semantic graphs and social networks

Camille Roth

Centre Zentrum Marc Bloch
Computational Social Science Team
cmb.huma-num.fr

articulating structure and culture

from cultural sociology to computational social science

articulating structure and culture

cultural sociology

articulating structure and culture

cultural sociology

critique of
a strict social structuralism

- SNA: categories and social order derive from the interactional structure
  vice versa for anthropologists

- call in the 1990s to abandon the primacy of structure over culture
  (Emirbayer & Goodwin 1994)

- « the structural presence –or absence– of ties may have cultural explanations as well »
  (Pachucki & Breiger 2010)

articulating structure and culture

cultural sociology

critique of a strict social structuralism

- SNA: categories and social order derive from the interactional structure
  vice versa for anthropologists
- call in the 1990s to abandon the primacy of structure over culture
  (Emirbayer & Goodwin 1994)
- « the structural presence –or absence– of ties may have cultural explanations as well »
  (Pachucki & Breiger 2010)

dependence of networks on meaning

- « networks of meaning » (White 1992)
  each identity, interaction context configures a different network
  - « multi-purpose ties »
    (Fine & Kleinman, 1983)
- « netdoms »
  social *net*works and semiotic *dom*ains
  (Godart & White 2010)
- social networks and cultural structure seen, even, as secondary and joint manifestations of a « deeper set of forces »
  (Breiger 2010)

articulating structure and culture

cultural sociology

critique of a strict social structuralism

- SNA: categories and social order derive from the interactional structure
  vice versa for anthropologists
- call in the 1990s to abandon the primacy of structure over culture
  (Emirbayer & Goodwin 1994)
- « the structural presence –or absence– of ties may have cultural explanations as well »
  (Pachucki & Breiger 2010)

dependence of networks on meaning

- « networks of meaning » (White 1992)
  each identity, interaction context configures a different network
  - « multi-purpose ties »
    (Fine & Kleinman, 1983)
- « netdoms »
  social *net*works and semiotic *dom*ains
  (Godart & White 2010)
- social networks and cultural structure seen, even, as secondary and joint manifestations of a « deeper set of forces »
  (Breiger 2010)

need to introduce culture directly into SNA

- long tradition in SNA
  « membership networks »
  (Breiger 1974)
  - bipartite connection between social and semantic entities
- various operationalizations
  Carley 1986
  Leenders 1997, Snijders et al 2007
  Lazega & Pattison 1999
  Wang et al 2003
  Basov & Brennecke 2017
- yet still a relevant call
  (Ferguson et al. 2017)

articulating structure and culture

computational social science

characterizing social network dynamics

- from mathematical sociology on rather small case studies
- to complex system science aiming at finding universal laws of networks

articulating structure and culture

computational social science

characterizing social network dynamics

- from mathematical sociology on rather small case studies
- to complex system science aiming at finding universal laws of networks
- now a rich, specialized literature on social network structure and dynamics in specific application domains

articulating structure and culture

computational social science

characterizing social network dynamics

- from mathematical sociology on rather small case studies
- to complex system science aiming at finding universal laws of networks
- now a rich, specialized literature on social network structure and dynamics in specific application domains

characterizing semantic entities

- scientometrics as a precursor
  e.g., co-occurrence clusters, subfield maps

articulating structure and culture

computational social science

characterizing social network dynamics

- from mathematical sociology on rather small case studies
- to complex system science aiming at finding universal laws of networks
- now a rich, specialized literature on social network structure and dynamics in specific application domains

characterizing semantic entities

- scientometrics as a precursor
  e.g., co-occurrence clusters, subfield maps
- now, a myriad of denominations / approaches
  "text mining"
  "automated text analysis"
  "text-as-data methods"

articulating structure and culture

computational social science

characterizing social network dynamics

- from mathematical sociology on rather small case studies
- to complex system science aiming at finding universal laws of networks
- now a rich, specialized literature on social network structure and dynamics in specific application domains

characterizing semantic entities

- scientometrics as a precursor
  e.g., co-occurrence clusters, subfield maps
- now, a myriad of denominations / approaches
  "text mining"
  "automated text analysis"
  "text-as-data methods"
- and a myriad of notions and objects:

- scalar numbers, vectors
- fixed categories / ontologies (coding)
- named entity recognition

- word and n-gram bags, distributions
- topic modeling, LDA

- word, sentence, document embeddings

- sentence and pattern matching and clusters
- dependency parsing for syntactic roles
- claim and argument extraction

articulating structure and culture

computational social science

characterizing social network dynamics

- from mathematical sociology on rather small case studies
- to complex system science aiming at finding universal laws of networks
- now a rich, specialized literature on social network structure and dynamics in specific application domains

characterizing semantic entities

- scientometrics as a precursor
  e.g., co-occurrence clusters, subfield maps
- now, a myriad of denominations / approaches
  "text mining"
  "automated text analysis"
  "text-as-data methods"
- and a myriad of notions and objects:

- scalar numbers, vectors
- fixed categories / ontologies (coding)
- named entity recognition

- word and n-gram bags, distributions
- topic modeling, LDA

- word, sentence, document embeddings

- sentence and pattern matching and clusters
- dependency parsing for syntactic roles
- claim and argument extraction

in recent years: increasingly intertwining both

semantic graphs in social science

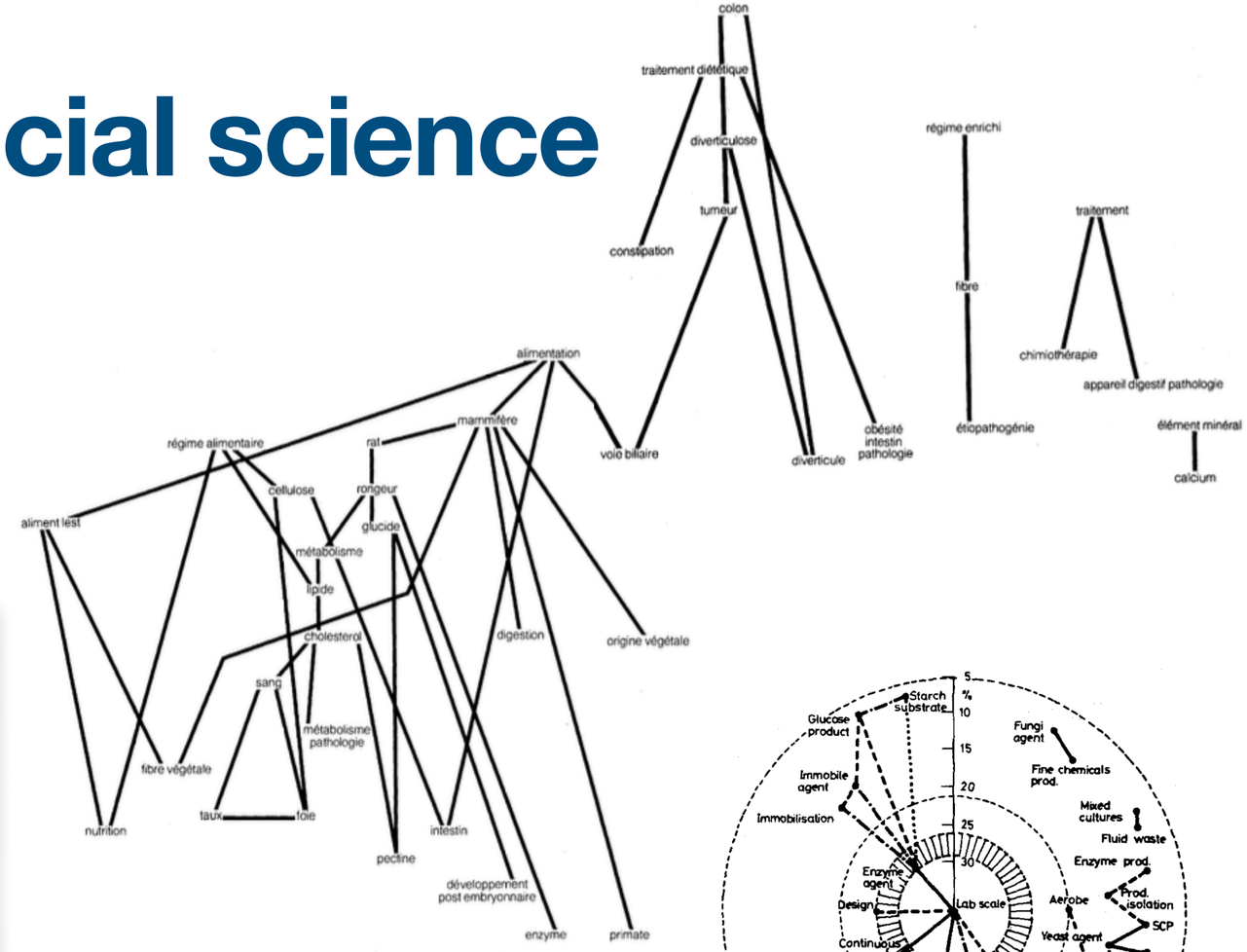
(in hybrid / qual-quant social science approaches)

semantic graphs in social science

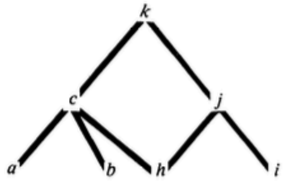
scientometrics

- *co-occurrence* maps in the broad sense
  - « Leximappe »
  - circular maps

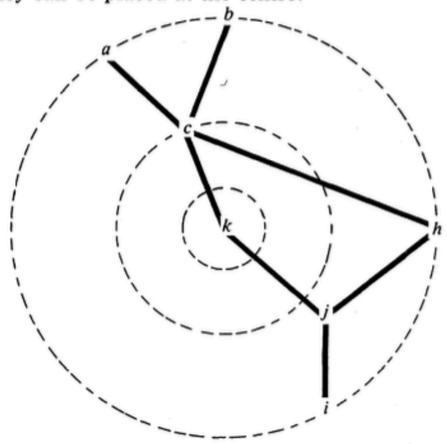
Callon, Courtial, Turner, Bauin, 1983



— the most frequently occurring interests — words — can be placed on top of the graph



— or they can be placed at the centre.



