Berkeley
UNIVERSITY OF CALIFORNIA

Modeling Informal Culture:

Conspiracy Theories, Hobbits, Fairy Tales and K-pop

Conspiracy Theories, Hobbits, Fairy Tales and K-pop
Hobbits, Conspiracies, and K-Pop

Timothy R. Tangherlini

Culture analytics?

- Data-driven analysis of cultural expressive forms and their dynamics
- Close integration of:
  - Data science theories & methods (mathematical models, machine learning, network science, etc)
  - Qualitative theories & methods related to cultural production, dissemination, and interpretation
    - E.g. literatures, languages, linguistics, history, art history, sociology, folklore, anthropology, archaeology, philosophy, ethnology, etc. etc.
  - So that the domain expertise across these domains complement each other
- Interdependence of these fields to address complex problems in the study of culture

Enabling Technologies

(how can we do this work?)

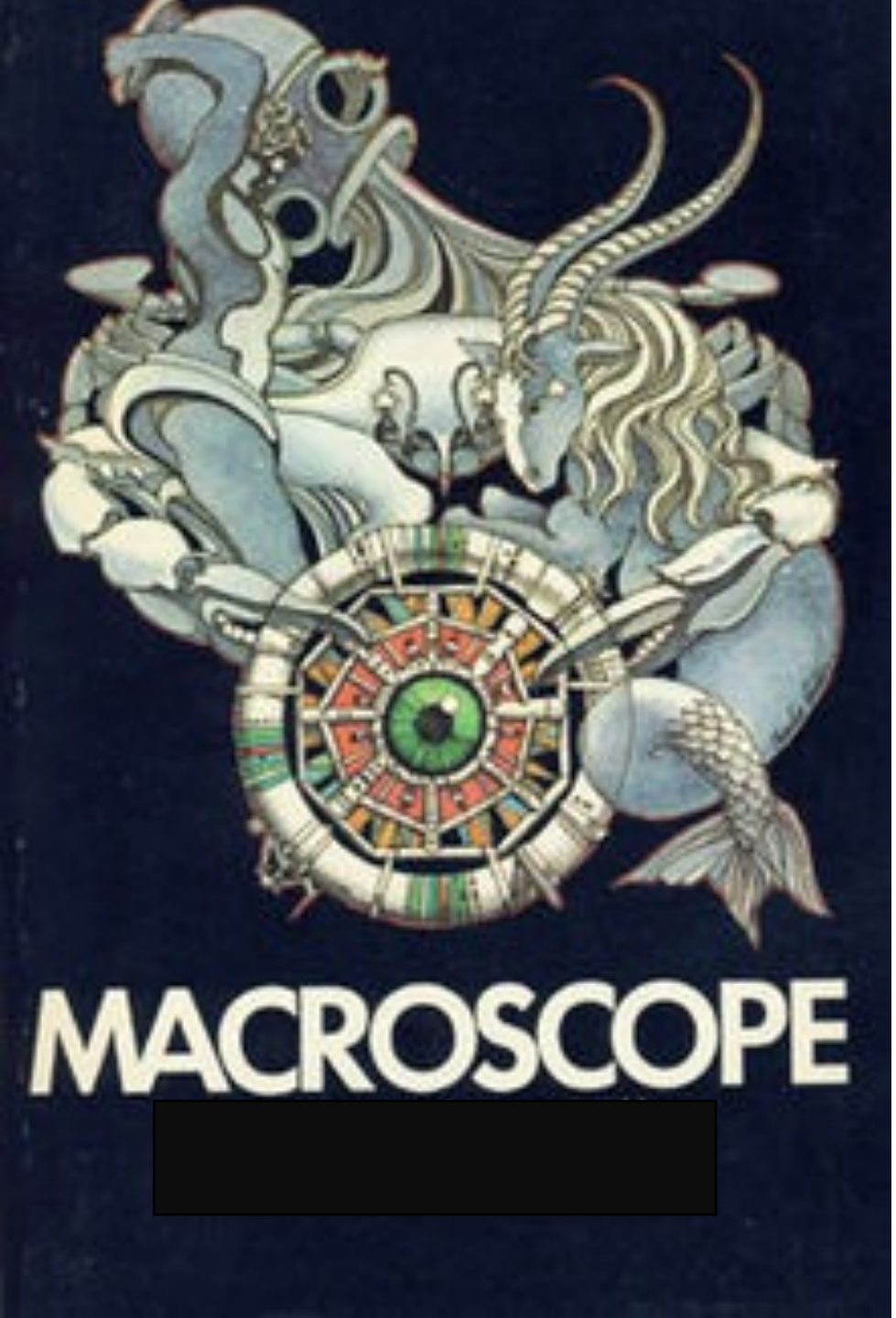
Algorithmic sophistication in:

- Natural Language Processing
  - Language models (BERT, GPT-3, etc)
- Machine learning
  - Supervised, unsupervised, semi-supervised methods
  - Deep learning
- (h)GIS – geographic information systems
- Graph analytics

Advances in digital collection development

- Archival collections made machine readable
- Geocoded resources
- Better methods for data storage, curation and retrieval

Broader access to eas(iesh)-to-use code and supercomputing



Finding Patterns: Macroscopic Analysis

- Macroscopes provide a ‘vision of the whole,’ helping us ‘synthesize’ the related elements and detect patterns, trends, and outliers while granting access to myriad details. Rather than make things larger or smaller, macroscopes let us observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend

-Katy Börner (2011)

Informal Culture & The Study of Folklore

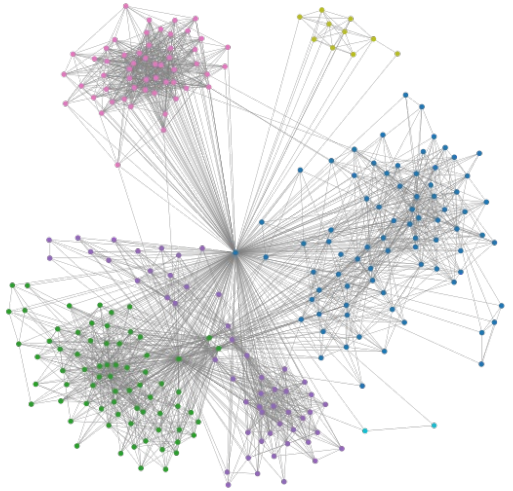
Four experiments and (tentative) results

Informal Culture

- Social media and “always on” internet provide us with an unusual opportunity to explore the dynamics of informal culture at very large scale
- Internet forums, video sharing sites, communities of practice
- How do we explore these phenomena at scale?
  - Do we need new methods and theories?
  - What can we learn?

Informal culture and the study of folklore

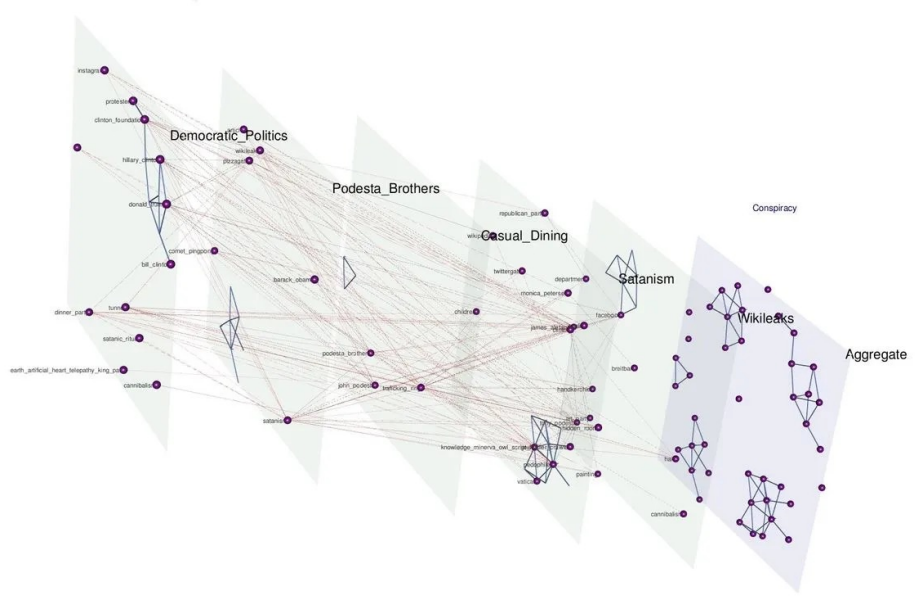
- Folkloristics focuses on understanding the dynamics of the productive dialectic between individuals and the groups to which they belong
- Folklore itself can be seen as informal expressive cultural expressive forms circulating on and across social networks
- Many of the domains of informal culture have features that are folkloric in nature



Four (macroscopic) experiments in Culture Analytics

- Conspiracy Theories
  - Noisy storytelling on social media
- Hobbits
  - Social Reading at Scale
- Danish Fairy Tales
  - Embedding ~1900 fairy tales in a classifier space
- K-Pop dance
  - Toward a dance search engine

Experiment #1



Conspiracy Theories

Conspiracy Theory in the Time of Covid-19

From Pizzagate to #StoptheSteal

With Vawni Roychowdhury and the Narrative Modeling group at UCLA

General problem

- Stories on social media are:
  - Noisy
  - Partial (incomplete)
  - Based on knowledge of the ongoing conversation
- We want to be able to estimate the parameters driving any conversation on social media so that we can understand:
  - The “boundaries” of the discussion
  - The structure of the conversation and its elements
  - The emergence of “narrative communities”
  - The robustness/resilience of narrative frameworks

Based on several (theoretical) propositions

- “In every discourse, whether of the mind conversing with its own thoughts, or of the individual in his intercourse with others, there is an assumed or expressed limit within which the subjects of its operation are confined” (George Boole [1854] 1958, 42)

- It should be possible to estimate an underlying story space, made up of actants and interactant relationships
  - We represent this space as a (dynamic) network graph, where actants are nodes and their interactions constitute edges

Semantic and Syntactic Challenges

- Users choose different words/phrases for a single entity or concept
  - No ideal semantic encoding
  - No standard/ideal method for similarity measurement
- Same actants appear in various domains
- They are referred to with different phrases/mentions
- Solution: Parsing at sentence level followed by aggregation via semantic encodings

Story Model (Part A)
Sentence Level Relationship Extraction

Sentence	arg1	Verb	arg2	Type
The argument is that 5G causes coronavirus.	{5G}	{causes}	{coronavirus}	SVO
Bioweapon escapes from Chinese laboratory	{Bioweapon}	{escapes}	from Chinese {laboratory}	SVP
this virus is the perfect bio weapon.	this {virus}	{is}	the perfect bio {weapon}	SVcop
They said millions were gonna die.	{millions}	were gonna {die}	-	SV

Story Model (Part B)

Actants to SuperNodes

- Actant: an entity or entities captured in arg1 or arg2 that serve the same or similar semantic roles
- Supernode: an automated grouping of highly related entities
- Entity Discovery:
  - Named entity Recognition discovered entities belonging to a set of predefined entity classes
  - Headwords in noun phrases
  - Frequency: number of phrases

Ranked Entities
China, people, virus, wuhan, trump, ccp, covid19, jews, government, cdc, lab, italy

Supernode frequency	Mentions
2966	virus, viruses, corona
1158	gate, gates, bill
520	doctor, doctors, nurse

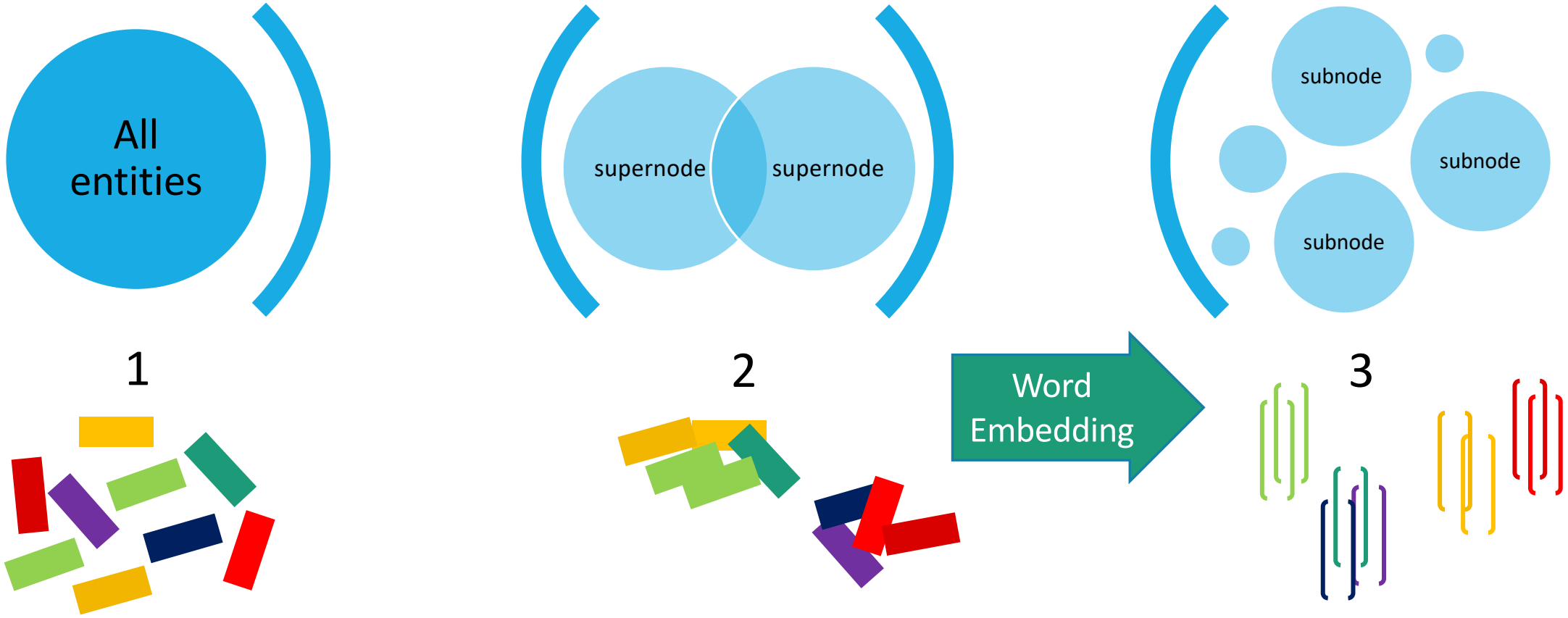
Story Model (Part C)

Computing Subnodes

- Subnode: same context subset of a supernode
  - Uses clustering methods in an embedding space (BERT).
- How do we generate subnode label?
  - Each cluster is assigned a label, combination of:
    - Most frequent words
    - High score nodes

Subnode	Phrase Example
death,flu	'a huge decrease in Flu deaths', 'The regular flu deaths', 'The regular flu deaths', '170 deaths from the Flu', 'flu and pneumonia deaths', 'the influenza rate', 'other flu outbreaks', 'the flu deaths', 'the flu deaths'
Flu, common	'this flu', 'to seasonal flu', 'the common flu', 'the normal seasonal flu', 'normal flu', 'a weaker flu', 'the normal flu', 'vs flu', 'the common flu', 'the common flu', 'to regular flu', 'this flu', 'this flu'

