14th AIChE
Annual Meeting, Atlanta, GA

UNIVERSITY OF
SURREY

Towards a Re-Usable Ontology for Waste Processing

Franjo Cecelja, Nikolaos Trokanas, Tara Raafat

www.surrey.ac.uk

14th AIChE
Annual Meeting, Atlanta, GA

UNIVERSITY OF
SURREY

Introduction

Reusing Ontologies: problem decomposition and case demonstration

Ontology: definition and reuse statement – design requirements

Reusability: framework definition

- Degree of confidence
- Evaluation algorithm

Validation: industrial symbiosis case

www.surrey.ac.uk

14th AIChE Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

Reusing Ontologies

Benefit Decomposition

- Reducing Cost
- Minimising Effort
- Sharing Knowledge
- Fulfilling Purpose

Not often demonstrated in practice, especially in the domain of **chemical/processing** engineering!

Proposed **framework** addresses practicality of **reusing ontologies** for non-programmers.

www.surrey.ac.uk

14th AIChE Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

Case: Industrial Symbiosis

Definition: industrial **networks** set to trade material, energy and water produced from **residues/waste** to gain **economic, environmental & social** benefits

Ad-hoc networks:

- Confined **geographical** environment (focus on local communities)
- Short** life-time, **unpredictable** quantities, **non-standardised** composition
- Difficult to predict **benefits**
- Strong **environmental** expectation
- Large number** of options which have to satisfy technological requirements
- Awareness** of benefits and partner existence

www.surrey.ac.uk

14th AIChE Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

IS Decision Support: Ontology (2)

$$O = \langle H^I, H_C, R_i^C, R^C, (f_R), S_i^I \rangle$$

$$f_{R(D)} = \text{dom} R_i^C |_{S_i^R} \rightarrow f_D: S_i^I \xrightarrow{f_D} S_i^D$$

$$f_{R(R)} = \text{rang} R_i^C |_{S_i^D} \rightarrow f_R: S_j^I \xrightarrow{f_R} S_j^R$$

Hence binary relationship:

- S_i^D and S_j^R **universal** and **existential** quantifiers over properties R_i^C of S_i^I ,
- S_i^D and $n, n \in \mathbb{N}$, **cardinality** quantifiers over properties P_i^{np} of S_i^I ,
- S_i^D and $v, v \in S_i \vee N$, **equality** quantifiers over properties P_i^{np} of S_i^I .

Raafat, T., F. Cecelja, N. Trokanas and B. Xrisha (2013). *An Ontological Approach Towards Enabling Processing Technologies Participation in Industrial Symbiosis*, Computers & Chemical Engineering, 59, pp 33-46.

ac.uk

14th AIChE Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

Prerequisites for Reusable Ontologies

Modularisation

Domain
Level

Naming Conventions

Coherent Modelling

External Resources
Data types

Minimum Restrictions

Detailed Annotations

Syntactic Heterogeneity

Loosely defined guidelines referring to a **good practice**, not **strict rules**.

www.surrey.ac.uk

* Trokanas, N., T. Raafat, F. Cecelja, A. Kokosis, and A. Yang. **Semantic Formalism for Waste and Processing Technology Classifications Using Ontology Models**. *Computer-Aided Chem. Eng.* 2012, 30: p. 167 - 171. www.surrey.ac.uk

www.surrey.ac.uk

Evaluation Algorithm

Vector Modelling

Cosine Similarity $CoSim(a, b) = \frac{\sum_{i=1}^n a_i * b_i}{\sqrt{\sum_{i=1}^n (a_i)^2} * \sqrt{\sum_{i=1}^n (b_i)^2}}$

Euclidean Distance $Eucl(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$

First Aggregation Step $S_{AB} = \frac{\sum_i^O Sim_i}{\sum_i^O w_i}$

Second Aggregation Step $C_{ij} = 0.7 * S_{ij} + 0.3 * doc_{ij}$

www.surrey.ac.uk

Vector Space Modelling

Terminology: $T_A = (t_1, t_2, \dots, t_i, \dots, t_n)$

Language: $L_A = (l_1, l_2, \dots, l_i, \dots, l_n)$

Data type: $E_A = (e_1, e_2, \dots, e_i, \dots, e_n)$

External resources: $R_A = (r_1, r_2, \dots, r_i, \dots, r_n)$

Lexical similarity:

$$LexSim(str_1, str_2) = \left(CC + \frac{2 * LCS(str_1, str_2)}{l_1 + l_2} \right)^{MSP}$$

