Automatic Metadata Enrichment and Linking for Event-driven Access to Distributed Collections

agora
eventing history

The Agora project aims to develop a social platform in which objects from cultural heritage organisations are presented in an explicit historical context. The historical context is provided by events, that are used as a flexible model to identify semantically relevant relationships between objects in diverse cultural heritage collections such as those of the Netherlands Institute for Sound and Vision and the Rijksmuseum Amsterdam. We present our first steps in creating this platform by showing how objects are linked to events and the integration of the event model with the collection presentation.

IDENTIFICATION OF:
- Events
- Actors
- Dates
- Locations

IDENTIFICATION FROM:
- Collection objects metadata
- External resources such as exhibition catalogues and Wikipedia

IDENTIFICATION BY:
- Named Entity Recognition adapted for the Dutch history domain
- Knowledge from Event Thesaurus

Opening of the New Guinea conference
Week number 50-49

Newsreels in which Dutch subjects of a certain week are presented.

type: video
length: 01:13
publication date: 4 December 1950
user: Nederlands Instituut voor Beeld en Geluid
source: WEEKNUMMER504-HRE00015911
tags: politiek, jaarthema 2010, Indonesië, geschiedenis

At the Round Table Conference in 1949 an agreement was reached on the sovereignty of Indonesia. The Dutch part of New Guinea was expressly left out of the agreement, in spite of strong pleas for the territory by the Indonesian Republic. It was agreed that the matter would be settled before 27 December 1950 by negotiations. On 4 December the Dutch and Indonesian representatives gather for their first official meeting in the Trèveszaal in The Hague. Present are a.o. the minister of Oversea Territories, mr. J.H. van Maarseveen, the minister of Foreign Affairs, mr. D.U. Stikker, the minister of Social Affairs mr. A.M. Joekes, prime minister dr. W. Drees, the Indonesian High Commissioner Mohammed Roem and other Indonesians.
SHOTS: - The gentlemen enter, shake hands and have an informal talk; - all are seated at the conference table and listen to the speeches of Van Maarseveen and Roem; - on a topographical map the position of New Guinea compared to Indonesia and Australia is shown; 00:06 Van Maarseveen (left); Roem (right); 00:15 from left to right behind the table Drees, Stikker, Joekes; 00:51 Drees (left); Stikker (right).

Associated Events:
- New Guinea Conference
- Round Table Conference

Associated Actors:
- J. H. van Maarseveen
- D. U. Stikker
- A. M. Joekes
- W. Drees
- Mohammed Roem

Associated Location:
- Trèveszaal, The Hague

Associated Date:
- 4 December 1950

EXPLORATION OF EVENTS:
- Through objects associated to events
- Through related events, e.g., super- or sub-event, and various temporal and semantic relations

ASSOCIATED METADATA:
- Explicit presentation of associated properties, e.g., events, actors, dates and locations
- Users browse through these hyperlinks to find related objects

Your Narrative:

NARRATIVE OF USER ACTIVITIES:
- objects viewed by user
- links followed by user

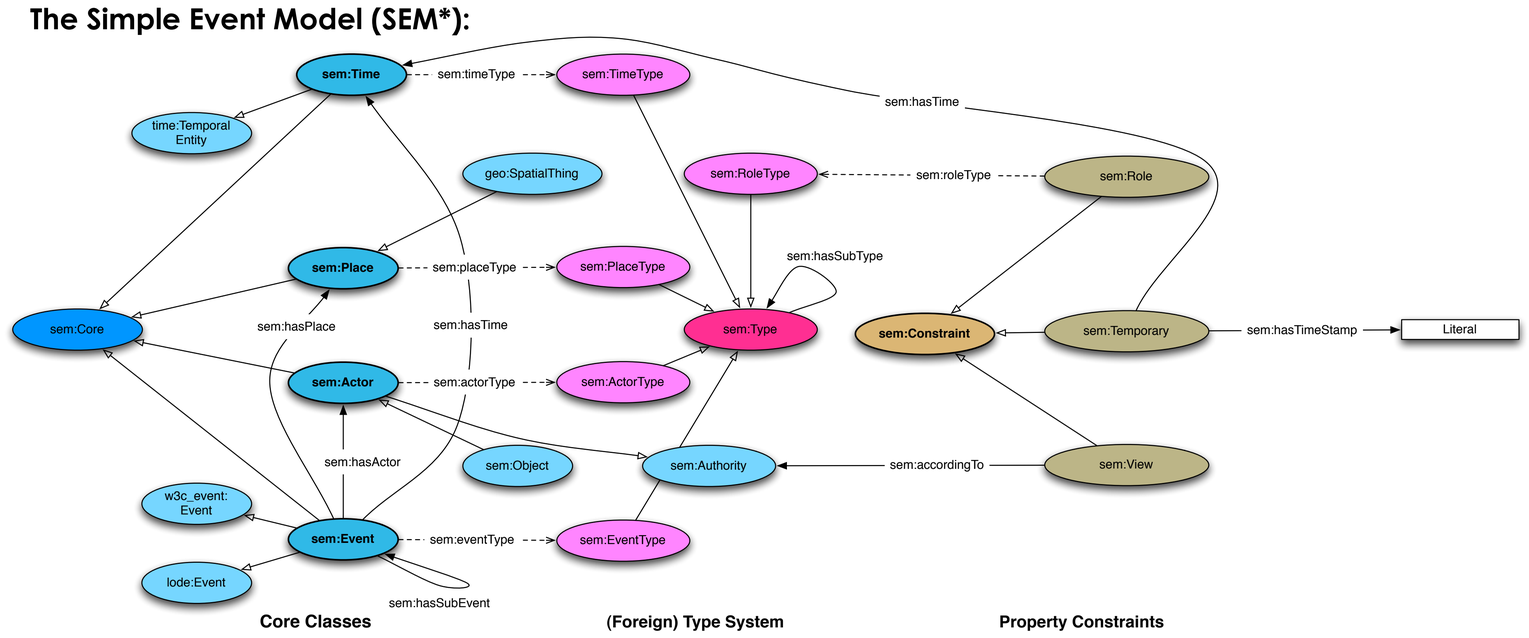
Research Questions:

1. How to model historical events and how to interpret events time-lines & narratives using semantics to derive/explain views, biases, contradictions, opinions and emotional reactions?

2. How to use domain knowledge (historical events), to improve search in semantically interoperable collections?

3. How can semantic web techniques allow for meaningful access to distributed collections?

The Simple Event Model (SEM*):



Chiel van den Akker, Lora Aroyo,Agata Cybulska, Marieke van Erp, Peter Gorgels, Laura Hollink, Cathy Jager, Susan Legêne, Lourens van der Meij, Johan Oomen, Jacco van Ossenbruggen, Guus Schreiber, Roxane Segers, Piek Vossen and Bob Wielinga

http://agora.cs.vu.nl/
http://www2.let.vu.nl/oz/cltl/semhis/

*Van Hage, W. R., Malaisé, V., De Vries, G., Schreiber, A. Th., Van Someren, M. (2010) Abstracting and Reasoning over Ship Trajectories and Web Data and the Simple Event Model (SEM) *Multimedia Tools and Applications*

Continuous Access To Cultural Heritage
a computer science research programme
NWO Netherlands Organisation for Scientific Research
vrije Universiteit
CAMERA
BEELD EN GELUID
RIJKS MUSEUM amsterdam

ArchiMind – A Semantic Wiki for Software Architecture Knowledge Retrieval and –Management

Klaas Andries de Graaf, Peng Liang, Hans van Vliet - Software Engineering group
Antony Tang – CS3 - Swinburne University, Melbourne, Australia

Software **documentation is essential for** the **success of** large and complex **software projects**.

The use of file-based software documents suffers from various **issues**, including:

- synonyms and homonyms
- spelling errors
- abbreviations
- ambiguity
- traceability

Document 1:
"Ordering config", "OCC", "Orders"

Document 2
"Order component", "order config component"

Document 3:
"Ordr_cnfg_cmpnt", "orderC-Int"

-Example of issues-

Proposed approach *(figure 1)*

- **Ontology** domain model for software knowledge (*figure 2*)
- **Semantic Wiki** (ArchiMind) tool for knowledge retrieval and management (*figure 3*)
  - Input of software document fragments as wikipages.
  - Semantic annotation of text on wikipages.

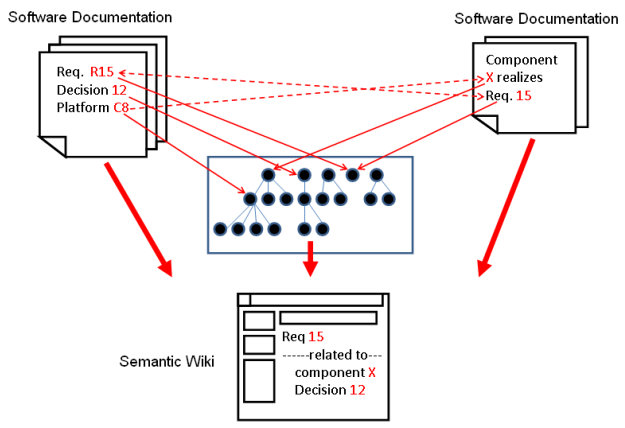


Figure 1- Overview of approach

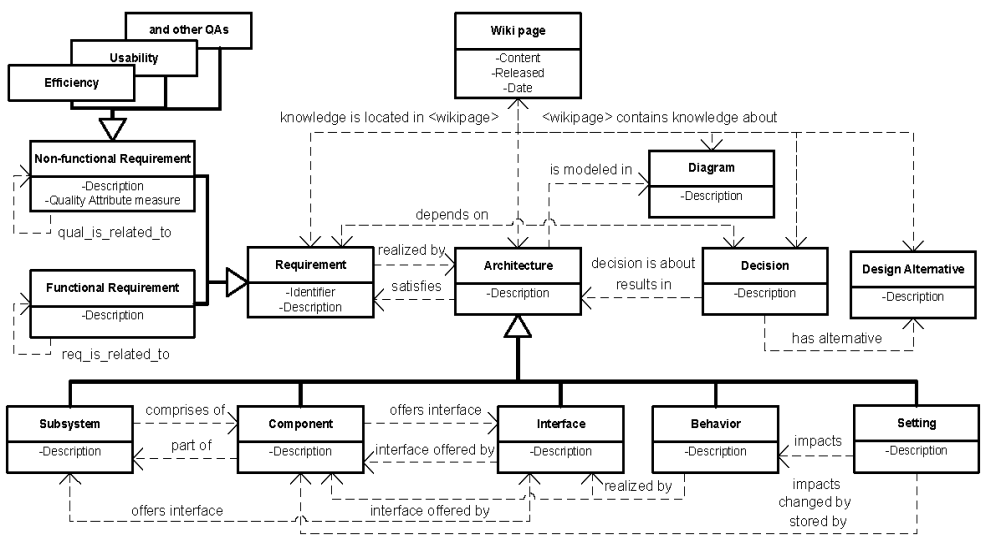


Figure 2 –Ontology for software architecture knowledge

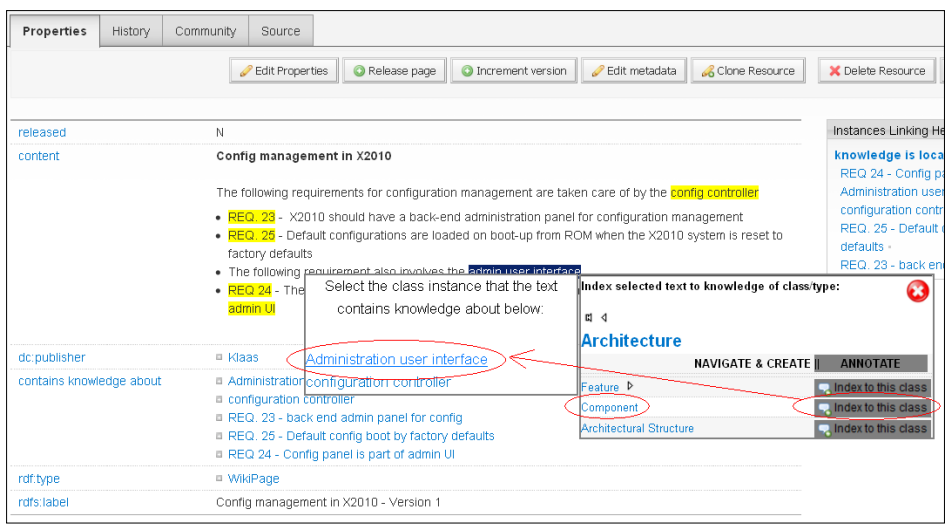


Figure 3 – semantic wiki

→ **Text** in software documents **is traceable** to formal knowledge and vice versa. Formal knowledge is interrelated.

→ **Quantitative evaluation**:
Increased time efficiency (~100%) and effectiveness (in Precision and Recall, ~34%) when comparing file-based documents (MS Word, diagrams) and semantic wiki for answering questions about software.

Statistical significance for questions such as:

"Which decisions have been made around component XX?"
"*Which settings have an impact on behaviour YY?"*

More **time saved** and **more correct** answers.

→ **Qualitative evaluation**:
"When searching for software knowledge, would you evaluate the semantic wiki, as compared to using normal documentation as:"
Better - 24 (92.3%) Worse - 0 (0%) Making no difference - 2 (7.7%)

Contact: k.a.de.graaf@vu.nl

Amsterdam Research on Talk & Text in Institutional Settings

Cienki, A., Lamerichs, J., Schasfoort, M., Sliedrecht, K.Y., Stommel, W., van Charldorp, T., van der Houwen, F.

WHO IS ARTTIS?

ARTTIS-VU is a group of researchers who focus on talk, text, and intertextuality as it is produced in different institutional settings:
- courtroom interaction
- police interrogations
- news interviews
- online health interaction
- job interviews
- televised political debates
- trauma interviews

In order to better understand such settings we approach our research materials from different perspectives and methods in the domain of interaction analysis, including but not limited to conversation analysis, discursive psychology, (critical) discourse analysis, and framing analysis.

We make an effort to apply the findings of our research in order to improve current communicative practices of the different target groups we study.

We meet at least once a month at VU University Amsterdam to discuss data and theoretical notions relevant to our fields of research. **If you want to join one of our ARTTIS data sessions, send us an email!**

From left to right and top to bottom:
Marca Schasfoort: Lecturer, Language & Communication, & founder of Schasfoort Communication consulting and training

Joyce Lamerichs: Assistant professor, Language & Communication

Fleur van der Houwen: Assistant professor, Language & Communication

Wyke Stommel: Postdoc, Language & Communication

Alan Cienki: Associate professor, English Linguistics

Keun Young Sliedrecht: PhD student, Language & Communication

Tessa van Charldorp: Assistant professor, Language & Communication

Martha Komter: Former member of ARTTIS, currently researcher at NSCR

How do we do Interaction Research?

DATA PREPARATION

First we make recordings (video or audio)

Then we make transcripts:

P= police officer

V= suspect

```
1. P: =vertel,
2.    (1)
3. V: °(ver)tel°.
4.    (0.4)
5.    hoe het verhaal is
      gegaan zeg maar?
6.    waar[om ik ben nu ben opgepakt,
7. P:     [↑ja,
```

ANALYSIS:

We look at ACTIONS that are accomplished:

In line 1, the police officer **invites** the suspect to tell a story. The suspect does a **pre-beginning** to his story in lines 3-6. He is also already **doing being innocent** in lines 5 & 6.

The symbols show the intonation details:

=	latching (utterance is latched onto previous utterance)
(1)	one second pause
__	underline indicates which syllable(s) is/are stressed
°	words in between ° signs is spoken softly
[	brackets indicate overlap
↑	indicates rising intonation

We look at HOW the actions are accomplished:

The officer tells the suspect to tell the story by using an imperative. The suspect treats this imperative as ambiguous. He repairs the question by repeating it and by rephrasing it twice. He rephrases by saying 'how the story went' and secondly, by saying 'why I've now been arrested.' His story therefore, will be one of being arrested, not one of what crime he committed. What we often see is that suspects are experienced at 'doing being ordinary' or 'doing being innocent.'

Scientific interest: how talk in institutional settings (in this case a police interrogation) is constructed and organised by participants in interaction.

Applied interest: advise police academy or police officers on what the effects are of certain ways of conducting interviews; help them reflect on ways of interviewing.

ARTTIS Project Examples

Keun Young Sliedrecht is writing her thesis about interaction in institutional settings. The data consists of 60 audio recordings of **job interviews, police interrogations and news interviews**. From a conversation-analytic perspective she studies how participants in these three institutional settings organize 'formulations' (Heritage & Watson, 1979) or summaries. With the analysis of formulations she sheds light on the way participants organize interactions in these institutional encounters and she compares her findings to the advice given to the institutional participants in handbooks.