Story Model (Part B and C)
Pretrained Phrase Embeddings with BERT

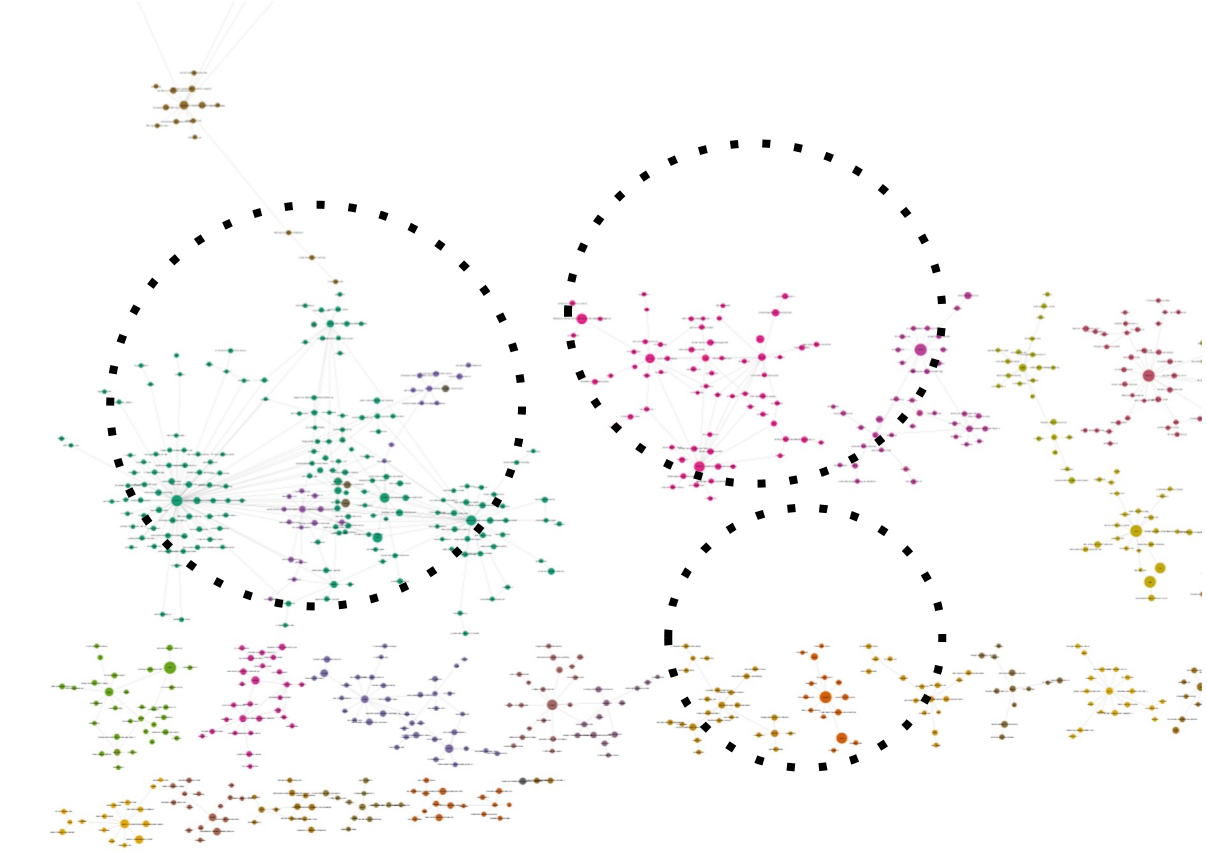


Story Model (Part D)
Edge Triplets in our Final Network

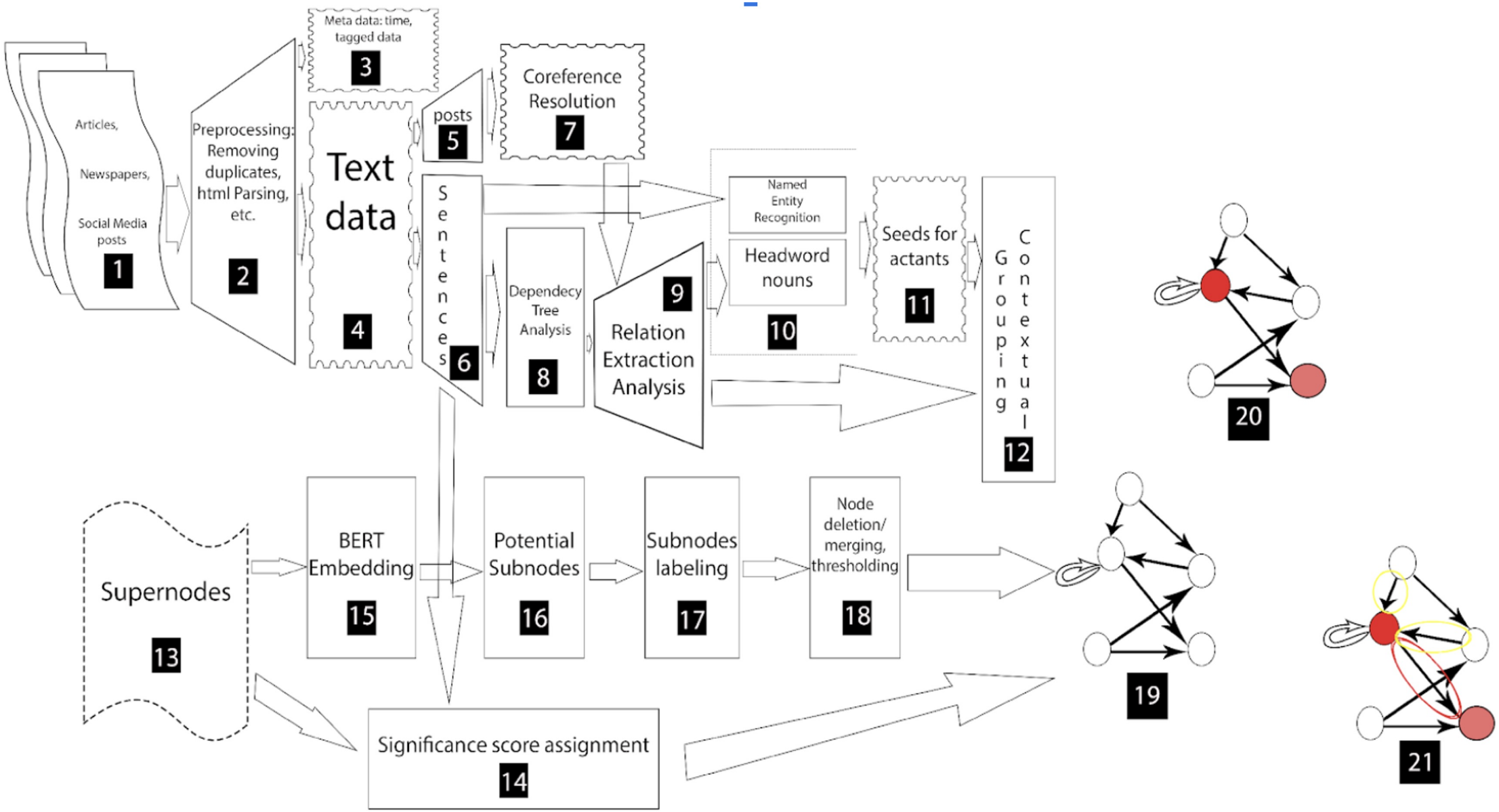
- For each subnode pair, we aggregate sentence level relationships
- We connect subnodes with the relationship triplets.
  - Examples:
  - Gates {obsessed} with exploding,population
  - Gates {funded} facility,faulty,practice,research
  - Gates {create/patented/funded/injected/programmed} coronavirus
  - Coronavirus {is} bioweapon
  - 5g wave {created/carry/cause} virus

Story Model (Part E)
Community Detection in Our Final Network

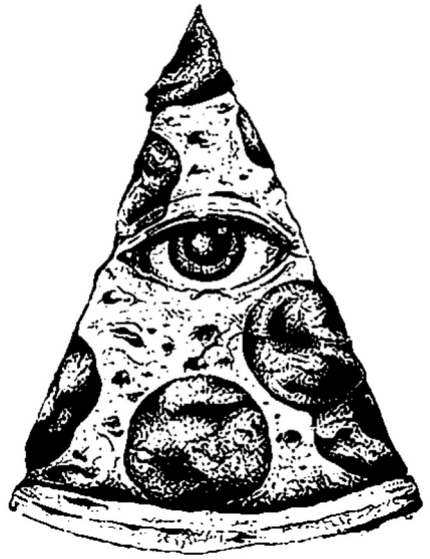
- Social media network consists of numerous narratives
- Community detection algorithm help us to find dense part of the graph
- A statistical method using heuristic community detection (Newman) helps us find stories evolving on social media



Bird's Eye view of our pipeline



Some noteworthy findings



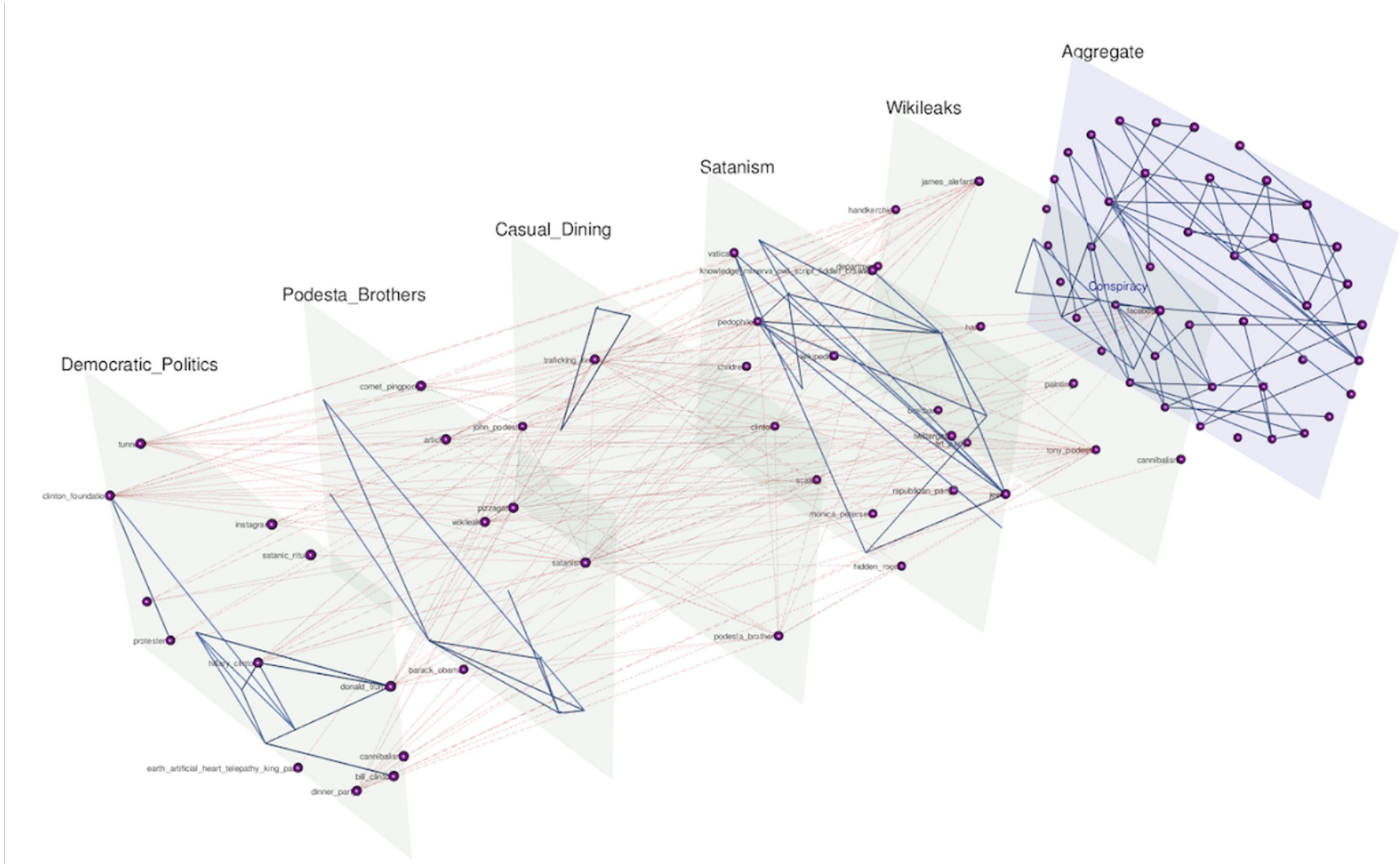
Pizzagate

Some noteworthy findings

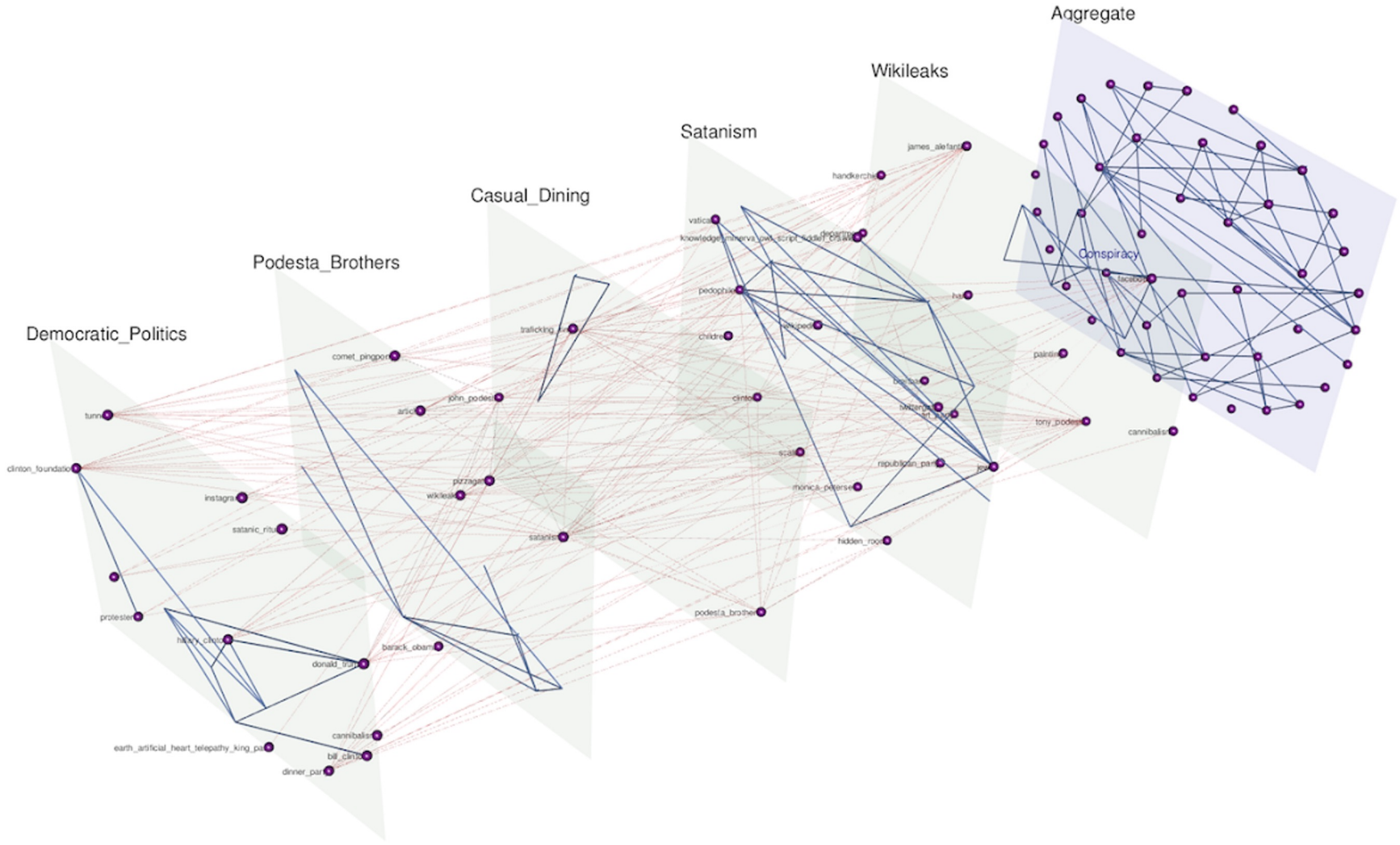
Topological feature of a conspiracy theory?

- Conspiracy theories are fragile
  - They are not robust to deletion!
  - There exists a small, separating set that allows one to attack the conspiracy theory's narrative framework graph
- Unfortunately, they are resilient
  - It is very easy to "add" the deleted edges/nodes back into the narrative at the next retelling!

The narrative network of a Pizzagate WITH Wikileaks



The narrative network without WikiLeaks



Very easy to add the links back in

Danish witchcraft legends

Some noteworthy findings

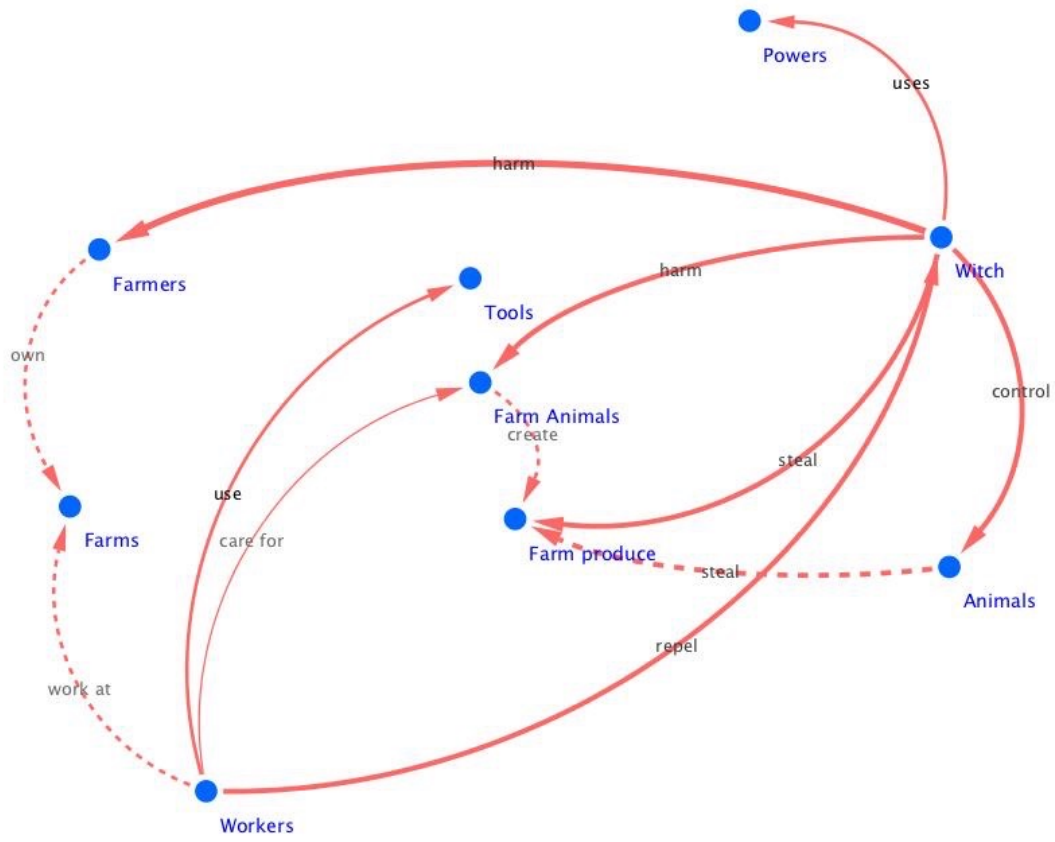
Narrative framework structures through time

- Topological features of the networks of interlocking stories have stability across time and language/cultures
- Simple threat narratives can be linked together to form complex representations of threatening groups
- That can have considerable impact on various communities
- These frameworks appear to have considerable stability across time once they are established

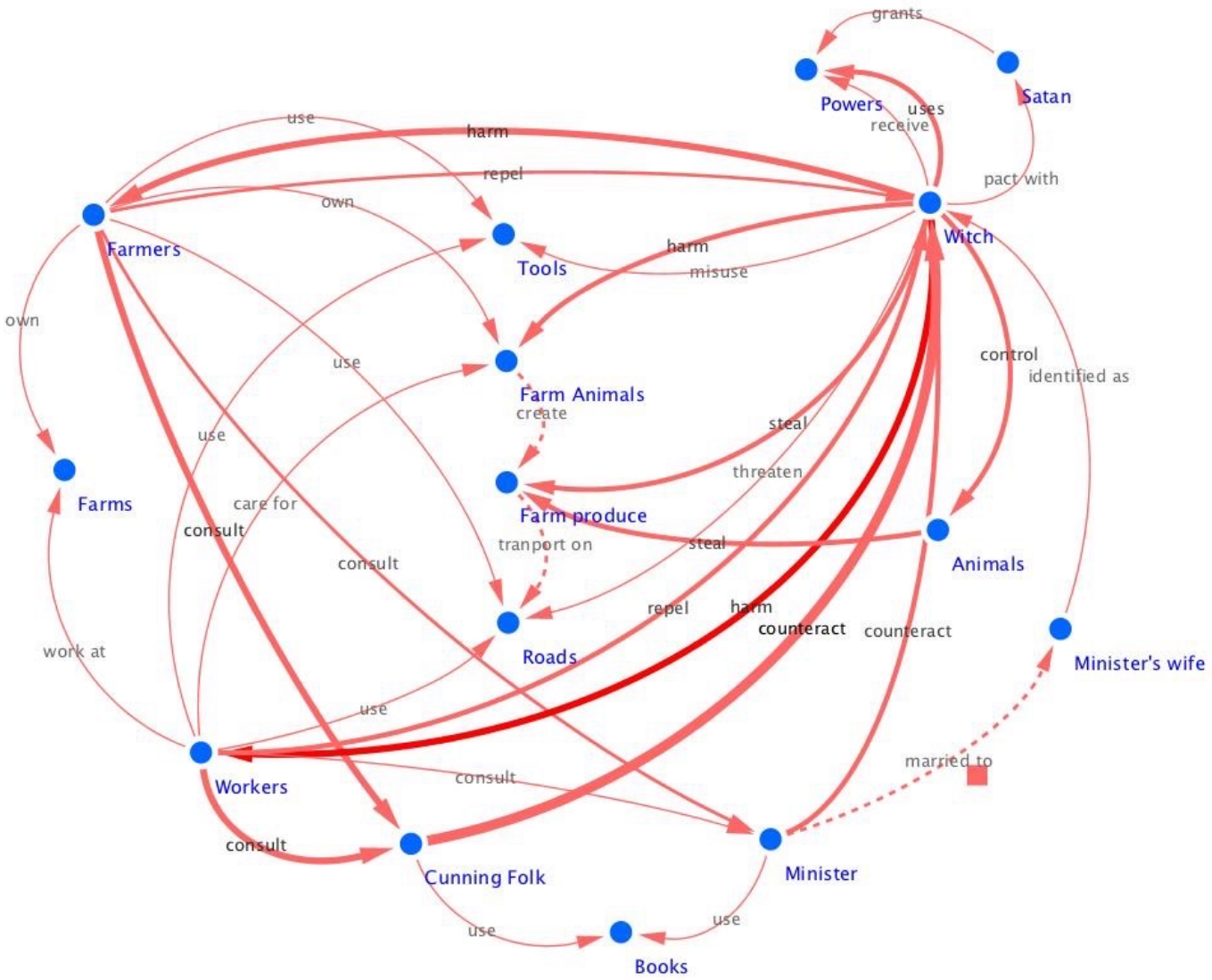
Select a legend from the domain

Phrase (observed)		Structure	Role	id (act-rel)	Specification
Et sted i Sundeved	}	Orientation:	Where:	Farm	Sundeved
boede en kone, som hed Else Mikkels	}	Orientation	Who:	Person	Else Mikkels
Denne kone gik omkring og gjorde så megen fortræd,	}	CA: Threat	Harm agent:	Witch	Threatens people
hun tog mælken fra koerne	}	CA: Threat	Harm agent:	Witch	Steals milk from cows
eller listede sig ind i husene og kastede noget for svinene, så de foer forstyrrede og forvildede omkring	}	CA: Threat	Harm agent: Harm loc:	Witch Inside	Harasses swine In the houses
Kom så nogen, skabte hun sig om til en hare,	}	CA: Threat	Harm agent:	Witch	Turns into hare
og når pigerne om morgenen kom ud i marken for at malke	}	Orientation	Who: Where: When:	Hired girls Fields Morning	
kunde de se hende gå og luske omkring henne i et hjörne af marken, og da var hun en hare.	}	CA: Threat	Harm agent: Harm loc:	Witch Fields	Runs in field as hare
En dag hun som sådan løb omkring,	}	CA: Threat	Harm agent:	Witch	Runs in fields
blev hun skudt i det ene ben	}	CA: Strategy	Protection:	Shoot	Shot in leg
og hun måtte da ligge til sengs i lang tid	}	Result	Harm: positive	Injury	Bed long time

