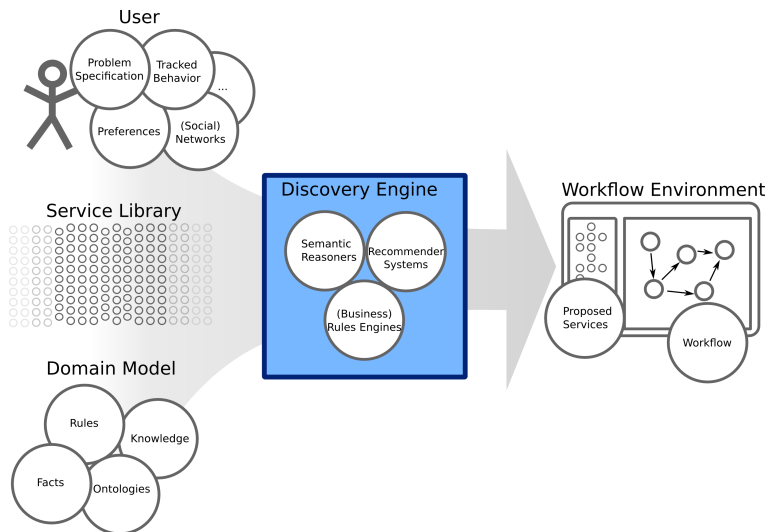


Projektgruppe: Situation-Aware Semantic Service Discovery

Veranstalter:
Anna-Lena Lamprecht
Stefan Naujokat
Lehrstuhl für Programmiersysteme

Outline

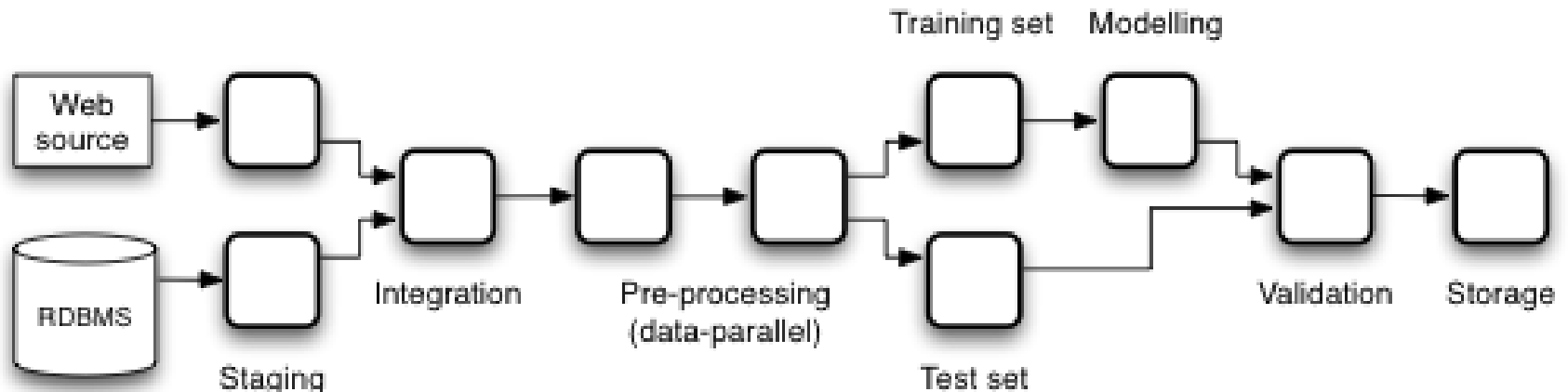


- 1) Workflows
- 2) Discovery
- 3) Recommender Systems
- 4) Example Domain
- 5) Organisatorisches

Workflows & Services

Workflows / Business Processes

- (one) definition of “**workflow**”:
“**an abstract description of steps required for executing a particular real-world process, and the flow of information between them**”

