



# Applications, Remaining Challenges, Future Directions and Resources

Event-Centric Natural Language Understanding (Part V)

Heng Ji and Dan Roth

Department of Computer Science

University of Illinois at Urbana-Champaign

**Feb 2020**

**AAAI Tutorials**

**Recent Advances in Transferable Representation Learning**

- Applications
  - Writing History Book
  - News Understanding and Recommendation
  - Disaster Relief
  - Intelligence Analysis
  - Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - External Knowledge Acquisition, Reasoning and Incorporation
  - Document-Level Event Extraction
  - Multimedia Information Extraction and Verification
- Grand Vision
- Resources

- Problems on Human Written History Books: Highly Biased (History is written by the victors)
- Translated Japanese comments about Nanjing Massacre:
  - 1. 日本的“南京大屠杀论争”

日本著名的“南京大屠杀论争”，是以南京大屠杀这一事件是否存在、其规模之大小为焦点的论争。随着中日关系的变化，论争不断受到政治的不同影响。

首先来看一下日本网民的看法。和中国一样，在日本的网络贴吧里，也有很多关于南京大屠杀的讨论。2013年2月，在日本某名为“大舰巨炮主义”的贴吧里，“[中国BBS]結局、日本人は南京大虐殺をどう考えているのか？ 2013年02月10日 20:32” ([military38.com/archives...](http://military38.com/archives...)) 这一帖子以“中国人对日本人不反思历史进行批判”这一事实为引子，引出众多网友对南京大屠杀的跟帖议论。

跟帖的人中，有一部分完全否认大屠杀，更大多数人承认在南京发生过大规模杀人事件的。但大部分人虽然承认，却否认30万死者的说法。

来自埼玉县的チーター：大屠杀的数字肯定是捏造啊，30万人从物理角度看都是不可能的。

来自大阪のシンガプーラ：我问过实际去过南京的人的孩子（超过70岁了），据他爸爸说，当时日本军队并不拥有能够杀死几十万中国人的军队人数与武器，30万的数字必然是捏造的。

来自福岡のパンパスネコ：我们要盯着过去到什么时候？我出生的时候战争已经结束了，所以老实说我对大屠杀完全不知道。如果总是拿上一代的责任说事，历史这个东西就会变得一团糟了。
- As my grandma would say, This is why we need to EDUCATE people
- It's international human right to know what happened in the history
- We aim to assist historians and librarians at writing more complete and authentic history books

## Conspiracy theories blaming Bill Gates for the coronavirus pandemic are exploding online

Isobel Asher Hamilton Apr 17, 2020, 6:55 AM



- **As the coronavirus pandemic has spread around the world, with millions infected and thousands dead, billionaire Microsoft co-founder and philanthropist Bill Gates has pledged a quarter billion dollars to combat the disease through his foundation.**

- **Gates has been an advocate for pandemic preparedness for years, and his Bill and Melinda Gates Foundation is contributing financing to several coronavirus vaccine initiatives. He famously gave a 2015 TED talk warning of the potential devastation caused by — and urged readiness for — a worldwide pandemic.**

- Why would anyone ever believe these ridiculous rumors?
- Because humans are very good at connecting dots
- And perhaps too good →

# What will such a History Book look like?

- Organize chapters by major events clusters and order them on a timeline
- Each chapter looks like a Wikipedia page
  - The description is organized by multimedia timeline with detailed source and evidence information, links to original news articles
  - Detailed participants (arguments) and their roles, and their connections and relations
  - Infobox shows event-event relations: temporal, causal and hierarchical
- Never-ending updating over time; put up to the wild for human editing and curation

**Protests against the government** [ edit ]

**4 December 2011** [ edit ]

On 4 November 2011, during the annual [Russian March](#) event, representatives of "[The Russians](#)" movement declared a protest action, unapproved and took place on 4 December at 21:00 in Moscow. The statement of non-recognition of electoral results spread widely. [Citizens](#) said to have occurred during the elections. [Alexander Belov](#) declared the beginning of the "Putin, go away!" campaign.<sup>[37]</sup> The protest also included [Alexander Belov](#), [Dmitry Dyomushkin](#), [George Borovikov](#) were arrested along with dozens of other nationalists. The head of the banned [Nationalist Front Party](#) was also arrested. Mass detentions of other public organizations occurred in Moscow. According to police some 258 persons have been detained.

**5–7 December 2011** [ edit ]

On 5 December, around 5,000 opponents of the government began protesting in Moscow, denouncing [Vladimir Putin](#) and [his government](#). [Putin](#) step down, whilst some demanded revolution.<sup>[15][39]</sup> [Alexey Navalny](#), a top blogger and anti-corruption activist who branded [Putin](#) posts on his [LiveJournal](#) blog and Twitter account. Navalny's agitation was denounced by United Russia as "typical dirty self-promotional account."<sup>[40][41]</sup>

Many pro-government supporters, including the pro-Putin youth group [Nashi](#), were mobilized on 6 December at the site of the planned [Nashi](#) on [Manezhnaya Square](#)<sup>[43]</sup> and an 8,000-strong rally of the [Young Guard](#) on [Revolution Square](#).<sup>[44]</sup> About 500 pro-United Russia expected anti-government protests. It emerged that 300 protesters had been arrested in Moscow the night before, along with 120 in [St. Petersburg](#).<sup>[45]</sup> Slogans against [Putin](#),<sup>[15]</sup> whilst anti-government protesters at [Revolution Square](#) clashed with riot police and [interior ministry troops](#). [Troops](#) on [Revolution Square](#) and dozens of arrests were reported, including [Boris Nemtsov](#), an opposition leader and former deputy prime minister,<sup>[48]</sup> and [Alexey Navalny](#) was also charged. At least one Russian journalist claimed he was beaten by police officers who stamped on him and hit his legs with batons.<sup>[50]</sup> The protest ended in [Bolotnaya Square](#). After three and a half hours, the Moscow protest came to an end.<sup>[51]</sup>

Attempts to stage a large protest in Moscow on 7 December fizzled out due to a large police presence in the city.<sup>[13]</sup>

**10 December 2011** [ edit ]



Rally in [Bolotnaya Square](#) in [Moscow](#) on 10 December 2011

Via a Facebook group "[Суббота на Болотной площади](#)" (Saturday at Bolotnaya Square),<sup>[52]</sup> [Bolotnaya Square](#) demonstration was authorized by the Moscow government for the demonstration which took place on 10 December.<sup>[53][54]</sup> Prior to the demonstration newspapers commented that tens of thousands of protesters gathered in [Moscow](#),<sup>[55][56]</sup> and, similarly, over 5,000 in [St. Petersburg](#).<sup>[57]</sup> A permit had originally been issued for a demonstration on [Bolotnaya Square](#). By 8 December, more than 30,000<sup>[52]</sup> had accepted the Facebook invitation to attend. The demonstration was authorized by the Moscow government for the demonstration which took place on 10 December. The demonstration was peaceful and without attempts by the state to prevent or disrupt it.<sup>[59][60]</sup> Rapper [Noize MC](#) performed at the demonstration, specially from [Paris](#) for the occasion.<sup>[61]</sup> [Guerrilla theater](#) by [FEMEN](#) and the circulation of a pamphlet were also reported.<sup>[62][63]</sup>



# Example Chapter: 2019 Hong Kong Protests

## The first big protests

On **9 June**, an estimated one million people marched to the government headquarters to show they were against the proposed bill.



## Protesters storm parliament

On **1 July**, the anniversary of Hong Kong's handover from the UK to China, the Legislative Council (LegCo) building was stormed by protesters who sprayed graffiti on the walls, displayed the colonial-era flag and defaced Hong Kong's regional emblem.

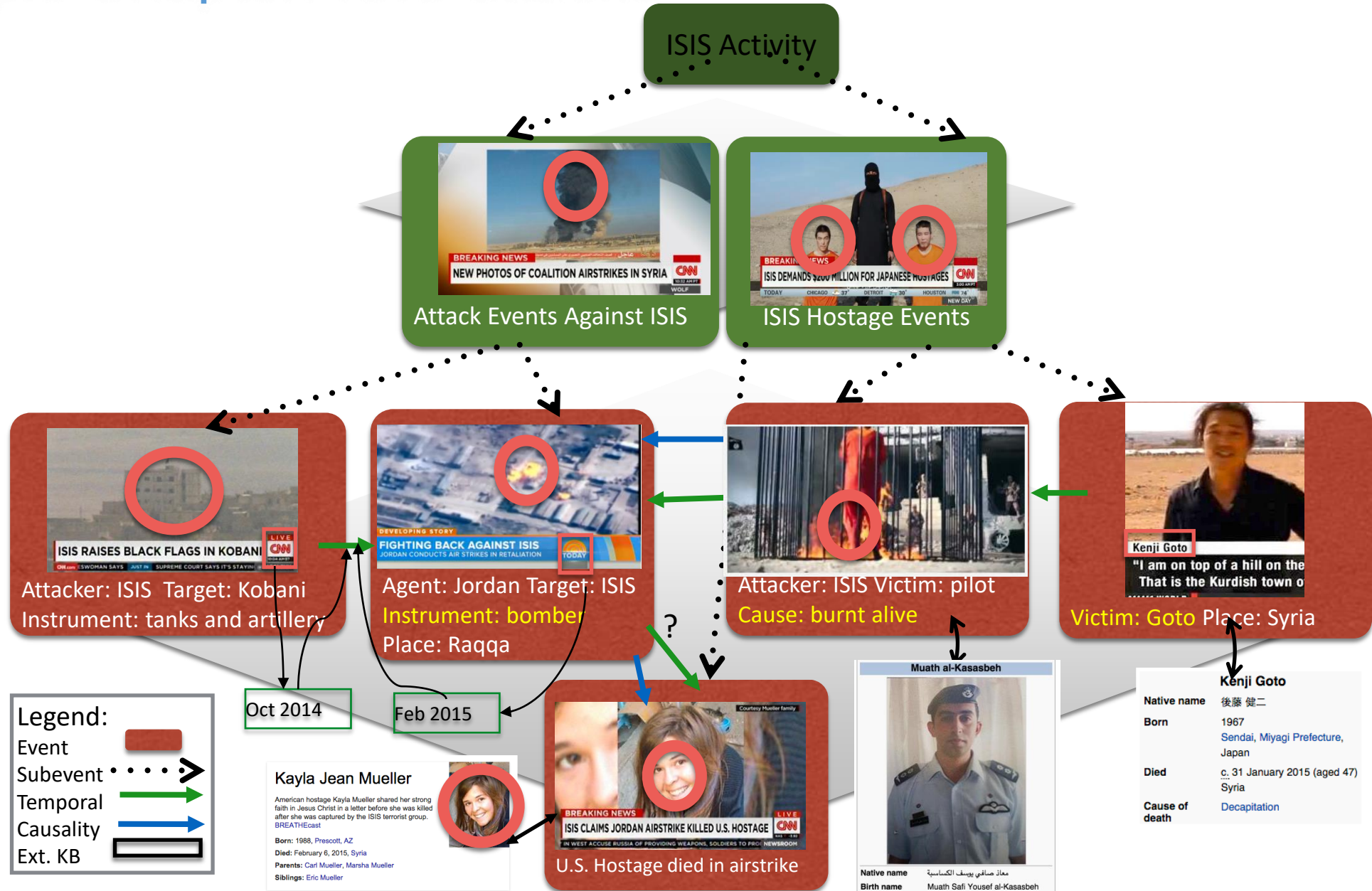


One week later, on **7 July**, tens of thousands marched in Kowloon - an area popular with mainland tourists - in a bid to explain their concerns. Until this point the protests had received little if any coverage in state-run mainland media.

On **9 July**, Carrie Lam reiterated that the extradition bill was "dead" urging protesters to stop their actions. She still refrained from fully withdrawing the bill.

