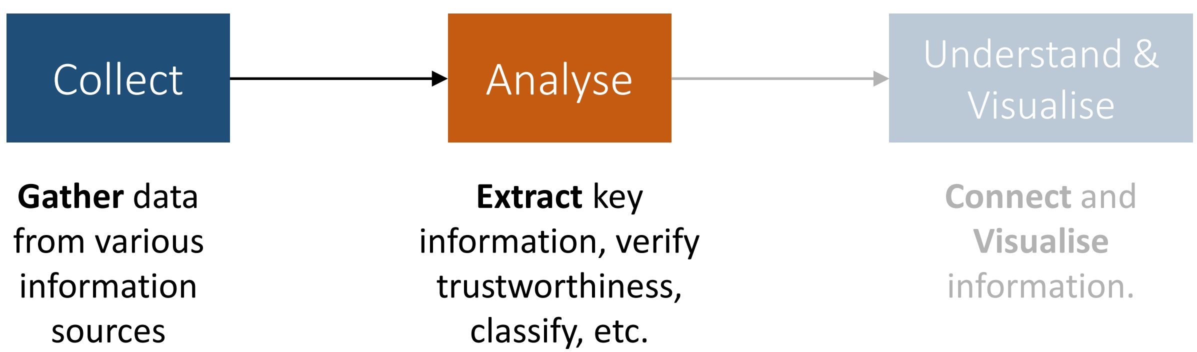
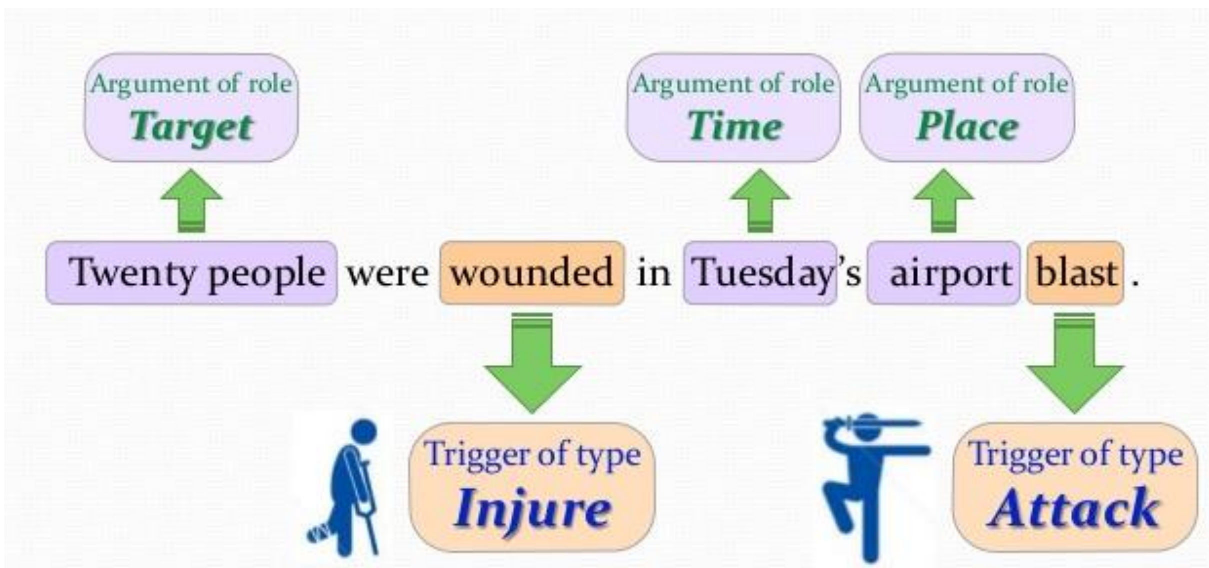
SMASAC - Event Extraction

GRÉGOIRE BUREL, MAYANK KEJRIWAL (PEDRO SZEKELY) AND PRASHANT KHARE



Simple Motivating Example



(More Complex) Motivating Example

<Entities and spatio-temporal info are *italicized*>

Event 1 anchor

> "On *Thursday*, there was a massive *U.S.* **aerial bombardment** in which more than 300 *Tomahawk cruise missiles* rained down on *Baghdad*. *Earlier Saturday*, *Baghdad* was again **targeted**. ..."