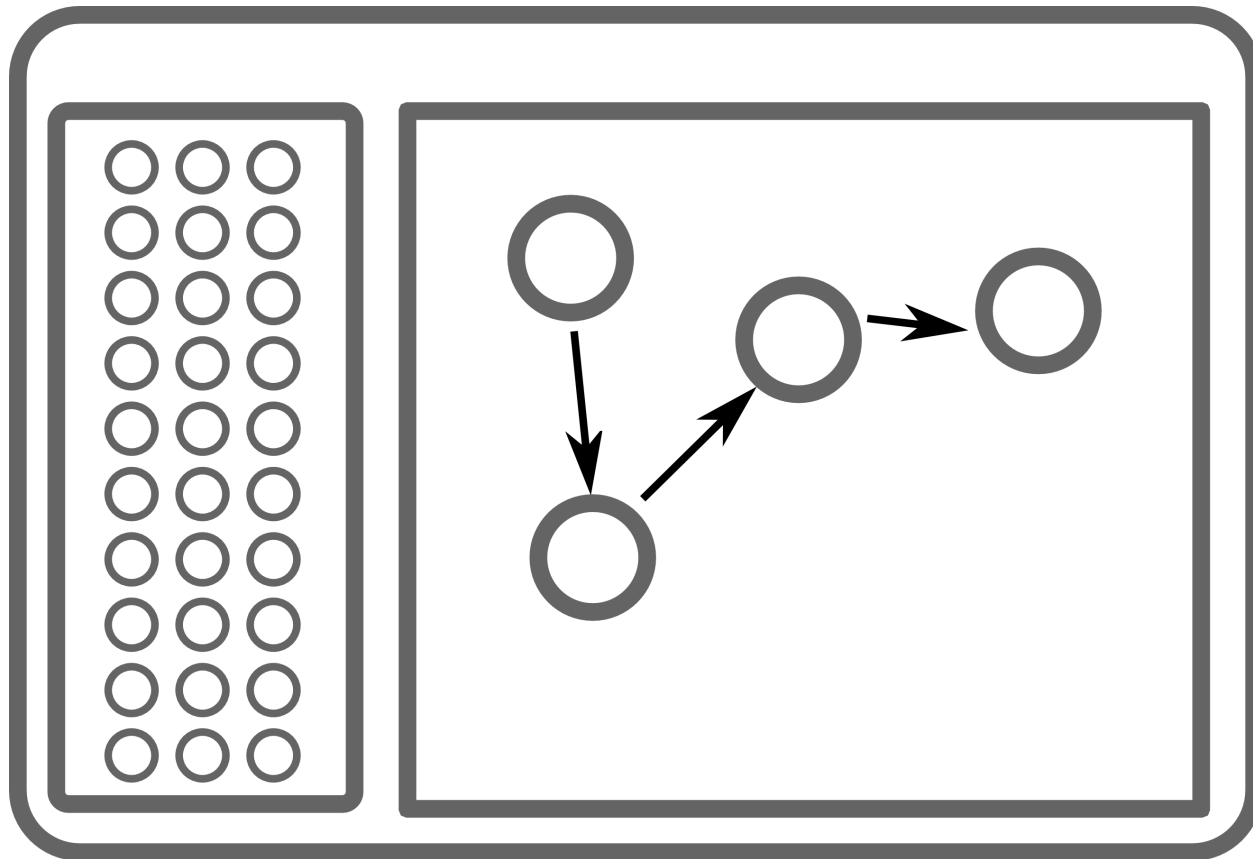


- from: Ghanem, M. & Curcin, V. *Scientific workflow systems - can one size fit all?* Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International, 2008, 1-9

Service orientation

- **Service:**
self-contained, well-defined unit of functionality with a well-defined interface
- **Workflows orchestrate distributed services:**
 - Abstract description of a series of analytical steps that constitute a complex computational experiment, and the flow of information between them
 - Essentially providing the “glue code” for distributed services
- **Loosely coupled, distributed services:**
 - Client independent of the service
 - High degree of **flexibility** and **scalability** for **evolving applications**

