Refining Imprecise Spatio-temporal Events: A Network-based Approach

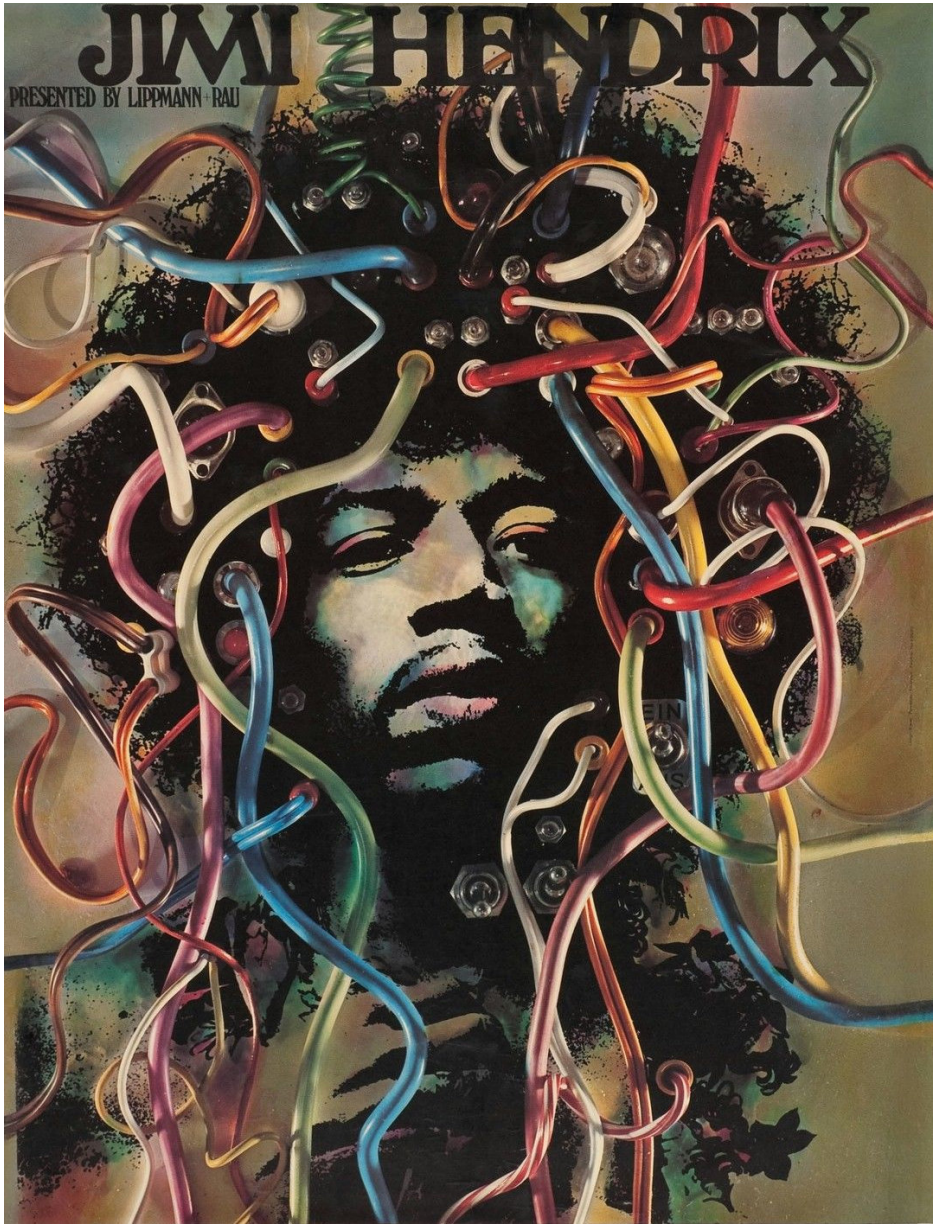
Andreas Spitz[1], **Johanna Geiß**[1], **Michael Gertz**[1], **Stefan Hagedorn**[2] **and Kai-Uwe Sattler**[2]

[1] Heidelberg University, Institute of Computer Science
Database Systems Research Group, Heidelberg

[2]Technical University of Ilmenau
Databases and Information Systems Group, Ilmenau

{spitz, geiss, gertz}@informatik.uni-heidelberg.de
{stefan.hagedorn, kus}@tu-ilmenau.de

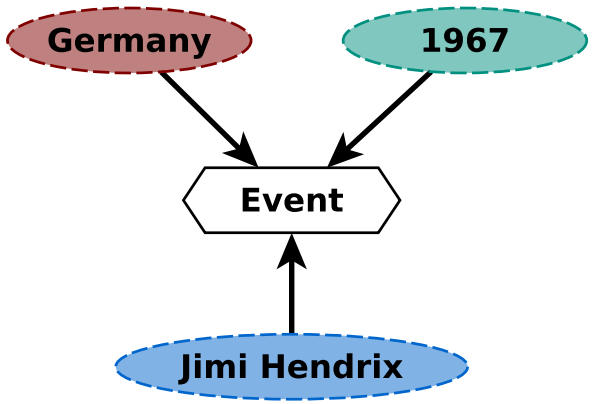
10th GIR Workshop
San Francisco, HALLOWEEN, 2016


JIMI HENDRIX
PRESENTED BY LIPPMANN · RAU

What is an Event?

Event

"The Jimi Hendrix Experience toured Germany in 1967."



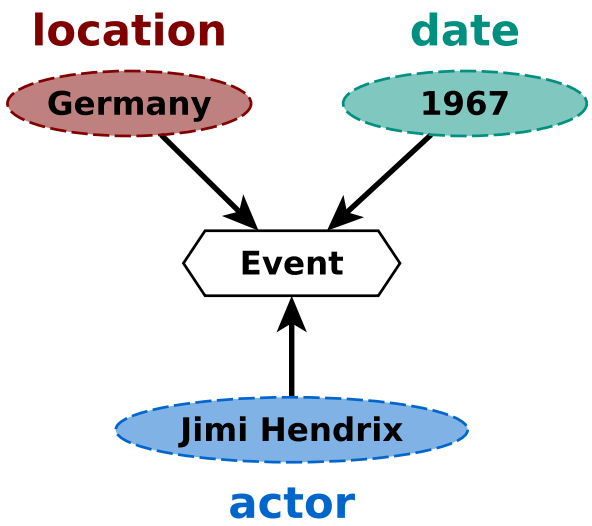
What is an Event?

Event

"The Jimi Hendrix Experience toured Germany in 1967."

Definition: Event

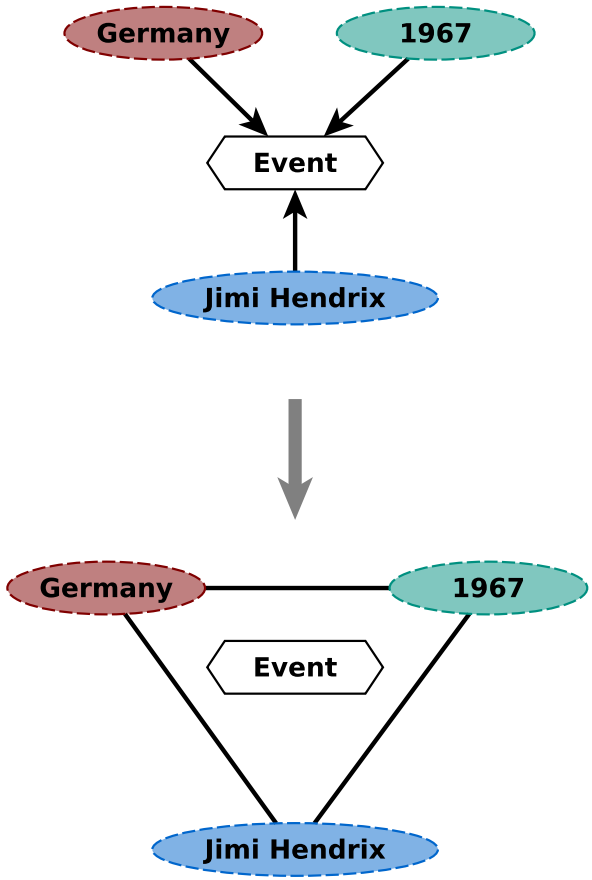
"Something that happens at a given **place** *and* **time** *between a group of* **actors**." [All02]



