VIS 2021

Mining and Understanding Stories in Text Sequences with Narrative Visualization

Award for Innovative Narrative Visualization and Analysis Methodology

Zeyu Li, Teng Wang, Ruizhi Shi, Zhaohui Li, Jiawan Zhang
Tianjin University

Agenda

- Overview
- Solutions
  - Q1
    - a classification system
    - a system used for verifying and exploring the classification
  - Q2: a narrative visualization creation system
  - Q3: a map with trajectory
  - Q4: summarized by human
- Conclusion

Overview: data

- Microblog messages
  - date
  - author
  - message
  - location(optional)

text sequences with timestamp

- Call center data
  - date,
  - message
  - location

spatiotemporal data

Overview: questions

- Organize, understand and summarize the content of messages from a retrospective perspective
  - Q1: classify messages and identify the feature of each class
  - Q2: analyze the evolution of the level of the risk

- Compare the differences between the retrospective analysis and real-time analysis in terms of decision making and visualization solution
  - Q3: determine a dispatch location for first responders
  - Q4: compare the previous and current solution

Overview: questions

Q1: classify messages and identify the feature of each class

Q2: analyze the evolution of the level of the risk

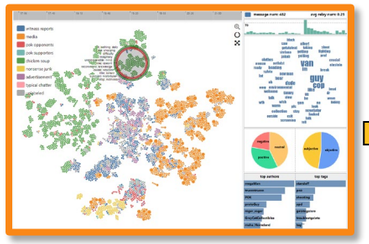
Q3: determine a dispatch location for first responders

Q4: compare the previous and current solution

#1

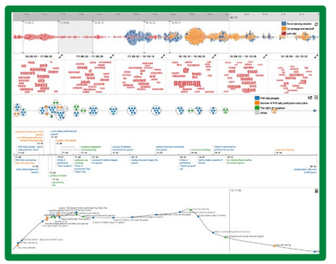
classified messages

#2



messages of three main events

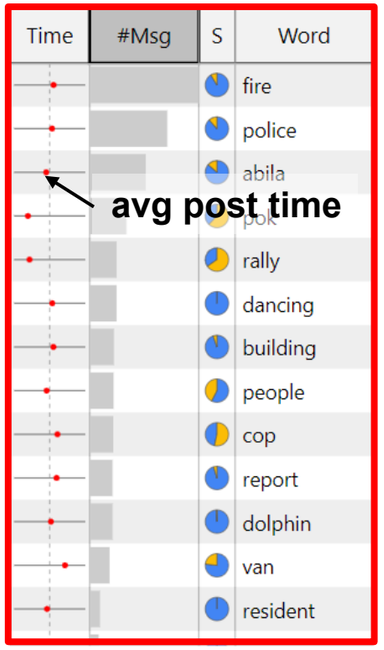
#3



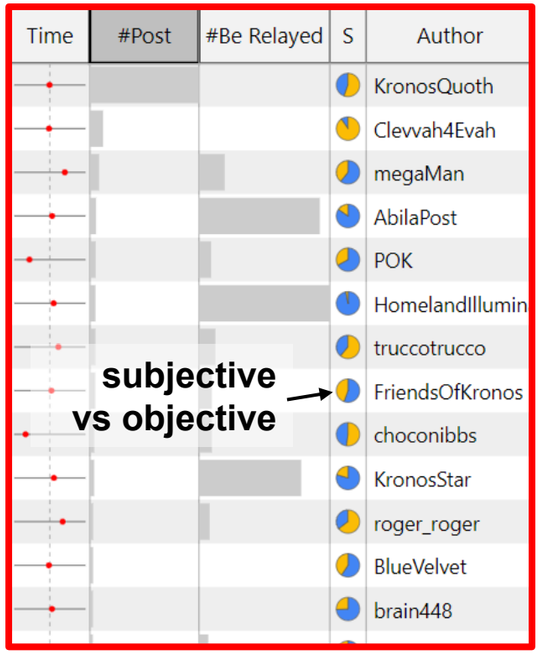
Q1: classify messages according to their content and to identify the features of each class