Workflow Modeling



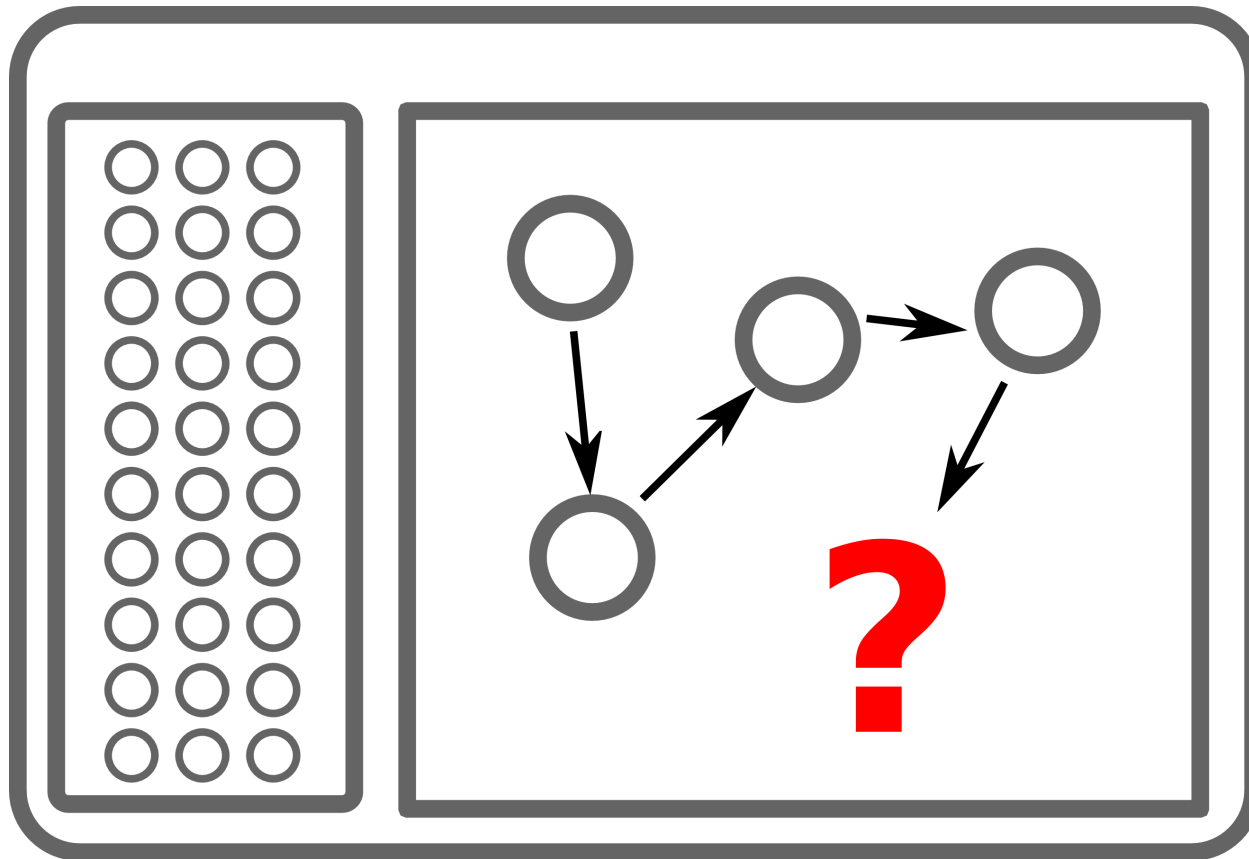
Workflow Systems

- envisaged: **Eclipse4Bio**
 - open-source, developed with Eclipse Modeling Technology
 - textual and graphical workflow representation
 - platform-independent service definition
 - execution with Spring Batch framework

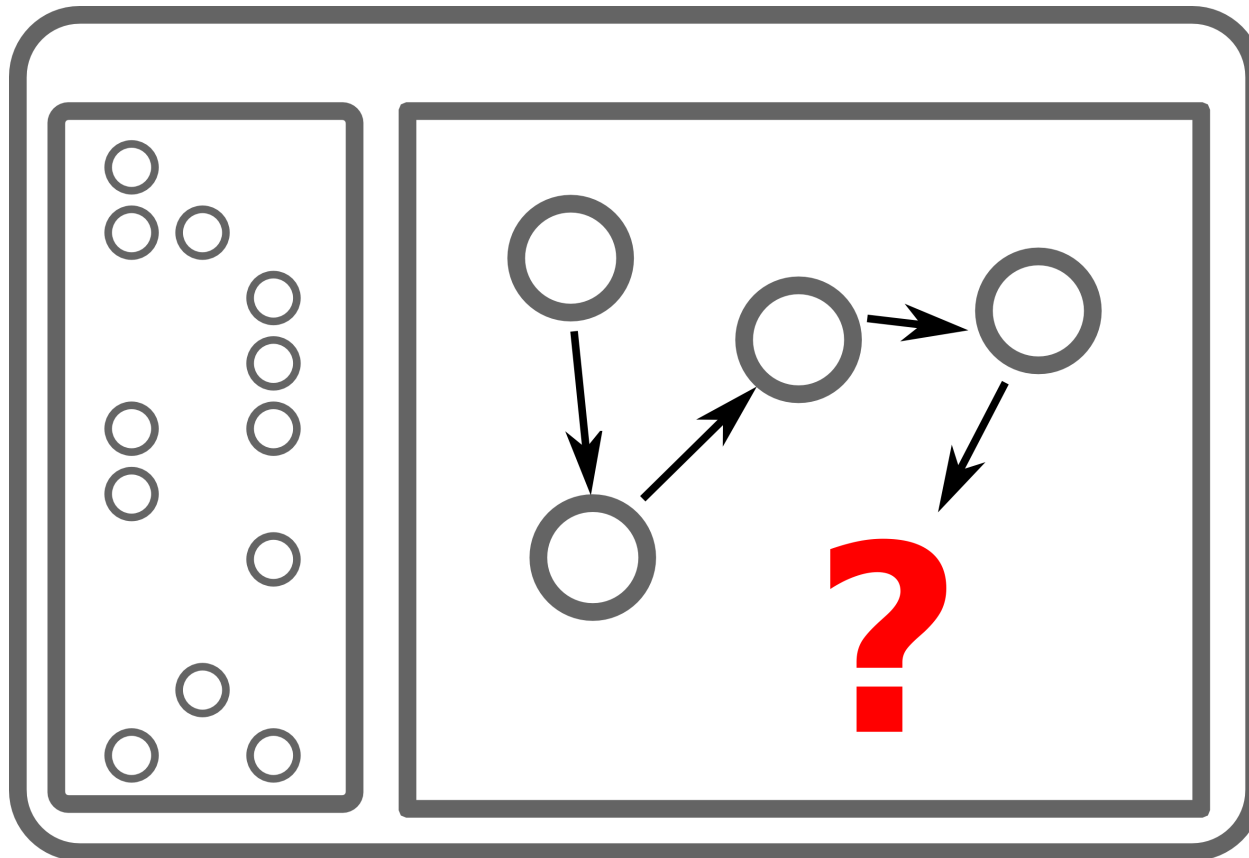
- possible alternatives (e.g.): **jABC, Taverna, ...**

