EventKG: A Multilingual Event-Centric Temporal Knowledge Graph

eventkg.l3s.uni-hannover.de

Simon Gottschalk, Elena Demidova
gottschalk@L3S.de, demidova@L3S.de
Leibniz University Hannover (Germany)

L3S Research Center
The wedding of Harry and Meghan had threatened to go awry all week – the Markle Debacle palace insiders were calling it ...

https://www.telegraph.co.uk/women/life/walt-disney-could-not-have-dreamt-better-princess-meghan-markle/

Wedding of Prince Harry and Meghan Markle

point in time: 19 May 2018

Harry, spouse, Meghan, Duchess of Sussex

start time: 19 May 2018
What is an Event?

- Wikipedia: “Event may refer to … “
- Merriam-Webster: “Something that happens“
- Oxford: “A thing that happens or takes place, especially one of importance.”
- TDT\(^1\): “Something that happens at a particular time and place”

<table>
<thead>
<tr>
<th>Text</th>
<th>Name</th>
<th>Relation</th>
<th>Location</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>The wedding of … takes place at …</td>
<td>Wedding of Prince Harry and Meghan Markle</td>
<td>(&lt;Prince Harry of Wales, spouse, Meghan Markle&gt;)</td>
<td>St. George's Chapel</td>
<td>19 May 2018</td>
</tr>
</tbody>
</table>

Current State: Events in Knowledge Graphs

- not all events are identified as such
  - UK European Union membership referendum, 2016
  - Wikidata: instance of referendum $\rightarrow$ voting $\rightarrow$ event
  - DBpedia (en): type owl:Thing

- many important events are not covered
  - April 27 – Kim Jong-un crosses into South Korea to meet with President Moon Jae-in.

- many important entity-event relations are missing
  $<$Angela Merkel, ???, European migrant crisis$>$
Current State: Events in Knowledge Graphs

- limited set of temporal relations

- disagreement between knowledge graphs
  - locations of the Battle of Gibraltar:

<table>
<thead>
<tr>
<th></th>
<th>DBpedia (en)</th>
<th>DBpedia (pt)</th>
<th>Wikidata</th>
<th>YAGO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strait of Gibraltar</td>
<td></td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cape Trafalgar</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Atlantic Ocean</td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
</tr>
</tbody>
</table>
Current State: Events in Knowledge Graphs

- problems of uniquely identify events in news
  - “Reducing Plastic Use as a Family Is Easy. Here's How.”

- very fine-grained NLP-based events → still important?

- events identified through peak detection
  - news publication time ≠ event happening time
  - limited to the time span of the source
  - Google Trends for “Obama”:

Google Trends started in 2004
EventKG: Goal

- provide a knowledge graph that
  - contains temporal and event-centric information
    - events
    - temporal relations
  - integrates data from different sources into a common schema
  - fuses data coming from disagreeing sources
  - interconnects entities and events
  - is multilingual
  - is openly accessible

<table>
<thead>
<tr>
<th></th>
<th>Wikidata</th>
<th>YAGO</th>
<th>DBpedia EN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events with Location</td>
<td>11.70%</td>
<td>26.61%</td>
<td>6.21%</td>
</tr>
<tr>
<td>Events with Time</td>
<td>33.00%</td>
<td>39.02%</td>
<td>7.00%</td>
</tr>
</tbody>
</table>

Coverage of Event Information in Different Sources
A Multilingual Event-Centric Temporal Knowledge Graph

- [eventkg.l3s.uni-hannover.de](http://eventkg.l3s.uni-hannover.de)

- 690K events and 2.3M temporal relations
  - 2018 FIFA World Cup
  - “The Space Shuttle Challenger is launched on its maiden voyage”
  - <Jennifer Aniston, married to, Brad Pitt, [2000-07-29,2005-10-02]>

- integrates data in five languages: EN, FR, DE, RU, PT
- provides provenance information
- publicly accessible: code, data and query endpoint
Events and Relations in EventKG

- **events in EventKG:**
  - no strict constraints on the representation of events
  - events can have a name, description, time, location, ...

- **relations in EventKG:**
  - temporal relations
  - relations between entities/events and events
  - link information:
    - how often are entities and events interlinked?
EventKG Schema

- EventKG builds upon the Simple Event Model (SEM)

Extensions of SEM:
- Existing vocabulary
- EventKG vocabulary
EventKG Representation

- EventKG is provided as an RDF store with N-Quads
- namespaces
  - eventKG-r: resource
  - eventKG-s: schema
  - eventKG-g: graph
- statements come with a named graph for provenance information
  - eventKG-g:yago
  - eventKG-g:wikipedia_fr
  - eventKG-g:eventKG
  - ...

