDFS is a W3C-Standard, the vocabulary may be exchanged with Thomas' colleagues and applications, i.e. it may be provided via a server

UAB

works-for

has-attachment

Agenda #18

DFKI

sent-to

has-address

Josep Llados

Invitation #432

works-for

08193 Bellaterra Spain

sent-from

Thomas Mustermann

has-address

has-task

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda. I am glad to accept and confirm my participation in the Semantic Desktop course on October 21,

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

Semantic Desktop Course

has-date

has-location

Trippstadter Straße 122 67663 Kaiserslautern

21-10-09

„Vitor confirms invitation"

Thomas makes use of an ontology-based document understanding system helping him to extract relevant facts from all resources

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda. I am glad to accept   and confirm my participation in the Semantic Desktop course on October 21,

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

Incoming Email

Ontological Knowledge

Ontology-Based
Document Understanding

PIMO

New Facts

Information extraction stepwize transforms the contents of documents into knowledge relating it to the existing ontology

Step	Description
Segmentation	Segmentation of the text into paragraphs, sentences, and words
Indexing	Description of the document using intrinsic text features
Categorization	Document classification
Attribute Assignement	Identification of potential attributes
Urification	Matching of attributes with known classes and instances
Contextualization	Extraction of properties and new instances

In a first step, the meta data is extracted from the header of the email and the document is classified

INSTICC

works-for

has-attachment

Agenda #18

sent-to

DFKI

has-address

Vitor Pedrosa

Invitation #432

Av. D. Manuel I, 27A 2 esq. 2910-595 Setùbal Portugal

sent-from

works-for

Segmentation

Indexing

Thomas Mustermann

Categorization

Attribute Assigment

Urification

has-task

Contextualization

Semantic Desktop Course

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda. I am glad to accep
and confirm my participation in the Semantic Desktop c
on October 21,

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

has-date

has-location

From:

Vitor Pedrosa

To:

Thomas Mustermann

Time:

12-10-2009 11:33 am

Trippstadter Straße 122 67663 Kaiserslautern

21-10-09

Confirmation

Subsequently, potential attributes are recognized and it is checked whether they fit into the existing knowledge

INSTICC

works-for

Agenda #18

has-attachment

DFKI

sent-to

Vitor Pedrosa

Invitation #432

works-for

Segmentation

Indexing

Categorization

Attribute Assigment

Urification

Contextualization

has-address

Av. D. Manuel I, 27A 2 esq.
2910-595 Setùbal
Portugal

sent-from

Thomas Mustermann

has-task

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda I am glad to accept and confirm my participation in the Semantic Desktop course on October 21.

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setùbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

e.g.

Semantic Desktop Course

has-date   *has-location*

Trippstadter Straße 122
Kaiserslautern

21-10-09

From:
Vitor Pedrosa

To:
Thomas Mustermann

Time:
12-10-2009
11:33 am

Confirmation

Based on that new relations (properties) may be extracted and incorporated into the ontological context of the message

From: Vitor Pedrosa <vitor@insticc.org>
Date: October 12, 2009 11:33:17
To: thomas.mustermann@dfki.de
Subject: Re: Semantic Desktop

Hi Thomas -

its nice to hear from you.

Many thanks for sending the agenda. I am glad to accept and confirm my participation in the Semantic Desktop course on October 21,

Warm regards,
Vitor

Vitor Pedrosa
INSTICC
Av. D. Manuel I, 27A 2 esq.
2910-595 Setúbal
Portugal
Tel: 34 93 581 24 03
Fax: 34 93 581 16 70
Email: vitor@insticc.org

The email content is related to the existing knowledge via semantic hyperlinks
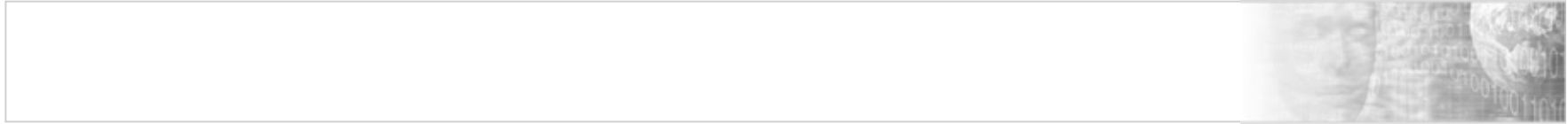
Let's have a look for the real technology!

