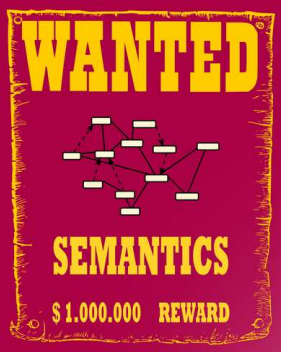
Hybrid Approach to Automated Domain Model Creation for Adaptive Social Learning System

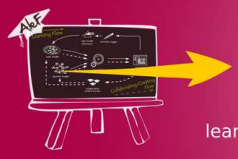
Marián Šimko supervised by **Professor Mária Bieliková** simko@fiit.stuba.sk

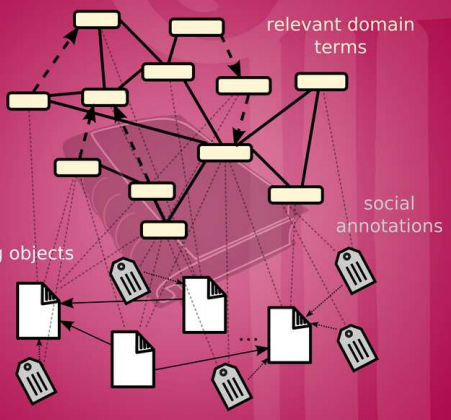


Bottlenecks

Semantics creation
- demanding
- time-consuming
- tedious
- complex
- almost imposible

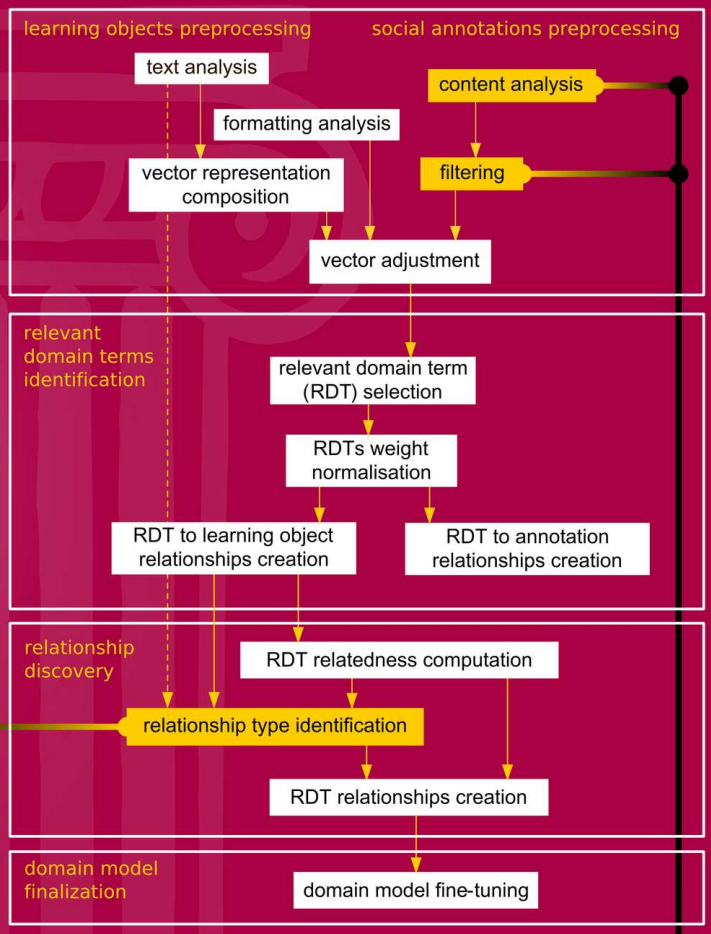
Semantics update
- student generated content!
  - comments
  - feedback
  - discussions





Adaptive Web-Based Learning 2.0

simplified domain modelling
with respect to:
- automate certain domain model parts creation
- collaborative social aspects
  (e.g. the need to modify or alter domain model by students themselves)



Hyponymy relationship discovery

term subsumption (A)

"*čistý výraz*" -> "*výraz*"
"*rekurzia na chvoste*" -> "*rekurzia*"

RDT phrases (B)

"Okrem *funkcie CONS* sa v lispe často používajú ..."

lexico-syntactical patterns (C)

1. find lexical candidates

   "*Pod pojmom forma rozumieme taký výraz, ktorý...*"
   `<-----LHS-----> delimiter <-----RHS-----.>`

2. `foreach (delimiter, rule):`
   a. filter LHS and RHS
   b. match RDTs for LHS and RHS - lemma match ratio
   c. generate relationship candidates

I. lexico-syntactical analysis

output: `(RDT1, RDT2, confidence)`

co-occurrence analysis (D)

- set theoretical basis
- "Concept A is a superconcept of concept B if the set of entities classified under B is a subset of the entities under A"

II. course metadata association utilization

course structure traversal (E)

- confidence boost based on "related-to" relationship

best candidates selection

Social annotation preprocessing

	comment	tag	ext. resource	highlighting
clear text extraction			X	
text segmentation	X		X	X
lemmatization	X	X	X	X
weights computation	X		X	X

Experiments

HYPONYMY RELATIONSHIP DISCOVERY

- lisp course
- 79 learning objects (explanations)
- comparison with gold standard:
  - 162 relevant domain terms
  - 126 is-a relationships
  - 76 related-to relationships
- result: R = 70.6 %

	#	$\#_{trans.}$
TOTAL	89	30
A	28	0
B	39	19
C	30	13

SOCIAL ANNOTATIONS STATS

	PrPr (2010)	*PSI (2011)*
comments:	233	238
tags:	2272	798
ext. resources:	1898	178
highlights:	-	> 10 000 (!)

FUTURE WORK

- recommendation simulation/prediction
- long-term live experiment