Rip, Courtial, 1984

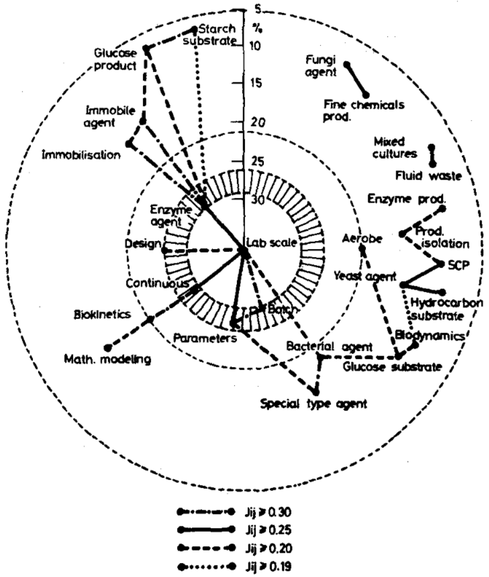


Fig. 1. Circular map, Jaccard, 1970–74
Jaccard links with keywords below 5% frequency have been deleted

semantic graphs in social science

scientometrics

- *co-occurrence* maps in the broad sense
  - co-citation, featuring actor co-occurrence
  - MDS visualization and validation

McCain, 1986

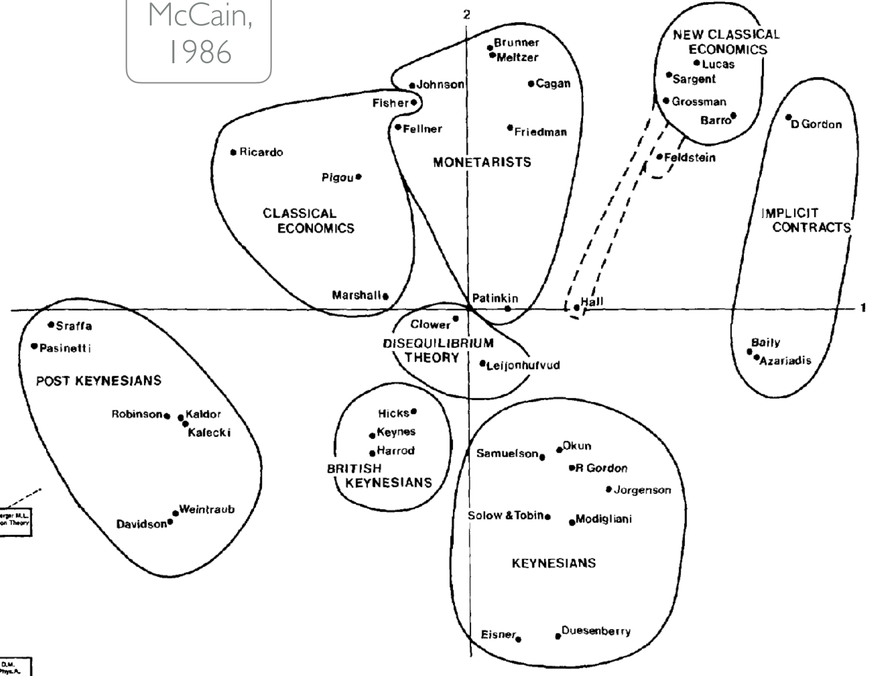
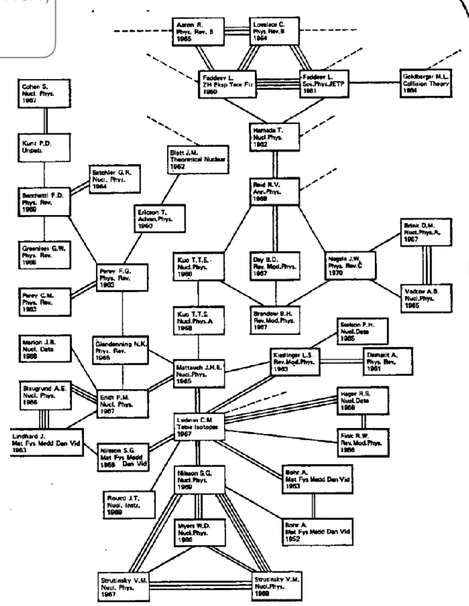


FIG. 2. Map of similarity judgments for forty-one authors in macroeconomics.

Small & Griffith, 1974

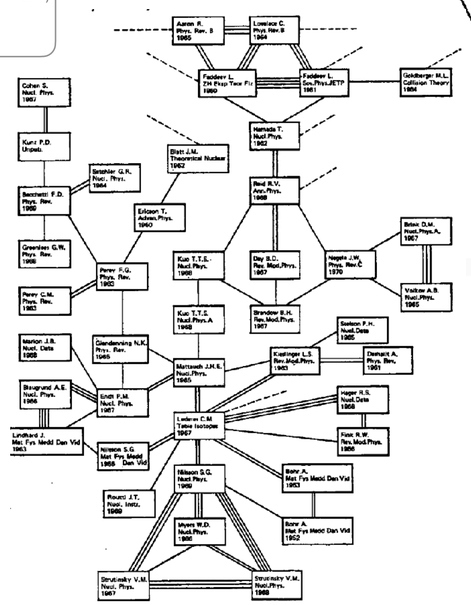


semantic graphs in social science

scientometrics

- *co-occurrence* maps in the broad sense
  - co-citation, featuring actor co-occurrence
  - MDS visualization and validation

Small & Griffith, 1974



McCain, 1986

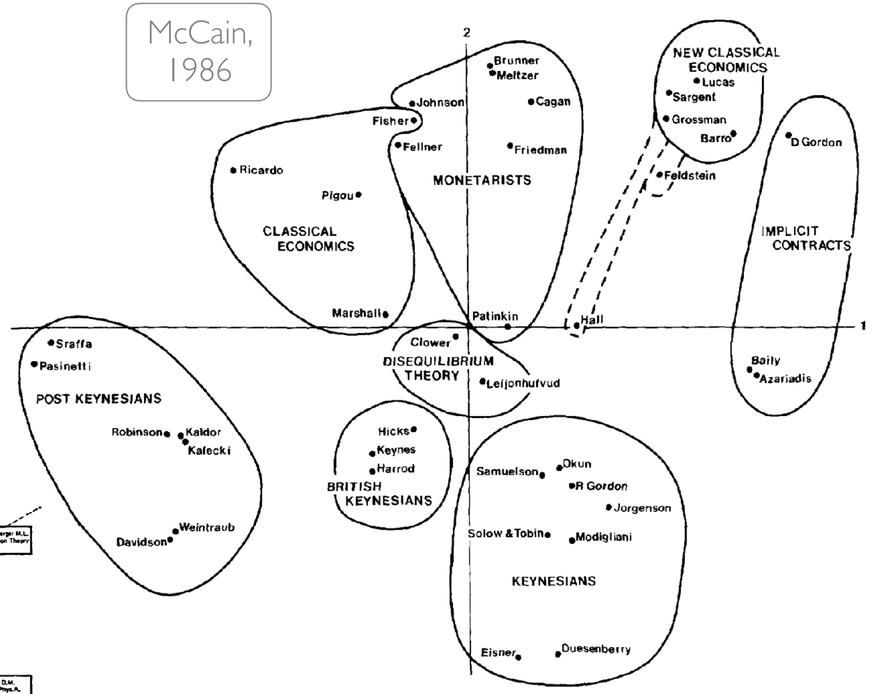


FIG. 2. Map of similarity judgments for forty-one authors in macroeconomics.

Leydesdorff 1989, 2010

\+ *clustering*

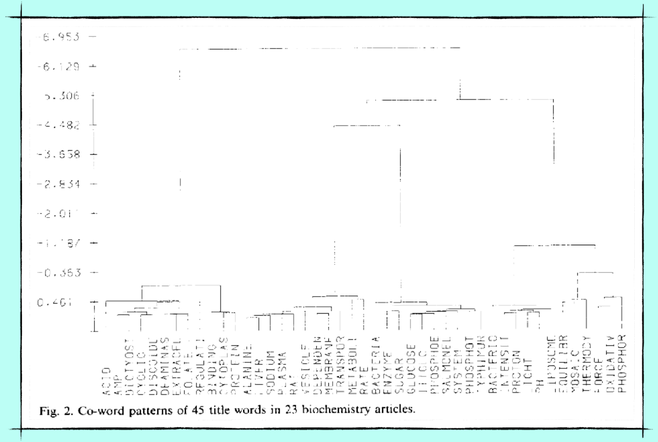


Fig. 2. Co-word patterns of 45 title words in 23 biochemistry articles.