Tessa van Charldorp completed her dissertation 'From police interrogation to police record' in 2011. In this dissertation she outlined how a written police record comes to be constructed through talk and typing. Her current research project is titled: '**Writing, blogging, facebooking, rapping, tweeting about criminality**. An insight into the life world, norms and values of the Dutch youth through websites, YouTube, Hyves, Facebook & Twitter.' Tessa looks at how Dutch youth talk about criminal activities online and what their language shows us about their norms and values. This project is financed by the Police Science & Research Programme.

Joyce Lamerichs & Marca Schasfoort are working together on a project on **trauma interviews with children**. They are using the perspective of conversation analysis and discursive psychology to conduct a fine- grained analysis of 25 interviews with children (age 8-12). The children have been interviewed by a trained psychologist who investigated how children have experienced single incident trauma (e.g. death in the family, a serious accident, murder) and what their strategies for coping are. The focus of Joyce and Marca's research is on ways in which questions by the psychologist are formulated to get at particular answers, and how the child's answers are subsequently taken up by the professional, for example to encourage the child to tell more. This project is being conducted in collaboration with Eva Alisic, a trained psychologist who works for the University Medical Centre (UMC) Utrecht.

Alan Cienki is working on a project titled '**The multimodal nature of talk in contexts of persuasive communication'**. This research considers the interplay of both audio and visual elements of spoken language use, such as how prosody and gesture with speech are deployed by speakers to frame their message as well as themselves. Specific contexts have included televised debates by politicians and pitches by entrepreneurs to persuade potential investors.

Fleur van der Houwen was trained as a sociolinguist specializing in the analysis of language variation and change. She draws on various methods, including speech act theory, conversation analysis, and critical discourse analysis. Her research focuses on **language use in criminal courts and e-health**, as well as on how institutions are represented in the media. Fleur is registered with the LDM as an expert in (forensic) linguistics and advises on cases that are sent to her for analysis. Cases vary from an analysis of secretly taped conversations, case file documents from a rape trial, authorship of letters of threat as well as other types of writing.

Wyke Stommel has recently been granted an NWO-grant (program Comprehensible Language And Effective Communication) for **a comparative study of chat and telephone Alcohol and Drugs help line interactions**. This project is being conducted in cooperation with Trimbos Institue, CAMeRA and Sensoor.

Sample Publications

Charldorp, T.C. van (2011). *From police interrogation to police record.* Oisterwijk: Boxpress.

Charldorp, T.C. van (2011). The interactional construction of the police record via the coordination of talking and typing in police interrogations. *Crossroads of Language, Interaction and Culture.*

Cienki, A. Multi-modal metaphor analysis. In Lynne Cameron (ed.), *Metaphor Analysis: Research Practice in Applied Linguistics, Social Sciences and the Humanities*. London: Equinox.

Cienki, A., & C. Müller (eds.). 2008. *Metaphor and Gesture*. Amsterdam: John Benjamins.

Houwen, F. van der. (2009). Formulating disputes. *Journal of Pragmatics, 41*, 10, 2072-2085.

Houwen, F. van der, Charldorp, T.C. van & Tavecchio, L. (*in preparation*) *Forensische linguïstiek: een introductie*. Boom Juridische Uitgevers

Lamerichs, J., Koenen, M., te Molder, H. (2009). Turning adolescents into analysts of their own discourse: Raising reflexive awareness of everyday talk to develop peer-based health activities. *Qualitative Health Research, 19*, 8, 1162-1175.

Lamerichs, J., & Molder, H.F.M. te (2009). `And then I'm really like . . .': `preliminary' self-quotations in adolescent talk. *Discourse Studies, 11,* 4, 401-419.

Lamerichs, J. & Molder, H.F.M. te (2011). Reflecting on your own talk: the Discursive Action Method at work. In C. Antaki (Ed.). *Applied Conversation Analysis. Intervention and Change in Institutional Talk*.

Müller, C., A. Cienki, E. Fricke & D. McNeill (eds.). In preparation. *Körper–Sprache–Kommunikation / Body–Language–Communication.* 2 vols. For the series *Handbücher zur Sprach- und Kommunikationswissenschaft*. Berlin: Mouton de Gruyter.

Onrust, M., Schasfoort, M. & Torck, D. (2002). Verbale communicatie: praktijken van gesproken en geschreven taalgebruik. In: T. Janssen (red), *Taal in Gebruik*. Den Haag: Sdu Uitgevers, 181-197

Sliedrecht, K.Y. & Charldorp, T.C. van (2011). Tussen spraak en schrift: de rol van samenvattingen in het politieverhoor. *Tijdschrift voor Taalbeheersing, 33,* 34-55.

Stommel, W. (in preparation) "This cybernetic remote control is not what I need" An exploratory study of online, direct complaining in institutional interaction.

Stommel, W. & van der Houwen, F. (forthcoming). Formulations in "trouble" chat sessions. In: Kupferberg, Irit & Izhak Gilat (eds.) *Computer-Mediated troubled talk*. Special Issue: Language@Internet

Stommel, W. & Lamerichs, J. (forthcoming, 2012) Communication in Online Support Groups: Advice and beyond. In: Hamilton, Heidi & Chou, Wen-Ying Sylvia (eds.) *Handbook of Language and Health Communication*, Routledge: New York.

PLEASE CONTACT US IF YOU ARE INTERESTED IN ANY OF THESE (or other) PUBLICATIONS!

Contact Information

ARTTIS, Department of Language & Communication
Faculty of Arts
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Phone Number: +31(020) 5983837
E-mail: t.c.van.charldorp@vu.nl

CLIF: A Method to Formalise Clinical Quality Indicators

Kathrin Dentler, Annette ten Teije, Ronald Cornet and Nicolette de Keizer
Academic Medical Center - University of Amsterdam & Vrije Universiteit Amsterdam

If you can not measure it, you can not improve it. Lord Kelvin

Problems, Goal and Requirement

Problems

- Indicators that are released in natural language are *ambiguous*
- *Increasing number* of indicators makes their manual calculation unfeasible

Goal

Ideally, indicators would be *calculated automatically* based on patient data that is captured during the care process.

Requirement

To derive comparable values, a *formalisation method* is needed to transform indicators from natural language into an *unambiguous, machine-processable, formal representation*.

Example Quality Indicator

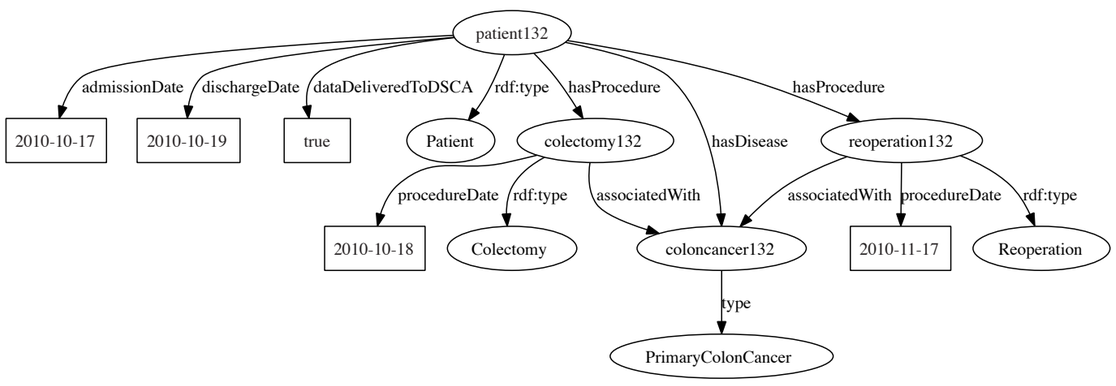
Number of examined lymph-nodes after resection

Numerator: Number of patients who had 10 or more lymph-nodes examined after a resection of a primary colon carcinoma.

Denominator: Number of patients who had lymph-nodes examined after a resection of a primary colon carcinoma.

Exclusion Criteria: Patients with previous radiotherapy; recurrent colon carcinoma.

Example Patient

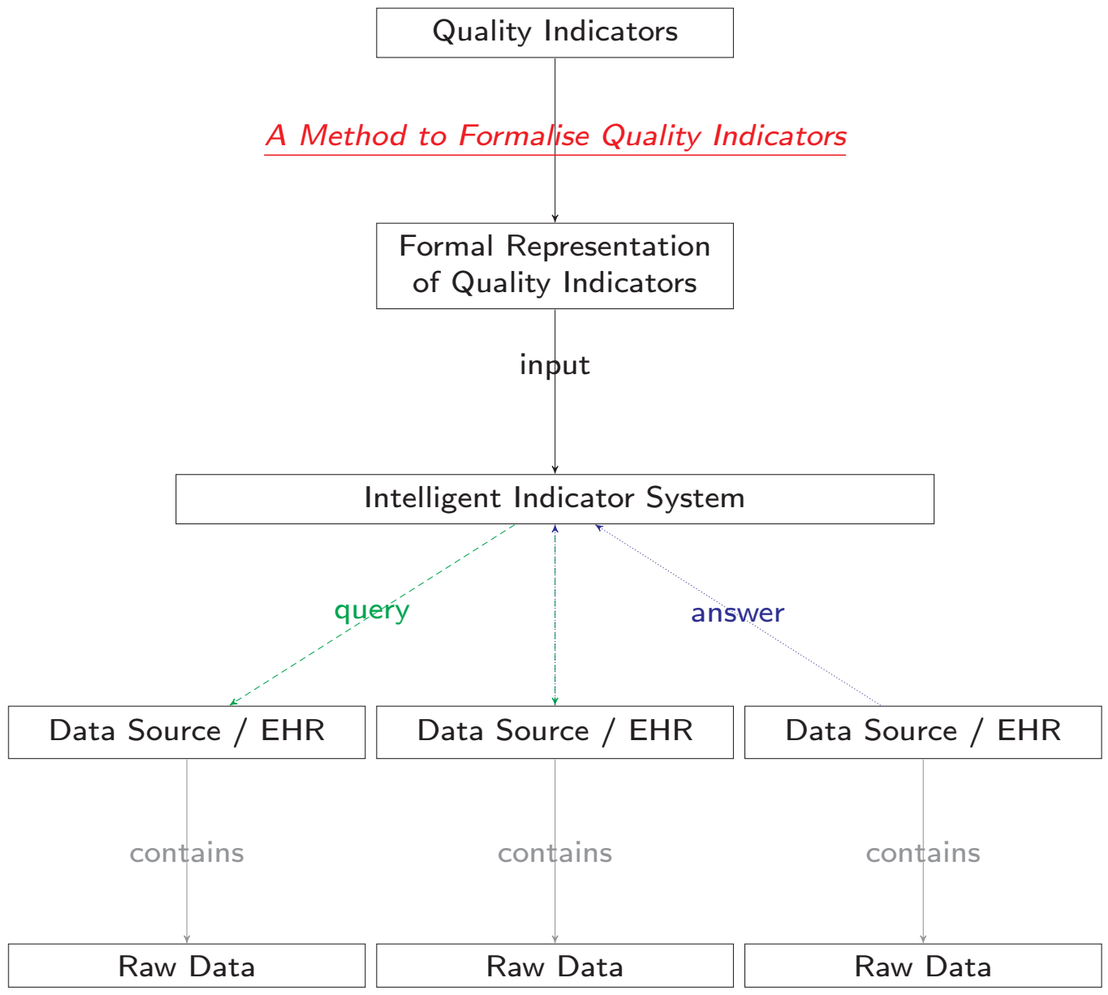


Definition Quality Indicator

A quality indicator is "a measurable element of practice performance for which there is evidence or consensus that it can be used to assess the quality, and hence change in the quality, of care provided."

(Lawrence 1997)

The Big Picture: Formalise indicators to calculate them automatically



Method to formalise Quality Indicators into SPARQL Queries

Step 1)
Encode relevant concepts (diagnoses and procedures) from the indicator by concepts from a terminology

?colectomy a sct:SCT_23968004 .

Step 2)
Define the information model
?patient schema:hasDisease ?coloncancer .

Step 3)
Formalise temporal constraints (FILTER - SPARQL FILTERs restrict solutions to those for which the filter expressions evaluate to true)
FILTER (?lymphnodeexaminationdate > ?colectomydate)

Step 4)
Formalise numeric constraints (FILTER)
FILTER (?numberexaminedlymphnodes ≥ 10)

Step 5)
Formalise boolean constraints (FILTER)
FILTER (?deliveredToDSCA = true)

Step 6)
Group constraints by boolean connectors
(coloncancer AND colectomycode) OR (rectumcancer AND resectionrectum)

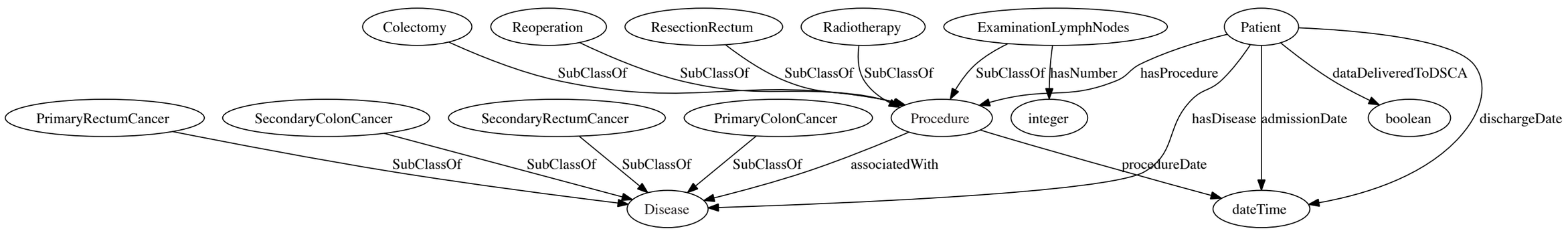
Step 7)
Mark exclusion criteria
FILTER (?lymphnodeexaminationdate > ?radiotherapydate) }

Step 8)
Mark constraints that only aim at the numerator
FILTER (?numberexaminedlymphnodes ≥ 10)

Resulting SPARQL query (Numerator)

```
SELECT ?patient WHERE

# Step 1)
?patient a sct:SCT_116154003 .
?coloncancer a sct:SCT_93761005 .
?colectomy a sct:SCT_23968004 .
?lymphnodeexamination a sct:SCT_284427004 .

# Step 2)
?colectomy sct:SCT_47429007 ?coloncancer .
?patient schema:hasDisease ?coloncancer .
?patient schema:hasProcedure ?colectomy .
?colectomy schema:procedureDate ?colectomydate .
?patient schema:hasProcedure ?lymphnodeexamination .
?lymphnodeexamination schema:procedureDate ?lymphnodeexaminationdate .
?lymphnodeexamination schema:hasNumber ?numberexaminedlymphnodes .

# Step 3)
FILTER ( ?lymphnodeexaminationdate > "2010-01-01T00:00:00+02:00" )
FILTER ( ?lymphnodeexaminationdate < "2011-01-01T00:00:00+02:00" )
FILTER ( ?lymphnodeexaminationdate > ?colectomydate)

# Step 4); needs to be removed to construct the denominator (# Step 8))
FILTER ( ?numberexaminedlymphnodes ≥ 10 )

# Step 7)
FILTER NOT EXISTS
?radiotherapy a sct:SCT_108290001 .
?patient schema:hasProcedure ?radiotherapy .
?radiotherapy schema:procedureDate ?radiotherapydate .
FILTER ( ?lymphnodeexaminationdate > ?radiotherapydate)
```

OWL Schema for Patient Data



mail: k.dentler@amc.uva.nl & k.dentler@vu.nl

A culture of safety

or the secret organisation of process and protocol

David Passenier
D.F.Passenier@vu.nl
Organisation Science
& the Metaphor Lab
www.metaphorlab.vu.nl
Vrije Universiteit Amsterdam

Oil & Gas

In 1988, **the Piper Alpha oil rig exploded** into a horrible fire. The disaster cost many lives. The investigation revealed critical **Human Errors**: errors and **violations of protocol** that showed commitment to safety was compromised. The industry became acutely aware that it had to **manage human errors**. Shell took the lead, developing the Reason Model of Accidents to **diagnose latent errors** that might one day cause disaster. From this model a **Safety Management System** was developed that organisations can use to manage their risks better. As industry leader, Shell now prides itself with its **uncompromising safety culture**.

The concept of 'safety culture' raises expectations that it cannot fulfil.

It suggests a 'strong culture' which somehow enables participants to know when to stick to procedure, and what situations are so extraordinary that improvising is allowed.

This image masks the benign uncertainties and routine improvisations that occur in daily practice. They usually go right, but sometimes they go wrong.

Societal responses to accidents are fierce, harming carefully constructed safety cultures with familiar politics of blaming.

