Two Protégé plug-ins for supporting document-based ontology engineering and ontological annotation at document-level
- aka Domain Modeling Tool (DMT)

Viktoria Pammer, Peter Scheir, Stefanie Lindstaedt

aposdle – New ways ...

... to work, learn and collaborate!

Outline

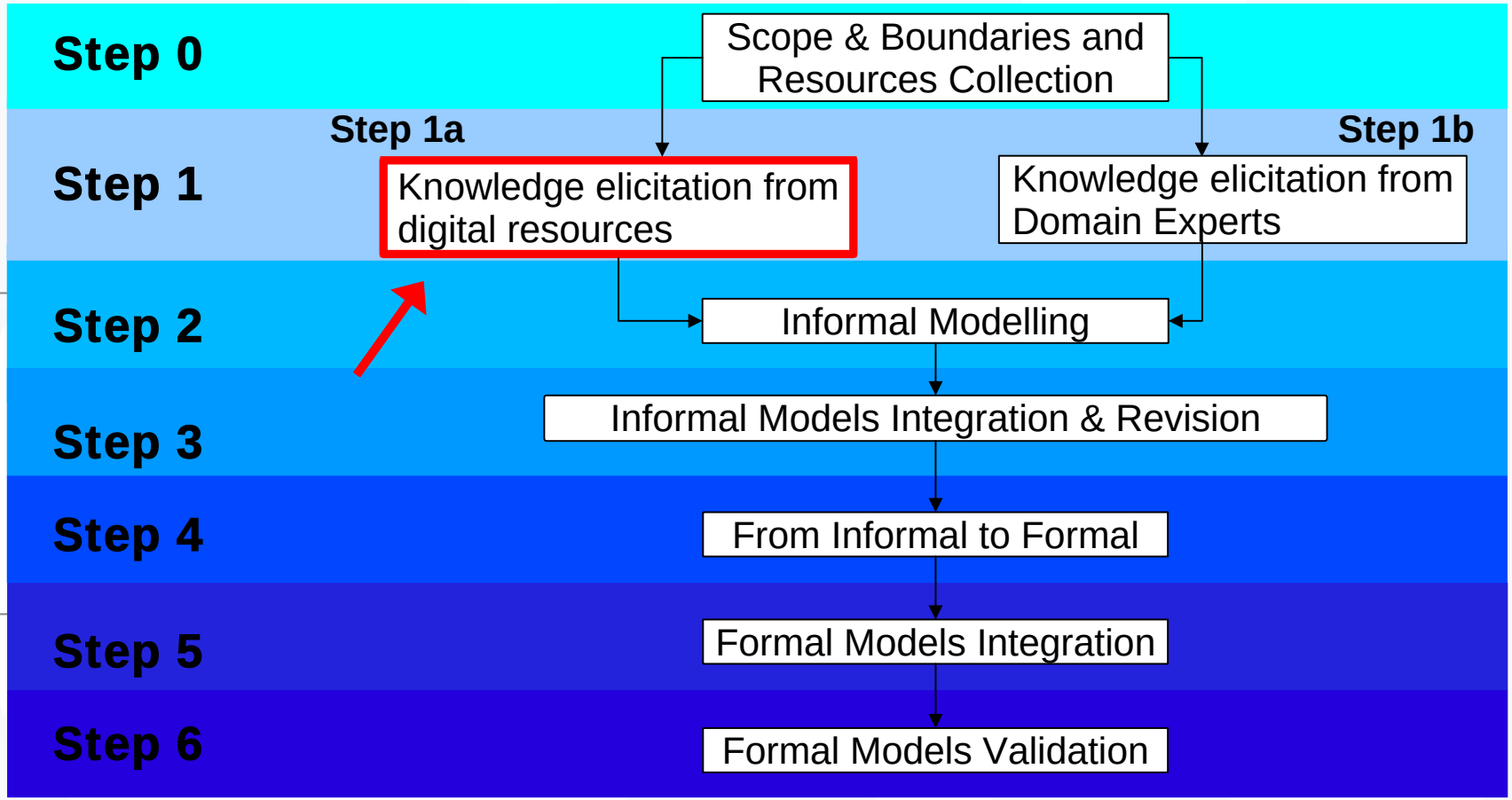
- Motivation
- Document-based ontology-engineering
- Ontology-based document annotation
- Intermediate résumé
- Conclusion

Introduction

- **APOSDLE**
  - System supporting work-integrated learning
  - "**Knows**" about users, tasks, task and application domain knowledge, competencies, resources
  - "**Understands**" resources
  - http://www.aposdle.org

Introduction cont'd

- **Where do we have knowledge?**
  - Heads of experts
  - Documents (FAQ, guidelines, project reports, MATLAB simulations etc.)
- **In what form do we need it?**
  - Formal language (OWL)
- **How do we make use of formal knowledge?**
  - Annotate resources and experts
  - Some other things...