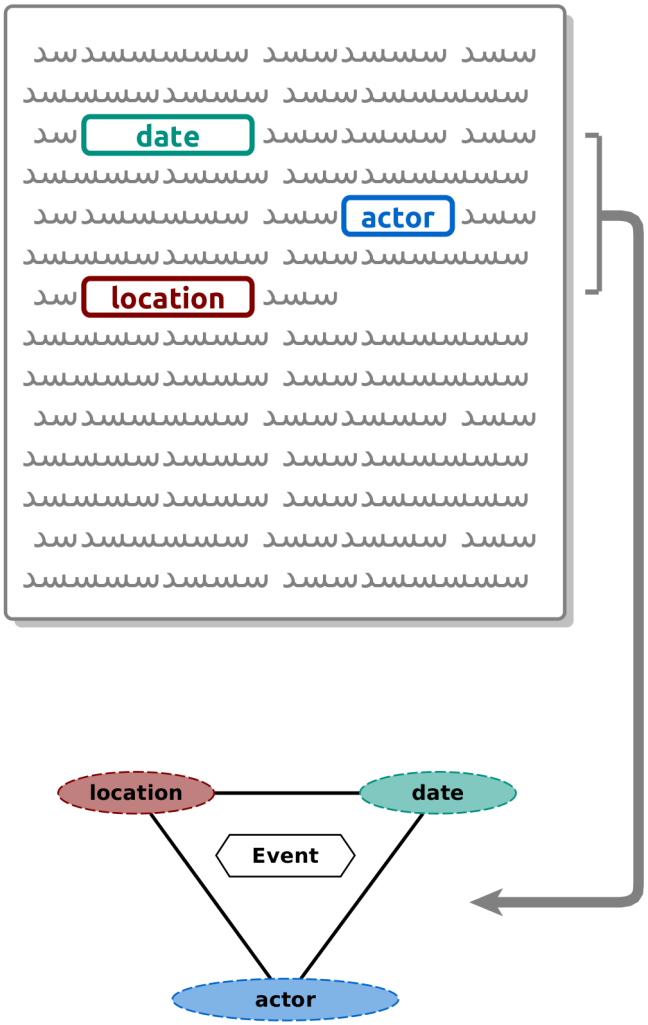
Event Triangles

Intuition:

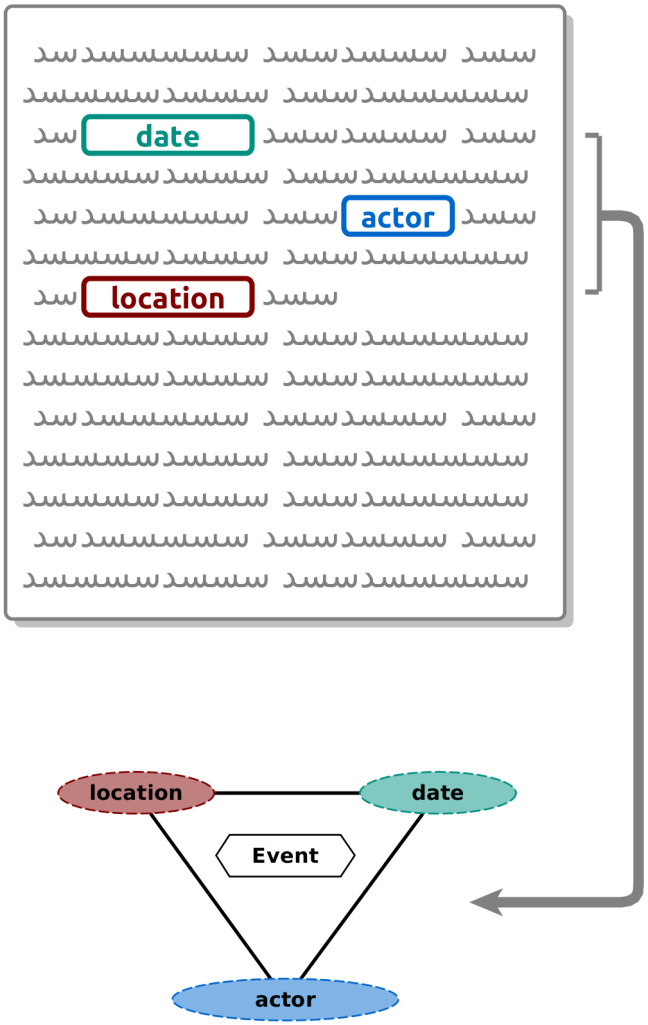
- Events correspond to triangular structures of entities
- Participating entities can be used to extract events
- Linguistic, sentence-based event detection can improve results



Event Extraction



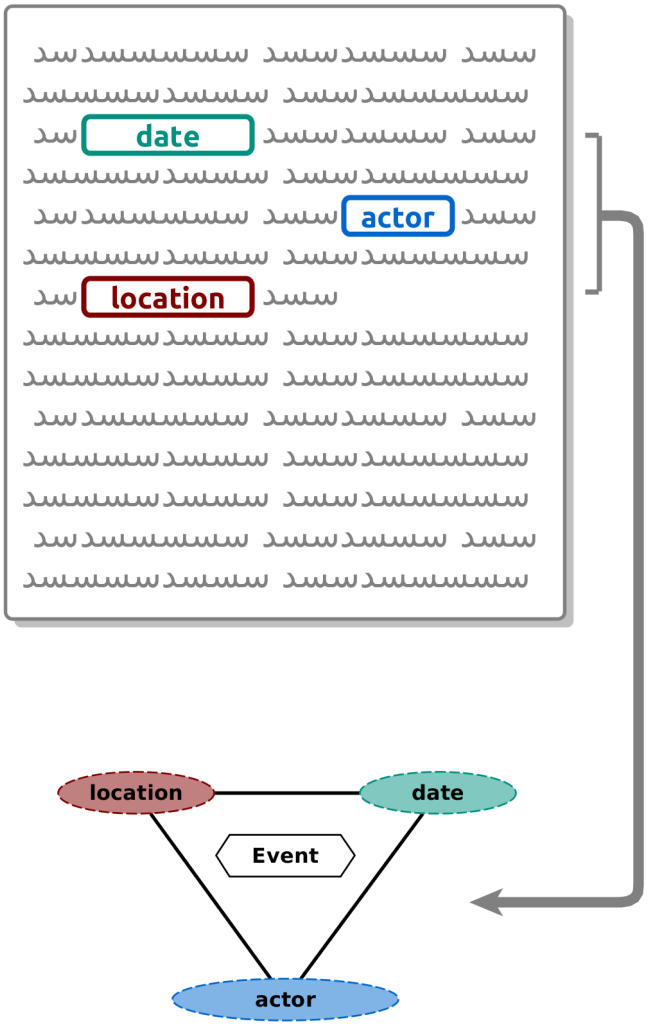
Event Extraction



Generating weights for events:

- Similarity of entities [GSG15]:
  $\phi(r, r') := \exp\left(-\frac{dist(r,r')}{2}\right)$