Event 2 anchor

Event 1 provides **context** for Event 2

(Yang and Mitchell, 2016)

Crisis Examples from Twitter

RT @jc_stubbs: Another tragic day in #Ukraine – More than 50 rebels killed as new leader unleashes assault: http://t.co/wcfU3kyAFX

RT @Stump4TrumpPAC: These are the 58 people killed in the Las Vegas massacre, the deadliest mass shooting in modern U.S. history.. 💐🙏 http…

RT @dragoner_JP: うげえ……「ラスベガス銃撃の(カナダ人)犠牲者が巨額の医療費に直面」→ Las Vegas shooting victims facing large medical bills https://t.co/1F3bgH17Jv

Natural Disaster Examples from Social Media 'Reporting'

The Associated Press @AP · Apr 18

BREAKING: Island-wide blackout hits Puerto Rico, which suffering unstable power grid following **hurricane**.

450 6.8K 6.0K

David Begnaud @DavidBegnaud · 3h

Puerto Rico's power company says it has restored electricity to all customers affected by an island-wide blackout that was caused by an excavator hitting a transmission line. Ten of thousands of families remain without normal service 7 months after **hurricanes** Irma & Maria (@AP)

10 167 207

AJ+ @ajplus · 9h

Puerto Rico restored power to about 70% of people after an island-wide blackout yesterday.

Parts of the island have been without power since **Hurricane** Maria hit in September 2017, making it the biggest blackout in U.S. history.

Definition: Event Extraction

- Given an *event ontology O*, and a text corpus of documents, **event extraction** is the problem of automatically extracting *instances* ('events') in terms of the *event classes* in *O*
  - Because of the ontology, events are structured representations amenable to querying/analytics
  - What is an 'event ontology'?
  - What do real-world event ontologies 'look' like?
- In many versions of the problem, the ontology is implicit, not known or not required by the extractor
  - But still in the background, conceptually

Ontology Example 1: ACE

- Much flatter than Semantic Web ontologies, which contain rich sets of classes, properties and axioms
- Contains eight event classes, each of which has sub-classes (indicated in parantheses)
  - Life (Be-Born, Marry, Divorce, Injure, Die)
  - Movement (Transport)
  - Transaction (Transfer-Ownership, Transfer-Money)
  - Business(Start-Org, Merge-org, Declare-Bankruptcy, End-Ord)
  - Conflict (Attack, Demonstrate)
  - Contact (Meet, Phone-Write)
  - Personnel (Start-Position, End-Position, Nominate, Elect)
  - Justice (Arrest Release, Fine, Execute, Extradite…)
- Events have four other attributes: modality, polarity, genericity and tense

Ontology Example 2: CAMEO

CAMEO	022
Name	**Appeal for diplomatic cooperation (such as policy support)**
Description	Make an appeal for, request, or suggest expansion of diplomatic ties or co-operation.
Usage Notes	This event form is typically, although not exclusively, a verbal act. It refers to appeals for expanded diplomatic ties and non-tangible support on particular policies. Appeals for more specific forms of diplomacy, such as mediation and negotiation, are coded elsewhere within category 02.
Example	North Korean state media have called on the United States to forge "ties of confidence" with Pyongyang ahead of six-party nuclear talks expected to be
Examp	
Examp	

CAMEO	0241
Name	**Appeal for leadership change**
Description	Make an appeal for, request, or suggest change in leadership or power.
Usage Notes	This event form refers to verbal and non-threatening appeals. More forceful "demands" for leadership change are coded as 1041; demonstrations, protests, etc. demanding change in leadership/power are coded under category 14. Note that even though calls for the target to resign or relinquish power are forms of yielding, they are still coded here. Also code appeals for elections here.
Example	Members of parliament from Kenya's Liberal Democratic Party called on

CAMEO	0234
Name	**Appeal for military protection or peacekeeping**
Description	Make an appeal for, request, or suggest deployment of peacekeepers or other military forces to preserve peace, enforce ceasefires, or protect civilians.
Usage Notes	This event form is typically, although not exclusively, a verbal act. The source actor could be making the appeal for itself or on behalf of another party; the target should represent the actor who is expected to provide the forces.
Example	A group of prominent Liberians have written to President George Bush urging him to send U.S. peacekeeping troops to their capital Monrovia.