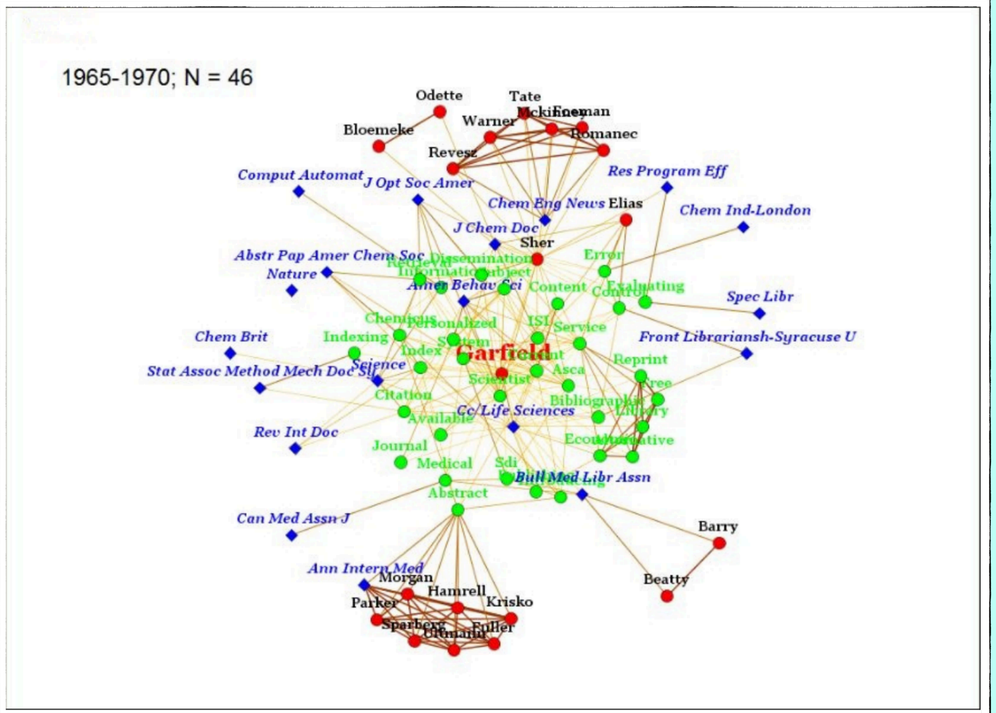


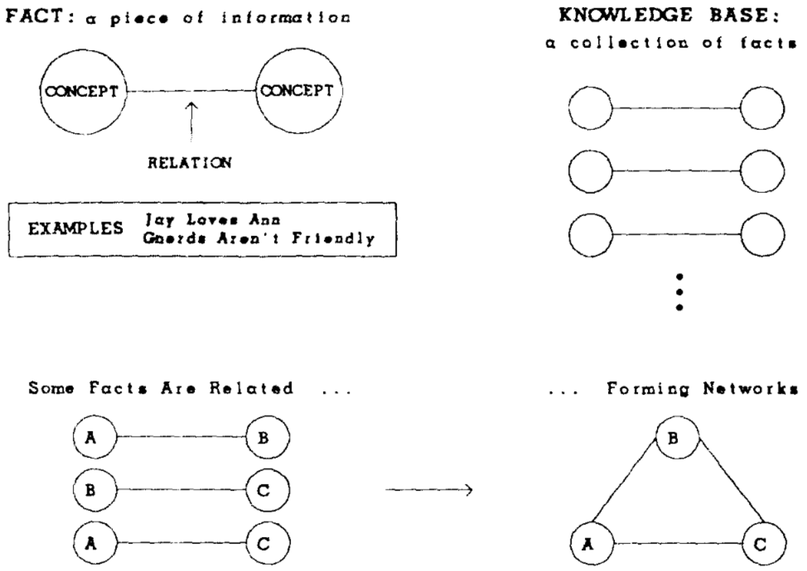
Figure 3: Development during the 1960s. Garfield uses his semantics in different directions in co-authorship relations with other groups. The pattern of publications in many journals is established in the second period (1965-1970).

semantic graphs in social science

cultural sociology

- *hybrid and actor-centric networks / cognitive maps*

Carley, 1986 | Carley, 1994 | Carley, 1997

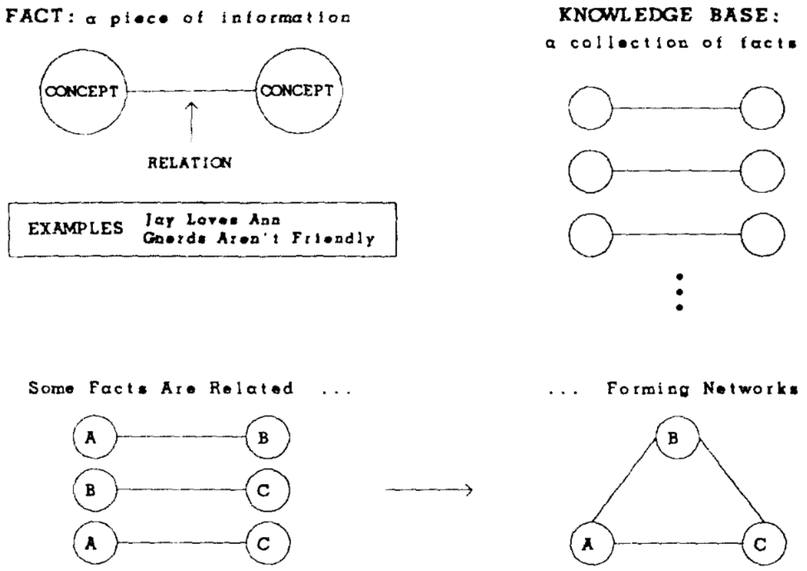


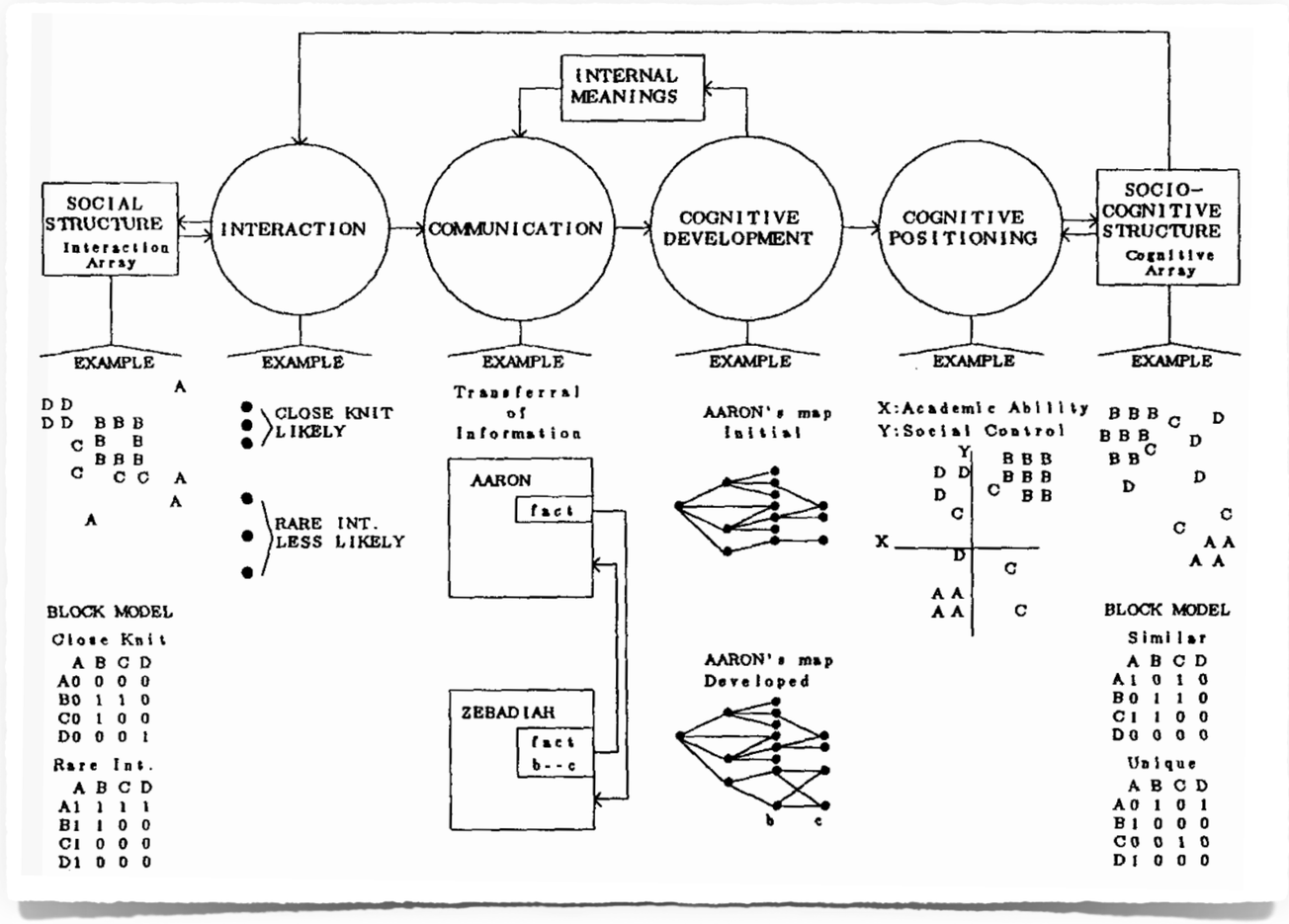
semantic graphs in social science

cultural sociology

- *hybrid and actor-centric networks / cognitive maps*

Carley, 1986 | Carley, 1994 | Carley, 1997





Typical SNA concepts,
such as centrality, transitivity and block-models,
are computed on both networks

semantic graphs in social science

cultural sociology

- *hybrid and actor-centric networks / cognitive maps*

Carley, 1986 | Carley, 1994 | Carley, 1997

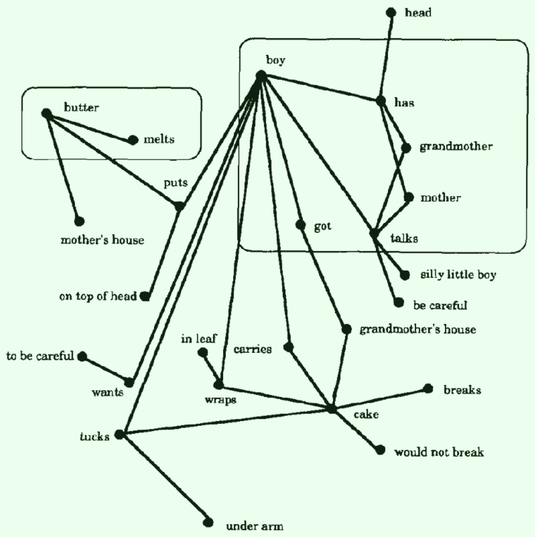


Fig. 6. Shared view of story.