Aviation

In 1977, **Human Error** caused the worst aircraft accident in history. On the foggy runway of **Tenerife Airport**, a KLM B747 and Pan Am B747 **crashed, killing 583** people. The investigation revealed that **the KLM Captain had violated a procedure and ignored his crewmembers'** correct reservations to commence takeoff. The industry realised it had to **flatten cockpit hierarchy** and create an **open atmosphere** towards errors. **Crew Resource Management** courses and training schemes since taught generations of pilots about the dangers of **complacency and hubris**. Aviation developed a **unique safety culture.** The majority of pilots will admit they are **only human**, make mistakes, and need others to correct them. Organisations also began to adopt and integrate explicit **Safety Management Systems**. They diligently update and apply **Standard Operating Procedures** to keep air transport safe.

Health Care

In 2001, an inquiry was completed on the **needless deaths** of many dozens of children in the **Bristol Royal Infirmary**. It was yet another shocking case of a **bad organisational culture** contributing to lethal **Human Errors** in a hospital. These findings contributed to putting **patient safety** on the agenda. Some health care organisations are now beginning to **adopt aviation's Crew Resource Management training** methods to improve cooperation in medical teams. Yet reports keep accumulating that medical staff are **not complying with the protocols** that are devised to keep patients safe. Whereas pilots believe in their own fallibility, **surgeons are still prone to complacency and hubris** because they do not as readily admit that they are fallible. They are under mounting pressure to **change their organisational structures**.

This project aims to create an empirically supported ***framework for safety culture*** that explicitly acknowledges ambiguous aspects of practice.

The project is an ***interdisciplinary*** effort, uniting an ***organisational anthropology*** approach to culture with a ***linguistic*** approach to discourse.

Models, images and particularly ***metaphors*** are believed to play a central role in the (re)production of meanings like safety culture.

Textual, social network and observational data are collected and analysed with an innovative combination of methods.

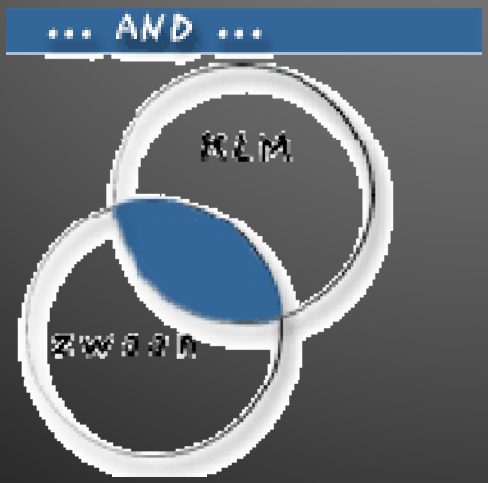
Design, semantics, and the brain:

Bo van Grinsven, VU University Amsterdam,
Communication Sciences

When does

become:

And for example:



Searching for the tipping point

- When does the composition of colors and shapes start activating brand association networks in consumers' minds?
- When do activated associations lead to recognition?
- Context: Research on logo (re)design

DutchSemCor: Targeting the ideal sense-tagged corpus

Piek Vossen[1], Attila Görög[1], Rubén Izquierdo[2], Antal van den Bosch[2]

1 VU University Amsterdam. Amsterdam. The Netherlands
2 Tilburg University. Tilburg. The Netherlands

Build a sense-tagged corpus for Dutch

- **Represent all senses of words**
- **Represent the variety of contexts**
- **Provide information on the sense distribution**

- **3000 most frequent Dutch words (also most polysemous)**
- **100 examples per sense**
- **3-4 avg. senses per word --> 1 million tokens**

Sense tagged corpus

- **Annotation method**
  - **Sequential tagging**
  - **Targeted tagging**
- **Textual coverage**
  - **All - words**
  - **Lexical sample**
    - **Balanced-sense**
    - **Balanced-context**
    - **Sense-Probability**

Project Methodology. 3 Phases

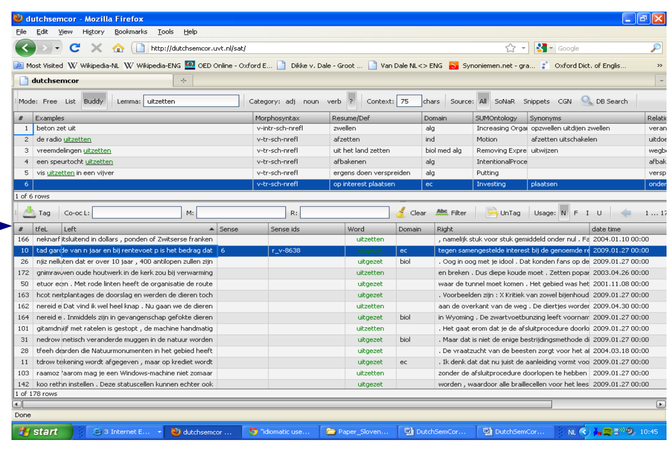
1)Human manual annotation

human + human

- Double annotation for each token
  - 282,503 tokens
  - Very clear and good examples selected

Goal: *Balanced-sense corpus*

SAT Tool - KWIC



SONAR CORPUS — **67 %**

Corpus Gesproken Nederlands — **5 %**

INTERNET — **28 %**

A 500-million-token corpus is not big enough to create a balanced-sense corpus !!!

- At the end of the annotation
  - 80% senses with 25 of more examples
  - 90% of lemmas with 25 examples per sense
  - This set is called INITIAL LEARNING

2)Active Learning

human + computer

- Only lemmas performing < 80 in accuracy are processed
- 50 examples per sense according to:
  - TIMBL confidence
  - Distance to the nearest neighbor
    - Low Distance (LD): similar examples
    - High Distance (HD): different examples

Goal: *Balanced-context corpus*

WSD System K-NN algorithm: TIMBL

Data	Token Accuracy	# Examples
Initial Learning	81.62	8641
IL + LowDist	78.87	13266
IL + LowDist_agree	85.02	11405
IL + HighDist	76.24	19055
IL + HighDist_agree	83,77	13359
IL + LowDist_agree + HighDist_agree	85.33	16123

3) Clustering

Goal: *sense-probability corpus*

- Similar to Word Sense Induction
- Clustering techniques different to WSD
- Cluster remain not tagged SONAR to:
  - Discover new senses
  - Use annotated instances to discover clusters with a predominant word sense and automatically tag the cluster

vrije Universiteit amsterdam

Tilburg University

Impacts of Multiple Stakeholders' Involvements on Higher Education Sector in Cambodia

Chanphirun Sam
Department of Organizational Sciences, Faculty of Social Sciences
c.sam@vu.nl
vrije Universiteit amsterdam

Introduction

- Over the last few decades, the Cambodian higher education has undergone a lot of reforms and restructuring to respond to the growing demand of access to higher education and changing labor market.
- The rapid expansion of HEIs has challenged the quality and efficiency of the education services.
- Most Cambodian universities are mainly teaching institutions though only a few public universities have started to prove some of their research activities.
- To secure their place in the global knowledge-based economy, universities are supposed to reconsider their role and relationship with various stakeholders.
- Due to the importance of entrepreneurship for the economic growth, the value of entrepreneurship education at HEIs has been widely recognized to develop entrepreneurial spirits among the students.
- The research studies on stakeholders' involvements and entrepreneurship education in Cambodian higher education have not yet been conducted.
- This study is intended to examine the Cambodian higher education sector by identifying the stakeholders, seeking insights into the impacts of multiple stakeholders' involvements, and look into the roles of higher education institutions in promoting entrepreneurship.

Research questions

Main question

- In what ways does the involvement of multiple stakeholders influence the higher education sector and entrepreneurship education in Cambodia?

Sub-questions

- Who are the stakeholders in the Cambodian higher education sector?
- What are the roles of the stakeholders in the Cambodian higher education sector?
- What are the stakeholders' interventions on the curricula and policies within the Cambodian higher education institutions?
- To what extent do higher education institutions promote entrepreneurship education?
- How does the involvement of multiple stakeholders affect entrepreneurship education in the Cambodian HEIs?

Methods

Research setting: Phnom Penh

Sampling methods

- Purposive sampling: Key informants from Ministry of Education, Youth and Sport (MoEYS), Concerned ministries, Accreditation Center of Cambodia (ACC), Donor Agencies, Higher Education Institutions
- Snowball sampling: Managers from local and international firms, NGOs

Data collection methods

- In-depth interview with key informants
- Focus group interview with government officials at MoEYS and concerned ministries, and ACC staff, and lecturers
- Stakeholder conference

Secondary data

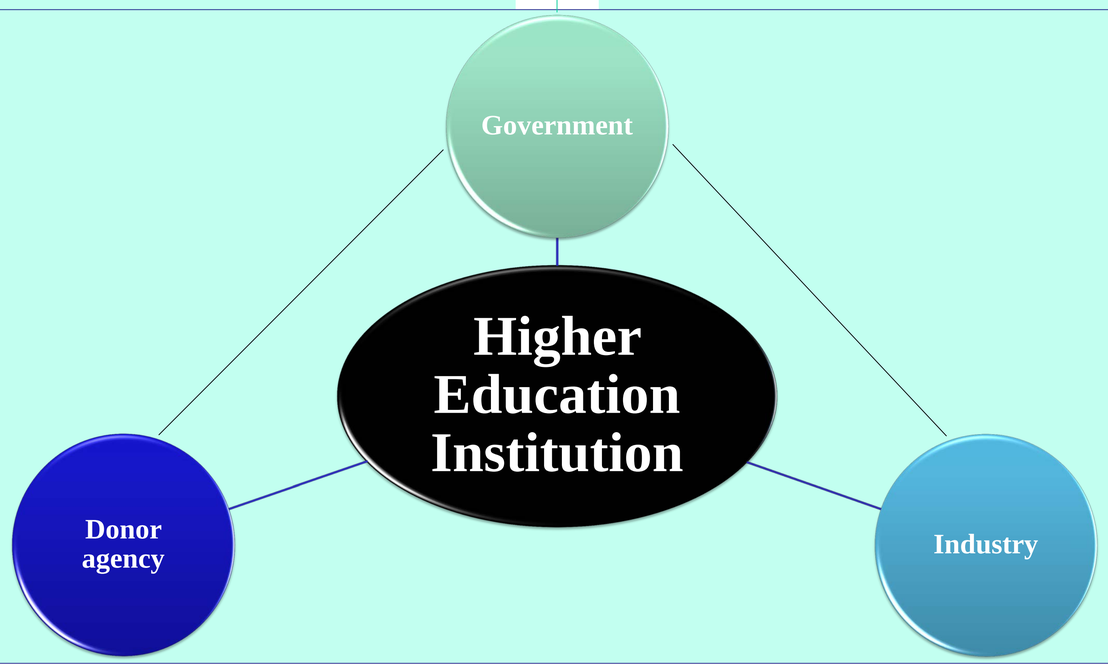
- Government development policy
- Education policy, Education Strategic Plan, Higher Education policy from MoEYS
- University curricula, internal policies and strategic plans from universities
- Accreditation policy on higher education

Data analysis methods

- Stakeholder analysis
- Content analysis
- Policy analysis
- Document analysis

Significance and expected outcomes

- To help universities categorize and determine the important stakeholders for partnership, important for their governance and accountability arrangements;
- To improve higher education curricula and policies within HEIs to respond to the needs of the knowledge-based economy;
- Research on entrepreneurship education at HEIs is limited in developing countries. No research on entrepreneurship education and the impacts of multiple stakeholders' involvements in the Cambodian higher education has been conducted. This study, therefore, contributes to the existing body of knowledge for developing countries, especially Cambodia.



Identifying and Analyzing Visual Metaphor

Ester Šorm

VU VRIJE UNIVERSITEIT AMSTERDAM
Faculteit der Letteren

The project

Researchers in various disciplines agree that metaphor may be found in various modalities and media of communication, including gesture, film, and still images (Cienki & Müller, 2008; Gibbs, 2008; Forceville & Urios-Aparisi, 2009). However, **non-verbal metaphor** has not received as much attention as metaphor in language.

The project aims at the development of a **method for visual metaphor identification and analysis**. It capitalizes on the method developed for verbal metaphor identification by Gerard Steen and colleagues (2010), and also on the five-step method for reconstructing conceptual mappings across domains (Steen, 2009).

An Example of Visual Metaphor

Le cabinet anthropomorphique by Salvador Dalí, 1936

Strength of the Method

- **Explicitness** of the procedures
- Testing the **reliability** of the procedures
- Applicable to **wide range of materials**, including:
  - Paintings
  - Cartoons
  - Advertisements
  - Educational illustrations

A Method for Visual Metaphor Identification and Analysis

1. A seven-step procedure for **identifying metaphor-related visual units.**

2. A five-step procedure for **analyzing conceptual mappings between domains.**

References

Cienki, A., & Müller, C. (2008). *Metaphor and gesture.* Amsterdam: John Benjamins.
Forceville, C., & Urios-Aparisi, E. (2009). *Multimodal metaphor*. Berlin / New York: Mouton de Gruyter.
Gibbs, R. (2008). *Cambridge handbook of metaphor and thought*. Cambridge: Cambridge University Press.
Steen, G.J. (2009). From linguistic form to conceptual structure in five steps: analyzing metaphor in poetry. In G. Brône & J. Vandaele (Eds.), *Cognitive poetics: Goals, gains and gaps* (pp. 197-226). Berlin/ New York: Mouton de Gruyter.
Steen, G.J., Dorst, A.G., Herrmann, J.B., Kaal, A.A., Krennmayr, T., Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam: John Benjamins.

The Metaphor Lab

The Metaphor Lab is an internationally unique expertise center for Metaphor Studies. Its mission is to stimulate interdisciplinary collaboration, innovation, and application in all areas of metaphor research inside and outside academic communities.

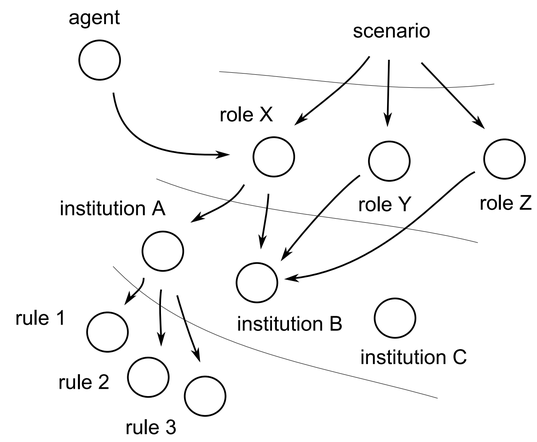
The Metaphor Lab performs fundamental research on metaphor in language, cognition, and communication is combined with applied research on metaphor in varied domains of discourse such as media use, organization, management, health etc.

The lab specializes in linguistics and discourse studies, psycholinguistics and cognitive science, sociolinguistics and social science, and applied linguistics. It seeks to develop new products and services for the non-academic community, to improve self-awareness, monitoring, training and testing of metaphor use in design, management, communication, and interaction.

Implementing the institutional stance

toward an agent-based modelling framework for legal, socio-economic scenarios

Giovanni Sileno - Leibniz Center for Law / UvA



Law changes, people change

Economic dynamics, historical changes, new social trends.. And still, the complexity of translating into regulations a given policy. The social relevance of evaluating an existing or a future regulative implementation, also considering non-compliance modes..

Let the people do what they do best

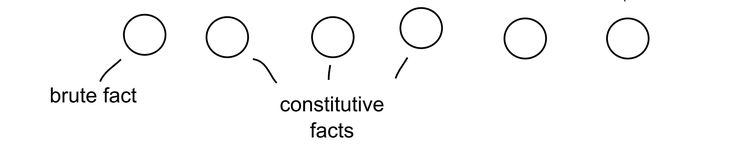
Identifying the contextual common patterns between individual cases, people (and among them, experts) can write down typical scenarios (as patterns of social interaction) and the typical (social) roles the agents play in them. Animating these stories means validating the problem and solution spaces given by the experts.

Administrative organizations and policy makers need new tools

A mixed reality agent simulation framework, integrating cognitive models ranging from individual practical interpretations of behaviour to simplified/ shared/monolithic conceptualizations..

The base of modelling

In empirical sciences, brute facts. In law? Real facts are processed *somehow* to count as legal facts, while legal reality intervenes *somehow* with normative intention on reality. Both *somehow*s are key concepts.



The institutional stance

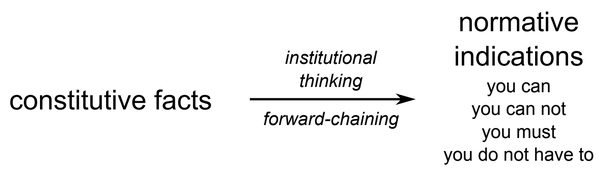
Proceeding along the path traced by Searle, an institution is defined by certain rules and some institutional facts. This conceptualisation unifies games, social informal norms and legal norms.