iDocument is a system for Ontology-based Document Understanding

- Documents are interpreted using ontologies as background knowledge

- New documents are semantically annotated and connected to the knowledge base

- New knowledge is created if contained in the document

relevant tokens

entity recognition

free-text document

fact recognition
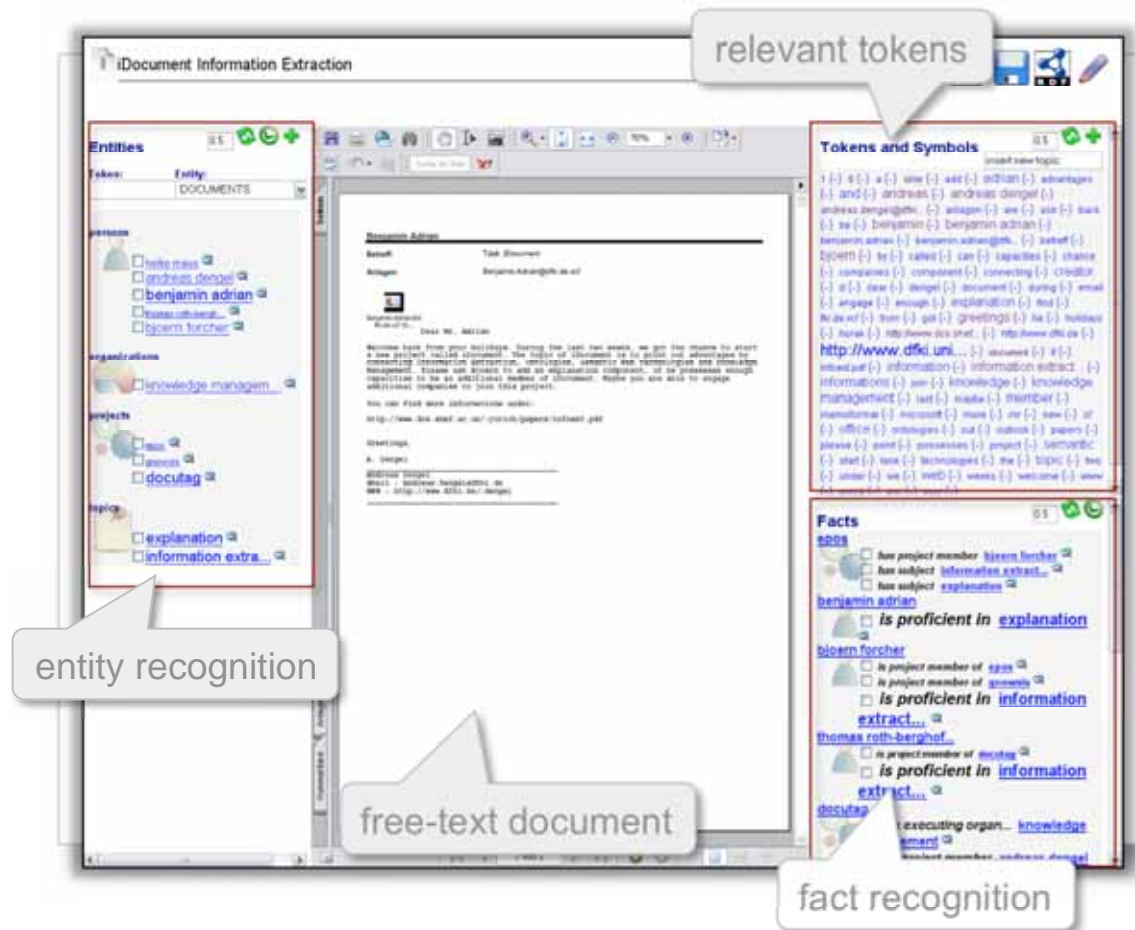
B. Adrian and A. Dengel, *Believing Finite-State cascades in Knowledge-based Information Extraction*
KI 2008: Advances in Artificial Intelligence, Springer, 2008, Springer LNAI 5243, 152-159

B. Adrian, M. Klinkigt, H. Maus, and, A. Dengel, *Using iDocument for Document Categorization in the Nepomuk Social Semantic Desktop*,
Proceedings 9th iSemantics 2009, Graz, Austria (Sept. 2009), accepted for publication.

PIMO

The Semantic Desktop is a means to manage all personal information across application borders based on a shared vocabulary

http://nepomuk.semanticdesktop.org

L. Sauermann, A. Bernardi and A. Dengel, *Overview and Outlook on the Semantic Desktop*, Proceedings International Semantic Web Conference, Galway, Ireland (Nov. 2005), pp. 1-19.

A. Dengel, *Knowledge Technologies for the Social Semantic Desktop*,
in: Z. Zhang and J. Siekmann (Eds.): Proc. KSEM 2007,, LNAI 4798, Springer Publ. (Nov. 2007), pp. 2-9.

Semantic Applications

PIMO Editor

Personal Wiki

Application Plugins

Tagging

Tagging

LogicalMediaType

loves

livesIn

Desktop Applications

web interfaces

Semantic Desktop Server

Domain Ontologies

Sesame2 Repository

RDF

Clustering

Ont. Matching

PIMO Store

Tagging

GUI invocation

Resource Store

Crawler

Desktop Search

Configuration Store

User Context

Service Store

Lucene Index

Aperture Crawlers

Outlook

Files

e-mail

Outlook

filesystem

e-mail server

Classes, instances, and properties of the underlying domain

PIMO

The Semantic Desktop is a means to manage all personal information across application borders based on a shared vocabulary

Relevant facts are extracted from all resources and proposed to the user for confirmation

http://nepomuk.semanticdesktop.org

L. Sauermann, A. Bernardi and A. Dengel, *Overview and Outlook on the Semantic Desktop*, Proceedings International Semantic Web Conference, Galway, Ireland (Nov. 2005), pp. 1-19.

A. Dengel, *Knowledge Technologies for the Social Semantic Desktop*, in: Z. Zhang and J. Siekmann (Eds.): Proc. KSEM 2007,, LNAI 4798, Springer Publ. (Nov. 2007), pp. 2-9.