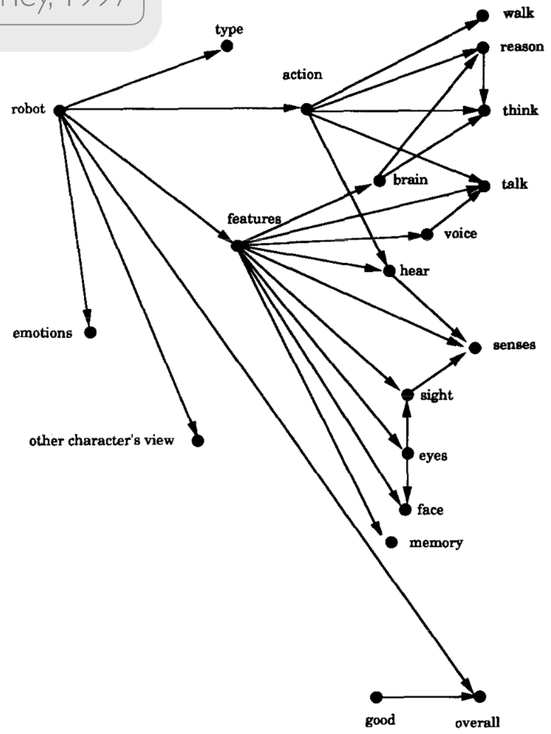


Fig. 2. Definition of robot as shared across three time periods.

semantic graphs in social science

cultural sociology

- *hybrid and actor-centric networks / cognitive maps*

Carley, 1986 | Carley, 1994 | Carley, 1997

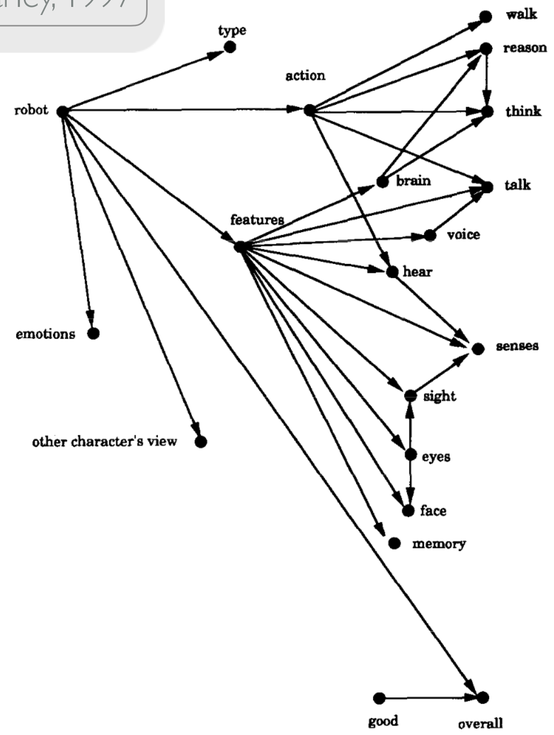


Fig. 2. Definition of robot as shared across three time periods.

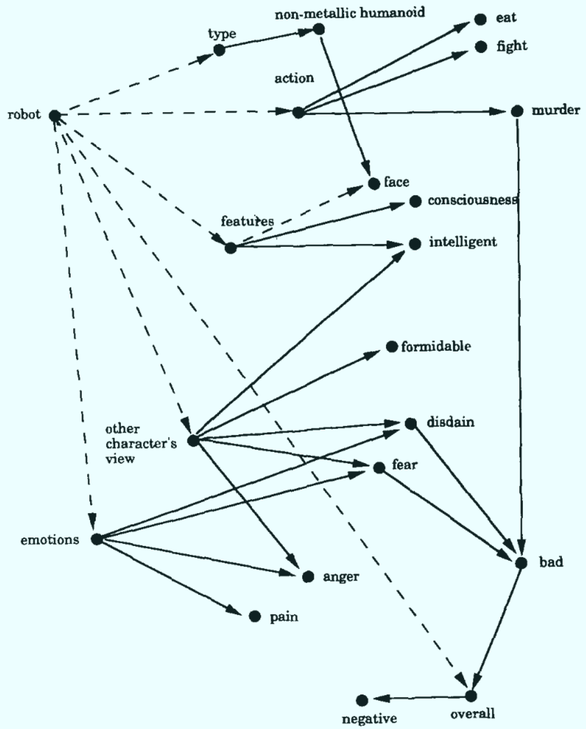


Fig. 3. Shared perception of robot unique to the pre-1950's.

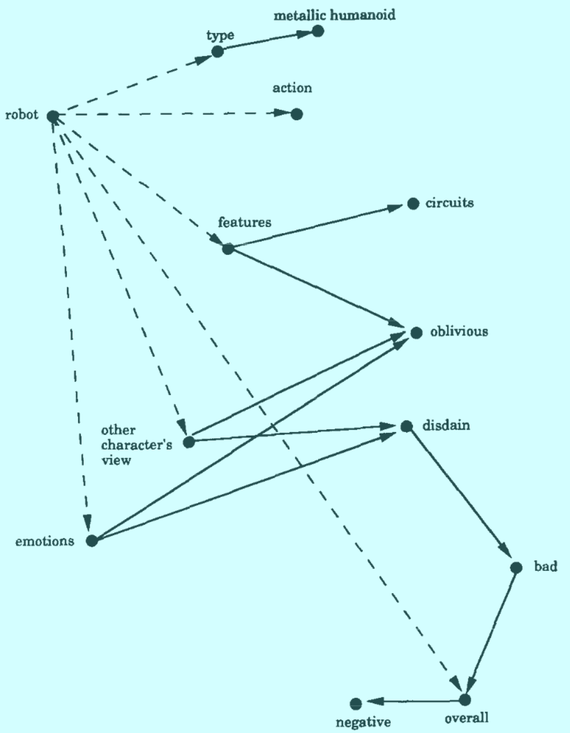


Fig. 4. Shared perception of robot unique to the 1950's and 1960's.