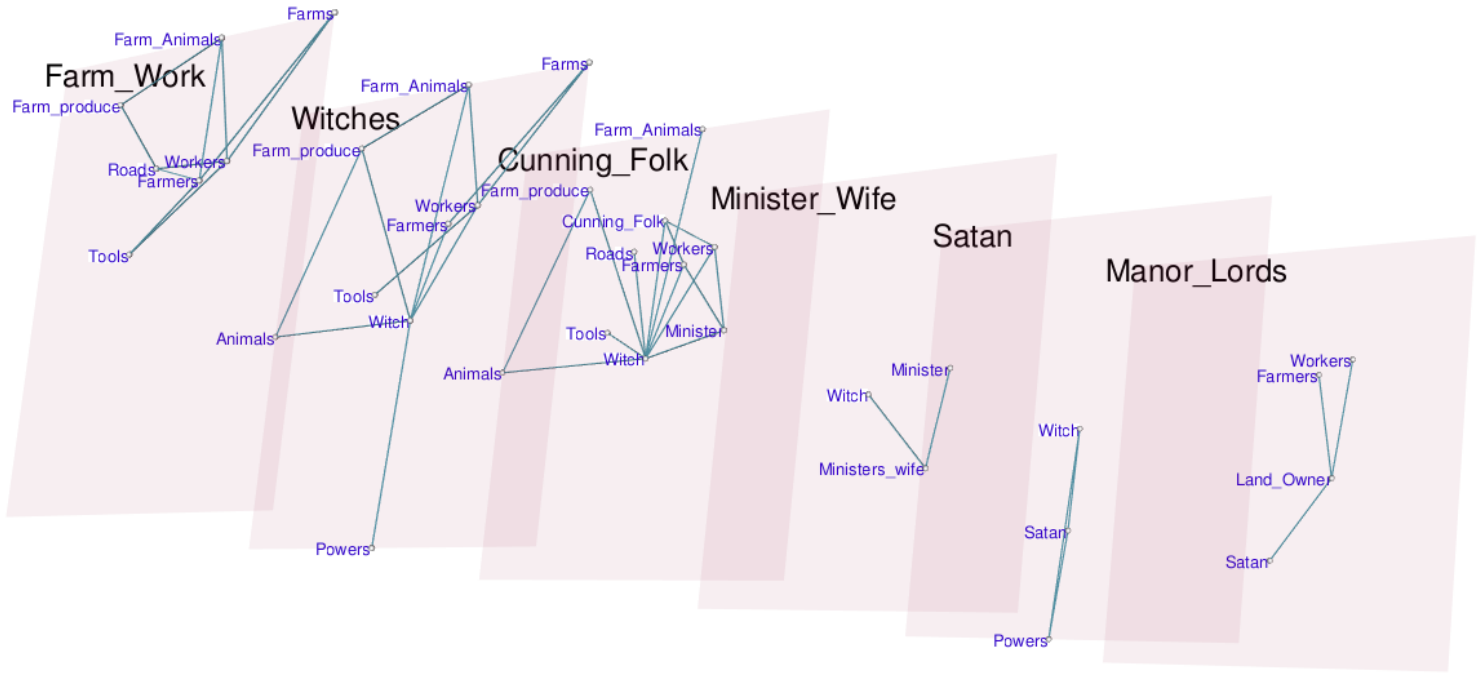
We find the actants and relationships in each story



We aggregate this over 100s of stories

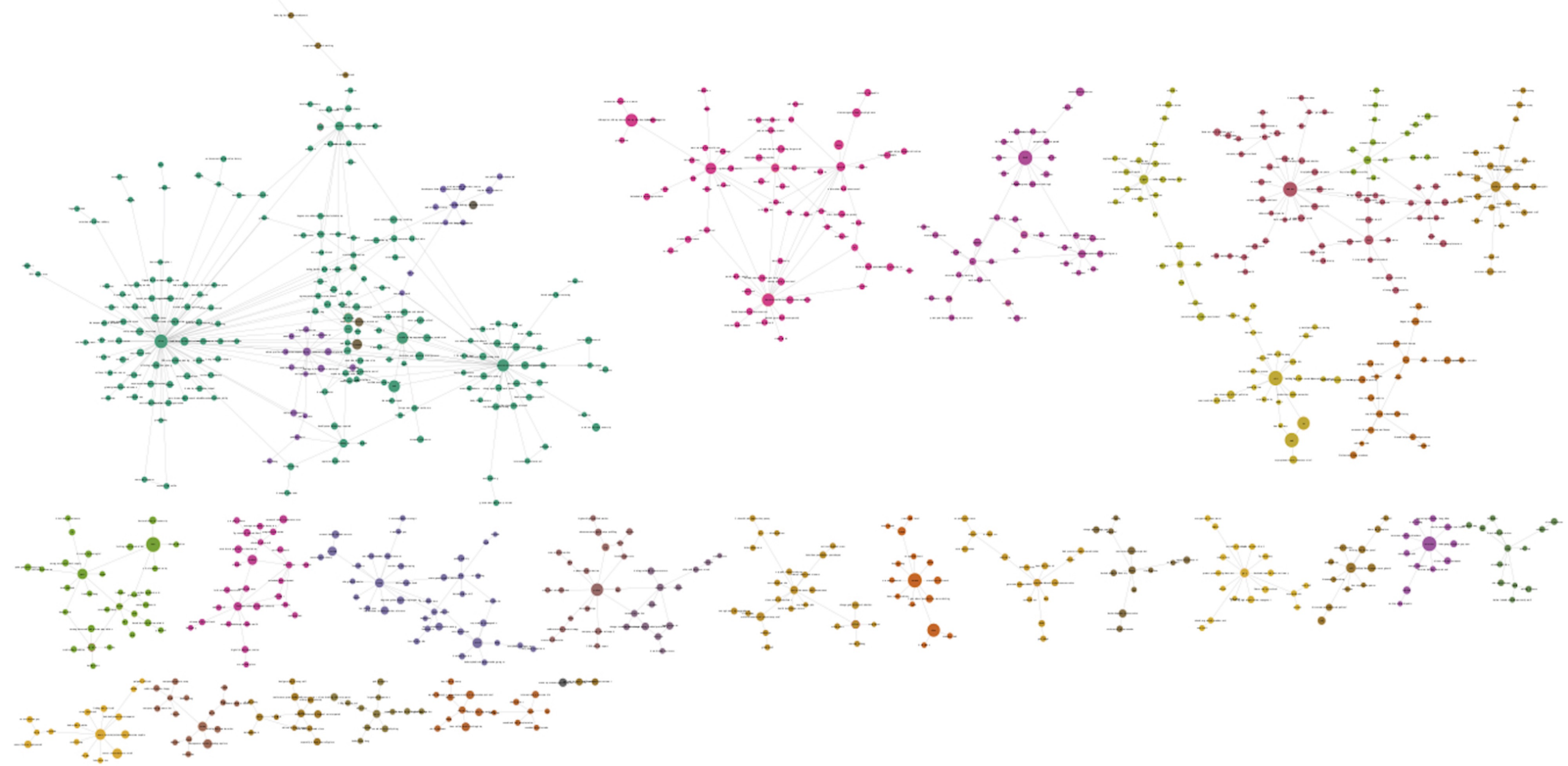


We discover a series of interdependent domains



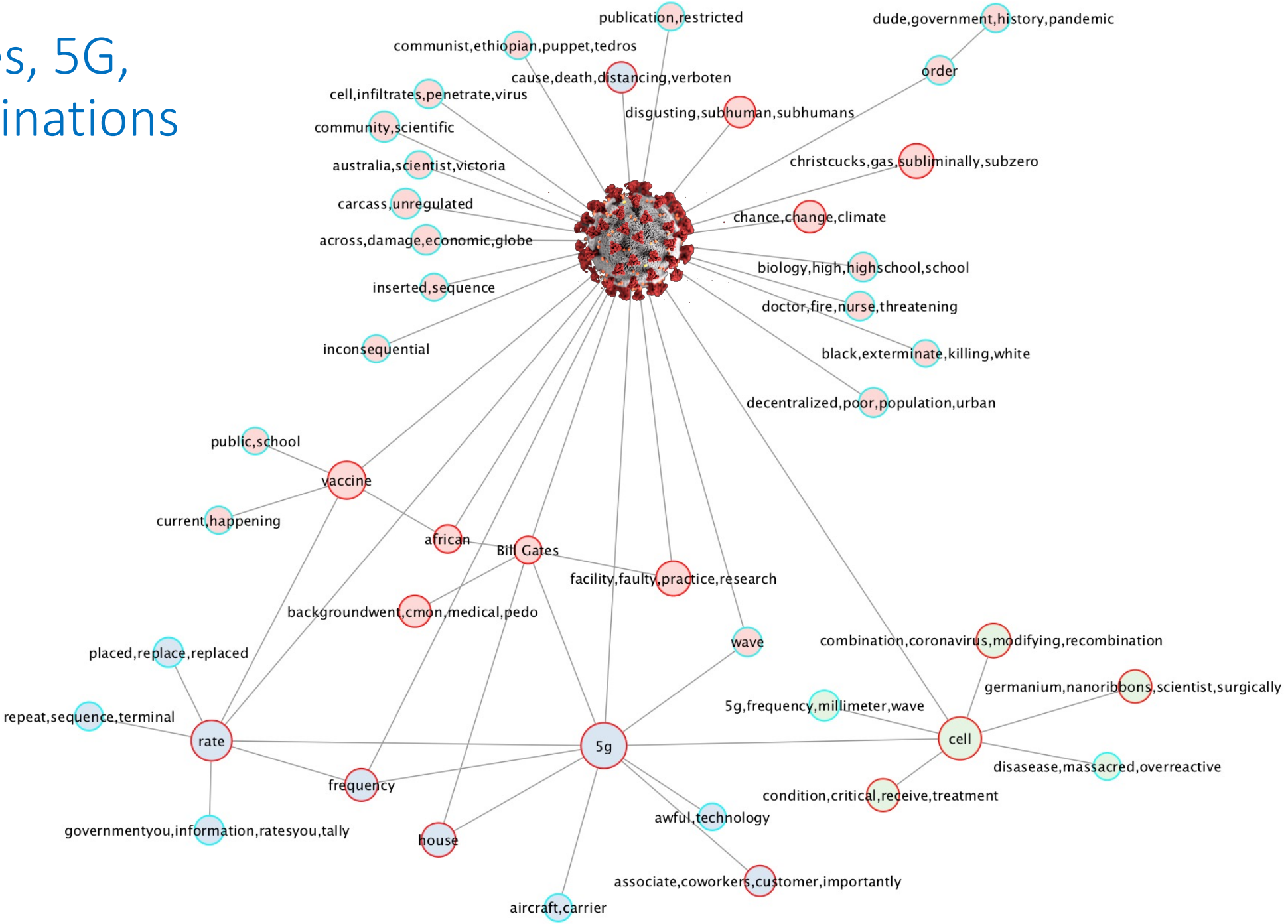
Conspiracy in the time of Covid

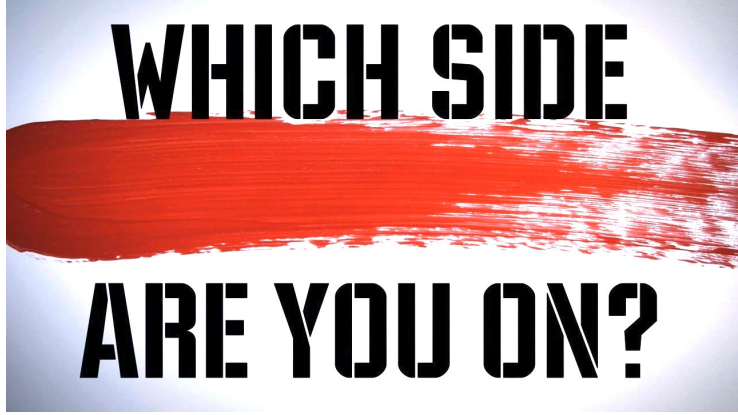
Some noteworthy findings



Conspiracy theories today seem to have ***highly separated communities*** with separate contexts.

Gates, 5G, Vaccinations





Current research on social media and narrative frameworks

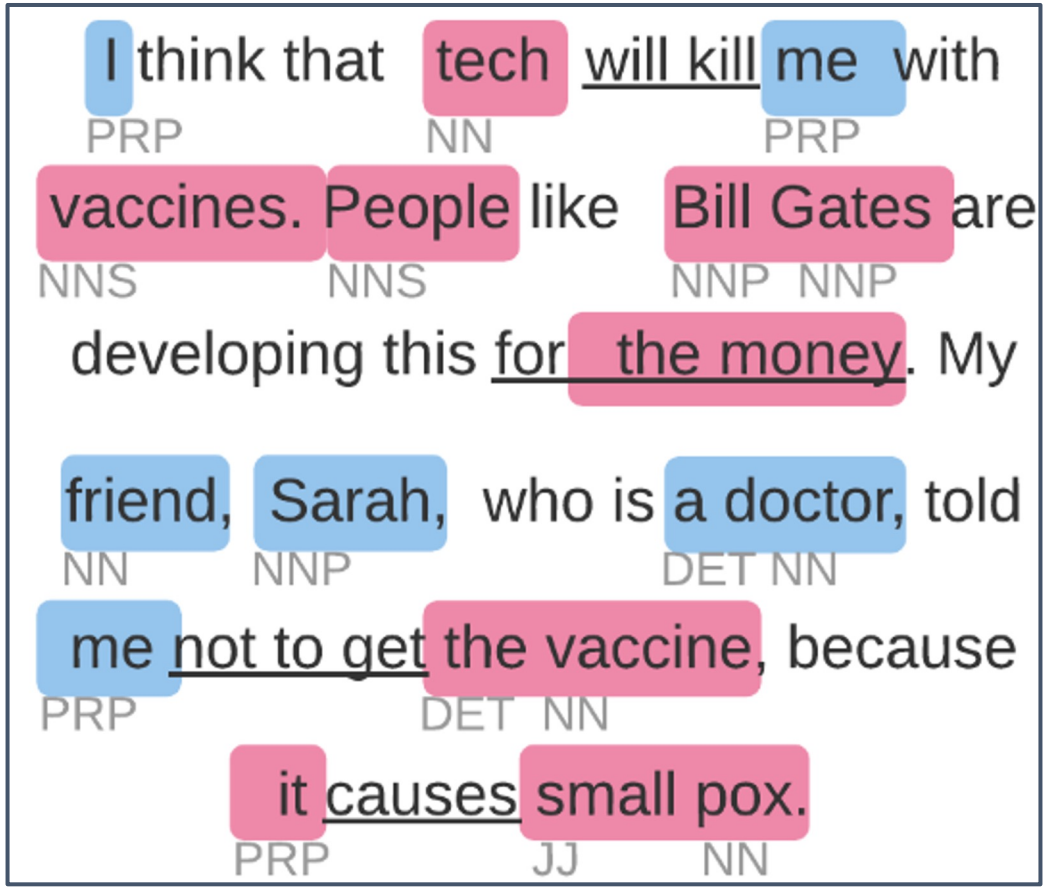
Conspiracy theories and threat narratives in contemporary society


Threat: “Stolen” Election
Strategy: Take over the Capital
Can I discover this in the noisy forums in places like *Parler?*

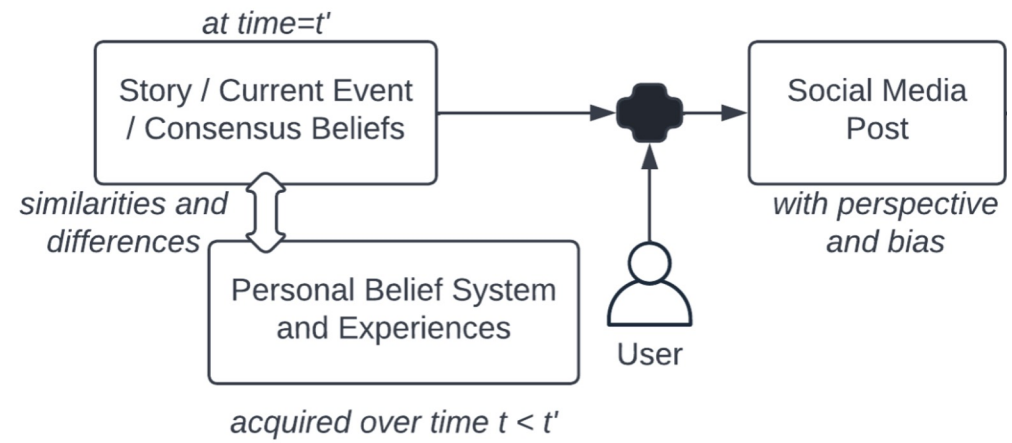
Some challenges

- With whom does the author of a social media post side?
- What do they oppose?
- The context of a post -- its language -- informs about the author’s opinions.