Institutions as internal agents

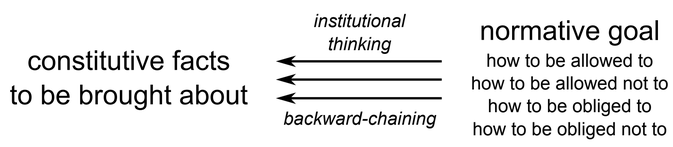
From the agent's point of view, to be complying to an institution means to: • be aware of the present institutional state (rules and facts) • reason and create the associated normative rules • behave accordingly. This institutional thinking naturally leads to implement institutions as agents with direct and unique communi-cation with the agent they superevene on.

Institutional thinking

Constitutive and institutional rules can be expressed in the form of normative conditionals: *if CONDITIONS then*^{n} *CONCLUSION*.



Every new institutional fact draws a new configuration of the jural relations between entities defined in that institu-tion. This new configuration is a result of a forward-chaining reasoning issued with an appropriate logic. However, it is interesting to consider also the other way around, the backward-chaining (conclusion, if condition).



Practical normative indications

Agents can be complying just because the norms exist, or because they can evaluate some positive outcome from them (in a sort of game-theory perspective). Both decision attitudes (*deontic* or *consequen-tialistic*) are strictly related to the decision-making cycle of the agent. Different atti-tudes could be associated to different components.

Conflicting plans

The modeler has to handle *norm vs desire* conflicts, altogheter with *desire vs desire* conflicts (caused for example by conflicts between roles), *norm vs norm* (between institutions), *rule vs rule* conflicts (for an internal institutional conflict).

Scenarios and roles

As a result, in the proposed framework roles, institutions and rules become basic modelling components for the agents.

Implementation in *Jason*

Jason is a popular multi-agent system development platform based on a variant of the AgentSpeak language and built on *Java*. It is based on logic programming and the BDI architecture for cognitive autonomous agents.

Three modelling exercises:

a mythological example of non-compliance
- *the story of Achilles avoiding the Trojan War,*

an institution as formalised by the law
- *the sale process as defined in common law,*

a story from an administrative organisation
- *a tax evasion scheme in real-estate transactions.*

Conclusions

Jason has proved to be an elegant and intuitive platform, but serves more as a middle-ware than a modelling platform that embodies an institutional perspective. In order to develop all the potentialities of the proposed conceptualisation, a strong extension is required.

Proposed developments

An ABM-oriented programming platform unifying: • the elegance and easiness of the syntax of Jason (AgentSpeak) • a stronger modularisation and concurrency capability • the perspective to integrate with expert systems or other knowledge technologies.

Preliminary design principles:

- BDI architecture
- dataflow programming oriented
- entity/architecture dichotomy
- forward-chaining operator (for inferences)
- backward-chaining operator (for plans)
- private encapsulation (internal agents)
- public inclusion (shared knowledge)
- explicit maintenance goals
- procedural and declarative memory
- synchronous and asynchronous comm.
- sequential and parallel constructs
- memory retrieval / conflict resolution plans
- three valued logic (false, true, unknown)
- belief annotations
- embedded time
- emb. provenance (percept, comm., inference)
- internal conceptualization of identity

Acknowledgments

This research is part of the AGILE project – Advanced Governance of Information services through Legal Engineering – targeted at the development of a design method, distributed service architecture and support tools that enable organisations to better govern their legislation and regulation based information services within in a networked envirodfsdsnment. AGILE is a cooperation between the Leibniz Center for Law and the Technical University Delft, O&i Management Partners, Be Informed and the Dutch Immigration and Naturalisation Service (IND).

Reasoning Service for the Support of Genome-Wide Association Studies

Zhisheng Huang[1], Gaston Tagni[1], Mark A. Greenwood[2], Angus Roberts[2], Raluca Brehar[3], and Mattias Johansson[4]

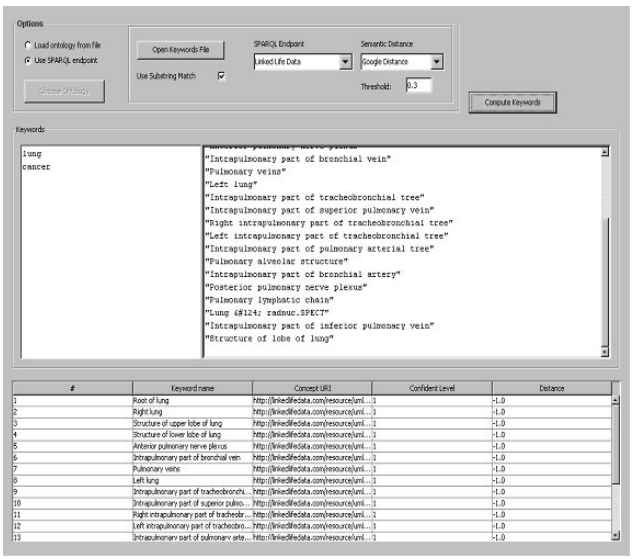
1.Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands

2. Department of Computer Science, University of Sheffield, UK

3. Technical University of Cluj-Napoca, Romania

4 International Agency for Research on Cancer (IARC), World Health Organization

Genome-Wide Association Studies have been widely used to identify common genetic factors that influence health and disease. We have developed a reasoning service to support the GWAS scientists in obtaining additional information which can be incorporated into the statistical association models to improve ranking techniques. The service is achieved with the support of the LarKC platform for reasoning over massive amounts of biomedical semantic data. Statistical and semantic distance measures, such as Google distance, are used to aid reasoning.



Pseudo Code For Deriving New Keywords

```
inputKeywords = initial set of keywords (defined by domain expert)

allFlag = true

threshold = the threshold defined by the user (eg. 0.7)

useSubstringMatch = true

FOR EACH keyword K in inputKeywords DO{
  concepts = search for ontology concepts whose name  is an exact
match with K's string value;

  FOR EACH concept C in concepts DO{
    IF (allFlag){
      outputKeywords += all the descendants of concept C;
    }ELSE{
      outputKeywords += only direct sub concepts of concept C; }

    IF (useSubstringMatch){
      otherConcepts = search for ontology concepts whose  name is a
super-string of K's string value

      FOR EACH concept D in otherConcepts DO{
        IF (distance(D,K) <= threshold){
          IF (allFlag){
            outputKeywords += all descendants of concept D
          }ELSE{
            outputKeywords += only direct sub concepts of concept D;
          }
        }
        outputKeywords += D
      }
    }
  }
}
return outputKeywords;
```

Noisy Semantic Data Processing in Seoul Road Sign Management System

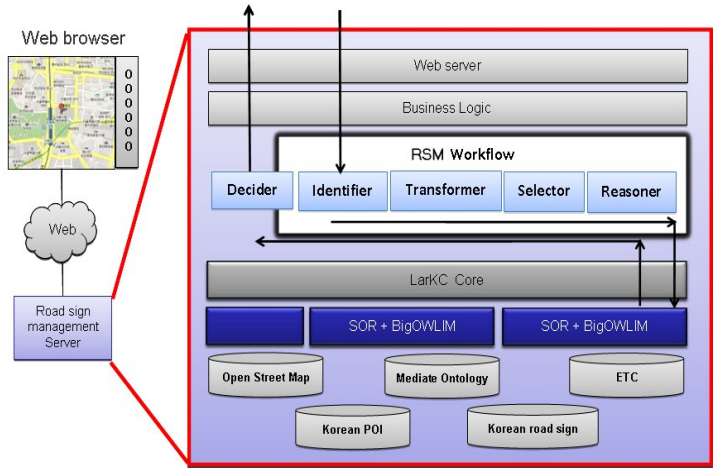
Zhisheng Huang[1], Jun Fang[2], Stanley Park[3], and Tony Lee[3]

1.Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands

2.School of Automation, Northwestern Polytechnical University, China

3. Saltlux Inc., Seoul, Korea

The Seoul Road Sign Management (RSM) is a system which provides the semantic integration of LOD's Linked Geo Data and Open Street Map with Korean POI data set. That is an attempt to develop intelligent road sign management system based on the LarKC platform. The RSM data set contains over 1.1 billion triples of semantic data. However, significant amount of the RSM data are noisy (e.g., inconsistent, partial, or erroneous). We have facilitated the RSM system with the capability of processing and reasoning with noisy semantic data, so that the RSM system is robust enough to return intended answers in spite of the poor quality of the semantic data.



Dataset	Features and Operations
Linked Geo Data (LGD)	1 billion triples in WGS84 coordinate Loading LGD full and extracted POIs written in English
Open Street Map (OSM)	Extracting all way information in WGS84 coordinate Selecting and importing 2 million triples for Seoul
POI Data in Korea (KPOI)	1 million POIs related with road signs written in Korean Around 4 million triples
Seoul road sign data (RSD)	Diverse data set of Seoul road signs Extracting half million triples from RDB Converting TM coordinate into WGS84 coordinate
Korean road sign regulations (RSR)	Around 30 Regulations of road sign about positioning and naming. Changing into SparQL for validation check
Mediate Ontology (MO)	Ontology linking between OSM, KPOI and RSD or other data Expressivity : subClassOf, subPropertyOf, sameAs, inverseOf

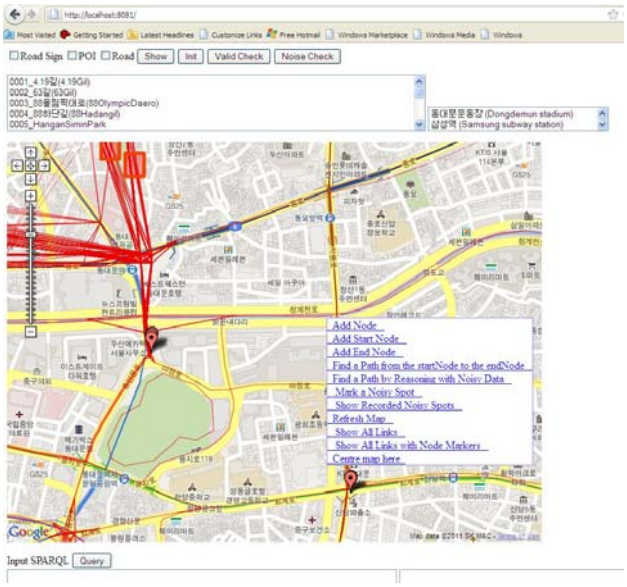
Table 1. RSM data set

(a) Partial Data

(b) Inconsistent Data

(c) Duplicated Data

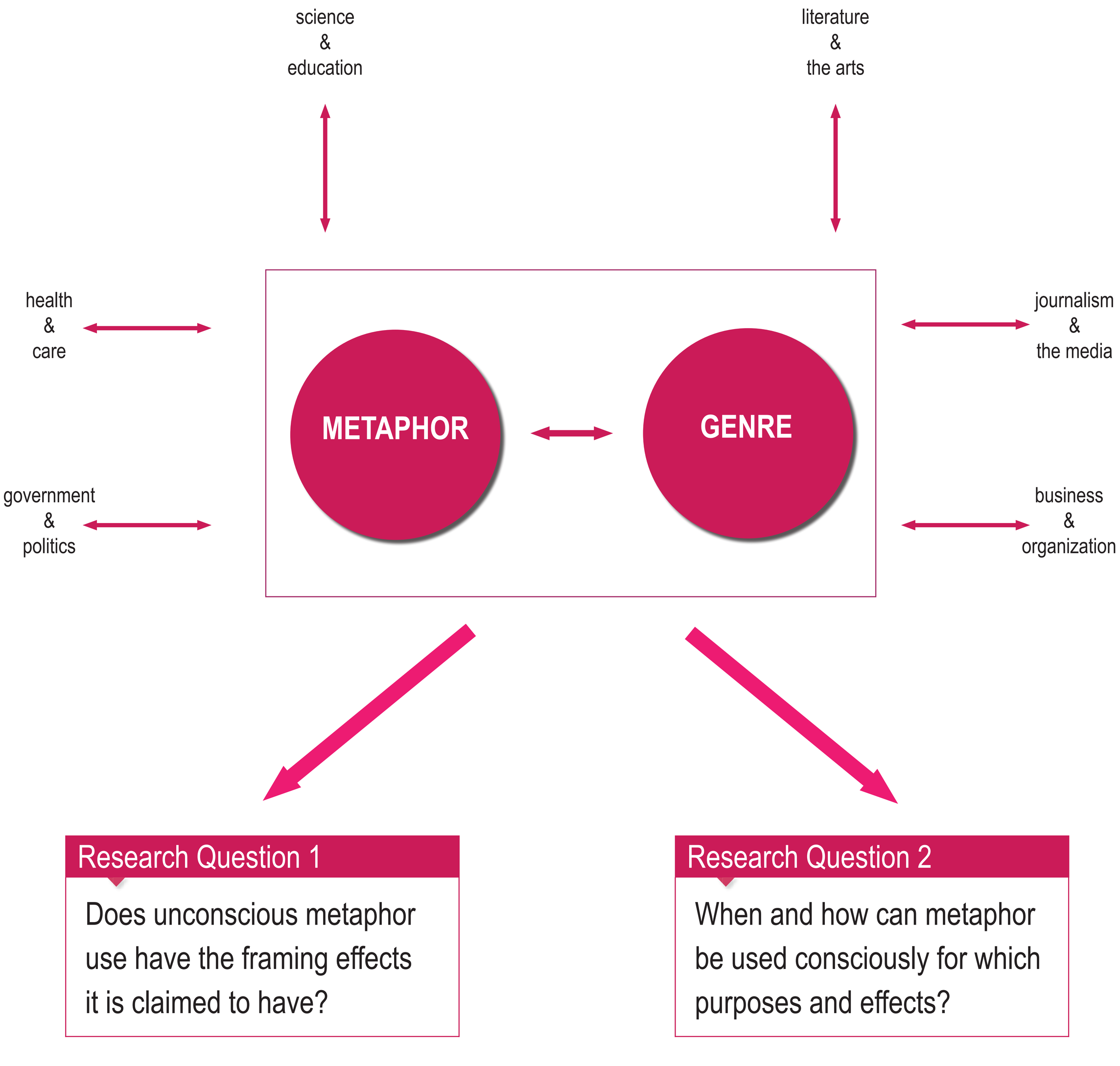
(d) Erroneous Data



Metaphor in Discourse: Time and Consciousness

Gerard Steen

VU UNIVERSITY AMSTERDAM | Faculty of Arts



Research Question 1

Does unconscious metaphor use have the framing effects it is claimed to have?

Research Question 2

When and how can metaphor be used consciously for which purposes and effects?

References

- Steen, G. (2011a). The contemporary theory of metaphor -- now new and improved! *Review of Cognitive Linguistics*, 9(1), 26-64.
- Steen, G. (2011b). Genre between the humanities and the sciences. In M. Callies, W. Keller, & A. Lohöfer (Eds.), *Bi-directionality in the cognitive sciences - Examining the interdisciplinary potential of cognitive approaches in linguistics and literary studies* (pp. 21-42). Amsterdam/Philadelphia: John Benjamins.
- Steen, G. (in press). Deliberate metaphor affords conscious metaphorical cognition. *Cognitive Semiotics, special issue on 30 years of Conceptual Metaphor Theory*.

The Metaphor Lab

The Metaphor Lab is an internationally unique expertise center for Metaphor Studies. Its mission is to stimulate interdisciplinary collaboration, innovation, and application in all areas of metaphor research inside and outside academic communities.

The Metaphor Lab performs fundamental research on metaphor in language, cognition, and communication is combined with applied research on metaphor in varied domains of discourse such as media use, organization, management, health etc.

The Lab specializes in linguistics and discourse studies, psycholinguistics and cognitive science, sociolinguistics and social science, and applied linguistics. It seeks to develop new products and services for the non-academic community, to improve self-awareness, monitoring, training and testing of metaphor use in design, management, communication, and interaction.

Modeling optimal timing for reinforcing decisions for elderly people

Iulian M. Covlescu, Peter Roelofsma,
VU University Amsterdam,The Netherlands

Objectives

Designing an **mathematic model** to compute an **optimal time** when to prompt elders to keep their decision of being more social active

Situation studied

An elder decides he wants to meet with someone from his social network at a certain time point in future.

We want to have a system able to answer the question:

What is the optimal time when should the system prompt him about his meeting, so that the probability to reach it increases.