semantic graphs in social science

cultural sociology

- *hybrid and actor-centric networks / cognitive maps*

Carley, 1986 | Carley, 1994 | Carley, 1997

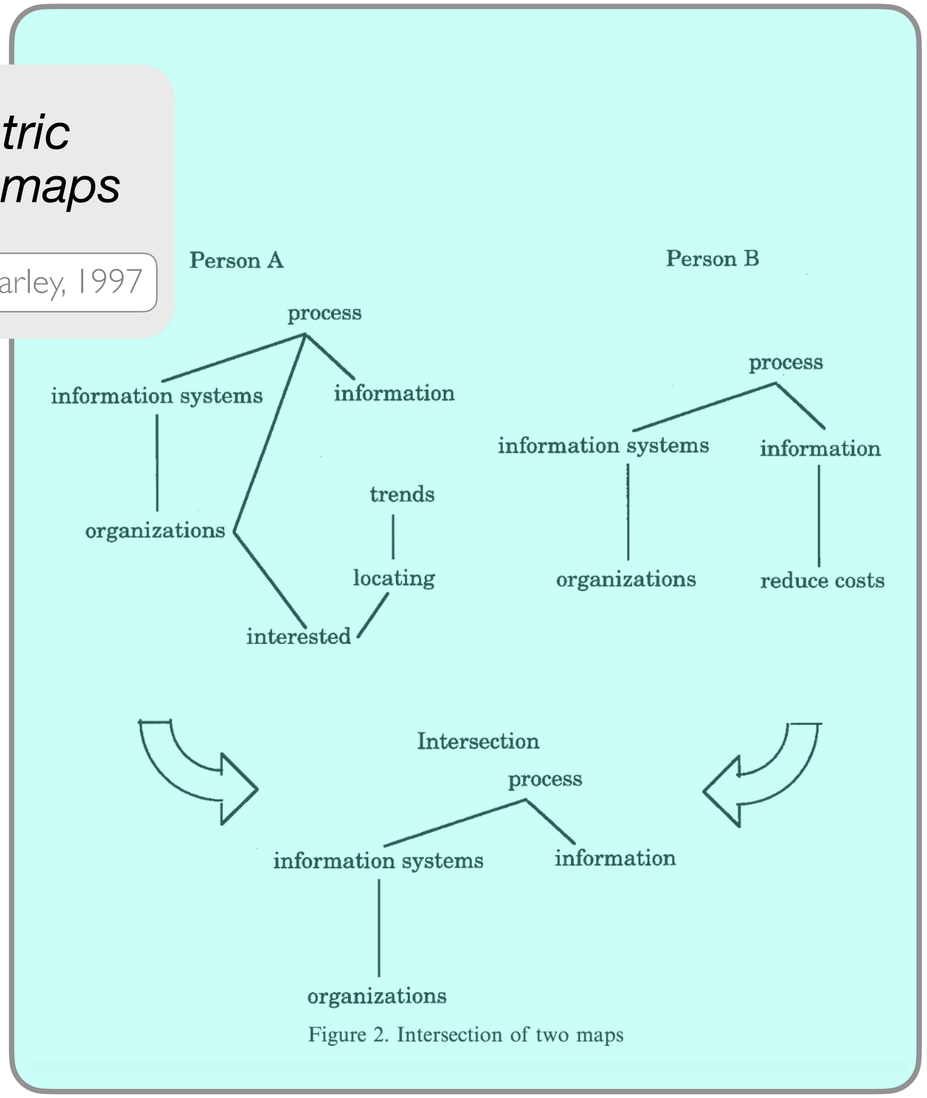


Figure 2. Intersection of two maps

semantic graphs in social science

cultural sociology

- *hybrid and actor-centric networks / cognitive maps*

Carley, 1986 | Carley, 1994 | Carley, 1997

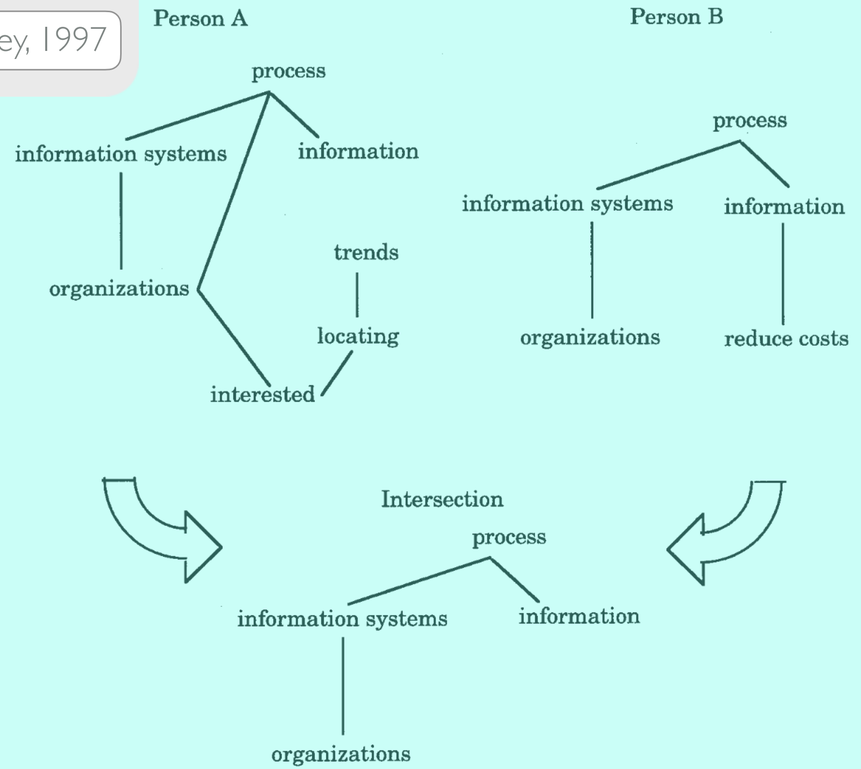


Figure 2. Intersection of two maps

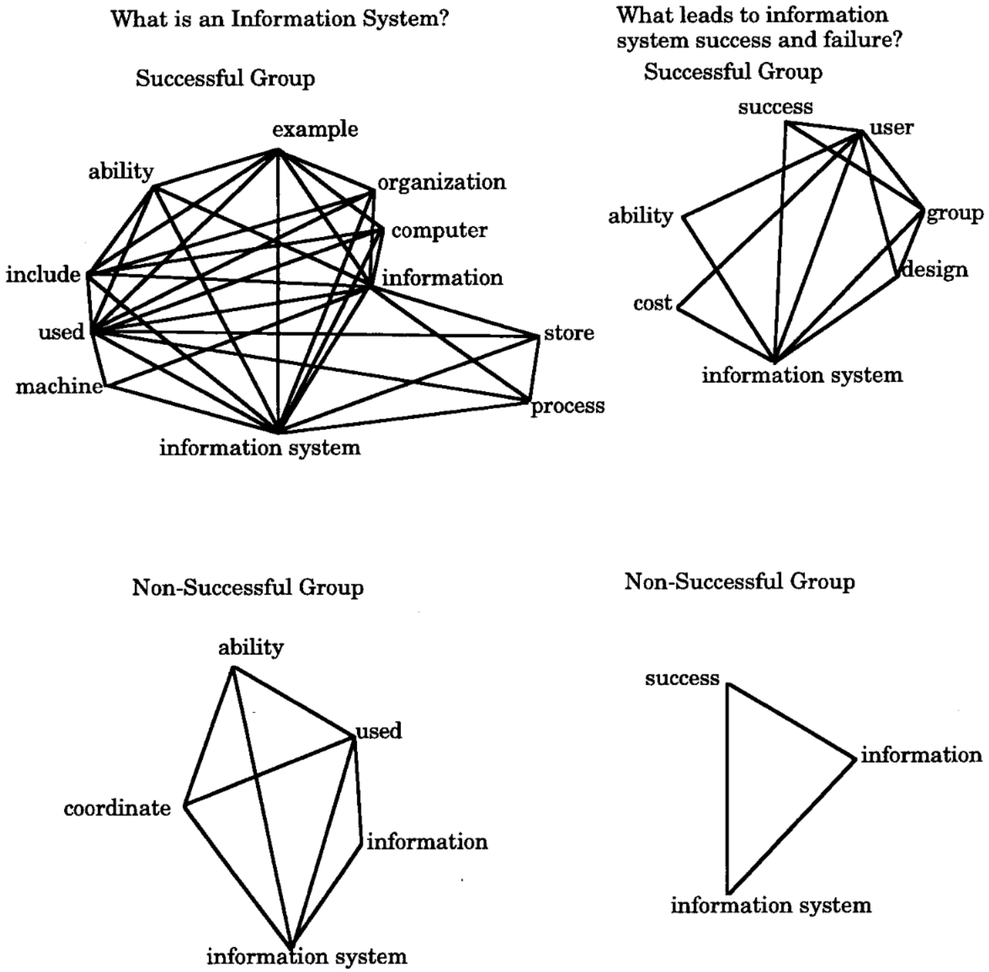


Figure 3. Illustrative team maps

semantic graphs in social science

cultural sociology

- *up to a...meta-matrix*

Butts, Carley, Krackhardt, Ren, 2001

	Personnel	Knowledge	Resources	Tasks	Organizations
Personnel	PP	PK	PR	PT	PO
Knowledge	KP	KK	KR	KT	KO
Resources	RP	RK	RR	RT	RO
Tasks	TP	TK	TR	TT	TO
Organizations	OP	OK	OR	OT	OO

socio-semantic visualization

conversation maps

Sack, 2000

"Conversation Map: An Interface for Very Large-Scale Conversations"

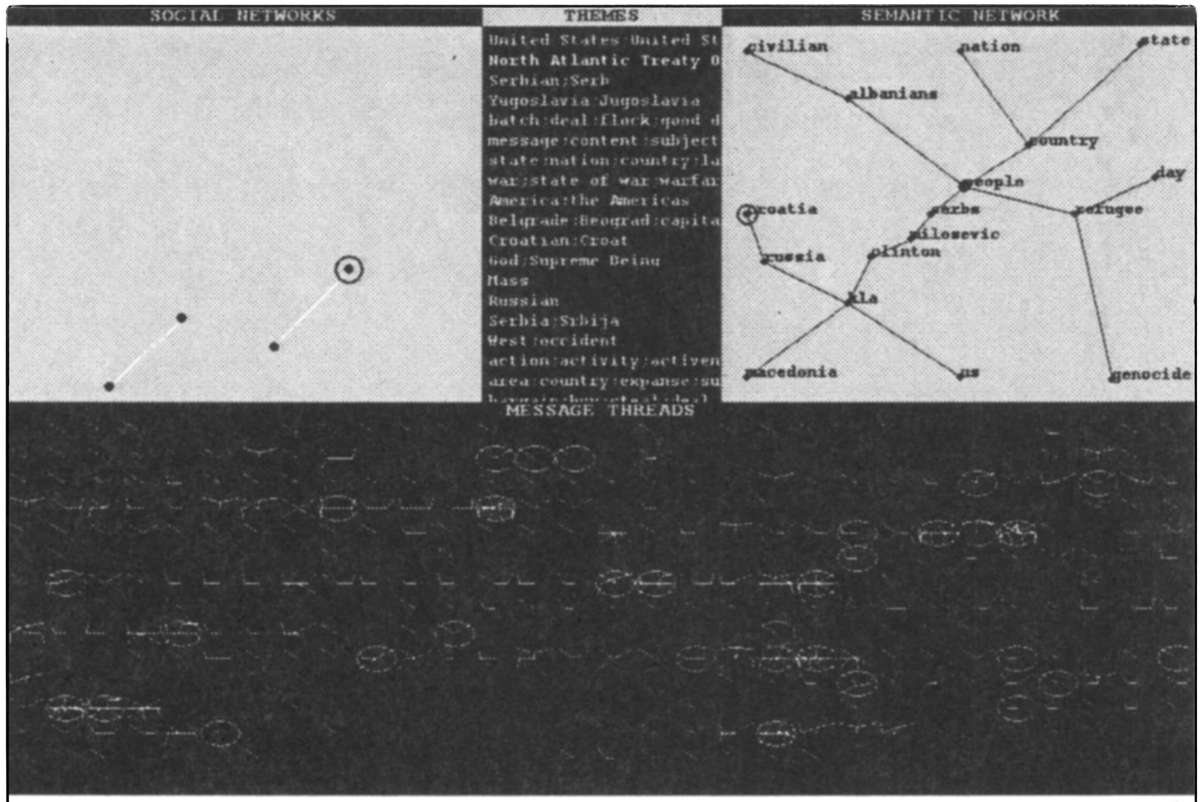


Figure 1. Example of a Conversation Map

socio-semantic visualization

conversation maps

Sack, 2000

"Conversation Map: An Interface for Very Large-Scale Conversations"

