

#### **Extractivism**



## MONA

Interdisciplinary research into activist events

Organizational Sciences Frank de Bakker lina Hellsten Maxine Kruijt Bibiana Armenta

Computer Science Lora Aroyo Thomas Ploeger Chun Fei Lung

**Computational Linguistics** Antske Fokkens

GREENFEACE



Activists shape the future by influencing opinions and policies

EEK JE UIT

#### **HOW?** Read news articles, identify the events activists participate in

Simple Questions Narrative Questions Interpretative Questions

## Problems

Answering these types of questions is difficult for various reasons

Processing and aggregation is time consuming

Individual news articles may lack context and present an incomplete and biased story



#### Automatically extract event mentions from multiple news sources

Store them in a structured format to enable aggregation

Present statistics and visual overviews such as timelines

#### Research Goals

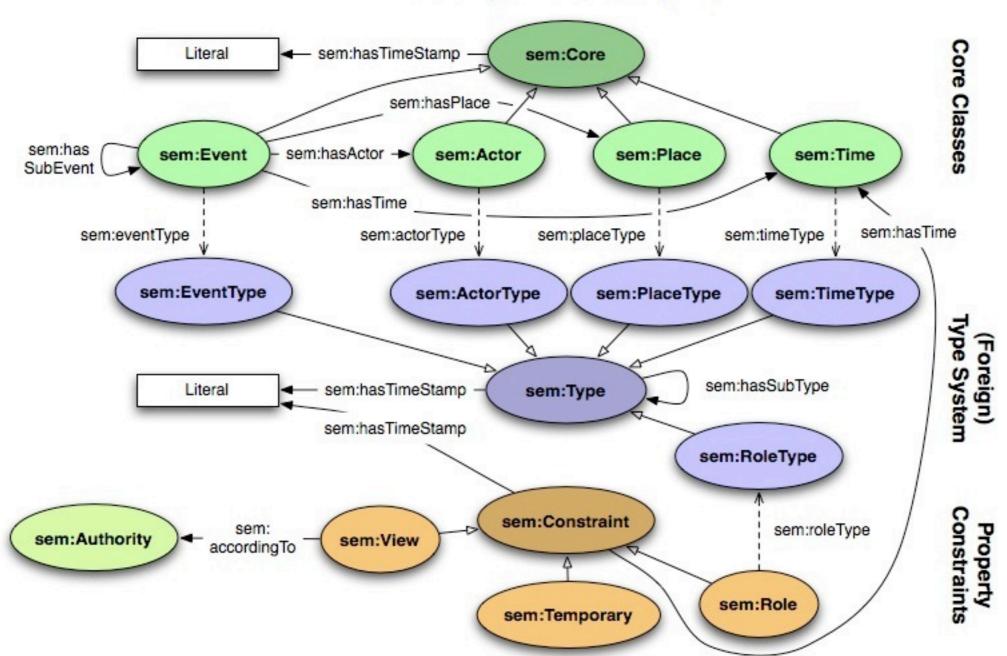
Explore semantic models for storing and interpreting events

Build (semi)-automated tools to support question answering

NLP Tools for extraction

Visual Analytics for browsing

#### **SEM** An ontology to describe events in terms of "Who did what, where and when?" (Van Hage 2011)



the Simple Event Model (SEM)