Discovery

Service Discovery

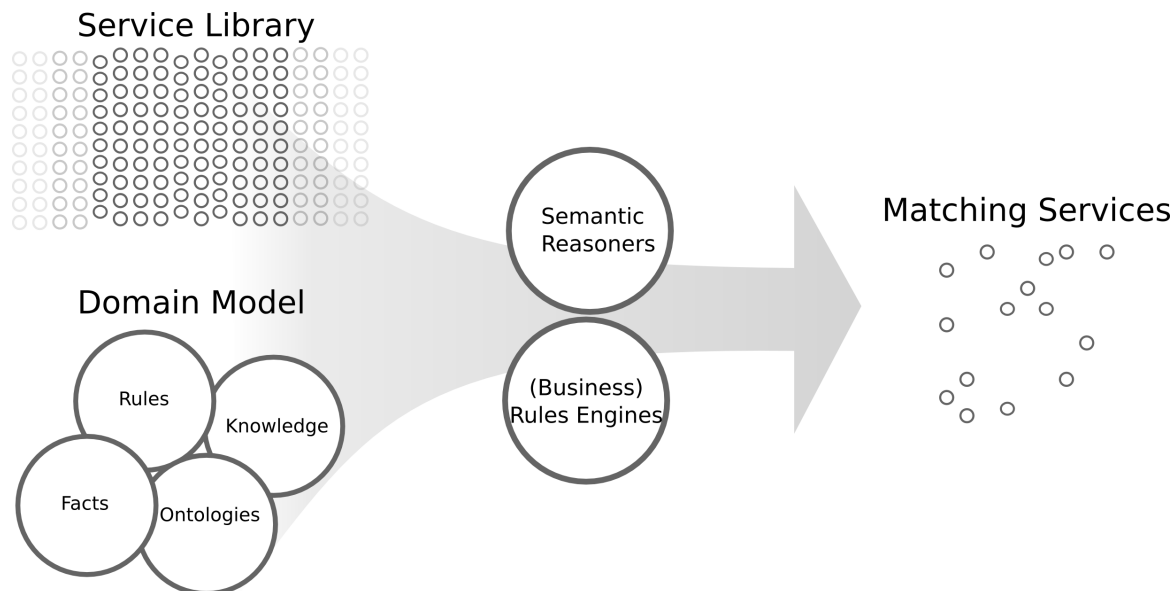


Service Discovery



Service Discovery

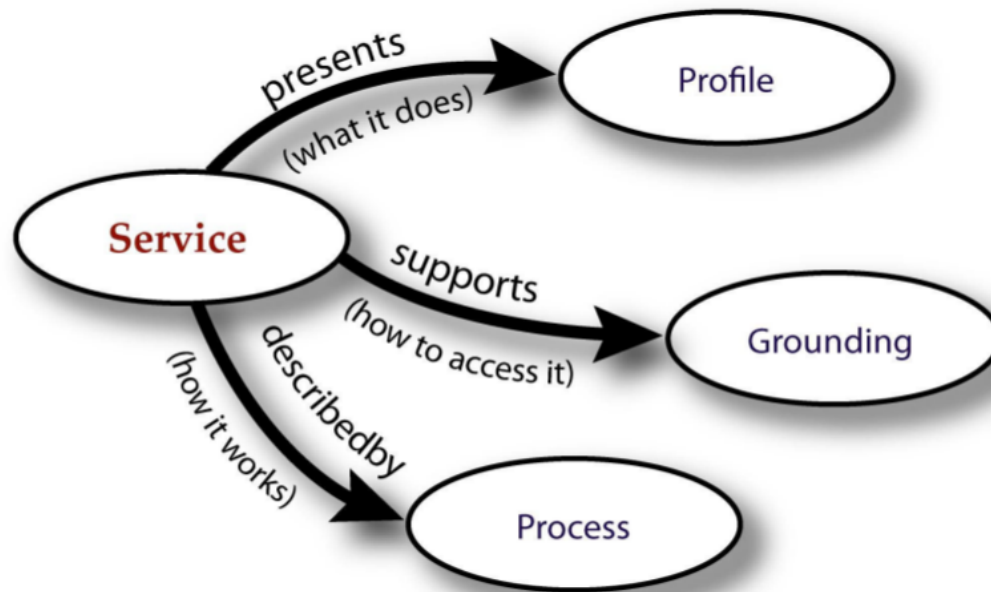
- Given: abstract specification (Query, Constraints, Rules)
- Find: concrete services that fulfill specification
- Syntactic vs. semantic discovery



Ontologies

- consists of
 - Concepts / Classes
 - Individuals / Instances
 - Relations
 - Rules, Constraints, Axioms
- Semantic Reasoning
 - Example: fatherOf(X,Y) relation
 - grandfatherOf(X,Y) can be inferred
- “Semantic Web”
- OWL – Web Ontology Language