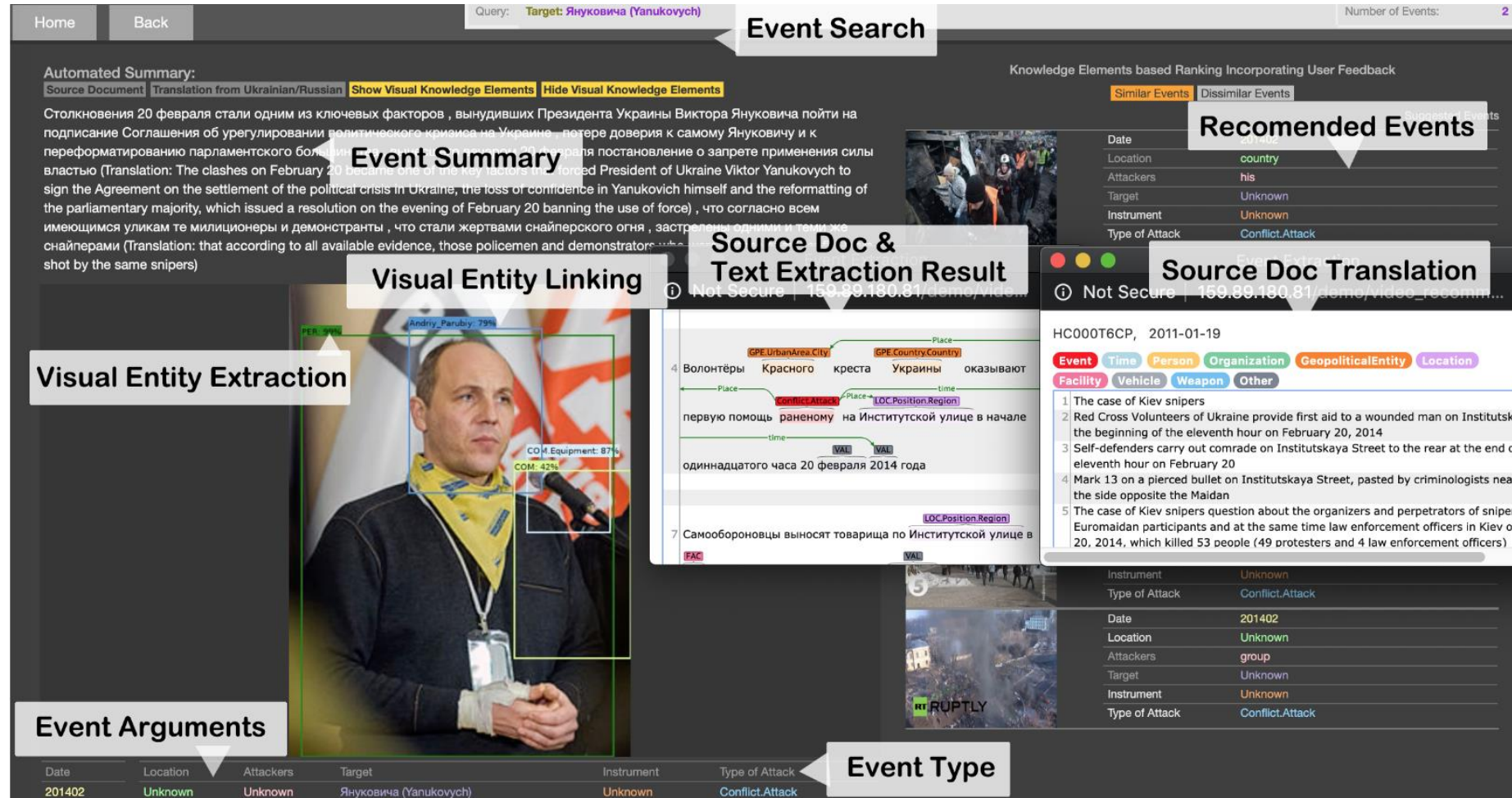
On **21 July**, protesters defaced China's Liaison Office in Hong Kong.

# Example Chapter: ISIS attacks





# Application 2: News Understanding and Recommendation (Li et al., ACL2020Demo)



The screenshot displays the GAIA-IE application interface with several key components:

- Event Search:** A search bar at the top with the query "Target: Януковича (Yanukovich)" and a "Number of Events: 2" indicator.
- Automated Summary:** A text block on the left providing a summary of the event in both Ukrainian/Russian and English. Callouts include "Event Summary" and "Source Doc & Text Extraction Result".
- Visual Entity Linking:** A central image of a man with callouts for "Visual Entity Extraction" and "Visual Entity Linking".
- Source Doc Translation:** A text block on the right showing the translated source document with callouts for "Source Doc Translation" and "Event Type".
- Recommended Events:** A table on the right displaying event details such as Date, Location, Attackers, Target, Instrument, and Type of Attack. Callouts include "Recommended Events" and "Event Type".
- Event Arguments:** A table at the bottom showing event arguments like Date, Location, Attackers, Target, Instrument, and Type of Attack.

(Li et al., ACL2020 Best Demo Paper Award)

GitHub: <https://github.com/GAIA-IE/gaia>

DockerHub: <https://hub.docker.com/orgs/blendernlp/repositories>

Demo: [http://159.89.180.81/demo/video\\_recommendation/index\\_attack\\_dark.html](http://159.89.180.81/demo/video_recommendation/index_attack_dark.html)



## ■ Identifying event orders and predicting future events

Heavy snow is causing disruption to transport across the UK, with heavy rainfall bringing flooding to the south-west of England. Rescuers searching for a woman trapped in a landslide at her home in Looe, Cornwall, said they had found a body.

**Q1: What events have already finished?**

A: searching trapped landslide said found

**Q2: What events have begun but has not finished?**

A: snow causing disruption rainfall bringing flooding

**Q3: What will happen in the future?**

A: No answers.

warm-up

**Q4: What happened before a woman was trapped?**

A: landslide

**Q5: What had started before a woman was trapped?**

A: snow rainfall landslide

**Q6: What happened while a woman was trapped?**

A: searching

**Q7: What happened after a woman was trapped?**

A: searching said found

User-provided

Ning, et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. EMNLP, 2020

- 3.2k news snippets with 21k human-generated questions querying temporal relationships

**Q: Who will drop Japan as a trading partner in August 2019?**

earlier than  
timestamp  
(2019-08-01)

(1/1/19) Apart from the fact of being one another's closest neighbors, the people of South Korea and Japan have a remarkable amount in common. Economically, they are among one another's biggest trading partners. And yet, time and again, relations between Seoul and Tokyo are marked, not by mutual support and co-operation but by anger, reproach and exasperation.

A) South Korea [0.41]    B) Syria [0.28]  
C) South Africa [0.15]    D) Portugal [0.16]

**Q: Will primary schools in Europe admit non-vaccinated children around September 2019?**

earlier than  
timestamp  
(2019-09-01)

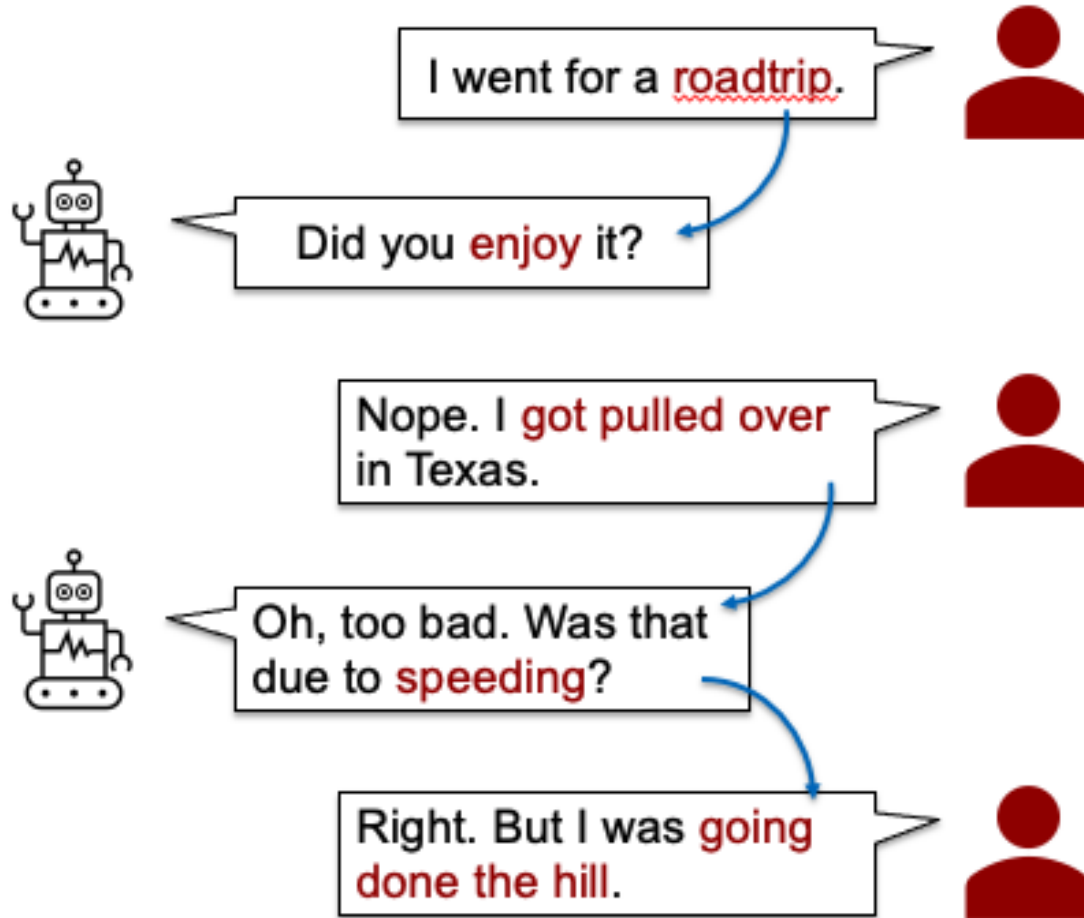
(3/8/18) Public officials and health experts had given several warnings: Do not allow a student in school if they had not been vaccinated against measles.

(6/27/19) Fines for parents refusing measles jab. Parents will be fined up to € 2,500 if they don't vaccinate their children against measles under draft legislation in Germany which also threatens exclusion from crèches, nurseries and schools.

Yes [0.38] / No [0.62]

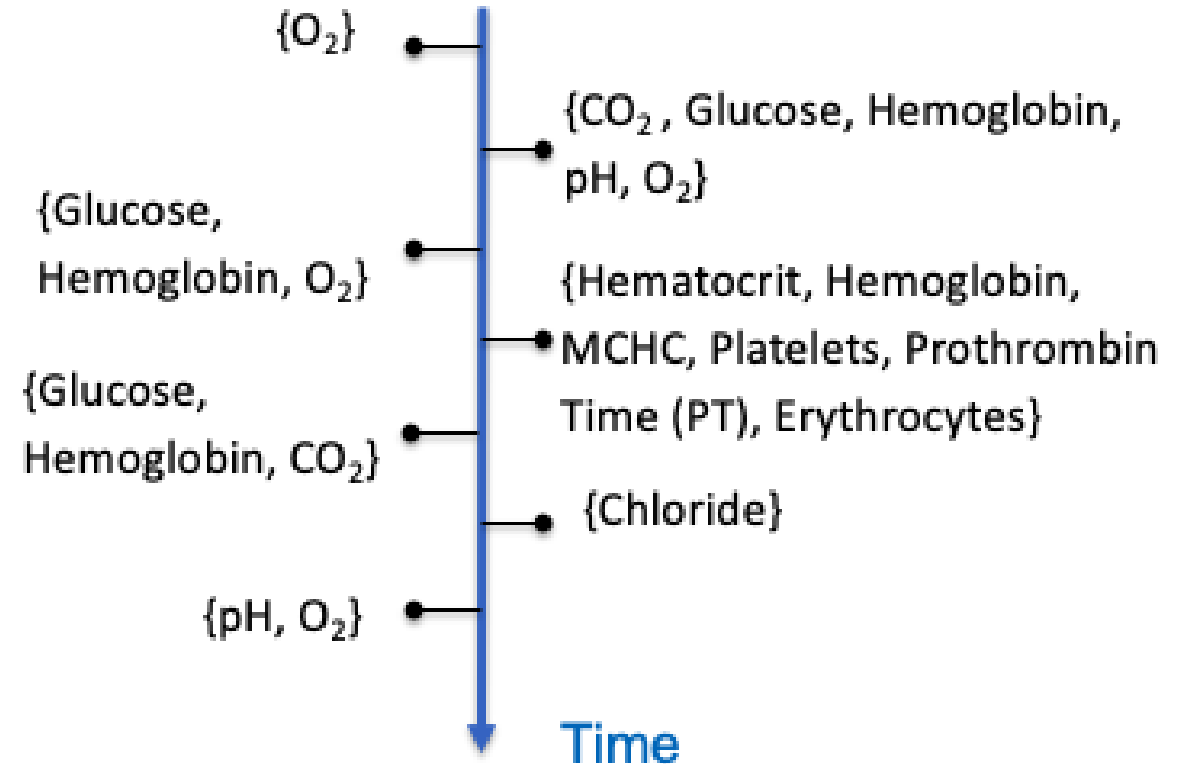
ForecastQA: A Question Answering Challenge for Event Forecasting