Symbols relationships, and facts may be explained
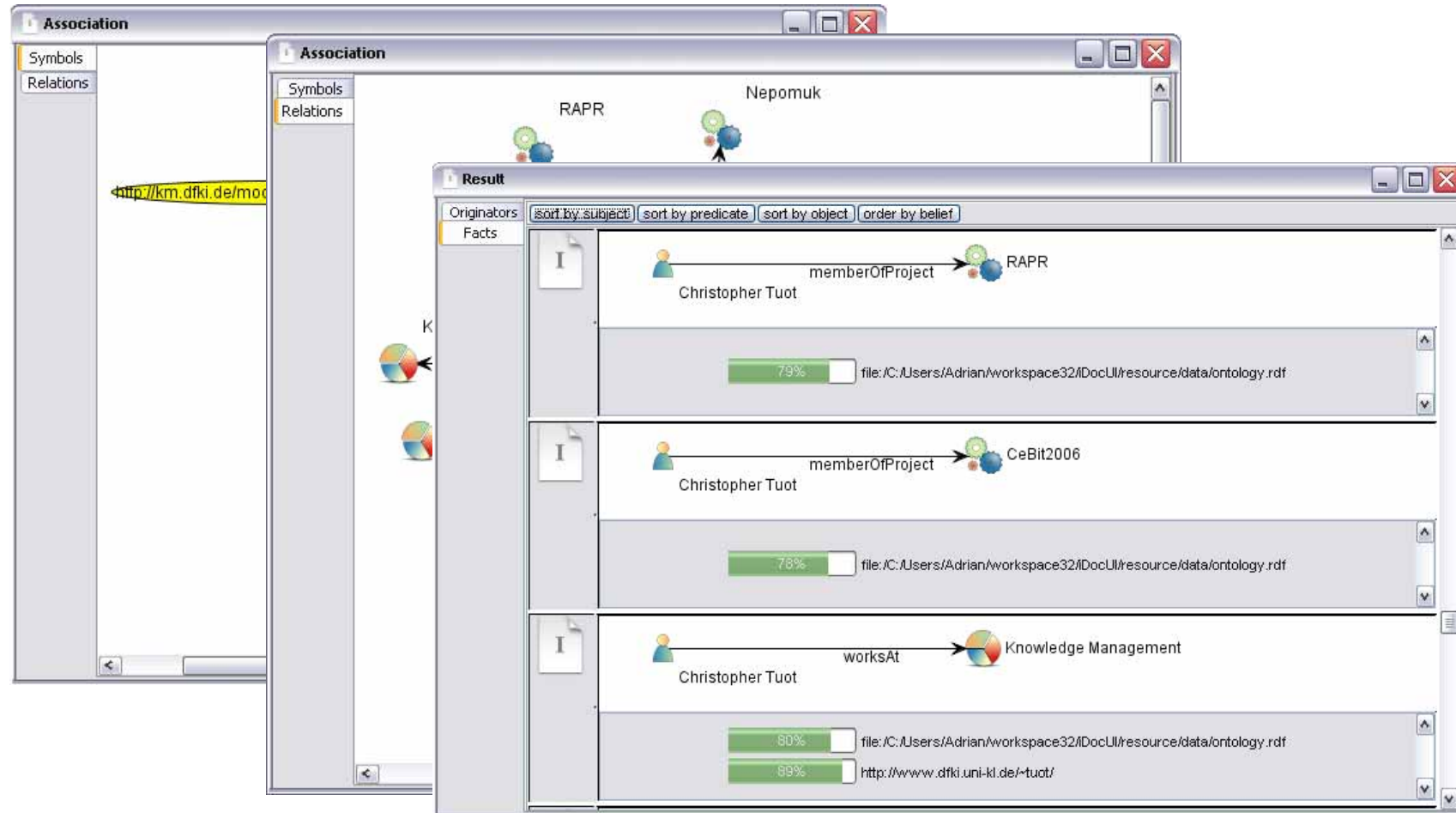
B. Adrian, B. Forcher, T. Roth-Berghofer, and A. Dengel, *Explaining Ontology-Based Information Extraction in the NEPOMUK Semantic Desktop*, Proceedings EXACT 2009, IJCAI Workshop on Explanation-Aware Computing, Pasadena, CA, (July 2009), accepted for pulication

The Semantic Desktop acts as a personal "information butler"

PIMO

http://nepomuk.semanticdesktop.org

- The Semantic Desktop is a means to manage all personal information across application borders based on a shared vocabulary

- Relevant facts are extracted from all resources and proposed to the user for confirmation

- By combining the PIMO with active user observation, the Semantic Desktops acts like an information assistant offering context-aware services

L. Sauermann, A. Bernardi and A. Dengel, *Overview and Outlook on the Semantic Desktop*, Proceedings International Semantic Web Conference, Galway, Ireland (Nov. 2005), pp. 1-19.

A. Dengel, *Knowledge Technologies for the Social Semantic Desktop*, in: Z. Zhang and J. Siekmann (Eds.): Proc. KSEM 2007,, LNAI 4798, Springer Publ. (Nov. 2007), pp. 2-9.