Credits go to Marco Rospocher (FBK) for this picture

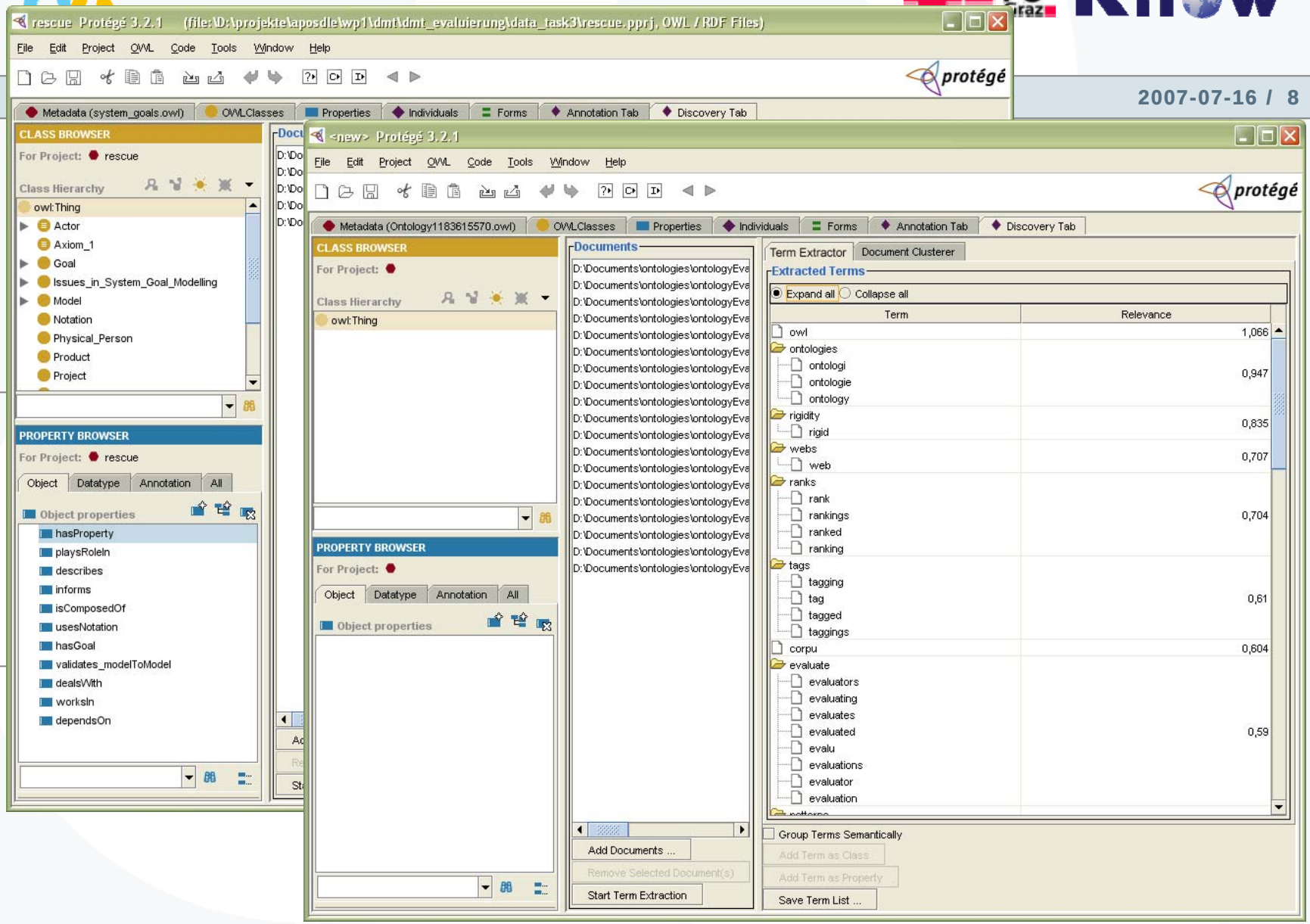
Document-based ontology engineering

- **Given (textual) documents about some domain, build an ontology**
  - The principle is: automatic extraction of a formal knowledge model from non-formal information.
  - Is full automation possible?
- **Question of technology: extract the same from multimedia data, from databases etc.**

