FAKULTÄT FÜR MATHEMATIK UND INFORMATIK
Institut für Informatik
Prof. Dr. Heiner Stuckenschmidt

# Results of the OAEI Anatomy Track 2009

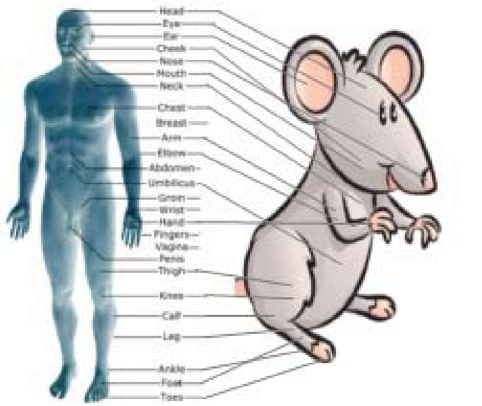
## … and a short outlook on the SEALS project

Christian Meilicke, Heiner Stuckenschmidt
University of Mannheim
{christian,heiner}@informatik.uni-mannheim.de

# Anatomy Track

- Matching ontologies describing mouse anatomy and human anatomy
- For the third time (2007-2009) same reference alignment generated by domain experts

3304 concepts in human ontology



2744 concepts in mouse ontology

1544 correspondences in reference alignment

# Specifics of dataset

- Ontologies have been refactored / harmonized in the process of creating the reference alignment
  - E.g. concept labels named in similar style
  - At least ~ 60% of the reference alignment are detectable by string <u>equality</u> after normalization (= trivial correspondences)

- Introduction of recall+ given an alignment M
  - $\text{Recall+} = |(R \cap M) - S| \,/\, |R - S|$
  - R is the reference alignment
  - S is the trivial part of R

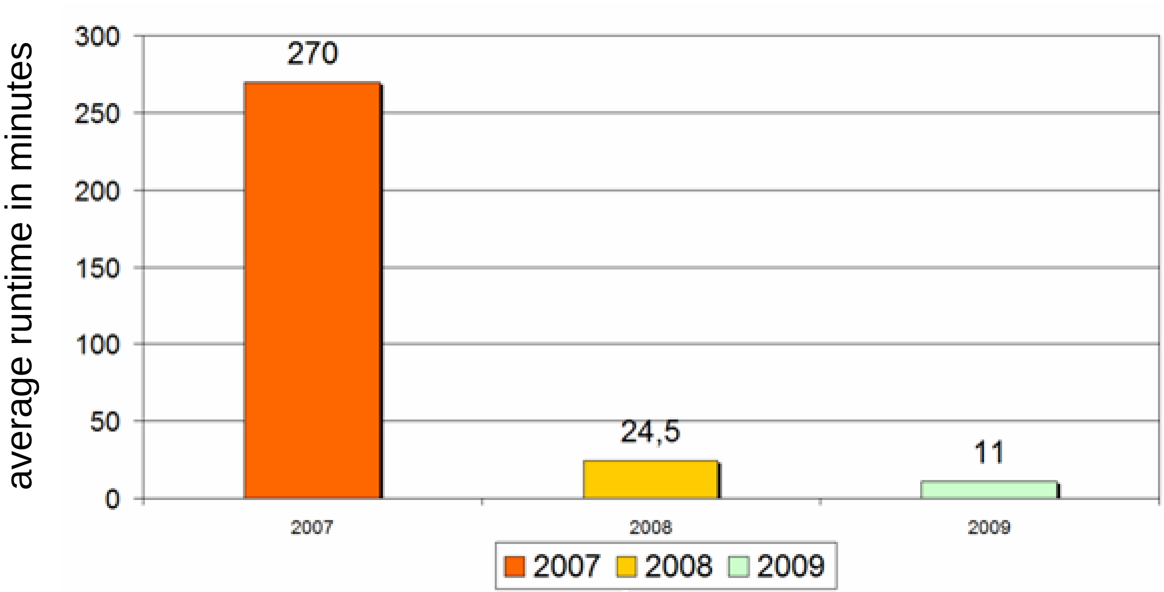
We would like to gratefully thank Martin Ringwald and Terry Hayamizu (Mouse Genome Informatics) who provided us with the reference mapping!