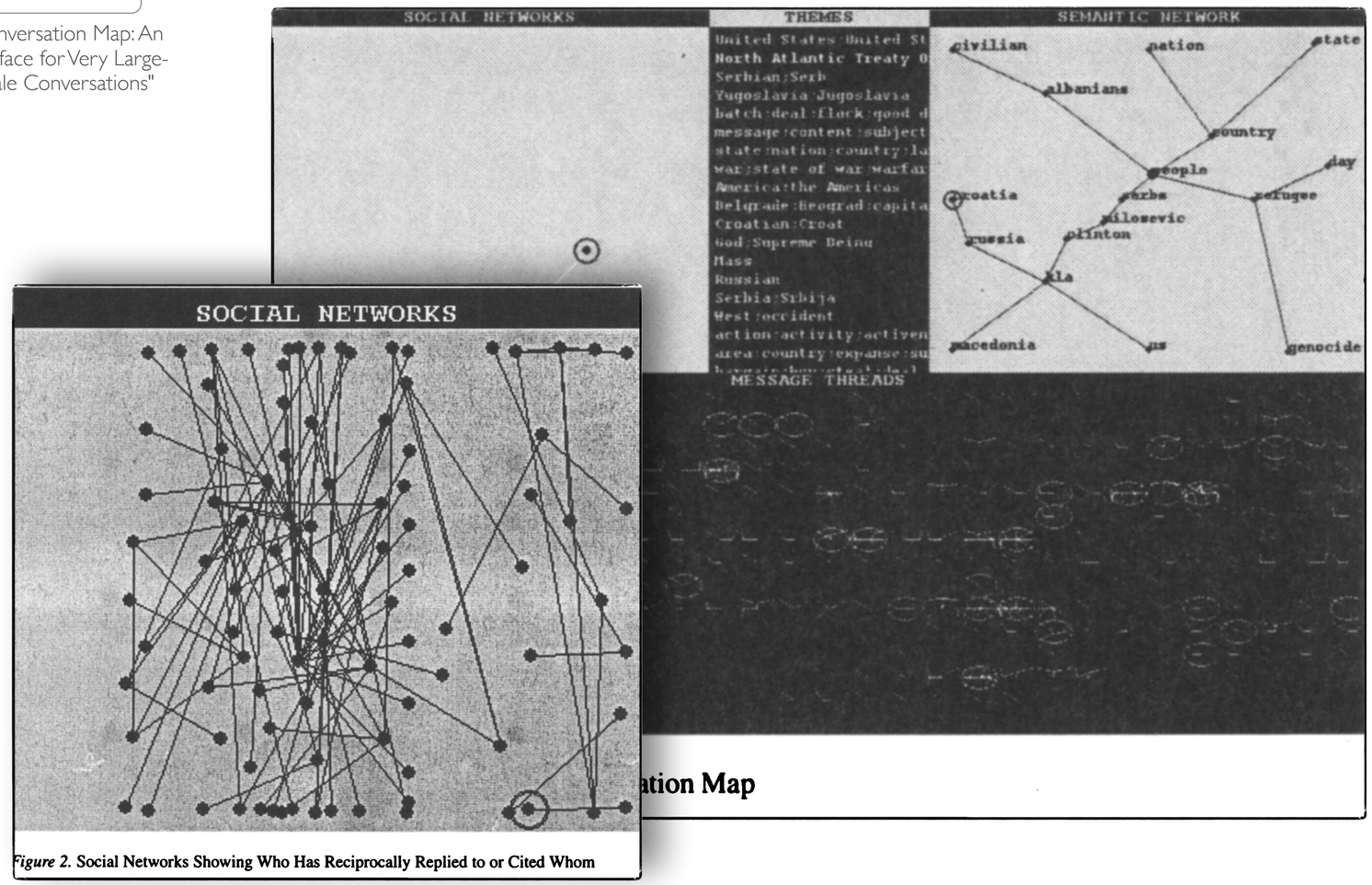


Figure 2. Social Networks Showing Who Has Reciprocally Replied to or Cited Whom

ation Map

socio-semantic visualization

conversation maps

Sack, 2000

"Conversation Map: An Interface for Very Large-Scale Conversations"

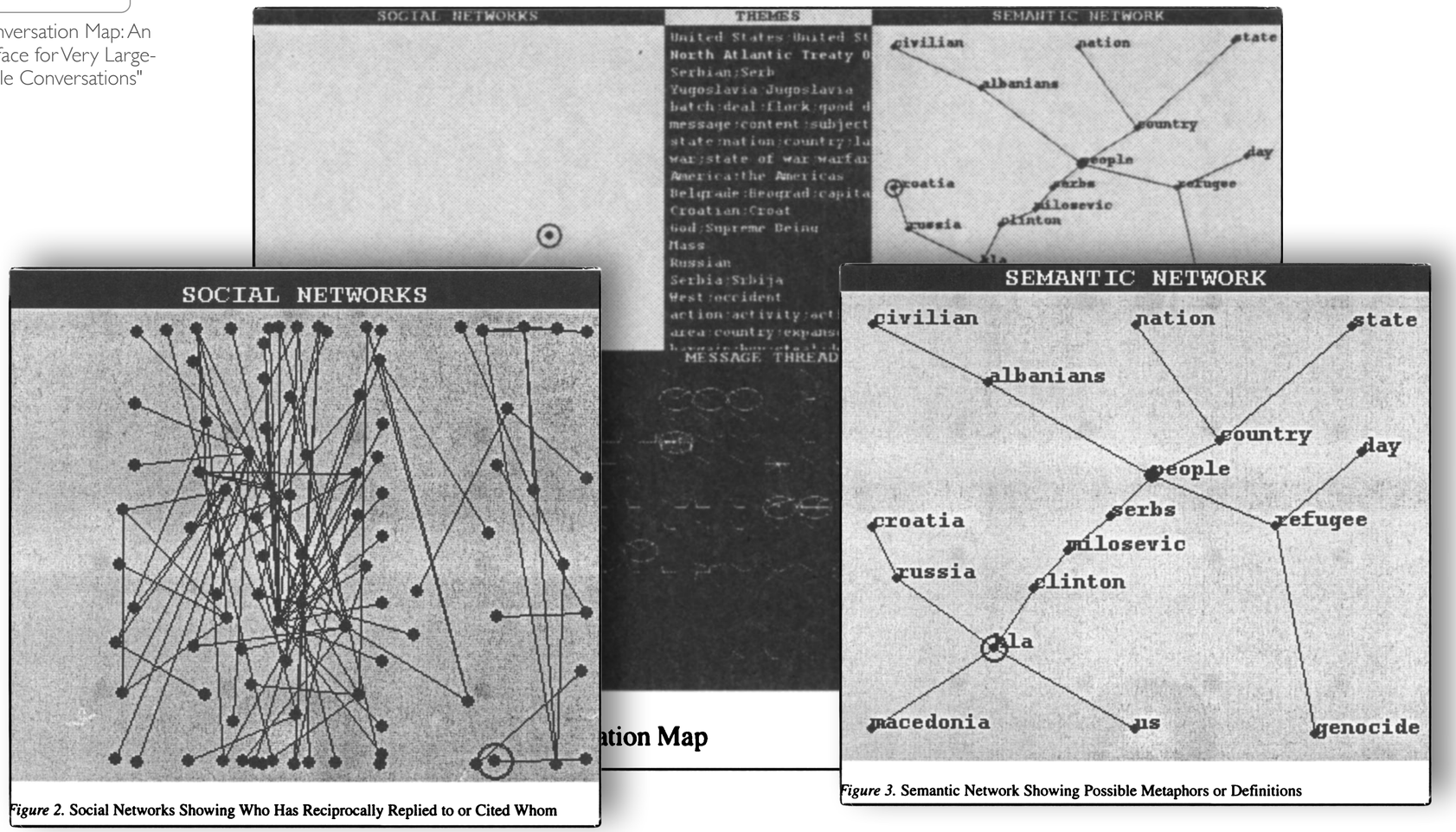


Figure 2. Social Networks Showing Who Has Reciprocally Replied to or Cited Whom

Figure 3. Semantic Network Showing Possible Metaphors or Definitions

socio-semantic visualization

conversation maps

Sack, 2000

"Conversation Map: An Interface for Very Large-Scale Conversations"