Model

The model uses 5 variables depending on the elder's behavior and his social network :

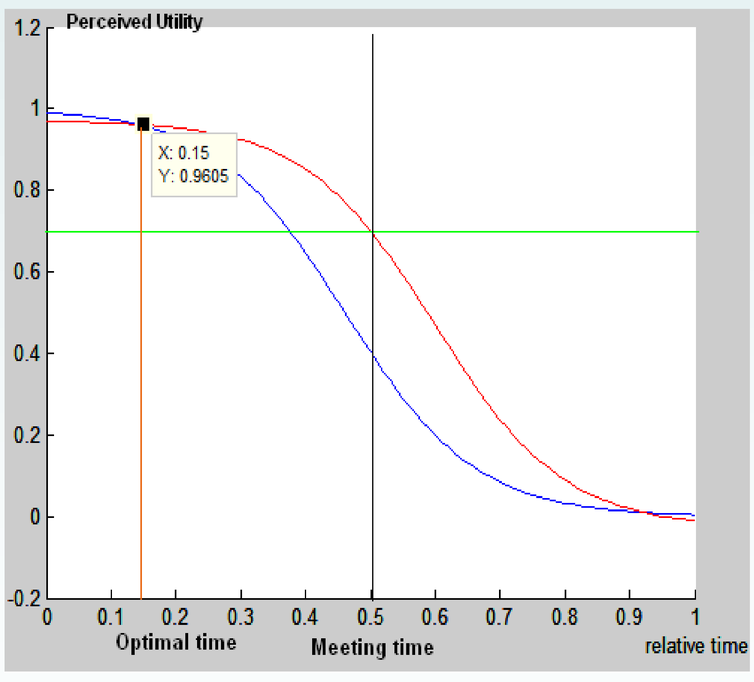
-***Perceived utility discount rate***

-***Attachment towards the goal***

-***Relative time of the meeting***

-***Perceived utility of meeting a person***

-***Personal Threshold***



Perceived utility after prompting

Perceived utility before prompting

Personal Threshold

Prompt

Fig1. Expected perceived utility variance, with one prompting at optimal time, and without.

Testing the Model

1 We are currently developing a **simulation** in the in a multi-agent modeling environment (*Netlogo*) that simulates a social network of an elder.(Fig2).

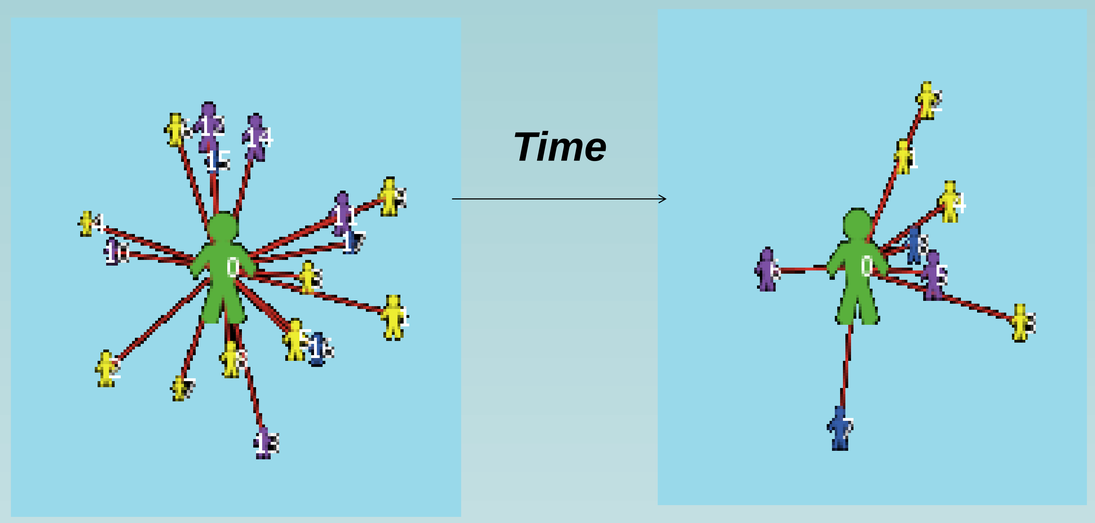


Figure 2. Evolution of social network of elders

The goal of the simulation is to compare the evolution of one's social network with regular prompting and with our prompting model deployed.

2. We have developed a **lesson** in the project ***Virtual Coach Reaches Out To Me*** that will implement the model we developed. This project is an ambient assisted living solution to prevent loneliness.
For more information please access:

or go to the website http://www.v2me.org/

Once data will be collected, we will be able to validate our model on it.

Future work and Conclusions

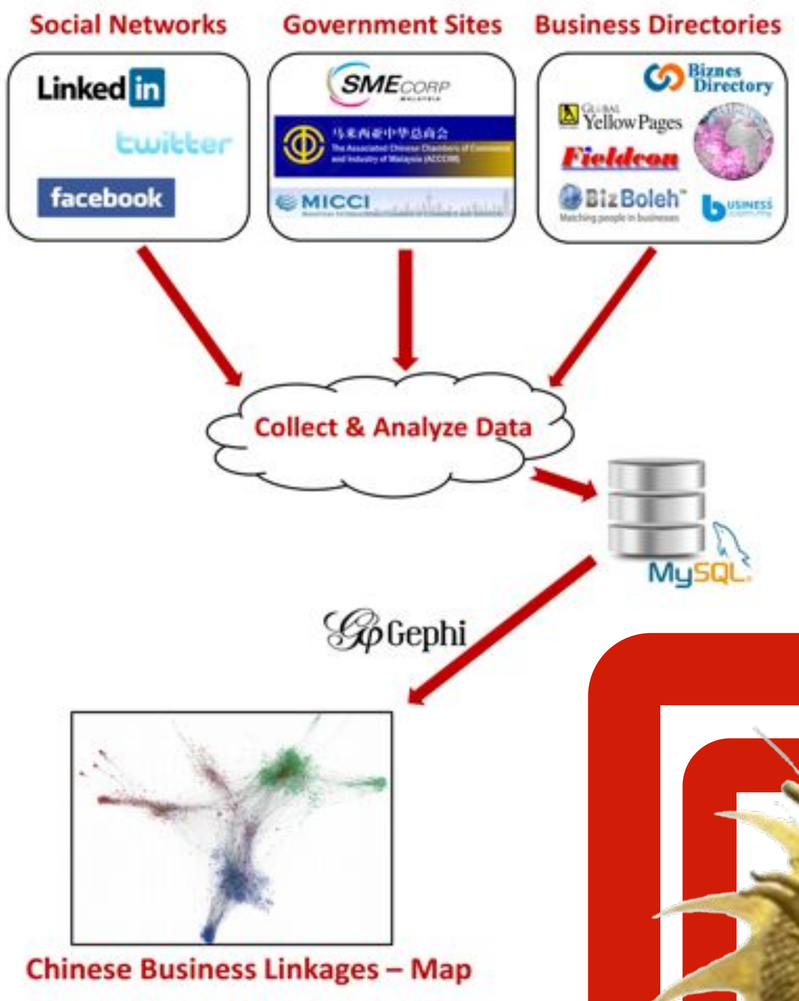
From what we are aware, this is the first attempt of automatic prompting based on the utility of meeting someone from a social network.

In our future work we aim to examine to what extend our model can be generalized to more situations.

Network Capitalism: Tracing Chinese Business Linkages through Web site Analysis

Cristina-Iulia Bucur, Katja Jönsas, Marieke van Erp, and Heidi Dahles

Workflow



Theory

It is argued that the success of ethnic Chinese businesses has to be explained in terms of the networking strategies of the ethnic Chinese, in particular the maintenance of close-knit and, at the same time, far-flung network relations within the worldwide ethnic Chinese 'community'.

While the argument about the Chinese networking virtuosity has come to be widely accepted, it has never been empirically tested.

We aim to test this argument by mapping Chinese network relations through linkages between ethnic Chinese business Web pages and Web page content.

First Findings

Ethnic Chinese run their business in a flexible and fluid manner:

→ Illustrated by hastily set up Web pages, disorganized layouts, spelling errors; sites are ill-maintained & quickly abandoned

The business directories indicate that the ethnic Chinese businessmen have diverse and overlapping networks:

→ To avoid a single point of failure for example through uncooperative bureaucrats

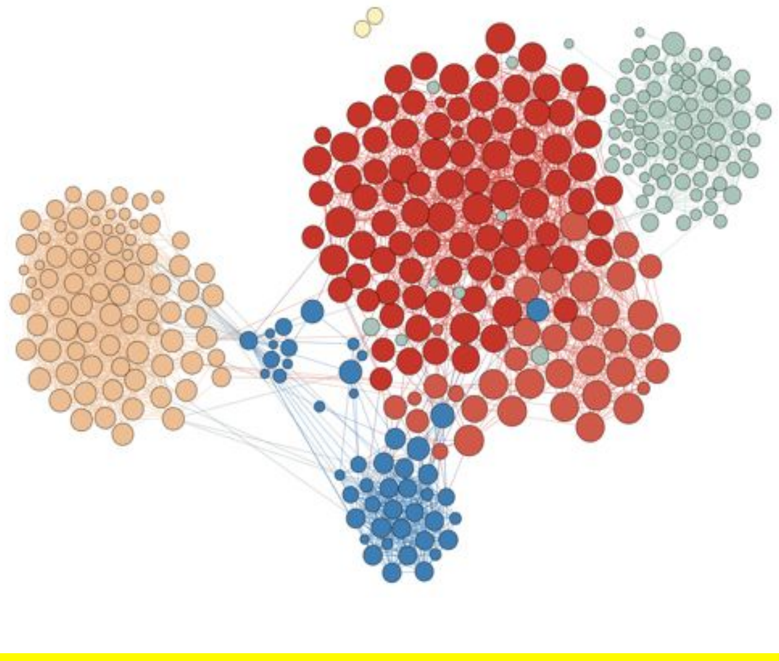
Low Web site traffic measurements and low rankings:

→ This may indicate that the Internet has not yet become a key part of ethnic Chinese businesses

→ Complicates automatic analysis

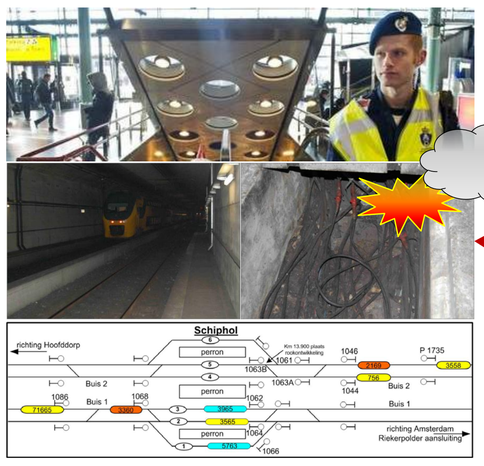
Network Analysis

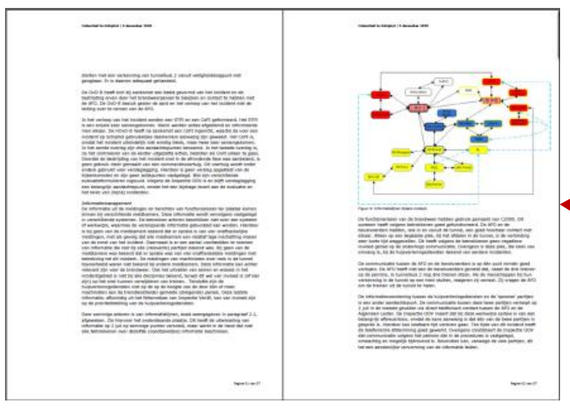
We represent the linkages between ethnic Chinese businesses in a layered network based on the economic domains in which they are active.



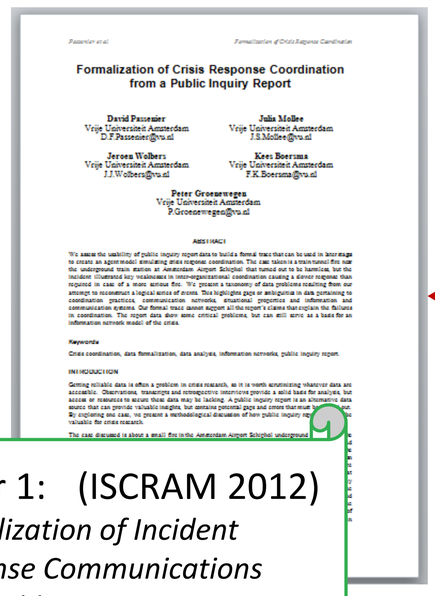
The Network Institute, VU University Amsterdam

Networks & Crises





Instigated by the case of the 2009 Schiphol train-tunnel fire, we analyze crisis situations from an interdisciplinary perspective. The domain of organization science is merged with artificial intelligence approaches, first by modeling the events of the incident response and second by modeling agents and decision making processes involved in crisis response. Our main data source consist of primary data and secondary data, i.e. reports, which are complemented by interviews with practitioners.



Formalization of Crisis Response Coordination from a Public Inquiry Report

David Passenier, Vrije Universiteit Amsterdam, D.F.Passenier@vu.nl
Julia Mollee, Vrije Universiteit Amsterdam, J.S.Mollee@vu.nl
Jeroen Wolbers, Vrije Universiteit Amsterdam, J.J.Wolbers@vu.nl
Kees Boersma, Vrije Universiteit Amsterdam, F.K.Boersma@vu.nl
Peter Groenewegen, Vrije Universiteit Amsterdam, P.Groenewegen@vu.nl

Paper 1: (ISCRAM 2012)
Formalization of Incident Response Communications from Public Inquiry Reports.

Incident

Report

Paper 5 (EXPECTED)

Matlab Tool

Formal Trace

Knowledge

Theory

Practitioner Interviews

Paper 3 (EXPECTED)

Paper 4 (EXPECTED)

Agent Model

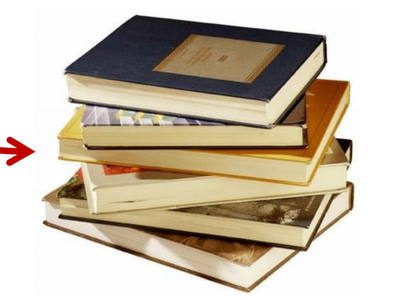
Formal Trace 1

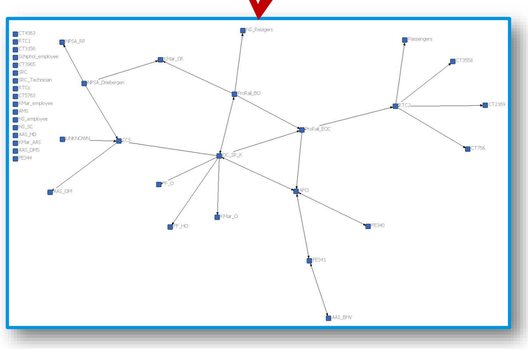
Formal Trace 2

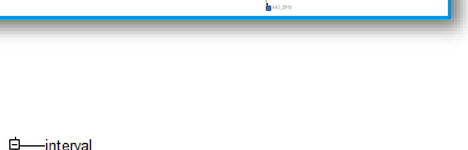
Formal Trace 3

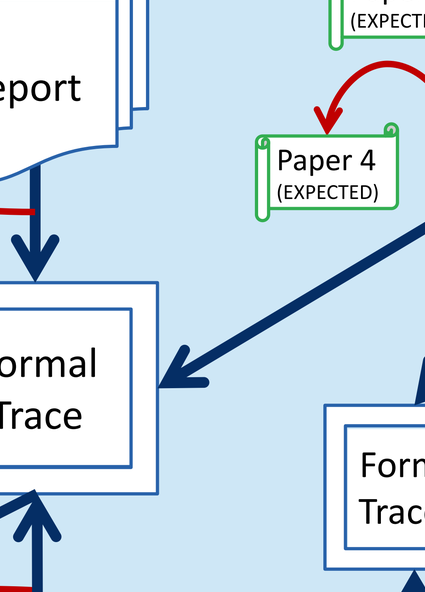
Communication Networks

Evaluation









CRISIS MANAGEMENT EVALUATION: FORMALISATION & ANALYSIS OF COMMUNICATION DURING FIRE INCIDENT IN AMSTERDAM AIRPORT TRAIN TUNNEL

Paper 2: (ECMS 2012)
Formalisation & Analysis of Communication during Fire Incident in Amsterdam Airport Train Tunnel.