First system: classify messages

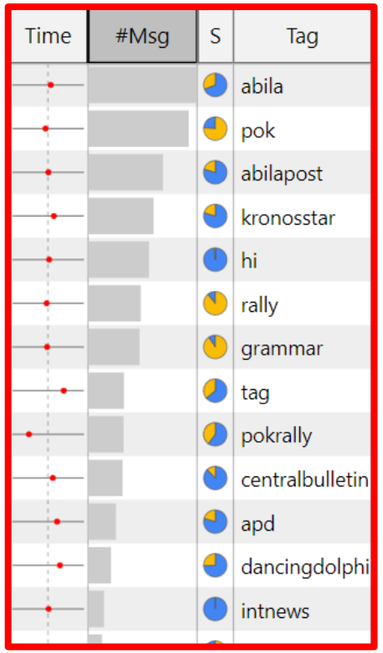
top keywords



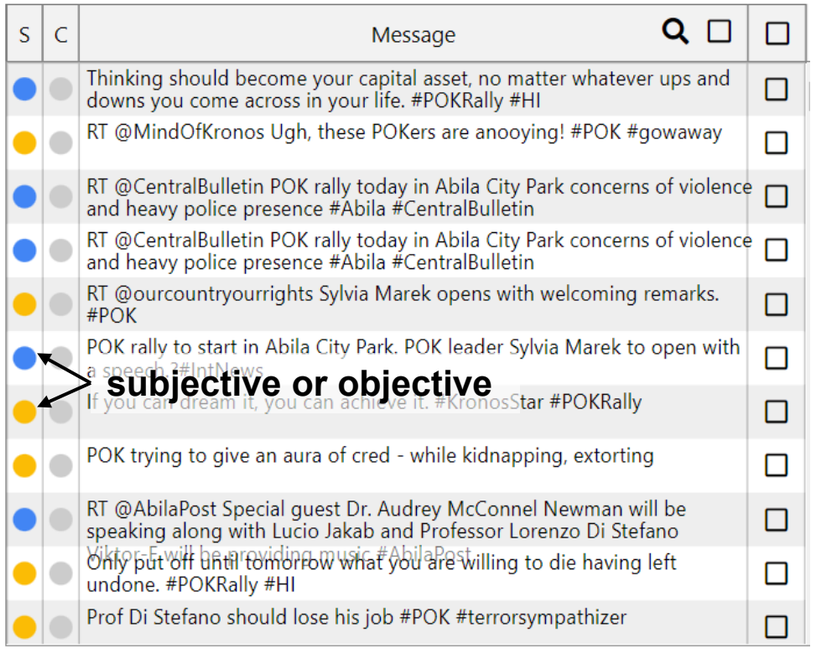
top authors

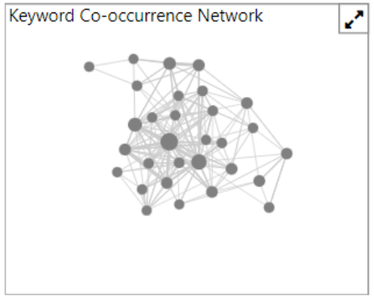


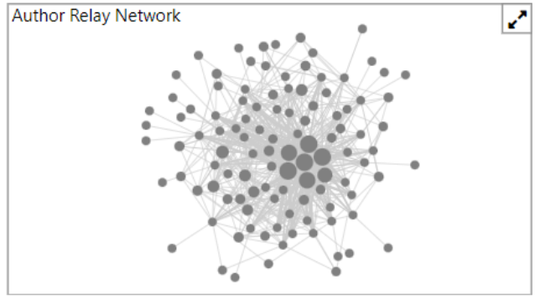
top tags

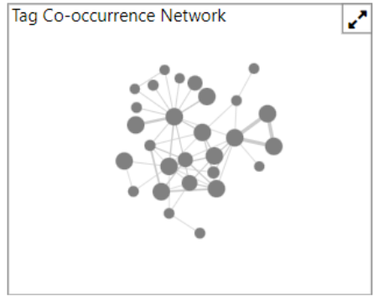


messages of an instance

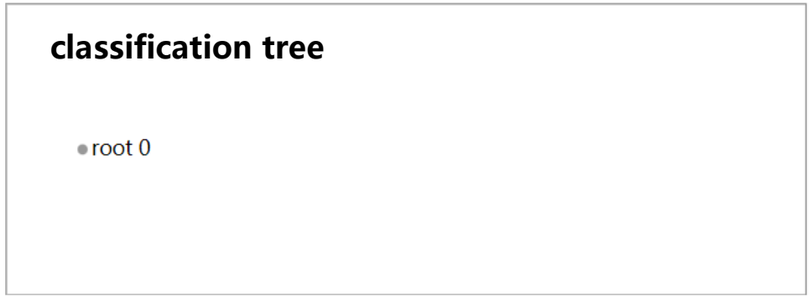






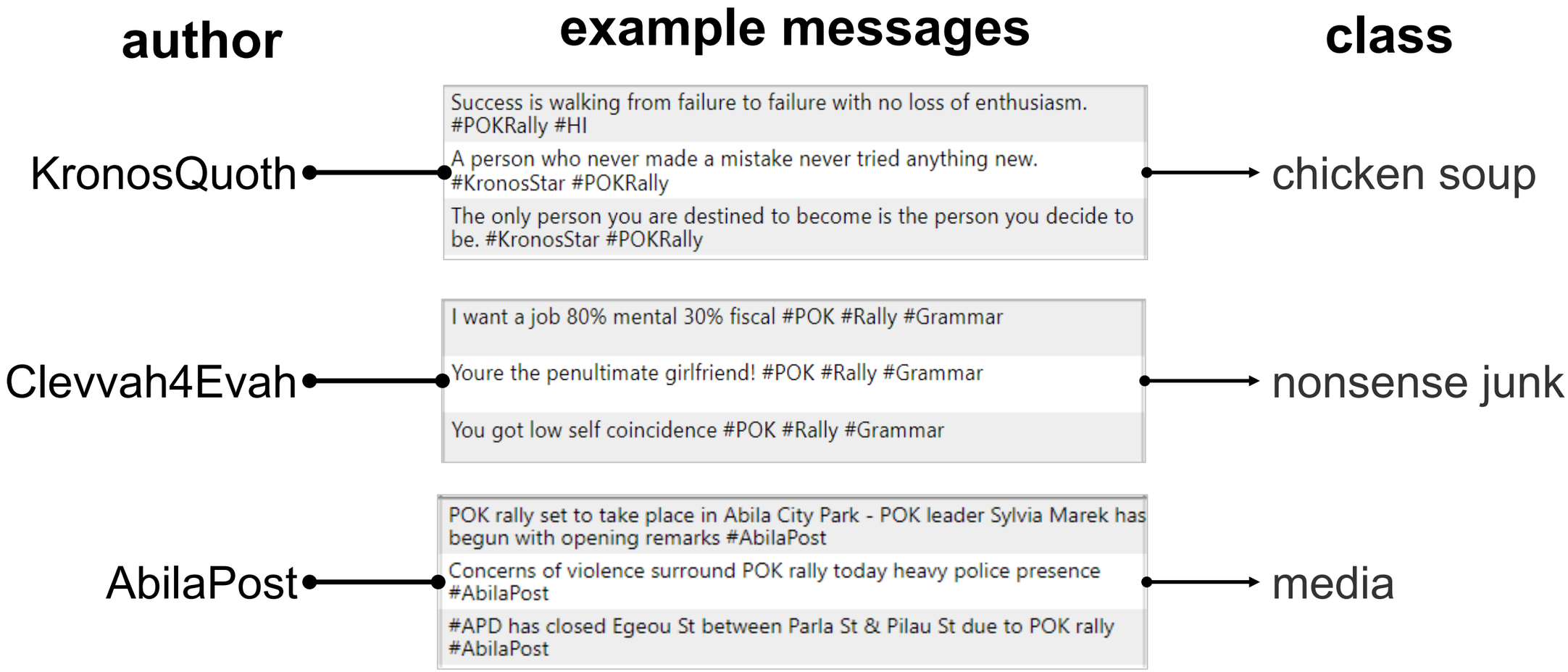


classification tree

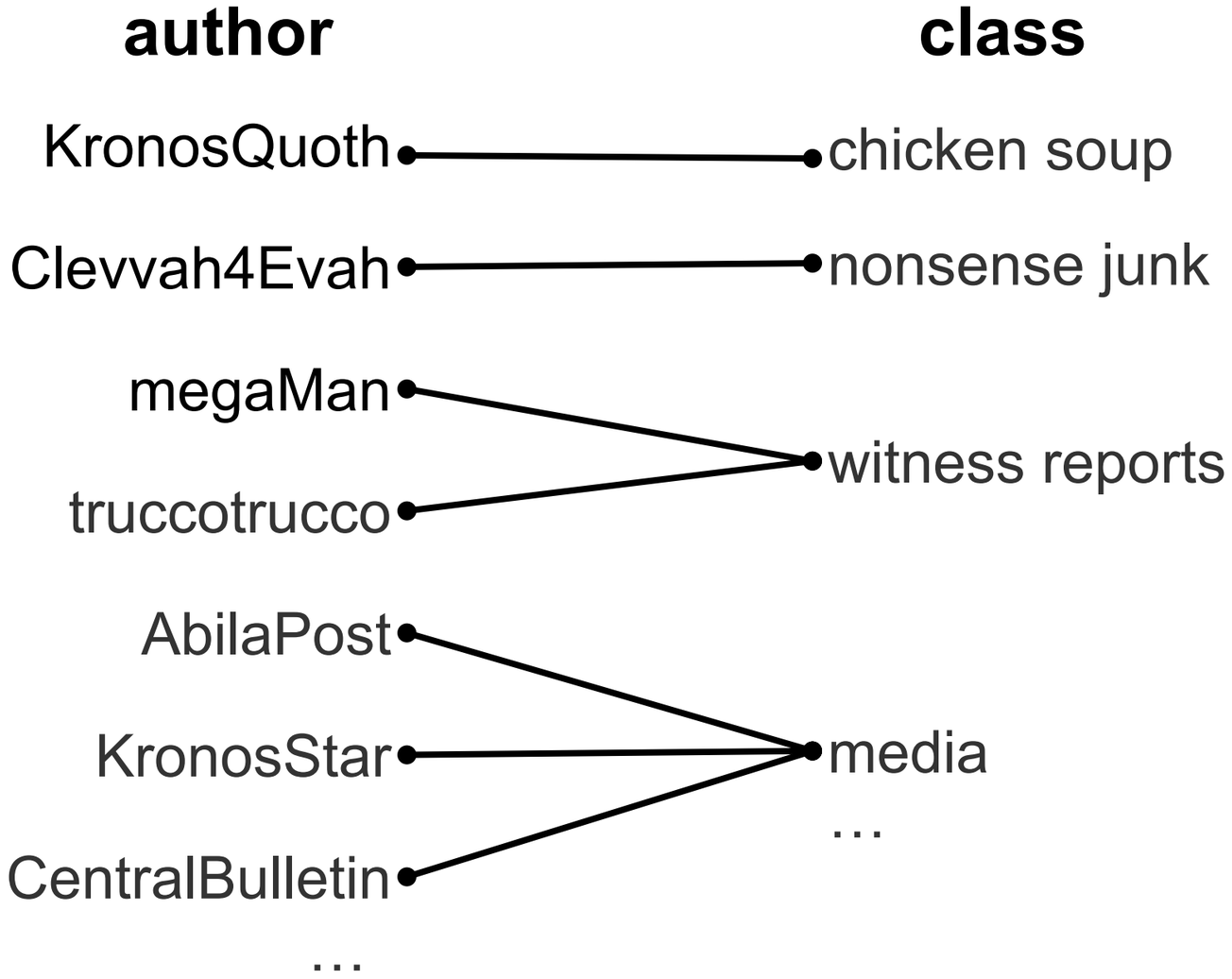


First system: classify messages

Mapping between authors and classes

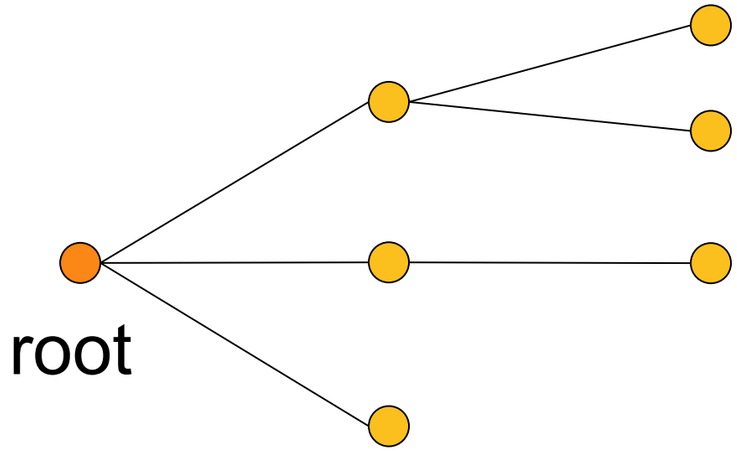


Mapping between authors and classes

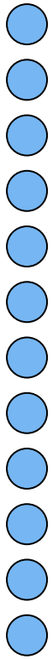


Build a classification tree

inner nodes: classes



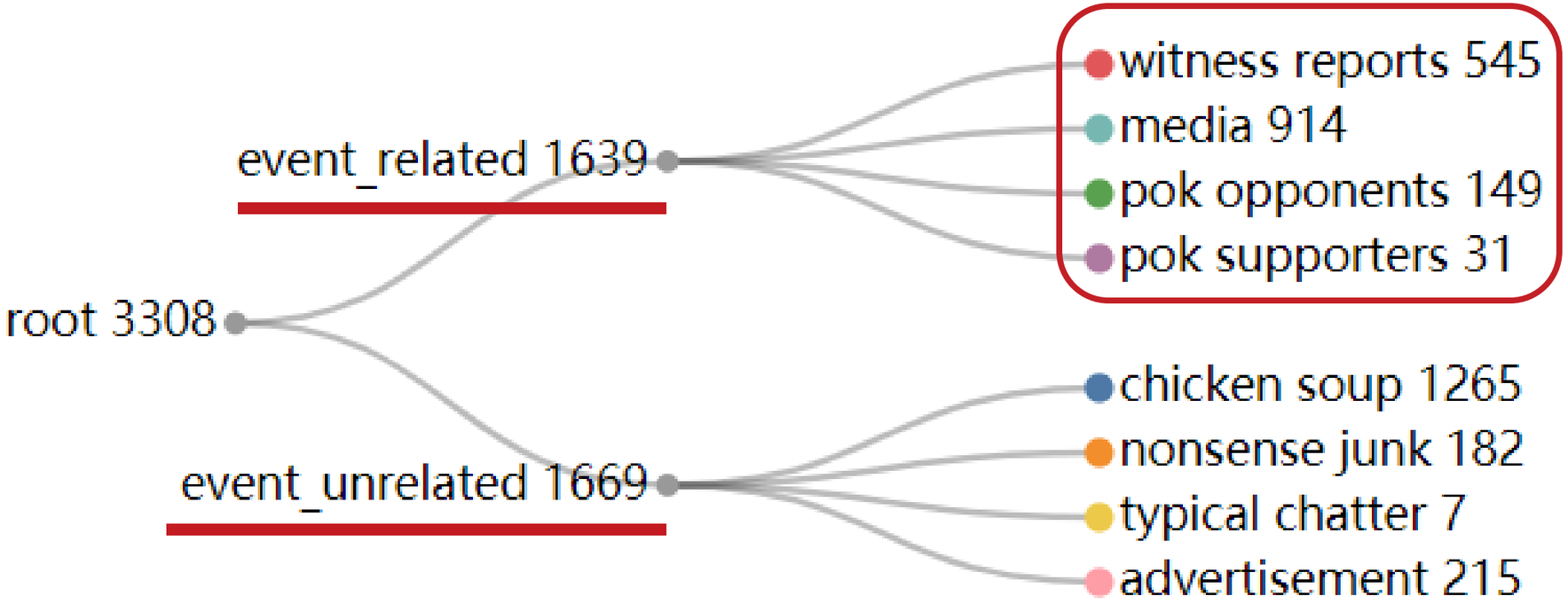
authors with messages



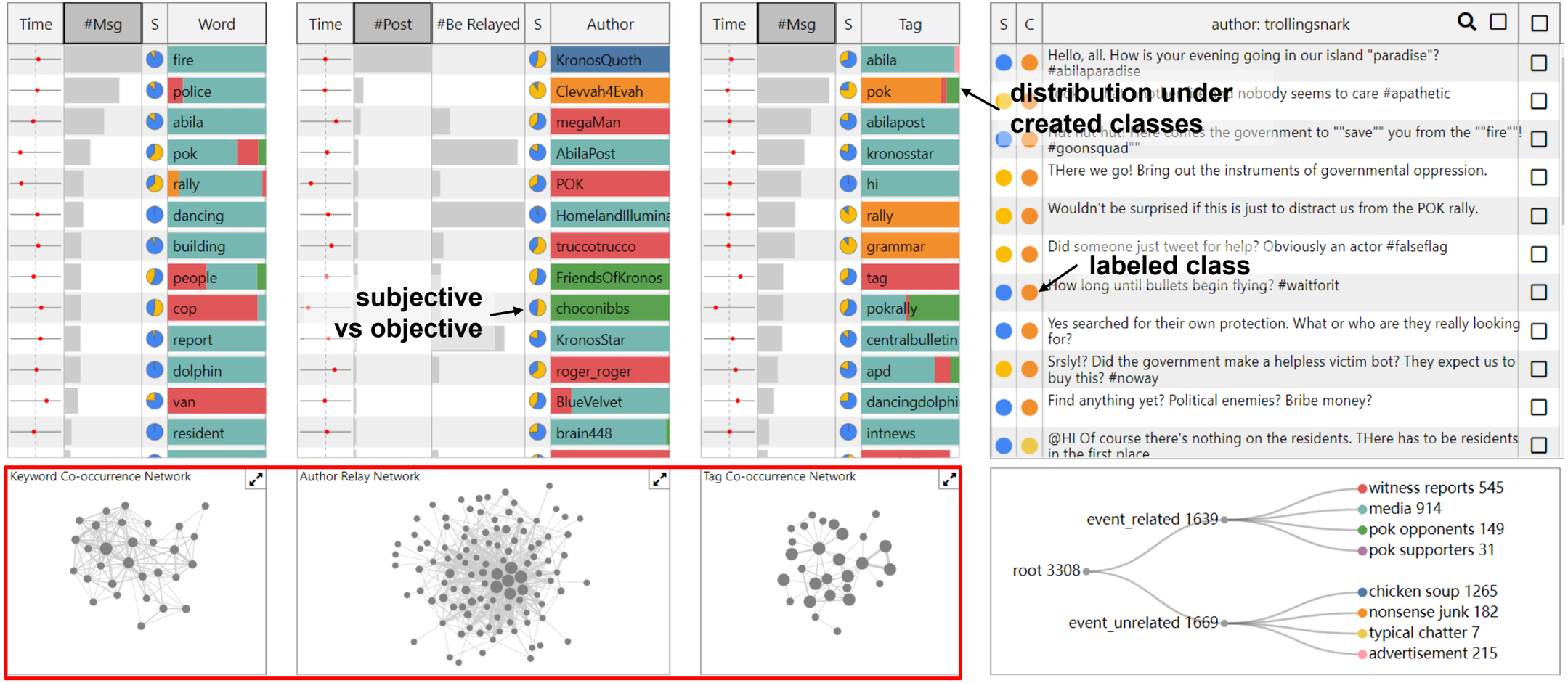
1. create a new class
2. assign an author’ messages to an class

A built classification tree

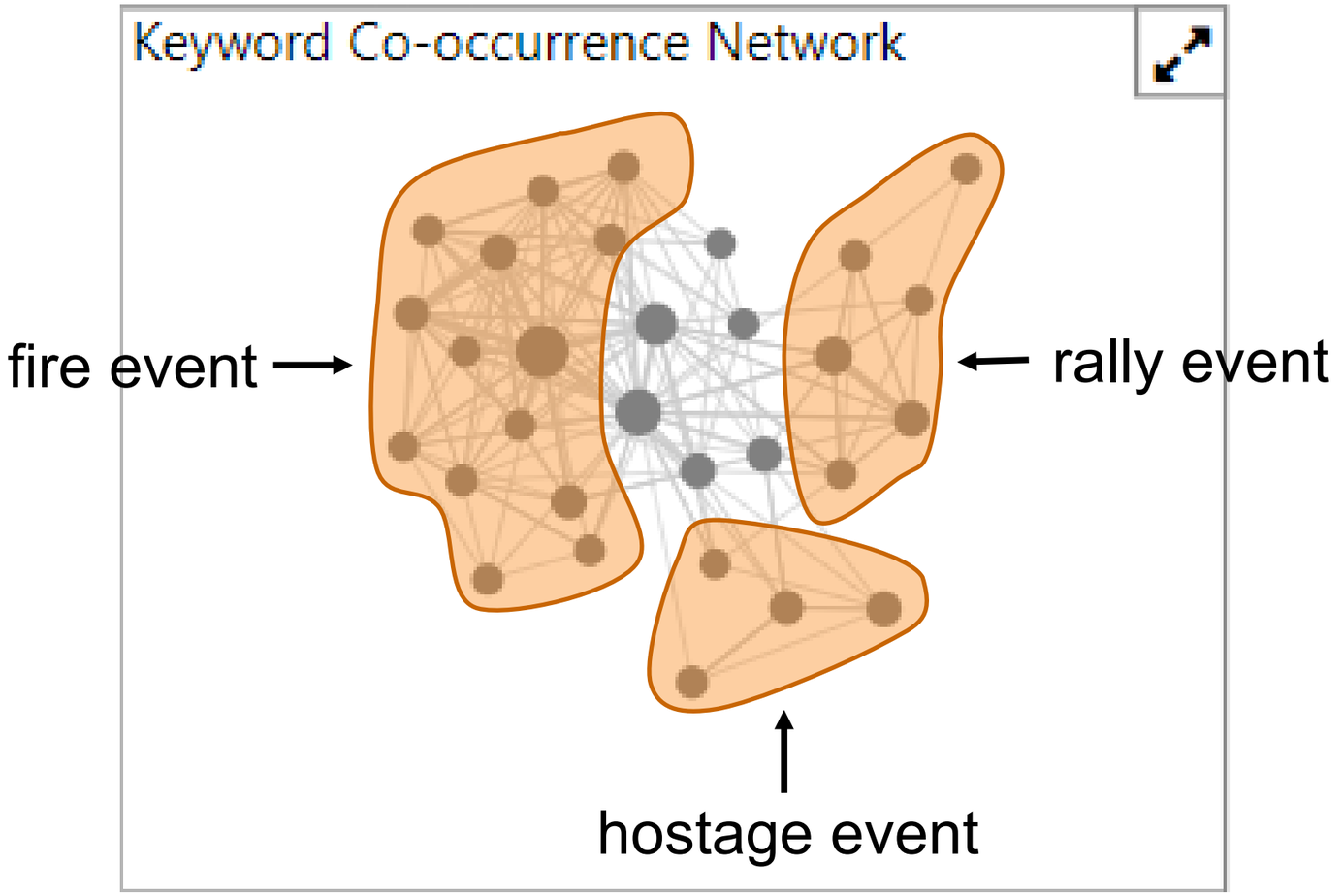
After labeling the top 30 authors, 95% of messages are assigned.