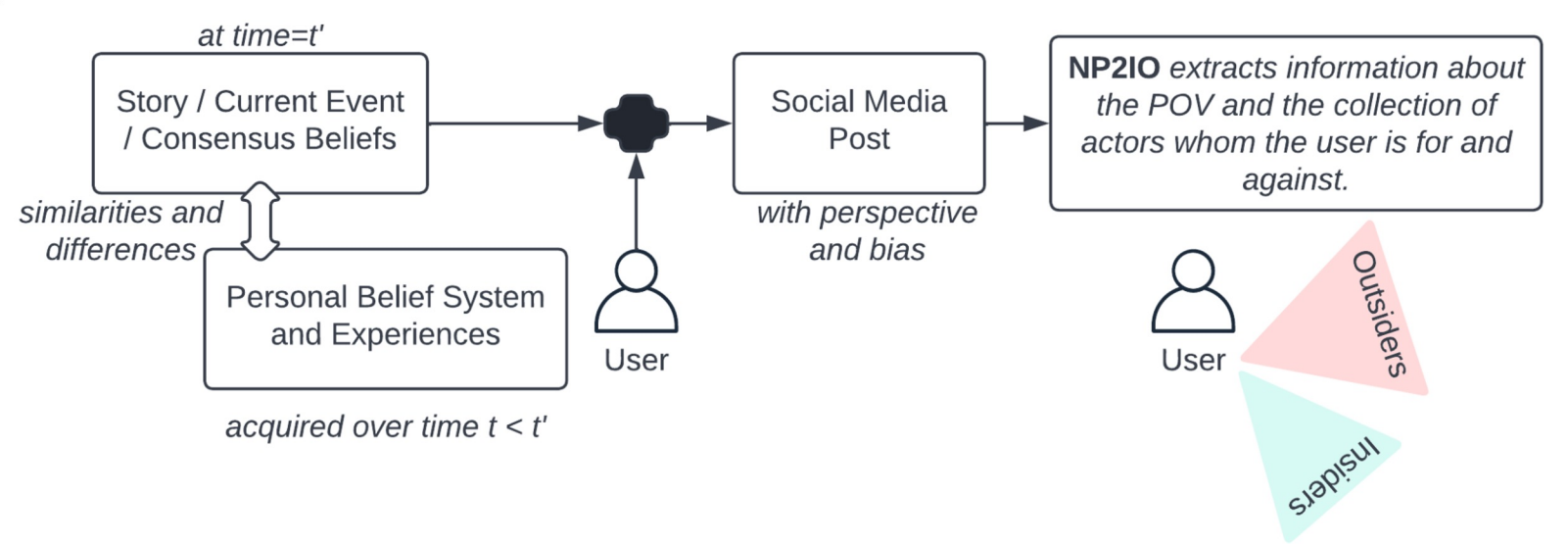
What does our framework do? A sample output.



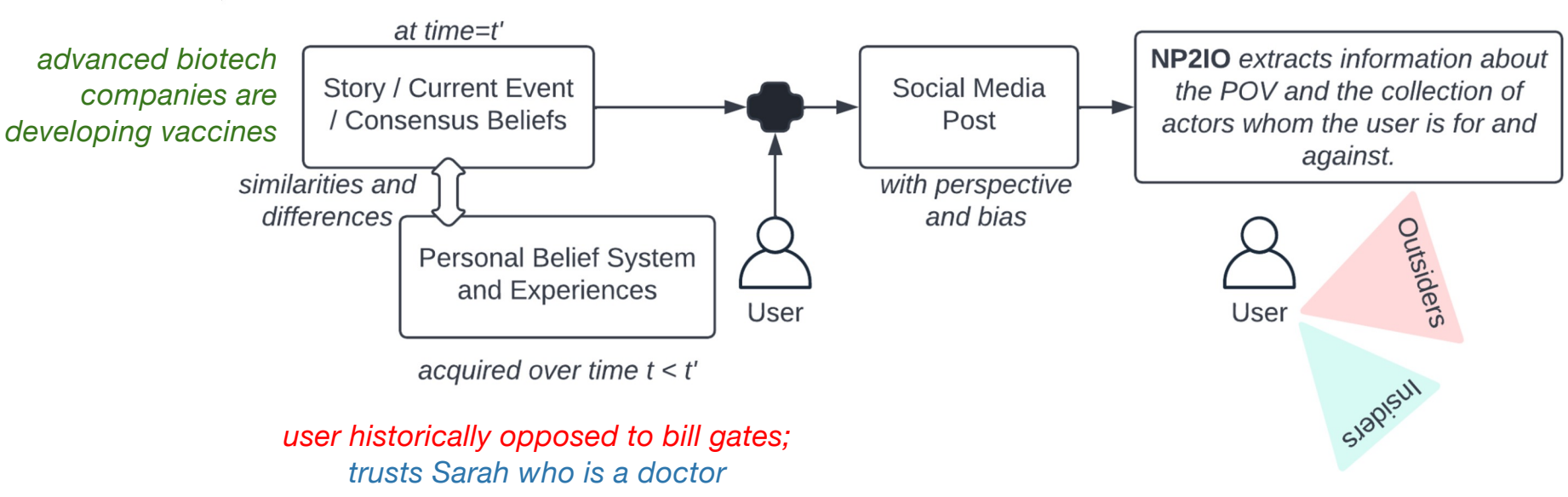
Defining Our Task - The Generative Model



Defining Our Task - Finding User Perspectives (NP2IO)

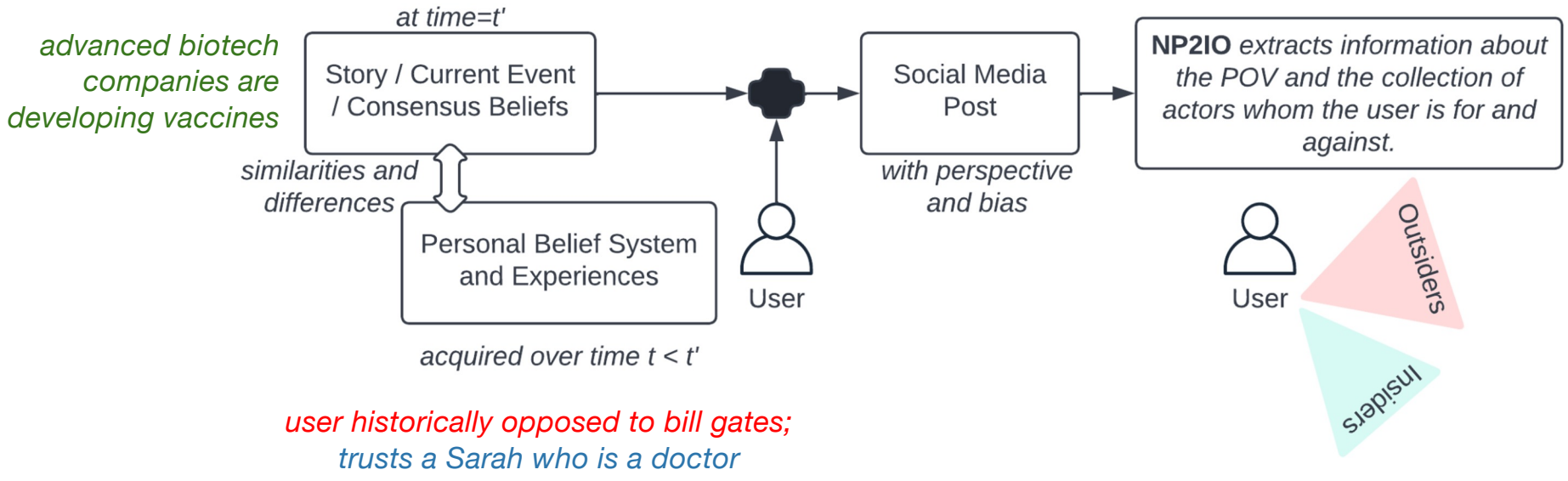


Defining Our Task - Use-Case



Defining Our Task - Use-Case

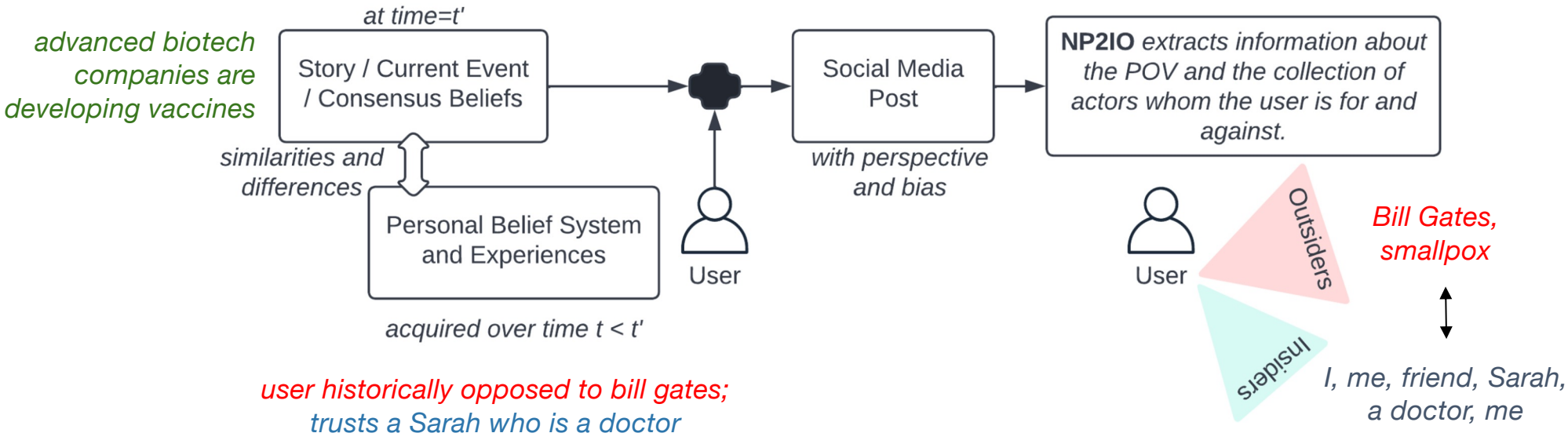
"I think that tech will kill me with vaccines. People like Bill Gates are developing this for the money. My friend, Sarah, who is a doctor, told me not to get the vaccine because it causes smallpox."



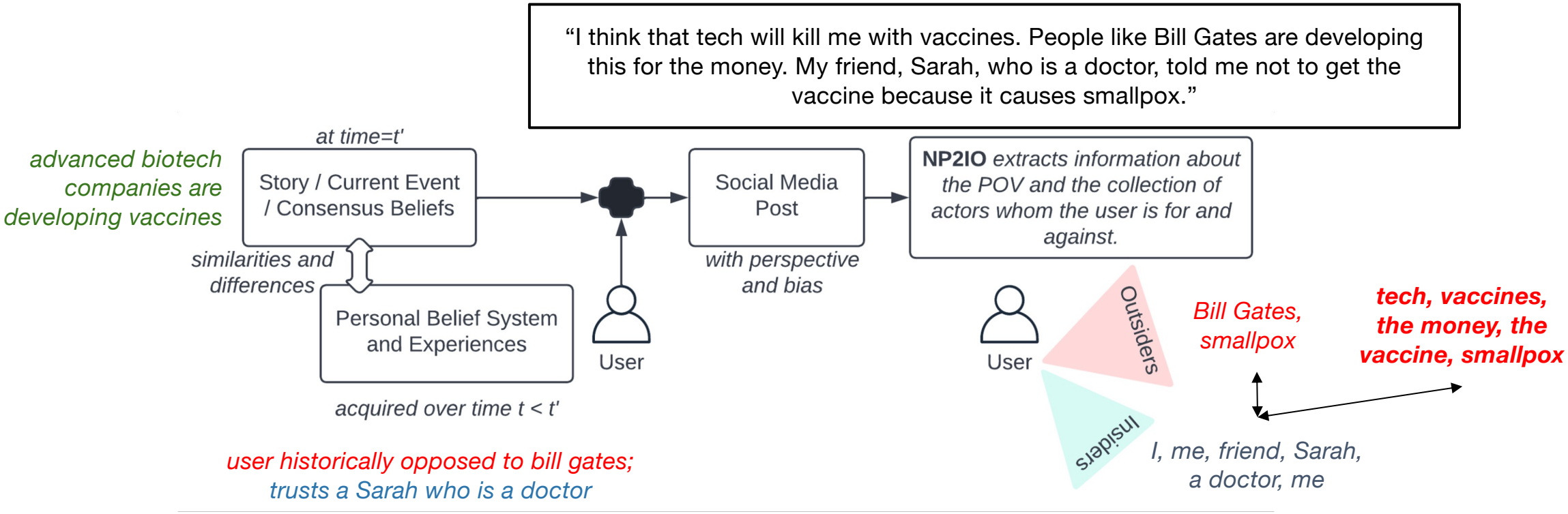
user historically opposed to bill gates;
trusts a Sarah who is a doctor

Defining Our Task - Use-Case

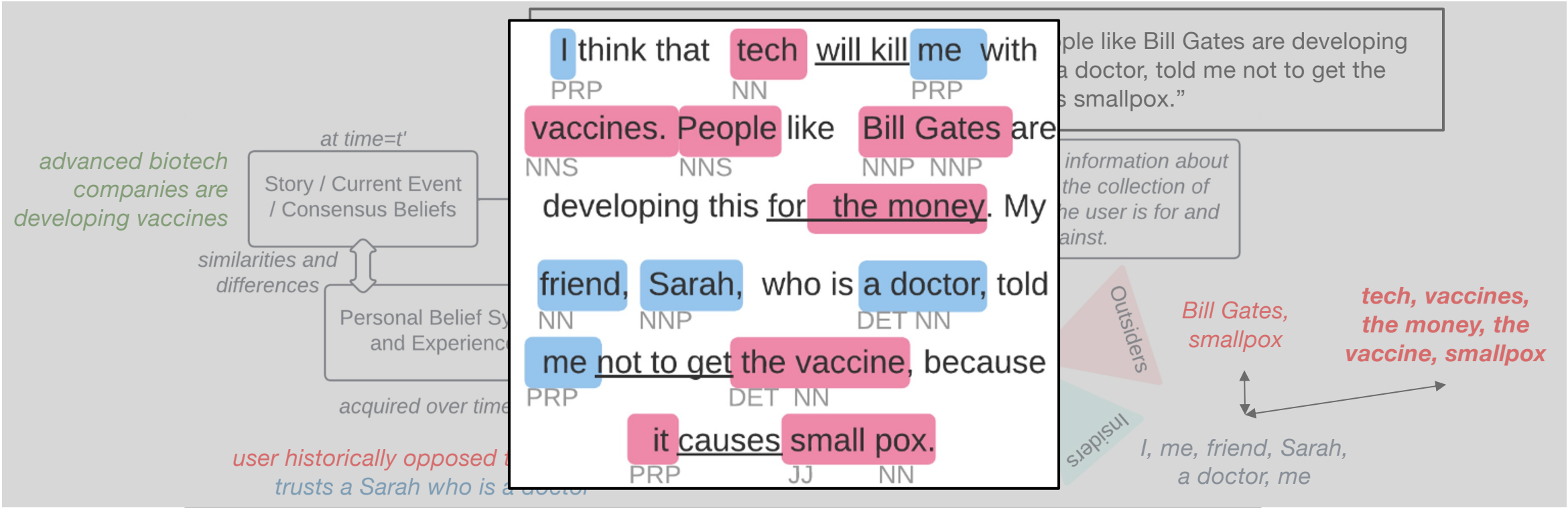
"I think that tech will kill me with vaccines. People like Bill Gates are developing this for the money. My friend, Sarah, who is a doctor, told me not to get the vaccine because it causes smallpox."



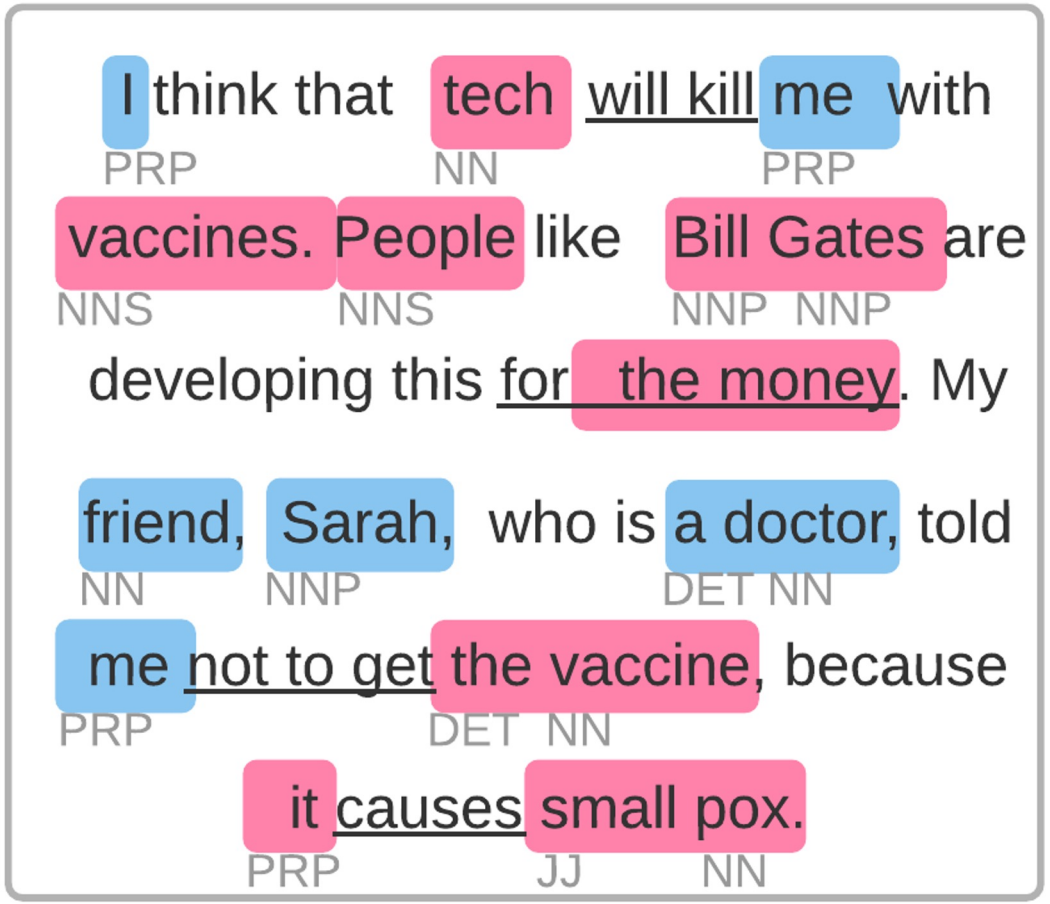
Defining Our Task - Use-Case



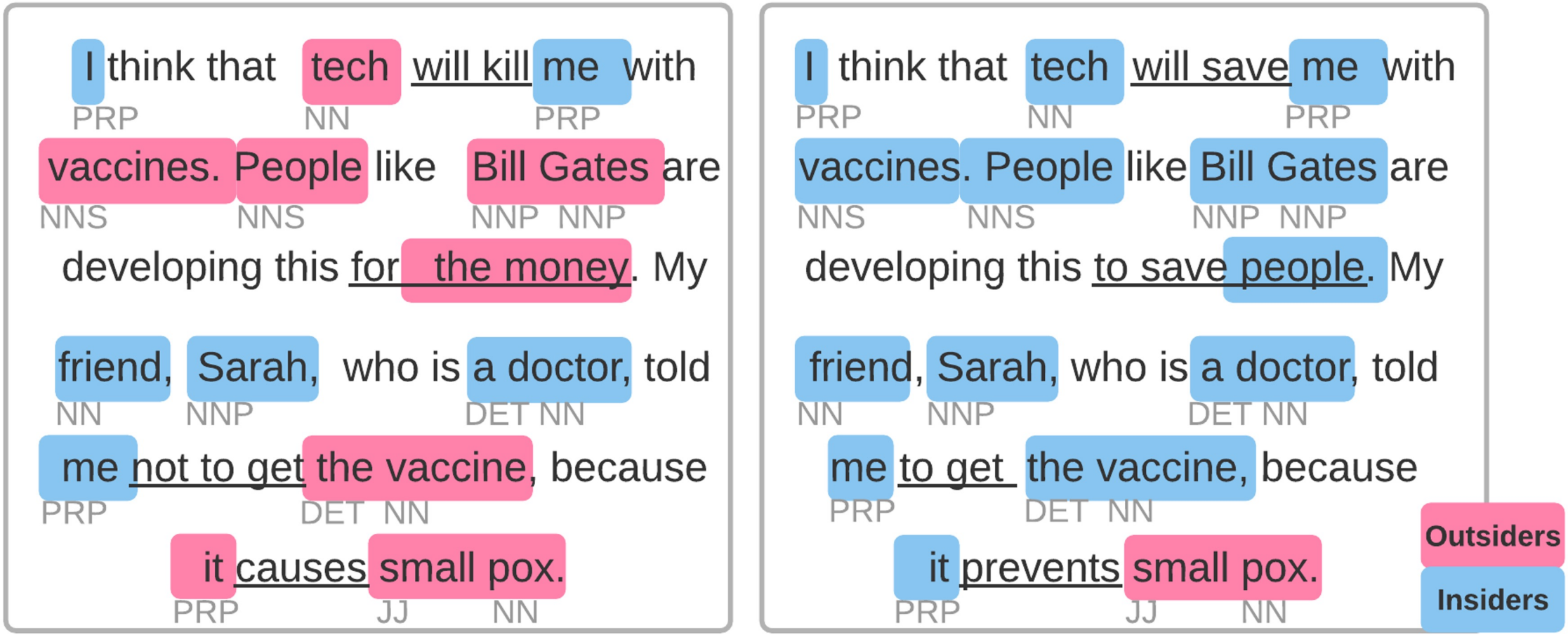
What does our model (NP2IO) do?