Example

eventKG-r: event_170890 rdfs: label "Perahera"@fr eventKG-g: wikipedia_fr .
EventKG Schema: Example

- the participation of Barack Obama in his second inauguration as US president in 2013 in EventKG
EventKG: A Multilingual Event-Centric Temporal Knowledge Graph

Web Science – Investigating the Future of Information and Communication

Extraction Pipeline

Input

{EN, FR, DE, PT, RU}

Identification and Extraction of Events

identify and extract events

Extraction of Event and Entity Relations

extract relations

Events

Events Relations Entities

Integration

integrate events and entities

Fusion

fuse times and locations

Output

EventKG
Event Identification

- use type hierarchy and category names
  - UK European Union membership referendum, 2016
    - Wikidata: instance of referendum → voting → event
    - DBpedia_{EN}: type owl:Thing
  - Martial arts
    - YAGO: self defence → protection → activity → act → event
  - Super Bowl XXVIII
    - Wikipedia_{EN}: January 1994 sports events
  - Volcanic impacts on the oceans
    - Wikipedia_{EN}: Volcanic events

- Manual evaluation on 100 events per source (Precision):

<table>
<thead>
<tr>
<th></th>
<th>Wikidata</th>
<th>DBpedia_{EN}</th>
<th>DBpedia_{RU}</th>
<th>DBpedia_{PT}</th>
<th>Wikipedia_{EN}</th>
<th>Wikipedia_{RU}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>96%</td>
<td>100%</td>
<td>100%</td>
<td>98%</td>
<td>94%</td>
<td>88%</td>
</tr>
</tbody>
</table>
Sources: Knowledge Graphs

- identify events using type hierarchies in DBpedia and Wikidata
- extract existence times, sub events, locations, ...

About: United States presidential election, 2016

Barack Obama (Q76)

spouse

Michelle Obama

start time

place of marriage

Trinity United Church of Christ
Barack Obama

From Wikipedia, the free encyclopedia

**Barack Hussein Obama II** (/bəˈrɑːk hjuːˈsɛn əˈbaːmə/ (listen);[1] born August 4, 1961) is an American politician who served as the 44th President of the United States from January 20, 2009 to January 20, 2017. The first African American to

In March, Obama's Treasury Secretary, **Timothy Geithner**, took further steps to manage the **financial crisis**, including introducing the **Public-Private Investment Program for Legacy Assets**, which contains provisions for buying up to two trillion dollars in

Barack Obama → Financial Crisis
Timothy Geithner ↔ Financial Crisis

**1965 in music**

List of notable events in music that took place in the year 1965.

**Events** [edit]

- **January 4** – Fender Musical Instruments Corporation is sold to CBS for $13 million.
- **January 12** – **Hullabaloo** premieres on NBC. The first show included performances by The New Christy Minstrels, comedian Woody Allen, actress Joey Heatherton and a segment from London in which Brian Epstein introduces The Zombies and Gerry & the Pacemakers.
Fusion of Times and Locations

- time fusion
  - majority voting
    - in case of equality
      - prefer times with fine granularity
      - prefer higher trusted sourcers

- event location fusion
  - union of locations from the different sources
  - exploit the so:containedInPlace relations to reduce this set
    - example: {Crete, Greece, Athens} → {Crete, Athens}

- fused times and locations are stored in the EventKG named graph
Number of Events and Relations in EventKG

EventKG provides more than 690K events and 2.3M temporal relations.

- Most events are identified in Wikidata and English Wikipedia.

<table>
<thead>
<tr>
<th></th>
<th>#Events</th>
<th>Known time</th>
<th>Known location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events from KGs</td>
<td>322,669</td>
<td>163,977</td>
<td>84,304</td>
</tr>
<tr>
<td>Events from semi-structured sources</td>
<td>367,578</td>
<td>362,064</td>
<td>-</td>
</tr>
<tr>
<td>Relations</td>
<td>88,473,111</td>
<td>2,331,370</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wiki-data</th>
<th>DBpedia</th>
<th>Wikipedia event lists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EN</td>
<td>FR</td>
</tr>
<tr>
<td>EN</td>
<td>266,198</td>
<td>60,307</td>
</tr>
</tbody>
</table>

Number of identified events per source.
Event Types

- EventKG covers a broad range of different events, including sports events, political events and military conflicts.

<table>
<thead>
<tr>
<th>Wikidata</th>
<th>EN</th>
<th>FR</th>
<th>DE</th>
<th>RU</th>
<th>PT</th>
</tr>
</thead>
<tbody>
<tr>
<td>dbo:type</td>
<td>Season</td>
<td>Military Conflict</td>
<td>Sports Event</td>
<td>Tennis Tournament</td>
<td>Military Conflict</td>
</tr>
<tr>
<td>Events, %</td>
<td>11.37%</td>
<td>6.31%</td>
<td>21.86%</td>
<td>33.00%</td>
<td>11.87%</td>
</tr>
</tbody>
</table>

Most frequent event type per source
Time and Location Integration

- integrating times and locations of events from different sources leads to a higher coverage of events with temporal and spatial information