Network of instances

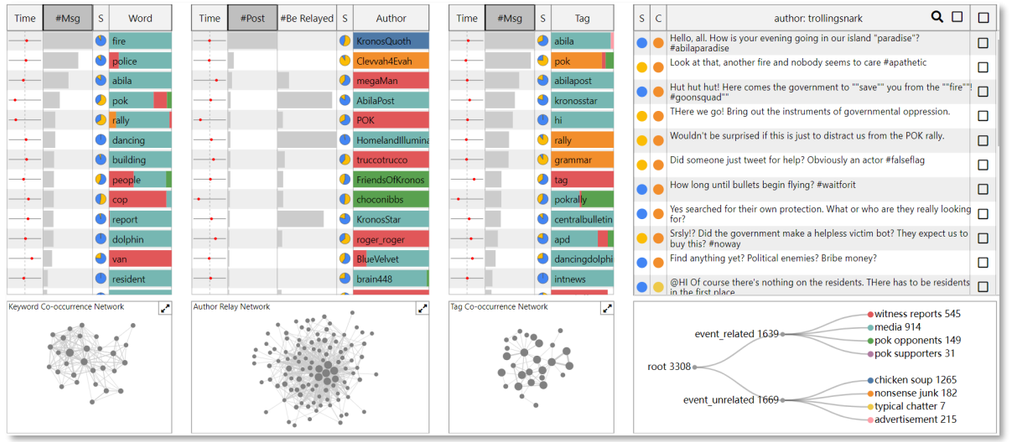


Network of instances: three main events



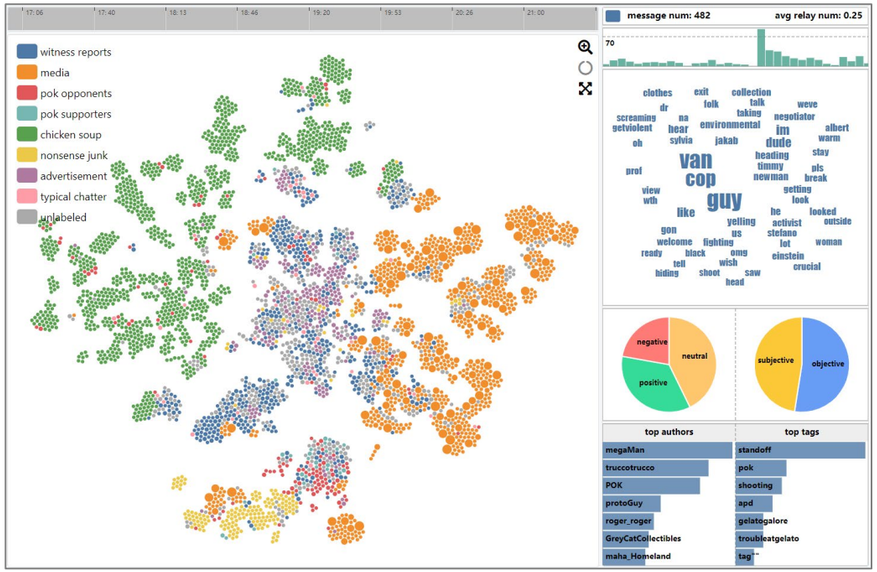
Second system

the first system



classified messages

the second system

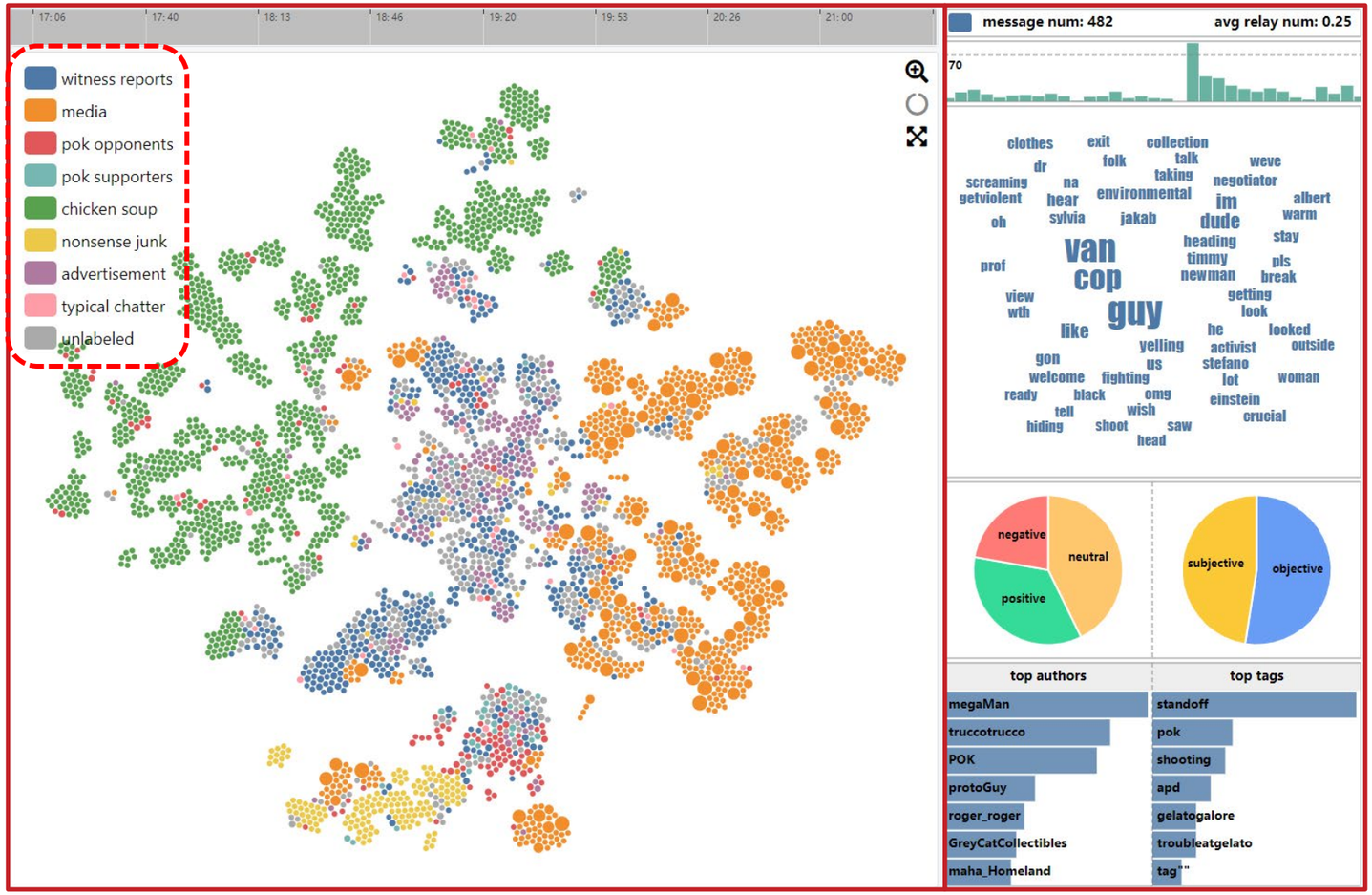


- check the reliability of the classification
- identify features of each class

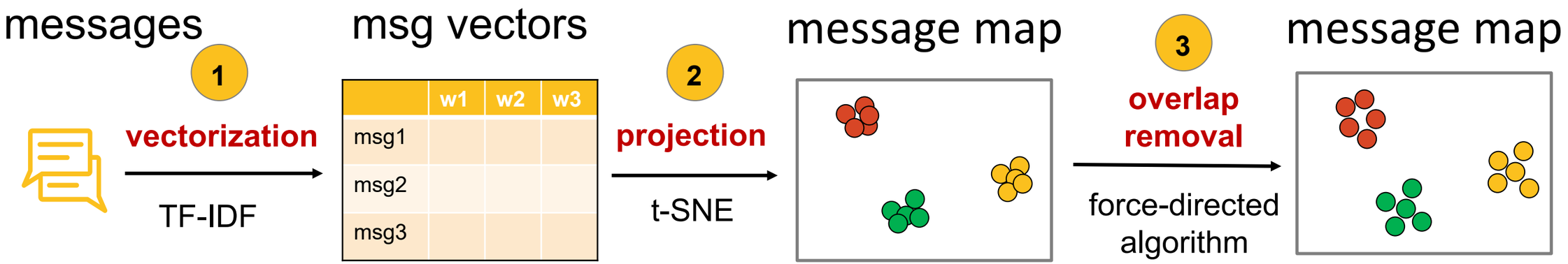
Second system

message map

feature checker

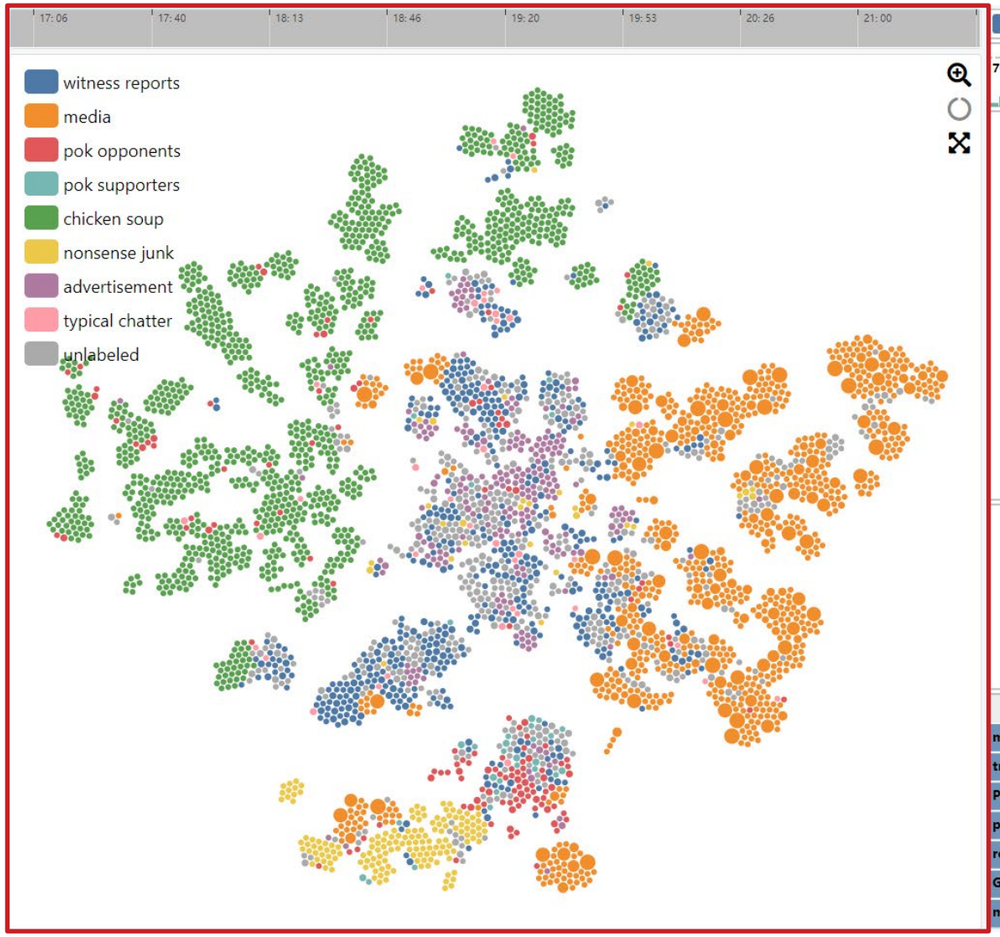


Three steps to create the message map

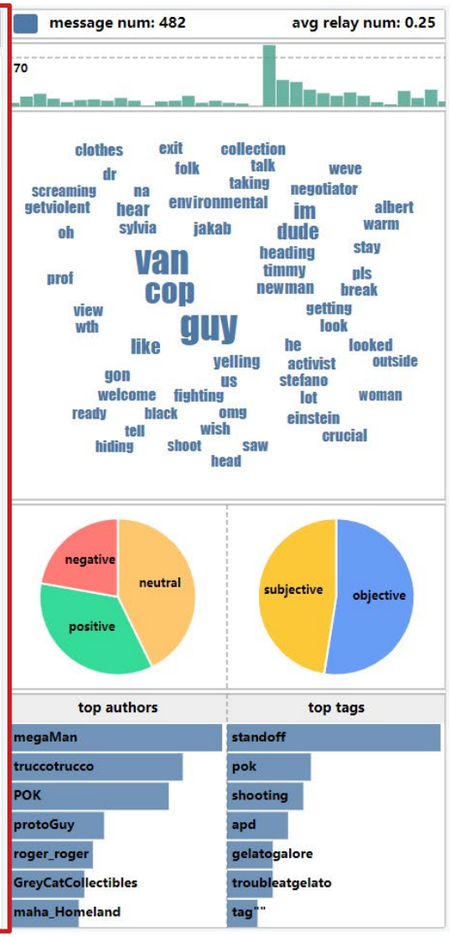


Second system

message map



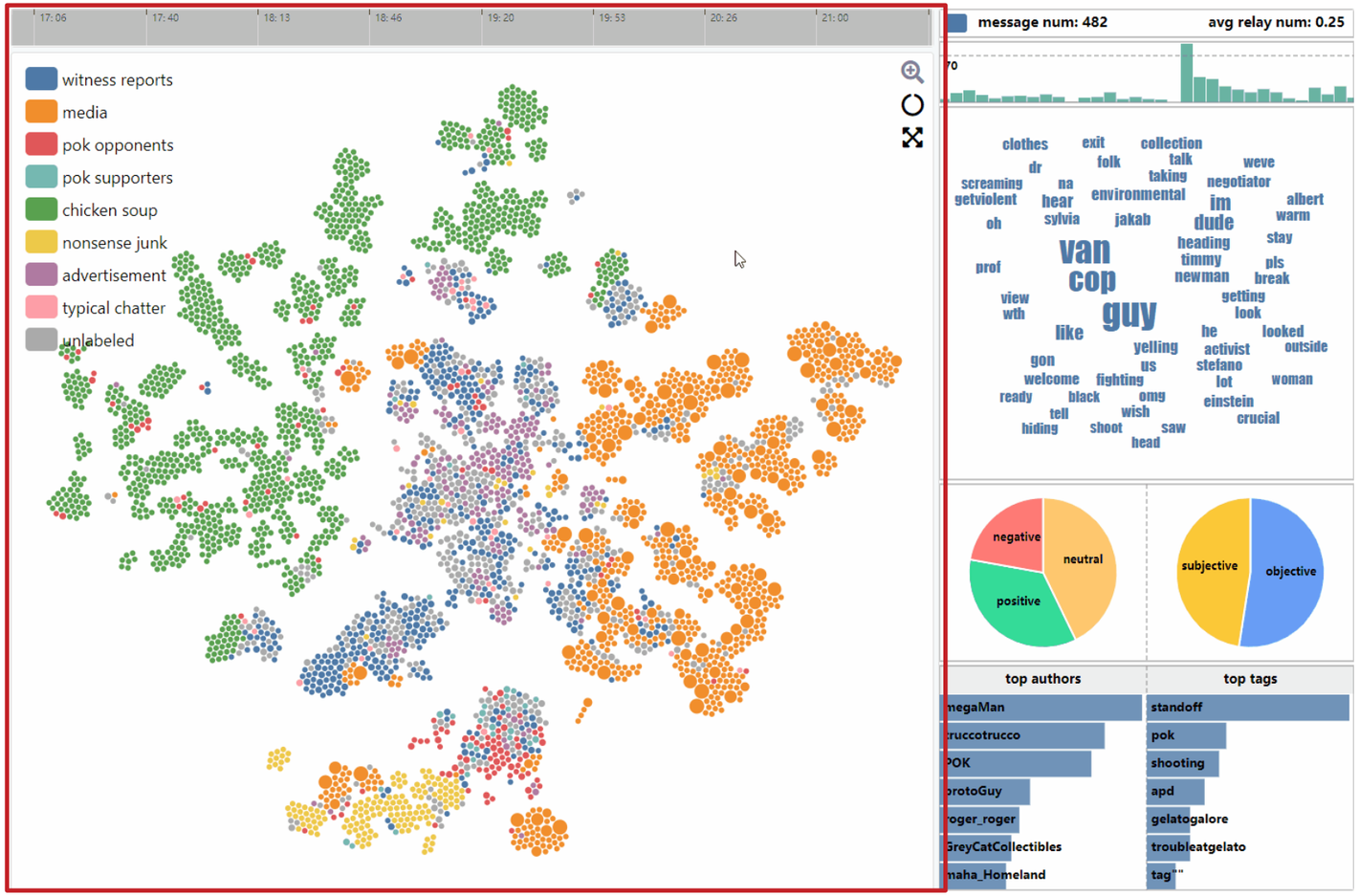
feature checker



Interactive lens

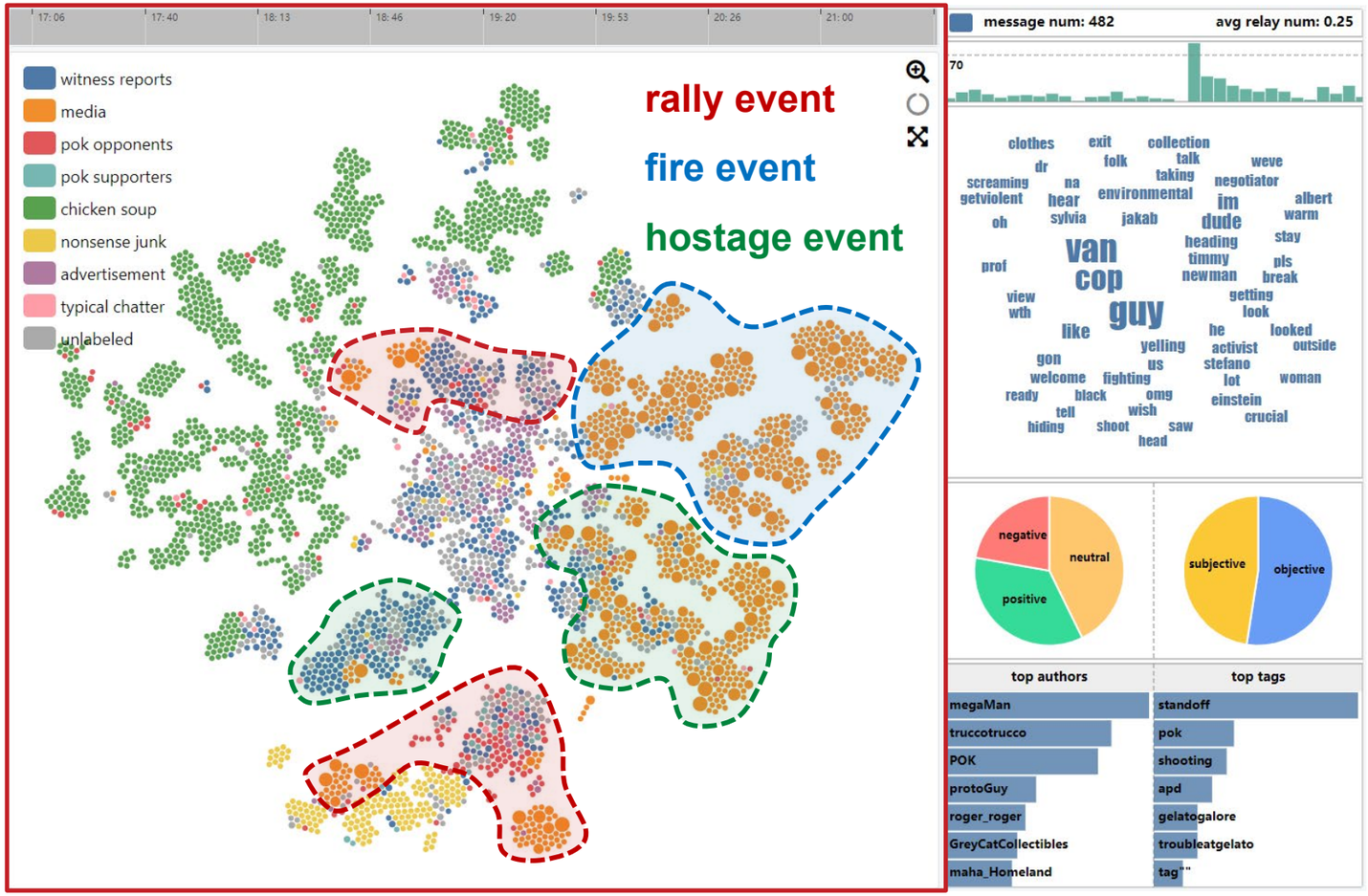
message map