# 4 (+1) Subtasks

- Task #1: Generate an alignment that is as good as possible with respect to the achieved f-value
- Task #2 / #3: Try to generate a highly precise alignment / an alignment with high recall
- Task #4 (new in 2008): Exploit a partial reference alignment as additional input

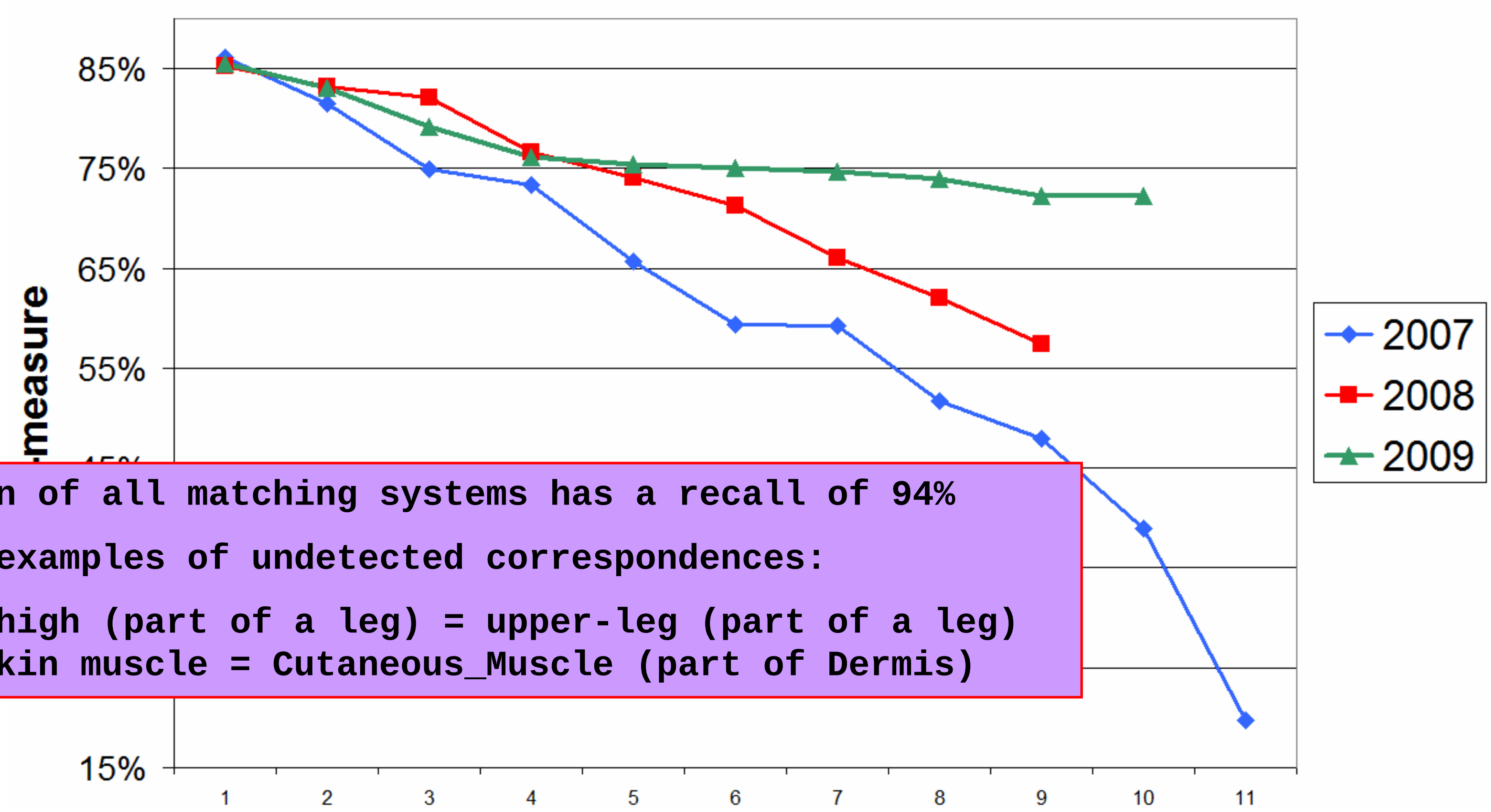
# Runtimes for #1

- Slowest system in 2007: 4 days



- Fastest system in 2009 (AFlood, with acceptable results): 15 seconds

# F-measure for subtask #1



**Union of all matching systems has a recall of 94%**

**Two examples of undetected correspondences:**

**1) thigh (part of a leg) = upper-leg (part of a leg)**
**2) skin muscle = Cutaneous_Muscle (part of Dermis)**

- Significant progress in average & reliable results of each matching system in 2009