Property P1B: Evacuation performed within 15 min at fire location.
For all time points *t1* and *t2*, all AGENTS *a* and *b* in trace γ, if at *t1* there is a fire at location tunnel 2A and there is no earlier time point at which there is a fire at location tunnel 2A, and at a later time point *t2*, AGENT *a* communicates to AGENT *b* that the tunnel is clear of trains, then interval *i* = *t2*-*t1* and *i* ≤ 30.

```
P1B_EVACUATION_PERFORMED_WITHIN_15MINUTES_AT_FIRE_LOCATION ≡
∀γ:TRACE , ∀t1,t2:TIME, ∀a,b: AGENT
state(γ, t1) |= world_state(at_location(fire, tunnel2A)) &
∀t0:TIME < t1 [state(γ, t0) |≠ world_state(at_location(fire,
tunnel2A))] &
state(γ, t2) |= communication_from_to(a,b, sign_clear(trains))
&t1≤t2
⇒
∃i: INTEGER
& i = t2-t1
& i ≤ 30
```

Contact	
Jan:	j.bim@vu.nl
Julienka:	j.s.mollee@vu.nl
Anne-Meike:	a.m.dewiljes@vu.nl

Social Science	Academy Assistants	Computer Science
Prof. Dr. Peter Groenewegen	Jan Bím	Dr. Natalie van der Wal
Drs. Jeroen Wolbers	Julienka Mollee	Dr. Tibor Bosse
	Anne-Meike De Wiljes	

VU UNIVERSITY AMSTERDAM
LOOKING FURTHER

Patterns of Clinical Trial Eligibility Criteria

Krystyna Milian[1], Annette ten Teije[1], Anca Bucur[2], Frank van Harmelen[1]
[1] Vrije Universiteit Amsterdam, The Netherlands; [2]Philips Research, The Netherlands

ABSTRACT

Medical research would benefit from automatic methods that support eligibility evaluation for patient enrollment in clinical trials and design of eligibility criteria. In this study we addressed the problem of formalizing eligibility criteria. By analyzing a large set of breast cancer clinical trials we derived a set of patterns, that capture typical structure of conditions, pertaining to syntax and semantics. We qualitatively analyzed their expressivity and evaluated coverage using regular expressions, running experiments on a few thousands of clinical trials also related to other diseases. Based on an early evaluation we conclude that derived patterns cover the language of eligibility criteria to a large extent and may serve as a semi-formal representation. We expect that extending the presented method for pattern recognition with recognition of ontology concepts will facilitate generating computable queries and automated reasoning for various applications.

CLASSIFICATION AND EXAMPLES OF PATTERNS

Based on our experiments, we identified that the following properties would improve the automatic reasoning capabilities with trial eligibility criteria:

- **Time independent status**: present, absent, conditional, potential, not selective
- **Temporal status**: historical, current, planned
- **Specification type**: time frame, including and excluding findings or therapies, value restrictions, purpose of a drug/treatment, co-occurrences, number of occurrences, confirmation, outcome constraint.
- **Medical content**: demographic data (age, gender), clinical data (pregnancy and nursing, menopausal status, adverse reactions), pathology and molecular data, interventions (prior and current therapies).
- **Data source of medical content**. This dimension is dependent on concrete EHR, it is meant to support automatic information extraction.
- **Variability and controllability**: stable, variable, controllable, subjective.
- **Subject:** candidate, family of a candidate

Dimension	%	Example of a pattern	Example of an instance
Time independent status	58		
Absent	16	No concurrent ()	No concurrent endocrine therapy.
Conditional	13	() allowed if ()	Multicentric breast tumors are allowed if all foci are ER-negative.
Temporal status			
Historical	26	No history of ()	No history of brain metastases
Current		Allergy to ()	Allergy to bisphosphonates.
Specification type	72		
Time frame	15	At least () since prior ()	At least 3 weeks since prior steroids.
Exclusions	5	No prior () except for ()	No prior malignancy, except for adequately treated basal cell.
Confirmation	8	confirmed by ()	No metastasis to brain (confirmed by CT or MRI)
Medical content	45		
Clinical	8	No pregnant	Negative pregnancy test
Pathology	7	T () stage	T1-3
Variability and controllability			
Stable	26	Gene () mutation	Known carrier of BRCA1 or BRCA2 mutation
Controllable	1	Must use contraception.	Patients must use effective non-hormonal contraception.
Subject			
Family of a candidate	5	family history of ()	Family history of colon cancer.

PATTERNS RECOGNITION

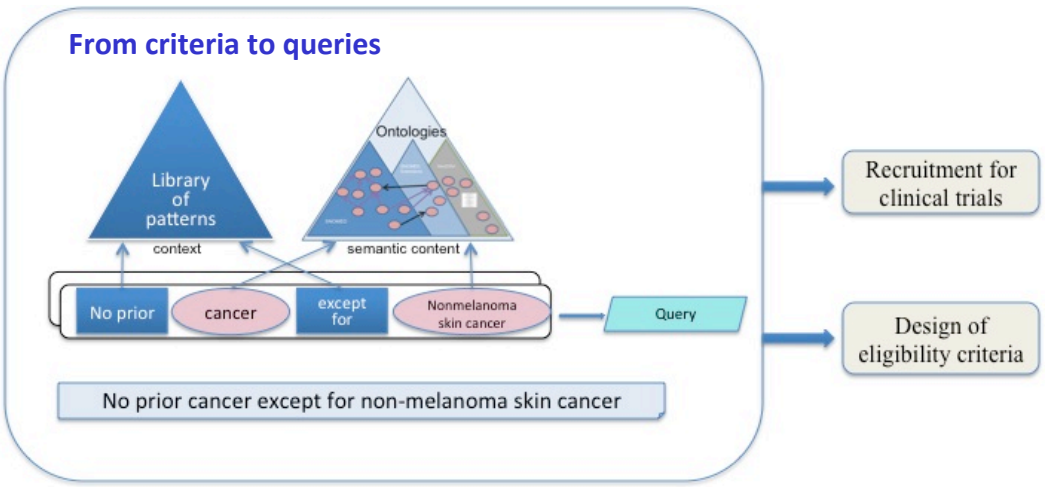
In order to evaluate the coverage of the defined set of patterns across medical domains we analyzed eligibility criteria from breast cancer, lung cancer and diabetes, published at ClinicalTrials.gov. We calculated a number of occurrences of each pattern in the set of eligibility criteria using 342 regular expressions.

	Breast cancer	Lung cancer	Diabetes
No. Trials	3905	2949	5499
No. of sentences processed	111334	119547	86526
Sentences with at least 1 identified patter	71 %	69 %	54%

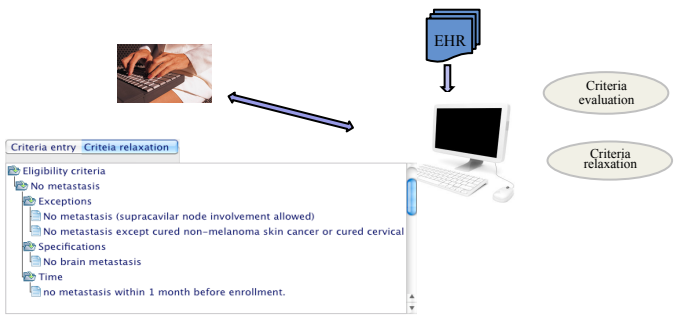
APPLICATIONS

We will use the patterns in the process of formalizing eligibility criteria. Based on annotations of criteria content with the patterns and ontology concepts, we can start generating computable queries.

From criteria to queries



Supporting design of eligibility criteria



SUMMARY

- ✓ We have investigated the possibility of capturing and formalizing the jargon of clinical trial eligibility criteria.
- ✓ We approached the problem by defining a set of 130 patterns that differ in the complexity level. We used 342 regular expressions to identify the patterns in eligibility criteria from breast cancer, lung cancer and diabetes clinical trials and were able to find at least one pattern in 71%, 69% and 54% of lines, respectively.
- ✓ We obtained a method for automatic classification of eligibility criteria according to fine-grained dimensions.

CONCLUSION

- ✓ Our findings indicate that the language used for expressing eligibility criteria is regular enough to be captured to a big extent by the set of defined patterns.

Processing large graphs in parallel

E. Krepska, T. Kielmann, W. Fokkink, H.Bal

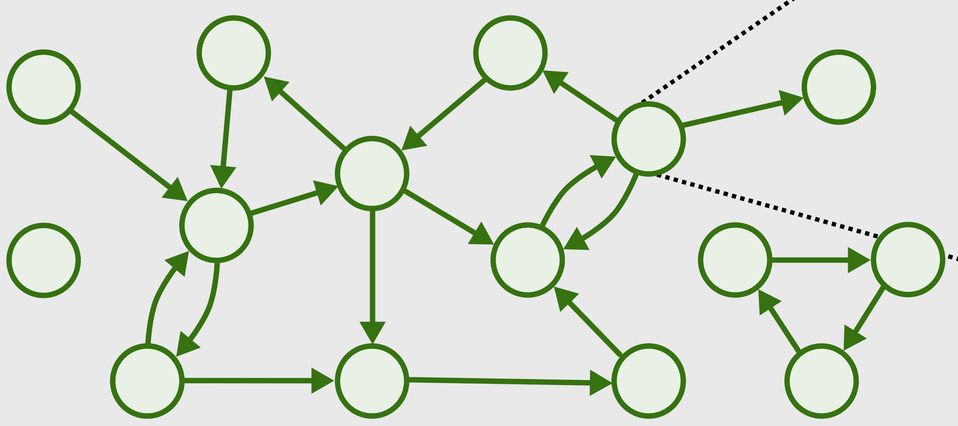
VU University Amsterdam, Netherlands

HipG

A framework to easily write distributed graph algorithms

- Push-button automatic parallelization
- Easy programming (exposed vertex/edge)
- Structure-driven fine-grained computations
- Handles billions of vertices and edges
- Efficient w.r.t. memory and computation

User implements graph vertices:



MyVertex

```
// custom data
boolean visited;

// custom methods
visit();
```

Visits all reachable vertices

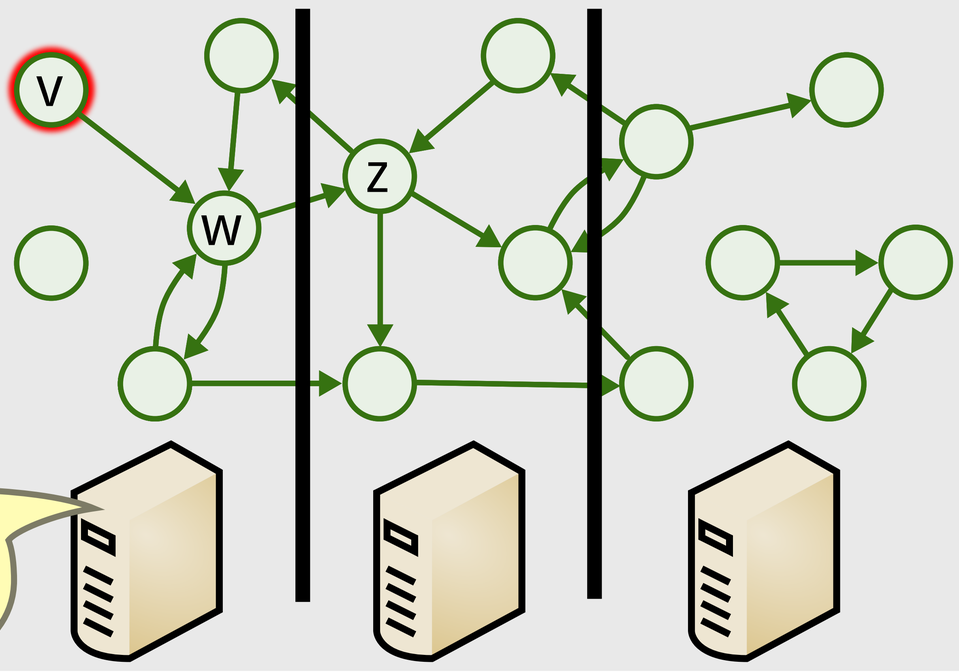
Java

```
public void visit() {
  if (!visited) {
    visited = true;
    while(hasNeighbor(i))
      neighbor(i++).visit();
  }
}
```

Gives work to other vertices

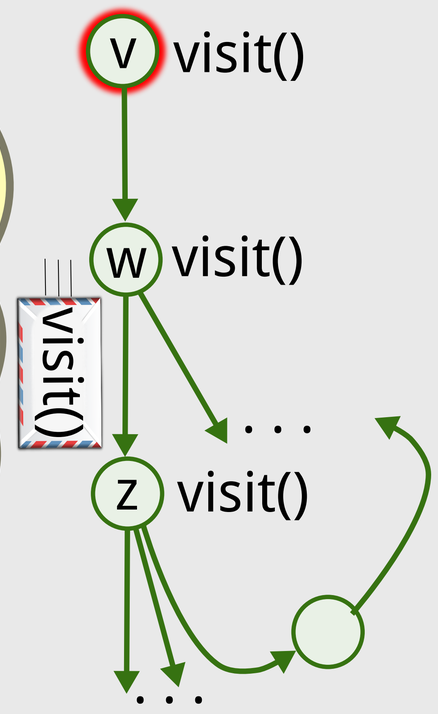
This is sequential code!

Automatic parallelization for a distributed computer:



Vertices distributed between memories of many machines

Methods on non-local vertices translated into asynchronous messages

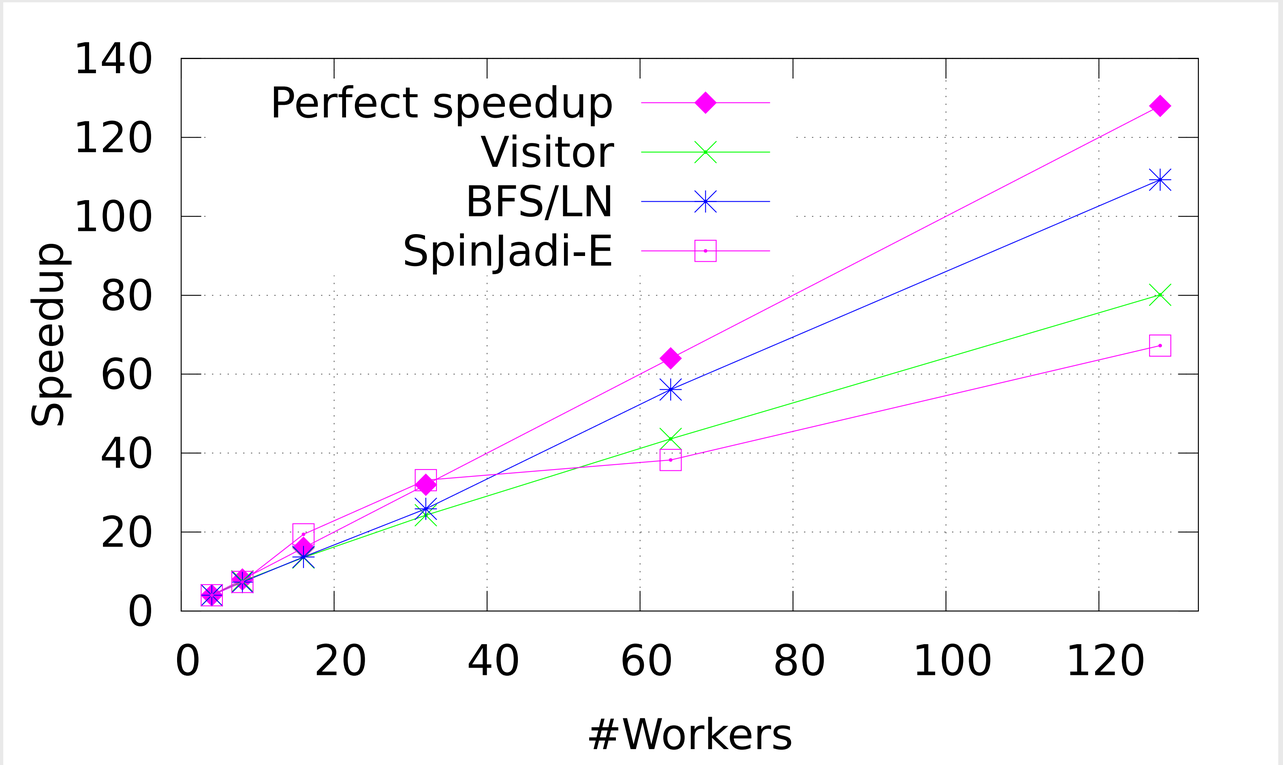


Termination detection:

```
v.visit();
barrier();
```

Results & examples

- Breadth-first search (BFS)
- Model checking (SpinJadi)
- Genetic networks



Speedup of selected HipG applications run using 4-128 workers (2 per machine) on the DAS-4/VU cluster, on graphs with up to 10^{10} of vertices and edges. On 64 machines obtained efficiency of 60-80%.

More features:

- Global operations
- Divide-and-conquer graph computations
- On-the-fly computations (concurrent with graph generation)

Talk to us!

We're looking for new applications

Contact: **e.krepska@vu.nl** Download/papers: **http://www.cs.vu.nl/~ekr/HipG**

Seeking collaboration

(w/ comp. linguistics, computer science)

Ivar Vermeulen, Anika Batenburg
Communication Science

1. Automated generation of word (fragment) based implicit measurement instruments (Ivar)

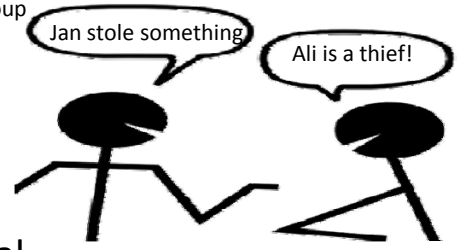
=> Implicit measurement: To measure what people **really** think **right now** (alt. for fMRI, questionnaires)
=> Usually word (fragment) based: people recognize words that are "on their mind" faster

E.g. word fragment completion	*E.g. word or no word?*	*E.g. word finder puzzles*
D E A _	HOSPITAL	
G R A _ E	NURSE	
C O F F _ _	AMBUANCE	
……		

Idea: automated generation of associated words or fragments using a thesaurus-based approach.
Result: tool, open access to researchers.