Context identification is a cornerstone of multi-context knowledge work support
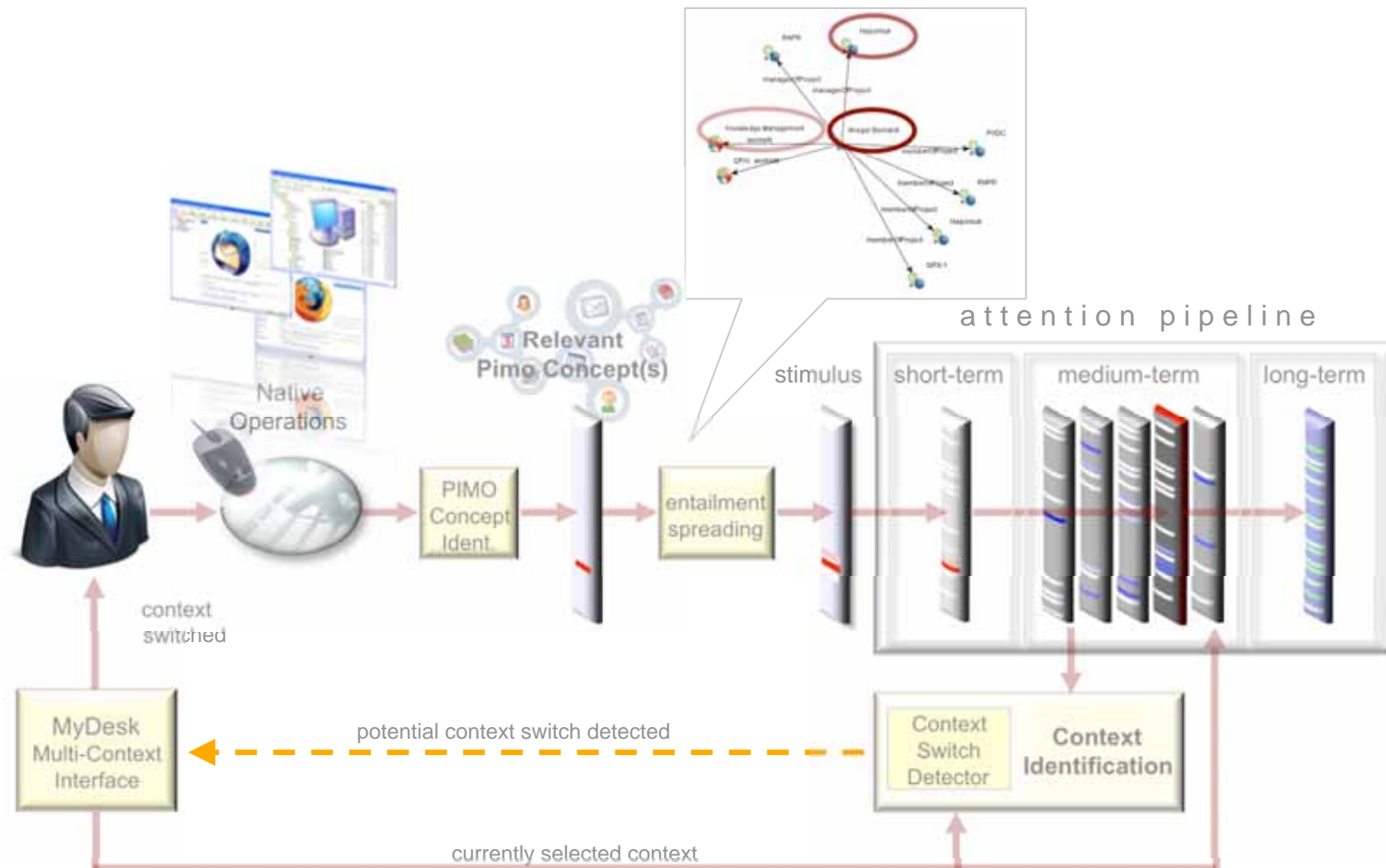
Sven Schwarz: *A Context Model for Personal Knowledge Management Applications*. In Modeling and Retrieval of Context, 2nd International Workshop, MRC 2005, Edinburgh, UK, July 31 - August 1, 2005
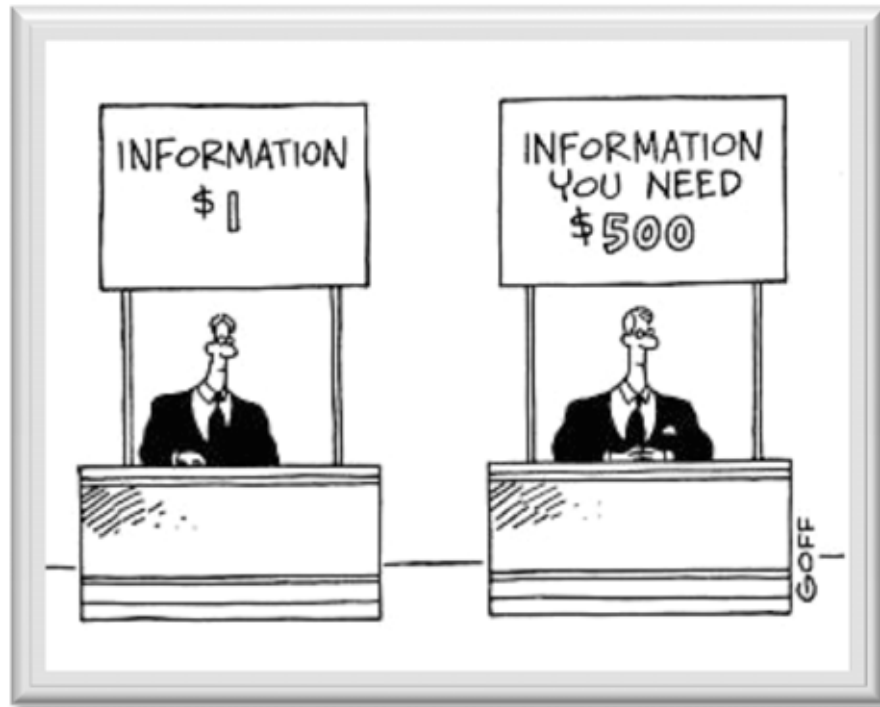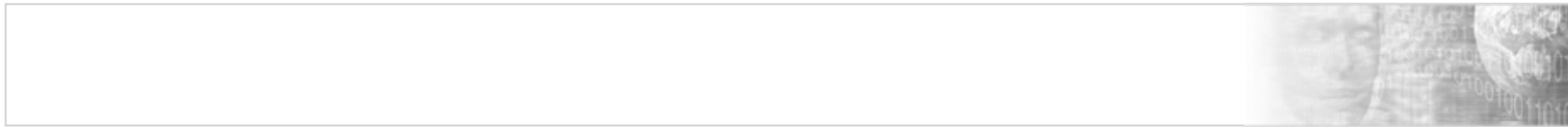
... what is the increased value of the Semantic Desktop

http://dynaq.opendfki.de

H. Holz, H. Maus, A. Bernardi, O. Rostanin, *From Lightweight Proactive Information Delivery to Business Process Oriented Knowledge Management*,
Journal of Universal Knowledge Management, JUKM, No. 5 (2005), pp. 101-127