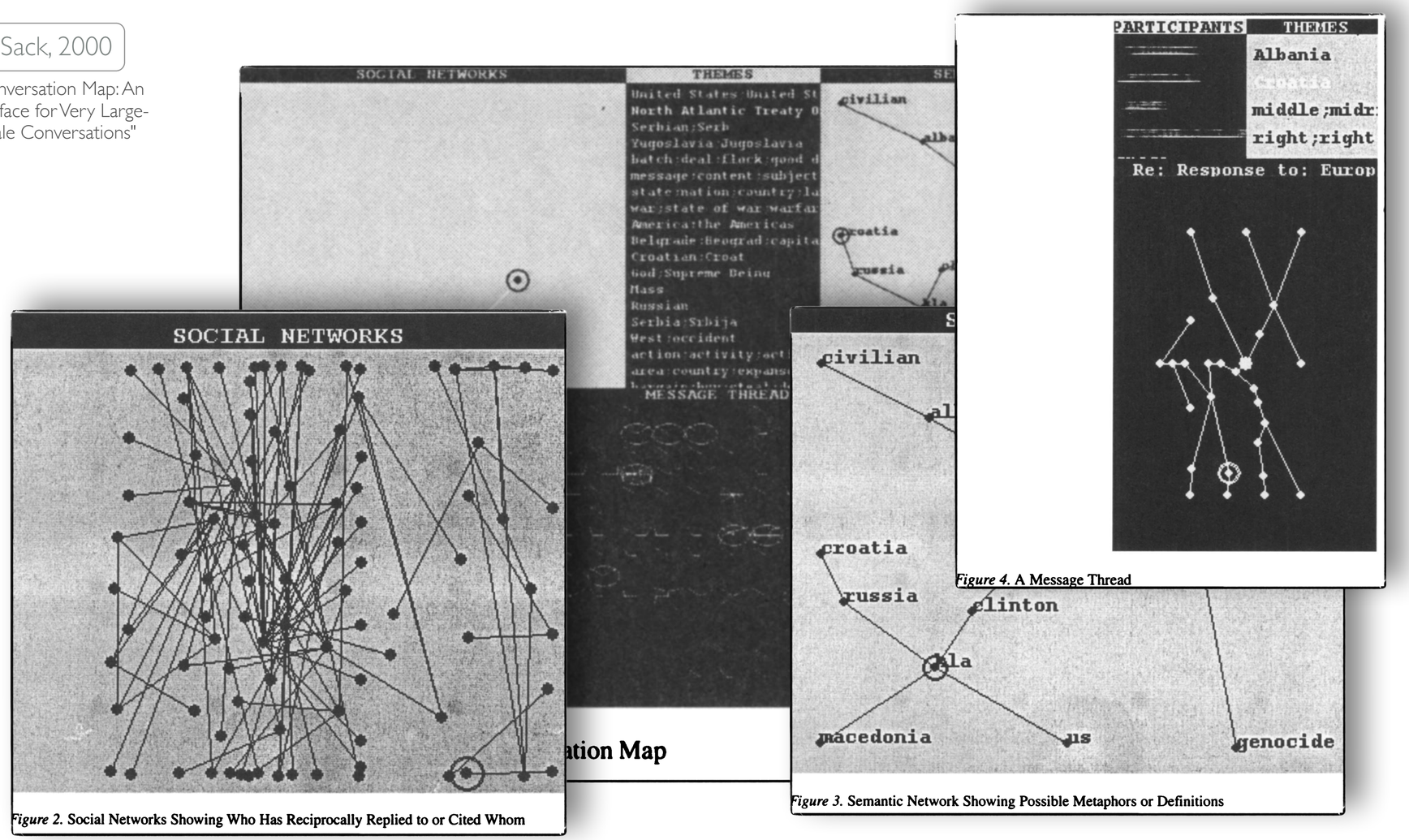


Figure 2. Social Networks Showing Who Has Reciprocally Replied to or Cited Whom

Figure 3. Semantic Network Showing Possible Metaphors or Definitions

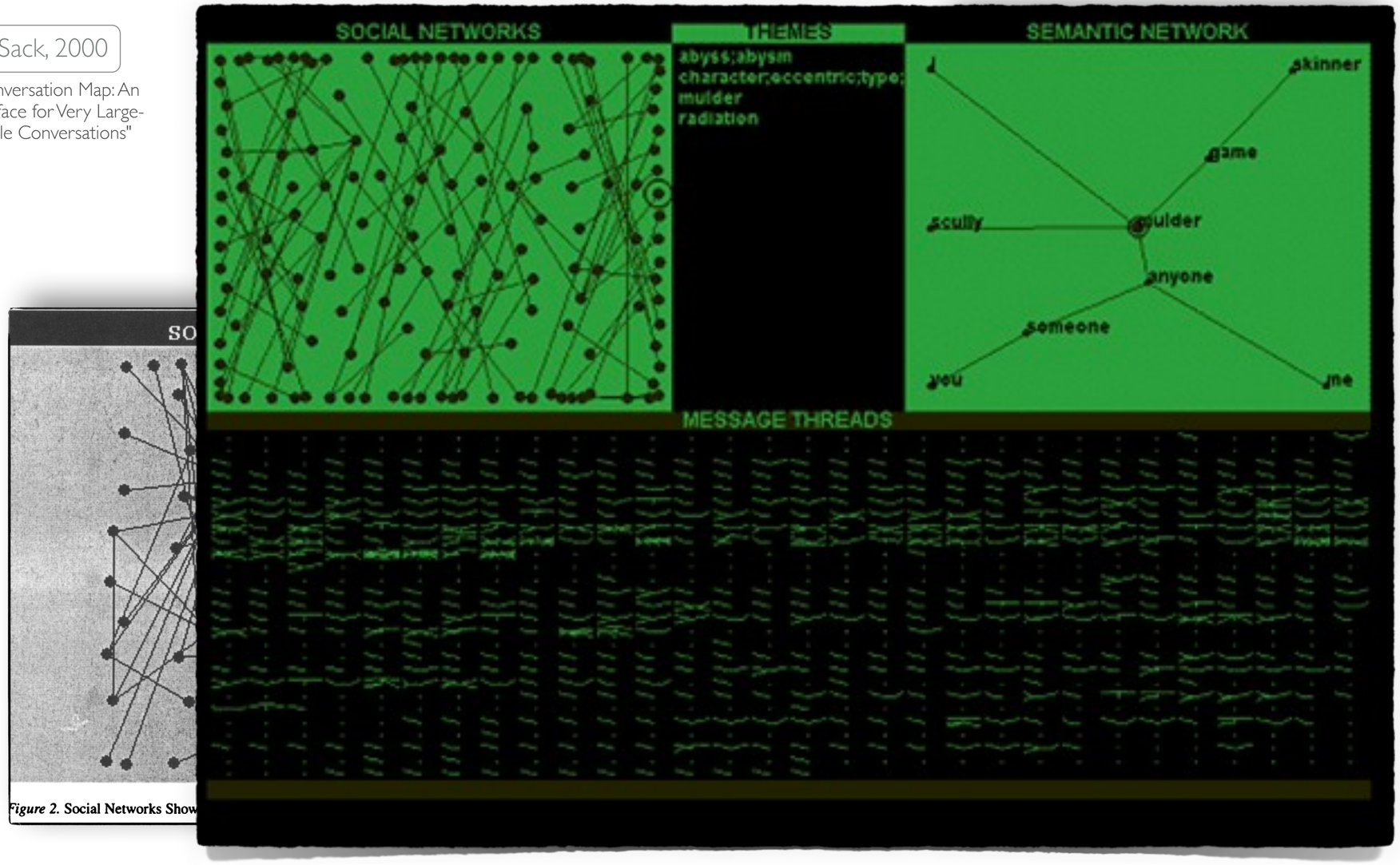
Figure 4. A Message Thread

socio-semantic visualization

conversation maps

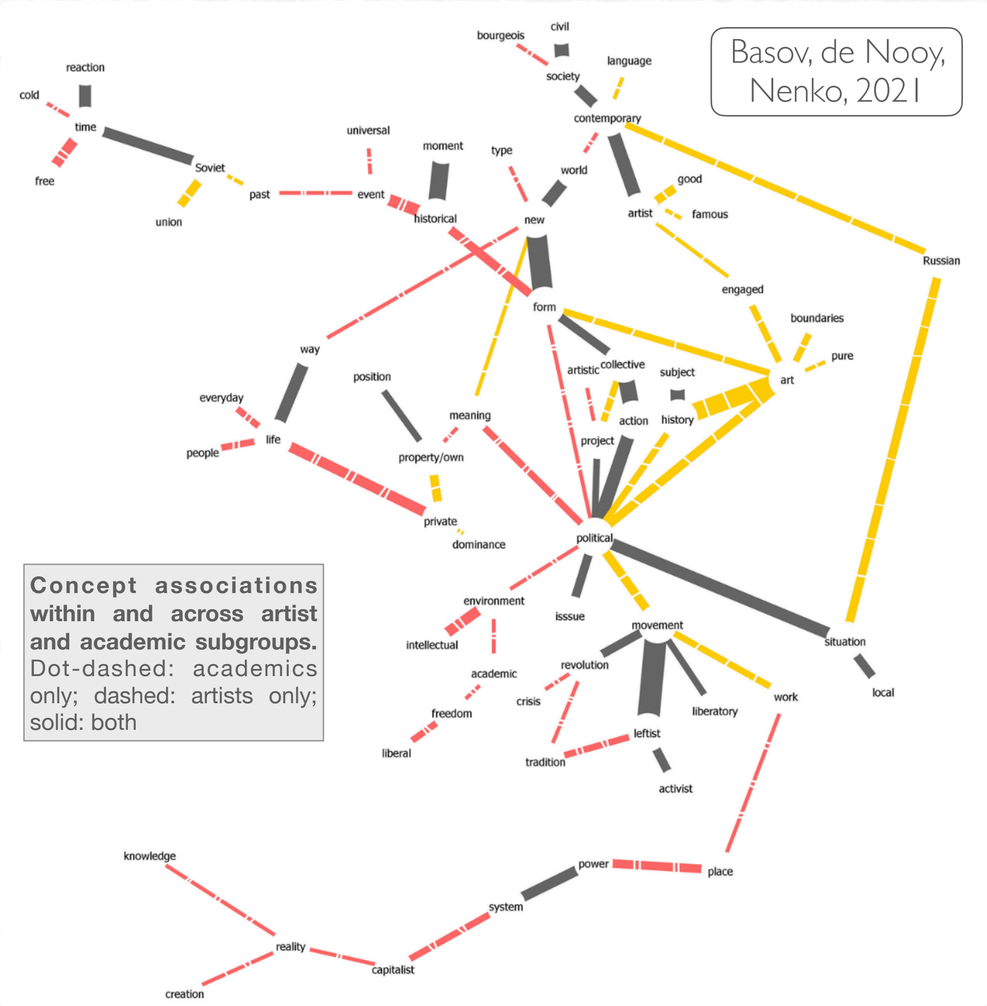
Sack, 2000

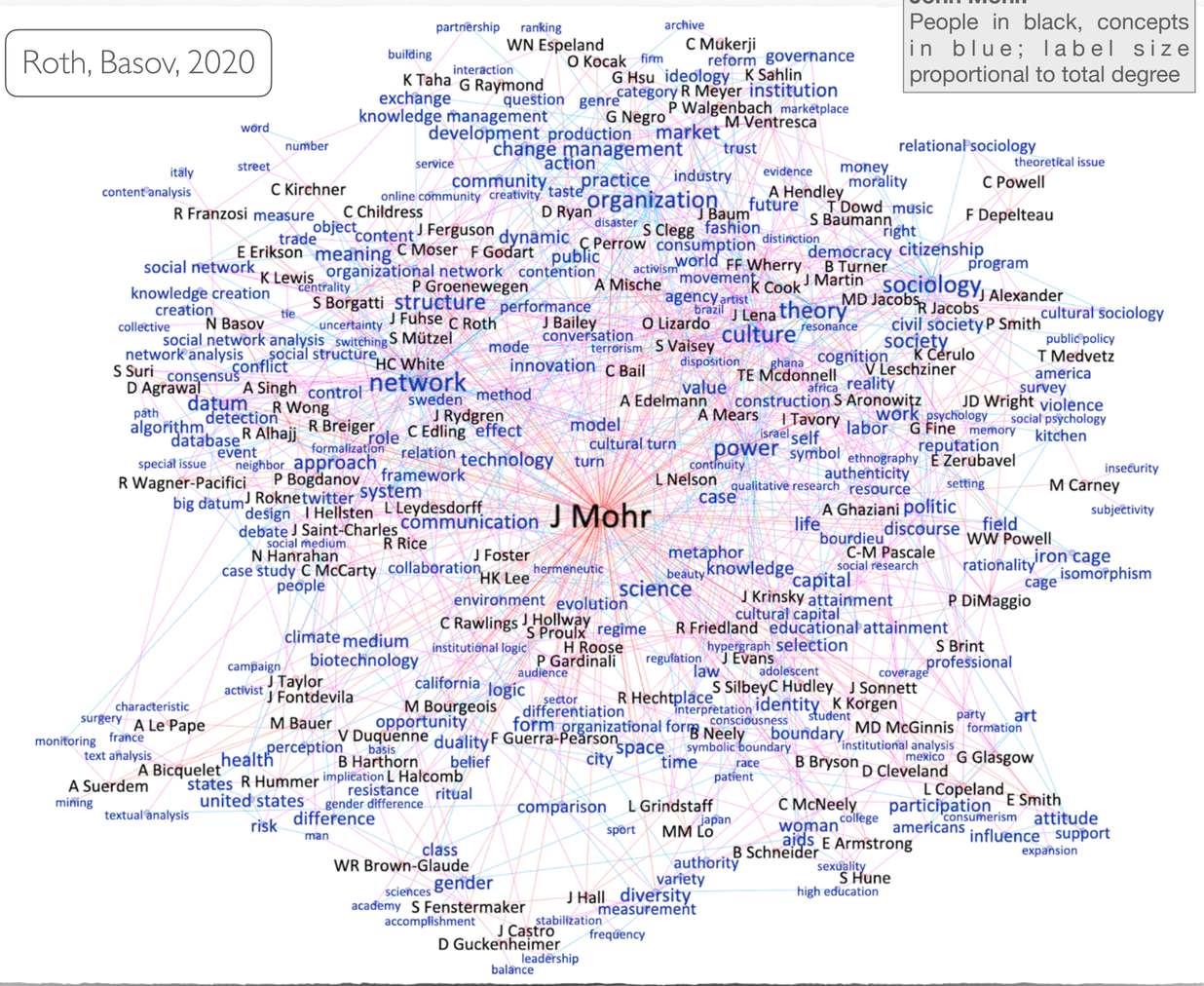
"Conversation Map: An Interface for Very Large-Scale Conversations"



semantic graphs in social science

cultural sociology and scientometrics





Roth & Bourgine 2003 (.....) Roth 2021

socio-semantic formalisms

ontology

asymmetry social vs. semantic ontology

social nodes
individuals,
rarely collective entities

semantic nodes
often abstracted from text corpora

- words, n-grams, terms?
- meaning stability across contexts, individuals, time?
- terms, topics, sentences, vectors, embeddings?