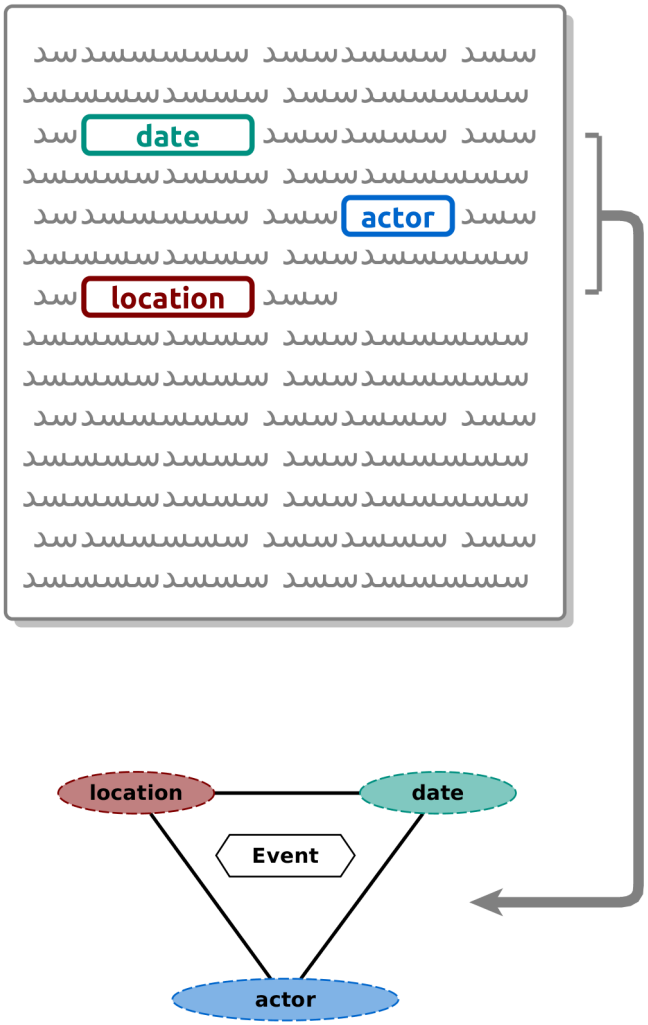
Event Extraction



Generating weights for events:

- Similarity of entities [GSG15]:
  $\phi(r, r') := \exp\left(-\frac{dist(r,r')}{2}\right)$
- Weight of individual events:
  $\omega(e_i) = \min\{\phi(l,t), \phi(l,a), \phi(t,a)\}$

Event Extraction



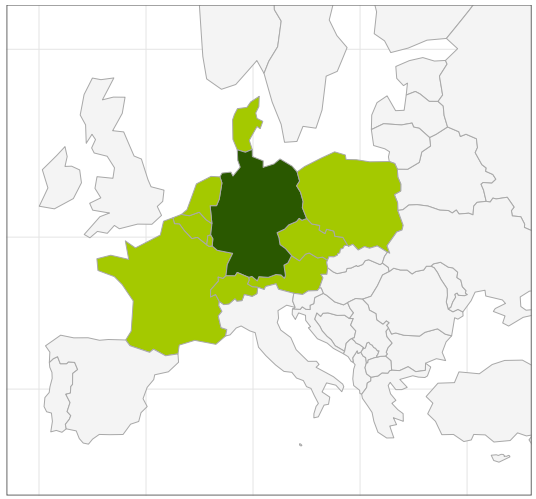
Generating weights for events:

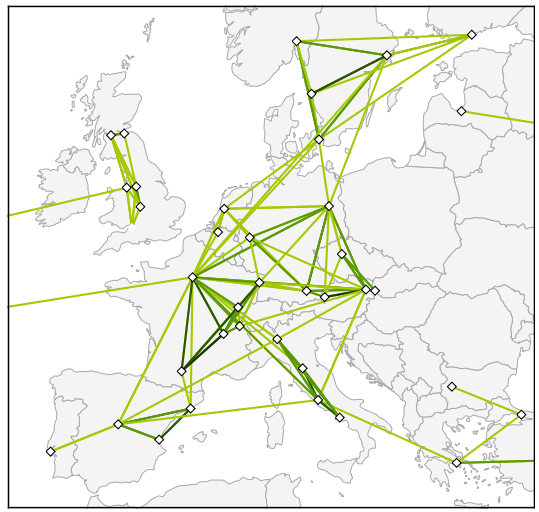
- Similarity of entities [GSG15]:
  $\phi(r, r') := \exp\left(-\frac{dist(r,r')}{2}\right)$
- Weight of individual events:
  $\omega(e_i) = \min\{\phi(l,t), \phi(l,a), \phi(t,a)\}$
- Weight of aggregated events:
  $\omega(e) := \sum_{i=1}^{k} \omega(e_i)$

Geographic Background Network

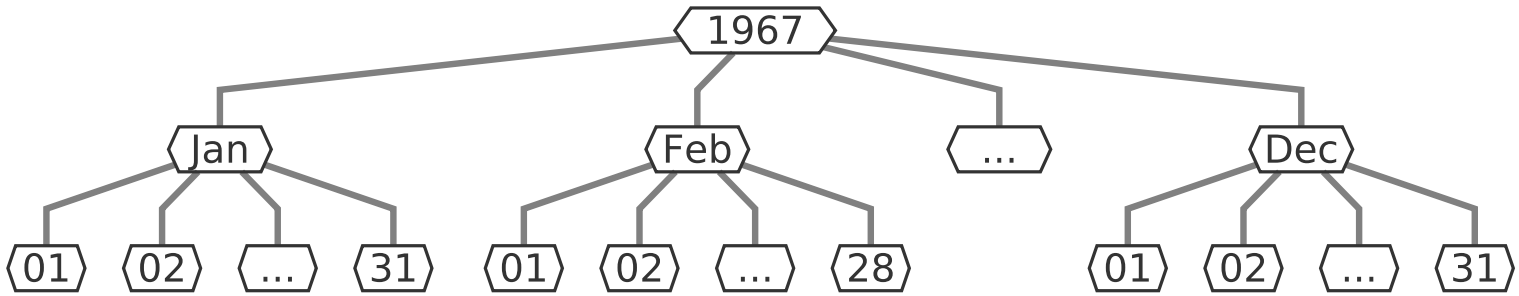
Locations can be connected in a location network G_L based on:

- Geographic containment
- Geographic neighbourhood
- Reachability
- Semantic similarity
- Context-dependent relations





Temporal Background Network



Dates are inherently connected in a temporal network G_T:

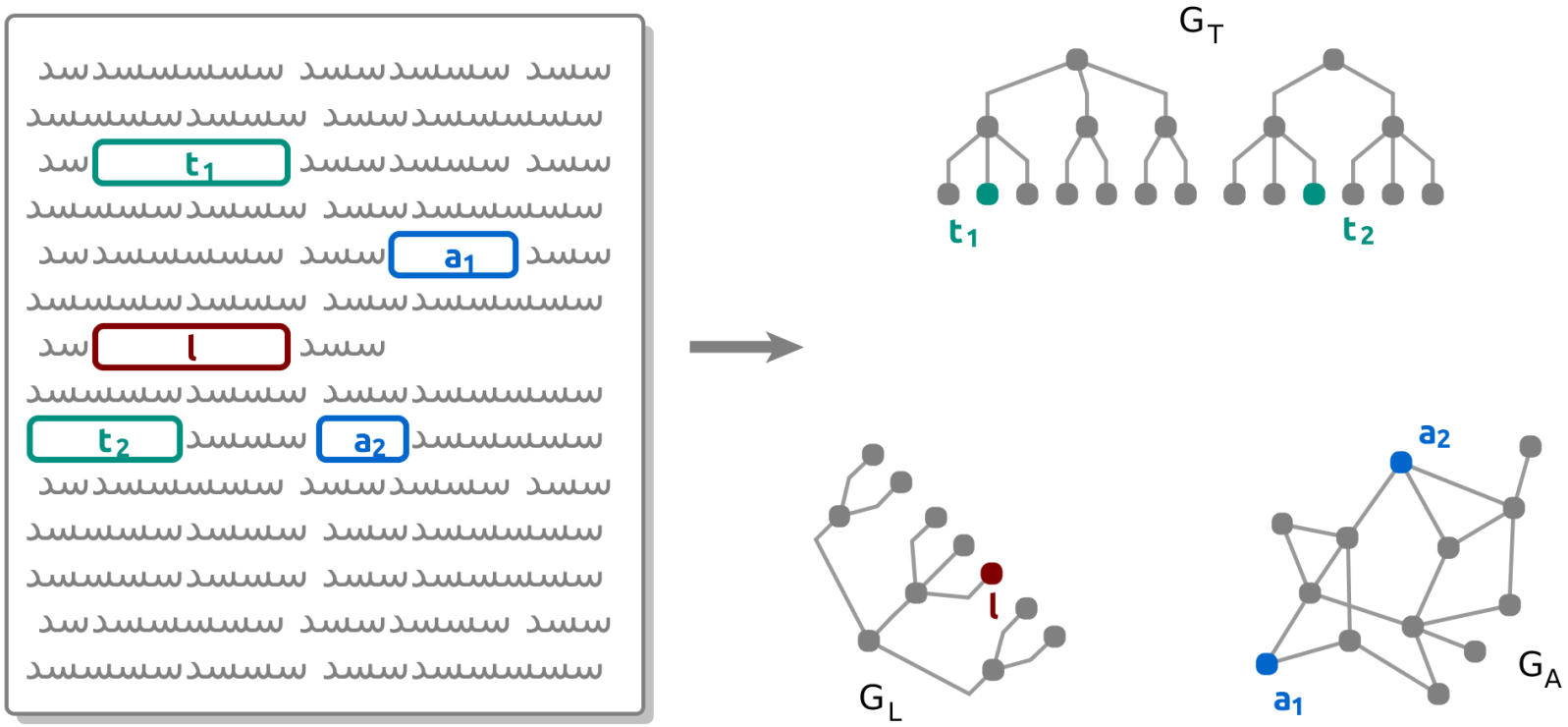
- Temporal containment is straightforward
- Heterogeneous granularity levels are possible