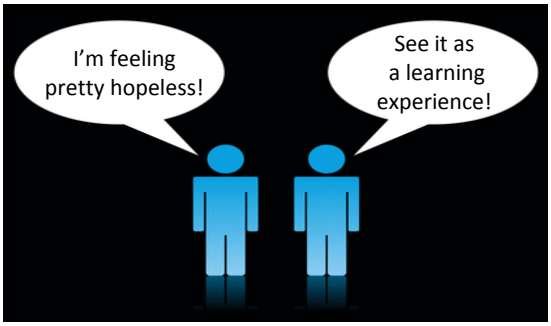
2. Identifying Linguistic Intergroup Bias online (Ivar)

=> **Positive** ingroup behaviors described more *abstractly*, positive outgroup behaviors decribed more *concretely*
=> **Negative** ingroup behaviors described more *concretely*, negative outgroup behaviors decribed more *abstractly*



Idea: combining Web retrieval techniques, semantic parsing, and linguistic categorizations using LIB's categories

3. Identifying instrumental and less instrumental writing styles in online support groups (Anika)



=> Cognitive reappraisal vs. emotional disclosure
=> Distinguishing initiations (sharing of own experiences) and response (attempts to help others)

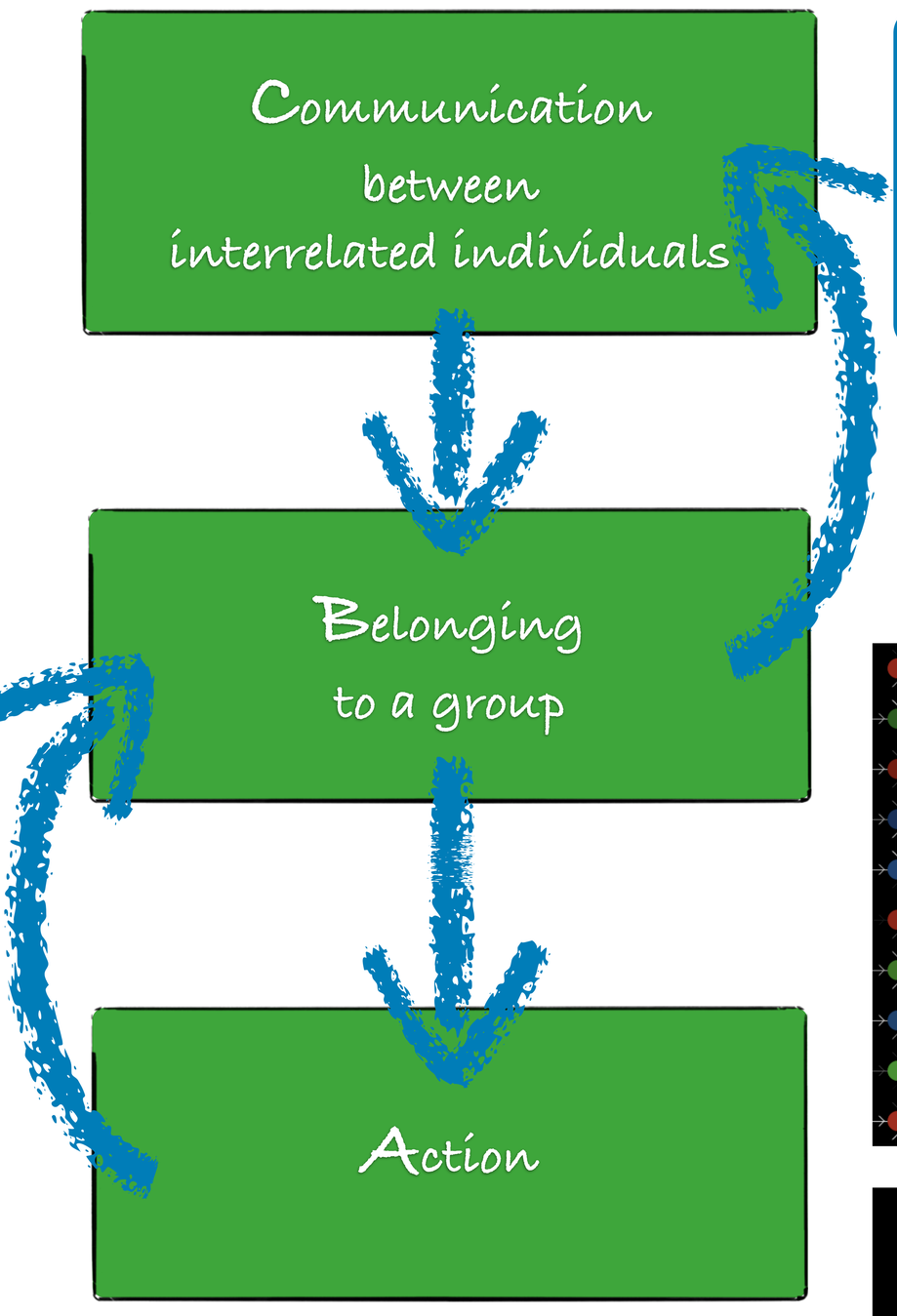
Idea: combining Web retrieval techniques, semantic parsing, thread parsing, and linguistic categorizations using Pennebaker's LIWC (or other methods?)

ie.vermeulen@vu.nl
ae.batenburg@vu.nl

Simulating Social Influence in Groups

Samantha Kneefel, Borre Mosch, Evert Haasdijk, Dick de Gilder, Tibor Bosse

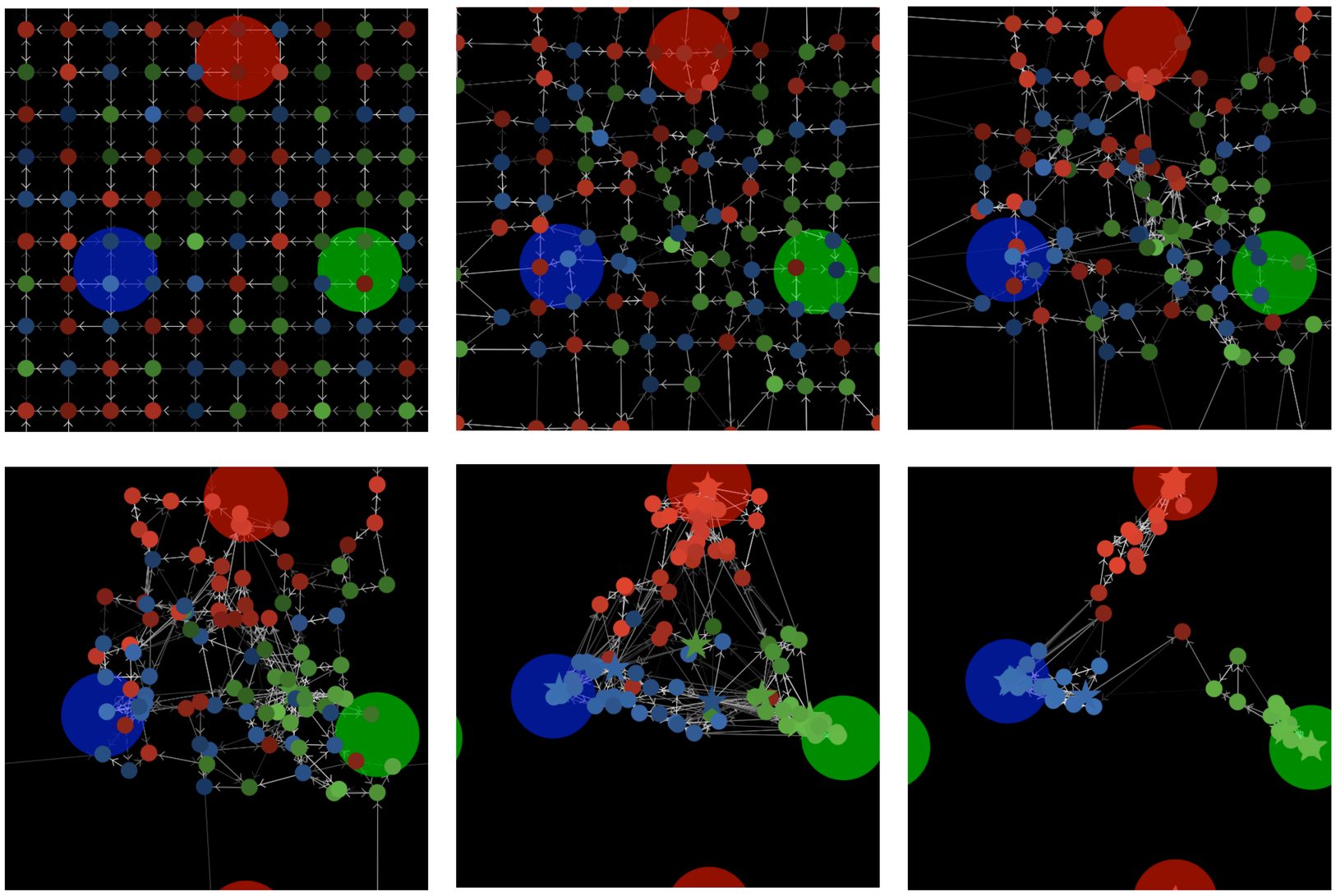
Model



Research Question

How do group identities develop, based on the dynamic interplay between network structure, inter-/intragroup interaction and group actions?

Simulation



Next Steps

- Validation & parameter estimation of the model
- Numerical analysis of simulation
- Simulate specific scenarios
- Towards an algorithm for emergent robot team composition

Part of the Network Institute's Academy Assistant programme

VU UNIVERSITY AMSTERDAM

T2PP

From Text to Political Positions

From Sentiments and Opinions to Political Party Positions

Annemarie van Elfrinkhof (Social Sciences), Bertie Kaal (Fac. of Arts), Isa Maks (Fac. of Arts)

T2PP combines contemporary theories and methods in linguistics with political science to develop an automated research tool for rich text-mining and opinion-mining. The transdisciplinary relevance of the project is that it will merge insights from political science with critical discourse analysis, cognitive processes and lexico-semantic-, computational- and corpus linguistics.

The convergence of Political , Linguistic, and Discourse Analysis

We apply three methods to position political parties on a Left/Right dimension and a Progressive/Conversative dimension, using Dutch election manifestos of 2010 as data. The aim of the study is to point out methodological constraints to party positioning and to show how a combination of methods can lead to a refinement of the individual methods.

Goldstandard Kieskompas 2010

Method I Words as Data (WD)

Wordscores is a quantitative text-analysis method developed by Laver, Benoit and Garry (2003). With regard to election manifestos, the underlying logic is that political parties do not use the same words with the same frequency.

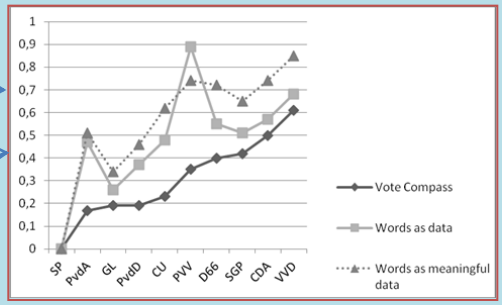
The relative frequency of words can be used to differentiate between parties' stances. The underlying logic is that political parties do not use the same words with the same frequency.

Ex. Political parties on the right propose "to cut taxes" in their election manifesto. Left-wing parties do not use the words "to raise taxes", although their plans might imply it. The word "tax" is therefore considered indicative of right-wing parties.

Method II Words as meaningful data (MWD)

This method combines Wordscores with a lexical semantic approach that selects deontic expressions in the election manifestos. Deontic expressions are relatively frequent in election manifestos and important in that they refer to a future ideal state and to some kind of actions toward that ideal.

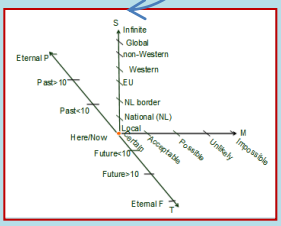
Ex. "*The Netherlands must be prepared to give some kind of humanitarian aid.*"



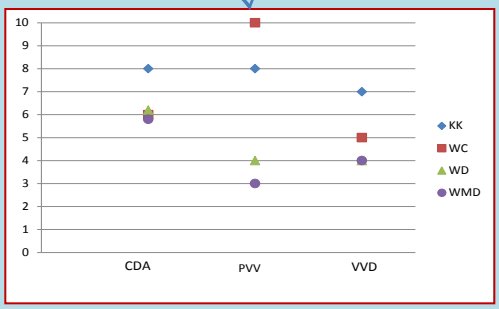
Results on Left/Right: Method II approximates the gold standard most

Method III Words in Context (WC)

Words in context takes a discourse approach to analyze the construction of stance taking particular to the motivation in party programs. The purpose is to differentiate between ground rationales of parties' worldviews on three basic parameters: Time, Space and Modality. We assume that Time and Space construct cognitive frames of reference that are consistent with ideologically motivated programs. Ex. "*Holland is full!*"



Time-Space-Modality axes



Results on Progressive/Conservative: Method III approximates the gold standard most

Conclusions

The three methods operate on different levels: word frames, deontic-semantic frames and cognitive-discursive frames.
We find that Method II gives better results on Left/Right than Method I and that Method III gives better results on Progressive/Conservative. Methods II and III complement each other for party positioning on the two political dimensions. They add a deeper level for opinion mining and take into consideration both qualitative and quantitative aspects of text and meaning.

Van Elfrinkhof, A., E. Maks and B. Kaal. Forthcoming. From text to political positions: The convergence of political, linguistic and discourse analysis. In Kaal, B., E. Maks and A. van Elfrinkhof (Eds) , *From Text to Political Positions: Converging approaches to estimating party positions.* Submitted for publication.

More information: http://www2.let.vu.nl/oz/cltl/t2pp/

D. Ceolin, P. Groth, W. R. van Hage, A. Nottamkandath and W. J. Fokkink

Trust Evaluation through User Reputation and Provenance Analysis

Would you drink this coffee?

It's made by a pro!

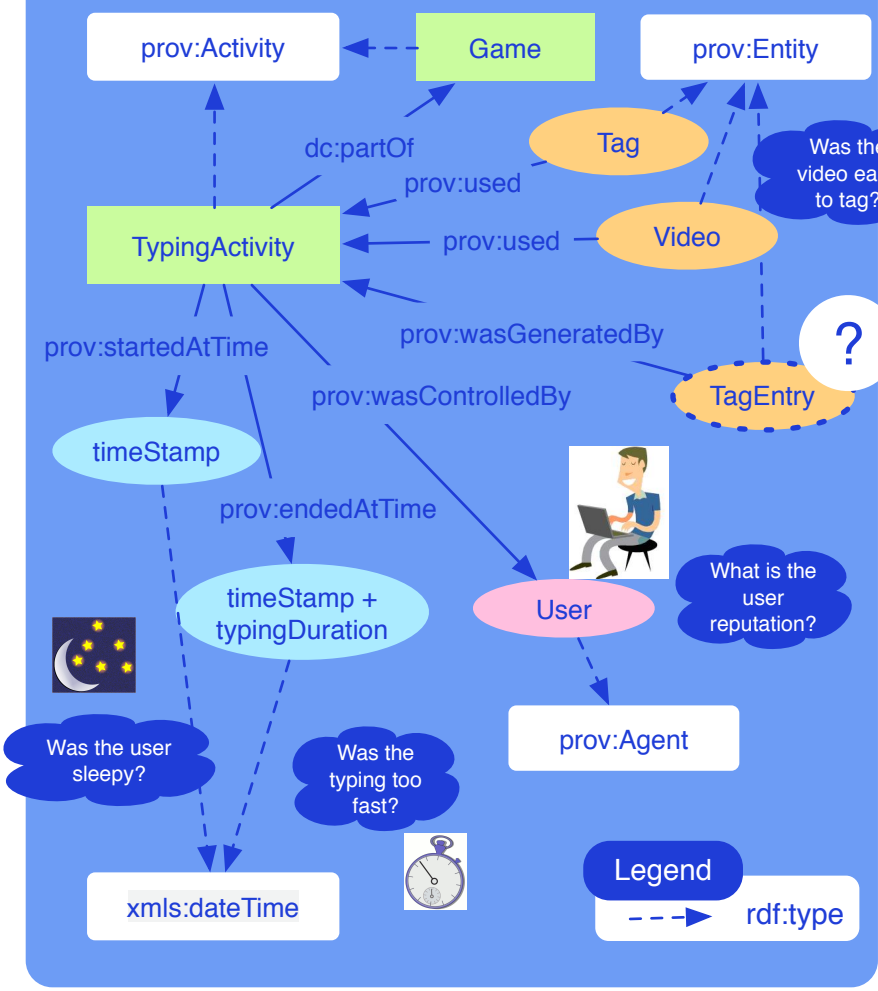
PS: take a look at the coffee maker...