feature checker



Find messages of the three main events

message map



1. label the circled messages as training data

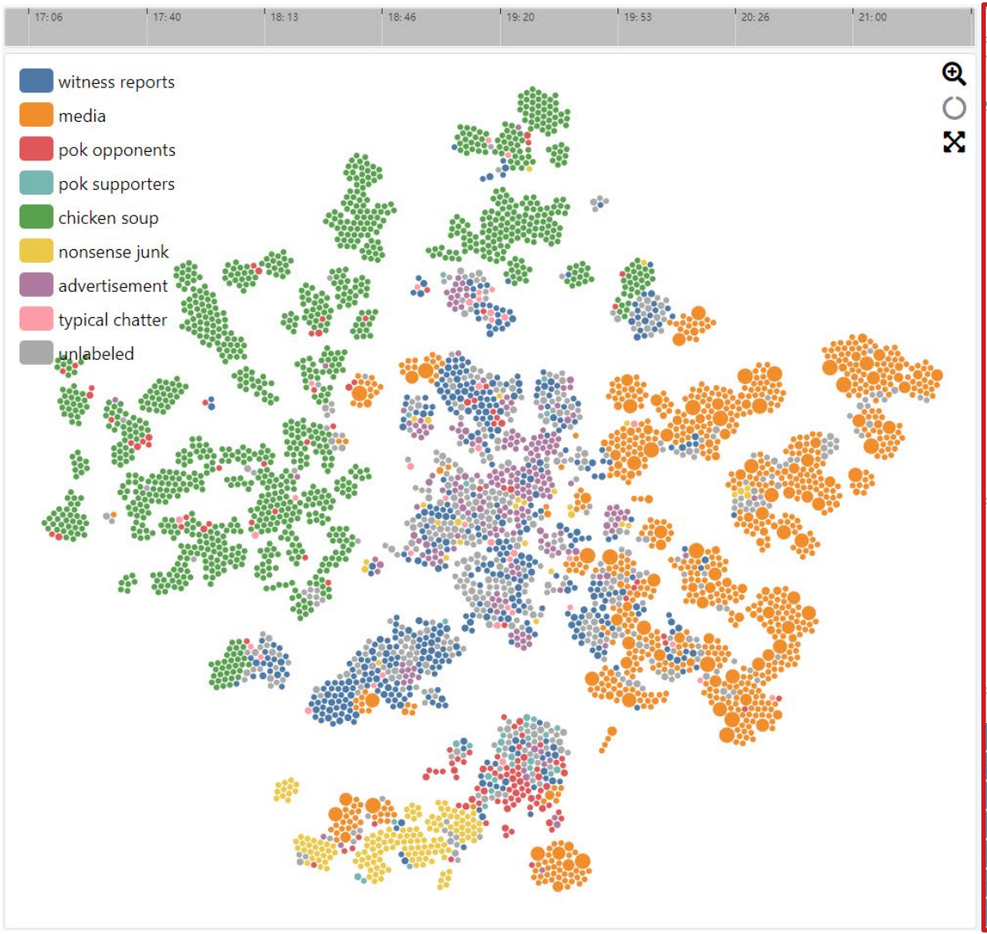
2. train a multi-class classifier

3. apply the classifier to the event-related messages

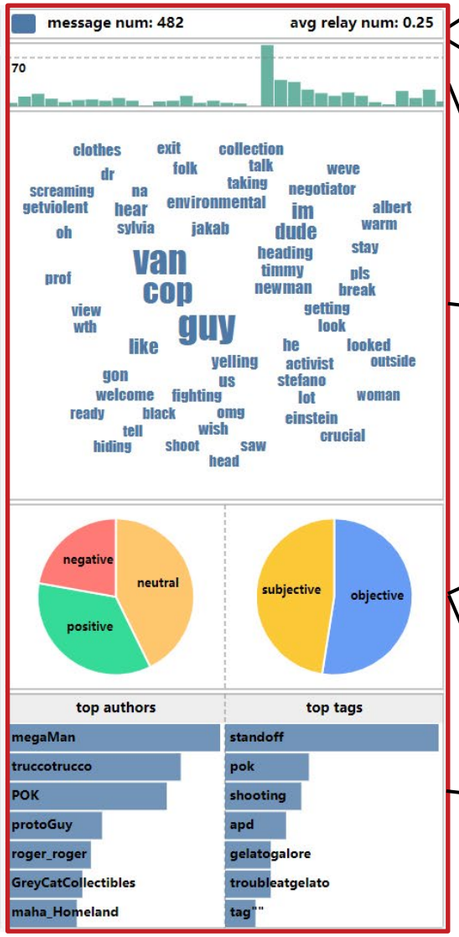
event-related messages
- rally event
- fire event
- hostage event

Features of a class

message map



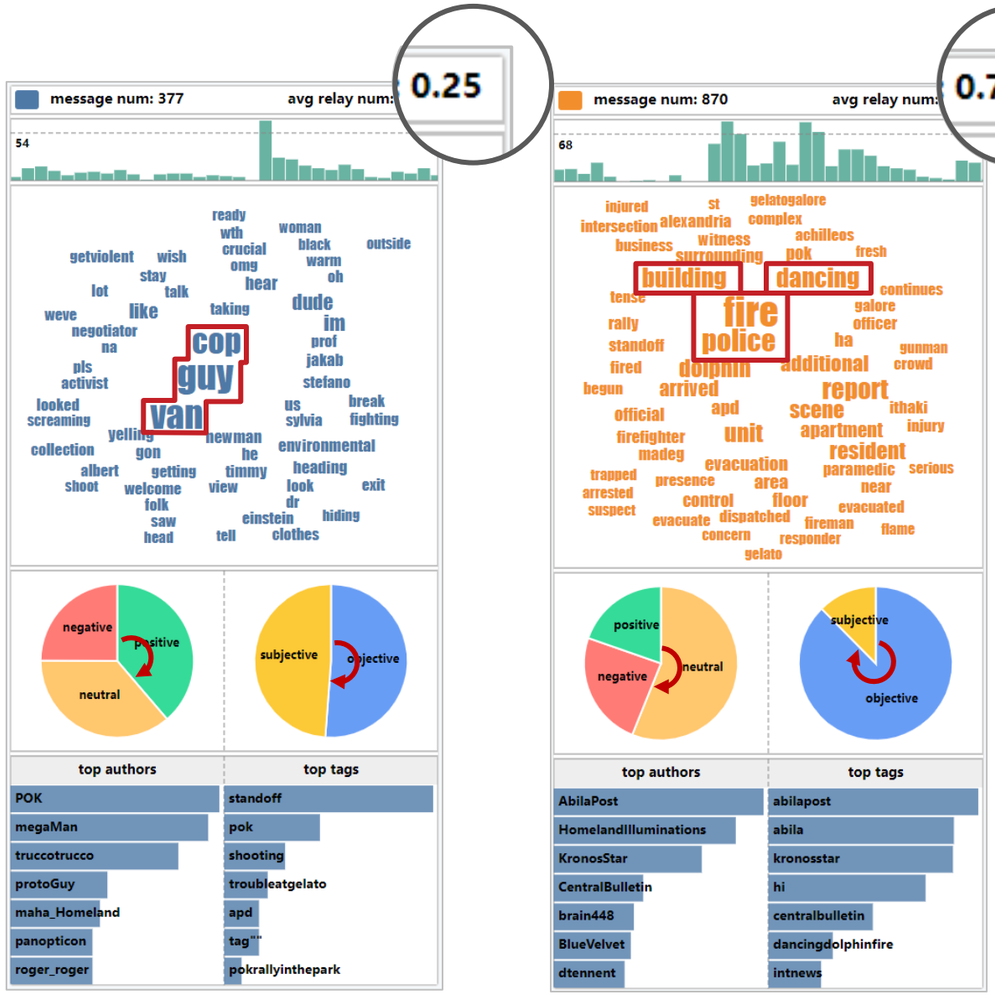
feature checker

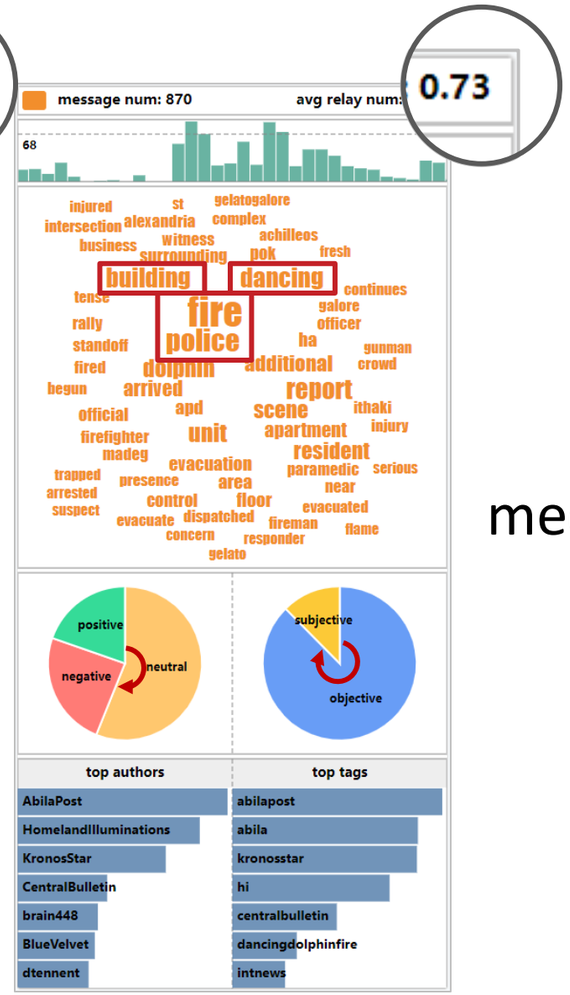


- number of messages
- forwarded times pre message
- distribution of messages over time
- feature words
- sentiment distribution (positive VS neutral VS negative)
- subjectivity (subjective VS objective)
- top authors and tags involved