Social Background Network

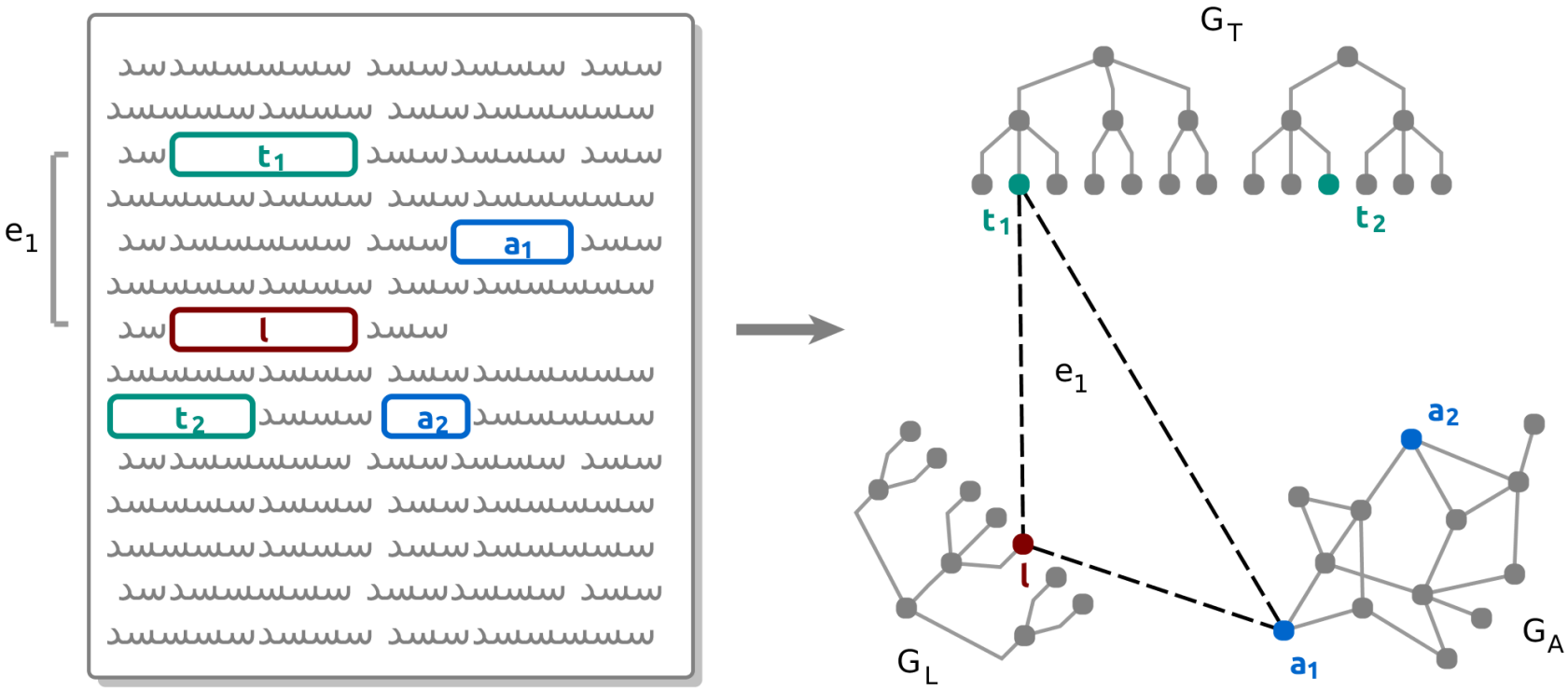
Actors form a social network G_A and are connected by:

- Acquaintance or Relation
- Collaboration
- Organization membership
- Context-dependent similarities