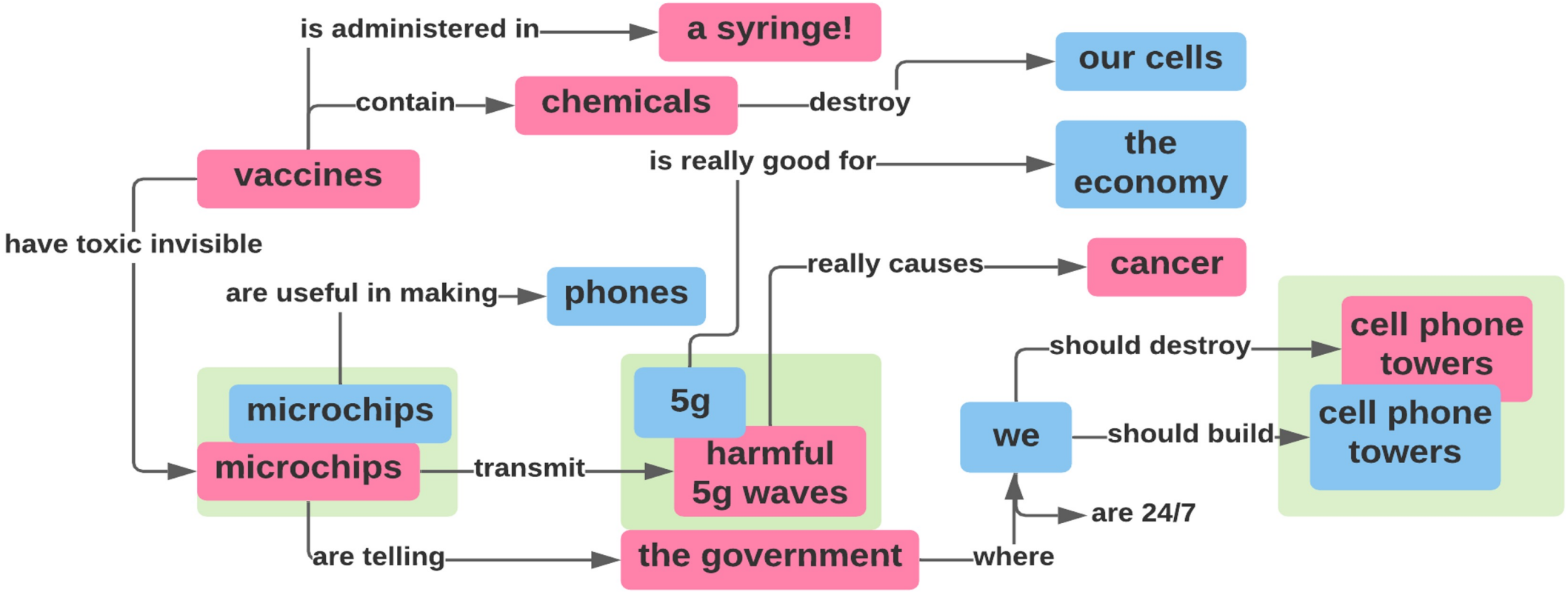
NP2IO is sensitive to context



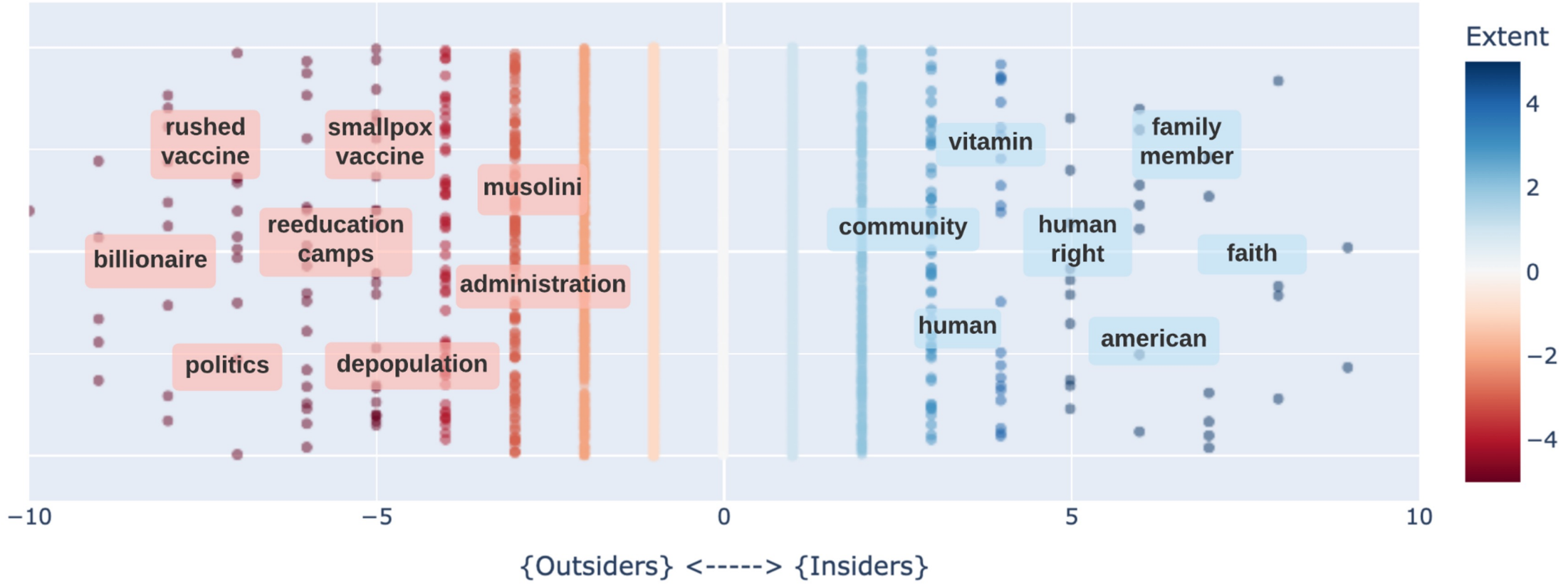
NP2IO is sensitive to context



Some noun phrases can take the role of either class depending on context.



Performance of majority voting on phrases not seen during training in any context (Zero-Shot)



Experiment #2

Modeling Hobbits

Memory and Forgetting on Goodreads

Holur, P., Shahsavari, S., Ebrahimzadeh, E., Tangherlini, T. R., & Roychowdhury, V. (2021). Modelling social readers: novel tools for addressing reception from online book reviews. *Royal Society open science*, *8*(12), 210797.

We applied our pipeline to book reviews!

- Why on earth would you do that???
- How do people remember at scale?
- Focused on corpus of 5 frequently reviewed books on Goodreads
  - Frankenstein
  - To Kill a Mockingbird
  - The Hobbit
  - Of Mice and Men
  - Animal Farm

Methods

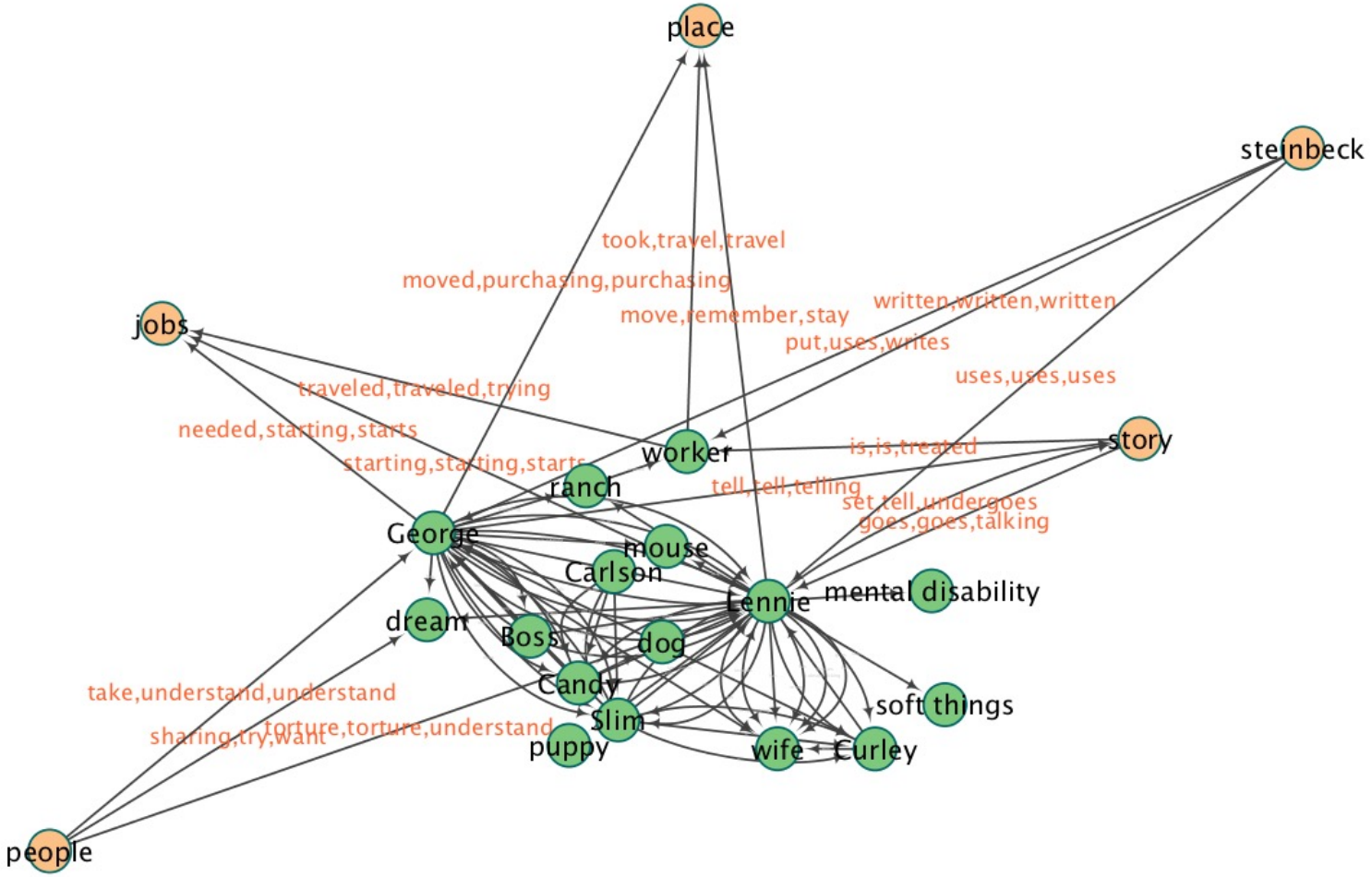
- Same methods as creating narrative frameworks for conspiracy theory work
  - Entity mention grouping (EMG)
  - Interactant relationship clustering (IARC)
- Two new parts of pipeline: REV2SEQ and Sent2Imp
    - We would like to capture consensus information to understand the **dynamics** of how readers imagine the sequence of events in a novel.
  - We also want to know what readers' impressions are of characters: Sent2Imp
  - Methods are detailed here: https://royalsocietypublishing.org/doi/full/**10.1098/rsos.210797**

Results

- (Expanded) story graphs that provide a consensus view of what readers remember of the target novel
- Sequenced graphs, revealing directed acyclic paths through the story network
- Representations of readers' impressions of story characters
  - Which can then be compared across novels

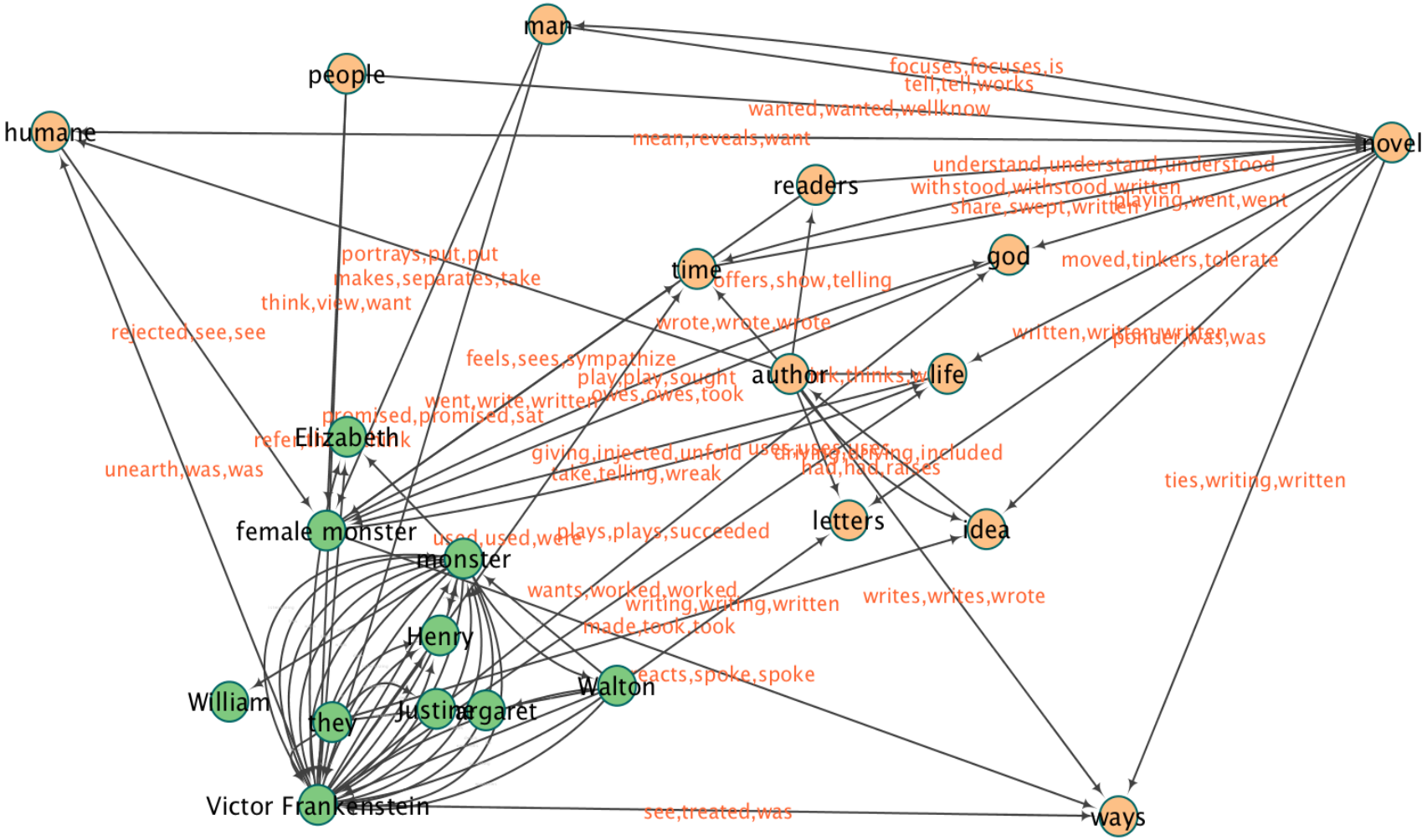
Expanded story graphs
Meta-text and interpretive communities

- *Of Mice and Men:*
- The node ``steinbeck'' has an in-degree of 0 suggesting readers' understanding of the author's impact on creating complex story characters, while the actants have no meaningful return engagement.
- Similarly, the ``place'' node cannot directly effect causal change in the story and as a result is very rarely found in the \textit{subject} part of a relationship.