Events involve many components (entities, relations…)

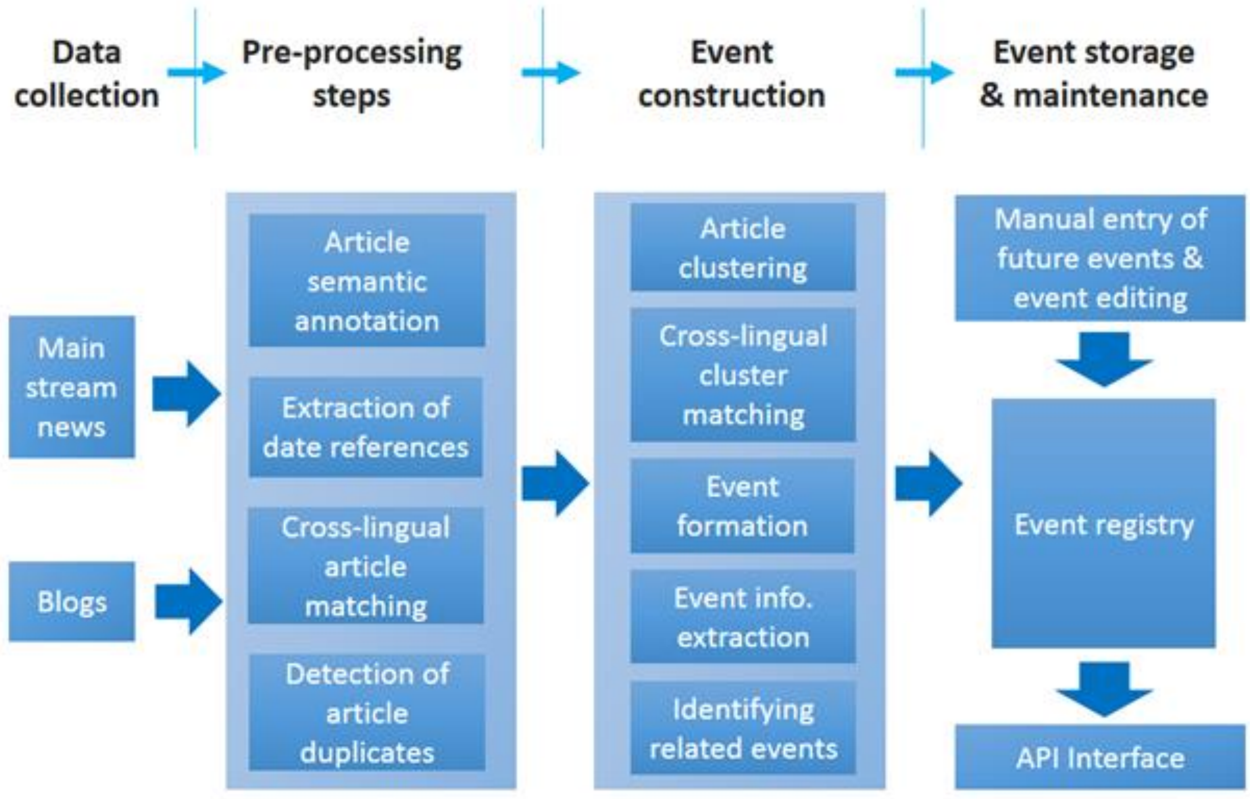
- A much harder problem than just extracting entities!
- Correctly classifying an extracted event with respect to an event type in ontologies like CAMEO (containing hundreds of types) also difficult
- Event co-reference resolution also a hard problem: F-measures well below 50% in state of the art systems

CAMEO	**0234**
Name	**Appeal for military protection or peacekeeping**
Description	Make an appeal for, request, or suggest deployment of peacekeepers or other military forces to preserve peace, enforce ceasefires, or protect civilians.
Usage Notes	This event form is typically, although not exclusively, a verbal act. The source actor could be making the appeal for itself or on behalf of another party; the target should represent the actor who is expected to provide the forces.
Example	A group of prominent Liberians have written to President George Bush urging him to send U.S. peacekeeping troops to their capital Monrovia.