CC – No of common characters, LCS – the least common substring, l_i – the length of the term, MSP – maximum of the Levenshtein distances of the suffixes and prefixes of the two terms

www.surrey.ac.uk

14th AIChE
Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

Reusing Examples

Required Knowledge for IS ontology:

Waste	no available ontologies
Materials	available ontologies very detailed
Products	available ontologies very detailed
Attributes	location, unit of measurement

```

graph TD
    A(( )) --> B(UO1 ontology)
    A --> C(MUO2 ontology)
    A --> D(QUTD3 ontology)
  
```

¹<http://unit-ontology.googlecode.com/svn/trunk/uo.owl> (units – physical)
²<http://idi.fundacionctic.org/muo/muo-vocab.html> (units of measurement)
³<http://qudt.org/> (Quantities, Units, Dimensions and Data Types) www.surrey.ac.uk

14th AIChE
Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

Reusing Example (continued)

Target Ontology

eSymbiosis ontology: ~2500 concepts

Candidate Ontologies

UO ontology: 2755 concepts

QUTD ontology: 216 concepts

MUO ontology: 10 concepts

www.surrey.ac.uk

Terminology

$$T_{eSymbiosis} = (t_1, t_2, \dots, t_i, \dots, t_n)$$

$$T_{qutd} = (t_1, t_2, \dots, t_i, \dots, t_n)$$

$$T_{muo} = (t_1, t_2, \dots, t_i, \dots, t_n)$$

$$T_{uo} = (t_1, t_2, \dots, t_i, \dots, t_n)$$

}

Terminology Similarity	eSymbiosis
QUTD	0.45
MUO	0.66
UO	0.25

www.surrey.ac.uk

External Resources

$$R_{eSymbiosis} = (r_1, r_2, \dots, r_i, \dots, r_n)$$

$$R_{qutd} = (r_1, r_2, \dots, r_i, \dots, r_n)$$

$$R_{muo} = (r_1, r_2, \dots, r_i, \dots, r_n)$$

$$R_{uo} = (r_1, r_2, \dots, r_i, \dots, r_n)$$

}

External Resources Similarity	eSymbiosis
QUTD	0.66
MUO	0.66
UO	0

www.surrey.ac.uk

Data types

$$E_{eSymbiosis} = (e_1, e_2, \dots, e_i, \dots, e_n)$$

$$E_{qutd} = (e_1, e_2, \dots, e_i, \dots, e_n)$$

$$E_{muo} = (e_1, e_2, \dots, e_i, \dots, e_n)$$

$$E_{uo} = (e_1, e_2, \dots, e_i, \dots, e_n)$$

}

Data types	eSymbiosis
Similarity	
QUTD	0.40
MUO	0.58
UO	0.52

www.surrey.ac.uk

Degree of Confidence

- Target Ontology
 - eSymbiosis ontology: ~2000 concepts
- Candidate Ontologies
 - UO ontology: 2755 concepts
 - QUTD ontology: 216 concepts
 - MUO ontology: 10 concepts

→

Degree of Confidence	eSymbiosis
QUTD	0.92
MUO	1.00
UO	0.00

www.surrey.ac.uk

14th AIChE Annual Meeting, Atlanta, GA

UNIVERSITY OF SURREY

Aggregation

	eSymbiosis
QUTD	1.00
MUO	1.00
UO	1.00
eSymbiosis	
QUTD	0.40
MUO	0.58
UO	0.52
eSymbiosis	
QUTD	0.45
MUO	0.66
UO	0.25
eSymbiosis	
QUTD	0.66
MUO	0.66
UO	0

Metadata Similarity	eSymbiosis
QUTD	0.62
MUO	0.72
UO	0.44

+

Degree of Confidence	eSymbiosis
QUTD	0.92
MUO	1.00
UO	0.00

=

Compatibility Score	eSymbiosis
QUTD	0.77
MUO	0.86
UO	0.22

www.surrey.ac.uk

Conclusions

Reusing Ontologies: stated the problem and defined a IS case for demonstrating ontology reusability

Ontology and Reusability:

- Defined the ontology design requirements
- Defined the reusability framework: confidence and ontology compatibility based on language, terminology, data type external resources and lexical similarity
- A full algorithm based on cosine and Euclidian similarity

Validation: industrial symbiosis case

www.surrey.ac.uk

Thank you

Towards a Re-Usable Ontology for Waste Processing

Franjo Cecelja, [Nikolaos Trokanas](#), Tara Raafat

www.surrey.ac.uk