Differences between witness reports and media

witness reports





media

forwarded times
pre message

Q2: analyze the evolution of the level of the risk to the public and consider the potential consequences of the situation and the number of people affected

Unit events

Main events

rally event

fire event

hostage event

Example of unit events

Sylvia welcomed the crowd

Roughly 1200 people

Lucio Jakab gave a speech

Viktor-E performed ‘River soldiers’

Cops were leaving

desired unit events:

- clear semantics
- appropriate granularity

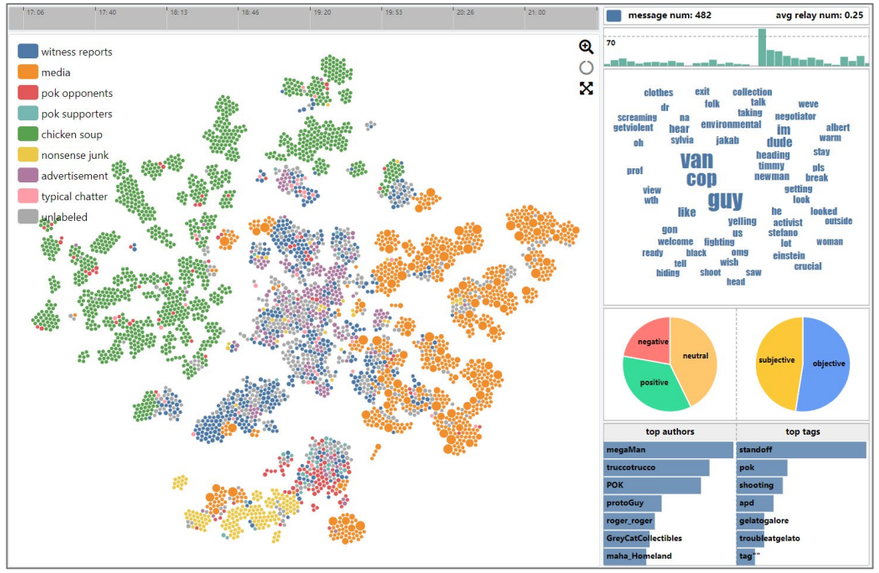
extract by algorithms?

human

- reasoning
- assessment

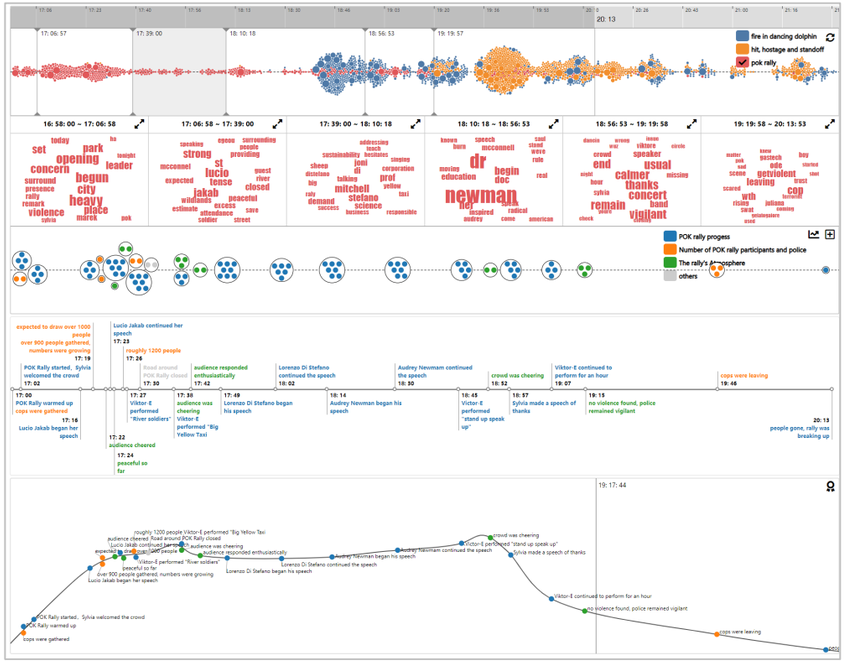
Third system

the second system

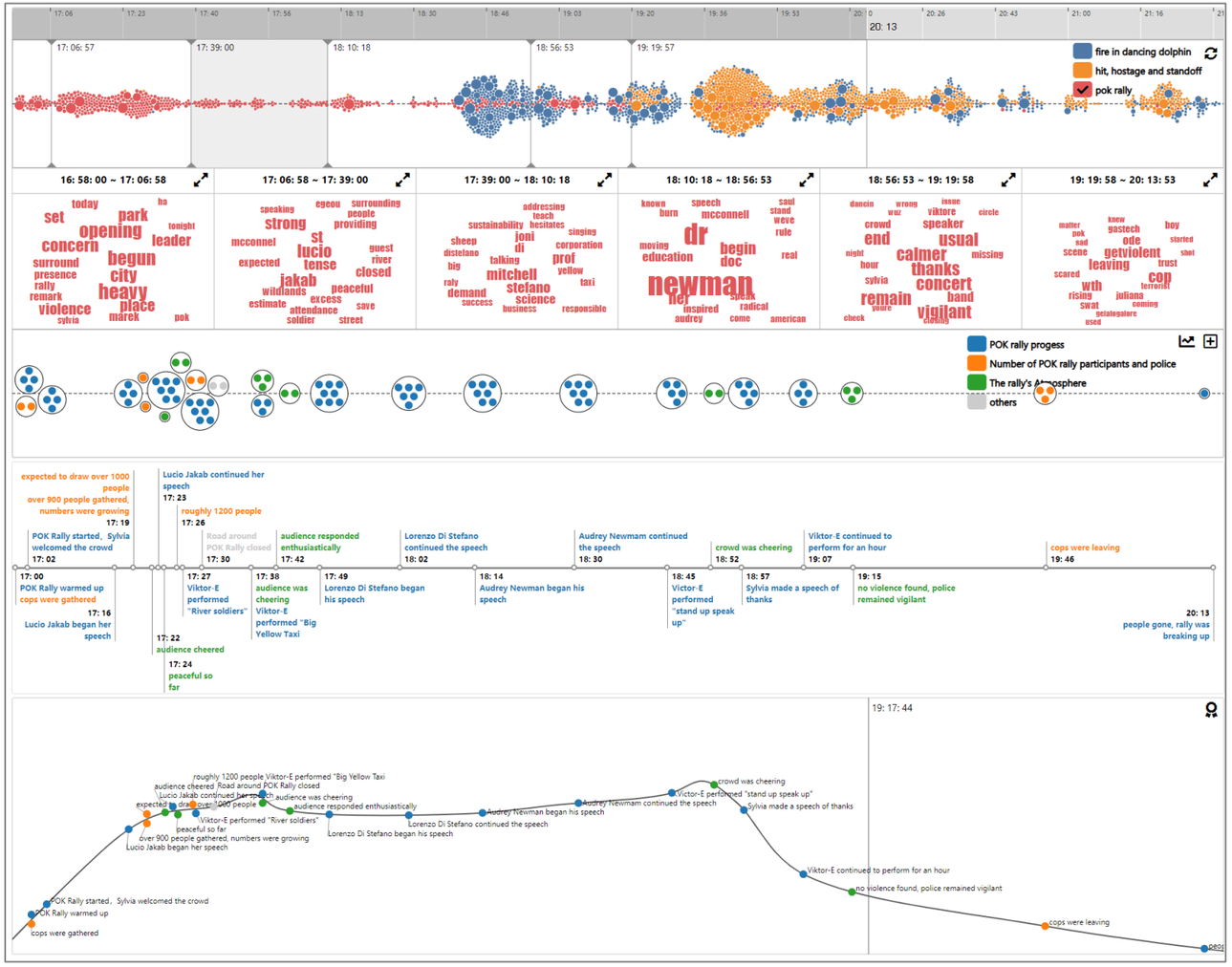


messages of the three main events

the third system



Third system



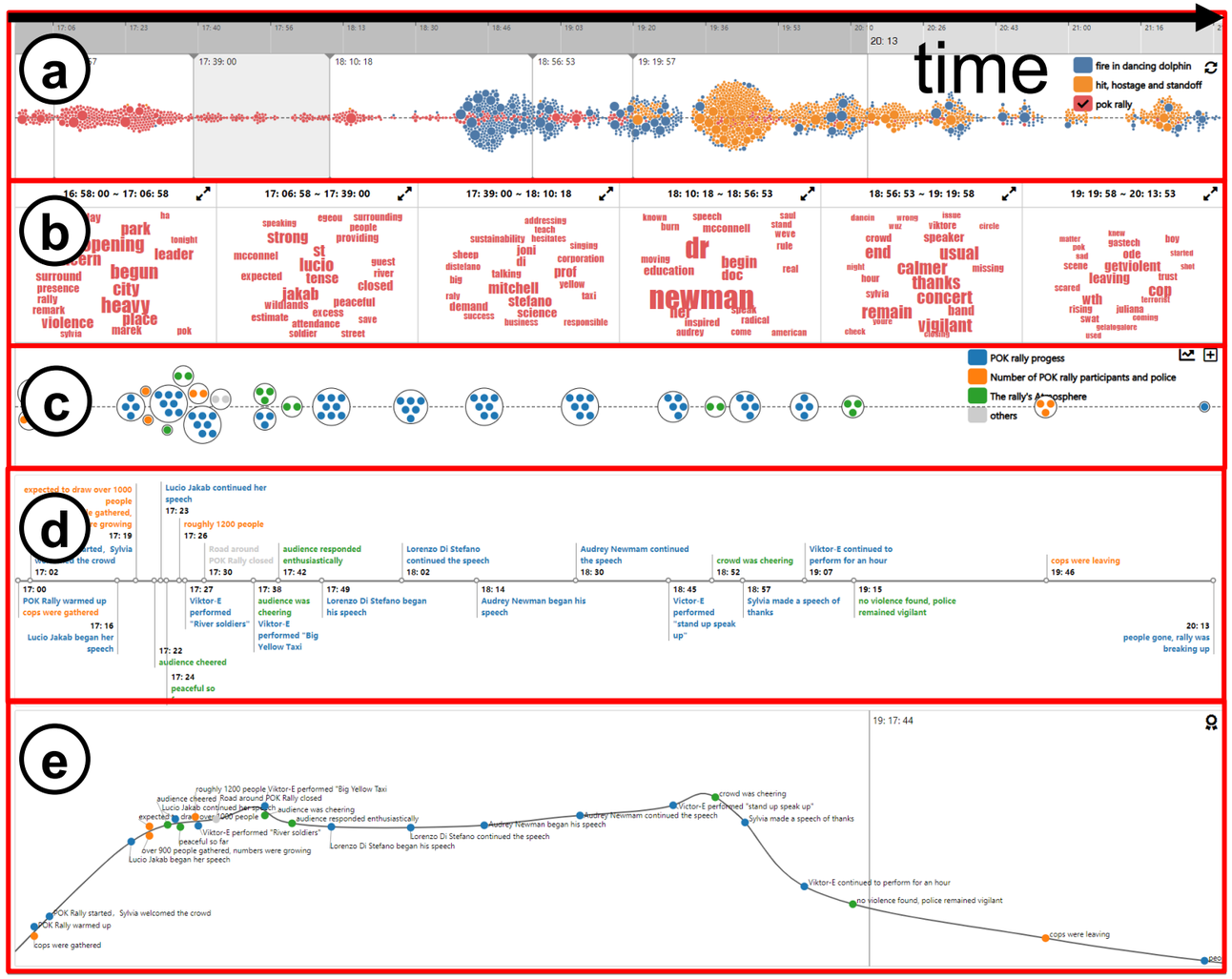
System properties:

- top-down
- human-led
- narrative visualization creation system

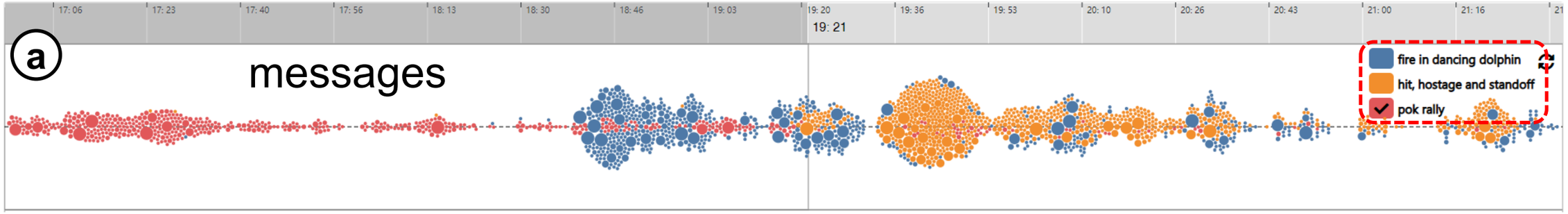
Steps:

- create unit events
- assess the level of risk

Third system



Migration of the risk



the risk was only from the rally event

17:00 — 18: 39

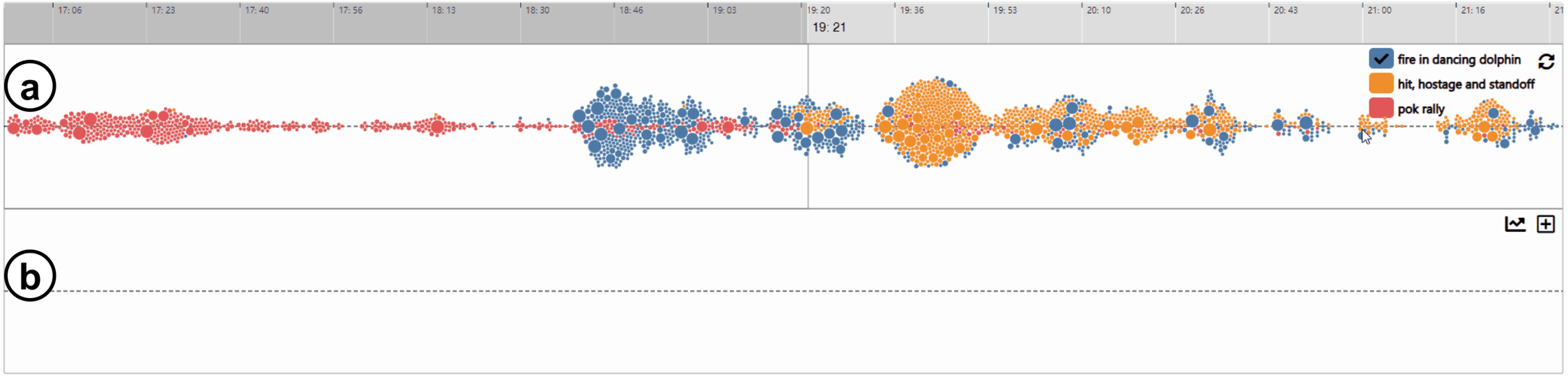
the risk was concentrated on the fire event