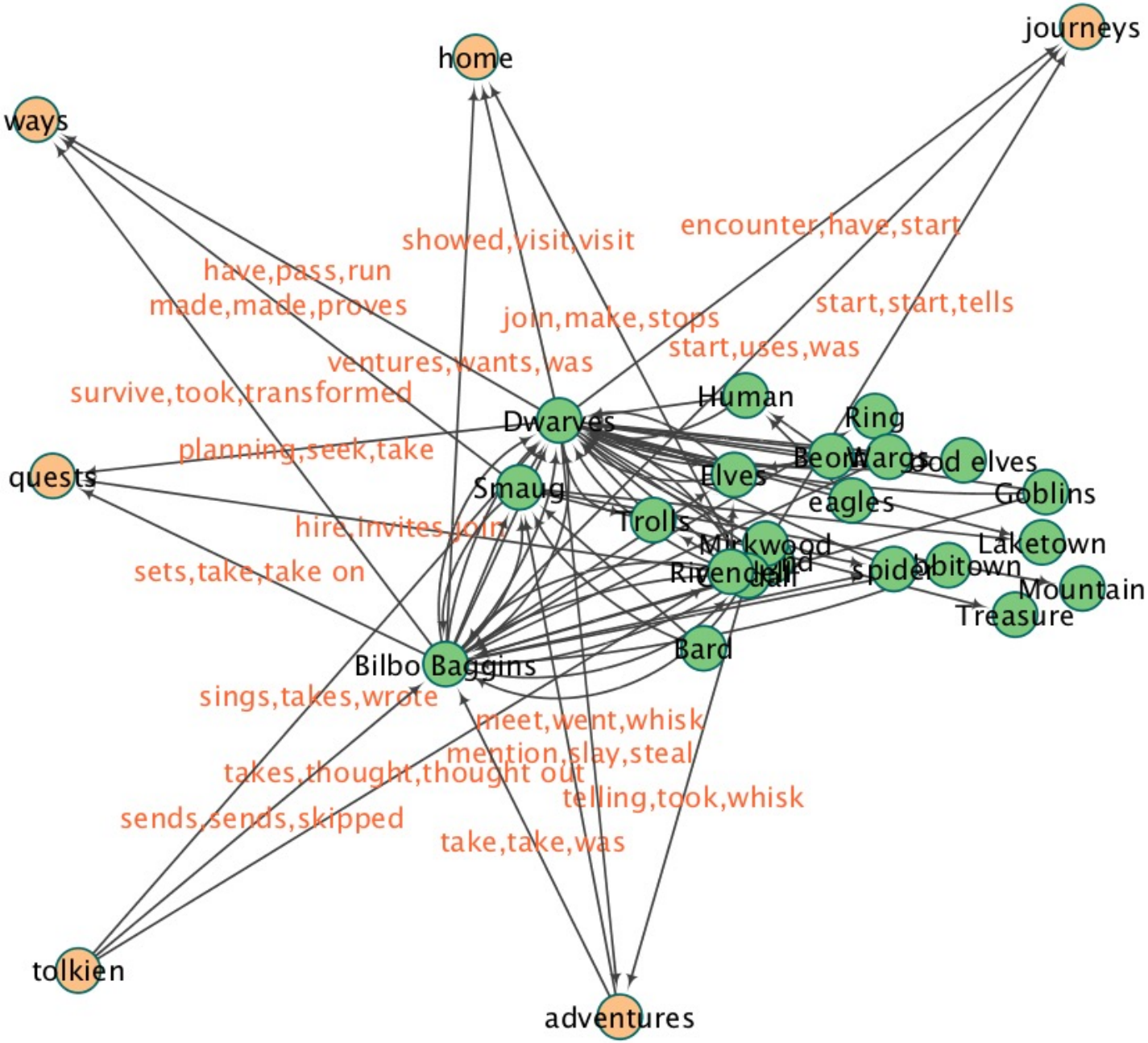
Frankenstein

- The subnetwork of “letters”, “author” and “novel” indicate that readers recognize the epistolary nature of Frankenstein.
- The common node “people” (which is found in most of the graphs) represents the reviewers’ perception of other reviewers.

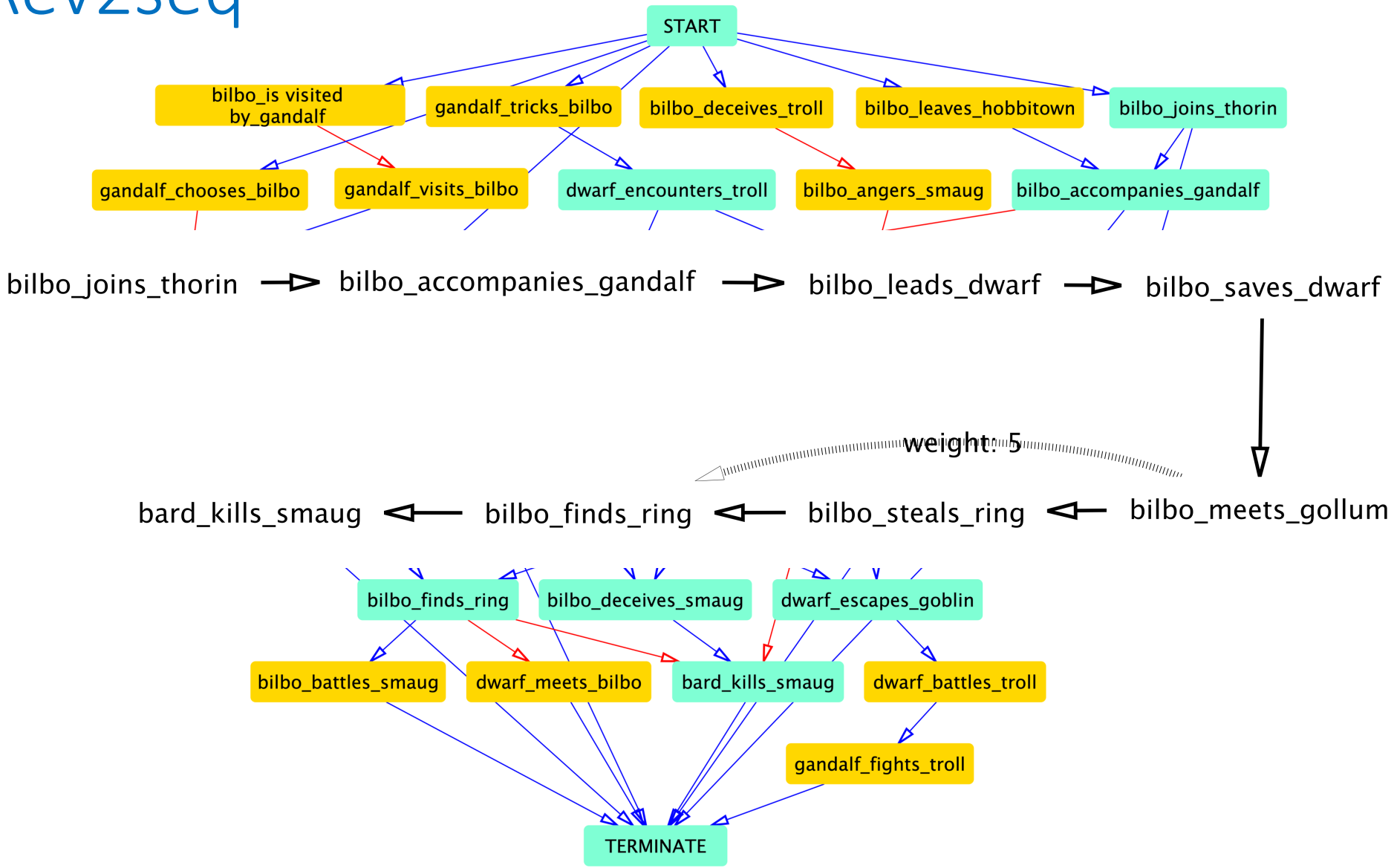


The Hobbit

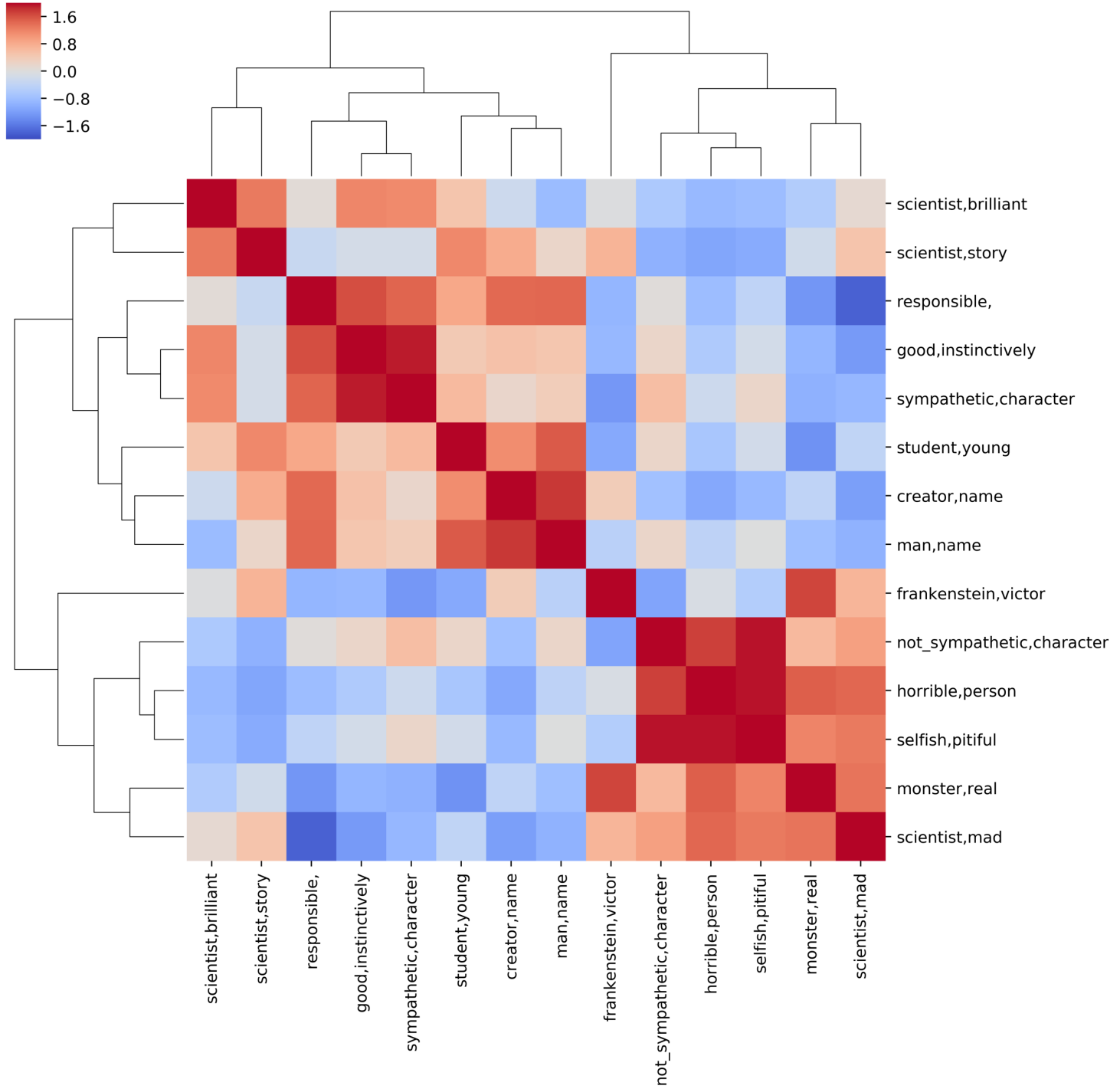
- The readers' classification of subgenre is immediately apparent in the nodes "adventures", "quest", "home" and "journeys".
- Inanimate actants such as "home", "journey", "quest" and "ways" typically have a very low out-degree (in this case 0)
- "tolkien" has a very low in-degree.
- The node "ways" signals strategy: "Dwarves" have "ways" or "Bilbo Baggins" took "ways".



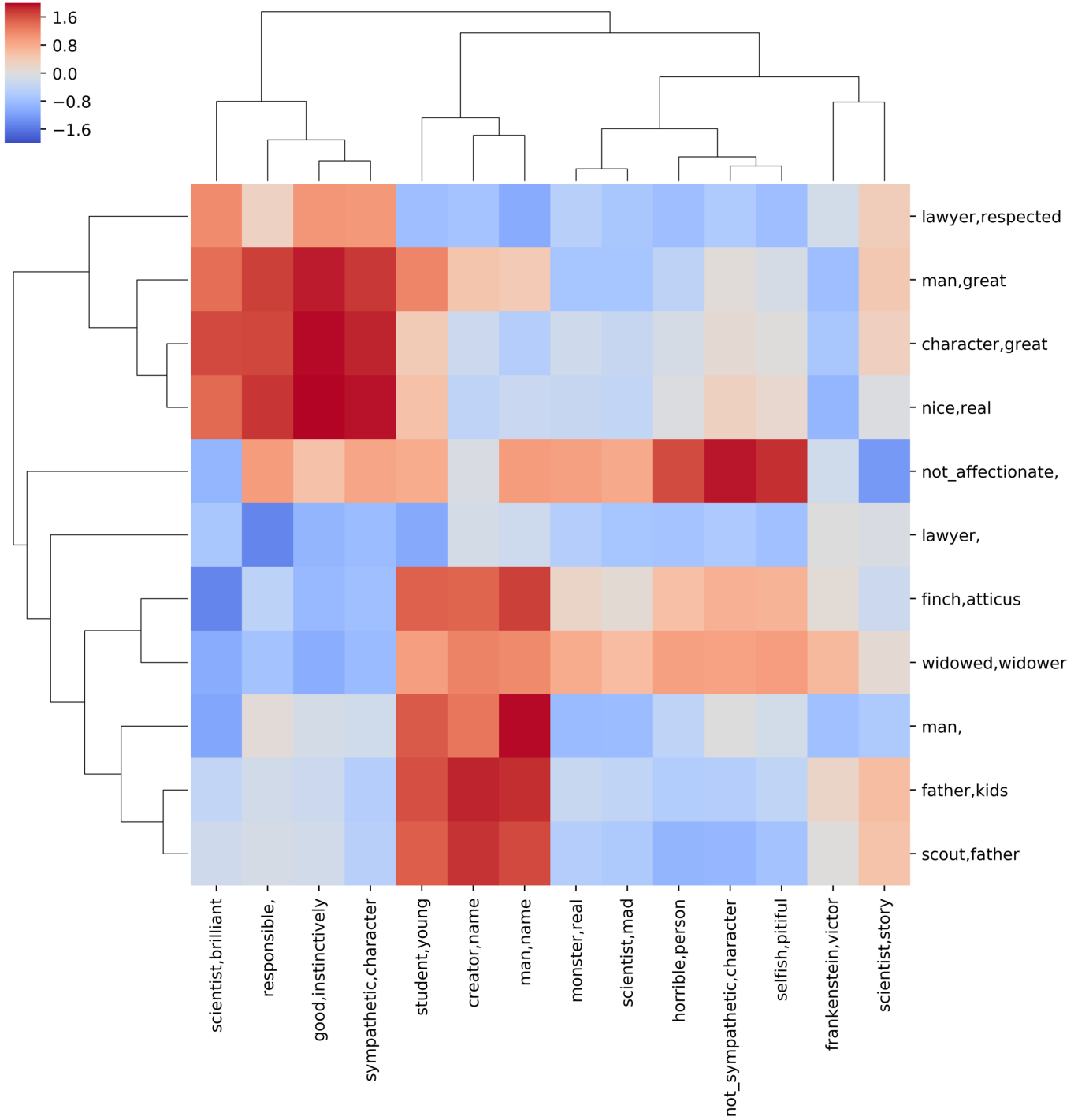
Rev2seq


START
bilbo_is visited by_gandalf
gandalf_tricks_bilbo
bilbo_deceives_troll
bilbo_leaves_hobbitown
bilbo_joins_thorin
gandalf_chooses_bilbo
gandalf_visits_bilbo
dwarf_encounters_troll
bilbo_angers_smaug
bilbo_accompanies_gandalf
bilbo_joins_thorin
bilbo_accompanies_gandalf
bilbo_leads_dwarf
bilbo_saves_dwarf
weight: 5
bard_kills_smaug
bilbo_finds_ring
bilbo_steals_ring
bilbo_meets_gollum
bilbo_finds_ring
bilbo_deceives_smaug
dwarf_escapes_goblin
bilbo_battles_smaug
dwarf_meets_bilbo
bard_kills_smaug
dwarf_battles_troll
gandalf_fights_troll
TERMINATE

Sent2Imp
Frankenstein-Frankenstein Impression Heatmap



Frankenstein-Atticus Impression Heatmap



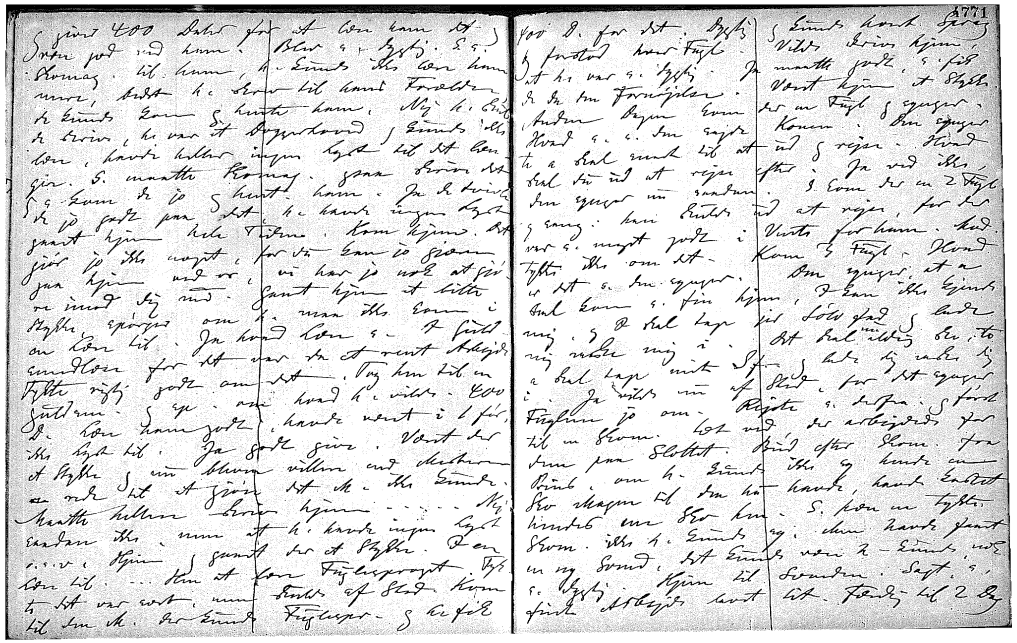
Experiment #3

Danish Fairy Tales at Tradition Scale

Embedding methods for understanding a folklore collection at very large scale

Abello, Broadwell, Tangherlini & Zhang. “Taming the Hairball”. *Fabula*. *forthcoming*

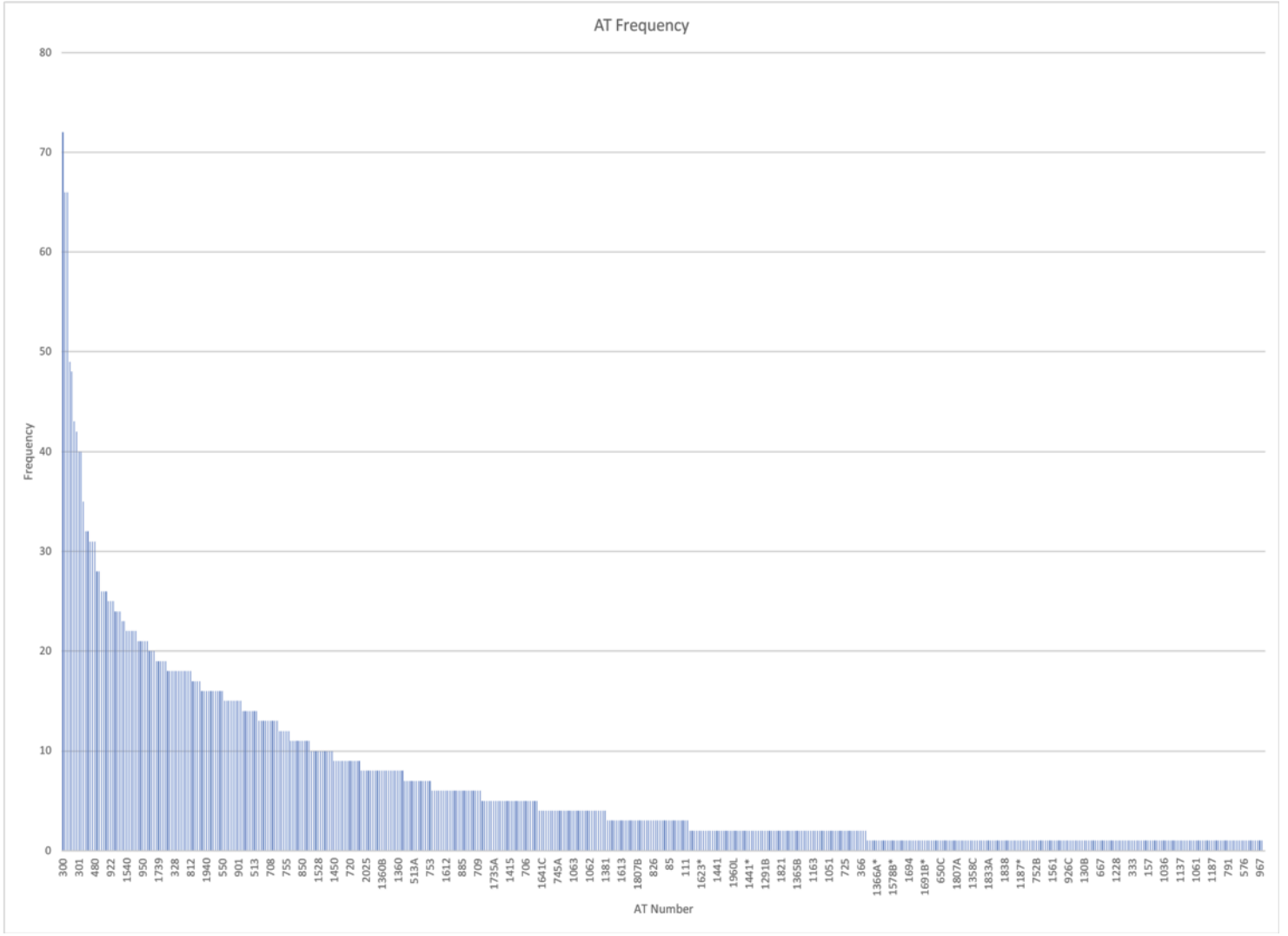
The Data & the Problem



- 2606 fairy tales
  - collected by Evald Tang Kristensen
  - 738 storytellers
  - 491 places
- Indexed according to the AT(U) index and the Motif Index
  - ATU index has 2875 types
  - MI has 46,500 motifs
- Many collections have sparse meta-data, and non-machine readable texts
- What can I learn about Danish fairy tale tradition from this sparse representation of a snapshot of Danish culture ~1870-1920?