A DropBox provides a service for the Semantic Desktop for filing and conceptualization
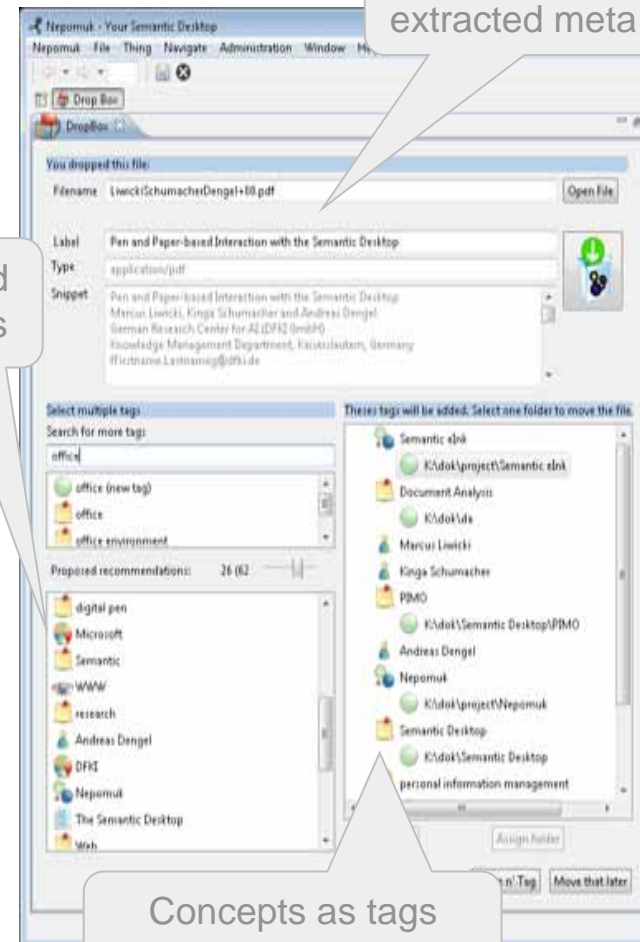
Document and extracted meta data

proposed concepts

Concepts as tags and connected folders

In addition we use a Semantic Wiki as part of a personal knowledge space

M. Kiesel and L. Sauermann, *Towards Semantic Desktop Wikis*,
RADE, Special Issue on "The Semantic Web", Volume VI, pp. 30-34

A user may make use of our gnowsis semantic search

User's information sources on his/her desktop

L. Sauermann, G.A. Grimnes, M. Kiesel, C. Fluit, D. Heim, D. Nadeem, B. Horak and A. Dengel, *Semantic Desktop 2.0: The Gnowsis Experience,* Proceedings 5th Int'l Semantic Web Conference, Athens, GA, USA LNCS 4273, Springer Publ. (Nov. 2006), pp. 887-900.
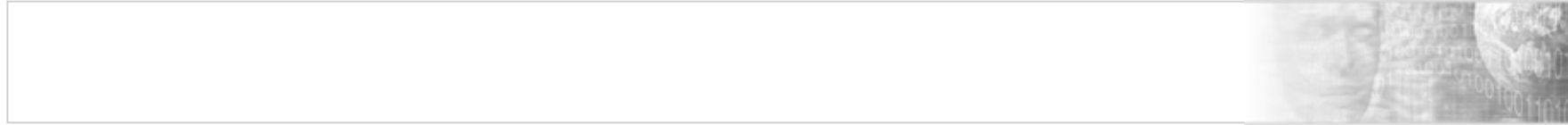
How to integrate Gutenberg's World?

Combing the PIMO with OCR technology allows to recognize entities in printed documents

H. Maus and A. Dengel, *Semantic Annotation of Paper-Based Information*,
Proceedings CBDAR 2008, Curitiba, Brasil (Sep. 2007), pp. 158-160.