## ■ Chatbots



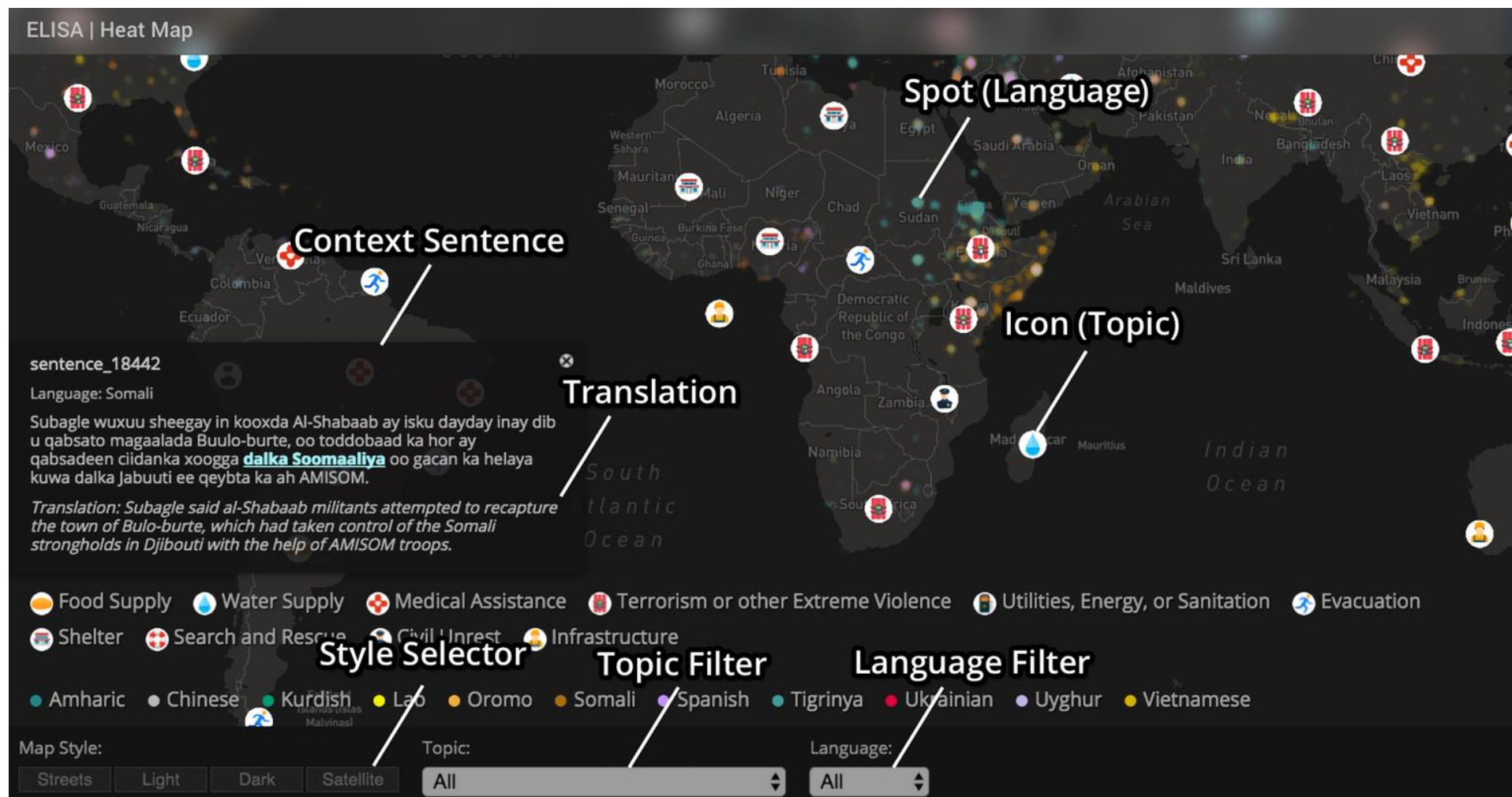
Can event processes improve the consistency of utterance generation/retrieval?

## ■ Clinical event processes



Diagnostic prediction, phenotype prediction, ...

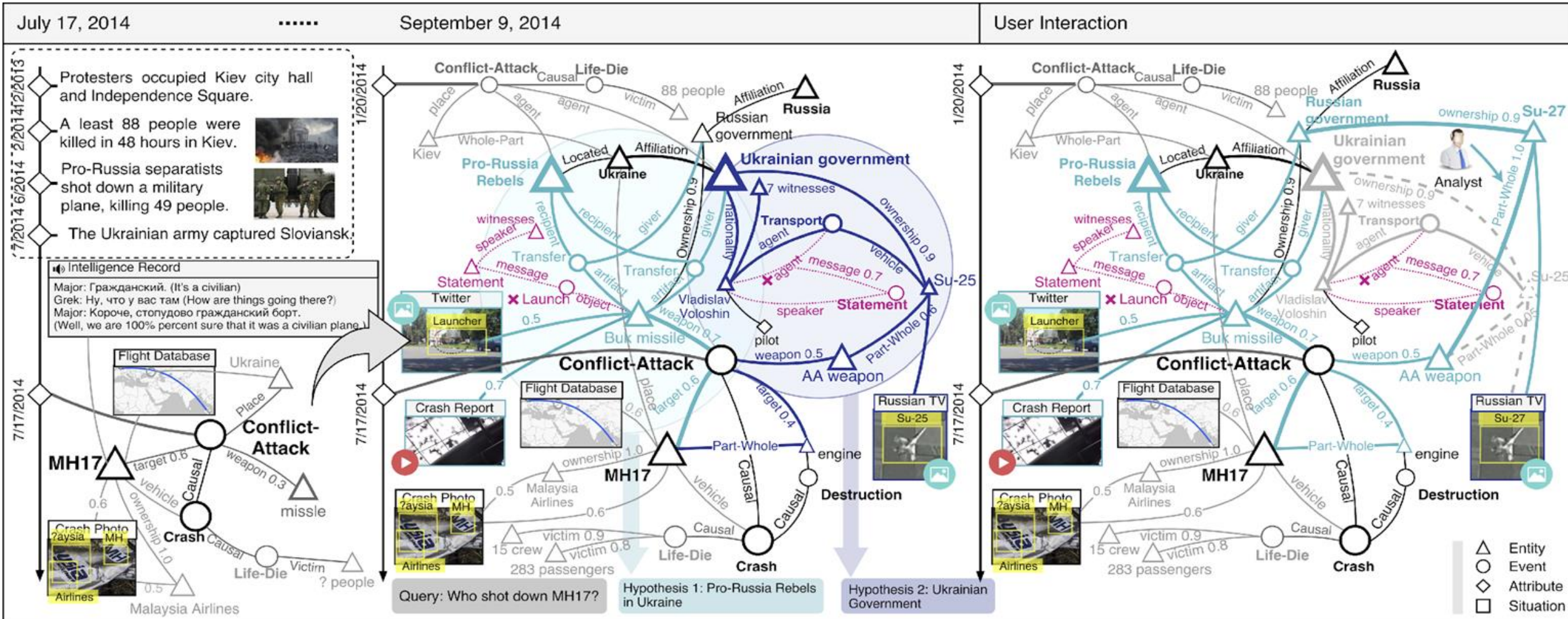
- Transfer learning can be important (naturally lack of data)
- Structured prediction can be important (dependency of phenotypes, disease labels)



- Re-trainable Systems: [http://159.89.180.81:3300/elisa\\_ie/api](http://159.89.180.81:3300/elisa_ie/api)
- Demos: [http://159.89.180.81:3300/elisa\\_ie](http://159.89.180.81:3300/elisa_ie)
- Heat map: <http://159.89.180.81:8080/>

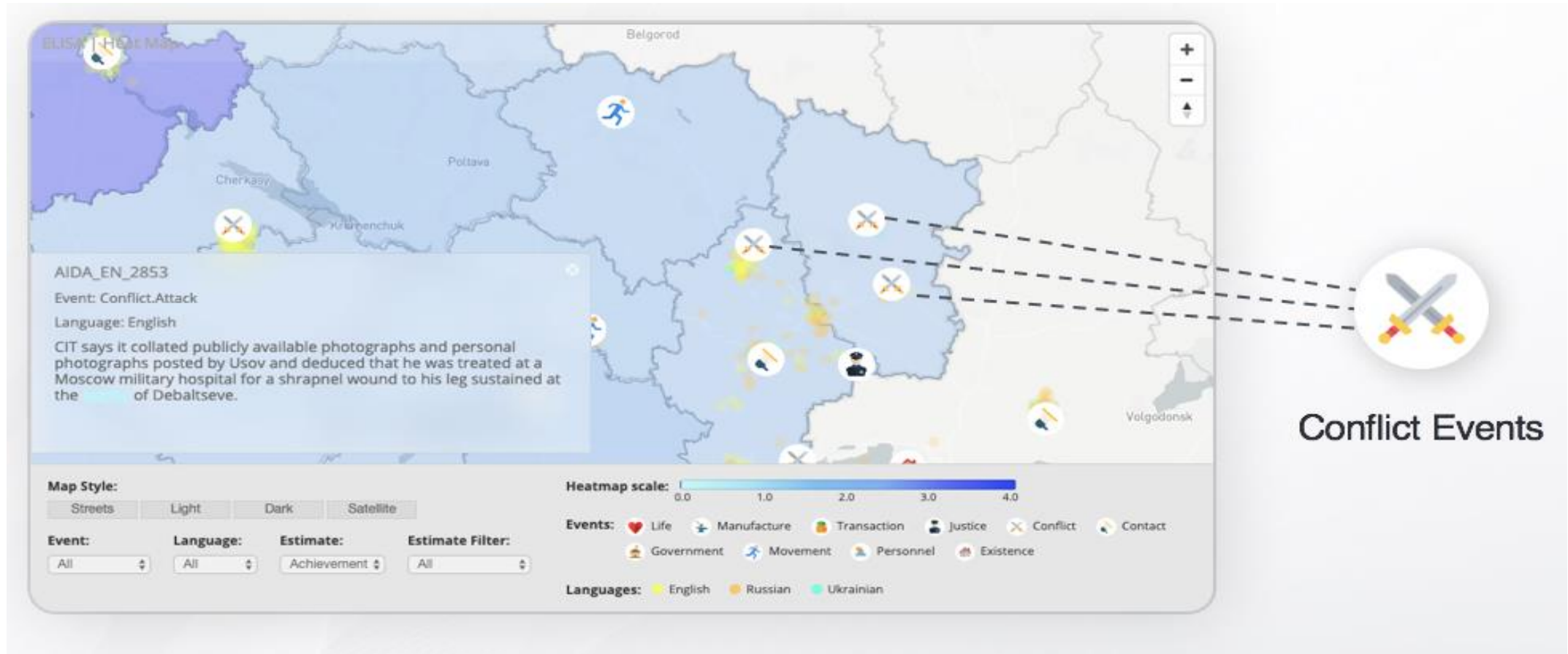


# Applications 4: Intelligence Analysis





# Applications 4: Intelligence Analysis



<http://162.243.120.148:8080/>

- Achievement, Benevolence, Conformity, Hedonism, Power, Security, Self-direction, Simulation, Tradition and Universalism (Schwartz, 2012)

# Application 5: Accelerating Scientific Discovery

DATA

NeurIPS'19, WWW'20,  
KDD'20, ACL'20

SEMANTICS

NAACL'19, EMNLP'19,  
ACL'19, ACL'20

KNOWLEDGE BASE

AAAI'19, ACL'20

**Image**

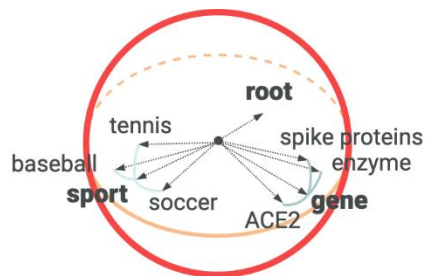
**Text**

The **Ni** [Element] catalyzed Suzuki coupling reaction [Reaction, Combination] also allowed a number of compounds that worked worse for the **palladium** [Element] catalyzed system than the **nickel** [Element] catalyzed system.