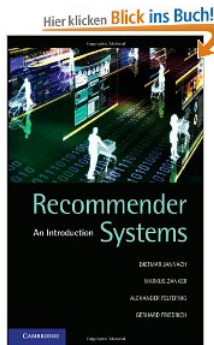
Service Ontologies (e.g. OWL-S)



from: Martin, D.; Paolucci, M. & Wagner, M., Towards Semantic Annotations of Web Services: OWL-S from the SAWSDL Perspective, OWL-S Experiences and Future Developments Workshop at ESWC 2007

Recommender Systems

Amazon's Product Recommendations



Recommender Systems: An Introduction [Gebundene Ausgabe]

[Dietmar Jannach](#) (Autor), [Markus Zanker](#) (Autor), [Alexander Felfernig](#) (Autor), [Gerhard Friedrich](#) (Autor)

★★★★★ (2 [Kundenrezensionen](#)) [Gefällt mir](#) (2)

Unverb. Preisempf.: **EUR 52,22**

Preis: **EUR 38,95** [kostenlose Lieferung](#). [Siehe Details](#).

Sie sparen: **EUR 13,27 (25%)**

Alle Preisangaben inkl. MwSt.

Auf Lager.

Verkauf und Versand durch **Amazon.de**. Geschenkverpackung verfügbar.

Nur noch 9 Stück auf Lager - jetzt bestellen.

Lieferung bis Freitag, 25. November: Bestellen Sie innerhalb **9 Stunden und 40 Minuten** per **Morning-Express**. [Siehe Details](#).

[21 neu](#) ab EUR 38,95 [2 gebraucht](#) ab EUR 44,22

Wird oft zusammen gekauft

Kunden kaufen diesen Artikel zusammen mit [Recommender Systems Handbook](#) Gebundene Ausgabe **EUR 152,45**



+



Preis für beide: EUR 191,40

[Beides in den Einkaufswagen](#)

[Verfügbarkeit und Versanddetails anzeigen](#)

Kunden, die diesen Artikel gekauft haben, kauften auch



**Empfehlungssysteme:
Recommender Systems -
Grund...** von André Klahold

★★★★★ (1)

EUR 29,95



**Recommender Systems
Handbook**

★★★★★ (1)

EUR 152,45



**The Filter Bubble: What
The Internet Is Hiding Fr...**
von Eli Pariser

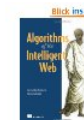
★★★★☆ (1)

EUR 12,95



**Music Recommendation
and Discovery: The Long
Tail...** von Oscar Celma

EUR 36,99



**Algorithms of the Intelligent
Web** von Haralampos Marmanis

EUR 29,95

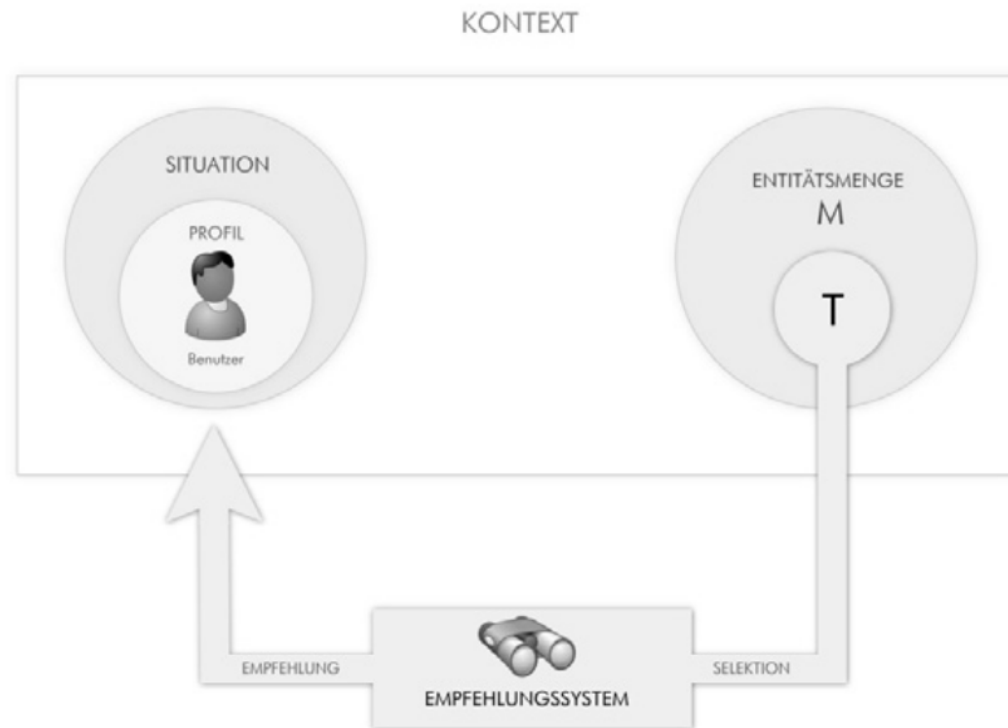


**Web Analytics: Metriken
auswerten, Besucherverh...**
von Marco Hassler

★★★★★ (6)