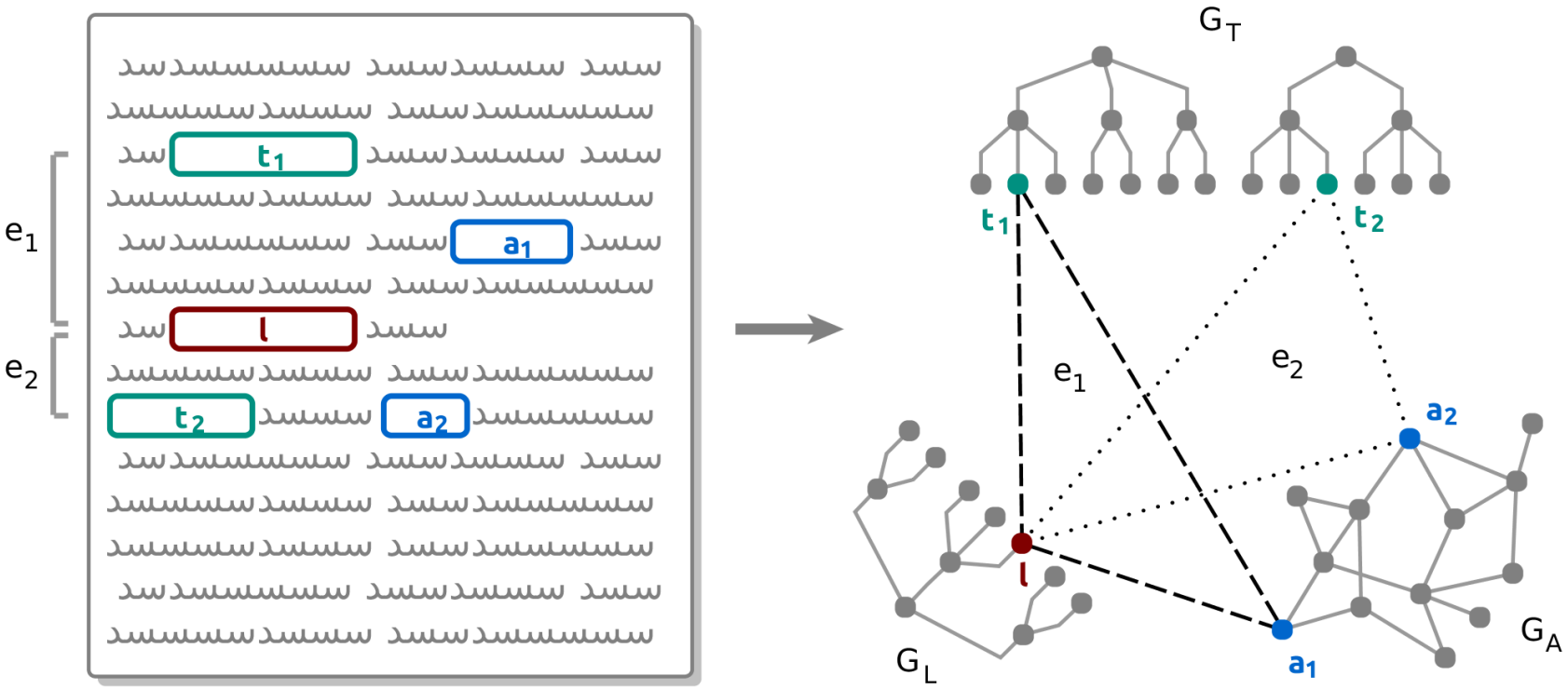
Event Hypergraph



Event Hypergraph

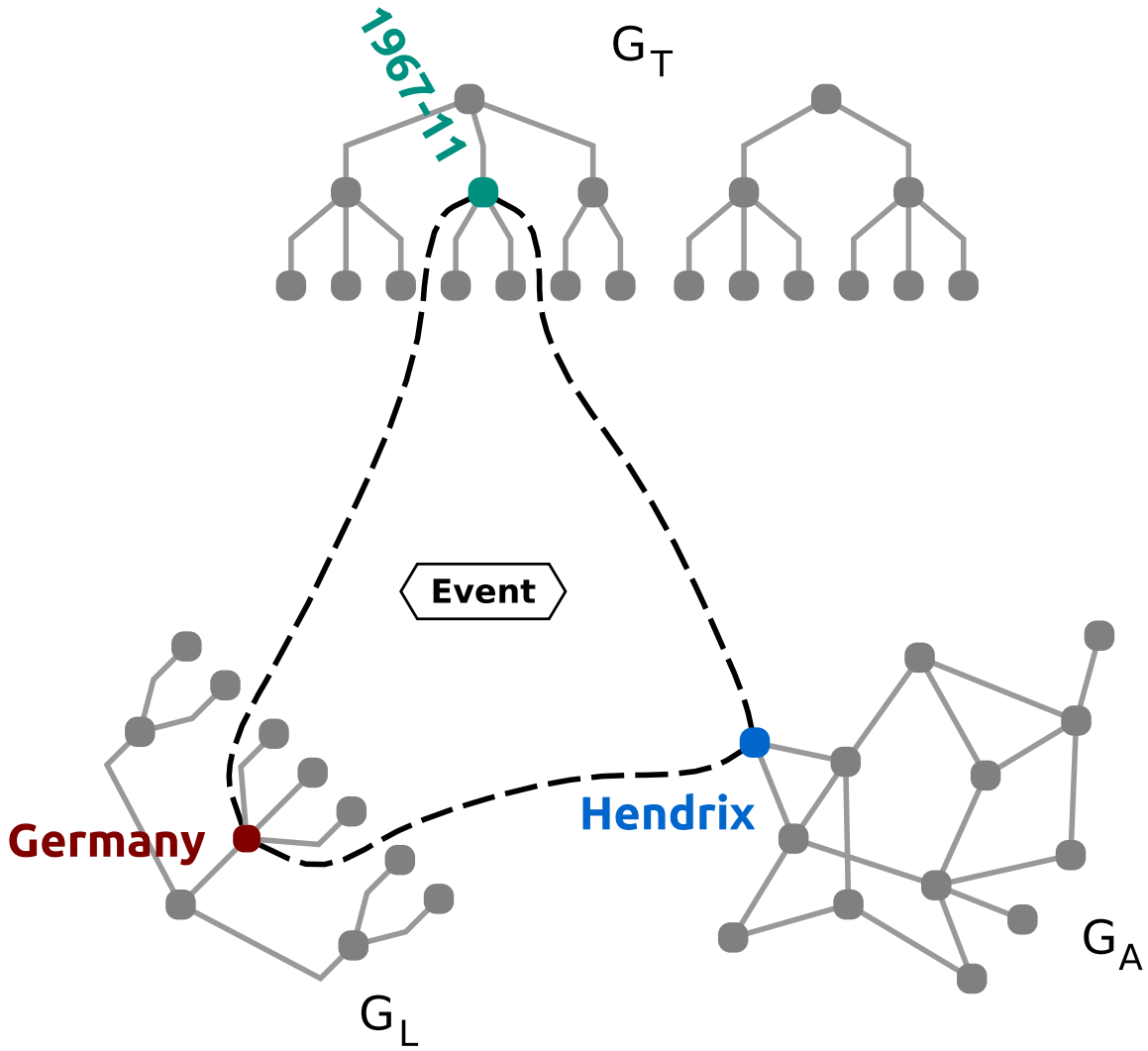


Event Hypergraph

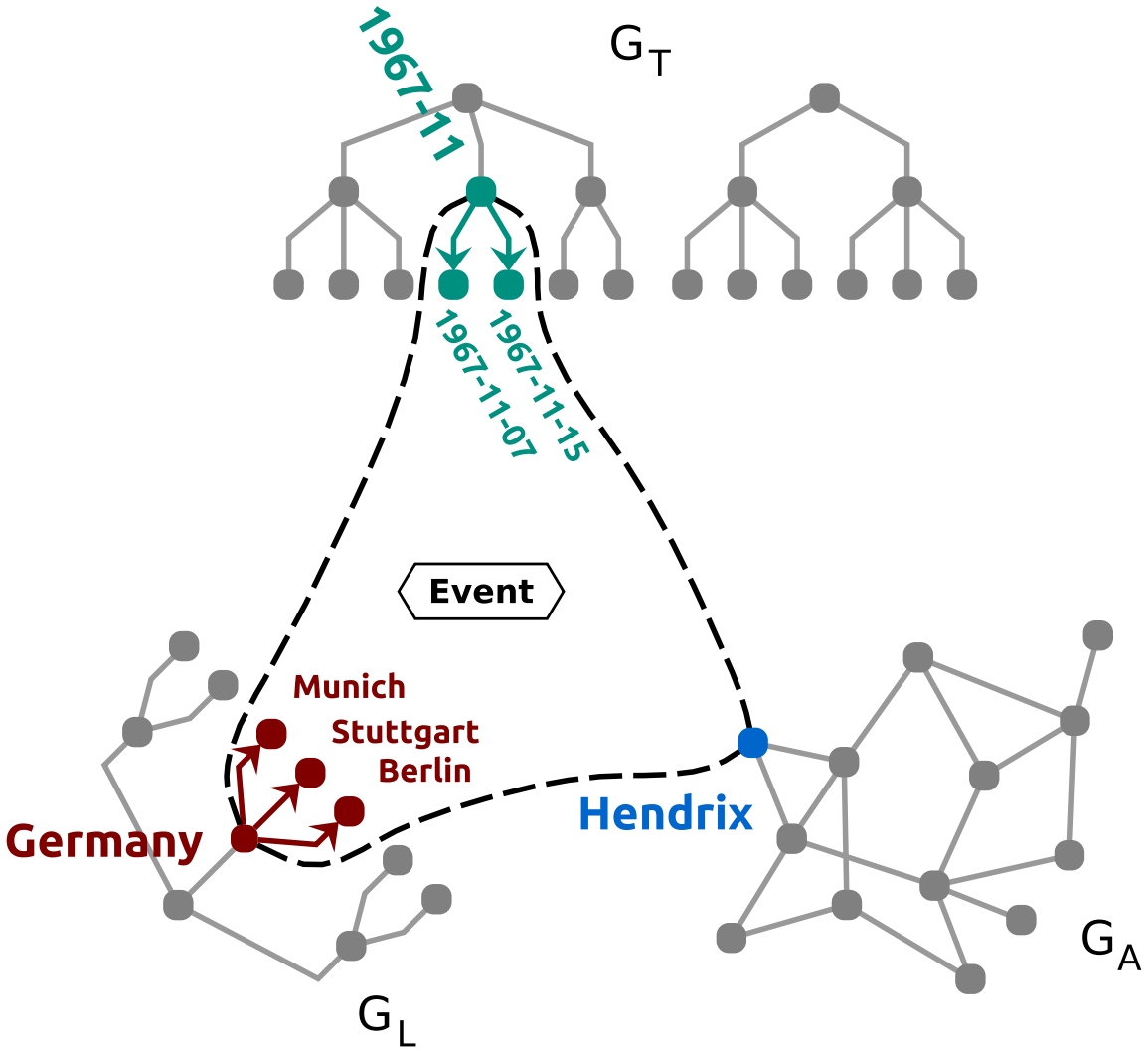


Event Refinement

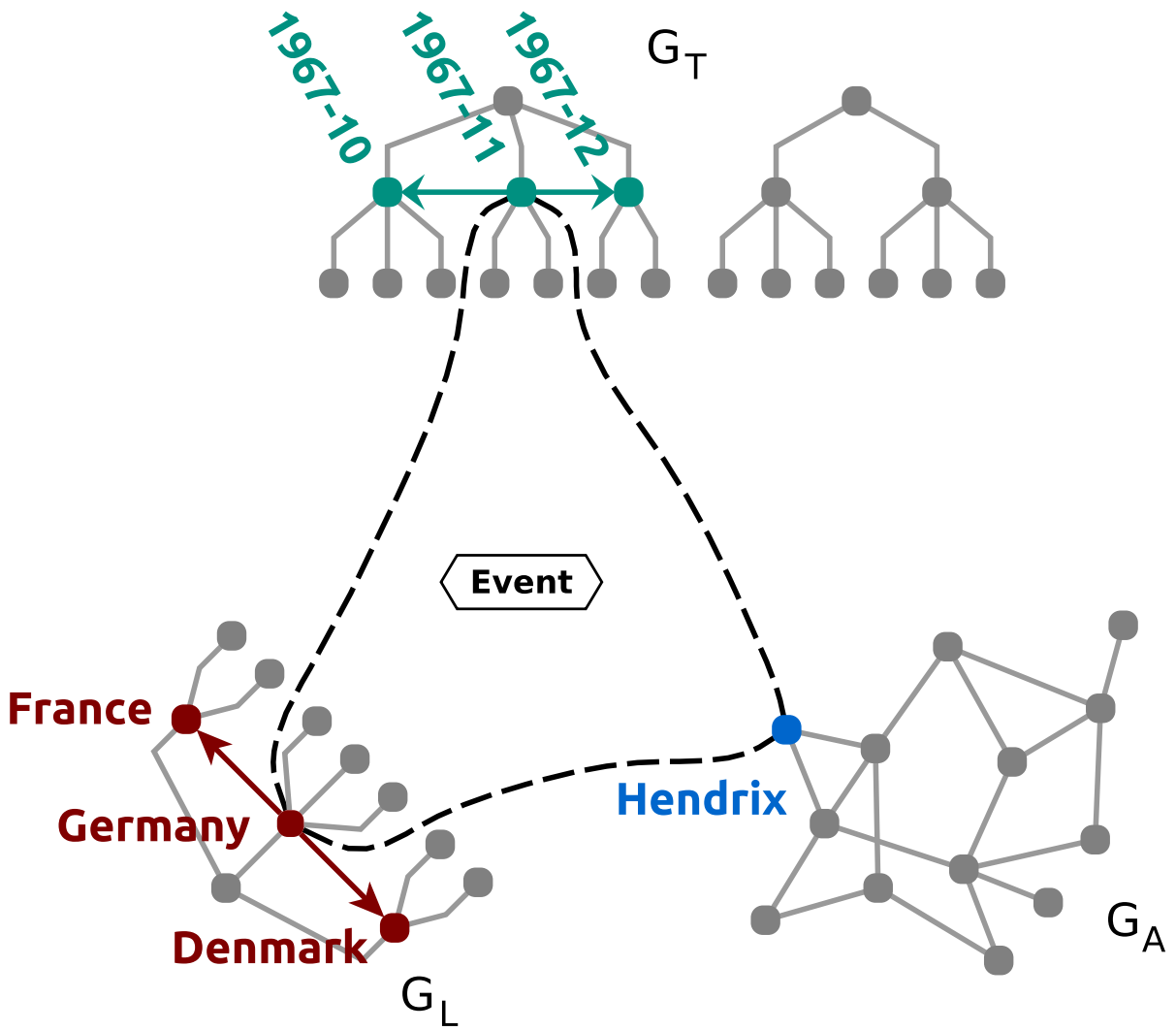
Event Refinement



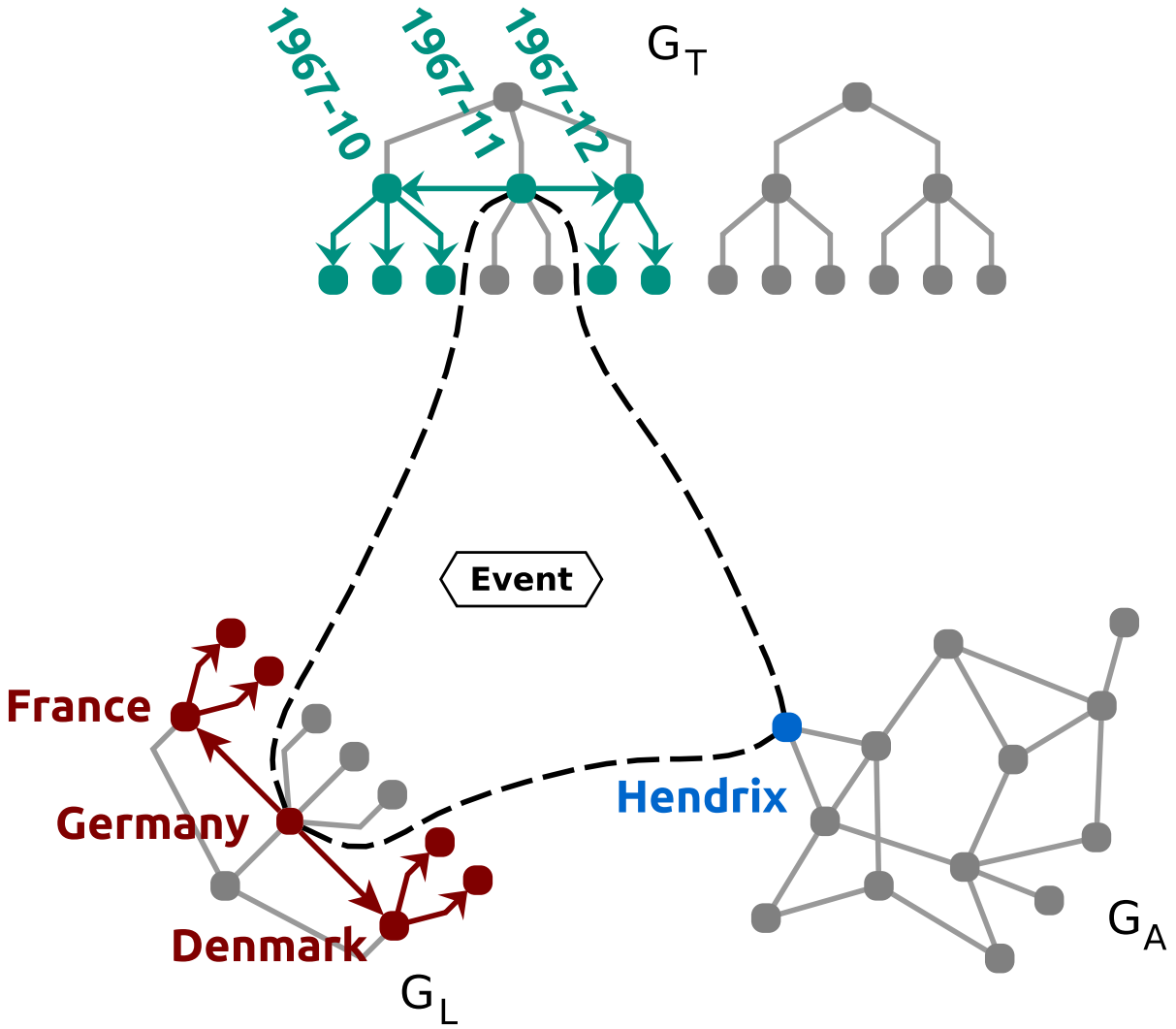
Granularity Refinement