Contents of paper documents may be easily intregated into the Semantic Desktop

- Transformation of printed documents combined with OCR and Google book search

- Knowledge acquisition from the document image

  1. Detection of instances (visualization of hypotheses, acceptance via click)

  2. Presentation of existing and potentially new relations between instances

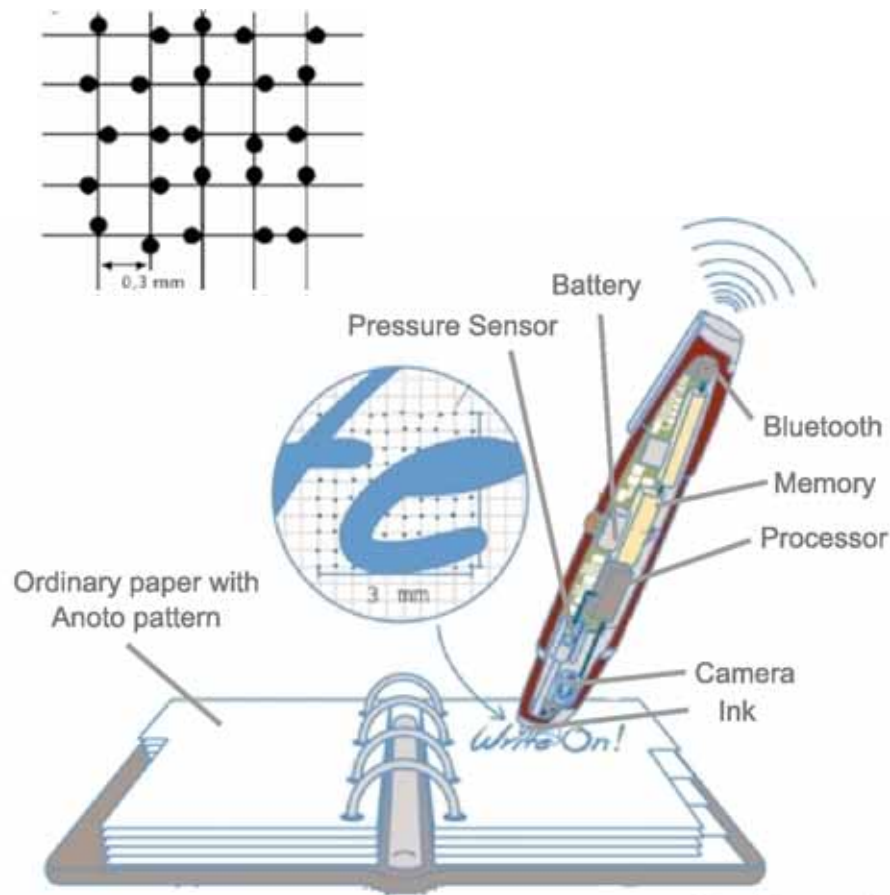 3. Generation of new instances
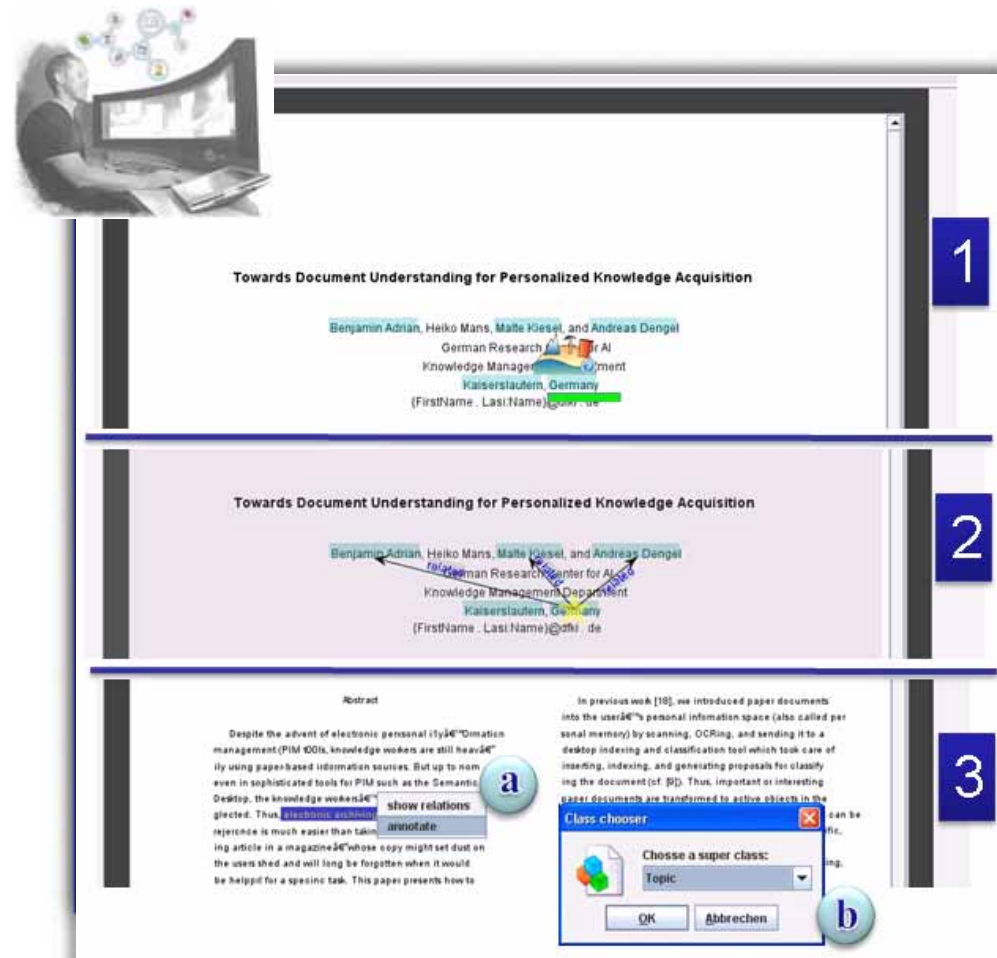
B. Adrian, H. Maus, M. Kiesel, and A. Dengel,
Towards Ontology-based Information Extraction and Annotation of Paper Documents for Personalized Knowledge Acquisition
Workshop on Personal Knowledge Management PKM09 at the WM 09, Solothurn, Switzerland

0,3 mm

Pressure Sensor
Battery
Bluetooth
Memory
Processor
Camera
Ink

Ordinary paper with Anoto pattern

3 mm

Write On!

- With digital pen annotations on paper, content is semantically
  linked to existing contacts, topics, and events
  - Make annotations for document or text parts
  - Use ontological concepts

Seen at the DAS 2008

Pen and Paper-based Interaction with the Semantic Desktop *Title*

Author Marcus Liwicki, Kinga Schumacher and Andreas Dengel
German Research Center for AI (DFKI GmbH)
Knowledge Management Department, Kaiserslautern, Germany
{Firstname.Lastname}@dfki.de