#### **Extraction** Are there any existing methods we can re-use?

Gene and protein interactions from molecular biology papers

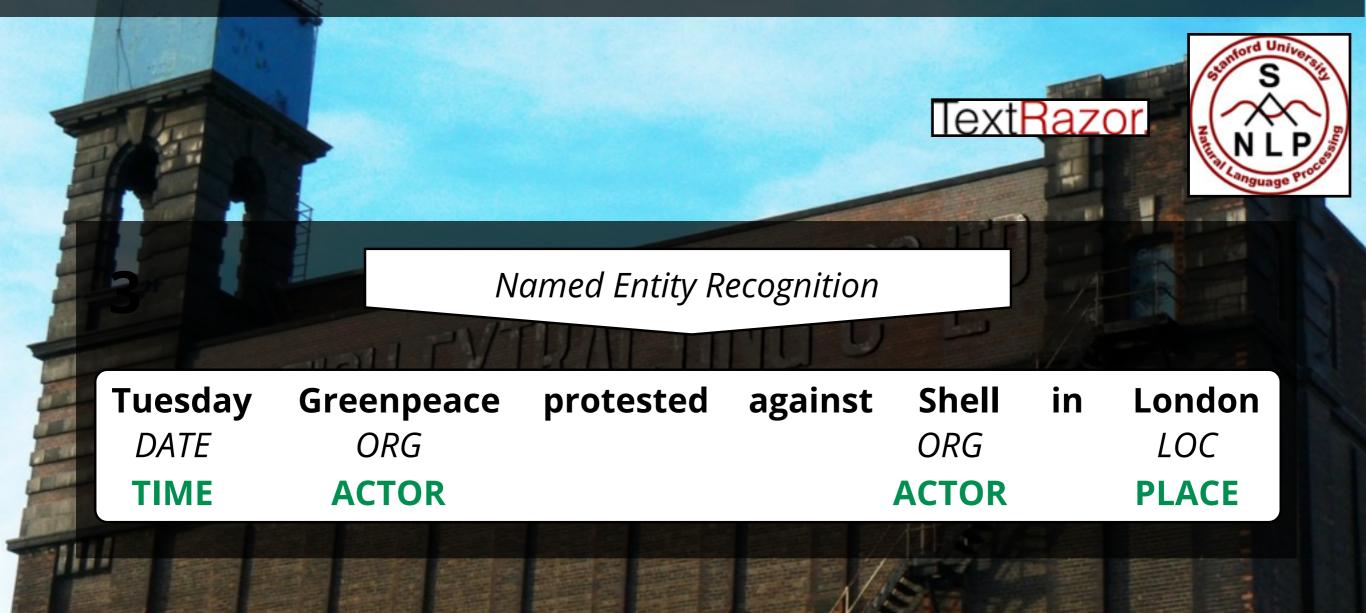
NI OTI

Illegal migration events from news articles for border security

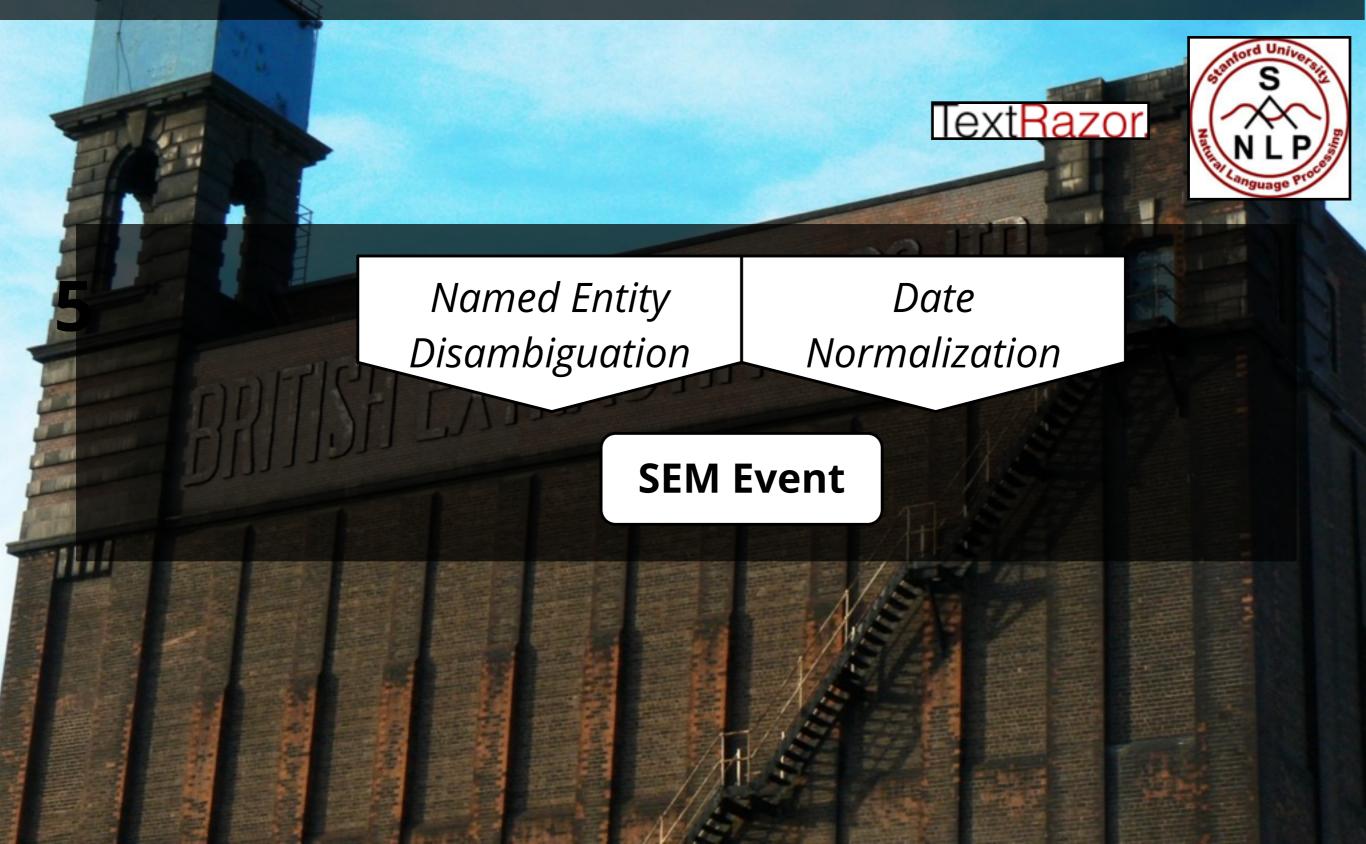
Football events for creating match summaries