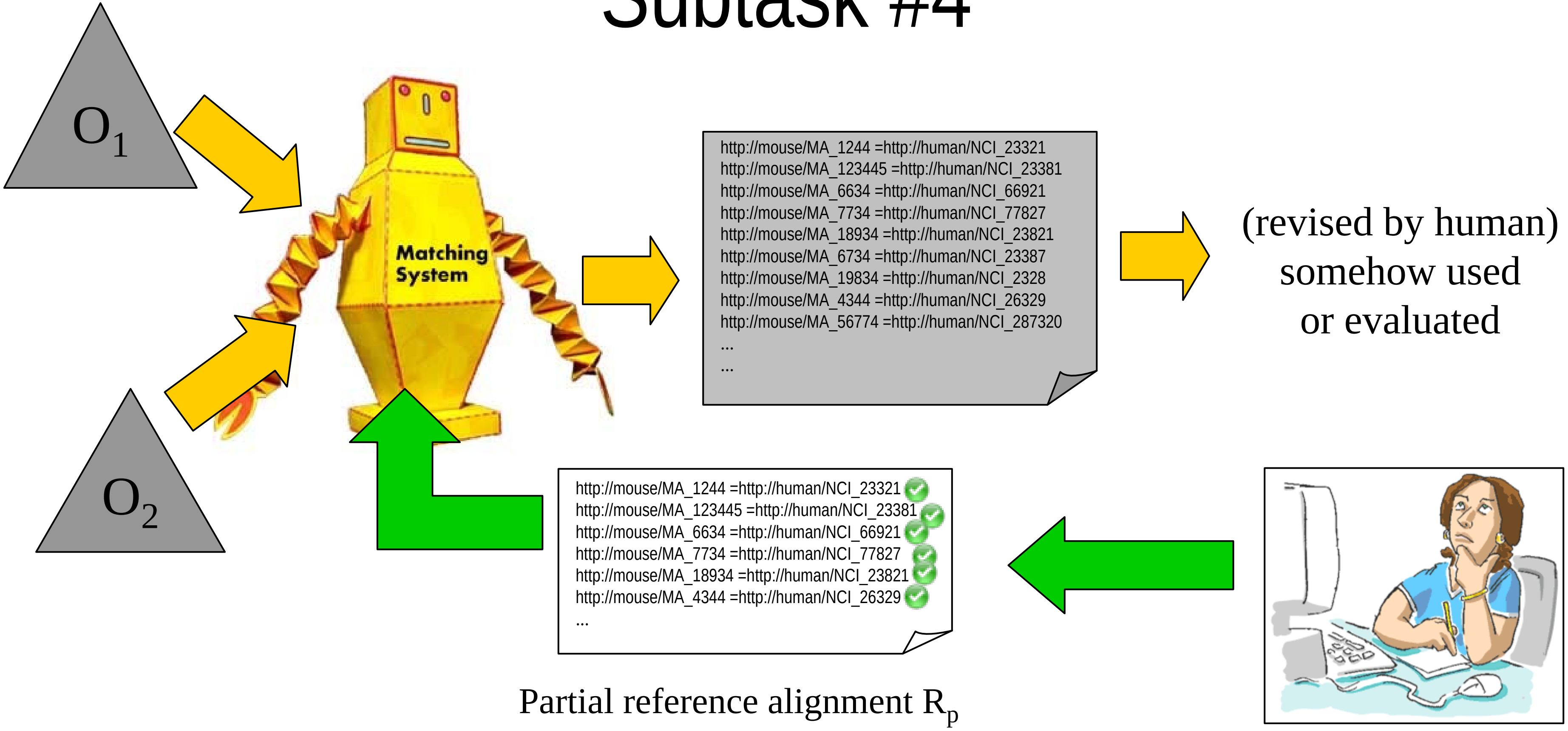
# Matcher Characteristics

| | UMLS | Prec | Rec | Rec+ | $\text{Rec+}_{\#3}$ | F |
|---|---|---|---|---|---|---|
| SOBOM | No | 0.952 | 0.777 | 0.431 | – | 0.855 |
| $\text{SAMBO}_{2008}$ | Yes | 0.869 | 0.836 | 0.586 | – | 0.852 |
| AgrMaker | Yes | 0.865 | 0.798 | 0.489 | 0.553 | 0.831 |
| RiMOM | No | 0.940 | 0.684 | 0.183 | – | 0.792 |
| TaxoMap* | No | 0.870 | 0.678 | 0.222 | 0.319 | 0.762 |
| … | | | | | | |

In favor of increased recall: ASMOV, Aflood, Lily

In favor of precision: AROMA, KosiMap

# Subtask #4



http://mouse/MA_1244 =http://human/NCI_23321
http://mouse/MA_123445 =http://human/NCI_23381
http://mouse/MA_6634 =http://human/NCI_66921
http://mouse/MA_7734 =http://human/NCI_77827
http://mouse/MA_18934 =http://human/NCI_23821