Scientific Literature



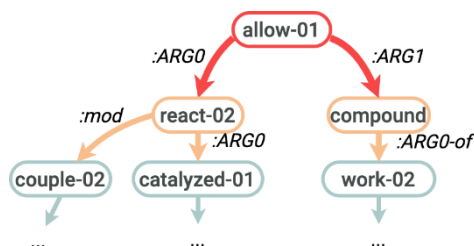
Chemical Ontology & Existing Databases



Hierarchical Spherical Embedding

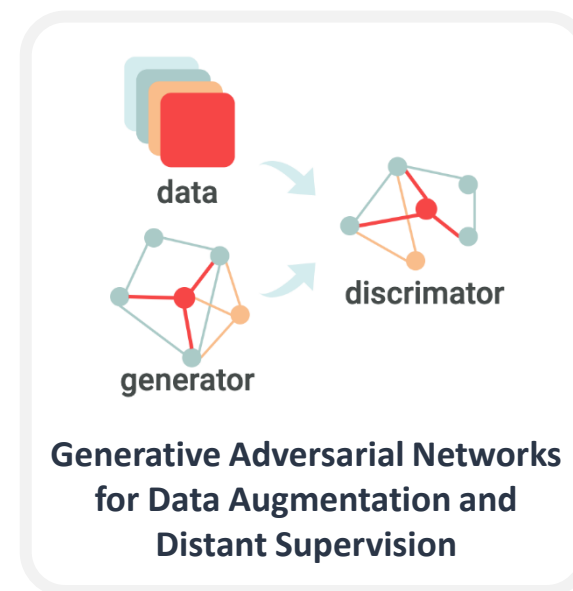
KB concept of Ni	
Name: Nickel	Element category: transition metal
Chemical Symbol: Ni	Appearance: lustrous, metallic, and silver with a gold tinge
Atomic Number: 28	
Type: chemical element	

Ontology Enriched Text Embedding



Cross-media Structured Semantic Representation

Graph neural networks  
Joint entity/relation/event extraction  
and ontology construction



Generative Adversarial Networks  
for Data Augmentation and  
Distant Supervision

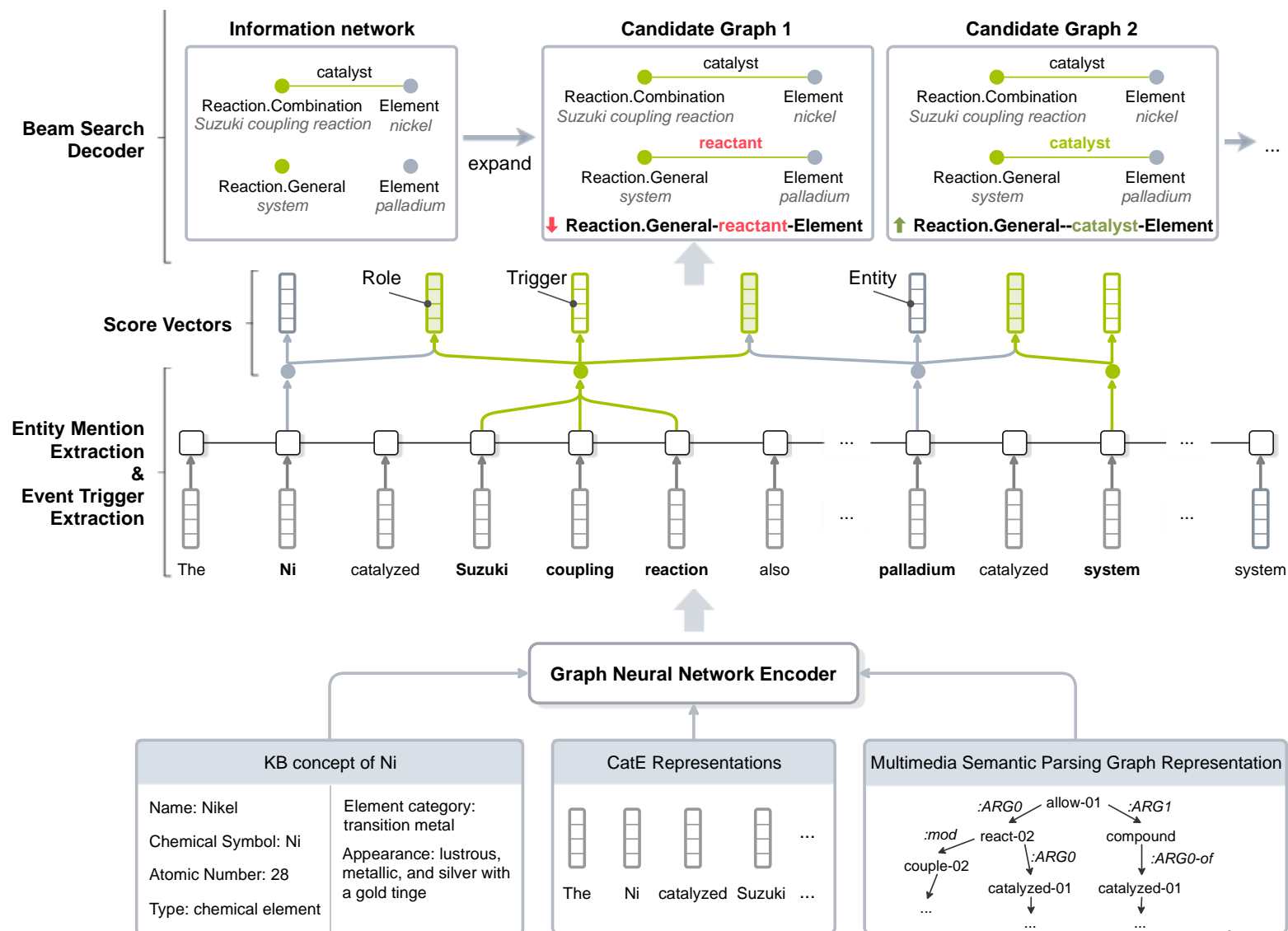


Multimedia Knowledge Base

Multimedia Search and Summarization

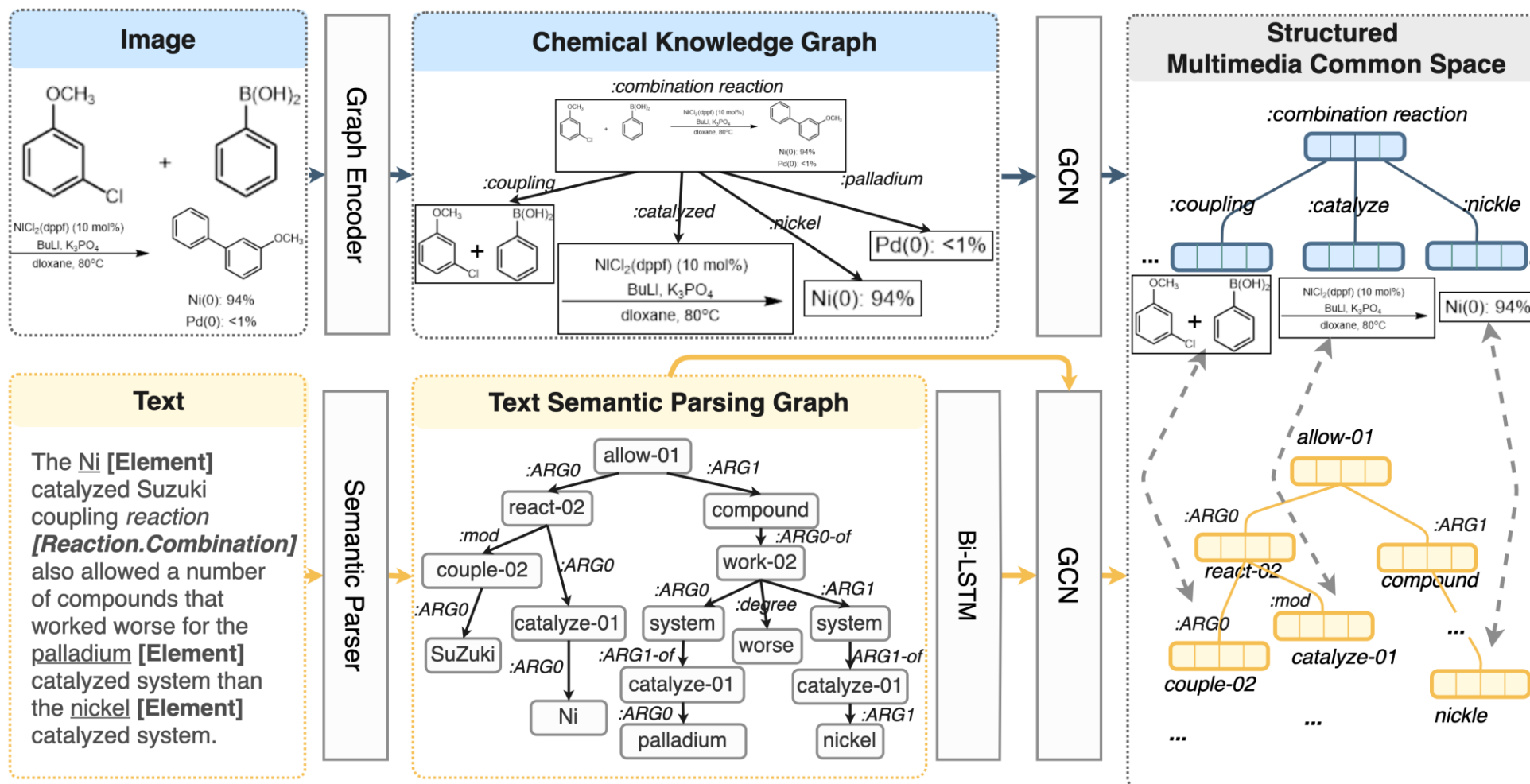
■ Extending to other scientific domains including Molecular Synthesis and Agriculture under two new NSF institutes

# Fine-grained COVID-19 Event Extraction



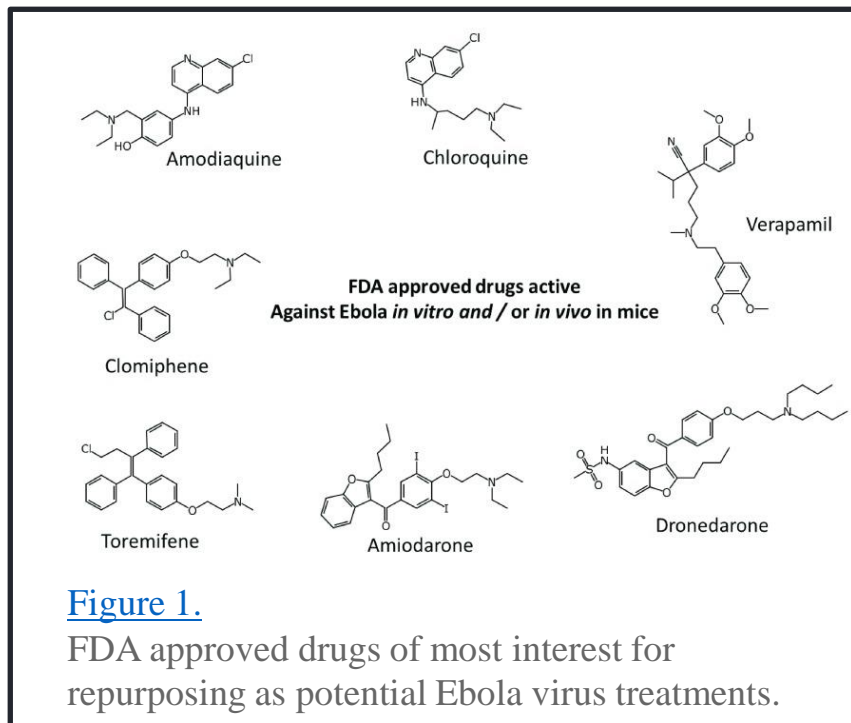
- Joint neural Information Extraction model is proven successful for news domain (Lin et al., ACL2020)

# Fine-grained COVID-19 Event Extraction



- Multimedia Common Semantic Space Construction is proven successful for news domain (Li et al., ACL2020)