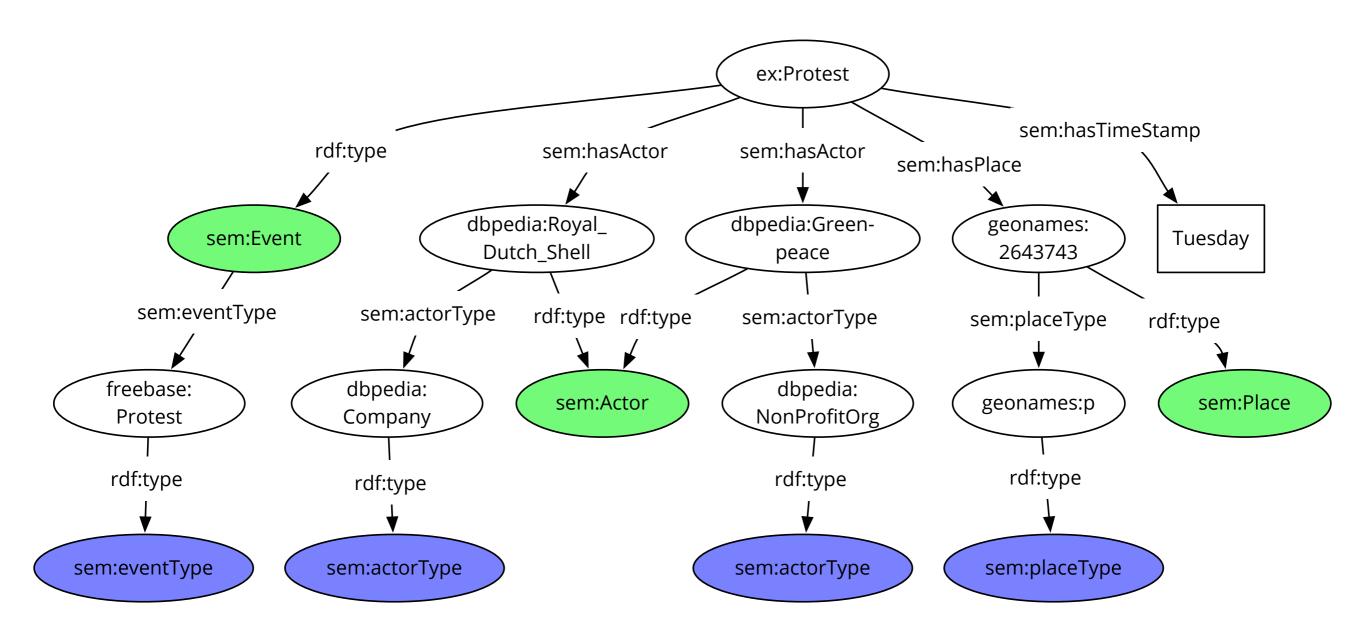






# **Extraction** Resulting SEM Event

"Tuesday Greenpeace protested against Shell in London."