http://mouse/MA_6734 =http://human/NCI_23387
http://mouse/MA_19834 =http://human/NCI_2328
http://mouse/MA_4344 =http://human/NCI_26329
http://mouse/MA_56774 =http://human/NCI_287320
...
...

(revised by human)
somehow used
or evaluated

http://mouse/MA_1244 =http://human/NCI_23321 ✓
http://mouse/MA_123445 =http://human/NCI_23381 ✓
http://mouse/MA_6634 =http://human/NCI_66921 ✓
http://mouse/MA_7734 =http://human/NCI_77827 ✓
http://mouse/MA_18934 =http://human/NCI_23821 ✓
http://mouse/MA_4344 =http://human/NCI_26329 ✓
...

Partial reference alignment $R_p$

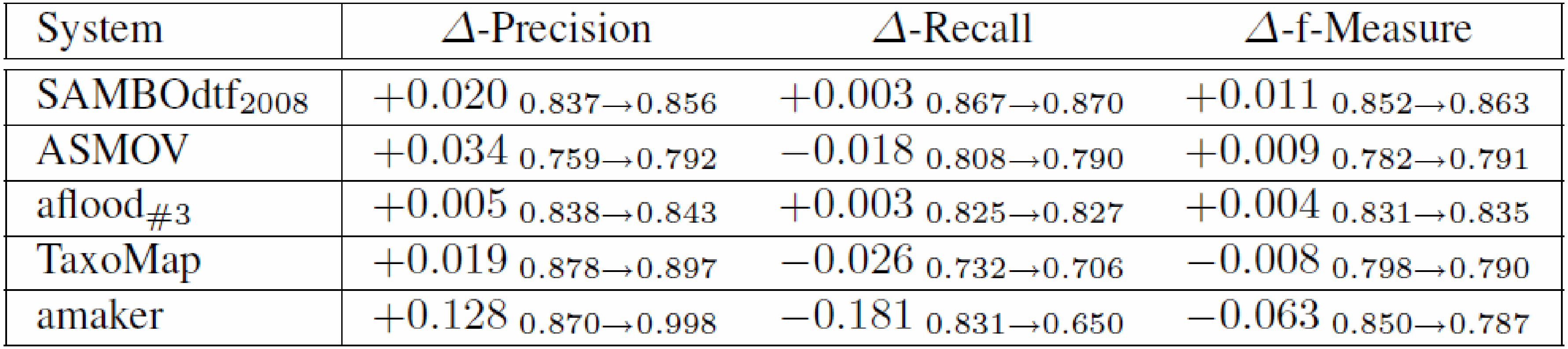
# The partial reference alignment PRA

- Consists of 988 correspondences
  - Trivial part of reference alignment (934)
  - 54 non trivial correspondences (detected by systems in 2007)