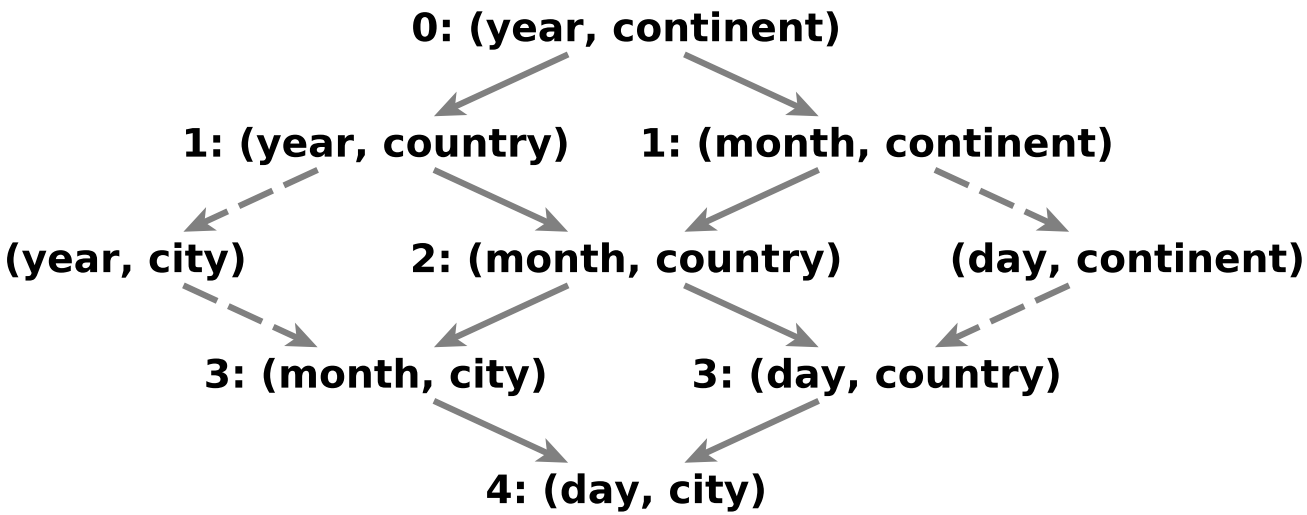
Neighbourhood Refinement



Neighbourhood Refinement



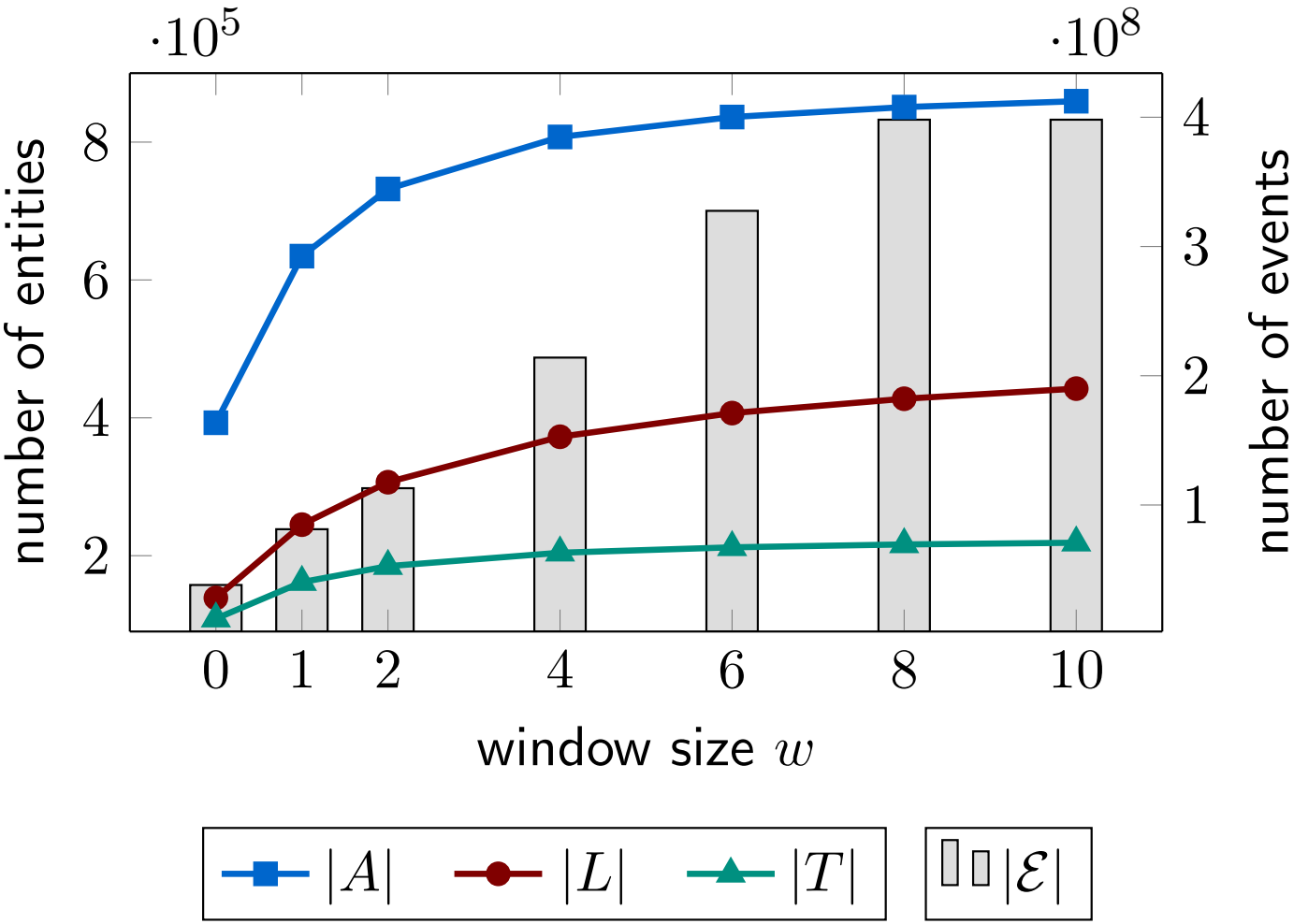
Event Refinement: Stratification



Evaluation

Extraction of Events from Wikipedia

Entities and Events by Window Size



Background Network Construction

Location Network:

- From Wikidata: continents, countries and cities
- Containment edges between hierarchy levels
- Neighbourhood edges between adjacent countries
- Neighbourhood edges between close cities

Temporal Network:

- Temporal Tagging with Heideltime [SG13]
- Containment edges between years, months and days
- Neighbourhood edges within granularity layers