DMT: Discovery Tab

- Cluster documents
- Extract terms (and rank them)
- Group terms according to synonymity (WordNet)

- Export extracted terms to textfile
- Create class and property from term

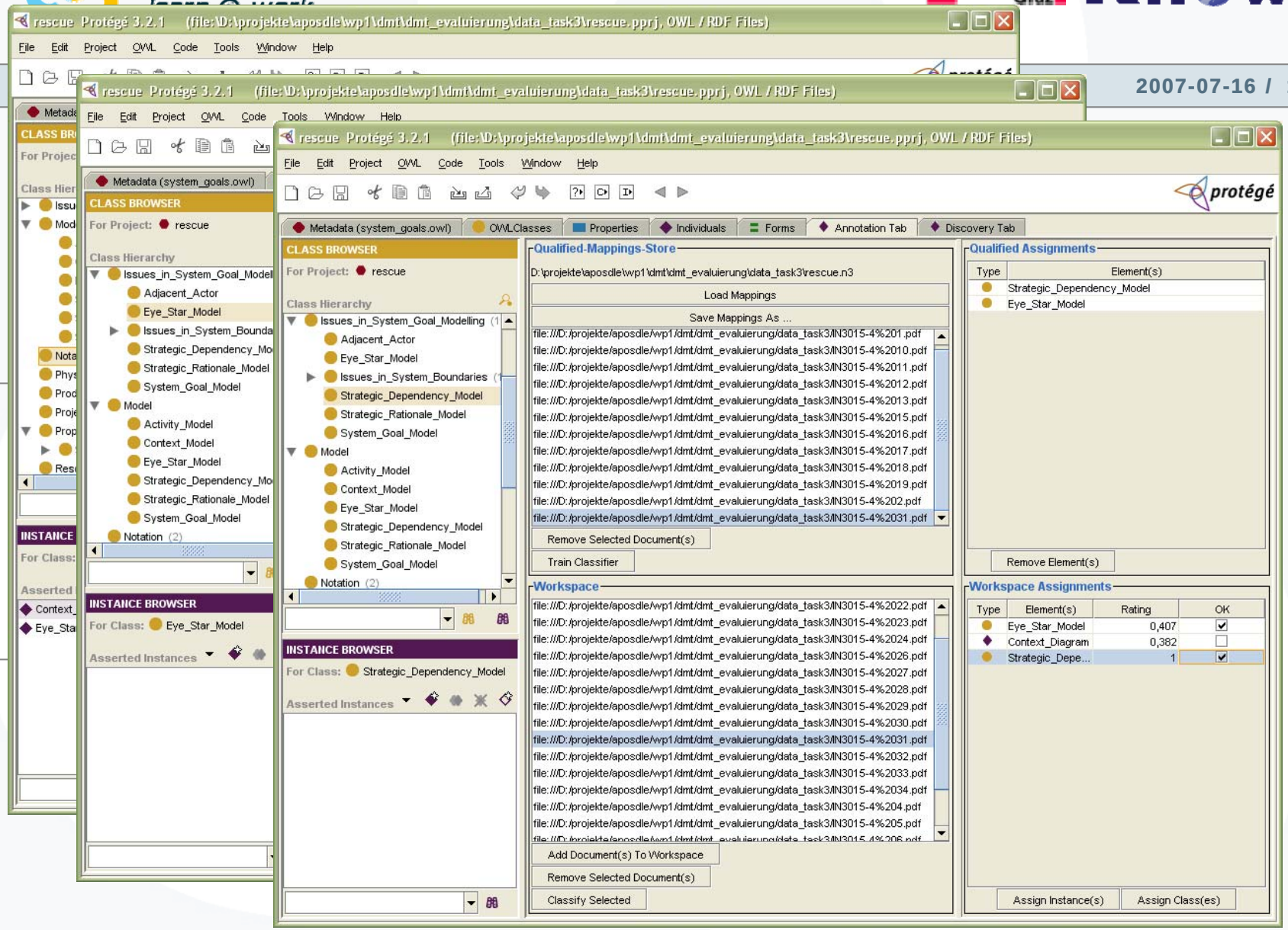

rescue Protégé 3.2.1 (file:\D:\projekte\aposdle\wp1\dmt\dmt_evaluierung\data_task3\rescue.pprj, OWL / RDF Files)
File Edit Project OWL Code Tools Window Help
protégé
Metadata (system_goals.owl) OWLClasses Properties Individuals Forms Annotation Tab Discovery Tab
CLASS BROWSER
For Project: rescue
Class Hierarchy
owl:Thing
Actor
Axiom_1
Goal
Issues_in_System_Goal_Modelling
Model
Notation
Physical_Person
Product
Project
PROPERTY BROWSER
For Project: rescue
Object Datatype Annotation All
Object properties
hasProperty
playsRoleIn
describes
informs
isComposedOf
usesNotation
hasGoal
validates_modelToModel
dealsWith
worksIn
dependsOn
<new> Protégé 3.2.1
File Edit Project OWL Code Tools Window Help
protégé
Metadata (Ontology1183615570.owl) OWLClasses Properties Individuals Forms Annotation Tab Discovery Tab
CLASS BROWSER
For Project:
Class Hierarchy
owl:Thing
PROPERTY BROWSER
For Project:
Object Datatype Annotation All
Object properties
Documents
D:\Documents\ontologies\ontologyEva
Add Documents ...
Remove Selected Document(s)
Start Term Extraction
Term Extractor Document Clusterer
Extracted Terms
Expand all Collapse all
Term Relevance
owl 1,066
ontologies
ontologi
ontologie
ontology 0,947
rigidity
rigid 0,835
webs
web 0,707
ranks
rank
rankings
ranked
ranking 0,704
tags
tagging
tag
tagged
taggings 0,61
corpu 0,604
evaluate
evaluators
evaluating
evaluates
evaluated
evalu 0,59
evaluations
evaluator
evaluation
Group Terms Semantically
Add Term as Class
Add Term as Property
Save Term List ...

Ontology-based document annotation

- **Annotate resources**
- **Use elements of a (formal) domain model for annotation**
  - Advantages over “tagging”: controlled vocabulary (facilitating automatic classification), navigation along ontology possible, **?**
  - Disadvantage: overhead for modeling

DMT: Annotation Tab

- **Manually annotate documents with classes and instances**
  - Store annotations
  - Annotations are in OWL Full
- **Classify new documents**
  - Annotations are suggested with a weight, user needs to accept before these annotations are stored


rescue Protégé 3.2.1 (file:\D:\projekte\aposdle\wp1\dmt\dmt_evaluierung\data_task3\rescue.pprj, OWL / RDF Files)
File Edit Project OWL Code Tools Window Help
protégé
Metadata (system_goals.owl)
OWLClasses
Properties
Individuals
Forms
Annotation Tab
Discovery Tab
CLASS BROWSER
For Project: rescue
Class Hierarchy
Issues_in_System_Goal_Modelling
Adjacent_Actor
Eye_Star_Model
Issues_in_System_Boundaries
Strategic_Dependency_Model
Strategic_Rationale_Model
System_Goal_Model