  $\Rightarrow$ 535 correct correspondences remain unknown

- Evaluation procedure:
  - Let $A_{\#1}$ be the submission to track #1, analog for $A_{\#4}$
  - Compare $A_{\#1} \cup$ PRA (baseline) and $A_{\#4}$ against the reference alignment R

# Results of subtrack #4



| System | $\Delta$-Precision | $\Delta$-Recall | $\Delta$-f-Measure |
|---|---|---|---|
| SAMBOdtf$_{2008}$ | $+0.020$ $_{0.837 \rightarrow 0.856}$ | $+0.003$ $_{0.867 \rightarrow 0.870}$ | $+0.011$ $_{0.852 \rightarrow 0.863}$ |
| ASMOV | $+0.034$ $_{0.759 \rightarrow 0.792}$ | $-0.018$ $_{0.808 \rightarrow 0.790}$ | $+0.009$ $_{0.782 \rightarrow 0.791}$ |
| aflood$_{\#3}$ | $+0.005$ $_{0.838 \rightarrow 0.843}$ | $+0.003$ $_{0.825 \rightarrow 0.827}$ | $+0.004$ $_{0.831 \rightarrow 0.835}$ |
| TaxoMap | $+0.019$ $_{0.878 \rightarrow 0.897}$ | $-0.026$ $_{0.732 \rightarrow 0.706}$ | $-0.008$ $_{0.798 \rightarrow 0.790}$ |
| amaker | $+0.128$ $_{0.870 \rightarrow 0.998}$ | $-0.181$ $_{0.831 \rightarrow 0.650}$ | $-0.063$ $_{0.850 \rightarrow 0.787}$ |

- PRA is not a representative sample of the reference alignment

- Structural techniques are hard to handle, due to missing links (subsumption, partOf) in the ontologies

Anatomy is a hard test case for this scenario!

# How to improve this evaluation?

- Systematically change the PRA
    - Chose random subset of x% correspondences from R and repeat evaluation many times
    - Compute mean values for different fraction x
- Problem: We want a blind evaluation!
- Find a way to automatically run the matching tools in some evaluation platform
    - Fair and reproducible runtime measurements
    - Supports more sophisticated evaluation scenarios
    - Evaluation services can be used aside from any campaign
    - Requires agreement on interfaces
    - Requires infrastructure that supports different runtime environments

