

SKIL

Studentenkonferenz
Informatik Leipzig
25. September 2012



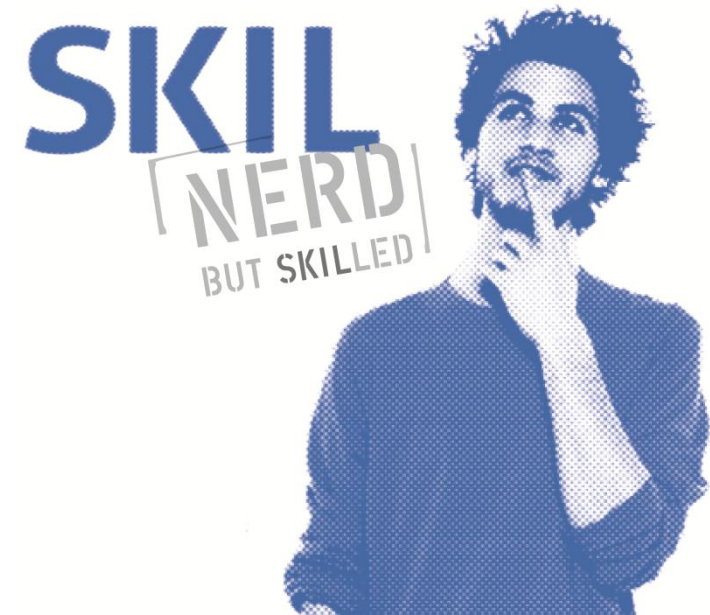
Visualisierung der Evolution von Ontologien in den Lebenswissenschaften

Victor Christen
Universität Leipzig
mam08bfa@studserv.uni-leipzig.de

Agenda

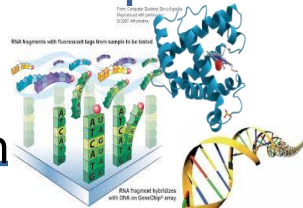
- Einleitung und Motivation
- Grundlagen
 - Ontologien in den Lebenswissenschaften
 - Evolution von Ontologien
 - Regionenalgorithmus
 - Problembeschreibung
- Hauptteil
 - REx - Region Evolution Explorer
 - Evaluation
- Zusammenfassung

EINLEITUNG UND MOTIVATION



Informationsquellen

- Experimentdaten
- Biologische Objekte (Gene, Proteine, etc.)
 - Anatomie
 - Krankheiten
- Klinische Daten



Annotation mittels Ontologie



Aufgaben

- Ähnlichkeitssuche
 - Analyse von Gensequenzen und Proteinstrukturen
- Homologien



Neue Erkenntnisse Anpassung

- Veränderungen?
- Ausmaß der Veränderung?
- Repräsentation?

Analyse der Evolution - Regionenalgorithmus

- Versionierung von Ontologien
 1. Diff- Berechnung zwischen aufeinanderfolgender Versionen
 2. Berechnung signifikanter numerischer Werte mittels des Regionenalgorithmus
 - Numerisches Kostenmodell für die Änderungsoperationen
 - Region eines Konzepts: Subgraph bestehend aus dem Konzept und den Nachfolgeknoten der is_a Relation
 - Lokale Kosten vs. Aggregierte Kosten vs. Avg_costs