Simple methods

AT Type	Count
300	72
400	66
326	66
852	49
935	48
302	43
303	42
301	40
313	40
1641	35
650A	32
330	32
1535	31
1539	31
480	31



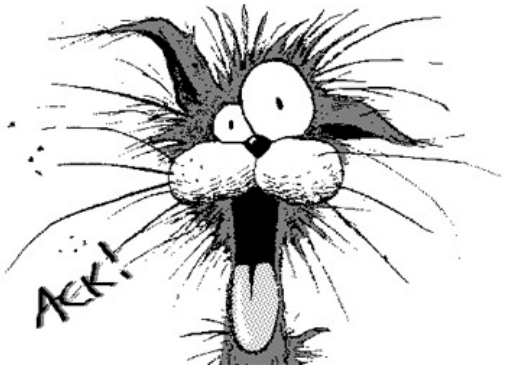
Problems, problems…

- Tells us only the frequency of tales
- Tells us nothing about the connection between tales
- Tells us nothing about the connection between storytellers and tales
- How can I take advantage of the latent information that exists in this collection, even in the absence of rich textual data?

Network methods?

- Make a network of the ATU index
- Make a network of the MI index
- Adjust them so that they are not only trees by drawing edges
  - between types that refer to each other and
  - between types and referenced motifs
- Add in the Danish stories
  - With links between stories and ATU number
  - Stories and storytellers
  - Storytellers and places of collection

Taming the hairball



- “Hairballs” are a common problem in Humanities network data / data visualizations
  - Standard methods attempt to reduce the hairball by clustering, calculating centrality measures, etc.
- Can I imagine a method of decomposing the graph so that I can discover the various “spines” holding the graph together?

Fixed Point of Degree Peeling!

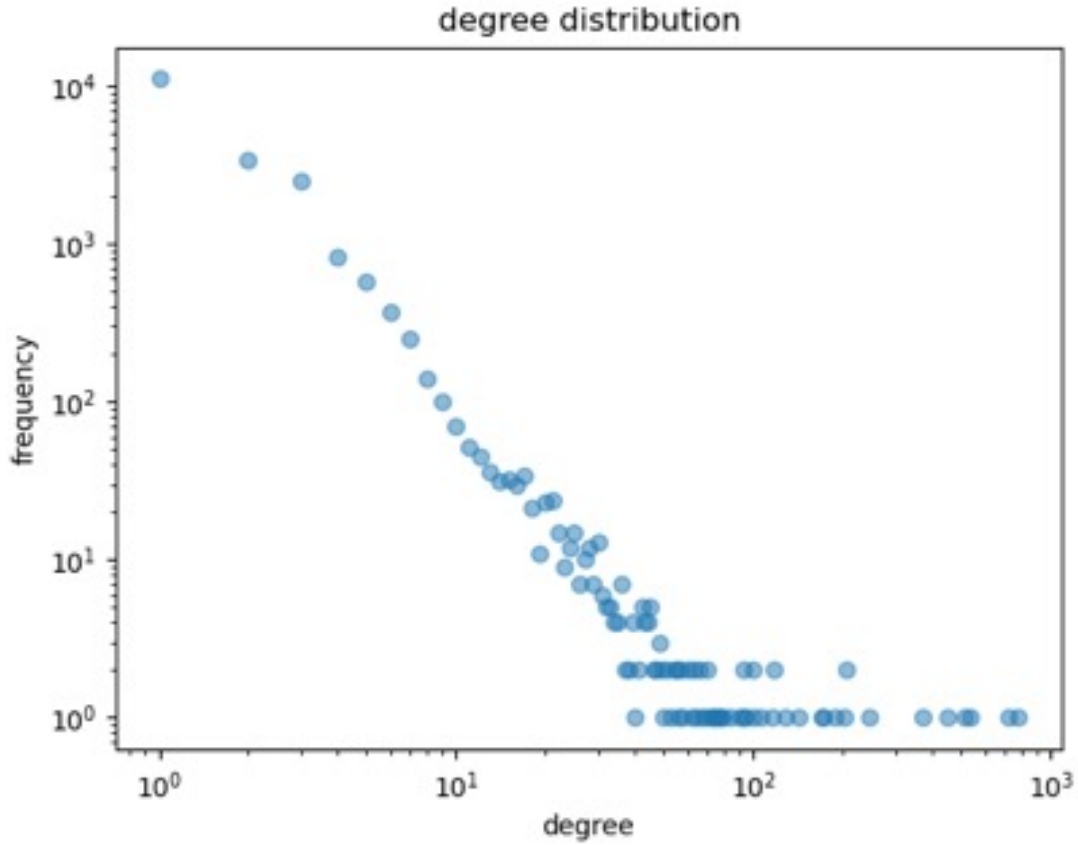
- Minimum Degree Peeling

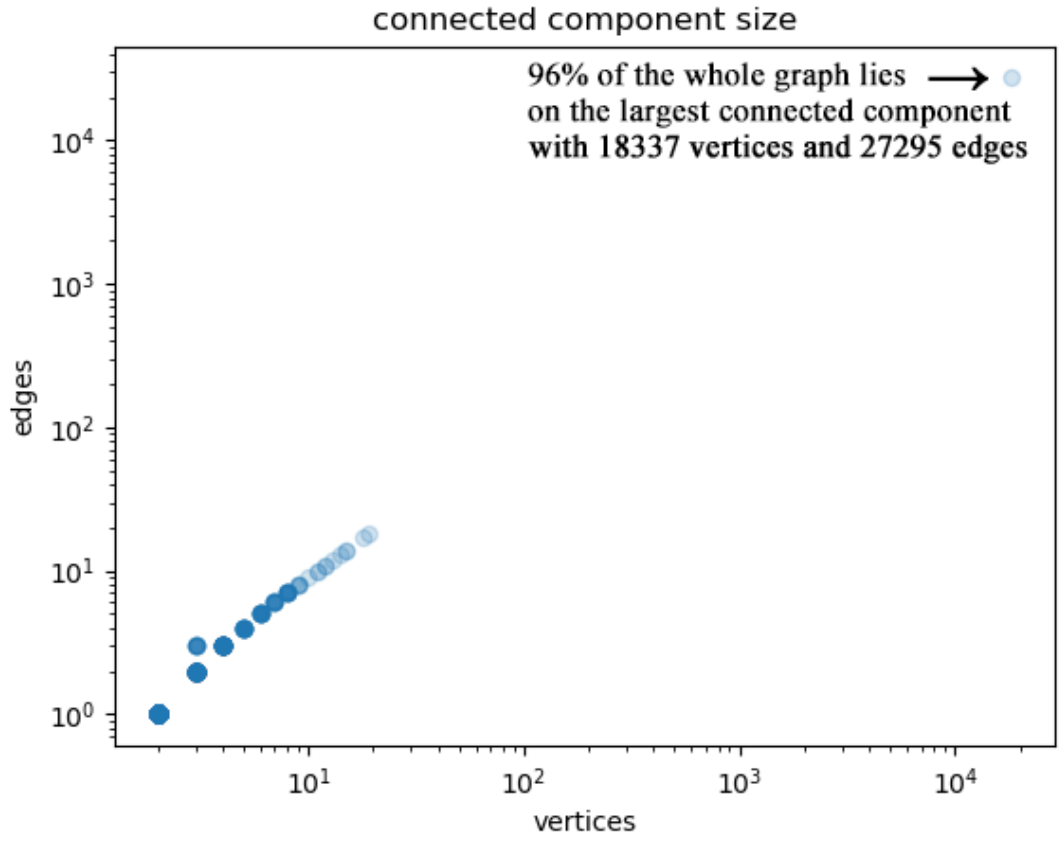
  Step 1. Set the current_peel_value = minimum degree;

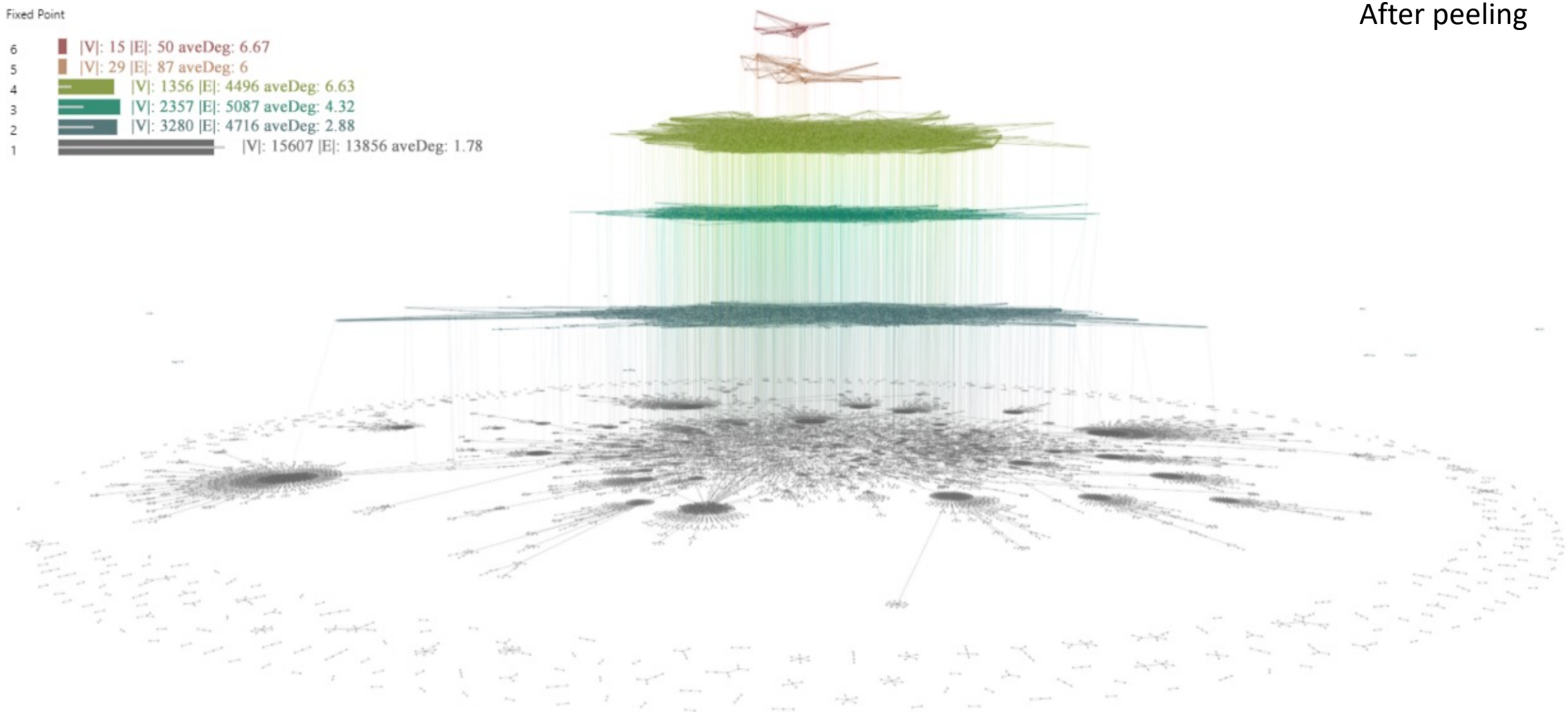
  Step 2. Iteratively remove all vertices of degree less than or equal to the current_peel_value updating the degrees of all neighbors of deleted vertices. Assign to each removed vertex, a peel value equal to the current_peel_value.

  Step 3. If the minimum degree is greater than the current_peel_value then output all the deleted vertices as vertices of peel value = current_peel_value and apply the same process of vertex peeling to the remaining graph with the next higher minimum degree.

- The core of G, denoted core(G), sometimes also called the k-core of G, is the subgraph induced by the maximal subset of vertices of G whose peel value is maximum


degree distribution
frequency
degree
10^0
10^1
10^2
10^3
10^4


connected component size
96% of the whole graph lies
on the largest connected component
with 18337 vertices and 27295 edges
edges
vertices
10^0
10^1
10^2
10^3
10^4



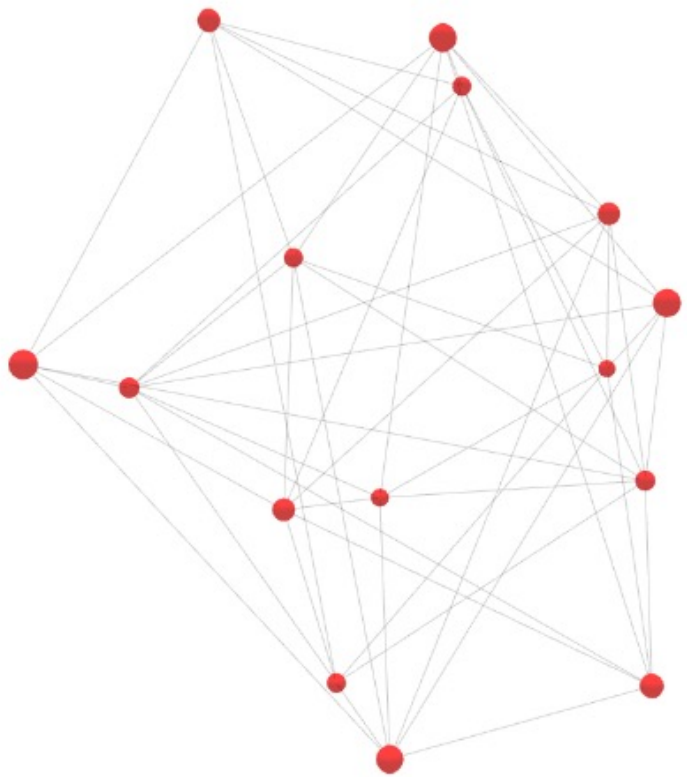
After peeling

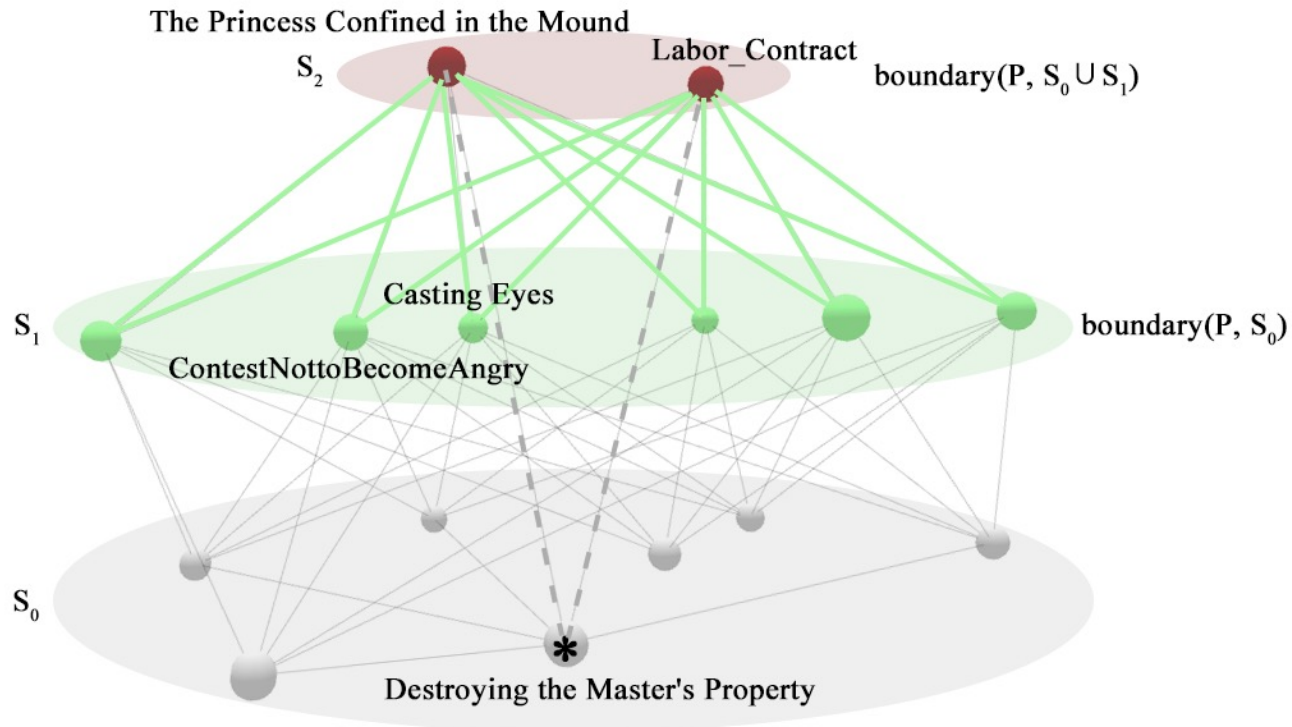
Decompose further?
Graph Fragments!