EUR 29,95

Context Components

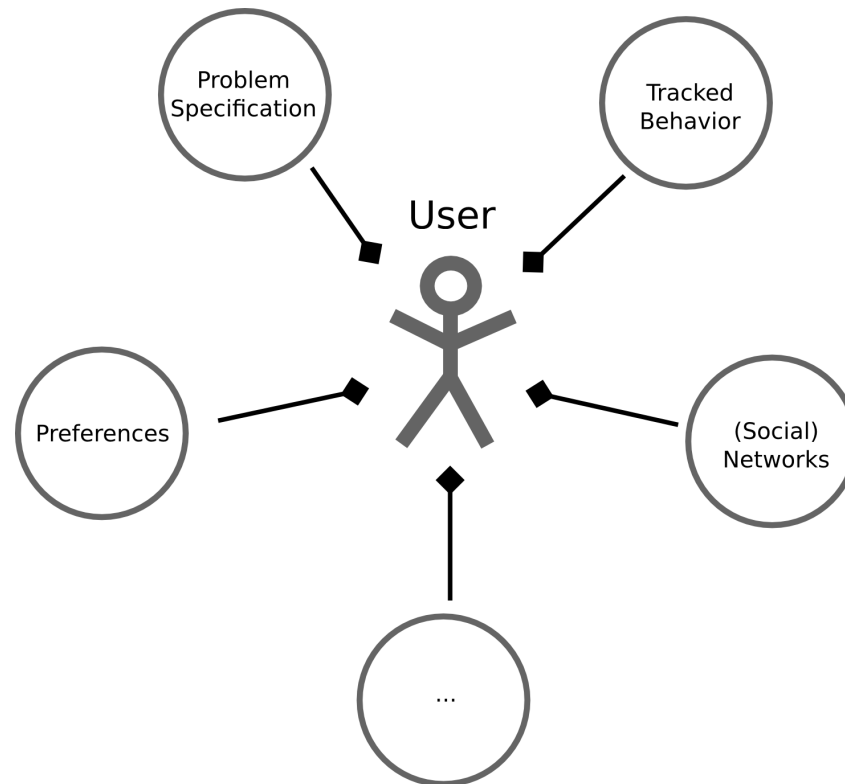


from: Klahold, A., Empfehlungssysteme - Recommender Systems - Grundlagen, Konzepte und Lösungen, Vieweg+Teubner, 2009

Types of Recommender Systems

- Collaborative Filtering
 - compare user profiles
 - similar profiles result in recommendations
- Content-Based Filtering
 - compare entities
 - similar attributes result in recommendations
- Hybrid Systems
 - most commonly used
 - combine both approaches

User-Specific Data



Ideas for Combination

- Parallel approach
 - Run recommender and reasoner independently
 - Intersect the results
- Sequential approach
 - First: Filter for possible services (discovery)
 - Then: Start recommendation with reduced entity set
 - Or vice versa: first recommend, then filter
- Combined approach
 - e.g.: extend ontology according to recommendations (dynamically or static)

Example Domain

EMBOSS



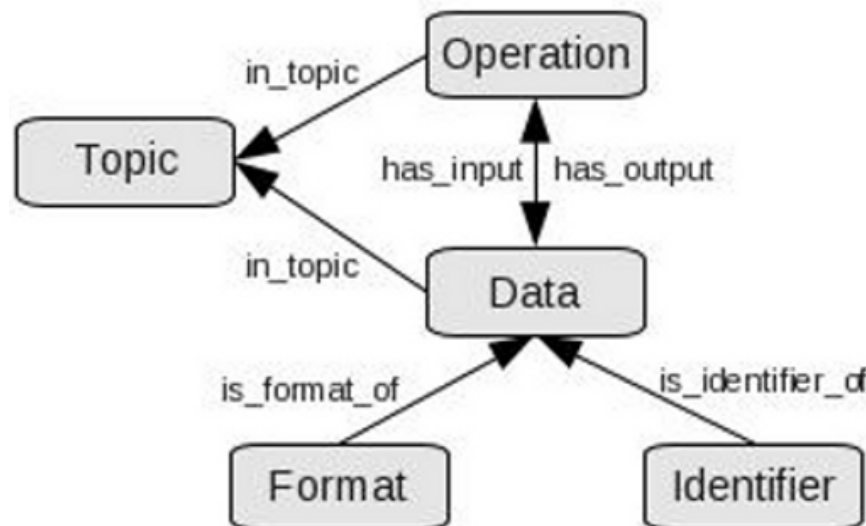
- European Molecular Biology Open Software Suite
- <http://emboss.sourceforge.net/>
- Open source software package
- Developed for the needs of the molecular biology user community

- Comprises around **430 tools** for different bioinformatics tasks:
 - Sequence alignment
 - Rapid database searching with sequence patterns
 - Presentation tools for publication
 - ...