18: 39 — 19: 32

the risk coexisted in the fire event and the hostage event

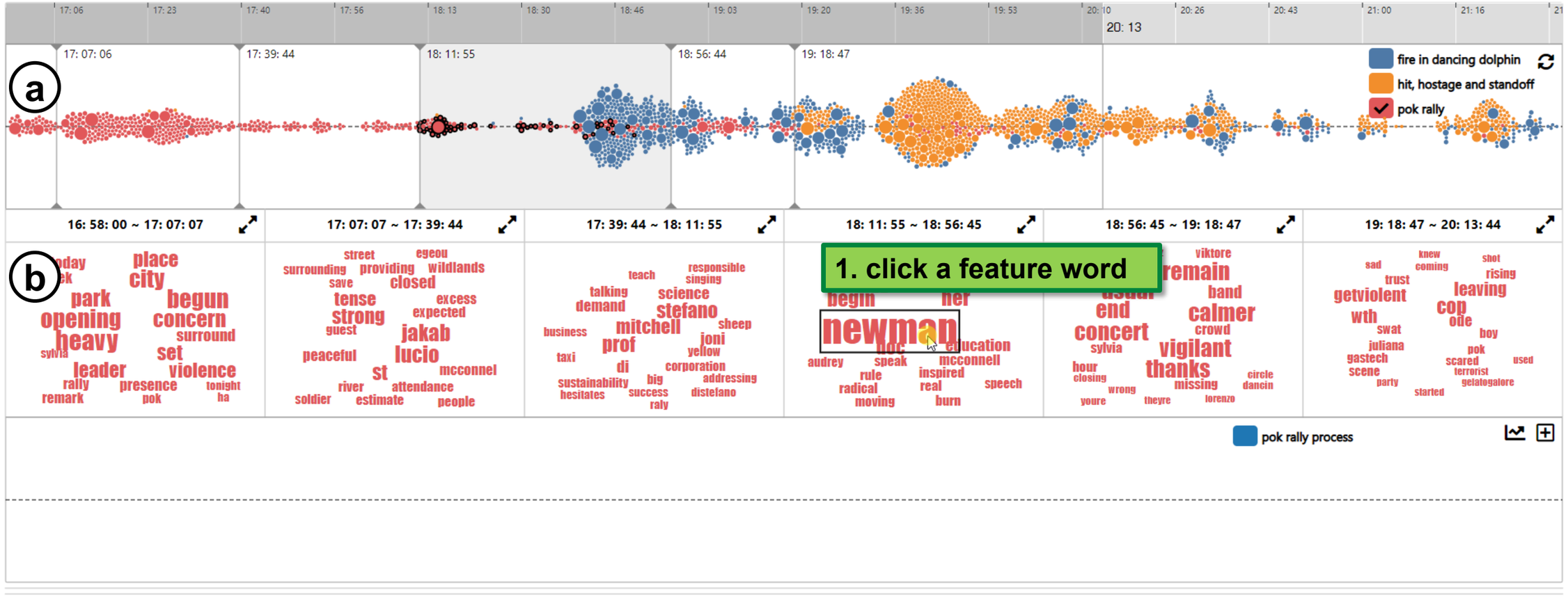
19: 32 — 21: 31

Create small multiples

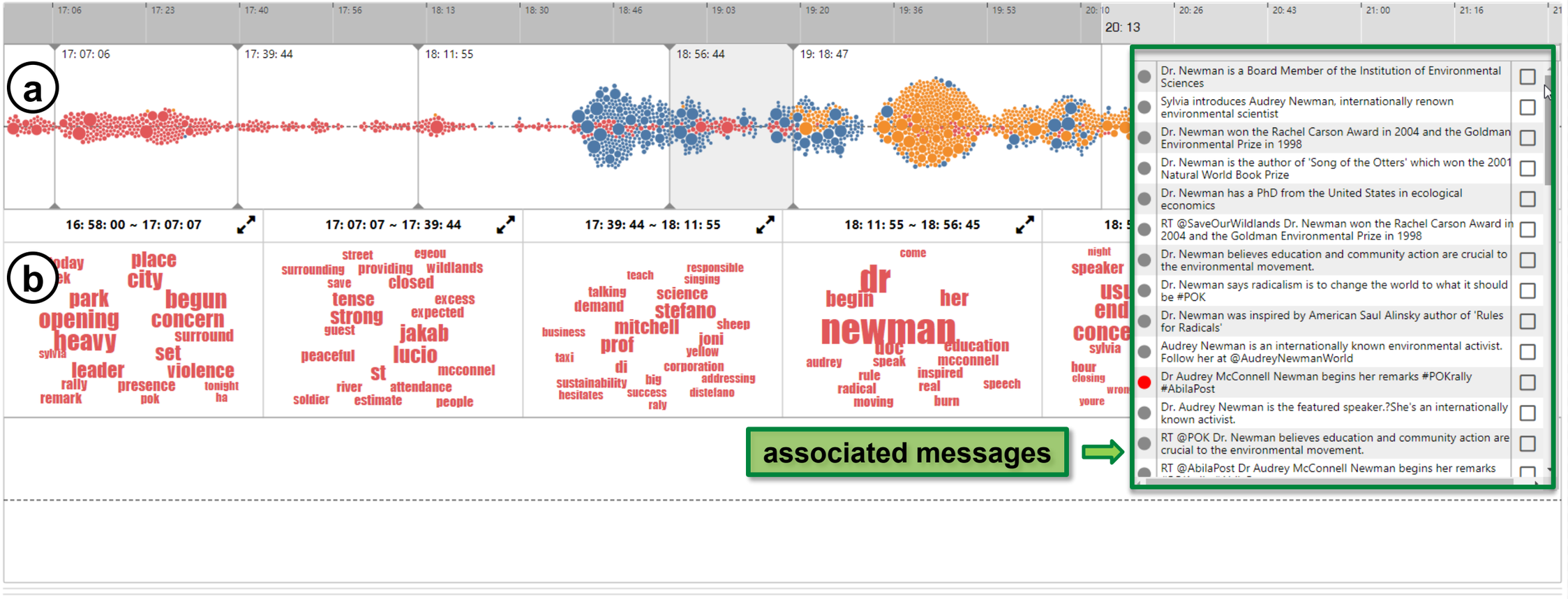


small multiples: feature words

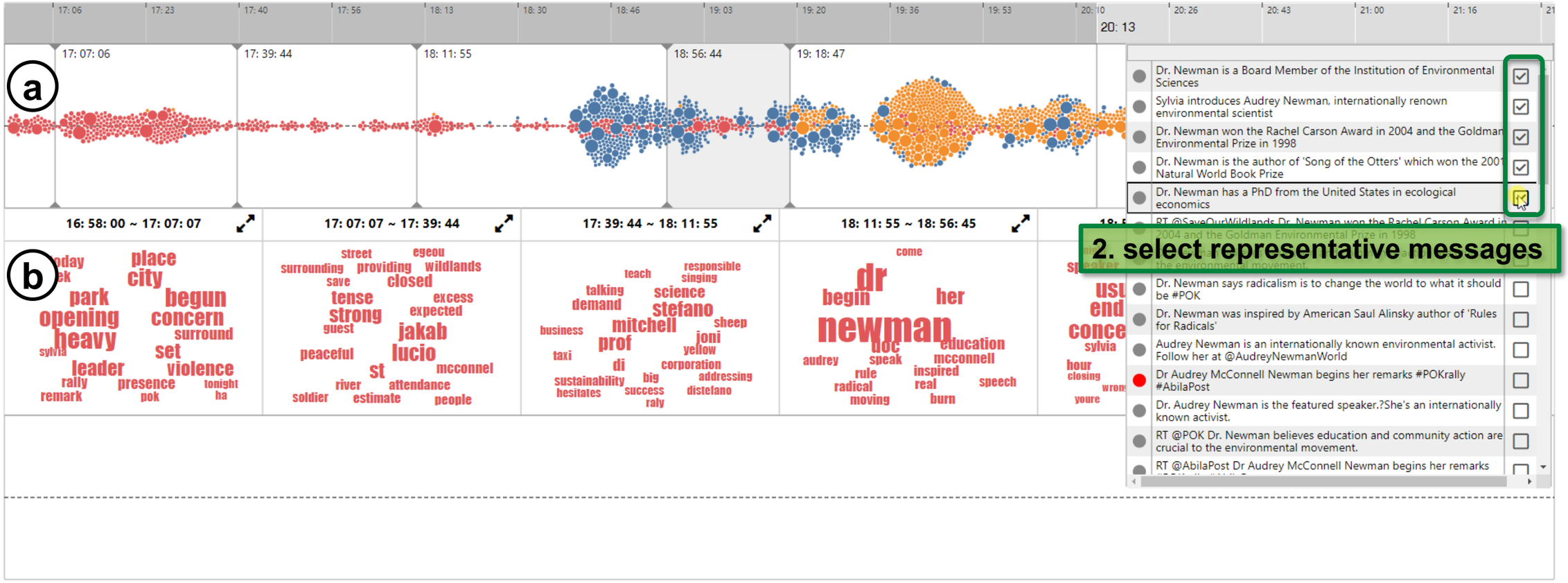
Steps to create a unit event



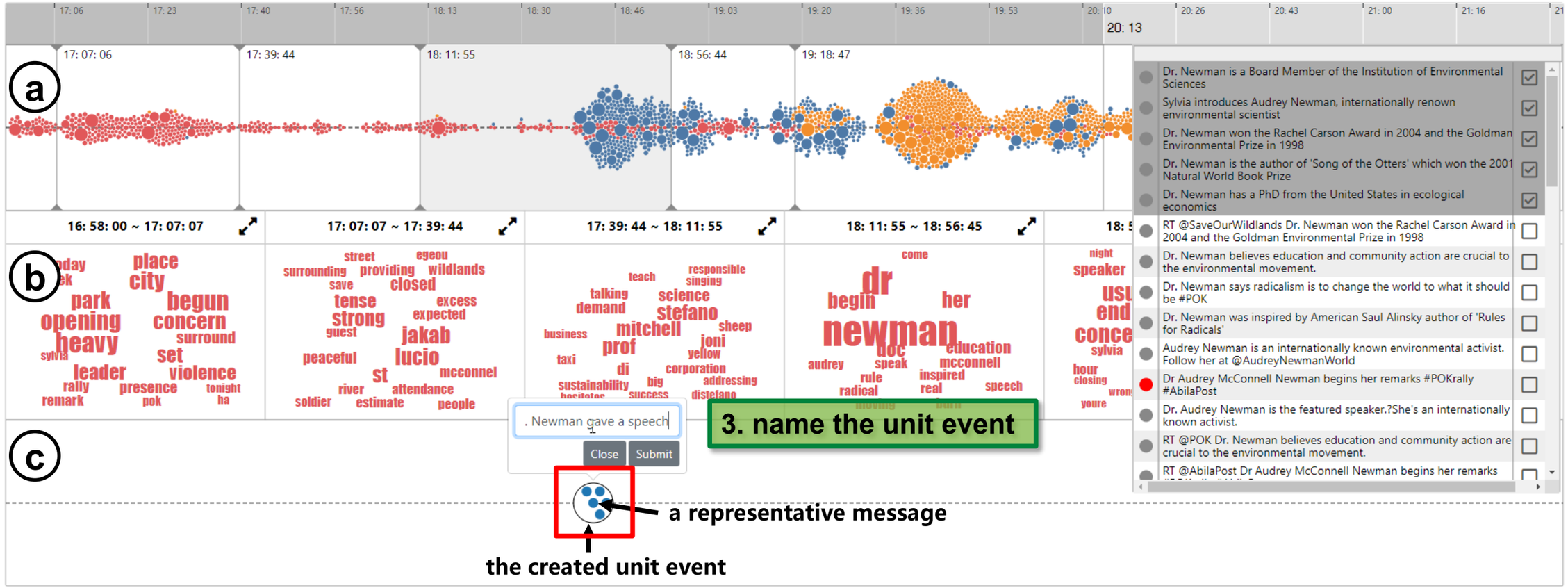
Steps to create a unit event



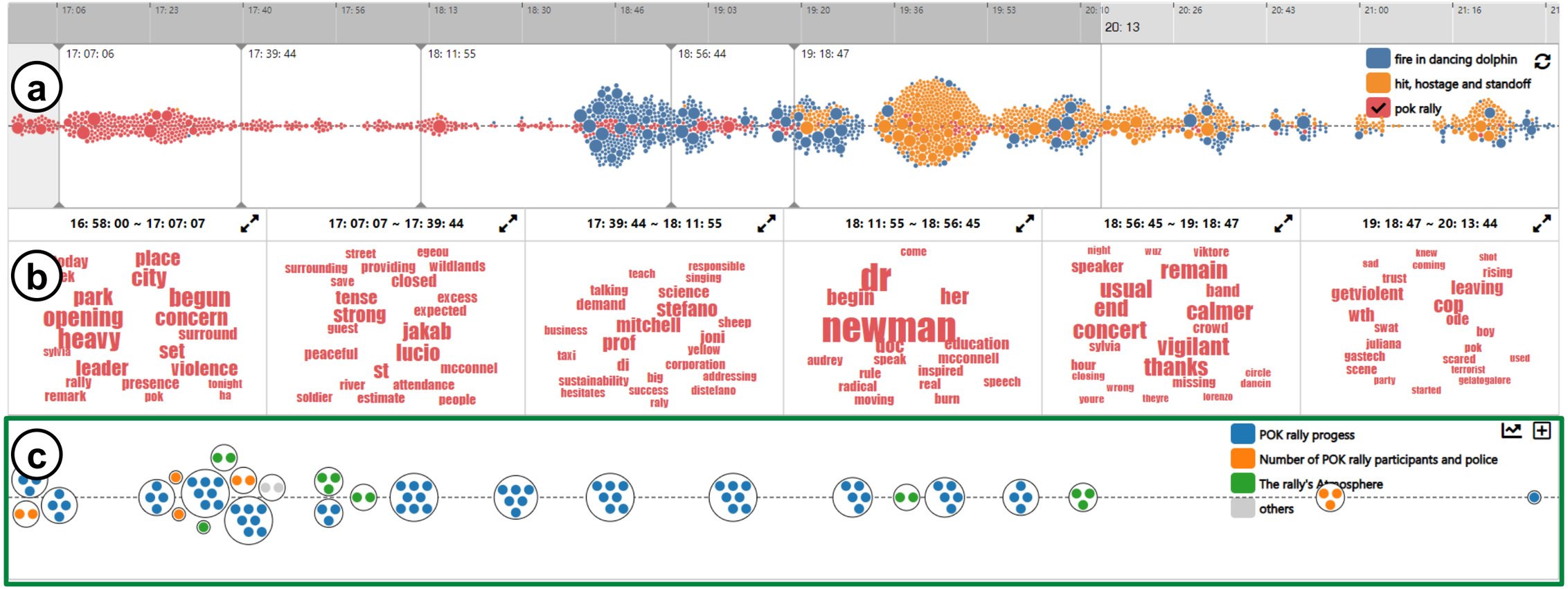
Steps to create a unit event



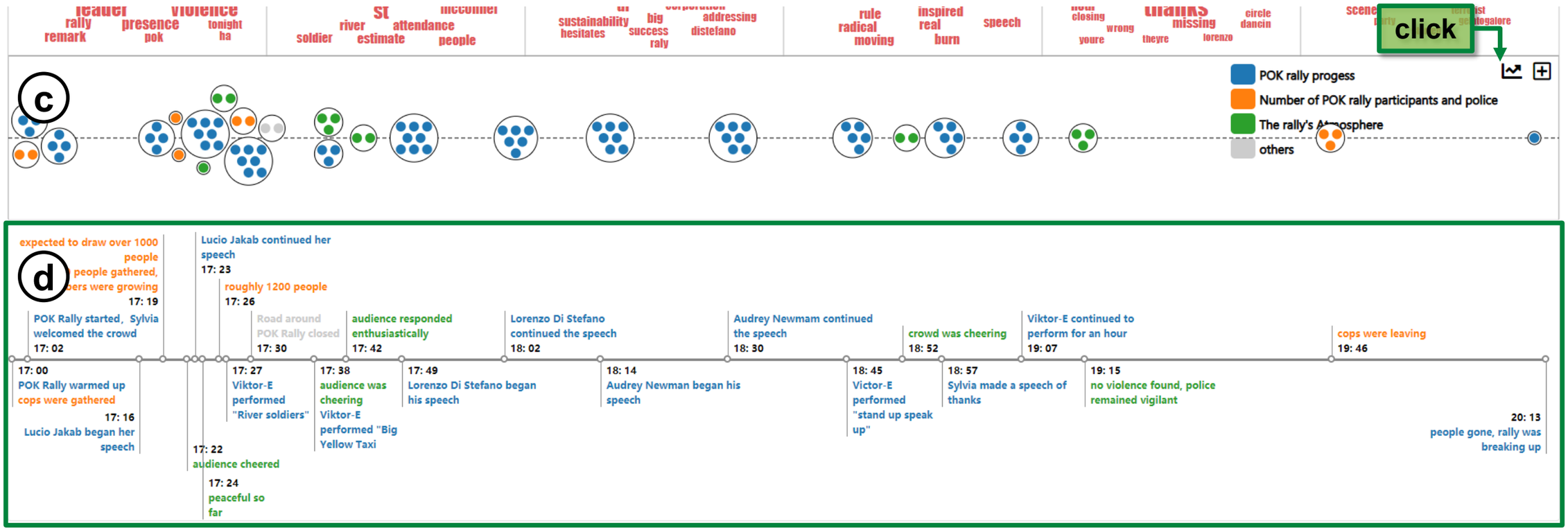
Steps to create a unit event



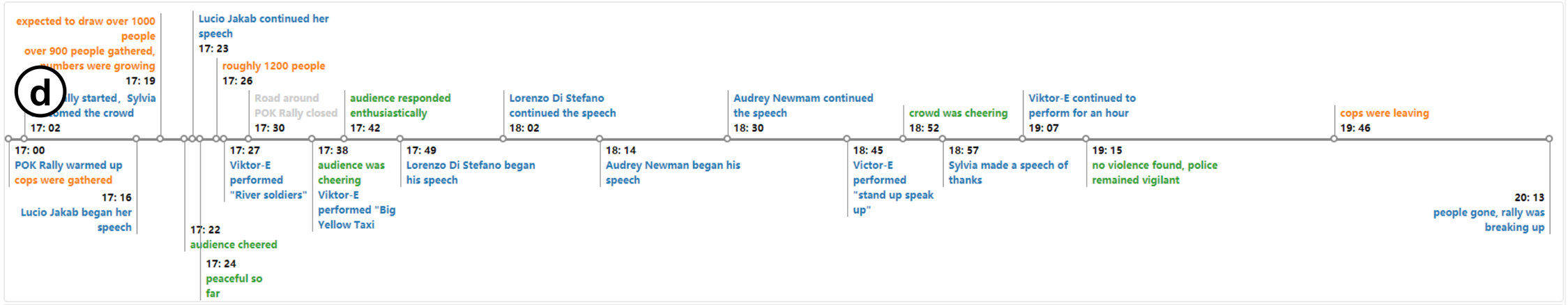
Narrative visualization: a timeline consisting of unit events

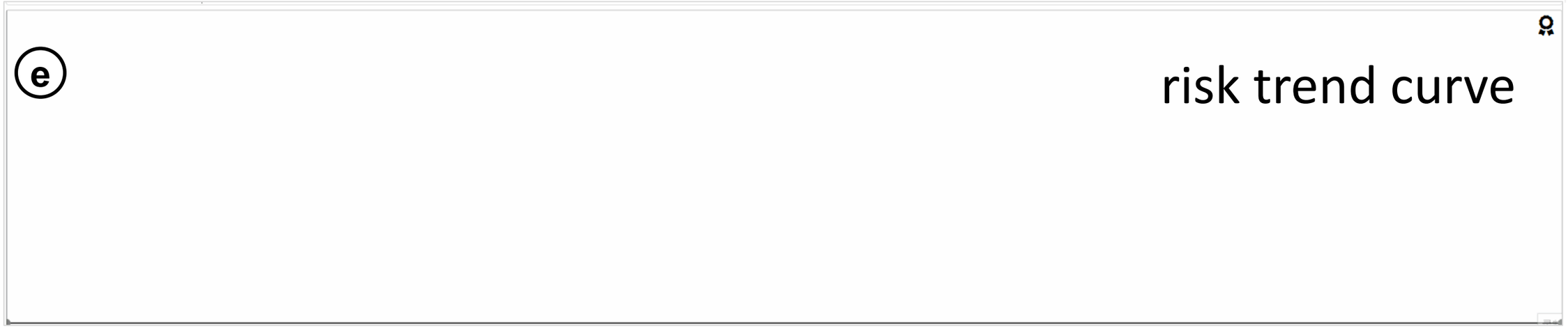


Narrative visualization: a timeline with explicit event names

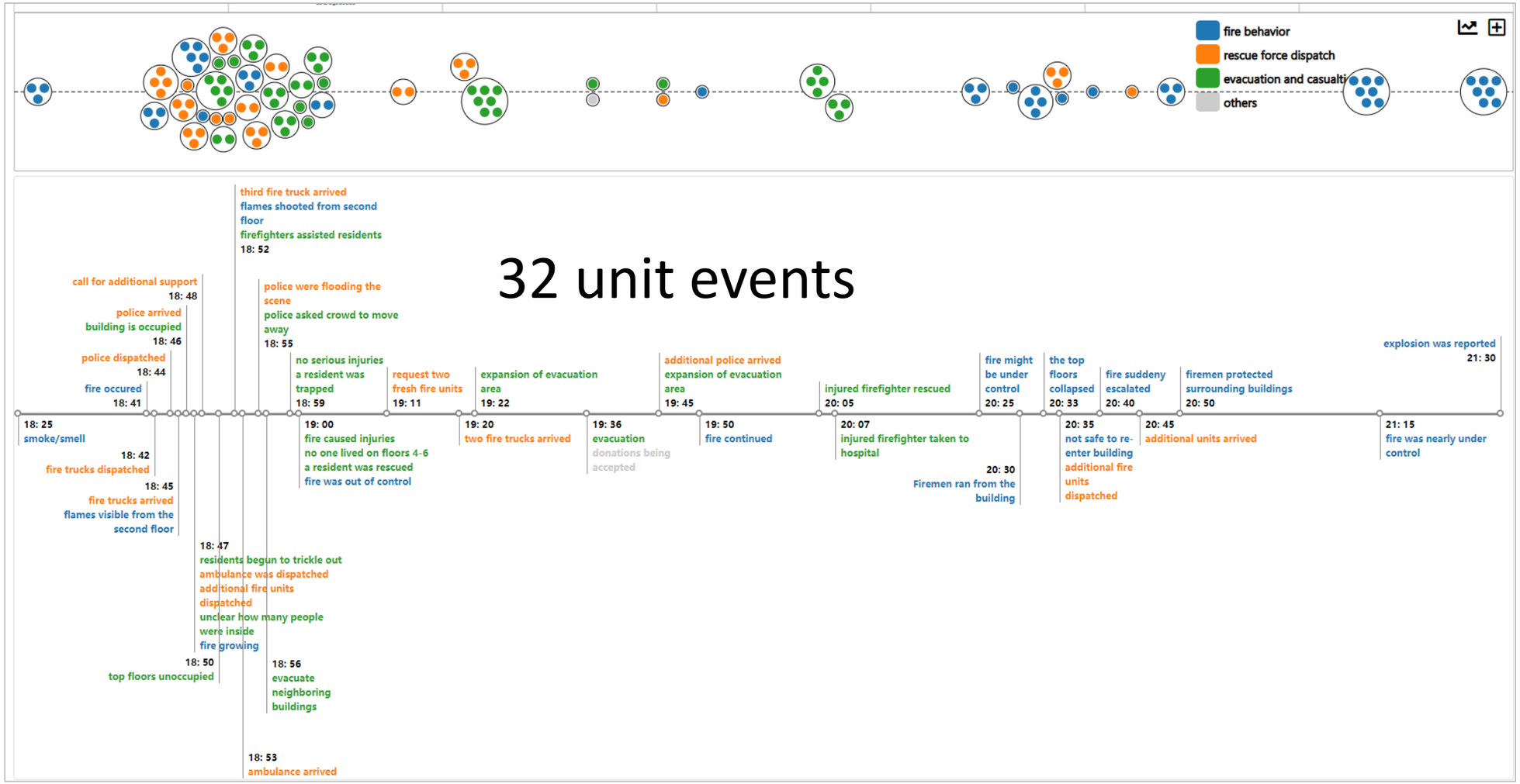


Narrative visualization: risk trend curve

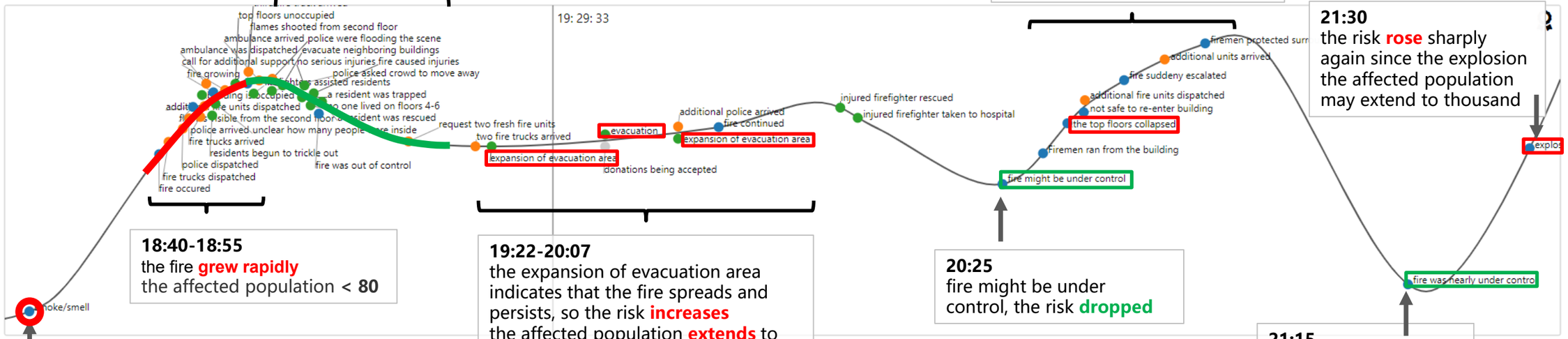




Risk assessment of the fire event



Risk assessment of the fire event



18: 55-19:11
the risk **dropped** as residents were evacuated and there were no serious injuries
the affected population < **10**