# **Evaluation** Experimental Setup

## 45 articles concerning arctic oil exploration

#### A variety of sources, including NYT, Guardian, and Amnesty

#### Annotated by domain experts to create a gold standard

## **Evaluation** Inter-rater agreement

	anti es	All and the second seco	1 95 (Comments) 20 E
Event	Actor	Place	Time
46%	35%	36%	28%
34%	32%	28%	43%
20%	33%	36%	20%
A DE		a Ville	
	46% 34% 20%	46%       35%         34%       32%         20%       33%	46%       35%       36%         34%       32%       28%         20%       33%       36%

#### Results Performance Measurements

#### We extracted 1829 events, 1609 actors, 772 places, and 456 times

 Event
 Actor
 Place
 Time

 Precision
 71.00%
 46.00%
 36.00%
 65.00%

 Recall
 58.00%
 35.00%
 54.00%
 60.0%

#### **ETTORS** Low Place Precision Low Actor Recall

#### "Shell is working with wary United States regulators."

MBONG

"The **European Union** has long shown leadership on the issue of climate change"

#### Errors Low Actor Recall

"The **Kulluk Oil Rig** was used for test drilling last summer." "The Coast Guard flew over **the rig** for a visual inspection."

"Scientists say the drilling may influence..."

#### **Errors** Gold Standard Annotation Instructions

#### Task was too loosely defined, resulting in different interpretations

Annotator	Event F1	Actor F1	Place F1	Time F1
1	63%	41%	49%	57%
2	54%	44%	19%	35%
3	60%	43%	48%	61%

### Visual Analytics

Greenpes By numb Coople Coople Tems of Use Tems of Use Coople Coo	actors with I
Image: state strate strate     Image: strate strate strate     Image: strate strate strate strate       Map     Timeline     Currently	
Coogle Terms of Use 2010-01-01	
Show large map     Only show       Map     Timeline	2013-02-0
	v entities with time
Groon	active filters
Events (73) Actors (188) Places (62) Times (18) Articles (18) About Greenpeace	68
Actors which are mentioned together with Greenpeace Greenpeace Greenpeace	actors with I -
Royal Dutch Shel (3)       Petroleum (4)       Gazprom (4)       Earth (4)       Tuberculosis (3)       Alaska (3)       Apple Inc. (3)       North Carolina (3)       Asia (2)       Apple Store (2)         International Fund for Animal Weitare (2)       Chinadialogue.net (2)       Kumi Naidoo (2)       Oceanography (2)       Advocacy group (2)       European Union (2)       Russia (2)       Popularity	ith I Greenpeace
Conservation (principal) internet (p) internet (p) internet (p) operating (p) internet (p) internet (p)	anizations, and places
United States (2) Offshore drilling (2) Vancouver (2) Gabe Waniewski (1) Think tank (1) Golden Agri-Resources (1) Phil Radford (1) No colors International Climate Politics (1) Public relations (1) Gazprom Nett Shelf (1) Reuters (1) BP (1) San Francisco (1) Mr. Kaiser (1)	
Premational Camate Motios (1) Municipalitions (1) Estadorem Net Specific Review (1) Review (1) Markaway (1)	