		<i>accession</i>	<i>name</i>	<i>abs_size</i>	<i>rel_size</i>	<i>avg_costs</i>
GO	<i>unstable</i>	GO:0005102	receptor binding	408	4.31%	0.95
		GO:0009653	anatomical structure morphogenesis	583	3.22%	1.22
		GO:0048856	anatomical structure development	566	3.13%	0.91
		GO:0033643	host cell part	77	2.81%	1.90
		GO:0003676	nucleic acid binding	241	2.55%	0.86
		GO:0048646	anatomical structure formation involved in morphogenesis	253	1.40%	0.92
	<i>stable</i>	GO:0031300	intrinsic to organelle membrane	36	1.32%	0.000
		GO:0030054	cell junction	31	1.13%	0.000
		GO:0050865	regulation of cell activation	184	1.02%	0.012
		GO:0075136	response to host	181	1.00%	0.019
		GO:0000151	ubiquitin ligase complex	25	0.91%	0.000
	GO:0016860	intramolecular oxidoreductase activity	71	0.75%	0.000	
NCIT	<i>unstable</i>	C28428	Retired Concept	3,264	4.21%	3.49
		C53791	Adverse Event Associated with Infection	1,186	1.53%	2.36
		C45678	Industrial Aid	889	1.15%	1.40
		C74944	Clinical Pathology Procedure	747	0.96%	0.84
		C66892	Natural Product	708	0.91%	1.35
		C53543	Rare Non-Neoplastic Disorder	504	0.65%	1.22
	<i>stable</i>	C64389	Genomic Feature Physical Location	1,026	1.32%	0.000
		C23988	Mouse Neoplasms	886	1.14%	0.000
		C48232	Cancer TNM Finding	742	0.96%	0.000
		C53798	Adverse Event Associated with Surgery & Intra-Operative Injury	707	0.91%	0.000
		C43877	American Indian	555	0.72%	0.000
	C53832	Infection Adverse Event with Unknown Absolute Neutrophil Count	386	0.50%	0.000	

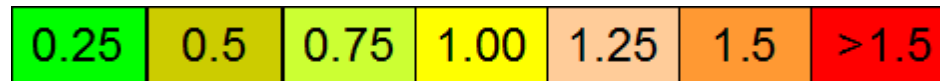
REx- Region Evolution Explorer

- Was?
 - Webapplikation für die Visualisierung von Regionen
 - Werkzeug für die visuelle Datenexploration
- Warum?
 - Anpassung der Annotationen ausschließlich für Regionen mit hoher Veränderung
 - Identifikation interessanter Gebiete → Forschung