In Social Media…

- Additional problem of first **detecting** whether the tweet is describing an event
  - Can be defined as a binary classification problem: *does the tweet describe an event or not?*
- Short text of tweet makes this essential for performance
- Detection and extraction complement each other
  - Forces us to decide what an event really is, for the purposes of extraction/detection/inference

An Example Workflow



Event Extraction Sub-Problems

- Extracting *entity mentions*
- Extracting *time expressions*, including dates, times of day…
- Extracting *event mentions,* where a single event mention has *exactly one trigger*, and an *arbitrary* number of *arguments*
  - An event trigger (usually a verb or noun) is the *main word* that clearly expresses an *event occurrence*
- Extracting event arguments (roles), defined as *both* the entity mentions involved in an event (e.g., Baghdad) and their *relation* to the event (e.g., *Place*)

Some Tasks Harder Than Others

- Blind test set based on ACE newswire
- Performance on social media much worse
- Even humans don't do too well on role classification

performance / system/human	Trigger classification			Argument classification			Role classification		
	P	R	F	P	R	F	P	R	F
Sentence-level baseline system	67.56	53.54	59.74	46.45	37.15	41.29	41.02	32.81	36.46
Within-event-type rules	63.03	59.90	61.43	48.59	46.16	47.35	43.33	41.16	42.21
Cross-event statistical model	68.71	68.87	**68.79**	50.85	49.72	**50.28**	45.06	44.05	**44.55**
Human annotation1	59.2	59.4	59.3	60.0	69.4	64.4	51.6	59.5	55.3
Human annotation2	69.2	75.0	72.0	62.7	85.4	72.3	54.1	73.7	62.4

Classification of Techniques

- Data-driven event extraction
  - Supervised machine learning models
  - Clustering
  - Inference models
  - Drawbacks?

- Knowledge-driven event extraction
  - Lexico-syntactic patterns
  - Domain expertise
  - Knowledge bases
  - Drawbacks?

Hybrid systems

(Hogenboom et al., 2016)

Event Detection in Twitter: A Taxonomy

- Techniques tend to be classified along several dimensions including:
  - Whether the type of event is specified or unspecified
  - Whether the method is supervised or unsupervised
  - Whether the task is to detect retrospective events (RED) or new events (NED)

References	Type of event		Detection method		Detection task		Application
	Specified	Unspecified	Supervised	Unsupervised	NED	RED	
Sankaranarayanan et al. (2009)		x	x	x	x		Breaking-news detection
Phuvipadawat and Murata (2010)		x		x	x		Breaking-news detection
Petrović et al. (2010)		x		x	x		General (unknown) event detection
Becker et al. (2011a)		x	x	x	x		General (unknown) event detection
Long et al. (2011)		x		x	x		General (unknown) event detection
Weng and Lee (2011)		x		x	x		General (unknown) event detection
Cordeiro (2012)		x		x	x		General (unknown) event detection
Popescu and Pennacchiotti (2010)	x		x		x		Controversial news events about celebrities
Popescu et al. (2011)	x		x		x		Controversial news events about celebrities
Benson et al. (2011)	x		x			x	Musical event detection
Lee and Sumiya (2010)	x			x	x		Geosocial event monitoring
Sakaki et al. (2010)	x		x		x		Natural disaster events monitoring
Becker et al. (2011)	x		x			x	Query-based event retrieval
Massoudi et al. (2011)	x			x		x	Query-based event retrieval
Metzler et al. (2012)	x			x		x	Query-based structured event retrieval
Gu et al. (2011)	x			x		x	Query-based structured event retrieval

(Atefeh and Khreich, 2013)

Feature Representation

- In general, many features are necessary to achieve good performance
- See survey by Atefeh and Khreich (2013) for full list