- Latest release: 6.4.0 (15th July 2011)

EDAM

- EMBRACE Data and Methods Ontology
- <http://edamontology.sourceforge.net/>
- Ontology of general bioinformatics concepts
- Serves as controlled vocabulary for the **annotation** of, e.g., services.



from: <http://edamontology.sourceforge.net/>

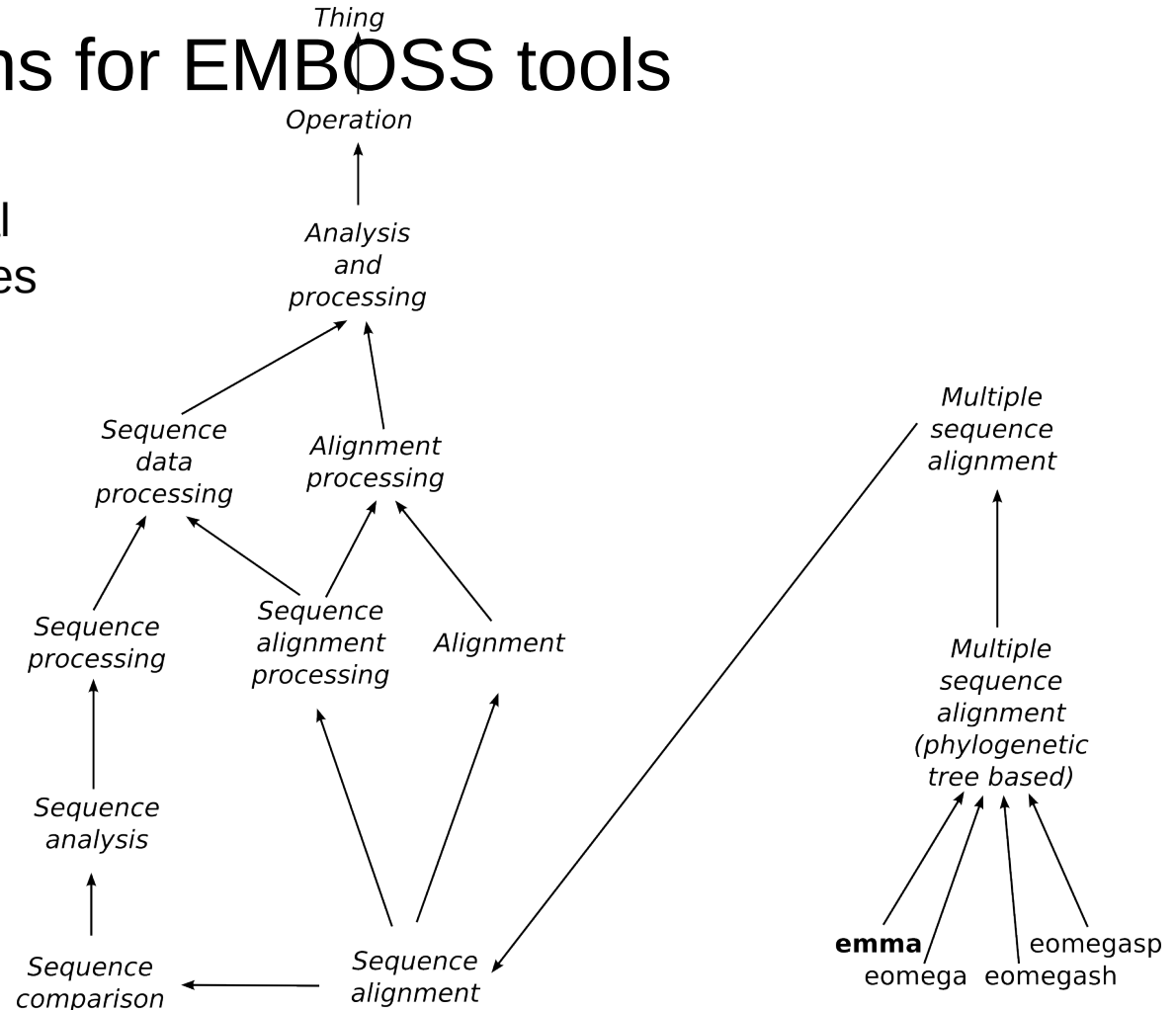
```

    section: output [
      information: "Output section"
      type: "page"
    ]
    section: input [
      information: "Input section"
      type: "page"
    ]
  application: emma [
    documentation: "Multiple sequence alignment (Clustalw wrapper)"
    groups: "Alignment:Multiple"
    gui: "yes"
    batch: "yes"
    cpu: "medium"
    external: "clustalw" ]
    relations: "EDAM:0000182 topic Sequence alignment"
    relations: "EDAM:0000499 operation Multiple sequence alignment"
    (phylogenetic tree-based)
    additional: "Y"
    default: "N"
    information: "Only produce dendrogram file"
    relations: "EDAM:0002527 data Parameter or primitive"
    knowntype: "dendrogram"
    relations: "EDAM:0000872 data Phylogenetic tree"
  ]
endsection: input ]
endsection: output