#### bit.ly/15We6q0

## Visual Analytics

🛆 номе				Green	npeace					Search
When did I actors wi	th I Greenpo	Greenpeace occur? 1x Number of events in period 7x								Filter 188   actors with   -
	2010 /	ipril July	October	2011 April	July	October	2012	April	July	Greenpeace
Greenpeace (76)			•					• •••	•••	
Royal Dutch Shell (6)									•	By number of mentions (popularity)
Petroleum (4)			•					-	•	
Gazprom (4) Earth (4)								•		
Kumi Naidoo (2)									· ·	1 76
Russia (2)										
Internet (2)								•		<ul> <li>Select timeframe</li> </ul>
Rainforest Action Ne (2)	•								-	2010-01-01 2013-02-
Cfishore drilling (2)										2010 01 01
Golden Agri-Resources (1										-
BP (1)			•							Only show entities with time
	'		_							Currently active filters
			Мар	Time	line					Greenpeace
🚰 Events (73)	Actors (188)	<b>?</b> Places	(62) 🔢 Ti	mes (18)	Articles	(18) Ab	out 🕴 🛛	reenpea	ice	
1										Color 188   actors with   Greenpeace
Actors which are me	ntioned toge	ther with I G	eenpeace							
Royal Dutch Shell (9)	Petroleum (4)	Gazprom (4) Ear	th (4) Tubercul	osis (3) Alaska	(3) Apple In	c. (3) North Ca	arolina (3)	Asia (2)	Apple Store (2)	Popularity with   Greenpeace
International Fund for A	nimal Welfare (2)	Chinadialogue.ne	t (2) Kurni Naid	oo (2) Oceano	graphy (2) A	kdvocacy group (	2) Europ	ean Union (2	Russia (2)	People, organizations, and places
Conservation (ethic) (2)	Internet (2)	Renewable energy	(2) Rainforest	Action Network (2	Sylvia Earl	e (2) Scientist	(2) Worl	d Wide Fund	for Nature (2)	No colors
United States	(2) Offshore d	iling (2) Vancouv	er (2) Gabe We	sniewski (1) Thi	ink tank (1)	Golden Agri-Res	ources (1)	Phi Radfo	rd (1)	
International	Climate Politics	(1) Public relation	s (1) Gazprom	Neft Shelf (1)	Reuters (1) E	IP (1) San Fra	ncisco (1)	Mr. Kaiser	(1)	Legend: number of mentions
Greenpeace activists b	part Kremin (1)	Marine conservati	on (1) Antarctic	Ocean Allance (	(1) Priraziom	nava (1) Gary	Cook (1)	Twitter Cas	son Trenor (1)	1 4 6

#### bit.ly/15We6q0

#### Lessons Learned

Off-the-shelf NLP tools yield useful, but imperfect results with little domain-specific tuning

Minimally constrained notion of event makes it difficult to instruct annotators

...and to fine-tune extraction process.

#### Future Work

#### FUTURE CITY

Implement and evaluate solutions to known errors, re-run evaluation

New gold standard with more rigorous annotator instructions and majority voting

Methods for identifying complementary events

#### Recap

Study activists, they shape the future by influencing opinions and policies

Used a pipeline of off-the-shelf NLP tools to extract events

Extracted events used in visual analytics to answer questions

Loose event definition is challenging to annotate and fine-tune

#### **Photo Attributions**

- 1 http://www.flickr.com/photos/greenpeacesuomi/
- 2 http://www.flickr.com/photos/greenpeace\_switzerland/
- 3 <u>http://www.flickr.com/photos/greenpeace\_nederland/</u>
- 4 http://www.flickr.com/photos/x1klima/
- 5 http://www.flickr.com/photos/oxfam/
- 6 http://www.flickr.com/photos/remonrijper/
- 7 http://www.flickr.com/photos/davepattern/
- 8 http://wrvh.home.xs4all.nl/wrvhage/
- 9 http://www.flickr.com/photos/cactusmelba/
- 10 http://www.flickr.com/photos/cactusmelba/
- 11 Own work
- 12 http://www.flickr.com/photos/liloh
- 13 <u>http://www.flickr.com/photos/liloh</u>
- 14 http://www.flickr.com/photos/dgoodphoto/
- 15 <u>http://www.flickr.com/photos/kaiban/</u>
- 16 http://www.flickr.com/photos/kaiban/
- 17 http://www.flickr.com/photos/kaiban/
- 18 \*
- 19 \*
- 20 http://www.flickr.com/photos/rich\_w/
- 21 http://www.flickr.com/photos/9948354@N08/
- 22 <u>http://www.flickr.com/photos/christinawelsh/</u>
- 23 \*
- 24 http://www.flickr.com/photos/39955793@N07/

## Discussion

#### Tentative example

#### Shell oil rig ran aground in the Gulf of Alaska on the 31st of December 2012

18 days later, United States government expresses concerns about Shell's operational accidents and impact on the environment

A few months later, Shell is granted a permit to begin drilling oil wells in Alaska, despite his earlier concerns