References	Detection techniques	General features	Twitter-specific features
Sankaranarayanan et al. (2009)	Naive Bayes classifier and online clustering	Term vector	Hashtags and timestamps
Phuvipadawat and Murata (2010)	Online clustering	Term vector, proper nouns (conventional NER)	Hashtags, #followers, #retweets and timestamps
Petrović et al. (2010)	Online clustering (based on locality sensitive hashing)	#tweets, #users and entropy of messages	–
Becker et al. (2011a)	Online clustering and support vector machine classifier	Term vector	Hashtags, multi-word hashtags with special capitalization, retweets, replies and mentions.
Long et al. (2011)	Hierarchical divisive clustering	Word frequency and entropy	Probability of word occurring in hashtags
Weng and Lee (2011)	Discrete wavelet analysis and graph partitioning	Individual words	–
Cordeiro (2012)	Continuous wavelet analysis and latent Dirichlet allocation	–	Hashtag occurrences
Popescu and Pennacchiotti (2010)	Gradient boosted decision trees	Correlation of target events (or entities) with the Web and traditional news media	Proportion of nouns, verbs, questions, bad words, etc.; #tweet, #retweets, #replies, #tweets per user, hashtags; proportion of tweets and hashtags involving buzziness, sentiment, controversy

(Atefeh and Khreich, 2013)

Case Study: Joint Event Extraction

- Key idea was to model event and entity extraction as a joint problem
  - Model the jointness using a graphical model such as factor graph
- Outperformed state-of-the-art systems at the time
- Joint models continue to be state-of-the-art

(Yang and Mitchell, 2016)

Case Study: Joint Event Extraction

- Features used in Yang and Mitchell's joint IE system

Category	Type	Features
Trigger	*Lexical resources:* WordNet Nomlex FrameNet Word2Vec	1. lemmas of the words in the trigger mention 2. nominalization of the words based on Nomlex (Macleod et al., 1998) 3. context words within a window of size 2 4. similarity features between the head word and a list of trigger seeds based on WordNet (Bronstein et al., 2015) 5. semantic frames that associate with the head word and its p-o-s tag based on FrameNet (Li et al., 2014) 6. pre-trained vector for the head word (Mikolov et al., 2013)
	Syntactic resources: Stanford parser	7. dependency edges involving the head word, both lexicalized and unlexicalized 8. whether the head word is a pronoun
Argument	*Lexical resources:* WordNet	1. lemmas of the words in the entity mention 2. lemmas of the words in the trigger mention 3. words between the entity mention and the trigger mention
	Syntactic resources: Stanford parser	4. the relative position of the entity mention to the trigger mention (before, after, or contain) 5. whether the entity mention and the trigger mention are in the same clause 6. the shortest dependency paths between the entity mention and the trigger mention
Entity	*Entity resources:* Stanford NER NELL KB	1. Gender and animacy attributes of the entity mention 2. Stanford NER type for the entity mention 3. Semantic type for the entity mention based on the NELL knowledge base (Mitchell et al., 2015) 4. Predicted entity type and confidence score for the entity mention output by the entity extractor described in Section 3.3

(Yang and Mitchell, 2016)

Experimental Results (ACE2005 test set)

- P=Precision, R=Recall and F1=F1-Measure

Model	Event Trigger Identification			Event Trigger Classification			Event Argument Identification			Argument Role Classification		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
JOINTBEAM (Li et al., 2013)	76.6	58.7	66.5	74.0	56.7	64.2	74.6	25.5	38.0	68.8	23.5	35.0
STAGEDMAXENT	73.9	**66.5**	70.0	70.4	63.3	66.7	**75.7**	20.2	31.9	**71.2**	19.0	30.0
WITHINEVENT	76.9	63.8	69.7	74.7	62.0	67.7	72.4	37.2	49.2	69.9	35.9	47.4
JOINTEVENTENTITY	**77.6**	65.4	**71.0***	**75.1**	63.3	**68.7**	73.7	**38.5**	**50.6***	70.6	**36.9**	**48.4***

(Yang and Mitchell, 2016)

Open Research Issues

- Context-based advantages of data-driven, knowledge-driven or hybrid approaches
- Understanding the limitations of specific event extraction techniques
- The domain-dependency of event extraction procedures, affecting both their flexibility and effectiveness
- The scalability of event extraction approaches when dealing with Big Data
- The complexity of extracted events
- Better extraction from noisy sources, including social media

Summary

- Event extraction is a difficult problem for which performance continues to be relatively poor (compared to NER)
- Some sub-tasks (e.g., event trigger identification) tend to be easier than others (e.g., argument role classification)
- Joint models have emerged as a solid choice for relatively clean text
- Twitter presents new challenges that continue to be the subject of ongoing research
- Techniques can be classified along many dimensions, depending on application, algorithm, inputs and outputs