REx- Graphrepräsentation

- Ontologie als Graph
 - Knoten \equiv Konzept
 - Kante \equiv Relation

- Darstellung der Intensität der Veränderung einer Region durch Farbskala



Geringe avg_costs

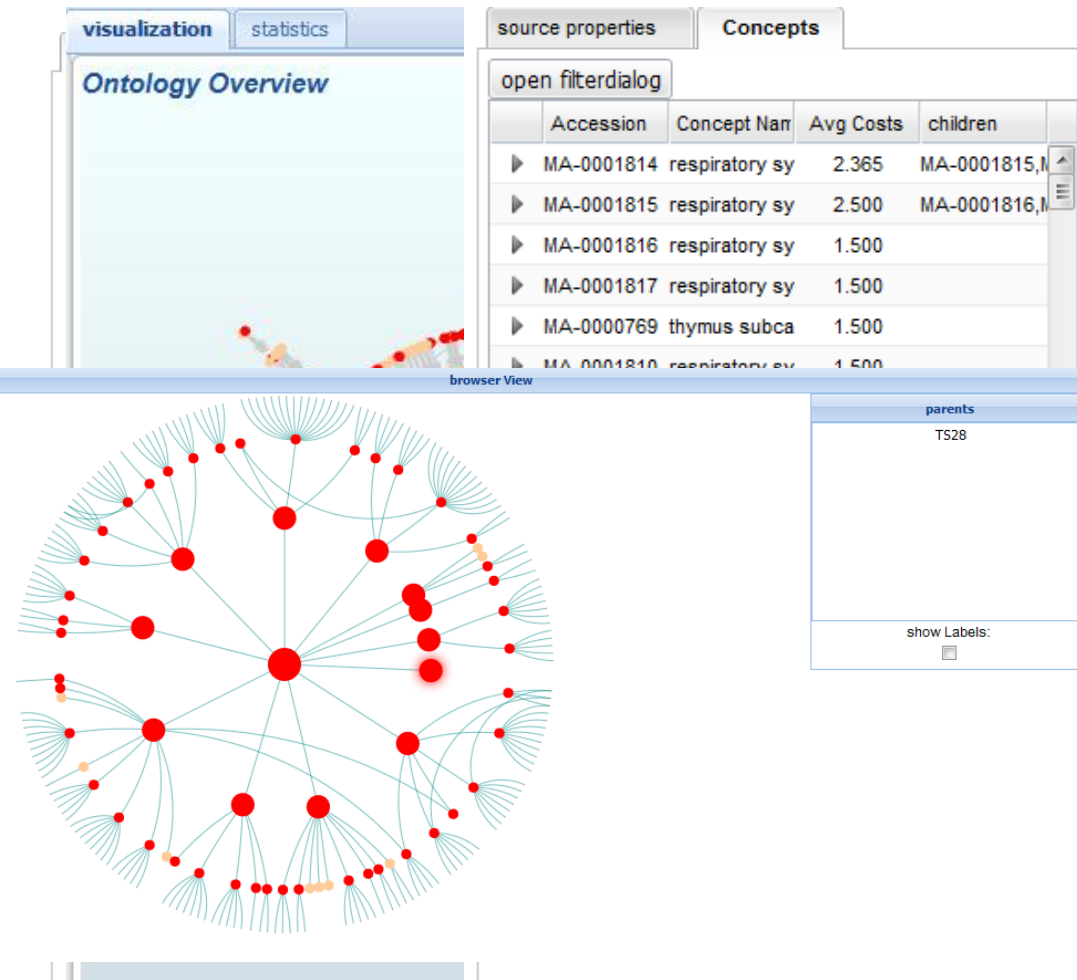
hohe avg_costs

Ablauf des Graph Workflows

1. Spezifikation der Eingabeparameter

2. Präsentation

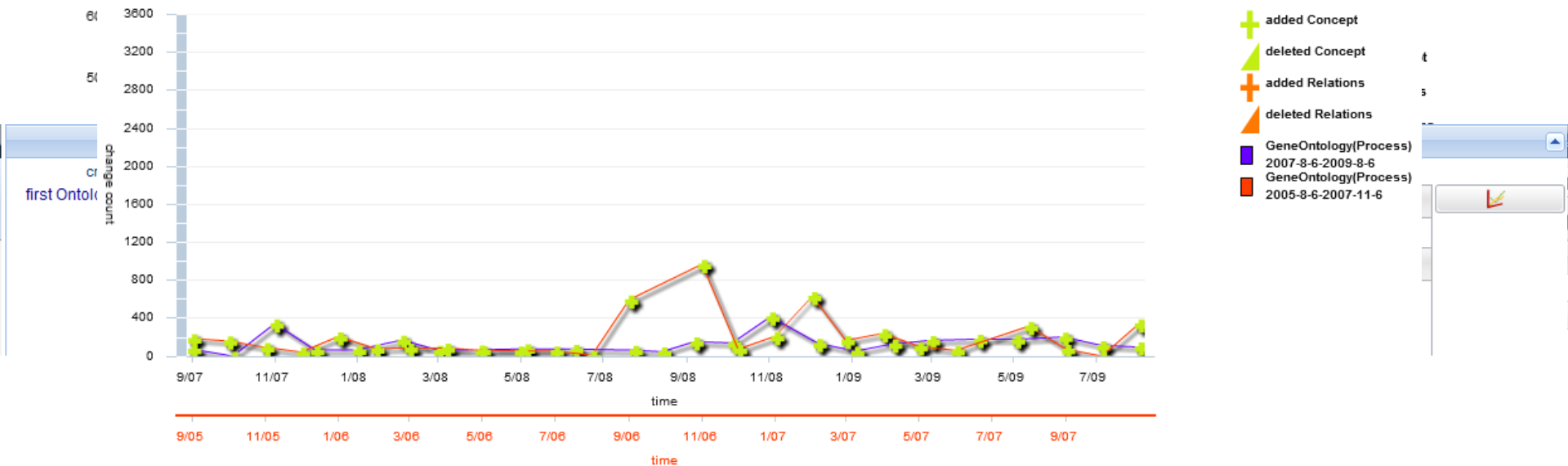
- Übersichtsgraph
 - Erkennen von interessanten Teilgebieten
- Tabelle
- Detail Browser Graph