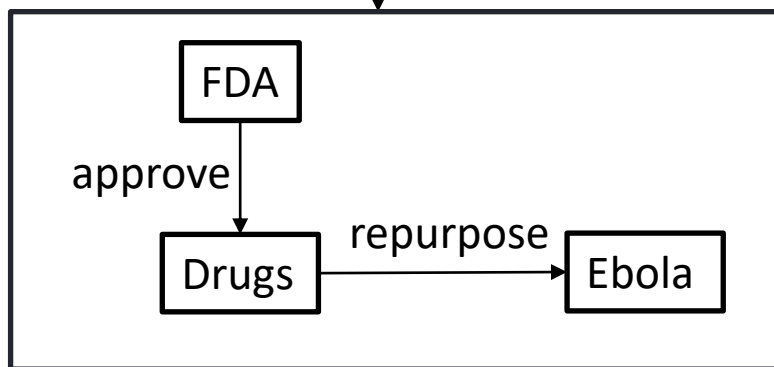




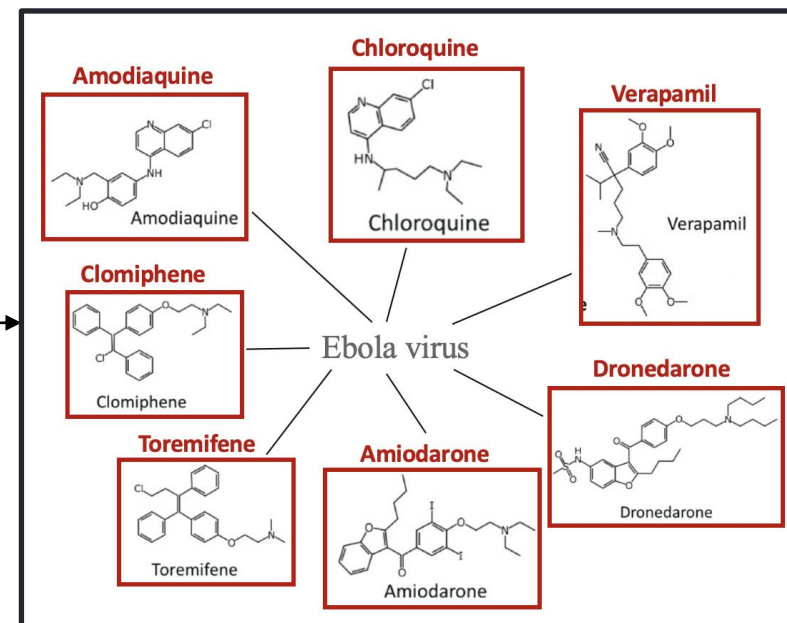
**Figure 1.**  
FDA approved drugs of most interest for repurposing as potential Ebola virus treatments.

Entity Grounding for Drug Molecular Structure Image

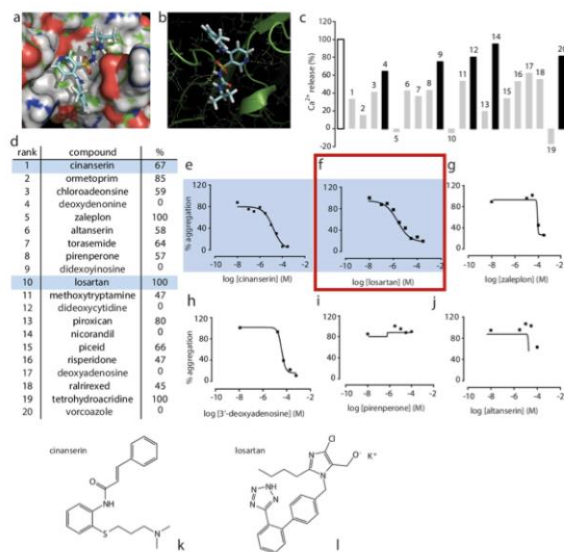
KG from caption text



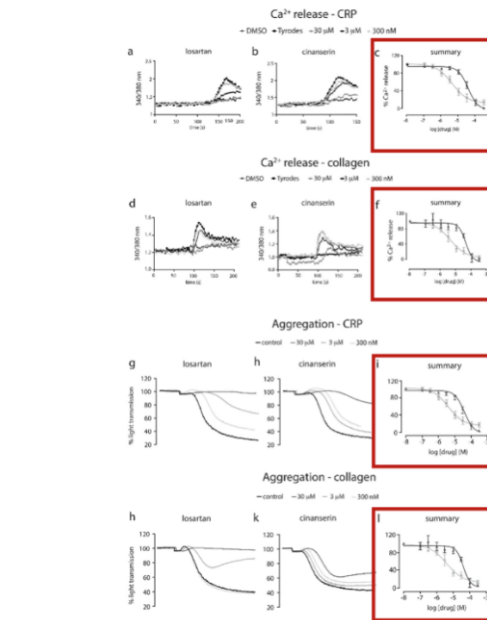
Multimedia Knowledge Graph Expansion



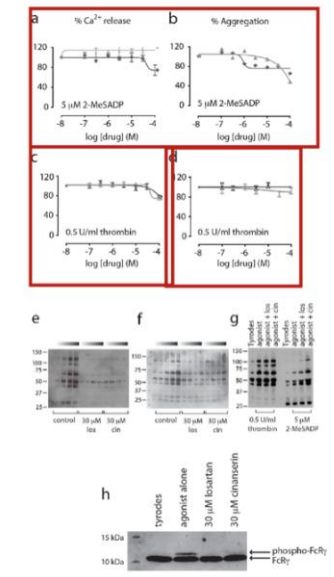
# Drug Repurposing Report Generation (Section 6: In vitro Data available)



Src: PMC4074120 Fig. 1. In silico identifies GPVI antagonist. Representative image capture of in silico docking into GPVI using Glide, with space filling model is shown in a, and H-bonding to relevant side chains is detailed in b. The 20 highest ranking compounds were screened for effects on Ca<sup>2+</sup> release by the GPVI-specific agonist CRP-XL (10 mg/ml) (c and d, % refers to percent inhibition of Ca<sup>2+</sup> release). Maximum Ca<sup>2+</sup> release is shown in white, compounds that inhibited Ca<sup>2+</sup> release by 50% or more are in grey, and the remainder in black. Commercially available compounds that inhibited CRP-XL-induced Ca<sup>2+</sup> release 50% were further screened by light transmission aggregometry to identify compounds displaying dose-dependent inhibition (e-j). Examples are shown of weak antagonism (g and h) and false positives (i and j). Cinanserin (i) and losartan (k) were taken on for further study.



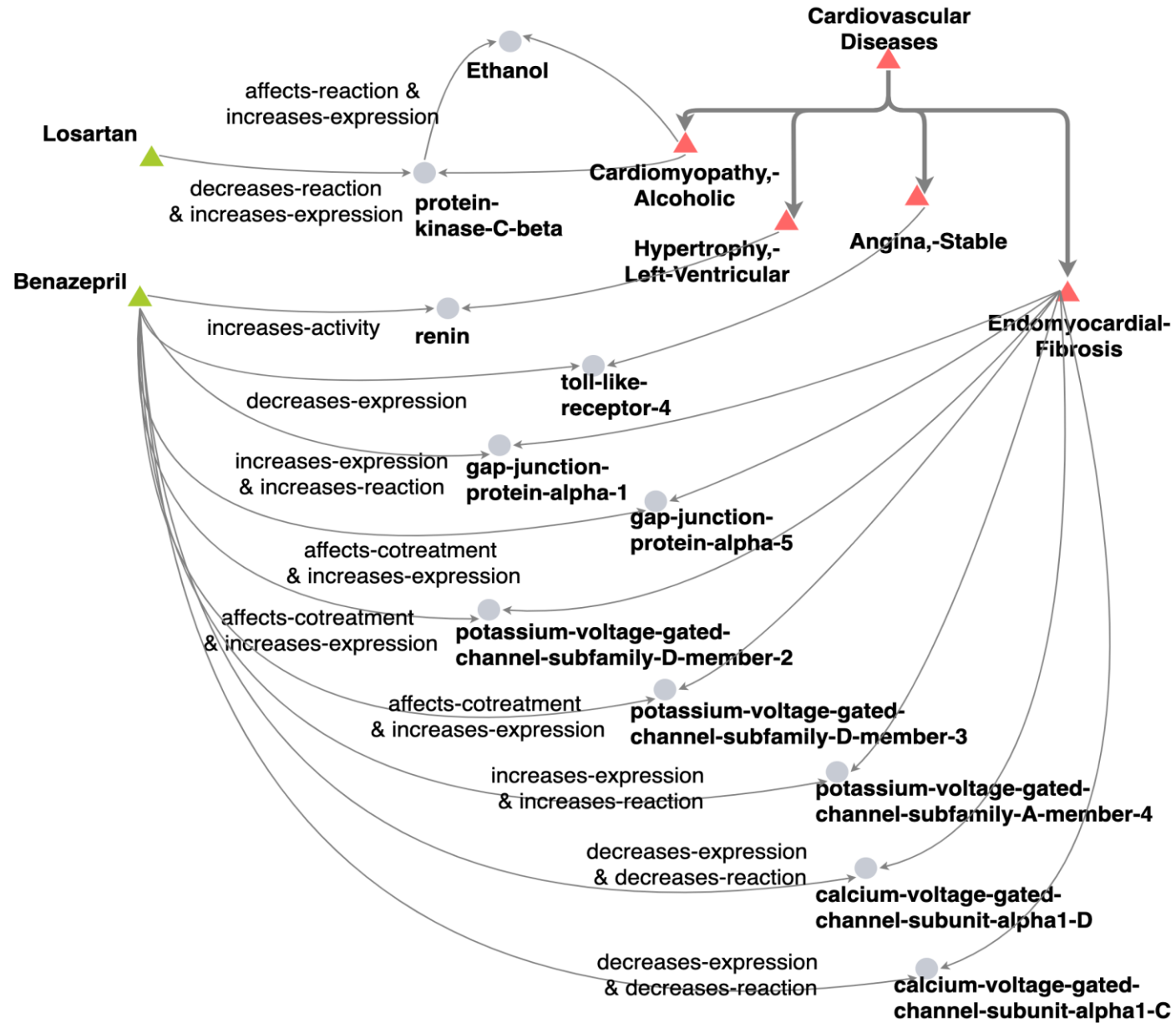
Src: PMC4074120 Fig. 2. Losartan and cinanserin inhibit GPVI-mediated cell activation. Washed human platelets were loaded with fura-2-AM and screened for drug-mediated inhibition of Ca<sup>2+</sup> release by 1 mg/ml CRP-XL (n = 3, 6 SEM, representative traces and summary, a-c) and 1 mg/ml collagen (n = 3, SEM, representative traces and summary, d-f); losartan (k) and cinanserin (m). To measure aggregation, washed human platelets were incubated with drug for one minute prior to the addition of 1 mg/ml CRP-XL (representative traces and summary shown in g-i) or 1 mg/ml collagen (representative traces and summary shown in j-l).



Src: PMC4074120 Fig. 3. Losartan and cinanserin demonstrate selectivity for GPVI. Ca<sup>2+</sup> release and aggregations were carried out with 5 mM of the P2Y<sub>12</sub> receptor agonist 2-MeSADP (a, Ca<sup>2+</sup> release and b, aggregation), or 0.5 U/ml of the PAR<sub>1</sub> and PAR<sub>4</sub> receptor agonist thrombin (c, Ca<sup>2+</sup> release and d, aggregation). Losartan (X); cinanserin (m), n = 3, 6 SEM. For global tyrosine phosphorylation, washed human platelets were incubated with drug or vehicle alone before addition of 1 mg/ml CRP-XL or collagen. Samples were collected at 10, 30, 60 or 90 seconds (as indicated by the graduated bars with time increasing to the right) in ice cold 26 lysis buffer and separated on 4-12% NuPage gels under reducing conditions. Tyrosine phosphorylation was visualized with 4G10 anti-phosphotyrosine antibody. Losartan and cinanserin reduce CRP-XL- (e) and collagen- (f) induced global tyrosine phosphorylation, but have no effect on thrombin or 2-MeSADP induced global tyrosine phosphorylation (g). Both drugs reduce FcγR2 phosphorylation (h), (unphosphorylated, lower band; phosphorylated, upper band).

Disease	Cardiovascular disease
PMID, PMCID	Evidence Sentences
22800722 <a href="#">PMC7102827</a>	The in vitro half-maximal inhibitory concentration (IC <sub>50</sub> ) values of food-derived ACE inhibitory peptides are about 1000-fold higher than that of synthetic captopril but they have higher in vivo activities than would be expected from their in vitro activities..... Germinal ACE depends on chloride to a lesser extent compared with the C domain of sACE. Cushman and Cheung reported an optimal in vitro ACE activity of rabbit lung acetone extract in the presence of 300 mM NaCl at pH 8.1-8.3...