Model
Activity_Model
Context_Model
Eye_Star_Model
Strategic_Dependency_Model
Strategic_Rationale_Model
System_Goal_Model
Notation (2)
INSTANCE BROWSER
For Class: Strategic_Dependency_Model
Asserted Instances
Qualified-Mappings-Store
D:\projekte\aposdle\wp1\dmt\dmt_evaluierung\data_task3\rescue.n3
Load Mappings
Save Mappings As ...
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%201.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2010.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2011.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2012.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2013.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2015.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2016.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2017.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2018.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2019.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%202.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2031.pdf
Remove Selected Document(s)
Train Classifier
Workspace
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2022.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2023.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2024.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2026.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2027.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2028.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2029.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2030.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2031.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2032.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2033.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%2034.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%204.pdf
file:///D:/projekte/aposdle/wp1/dmt/dmt_evaluierung/data_task3/IN3015-4%205.pdf
Add Document(s) To Workspace
Remove Selected Document(s)
Classify Selected
Qualified Assignments
Type Element(s)
Strategic_Dependency_Model
Eye_Star_Model
Remove Element(s)
Workspace Assignments
Type Element(s) Rating OK
Eye_Star_Model 0,407
Context_Diagram 0,382
Strategic_Depe... 1
Assign Instance(s)
Assign Class(es)

Intermediate résumé + conclusion

- First-cut evaluation:
  - **Users rate convenience-functionality (sorting lists, searching in terms etc.) very high**
  - **Quality of automation (term extraction, clustering, classification) is crucial for non-expert users**
  - **Informal modeling needs to be emphasized**
- A number of standard text-processing methodologies grouped together within one tool
- Embedded in a standard ontology development tool

Thank you!

- http://www.know-center.at/dmt
- **We welcome feedback!**

?,