```

EDAM annotations for EMBOSS tools

- Enable, e.g., ontological classifications of services and inputs/outputs



BioCatalogue



- A curated catalogue of Life Science Web Services
- <http://www.biocatalogue.org/>
- Currently 2269 services (also EMBOSS), 159 providers, 634 members

- Provides:
 - Services
 - Technical service (interface) information
 - Tags, categorization
 - User favouring
 - Links to similar services
 - Information on costs, licences, etc.
 - ...

myExperiment



- <http://www.myexperiment.org/>
- part of the myGrid consortium (the UK's e-Science initiative)
- (social) platform for sharing workflows etc.:
“myExperiment makes it easy to **find**, **use** and **share scientific workflows** and other **Research Objects**, and to build **communities**.”
- Originally designed for sharing Taverna workflows, meanwhile open for all kinds of workflows.
- Currently more than 2000 workflows available (also several using EMOSS).

[Home](#)
[Users](#)
[Groups](#)
[Workflows](#)
[Files](#)
[Packs](#)
[Services](#)
[Topics](#)
[Home](#) » [Workflows](#)
 [BOOKMARK](#)   

Workflows

Search filter terms

Filter by type

- ☐ Taverna 2 719
- ☐ Taverna 1 562
- ☐ RapidMiner 168
- ☐ Kepler 43
- ☐ Bioclipse Scri... 33
- ☐ GWorkflowDL 24
- ☐ LONI Pipeline 22
- ☐ BioExtract Ser... 16
- ☐ Trident (Packa... 10
- ☐ LabTrove Tem... 9

Filter by tag

- ☐ example 213
- ☐ mygrid 103
- ☐ localworker 100
- ☐ bioinformatics 98
- ☐ benchmarks 77

« previous **1** **2** **3** ... **167** next »

Sort by: Rank

Showing 1670 results. Use the filters on the left and the search box below to refine the results.

Search

Taverna 2



Pathways and Gene annotations for QTL region

(v7)

View

Download (v7)

Original Uploader

Created: 19/11/09 @ 18:18:52 | Last updated: 02/09/11 @ 11:44:57

Credits:  Paul Fisher

 Paul Fisher

License: Creative Commons Attribution-Share Alike 3.0 Unported License



This workflow searches for genes which reside in a QTL (Quantitative Trait Loci) region in the mouse, *Mus musculus*. The workflow requires an input of: a chromosome name or number; a QTL start base pair position; QTL end base pair position. Data is then extracted from BioMart to annotate each of the genes found in this region. The Entrez and UniProt identifiers are then sent to KEGG to obtain KEGG gene identifiers. The KEGG gene identifiers are then used to search for pathways in the KEGG path...

New/Upload

Workflow



GO

Log in / Register

Username or Email:

Password:

Remember me: ☐

OR

Use OpenID:

(eg: name.myopenid.com)

Log in

Need an account?
[Click here to register](#)
[Forgot Password?](#)

Popular Tags

25 tags

[\[All Tags\]](#)

Organisatorisches

Zeitplan (1)

- Vorbesprechung
- Seminarphase
 - 2 Tage (kurz vor oder zu Vorlesungsbeginn)
 - Präsentationen zu den Kernthemen
 - Kennenlernen
- Erstes Semester
 - Einarbeitung, Analyse, Entwurf, Implementierungsbeginn
 - Dokumentation, Zwischenbericht
- Zweites Semester
 - Implementierungsabschluss, Integration, Testen, Evaluation
 - Dokumentation, Endbericht
 - Abschlusspräsentation

Zeitplan (2)

1. Semester:

Einarbeitung in E4B Recommender Sys., Discovery-Verfahren und EMBOSS/EDAM	Anforderungen und Entwurf zur Realisierung ausarbeiten	Implementierung der zuvor entwickelten Discovery-Verfahren
---	--	--

Dokumentation der Analysen und entwickelten Konzepte / Zwischenbericht

2. Semester:

Integration der im ersten Semester implementierten Verfahren in Eclipse4Bio	Ausführliches Testen des Produkts sowie Evaluierung anhand der EMBOSS-Services	Projektabschluss, Reflexion und Präsentation
---	--	--

Technische Dokumentation und User Manual / Endbericht

Was ihr braucht und was ihr lernt

- Was ihr braucht:
 - Programmierkenntnisse
 - Interesse an logischen Systemen
 - **Motivation, sich in die Themen einzuarbeiten**
- Was ihr lernt:
 - Umgang mit verschiedenen (neuen) Technologien
 - Wissenschaftliches Arbeiten
 - Teamarbeit
 - Selbstorganisation und Projektmanagement
 - ...

Was ihr (von uns) bekommt

- PG-Schein :-)
- Erfahrungen im Umgang mit
 - Services und Workflowsystemen
 - Semantic Domain Modeling, Synthesis & Planning
 - Eclipse-Technologien
 - Bioinformatik-Workflows
 - ...
- Diplom-/Masterarbeitsthemen

Offene Fragen?

- ?