# Section 9: Has the drug shown evidence of systemic toxicity?



- Applications
  - Writing History Book
  - News Understanding and Recommendation
  - Disaster Relief
  - Intelligence Analysis
  - Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - External Knowledge Acquisition, Reasoning and Incorporation
  - Document-Level Event Extraction
  - Multimedia Information Extraction and Verification
- Grand Vision
- Resources



# Event Extraction Challenges: Scene Understanding

- Three Young Boys, ages 2, 5 and 10 survived and are in **critical** [**injure**] condition after spending in 18 hours in the cold.
- This was the Italian ship that was taken -- that was **captured** [**transfer-ownership**] by Palestinian terrorists back in **1985** [**die\_time**] and some may remember the story of Leon clinghover, he was in a cheal chair and the terrorists shot him and pushed him over the side of the ship into the Mediterranean [**die\_place**] where **he** [**die\_victim**] obviously, died.
- Then police say the baby's mother pulled out a kitchen **knife** [**die\_instrument**] opinion on the 911 tape you can hear Williams tape say "go ahead kill me."
- 



- Scenario = ISIS attack → Event = bombing instead of concert

## ■ Rare Triggers

- his men **back** to their compound
- A suicide bomber **detonated** explosives at the entrance to a crowded
- medical teams **carting** away dozens of wounded victims
- Today I **was let go** from my job after working there for 4 1/2 years.
- This morning in Michigan, a second straight night and into this morning, hundreds of people have been **rioting** in Benton harbor.

## ■ Rare Arguments

- We've seen in the past in Bosnia for example, you held elections and all of the old ethnic **thugs** **[election\_person]** get into power because they have organization and they have money and they stop the process of genuine building of democracy.
- He called the case and I'm quoting now, a judge's worst nightmare, but he noted that Maryland Parole Boards and the facility where Michael Serious was held, the institution made what the judge called the final decision on whether to release **serious** **[release-parole\_person]**.
- Last week Williamson, a mother of four, was found stabbed to death at a **condominium** **[die\_place]** in Greenbelt, Maryland.
- A source tell US Enron is considering suing its own investment bankers for **giving it bad financial advice** **[Sue\_crime]**.

- Ellison to spend \$10.3 billion to **get [org-acquisition]** his company.
- I want to **take [transport]** this opportunity to stand behind the Mimi and proclaim my solidarity.
- He's **left [transport]** a lot on the table.
- Stewart has found the road to fortune wherever she has **traveled [transport]**.
- And it's hard to win back that sort of brand equity that she's **lost [end-position]**.
- Still **hurts [attack]** me to read this.
- We happen to be at a very nice spot by the beach where this is a chance for people to **get [transport]** away from CNN coverage, everything, and kind of relax.
- He bought the machinery, **moved [transport]** to a new factory, rehired some of the old workers and started heritage programs.

- Super-event and subevent are often mistakenly identified as coreferential
  - ❑ These same imposters were later filmed **shooting** handguns and automatic weapons in the direction of the building just minutes after they had switched sides .
  - ❑ Right sector goons started the **fire** by throwing Molotov cocktails through the windows
- HC000030E: a sequence of subevents are often mistakenly considered as coreferential
  - ❑ Nor does the author speculate on why the police stood by while people hurled themselves from windows to escape the fire or were savagely **beaten** by right wing extremists on the pavement in front of the building .
  - ❑ Its true that Washington supports Neo-Nazi extremists who **burned down** the Odessa Trade Unions House .



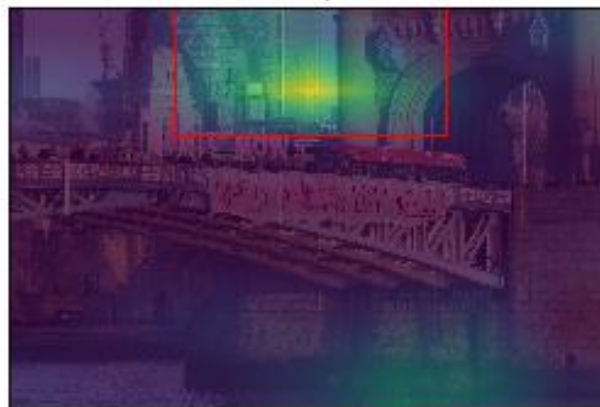
# Remaining Challenges for Multimedia Extraction: wrong localization

Predicted event: CONFLICT|DEMONSTRATE, ground truth: CONFLICT|DEMONSTRATE  
Predicted verb: parading

ENTITY: people



PLACE: open air



ENTITY: demonstrator



Predicted event: CONFLICT|DEMONSTRATE, ground truth: CONTACT|MEET  
Predicted verb: parading

ENTITY: people



PLACE: street



ENTITY: troops



Predicted event: CONFLICT | DEMONSTRATE, ground truth: CONFLICT | DEMONSTRATE  
Predicted verb: parading

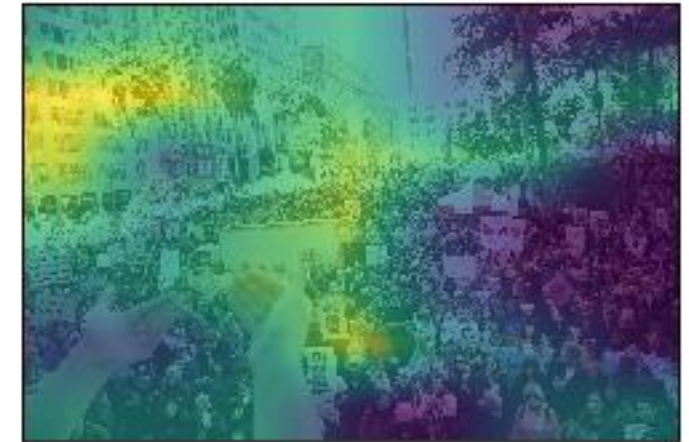
ENTITY: people



PLACE: street



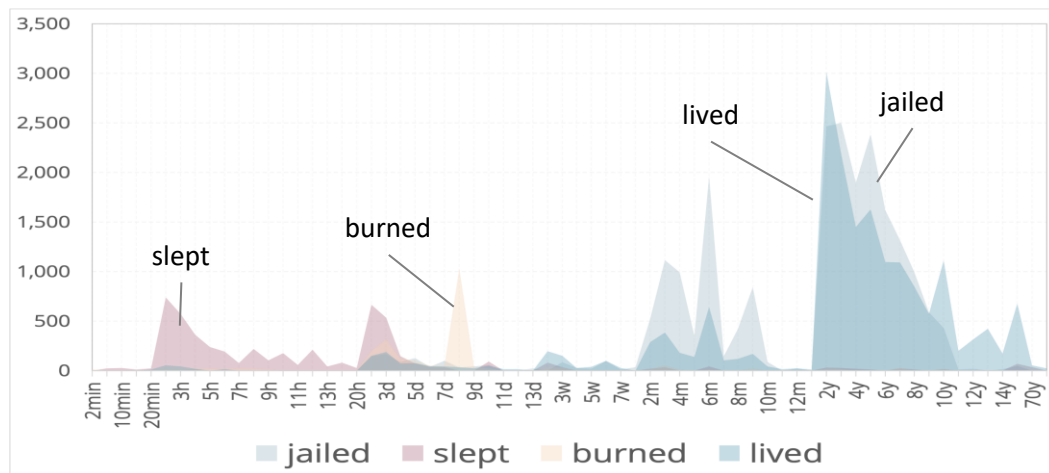
ENTITY: dissent





## 10 Discover from Google N-grams using lexical patterns

□ Verb + for + Time / Spent + Time + Verb-ing



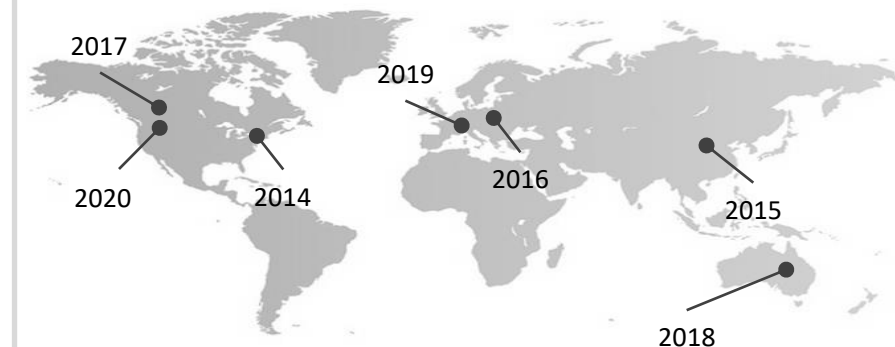
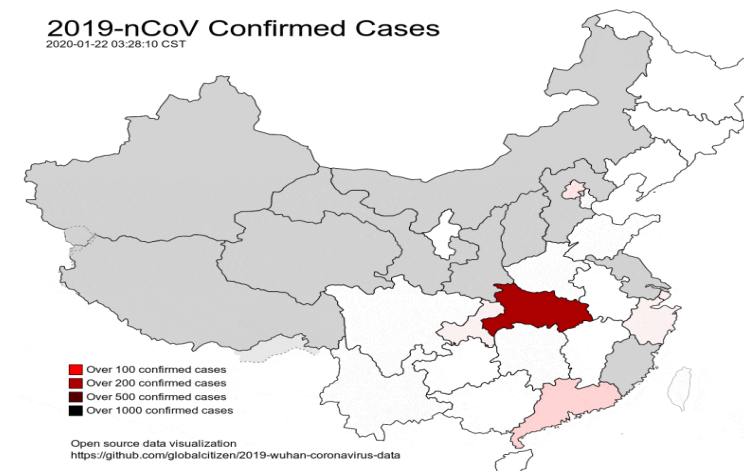
## 10 Discover from Wikipedia event infoboxes and categories

**Date** 1 September 1939 – 2 September 1945  
(6 years and 1 day)<sup>[a]</sup>

Categories: World Wars | **World War II** | Conflicts in 1939  
 Conflicts in 1940 | Conflicts in 1941 | Conflicts in 1942  
 Conflicts in 1943 | Conflicts in 1944 | Conflicts in 1945  
**Global conflicts** | Modern Europe | Modern history | Nuclear warfare

## Event Spatial Evolution Pattern

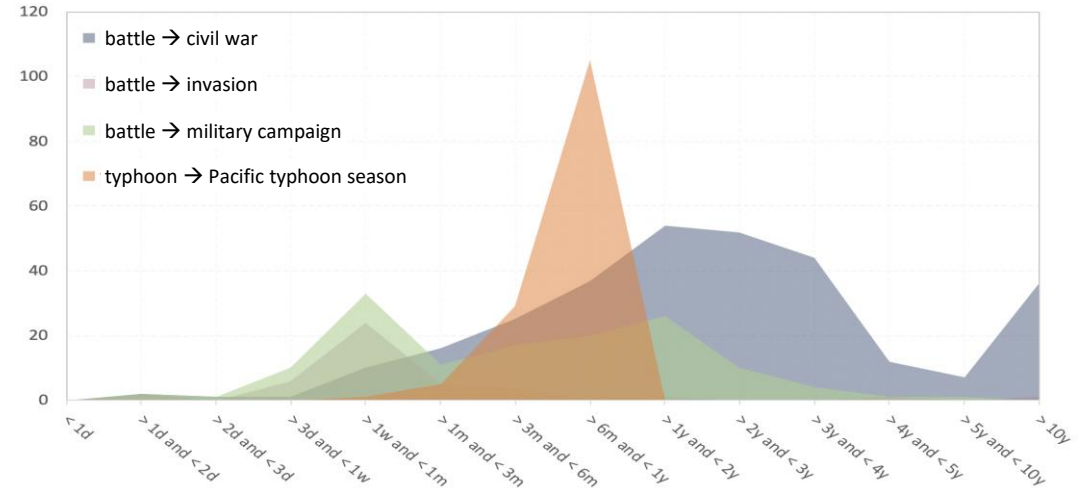
2019-nCoV Confirmed Cases  
2020-01-22 03:28:10 CST



# Future Direction 1: Knowledge Acquisition and Reasoning

- Exploit Wikipedia/Wikidata
  - layout, category (event type), part of relation, has cause, has immediate cause, has contributing factor, has effect, immediate cause of, start/end time
  - Mining external links in Wikipedia articles and edit history; measuring the similarity of two temporal histograms

## Cross-event Time Gap Pattern Discovery



## Hierarchical/Causal Composition

### Egyptian Crisis (2011–2014)

From Wikipedia, the free encyclopedia

"Egyptian crisis" redirects here. For the crisis of 1840, see [1840 Egyptian crisis](#).