Repräsentation als Statistiken

Quantitative Statistiken

- Basierend auf Insertion- und Löschoperation
- Auswahl der Vergleichskriterien und Spezifikation der Eingabe
 - zwei Ontologien
 - Zwei Zeiträume

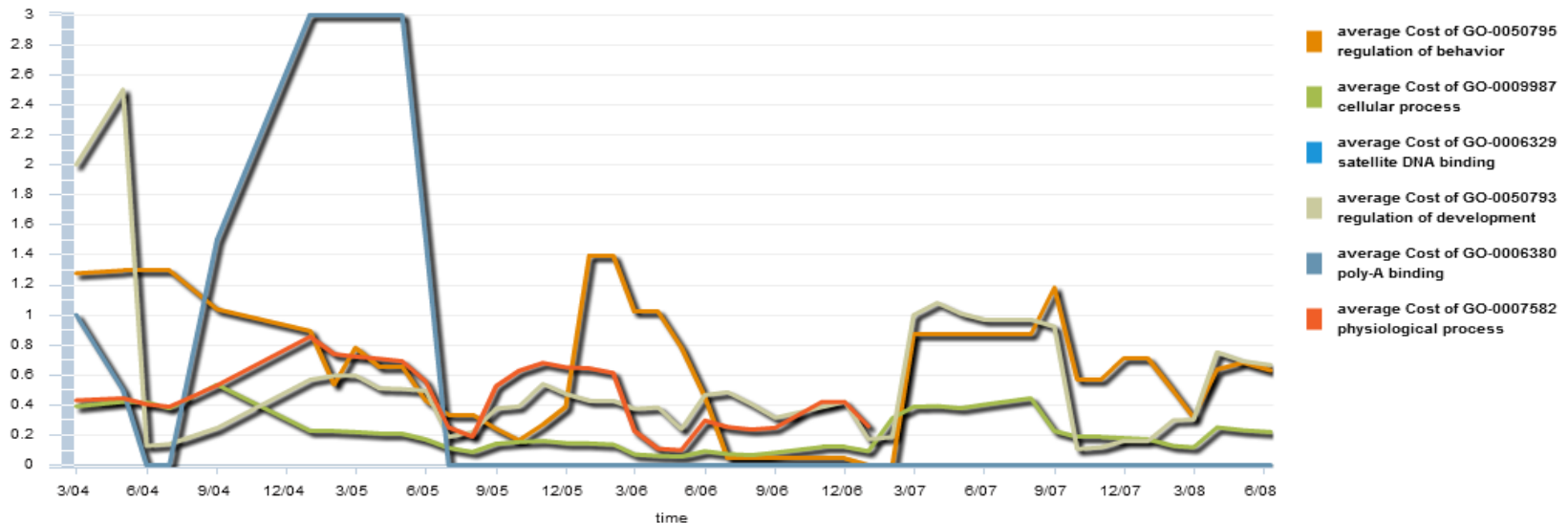


Repräsentation als Statistiken

Sliding- Window Statistik

Sliding – Window Verfahren

- Analyse der avg_costs für Konzepte einer Ontologie über einen Zeitraum
- Eingabe: Zeitraum t , Fenstergröße w , Schrittweite Δt
- Ausgabe: Avg_Costs für jedes Konzept für jeden Zeitpunkt am Ende des Fensters



Evaluation



Zusammenfassung

- Fazit
 - Analyse der Evolution von Ontologien notwendig
 - Vereinfachung von:
 - Validierungsprozess bestehender Annotationen
 - Ontologieentwicklung
 - Ontology Matching Tasks

- Future Work
 - Analyse jeder Ontologie
 - Explizite Darstellung der Änderungsoperationen bzgl. einer Region