Ground Truth Event Queries

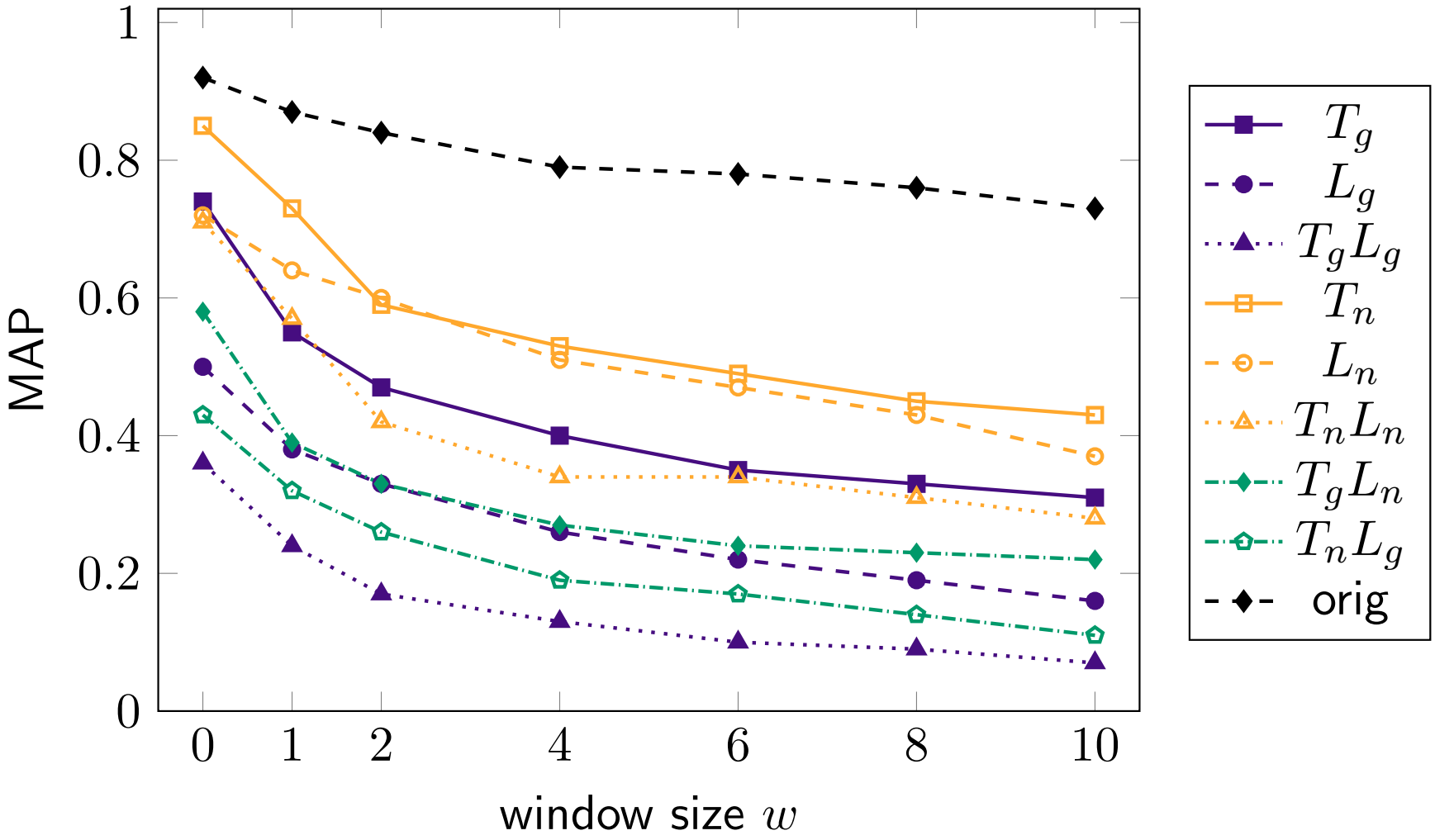
We obtain events by

- Named Entity Recognition in online news articles
- Extraction of the entity triples of events by hand

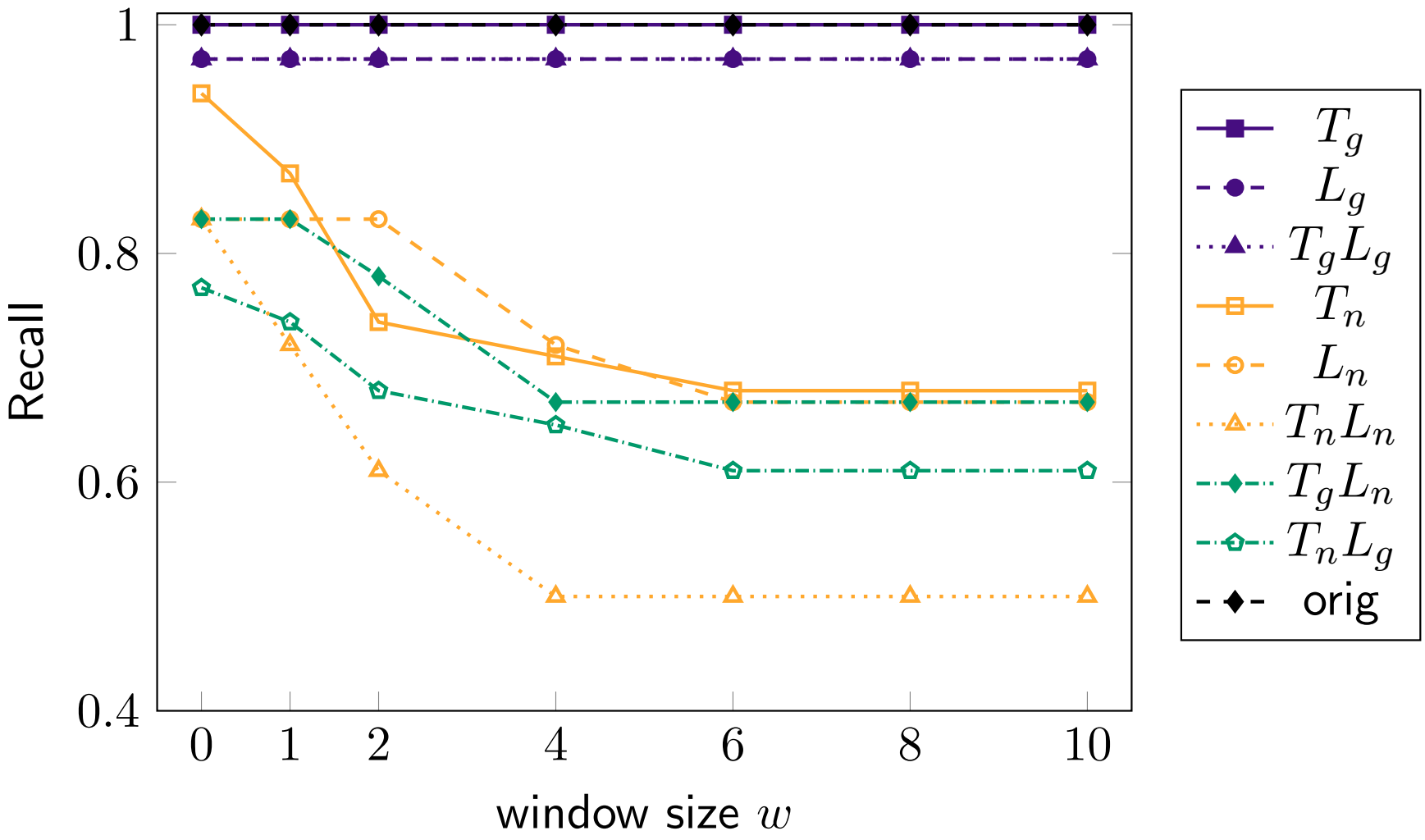
We generate queries by making events less certain in the combination of dimensions

- Location L
- Time T
- Granularity g
- Neighbourhood n

Evaluation Results (Precision)



Evaluation Results (Recall)



Summary

New method for event representation:

- As a hypergraph model
- Backed by underlying entity networks
- Compatible with any entity-based definition of event

Graph-based event refinement offers:

- Spatio-temporal refinement in two dimensions: neighbourhood and granularity
- Efficient computation due to localized queries
- Language independence

Ongoing Work

Directions for ongoing event refinement research:

- Include measures of granularity and neighbourhood in social background networks
- Include hierarchical or organization networks
- Add event structures beyond triangles

The event network and the background networks are available for download.

`http://dbs.ifi.uni-heidelberg.de/index.php?id=data`

The event network and the background networks are available for download.

`http://dbs.ifi.uni-heidelberg.de/index.php?id=data`

Thank you!
Questions?

Bibliography

James Allan.
Topic Detection and Tracking: Event-Based Information Organization, volume 12.
Springer Science & Business Media, 2002.

Johanna Geiß, Andreas Spitz, and Michael Gertz.
Beyond Friendships and Followers: The Wikipedia Social Network.
In *ASONAM*, 2015.

Jannik Strötgen and Michael Gertz.
Multilingual and cross-domain temporal tagging.
Language Resources and Evaluation, 47(2), 2013.

Denny Vrandečić and Markus Krötzsch.
Wikidata: A Free Collaborative Knowledgebase.
Communications of the ACM, 57(10), 2014.