20:30-20:50
the top floors collapsed, the fire suddenly escalated, so the risk was **rising** suddenly
the affected population > **200**

21:30
the risk **rose** sharply again since the explosion
the affected population may extend to thousand

18:40-18:55
the fire **grew rapidly**
the affected population < **80**

19:22-20:07
the expansion of evacuation area indicates that the fire spreads and persists, so the risk **increases**
the affected population **extends** to the surrounding areas

20:25
fire might be under control, the risk **dropped**

21:15
the risk **dropped** as the fire was nearly under control

18: 25
smelled smoke
the risk began to **emerge**
the affected population < **10**

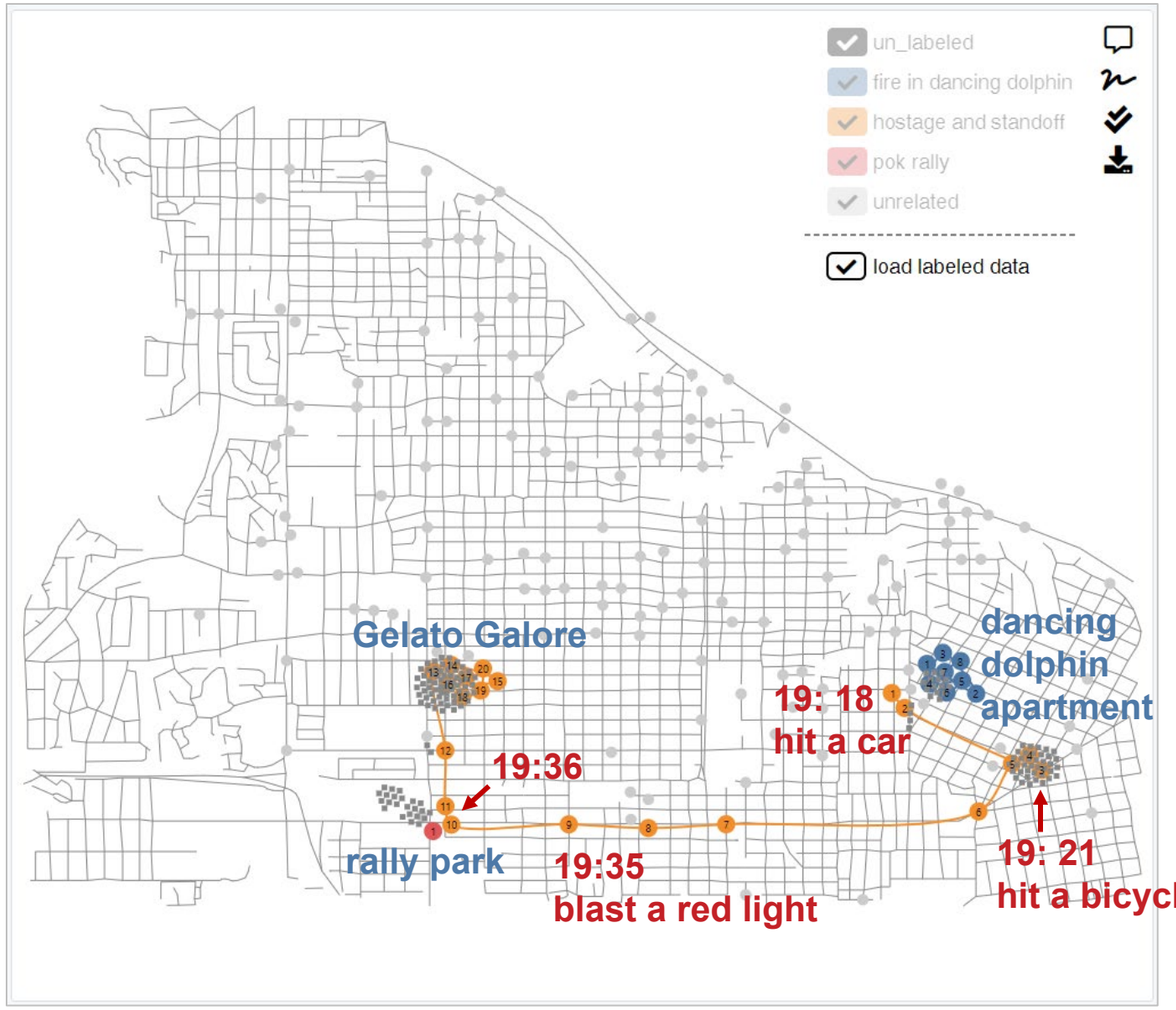
Q3: make decisions about where to dispatch a team of first responders in real-time analysis and retrospective analysis

From a retrospective perspective: fired apartment

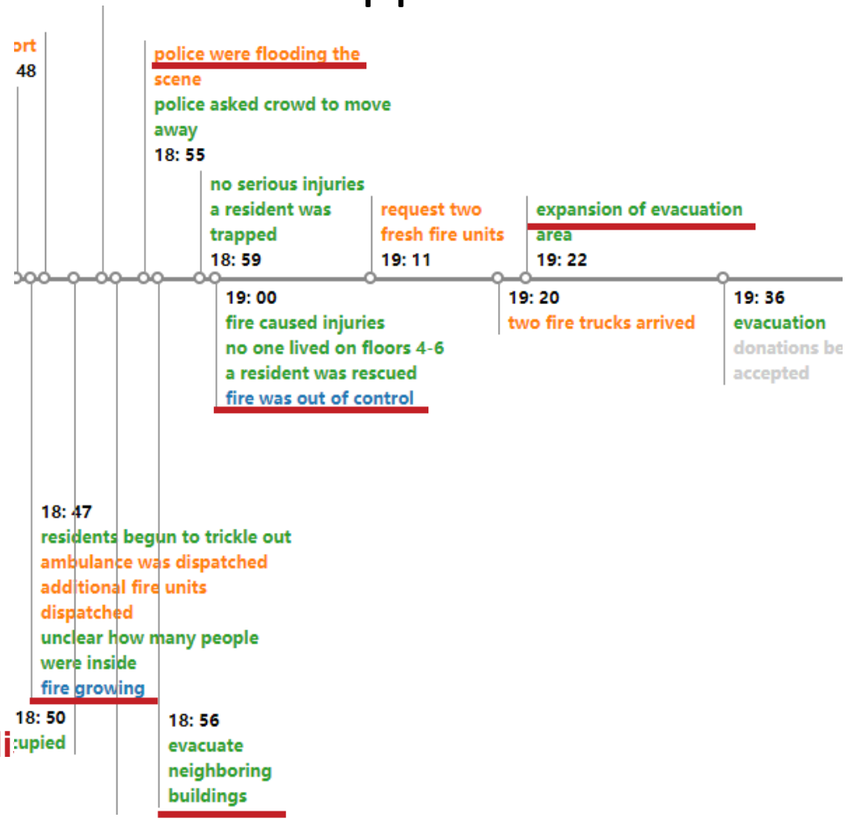
Principle: send the team to the location of the event with the most serious consequences

- real-time analysis: the potential and presumable consequences
- retrospective analysis: the actual consequences

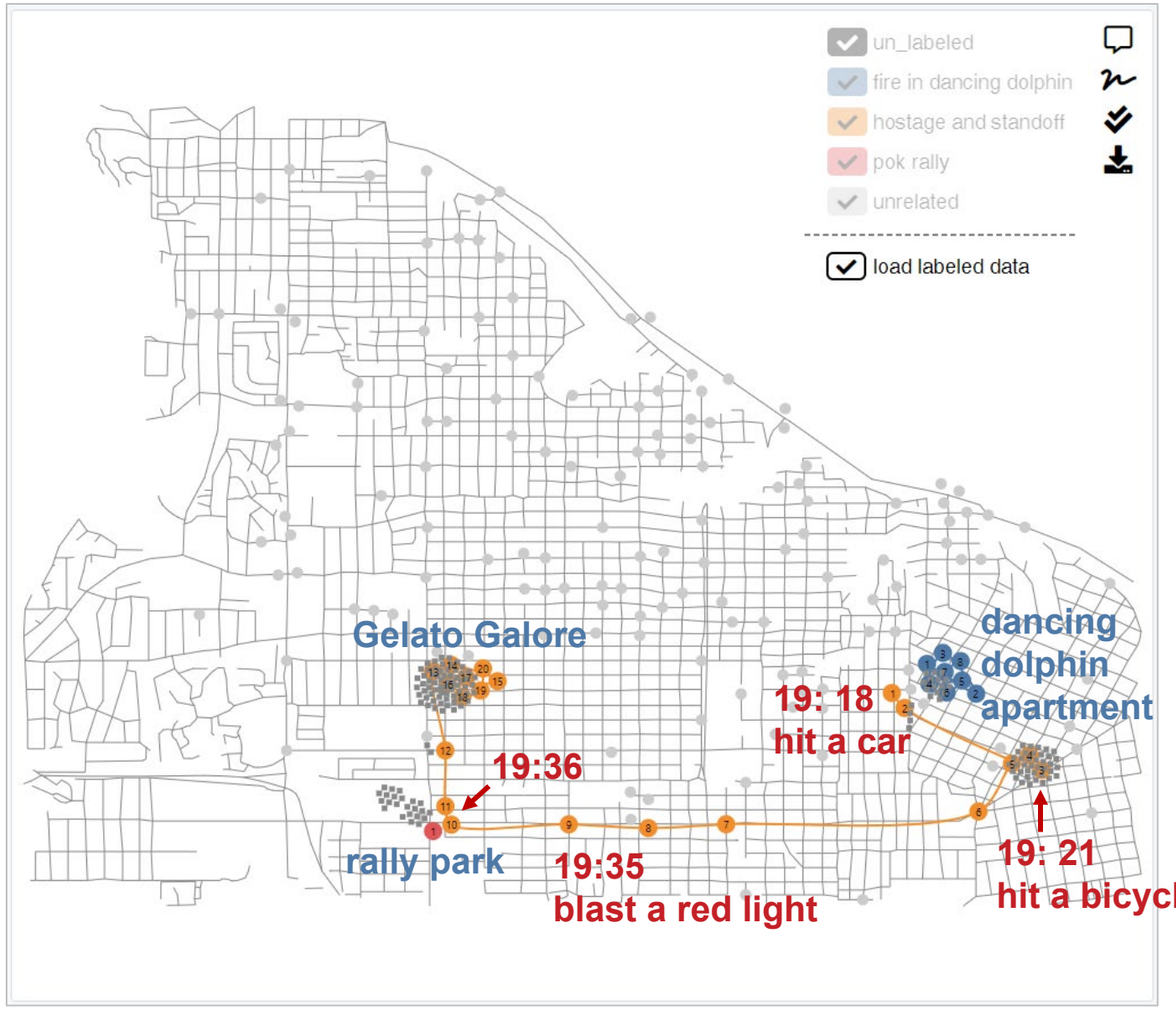
From a retrospective perspective: fired apartment



First, we suspect that the hostage event would not happen without the fire event.



From a retrospective perspective: fired apartment



First, we suspect that the hostage event would not happen without the fire event.

Second, compared to the other two events, the fire event actually has the most serious consequences.

From a real-time perspective: rally park

- **First,** the potential for mass violence at the rally is high.
- **Second,** in the event of mass violence, the cost would be enormous.
- **Third,** the rally precedes the other two events and its risk is predictable.

Q4: compare the solution of 2014 and the current solution

Differences between two solutions

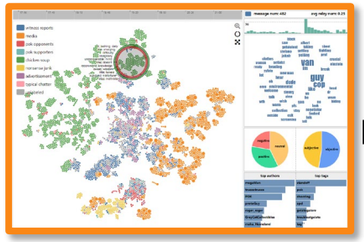
	solution in 2014	solution in 2021
Scenario	real-time, streaming data	retrospective， top-down
Who plays the lead role	system	analyst
What the system provides	visualizations in fixed template	creation tools for a narrative visualization
What does the analyst need to do	read and understand the given visualizations	structuring data as evidence, recording and analyzing evidence, and formulating conclusions as visualizations
How to record conclusions	remain in mind or be recorded in words	be drawn as a narrative visualization
Interactions	limited	rich

Conclusion

#1

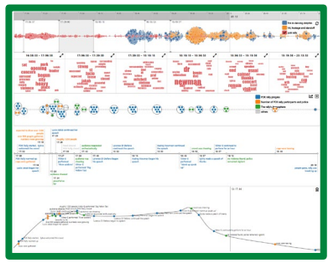
classified messages

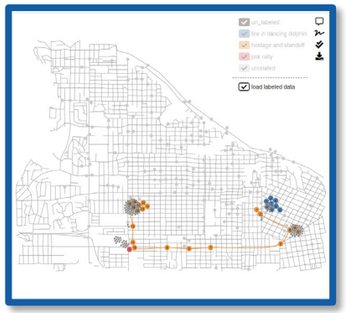
#2



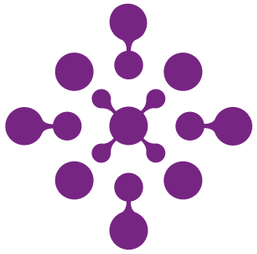
messages of three main events

#3





- Providing analysts with a flexible exploration environment at the beginning often yields twice the results with half the effort.
- The idea of interactive creation of narrative visualization is worth trying, especially for the problem that the required abstract semantics cannot be derived from algorithms.



VIS 2021

Thanks