The **Egyptian Crisis** began with the [Egyptian revolution of 2011](#), an ideologically and socially diverse mass protest movement that ultimately forced longtime president [Hosni Mubarak](#) from office.<sup>[1][2]</sup> A protracted political crisis ensued, with the [Supreme Council of the Armed Forces](#) taking control of the country until a

### 2011 Egyptian Revolution

Part of the Egyptian Crisis and the Arab Spring

**Date** 25 January 2011 – 11 February 2011 (2 weeks and 3 days)

**Location** Egypt

### 30 June 2013 Egyptian protests

Part of the 2012–13 Egyptian protests during the Egyptian Crisis

**Date** 30 June 2013 – 3 July 2013 (3 days)

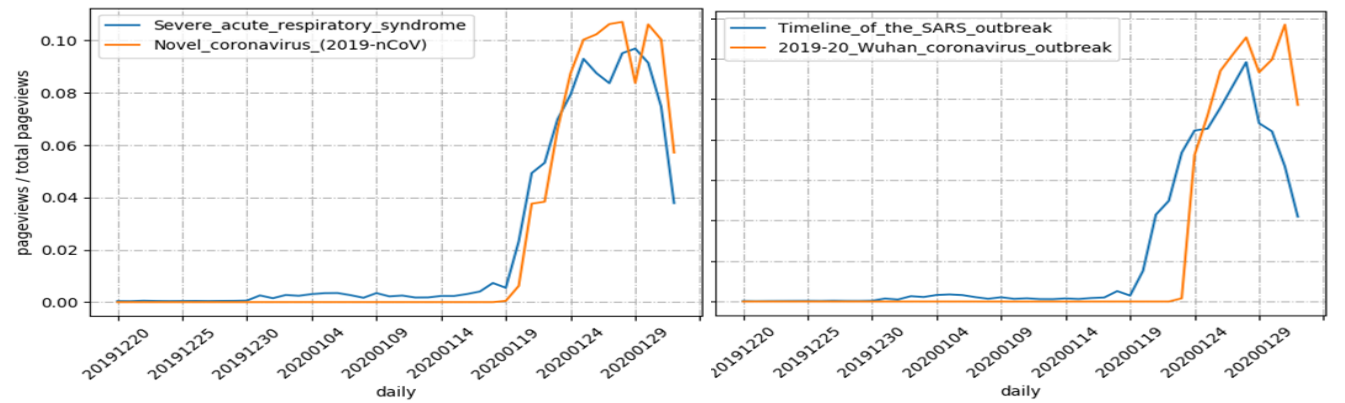
**Location** Egypt

### Sinai insurgency (2011–present)

Part of the [Insurgency in Egypt \(2013–present\)](#), the [Egyptian Crisis aftermath](#), the [Arab Winter](#), and [Terrorism in Egypt](#)

the following year, winning election to the presidency in a [landslide victory](#) described by EU observers as free but not necessarily fair.<sup>[5]</sup> Nonetheless, Sisi's election was widely recognized, and the political situation has largely stabilized since he officially took power; however, some protests have continued despite a government crackdown. The crisis has also spawned an ongoing [insurgency](#) led by [Ansar Bait al-Maqdis](#) in the [Sinai Peninsula](#), which became increasingly intertwined with the [regional conflict](#) against the [Islamic State of Iraq and the Levant](#) later in 2014.<sup>[6]</sup>

## Wikipedia Pageview Co-burst



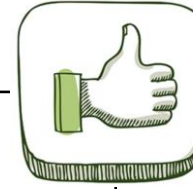




# But How to Encode such Symbolic Knowledge?





- We just need to encode them in an elegant way
- The following position slides are from a working group at Dagstuhl Seminar 2019 (Yoav Goldberg, Heng Ji, Dan Roth, Ido Dagan):

# Why? Because they are complementary



Representation	Pros 	Cons 
<b>Symbolic Semantics</b>	Easy to manipulate	Not trivial to define symbols
	Effective for ordering and composition	Hard to integrate incompatible representations across multiple data modalities or languages
	More explainable to users; Easy for users to specify needs / influence results	Has limited coverage; learning not scalable
		Complexity grows very fast with the number of symbols

# Why? Because they are complementary

Representation	Pros 	Cons 
<b>Distributional Semantics</b>	Easy to extend to multi-media and multi-lingual, can serve as a bridge for cross-media cross-lingual common representation	Not explainable
	provide continuous instead of discrete representation: allow soft decisions and soft matching	Hard to control the output and its properties
	more generalizable	hard/impossible to verify/validate the correctness of the result

1. Use symbolic semantics to represent input structure (e.g., edges in AMR graphs) and distributional semantics to represent nodes (e.g., concept nodes in AMR)
  - ❑ Use in composition based classifiers such as CNN and GCN
  - ❑ Use in composition based methods like Tree/Graph NN.
  - ❑ Match sub-structures using soft-similarity on nodes and hard similarity on edges/structure
2. Combine graph embedding (e.g., knowledge graph embedding, social graph embedding) with text embedding to:
  - ❑ Extend the coverage (Graph embedding provides additional 'neighbors'.)
  - ❑ Validate
  - ❑ Allow flexible matching



3. Use symbolic structure to enforce output structure
  - ❑ inter-dependency between labels and constraints (e.g., we use Bi-LSTM + CRFs for name tagging, where CRFs layer is used to capture the inter-dependency among labels, I-ORG cannot appear after B-PER, Lin et al., ACL2020)
4. Convert distributional semantic representation to symbolic semantic representation (to show to a user / to edit / to perform symbolic inference later)
  - ❑ ground/map the distributions of concepts to taxonomy/ontology/knowledge bases
  - ❑ perform hierarchical clustering on event trigger words based on their distributional semantic representations, and then select the centroid entity or predicate lexical dictionary in Propbank/Ontonotes/VerbNet to assign a name to the event type (Huang et al., 2016)
5. “reason with symbolic, compute with distributional”
  - ❑ Use local prediction/scoring based on distributional representation + global inference based on symbolic representation (Li et al., EMNLP2020)

Seven people **convicted** *last week* in *Vietnam*'s biggest-ever criminal trial, including two former senior government officials, have requested an **appeal** of the verdicts, a court official said Tuesday.

The trial by a *Ho Chi Minh City* court was seen as a litmus test of the communist government's resolve to fight widespread corruption.

The "godfather" of organized crime, *Truong Van Cam*, better known as Nam Cam, was convicted of seven crimes, including murder. He was **sentenced** to face a firing squad, and his lawyer has said he also plans to appeal.

*Hanh*, also a former member of the powerful Communist Party Central Committee, was **convicted** of receiving US\$8,500 in bribes from Nam Cam's family to secure the crime boss' early release from labor camp in 1990s. Hanh was **sentenced** to 10 years in jail.

*Chien* was **convicted** of receiving a stereo set worth 27 million dong (US\$1,750) from Nam Cam's family and **sentenced** to six years in jail.

—> Where was Hanh / Chien sentenced?

# Future Direction 3: Knowledge-element Level Information Consistency Checking

Attacker	Demand	Protester	Time	Location
police	democracy	130K	07-01-2019	Legislative Council Complex

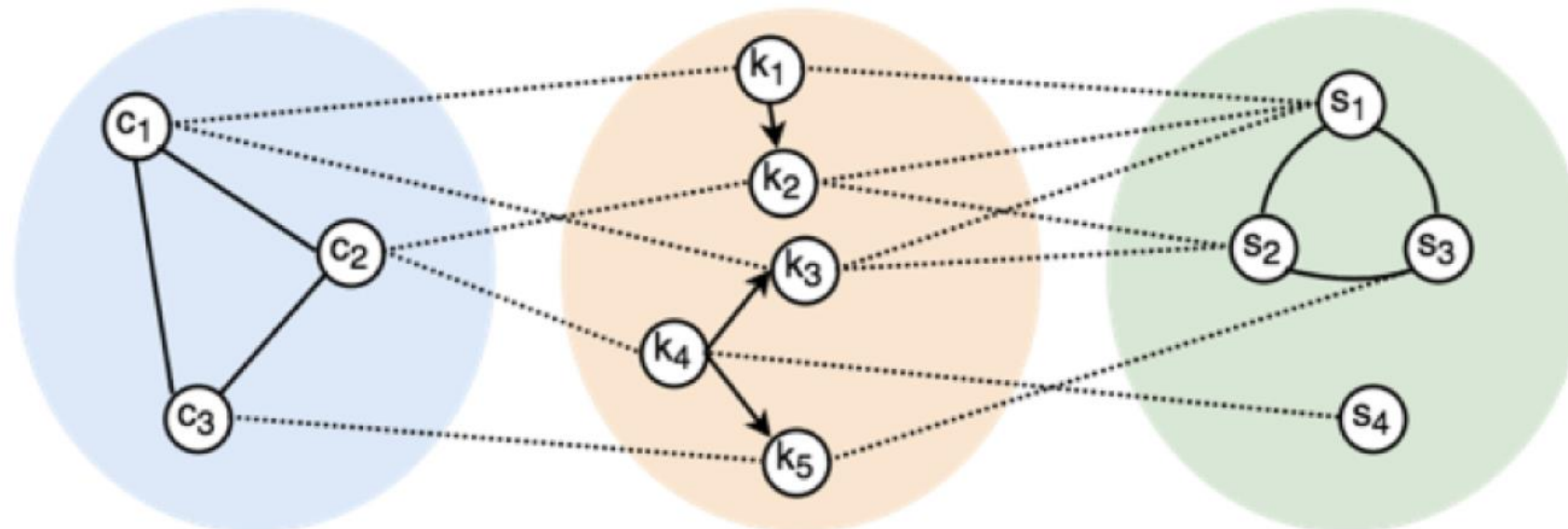


Type	Attacker	Possible Action	Protester	Time	Location
riot	protesters	Sending People's Liberation Army	550K	08-12-2019	Hong Kong International Airport



# Future Direction 3: Knowledge-element Level Information Consistency Checking

Context	Knowledge Element	Source (Narrator/User)
<ul style="list-style-type: none"><li>• Time</li><li>• Location</li><li>• Quality</li><li>• Perplexity</li><li>• Background</li><li>• Censorship</li></ul>	<ul style="list-style-type: none"><li>• Confidence</li><li>• Provenance</li><li>• Saliency</li><li>• Novelty</li><li>• Temporal/Causal/Subevent/ Conflict relations between knowledge elements</li></ul>	<ul style="list-style-type: none"><li>• Emotion</li><li>• Sentiment</li><li>• Aspect</li><li>• Confidence</li><li>• Political Stance</li><li>• Social Norm</li><li>• Language</li><li>• Perceived Control</li><li>• Moral value</li><li>• Power</li><li>• Expertise</li><li>• Deception</li><li>• Data Modality</li><li>• Human Feedback</li></ul>







- Applications
  - ❑ Writing History Book
  - ❑ News Understanding and Recommendation
  - ❑ Disaster Relief
  - ❑ Intelligence Analysis
  - ❑ Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - ❑ External Knowledge Acquisition, Reasoning and Incorporation
  - ❑ Document-Level Event Extraction
  - ❑ Multimedia Information Extraction and Verification
- Grand Vision
- Resources

# Detailed Event-Centric Challenges

# Representing Time

- Representing time of events for the purpose of reasoning about it is still an open problem
- There are other challenges that we still don't have good enough formulations for:
- Semantic clarity
  - What does it mean for a relation to be temporally qualified by some temporal argument T?
    - All time points? At least one? None outside?
      - In 1944 US and Germany were at war;
      - John got married in 1944;
      - Jim was born in 1944.
  - How informative are temporal relations?
    - The great depression started before FDR took office.
    - WWI started before FDR took office.
- Expressiveness
  - Multiple perspectives; Nesting; Negation; Uncertainty

- The lion had a large meal and slept for 24 hours. 
  - When did it eat? Exactly 24 hours?
- The lion didn't sleep after having a large meal.
  - [Negated]
- The lion may have had a large meal before sleeping.
  - [Uncertain]
- If the lion has a large meal, it will sleep for 24 hours.
  - [Hypothetical]
- The lion typically sleeps for 24 hours after having large meals. 
  - [Repetitive]
- After having a large meal, lions may sleep longer.
  - [Generic]

- Mary graduated from college in 1969 (or 1971)
- John believed in 1995 that Mary attended college in the 1960s
- Mary did not attend college in the 1970s (but at some other time)

# Planning & Reasoning about Events

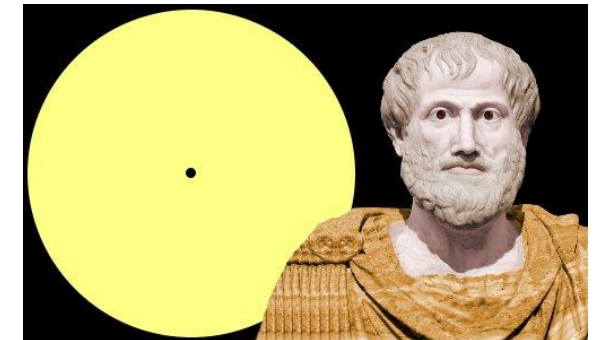
- Reasoning about Events often requires the ability to consider and plan around **implicit events** and reason about **quantities** and **time**.