# SEALS

*(Semantic Evaluation at Large Scale)*

## Service Infrastructure

- The SEALS Platform offering Evaluation Services

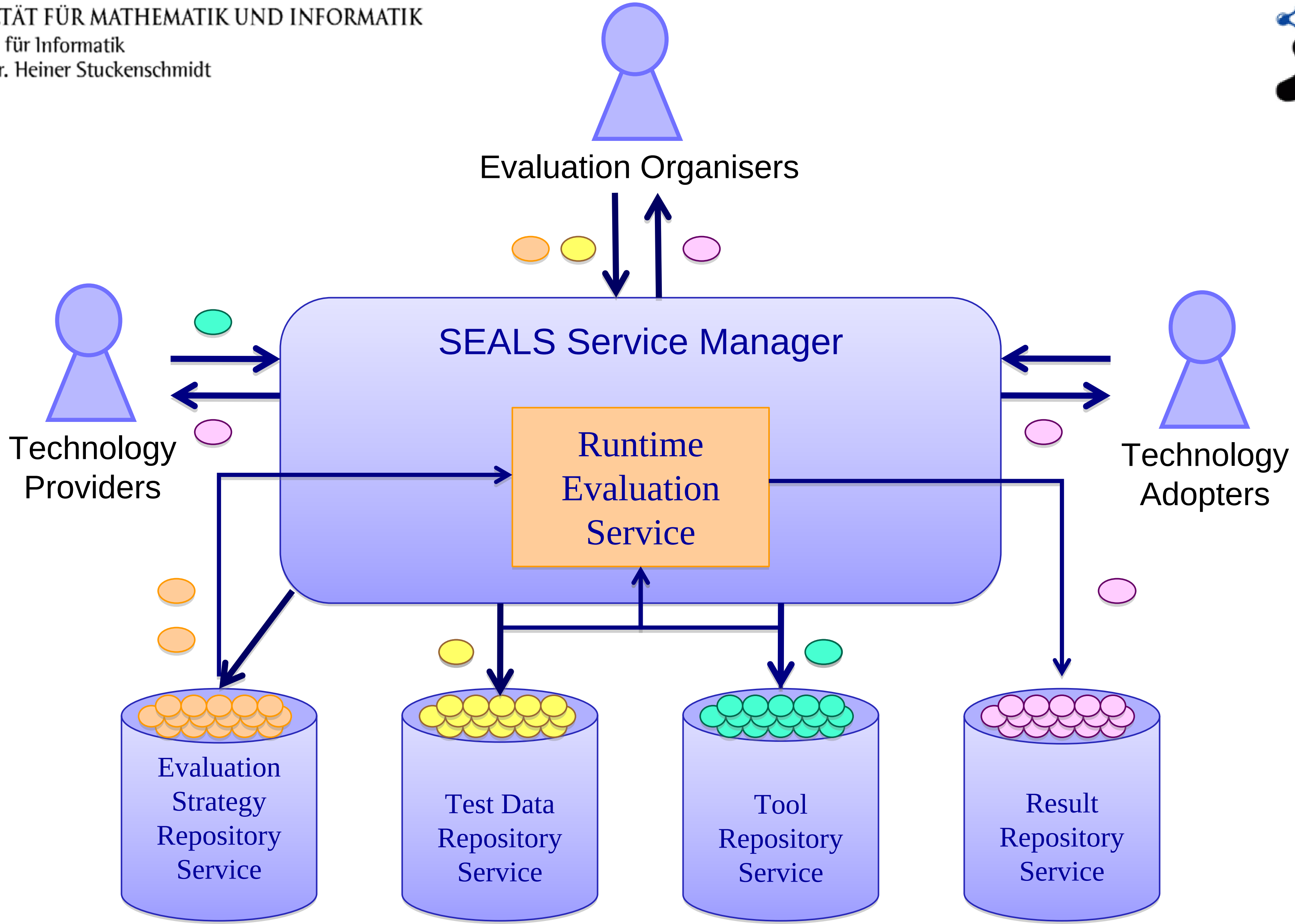
## Research Activities

- Evaluation Definitions
- Automatic Evaluation Execution

- Ontology Engineering Tools
- Storage and Reasoning Tools
- Matching Systems
- Semantic Search Tools
- Semantic Web Service Tools

## Community Building

- Benchmarking Methodology for Evaluation Campaigns
- Organize Evaluation Campaigns
- Semantic Technology Roadmaps (Recommendations)


Evaluation Organisers

SEALS Service Manager

Technology Providers

Technology Adopters

Runtime Evaluation Service

Evaluation Strategy Repository Service

Test Data Repository Service

Tool Repository Service

Result Repository Service

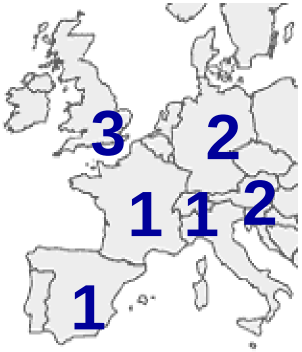
# Involved partners

## The SEALS Consortium

- Universidad Politécnica de Madrid (UPM, Spain)
- University of Sheffield (USFD, UK)
- Forschungszentrum Informatik (FZI, Germany)
- University of Innsbruck (UIBK, Austria)
- Institut National de Recherche en Informatique et en Automatique (INRIA, France)
- University of Mannheim (UMA, Germany)
- University of Zurich (UZH, Switzerland)
- STI International (STI2, Austria)
- Open University (OU, UK)
- Oxford University (OXF, UK)

## Core competencies

- Community building
- Evaluation campaigns
- Infrastructure development
- Ontology Engineering Tools
- Storage and Reasoning Systems
- Matching Tools
- Semantic Search Tools
- Semantic Web Service Tools



# Thanks for your attention

Any questions on Anatomy results?
Any questions on SEALS?