Trust

Who ≤ Who + How
Reputation ≤ Reputation + Provenance

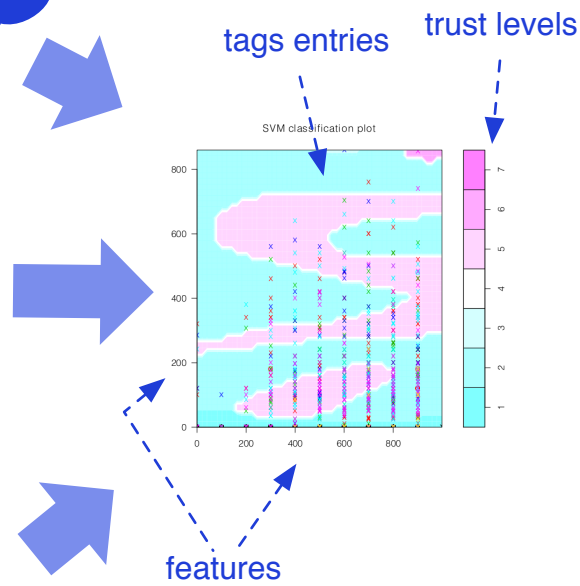
A video labeling game example

Waisda?



What features shall we use? How to evaluate them?

We let a classification algorithm decide!



vrije Universiteit amsterdam

The Network Institute

The Value of Deliberate Metaphor

Gudrun Reijnierse, MA

VU UNIVERSITY AMSTERDAM | Faculty of Arts

The Project

Metaphors are all around us in our daily lives: we *defend* a thesis, we live *in* 2012, and Dutch politician Geert Wilders speaks of a *tsunami* of Islamization.

Most people will recognize the last expression as metaphorical but not the first two. People thus do not always seem to be aware of the presence of metaphors in language use.

This project aims to find out what the difference is between deliberate (i.e. rhetorical) and non-deliberate (i.e. unconscious, automatic) metaphor.

It is the goal of this project to rehabilitate rhetorical and communication-oriented approaches to metaphor.

What is Deliberate Metaphor ?

Deliberate metaphor is the use of metaphor *as* metaphor in order to alter the addressee's perspective on a current discourse referent or topic.

(cf. Steen 2008, 2010, 2011)

The Value of Deliberate Metaphor: Phases of Analysis

MIPVU	Deliberate vs. Non-deliberate	5-Step method	Functional analysis	Quantitative analysis	Experiment
Linguistic forms	Distinction deliberate / non-deliberate	Conceptual structures	Communicative values	Distribution	Processing

How to Distinguish between Deliberate and Non-Deliberate Metaphor ?

Goal: Develop a formal, systematic procedure to identify deliberate metaphor in discourse

Characteristics of deliberate metaphor

- Presence of a signal (like, as, metaphorically speaking, "...")
- Novel metaphor (contextual sense not in dictionary)
- Direct metaphor (e.g. similes)
- Word play
- Clustering of metaphor

An Example of Deliberate Metaphor

After all, Mancunians and visitors to the Manchester conurbation are going to have to look at these mechanical millipedes for well into the twenty-first century

(A3M-fragment02)

References

Goatly, A. (1997). *The language of metaphors*. London: Routledge.
Goddard, C. (2004). The ethnopragmatics and semantics of "active metaphors." Journal of Pragmatics, 36(7), 1211-1230.
Krennmayr, T. (2011). *Metaphor in newspapers*. Utrecht: LOT.
Steen, G. (2008). The paradox of metaphor: Why we need a three-dimensional model of metaphor. Metaphor and Symbol, 23(4), 213-241.
Steen, G. (2010). When is metaphor deliberate? In N.-L. Johannesson, C. Alm-Arvius, & D. Minugh (Eds.), Selected Papers from the Stockholm 2008 Metaphor Festival (pp. 43-63).
Steen, G. (2011). The contemporary theory of metaphor -- now new and improved! Review of Cognitive Linguistics, 9(1), 26-64.

The Metaphor Lab

The Metaphor Lab is an internationally unique expertise center for Metaphor Studies. Its mission is to stimulate interdisciplinary collaboration, innovation, and application in all areas of metaphor research inside and outside academic communities.

The Metaphor Lab performs fundamental research on metaphor in language, cognition, and communication is combined with applied research on metaphor in varied domains of discourse such as media use, organization, management, health etc.

The lab specializes in linguistics and discourse studies, psycholinguistics and cognitive science, sociolinguistics and social science, and applied linguistics. It seeks to develop new products and services for the non-academic community, to improve self-awareness, monitoring, training and testing of metaphor use in design, management, communication, and interaction.

The Web of Radios

Introducing African Community Radio as an Interface to the Web of Data

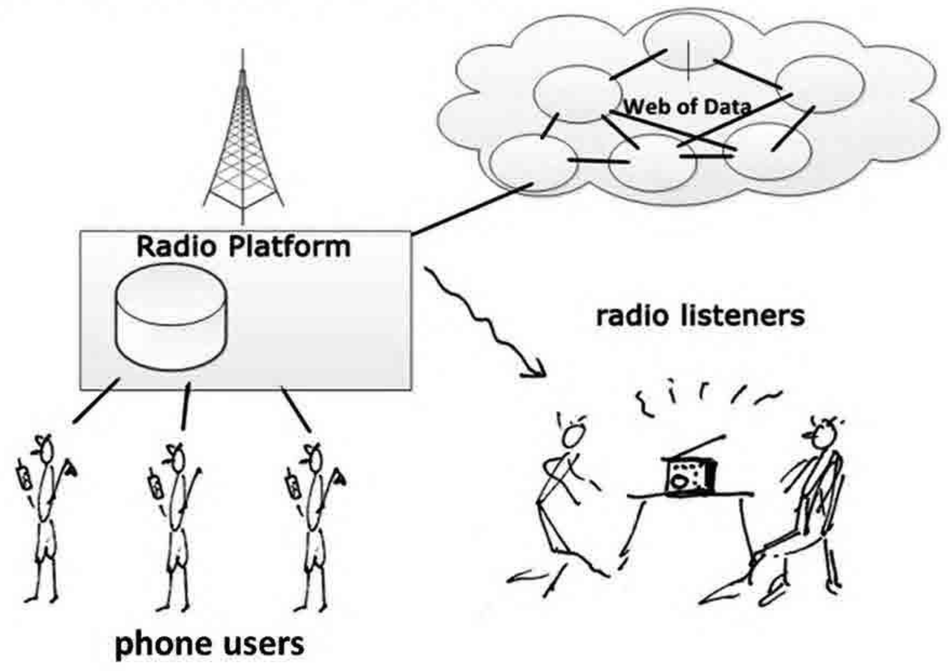
Anna Bon, Victor de Boer, Pieter De Leenheer, Chris van Aart, Nana Baah Gyan,
Max Froumentin, Stéphane Boyera, Mary Allen, Hans Akkermans

Can you access the Web, when you only have a radio and a simple mobile phone ?

Many people in remote rural areas, e.g in Mali, do not have computers or the Internet. Can you benefit from the Web, using only radio or a simple mobile phone? Can you use a voice-interface to the Web, in a local African language?

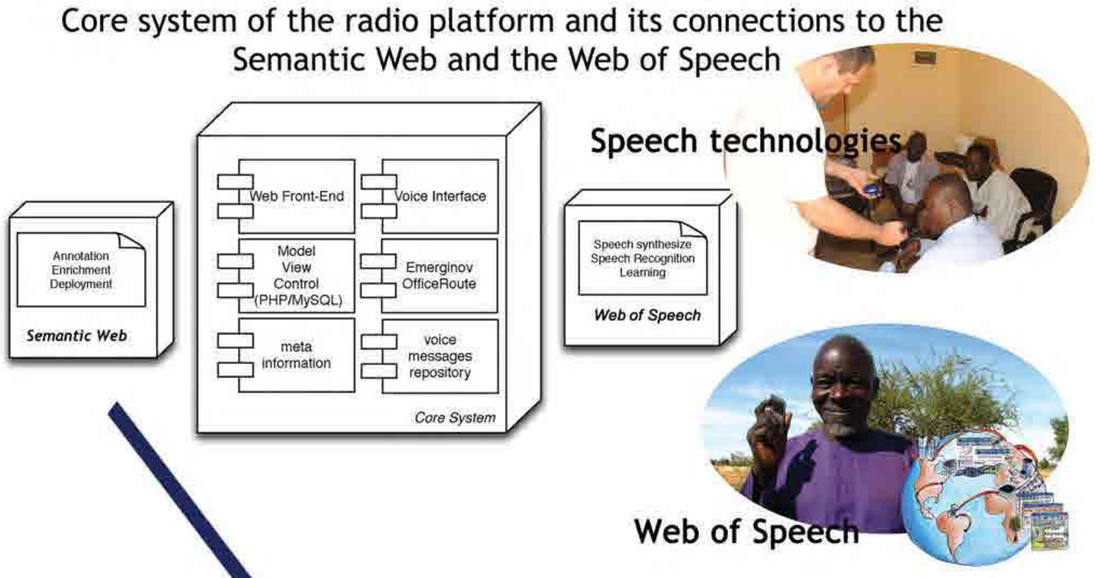
We propose the radio platform Foroba Blon

as a new, voice-based interface to the Web of Data,
for people who are out of reach of computers and the Internet
but do have radio and mobile phones



Architecture of Foroba Blon

Core system of the radio platform and its connections to the Semantic Web and the Web of Speech



Organizing the content

Citizen journalists phone in and leave a message for broadcasting, in **local African languages**. These are stored as audio-files.

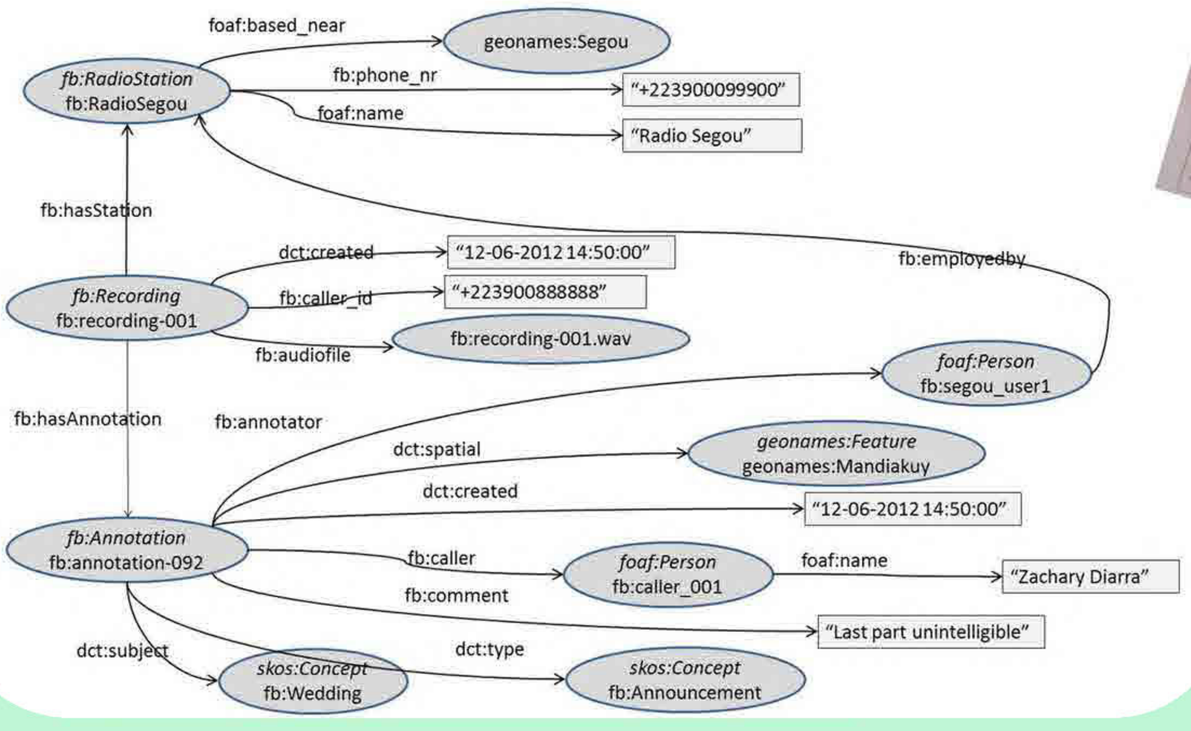
Initially, the data will be used locally. To ensure sharing and reuse of the radio data, a semantic datastore is built.

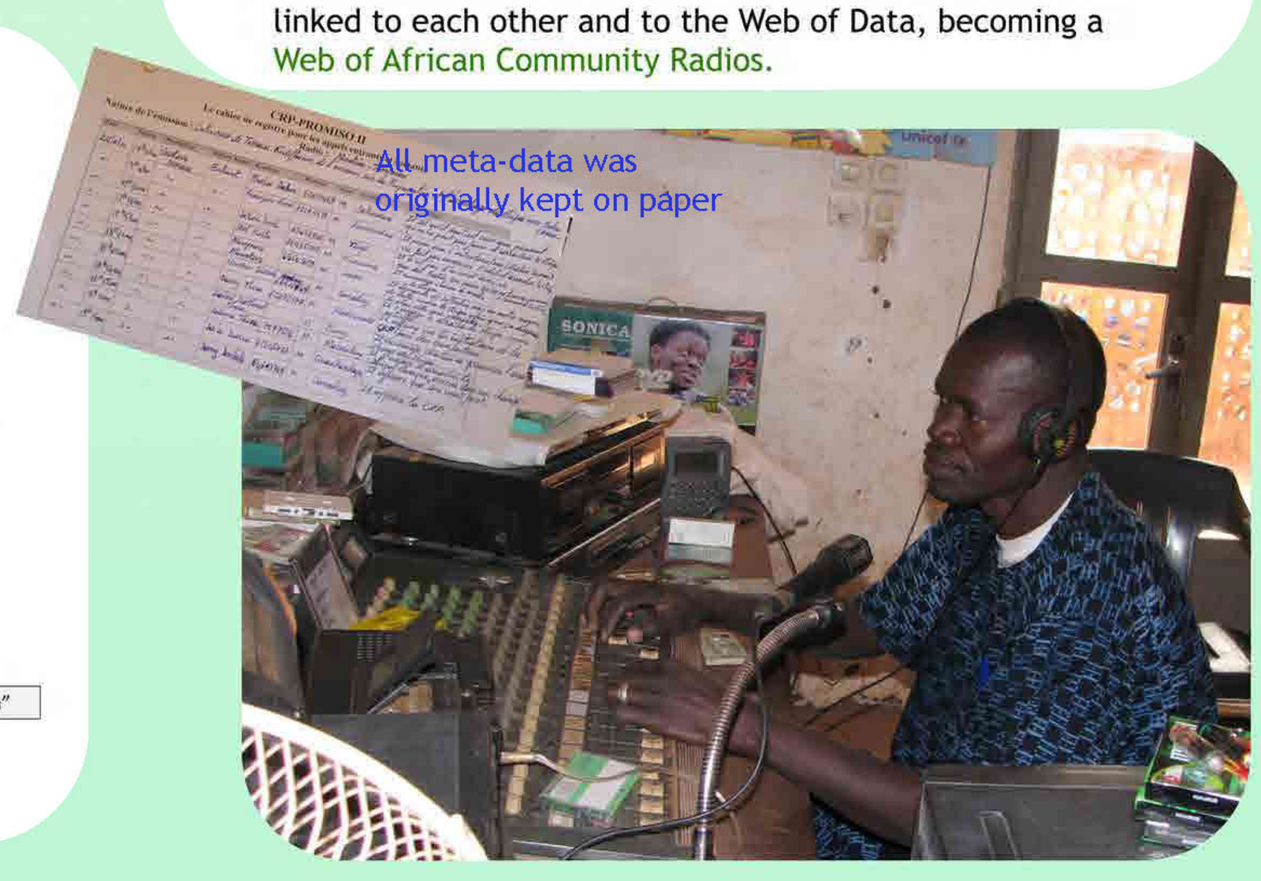
In the future,voice recordings may be shared across radio stations, and with other applications on the **Web**.

By exposing the radio data as **Linked Data**, we open the possibility for expected and future **unexpected re-use**.

Different instances of the Radio Platform can thus be linked to each other and to the Web of Data, becoming a Web of African Community Radios.

Semantic datamodel





This research is partially funded by the European Union through the Seventh Framework Program (FP7)
under agreement number 269954

and by the
International Press Institute
This project was winner in the
IPI News Innovation Contest 2011

Photography: Bruno van Moerkerken