<table>
<thead>
<tr>
<th></th>
<th>EventKG</th>
<th>Wikidata</th>
<th>YAGO</th>
<th>EN</th>
<th>FR</th>
<th>DE</th>
<th>RU</th>
<th>PT</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Events with</td>
<td>322,669</td>
<td>322,669</td>
<td>222,325</td>
<td>214,556</td>
<td>78,527</td>
<td>62,971</td>
<td>47,304</td>
<td>35,682</td>
</tr>
<tr>
<td>Location (L)</td>
<td>26.13%</td>
<td>11.70%</td>
<td>26.61%</td>
<td>6.21%</td>
<td>8.32%</td>
<td>4.03%</td>
<td>10.60%</td>
<td>6.15%</td>
</tr>
<tr>
<td>Time (T)</td>
<td>50.82%</td>
<td>33.00%</td>
<td>39.02%</td>
<td>7.00%</td>
<td>17.21%</td>
<td>2.00%</td>
<td>1.35%</td>
<td>0.08%</td>
</tr>
</tbody>
</table>

Comparison of the event representation completeness in the source-specific named graphs
Language-dependent Event Ranking

- EventKG provides link count information for its events extracted from five Wikipedia language editions.
- Example query: What are the most linked events in the English and the Russian Wikipedia?

<table>
<thead>
<tr>
<th>Rank</th>
<th>Event (en)</th>
<th>#Links (en)</th>
<th>Event (ru)</th>
<th>#Links (ru)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>World War II</td>
<td>189,716</td>
<td>World War II</td>
<td>25,295</td>
</tr>
<tr>
<td>2</td>
<td>World War I</td>
<td>99,079</td>
<td>World War I</td>
<td>22,038</td>
</tr>
<tr>
<td>3</td>
<td>American Civil War</td>
<td>37,672</td>
<td>October Revolution</td>
<td>7,533</td>
</tr>
<tr>
<td>4</td>
<td>FA Cup</td>
<td>20,640</td>
<td>Russian Civil War</td>
<td>7,093</td>
</tr>
</tbody>
</table>
Sub Event Search

- sub events of the Second World War in February 1941

<table>
<thead>
<tr>
<th>Start Time</th>
<th>Description (EN)</th>
<th>DBpedia link</th>
</tr>
</thead>
<tbody>
<tr>
<td>1941-02-12</td>
<td>Erwin Rommel arrives in Tripoli.</td>
<td></td>
</tr>
<tr>
<td>1941-02-17</td>
<td>The Battle of Trebeshina or the Battle of Mal Trebeshinë, was a series of engagements fought between the Greek and Italian armies in south-eastern Albania during the Greco-Italian War.</td>
<td><a href="http://dbpedia.org/resource/Battle_of_Trebeshina">http://dbpedia.org/resource/Battle_of_Trebeshina</a></td>
</tr>
</tbody>
</table>
Event Search

- Which revolts took place in Crete?

```sql
SELECT DISTINCT ?description {
  ?crete owl:sameAs dbr:Crete .
  FILTER regex(?description, "revolt", "i") .
}
```

- 1365: “A revolt against the Venetian rulers in Crete fails.”
- 1905-03-23: “Theriso revolt breaks out on Crete when about 1,500 men led by Eleftherios Venizelos meet ... to challenge the island's authoritarian government and press for its unification with Greece.”
- ...
### Relation Timeline

- Which temporal relations involve Barack Obama as the subject?

<table>
<thead>
<tr>
<th>Property</th>
<th>Object</th>
<th>Start time</th>
<th>End time</th>
</tr>
</thead>
<tbody>
<tr>
<td>residence</td>
<td>Tebet, South Jakarta</td>
<td>1967-01-01</td>
<td>1970-12-31</td>
</tr>
<tr>
<td>residence</td>
<td>Menteng</td>
<td>1970-01-01</td>
<td>1971-12-31</td>
</tr>
<tr>
<td>educated at</td>
<td>State Elementary School Menteng 01</td>
<td>1970-01-01</td>
<td>1971-12-31</td>
</tr>
<tr>
<td>residence</td>
<td>Honolulu</td>
<td>1971-01-01</td>
<td>1979-12-31</td>
</tr>
<tr>
<td>educated at</td>
<td>Punahou School</td>
<td>1971-01-01</td>
<td>1979-12-31</td>
</tr>
</tbody>
</table>
EventKG+TL: Creating Cross-Lingual Timelines from an Event-Centric Knowledge Graph

Simon Gottschalk, Elena Demidova. ESWC 2018 (Poster & Demo Session)

- an overview of events related to a query entity over time
  - pie chart size: the language independent event relevance
  - colored slices: the ratio of the relevance in a language

Example: Brexit
Thank you!
Questions?

eventkg.l3s.uni-hannover.de

Simon Gottschalk, Elena Demidova
gottschalk@L3S.de, demidova@L3S.de
Leibniz University Hannover (Germany)

L3S Research Center