Nadir Weibel, Beat Signer, and Moira C. Norrie
Institute for Information Systems, ETH Zurich, Switzerland *University*
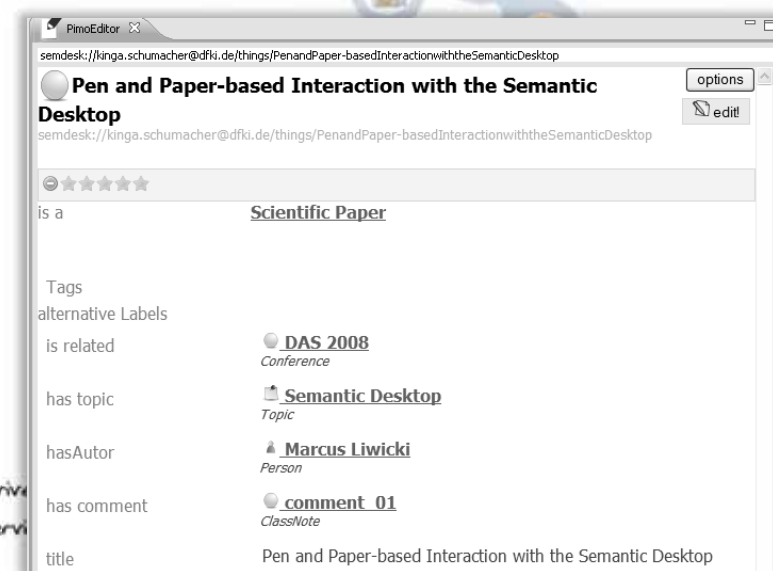{weibel,signer,norrie}@inf.ethz.ch

Semantic Desktop **Abstract**

In this paper we propose a system which recognizes and interprets the semantics of handwritten annotations on printed documents. The semantic information will be sent to the Semantic Desktop, the personal Semantic Web on the desktop computer, which supports users in their information management. This allows a seamless integration of interactive paper into the individual knowledge work. The current implementation of the proposed system works with

made notes and transform them into digital format.
In workflows like reviewing, where the annotations have a meaning for the marked text, the problem of mapping the paper to the digital counterpart arises. A variety of approaches have been investigated to enable this kind of paper-driven digital services. They use cameras, Wacom Graphics Tablets[1], ultrasonic positioning, RFID antennas, bar code readers, or Anoto's Digital Pen and Paper technology[2]. The Anoto technology is particularly interesting because it is based on regular paper and the recording of the

paper-driven
digital servi...

PimoEditor

semdesk://kinga.schumacher@dfki.de/things/PenandPaper-basedInteractionwiththeSemanticDesktop

Pen and Paper-based Interaction with the Semantic Desktop   options

semdesk://kinga.schumacher@dfki.de/things/PenandPaper-basedInteractionwiththeSemanticDesktop   edit!

|	
is a	**Scientific Paper**
Tags	
alternative Labels	
is related	**DAS 2008** *Conference*
has topic	**Semantic Desktop** *Topic*
hasAutor	**Marcus Liwicki** *Person*
has comment	**comment_01** *ClassNote*
title	Pen and Paper-based Interaction with the Semantic Desktop

M. Liwicki, K. Schumacher, A. Dengel, N. Weibel, B. Signer, M. Norrie, *Pen and Paper-based Interaction with the Semantic Desktop*
8th IAPR International Workshop on Document Analysis Systems (DAS 08)