Any non-regular fixed point can be assembled as a sequence of atomic subgraphs *Graph Fragments*.

To define graph fragments, we need to introduce the notion of the boundary of a vertex set.

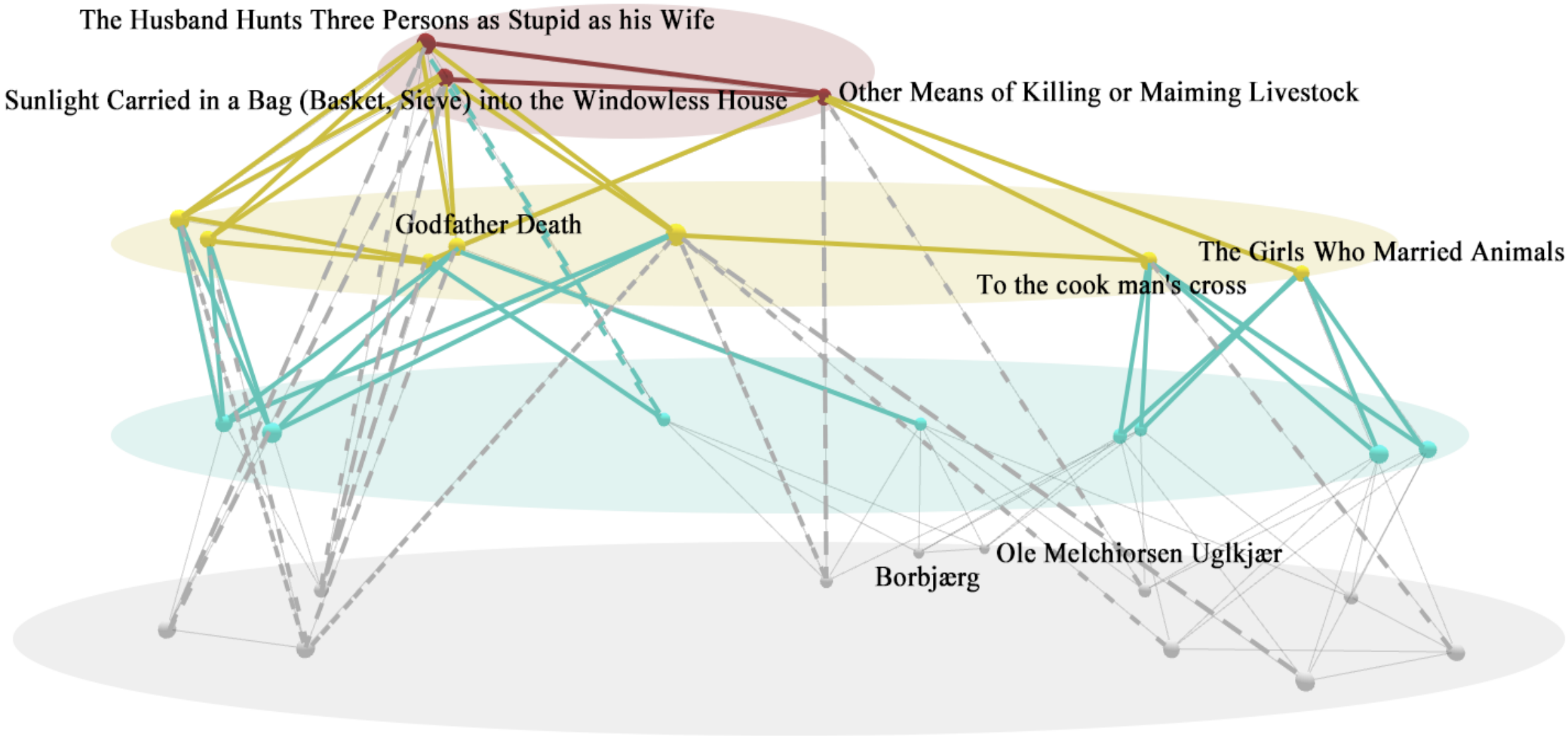
- i. **Boundary Vertices**. A vertex v is said to be in the boundary of a set of vertices S if v is not in S but it contains at least one neighbor in S. The set of vertices in the boundary of S is denoted by $\mathrm{boundary}(S)$.
- ii. **Graph Edge Fragments** associated with a set of vertices S. The collection of edges with at least one endpoint u or v in S is called the graph edge fragment associated with S and is denoted by $\mathrm{frag}(S)$.




The Princess Confined in the Mound
Labor_Contract
S_2
boundary(P, $S_0 \cup S_1$)
Casting Eyes
S_1
boundary(P, S_0)
ContestNottoBecomeAngry
S_0
*
Destroying the Master's Property

Some interesting findings

- The most “prominent” ATU vertices are
  - ATU 1000-1029: Labor Contract
  - ATU 870: The Princess Confined in the Mound.
- It is certainly interesting to note that
  - ATU 1006: Casting Eyes and ATU 1000: Contest Not to Become Angry have exactly the same neighborhood,
  - While the neighborhood of the star vertex ATU 1002: *Destroying the Master's Property (Formerly Dissipation of the Ogre's Property)*) is a subset of theirs.
- The top two vertices (ATU 1000-1029: *Labor Contract* and ATU 870: *The Princess Confined in the Mound*) also have identical neighborhoods.
- These are *all* tales about deception and manipulation of labor relationships


The Husband Hunts Three Persons as Stupid as his Wife
Sunlight Carried in a Bag (Basket, Sieve) into the Windowless House
Other Means of Killing or Maiming Livestock
Godfather Death
The Girls Who Married Animals
To the cook man's cross
Ole Melchiorsen Uglkjær
Borbjærg

Some unexpected discoveries!

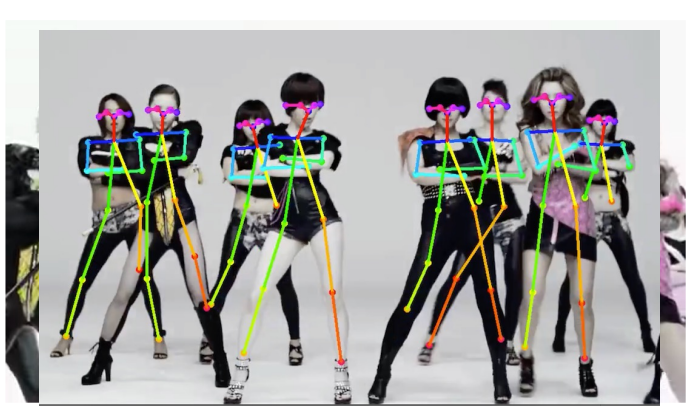
- At the very core of the Danish folktale tradition, stories about deceit, trickery and labor form the main spine of the tradition.
- Ole Melchiorsen stands out as an important storyteller
  - Mentioned only in passing in Holbek (1987)
  - But has a repertoire that he learned from
    - Many different people, both men and women
    - And from many different parts of Denmark
  - Although he tells some of the high-frequency tales
  - He tells more tales about labor, deceit and violence

Experiment #4

K-Pop Dance

Toward a pose-based search engine

Can we discover similarities automatically?



Brown Eyed Girls, “Abracadabra” (2009)

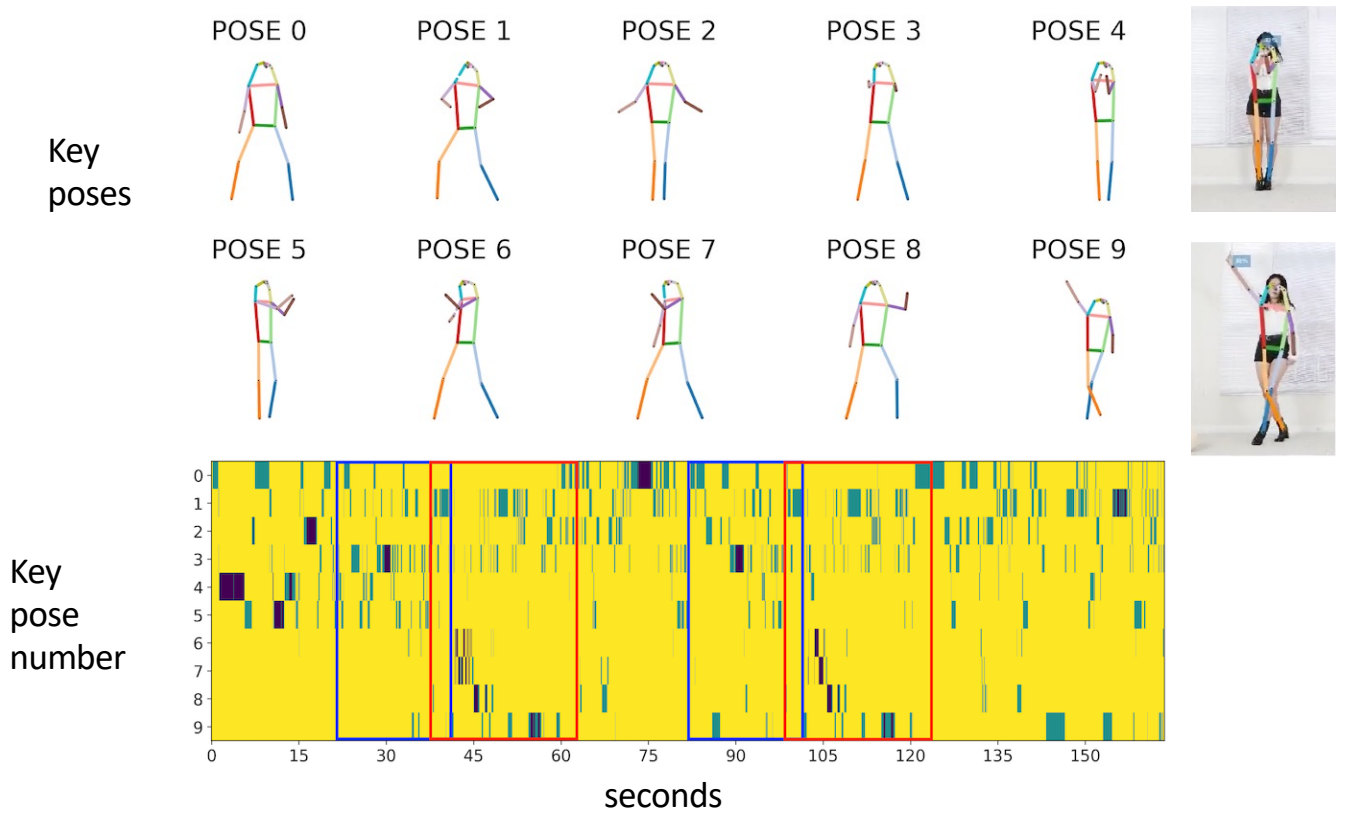
Psy (and Ga-In), “Gentleman” (2013)

- So that we can track patterns across thousands of K-pop dance sequences?
- Why?
  - Devise a search method based on the dance itself: body poses, and sequences of poses
  - A data-driven analysis of the similarities and differences in K-pop dance across performances and across time

Our challenges

- Identify dancers
  - And not lose track of them (or their body parts!)
- Understand their poses
  - Consider sequence of poses as a dance move
    - First do it for each individual dancer
    - Then do it for all of the dancers in a sequence of frames
  - Consider sequences of dance moves as a dance sequence
    - For single dancers
    - Across multiple dancers

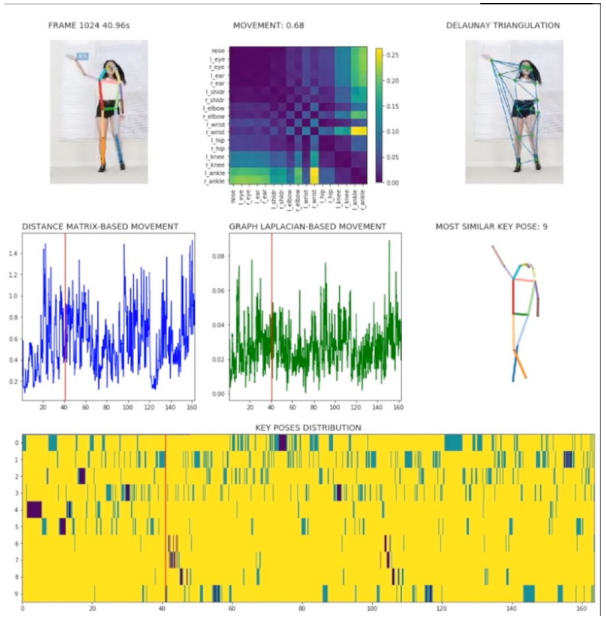
Pose clustering to find “key poses”

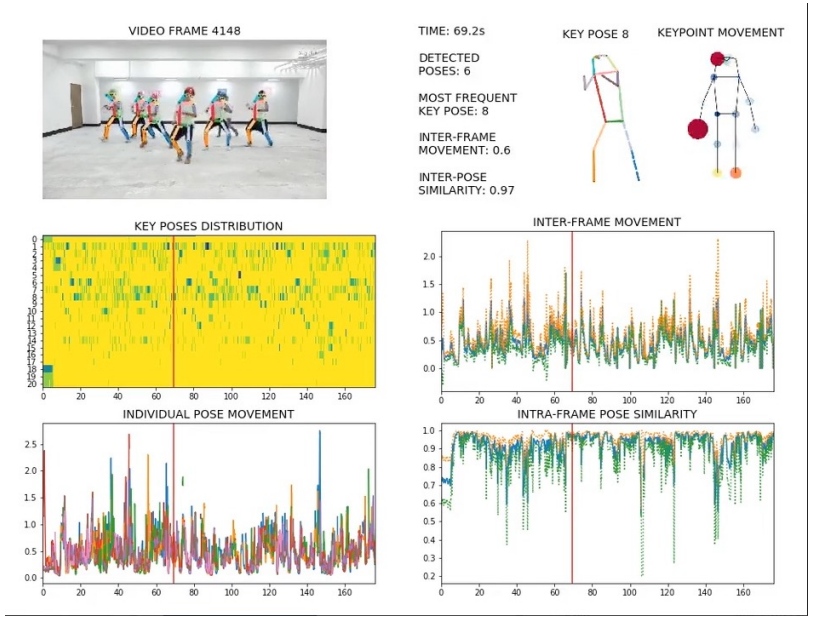


Putting it all together

- Jenny “Solo” – official video and our experimental interface

- BTS “Fire” – official video and our experimental interface





Why do all this?

The Challenges of Culture Analytics

Culture Analytics provides novel insights

- Even this brief exploration of some projects has shown that Cultural Analytics results are:
  - Novel
  - Reliant on a match between mathematical models and the underlying data
  - **Robust**
  - Not attainable by inspection (ie. **Not obvious**)

Future Directions

- AI, Deeplearning and the integration of System 1 and System 2 thinking
  - Joining neural and symbolic AI to solve complex problems in the dynamics of culture
  - “We probably won’t be able to achieve everything that we want in AI just with deep learning” (Dan Gutfreund, MIT-IBM Watson)
- There is an imperative to reach across the chasm that often separates the qualitative cultural fields and computation
  - Our problems are interesting
  - But we often cannot solve the problems ourselves
    - fundamental training in data and data science
  - Recognize that collaboration is a key to successful projects that do not fall prey to either
    - Restating the obvious
    - Or applying methods that do not match the data

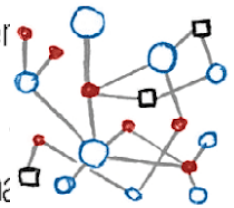
Future Projects / Challenges

- Climate Change
  - How is the debate shaped?
    - Fake news
    - Polarizing debates
    - Feedback between news and social media / comment threads
    - Do virtual communities influence real life decisions?
    - Local / national – global
- Global health
  - Not our last pandemic
  - Lessons learned?

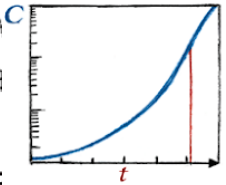
Arrowhead problems in Culture Analytics

The Arrowhead Problems in Culture Analytics

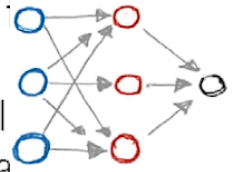
What are the basic objects of study and concepts subject to formalization?
Stable evidence-based mathematical formalization would allow for increased reproducibility and comparison across methods, datasets, and studies. Alternative formalizations may provide fundamentally different, but compler



What are the essential measures
Consistent, validated, and useful ma mental objects of study and concep mon measurements would also allo and biases in and across cultural da



What constitutes a successful ma
A mathematics of culture would be nature of culture. Identifying flexible foster the development of analytical quantitative and predictive cultural a



What are the fundamental structures of cultural interdependence?
A comprehensive understanding of cultural systems would make sense of cultural units and their subunits, groupings, connection types, and topologies, at multiple scales in space, time, and conceptual dimensions. It would further take into account systematic interdependence and overlap between subsystems of various types.

What are the fundamental dynamics of cultural change?
A deeper understanding of cultural dynamics and change would include appropriately sophisticated mathematical models of complex phenomena including cultural emergence, growth, percolation, diffusion, spreading, and evolution.

What are the algorithms that can detect the structures and dynamics of culture?
Scalable algorithms detecting structures and transitions in heterogeneous data would address low coverage in sparse data from historical sources or resource-poor areas, as well as massive real-time data streams as acquired through sensors, mobile platforms, or the Internet.

What are the ethical challenges in culture analytics?
Culture analytics is not without its risks. Algorithms, tools and methodologies developed for the data-driven analysis of culture must be carefully vetted and tested so that the results of this work are not only methodologically sound but also ethically sound. Further ethical challenges may arise from dataset collection and preservation bias that need to be carefully characterized.


THANK YOU
The Guardian

References

- Tangherlini, T. R., Roychowdhury, V., Glenn, B., Crespi, C. M., Bandari, R., Wadia, A., ... & Bastani, R. (2016). “Mommy blogs” and the vaccination exemption narrative: results from a machine-learning approach for story aggregation on parenting social media sites. JMIR public health and surveillance, 2(2), e6586.
- Tangherlini, T. R., Shahsavari, S., Shahbazi, B., Ebrahimzadeh, E., & Roychowdhury, V. (2020). An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, Pizzagate and storytelling on the web. PloS one, 15(6), e0233879.
- Shahsavari, S., Holur, P., Wang, T., Tangherlini, T. R., & Roychowdhury, V. (2020). Conspiracy in the time of corona: Automatic detection of emerging COVID-19 conspiracy theories in social media and the news. Journal of computational social science, 3(2), 279-317.
- Chong, D., Lee, E., Fan, M., Holur, P., Shahsavari, S., Tangherlini, T., & Roychowdhury, V. (2021, December). A real-time platform for contextualized conspiracy theory analysis. In 2021 International Conference on Data Mining Workshops (ICDMW) (pp. 118-127). IEEE.
- Holur, P., Wang, T., Shahsavari, S., Tangherlini, T., & Roychowdhury, V. (2022, May). Which side are you on? Insider-Outsider classification in conspiracy-theoretic social media. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 4975-4987).
- Tangherlini, T. R., Roychowdhury, V., & Broadwell, P. M. (2020). Bridges, Sex Slaves, Tweets, and Guns. Folklore and Social Media, 39.