Will we make it to dinner before the movie?



Did Aristotle have a laptop?

See [Geva et al. TACL'21]



Mayor Rahm Emanuel now has raised more than \$10 million toward his bid for a third term – more than five times the total raised by his 10 challengers combined, campaign finance records show.





# Grounding

- There is a need to **ground** in order to reason about events.

- Temporally, Spatially

- Multimodal: Where is this event? (and When did it happen)?

- Manchester Bomber Believed Muslims Were Mistreated, Sought Revenge

- The visit to Berlin took place just after the collapse of the wall.

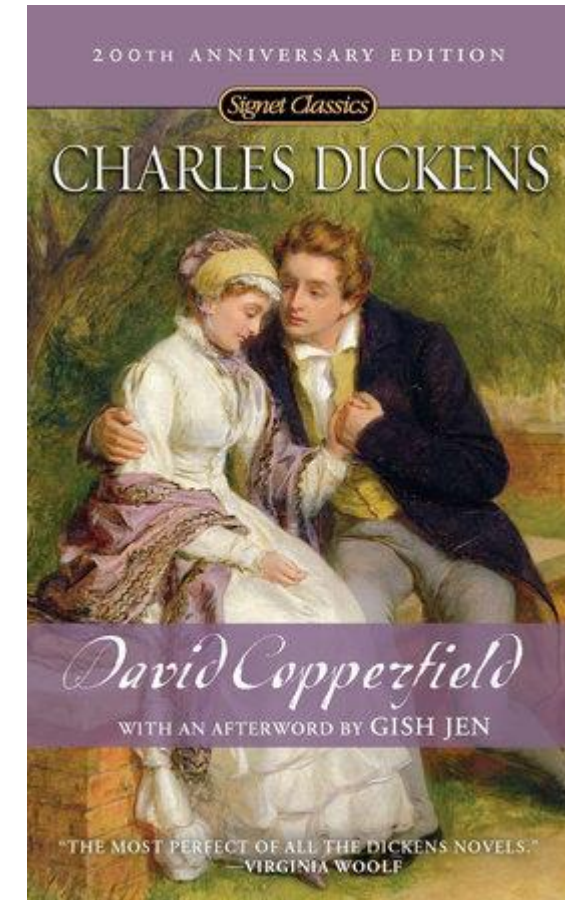


- What is the event?
- When/where did it happen?

- What is the event?
- When/where did it happen?

# High Level Event-Centric Challenges

- How can we support readers in better ways?
  - The novel features the character [David Copperfield](#), his journey of change and growth from infancy to maturity, as many people enter and leave his life and he passes through the stages of his development. (**Fiction, and you know it**)
  - London and England in the 19-th century; socio-economic state, child exploitation; schools, prisons, emigration to Australia (**True historical facts**)
- Currently, we can't even tell the difference
  - What are the computational tasks we should think about?



- The NYT put together a reconstruction video of the killing of Breonna Taylor.
  - Reconstruction is done by
    - Analyzing written evidence
    - Transcripts from interviews with witnesses, policemen, neighbors
    - Analysis of physical evidence collected in the scene
  - It took weeks to produce
- Can this be automated to form a reconstruction of a (chain of) events?
  - Ideally, this should be automated as a support tool for analysts
    - Requires multimodal capabilities
    - Understanding and reasoning about quantities & geometry
    - Reasoning about uncertainty
    - Reasoning about (possible) conflicting events/reports
      - Notion of multiple perspective & trustworthiness



Multimodal + 3D reconstruction:

<https://www.nytimes.com/video/us/100000007348445/breonna-taylor-death-cops.html>



- Joint Neural Information Extraction system
  - <http://blender.cs.illinois.edu/software/oneie/>
- Multimedia Event Extraction system and new benchmark with annotated data set
  - GitHub: [https://github.com/GAIA-AIDA/uiuc\\_ie\\_pipeline\\_fine\\_grained](https://github.com/GAIA-AIDA/uiuc_ie_pipeline_fine_grained)
  - Text IE DockerHub: <https://hub.docker.com/orgs/blendernlp/>
  - Visual IE repositories: <https://hub.docker.com/u/dannapierskitoptal>
- COVID-19 knowledge graph construction and drug repurposing generation
  - <http://blender.cs.illinois.edu/covid19/>

- [1] Ying Lin, Heng Ji, Fei Huang, Lingfei Wu. 2020. A Joint End-to-End Neural Model for Information Extraction with Global Features. Proc. ACL2020.
- [2] Manling Li, Alireza Zareian, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji, Shih-Fu Chang. 2020. Cross-media Structured Common Space for Multimedia Event Extraction. Proc. ACL2020.
- [3] Manling Li, Alireza Zareian, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, Heng Ji, Shih-Fu Chang, Clare Voss, Daniel Napierski and Marjorie Freedman. 2020. GAIA: A Fine-grained Multimedia Knowledge Extraction System. Proc. ACL2020 Demo [Best Demo Paper Award].
- [4] Heng Ji and Clare Voss. 2020. Low-resource Event Extraction based on Share-and-Transfer Invited Book Chapter for Events and Stories, Cambridge Press.
- [5] Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel and Clare Voss. 2018. Proc. The 56th Annual Meeting of the Association for Computational Linguistics (ACL2018).
- [6] Hongming Zhang, Haoyu Wang, Dan Roth. 2020. Unsupervised Label-aware Event Trigger and Argument Classification. arXiv:2012.15243.
- [7] Seth Ebner, Patrick Xia, Ryan Culkin, Kyle Rawlins and Benjamin Van Durme. 2020. Multi-Sentence Argument Linking. Proc. the 58th Annual Meeting of the Association for Computational Linguistics.
- [8] Xinya Du and Claire Cardie. 2020. Event Extraction by Answering (Almost) Natural Questions. Proc. the 58th Annual Meeting of the Association for Computational Linguistics.

# Event-Event Relation References



1. Algorithms for scoring coreference chains. Bagga & Baldwin, 1998.
2. Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms. Collins, 2002.
3. A Linear Programming Formulation for Global Inference in Natural Language Tasks. Roth & Yih, 2004.
4. Inducing temporal graphs. Bramsen et al., 2006.
5. Refining Event Extraction through Cross-document Inference. Ji & Grishman, 2008.
6. Unsupervised Learning of Narrative Event Chains. Chambers & Jurafsky, 2008.
7. Jointly combining implicit constraints improves temporal ordering. Chambers & Jurafsky, 2008.
8. Sentence Level Event Detection and Coreference Resolution. Naughton, 2009.
9. Graph-based event coreference resolution. Chen & Ji, 2009.
10. Using document level cross-event inference to improve event extraction. Liao & Grishman, 2010.
11. Evaluation Metrics For End-to-End Coreference Resolution Systems. Cai & Strube, 2010.
12. Predicting globally-coherent temporal structures from texts via endpoint inference and graph decomposition. Denis & Muller, 2011.
13. Minimally supervised event causality identification. Do et al., 2011.
14. Temporal evaluation. UzZaman & Allen, 2011.
15. Joint inference for event timeline construction. Do et al., 2012.
16. Semantic Relations between Events and their Time, Locations and Participants for Event Coreference Resolution. Cybulska & Vossen, 2013.
17. TEMPEVAL-3: Evaluating time expressions, events, and temporal relations. UzZaman et al., 2013.
18. HiEve: A Corpus for Extracting Event Hierarchies from News Stories. Glavas et al., 2014.
19. CATENA: CAusal and TEMPoral relation extraction from NATural language texts. Mirza & Tonelli, 2016.
20. Event Detection and Co-reference with Minimal Supervision. Peng et al., 2016.
21. Which Coreference Evaluation Metric Do You Trust? A Proposal for a Link-based Entity Aware Metric. Moosavi & Strube, 2016.
22. Story Comprehension for Predic3ng What Happens Next. Chaturvedi et al., 2017.
23. A Structured Learning Approach to Temporal Relation Extraction. Ning et al., 2017.
24. Joint Reasoning for Temporal and Causal Relations. Ning et al., 2018.
25. A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., 2018.
26. Improving Temporal Relation Extraction with a Globally Acquired Statistical Resource. Ning et al., 2018.
27. KnowSemLM: A Knowledge Infused Semantic Language Model. Peng et al., 2019.
28. Multilingual Entity, Relation, Event and Human Value Extraction. Li et al., 2019.
29. Joint Event and Temporal Relation Extraction with Shared Representations and Structured Prediction. Han et al., 2019.
30. An Improved Neural Baseline for Temporal Relation Extraction. Ning et al., 2019.
31. Fine-Grained Temporal Relation Extraction. Vashishtha et al., 2019.
32. Constrained Learning for Event-Event Relation Extraction. Wang et al., 2020.
33. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. Ning et al., 2020.

- [1] Zack, et al. Human brain activity time-locked to perceptual event boundaries. *Nature neuroscience*, 4(6):651–655.
- [2] Chambers and Jurafsky. Unsupervised learning of narrative event chains. *ACL*, 2008
- [3] Radinsky and Horvitz. Mining the Web to Predict Future Events. *WSDM*, 2013
- [4] Berant, et al. Modeling Biological Processes for Reading Comprehension. *EMNLP*, 2014
- [5] Chaturvedi, et al. Story comprehension for predicting what happens next. *EMNLP*, 2017
- [6] Rashkin, et al. Event2Mind: Commonsense Inference on Events, Intents, and Reactions. *ACL*, 2108
- [7] Liu, et al. Automatic event salience identification. *EMNLP*, 2018
- [8] Zhukov et al. Cross-task weakly supervised learning from instructional videos. *CVPR*, 2019
- [9] Zhang, et al. Analogous Process Structure Induction for Sub-event Sequence Prediction. *EMNLP*, 2020
- [10] Chen, et al. “What are you trying to do?” Semantic typing of event processes. *CoNLL*, 2020
- [11] Ning et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. *EMNLP*, 2020
- [12] Lyu, et al. Reasoning about Goals, Steps, and Temporal Ordering with WikiHow. *EMNLP*, 2020
- [13] Jindai, et al. Is Killed More Significant than Fled? A Contextual Model for Salient Event Detection. *COLING*, 2020
- [14] Fried, et al. Learning to Segment Actions from Observation and Narration. *ACL*, 2020
- [15] Zhang, et al. Diagnostic Prediction with Sequence-of-sets Representation Learning for Clinical Events. *AIME*, 2020
- [16] Surís, et al. Learning the Predictability of the Future. *arXiv:2101.01600*, 2021



- [1] WSC: Hector Levesque, Ernest Davis, and Leora Morgenstern. The winograd schema challenge. KRR 2012.
- [2] *ConceptNet*: Hugo Liu and Push Singh, ConceptNet - a practical commonsense reasoning tool-kit, BTTJ, 2004
- [3] ATOMIC: Maarten Sap, Ronan LeBras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, Yejin Choi, ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. AACL 2019
- [4] COMET: Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. COMET: commonsense transformers for automatic knowledge graph construction. ACL 2019.
- [5] LAMA: Fabio Petroni, Tim Rocktaschel, Sebastian Riedel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, and Alexander H. Miller. Language models as knowledge bases? EMNLP 2019.
- [6] ASER: Hongming Zhang, Xin Liu, Haojie Pan, Yangqiu Song, and Cane Wing-Ki Leung. ASER: A Large-scale Eventuality Knowledge Graph. WWW 2020.
- [7] TransOMCS: Hongming Zhang, Daniel Khashabi, Yangqiu Song, and Dan Roth. TransOMCS: From Linguistic Graphs to Commonsense Knowledge. International Joint Conference on Artificial Intelligence (IJCAI). 2020.
- [8] KnowlyWood: Niket Tandon, Gerard de Melo, Abir De, and Gerhard Weikum. 2015. Knowlywood: Mining Activity Knowledge From Hollywood Narratives. CIKM 2015.
- [9] Manling Li, Qi Zeng, Kyunghyun Cho, Heng Ji, Jonathon May, Nathanael Chambers, Clare Voss. Connecting the Dots: Event Graph Schema Induction with Path Language Mod