M. Liwicki, M. Weber, and A. Dengel, *Automatic Recognition and Interpretation of Pen-Based Document Annotations*,
Advances in AI, Lecture Notes in Artificial Intelligence, LNAI 5243, Springer Publ., Berlin-Heidelberg (Sept. 2009), to appear.
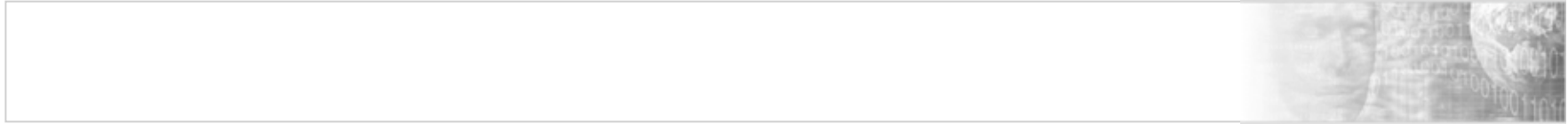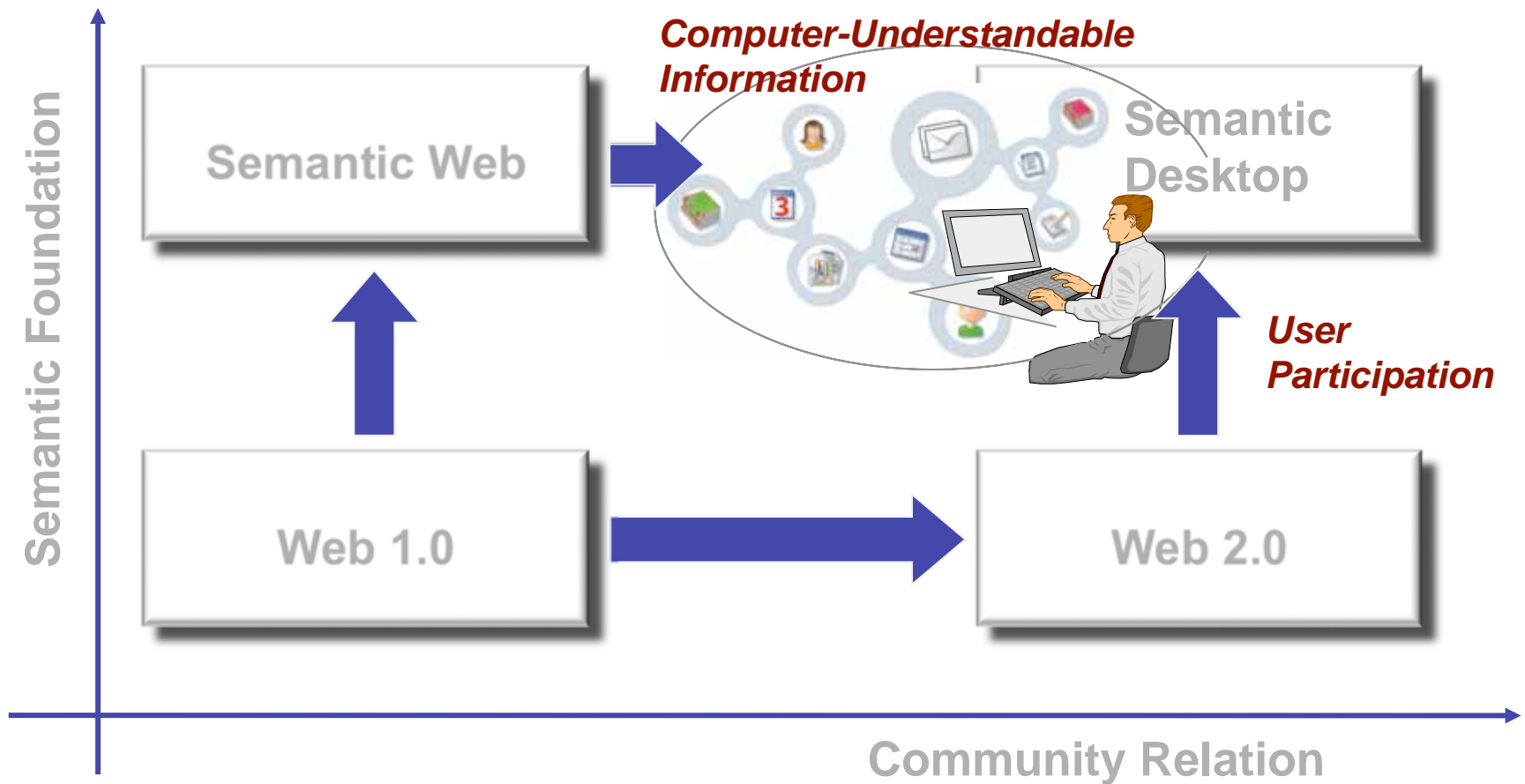
Let me summarize and conclude!
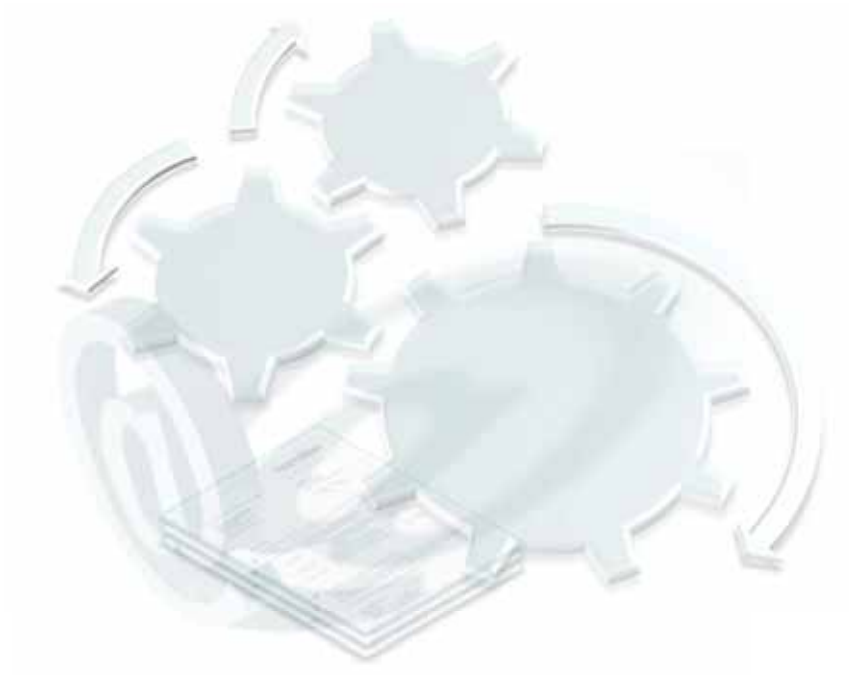
Summary

■ The traditional Web has recently undergone an orthogonal shift into a Web of People/Web 2.0 focusing on collective intelligence, which influences office work as well

■ The Semantic Desktop is a driving paradigm for desktop computing using Semantic Web standards but integrating native office applications and data

■ The Web became part of our thinking and part of our workspace, and the documents we generate at our workspace become part of the Web

■ Nowadays, a document is like a node in a network, a system of links to other documents, to events, locations, persons, or tasks

■ Trusted communities allow people to link with their colleagues (friends) and share information by making use of W3C Standards

... towards the Social Semantic Desktop

Semantic Foundation

Semantic Web

Computer-Understandable Information

Semantic Desktop

User Participation

Web 1.0

Web 2.0

Community Relation

W. Wahlster and A. Dengel*, Web 3.0: Convergence of Web 2.0 and the Semantic Web*,
in: Technology Radar, Feature Paper, Edition II (June 2006), pp. 1-23.

A. Dengel, *Knowledge Technologies for the Social Semantic Desktop*,
in: Z. Zhang and J. Siekmann (Eds.): Proc. KSEM 2007,, LNAI 4798, Springer Publ. (Nov. 2007), pp. 2

Thank you for your attention!

Address:
Prof. Dr. AndreasDengel
DFKI GmbH
P.O. Box 2080
D-67608 Kaiserslautern
email: andreas.dengel@dfki.de
http://www3.dfki.uni-kl.de/agd/dengel/content/index_eng.html