social links
often binarized, weights = intensity

semantic links
cooccurrences... or a myriad of similarity metrics

Roth & Bourgine 2003 (.....) Roth 2021

socio-semantic formalisms

ontology

asymmetry social vs. semantic ontology

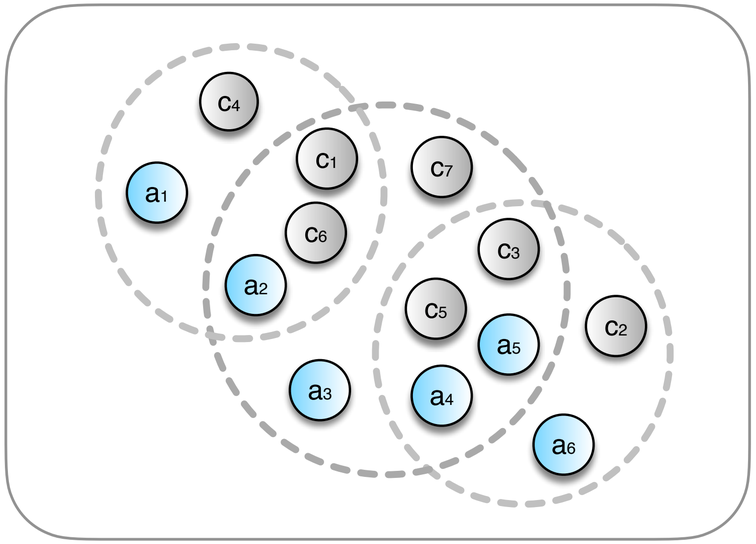
social nodes
individuals,
rarely collective entities

semantic nodes
often abstracted from text corpora

- words, n-grams, terms?
- meaning stability across contexts, individuals, time?
- terms, topics, sentences, vectors, embeddings?

social links
often binarized, weights = intensity

semantic links
cooccurrences... or a myriad of similarity metrics



socio-semantic hypergraph *X*

Roth & Bourgine 2003 (.....) Roth 2021

socio-semantic formalisms

ontology

asymmetry social vs. semantic ontology

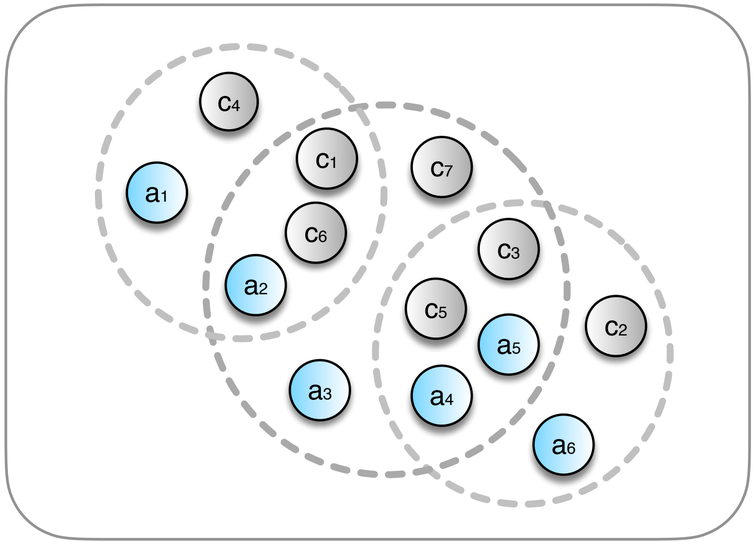
social nodes
individuals,
rarely collective entities

semantic nodes
often abstracted from text corpora

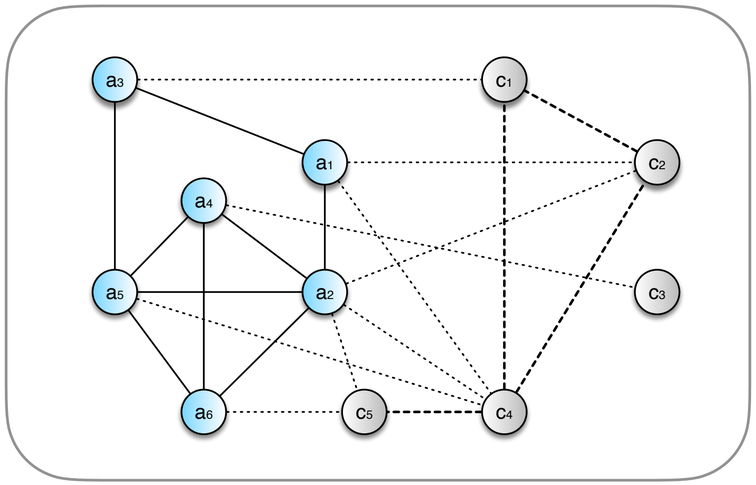
- words, n-grams, terms?
- meaning stability across contexts, individuals, time?
- terms, topics, sentences, vectors, embeddings?

social links
often binarized, weights = intensity

semantic links
cooccurrences... or a myriad of similarity metrics



socio-semantic hypergraph X

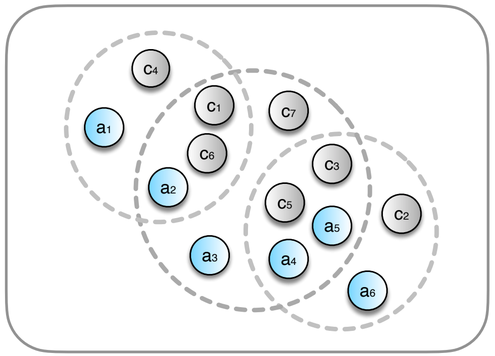


socio-semantic network $(s, \mathfrak{s}, \chi)$

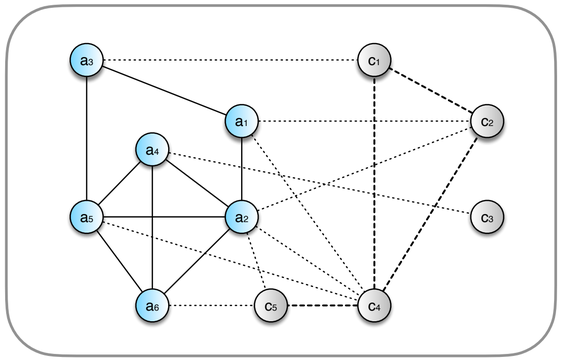
socio-semantic formalisms

ontology

Carley 1986	a common social network one semantic network per actor
Mohr & Duquenne 1997	connection between actors and concepts
Leenders 1997	attributed social network
Carley 2002	meta-network (actors, resources, talks, skills)



socio-semantic hypergraph X

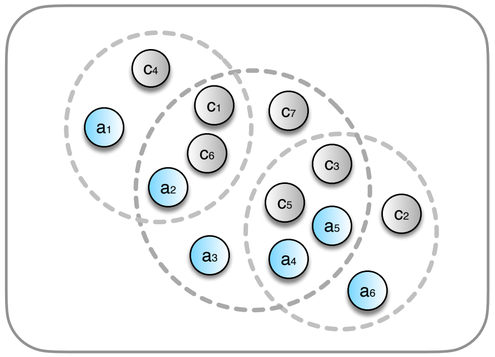


socio-semantic network $(s, \mathfrak{s}, \chi)$

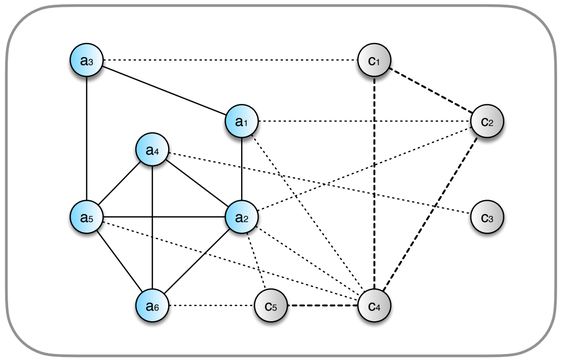
socio-semantic formalisms

ontology

Carley 1986	a common social network one semantic network per actor
Mohr & Duquenne 1997	connection between actors and concepts
Leenders 1997	attributed social network
Carley 2002	meta-network (actors, resources, talks, skills)



socio-semantic hypergraph X



socio-semantic network $(s, \mathfrak{s}, \chi)$

socio-semantic et al.

socio-semantic web	Zacklad et al., 2003
semantic-social networks	Mika, 2007
epistemic networks	Grim, 2009; Weisberg & Muldoon, 2009
semantic social networks	Gloor et al, 2009
attribute-augmented graphs	Zhou et al., 2009
bi-type information networks	Sun et al., 2009
content-based social networks	Cucchiarelli et al., 2010
social content networks	Wang & Groth, 2010
augmented social networks	Cruz et al., 2013
heterogeneous graphs	Liu et al, 2017

socio-semantic categories

blockmodels and lattices

Mohr 1994 — "Soldiers, mothers, tramps and others: Discourse roles in the 1907 New York City charity directory"

Mohr & Duquenne 1997 — "The Duality of Culture and Practice: Poverty Relief in New York City, 1888-1917"

Table 2. Poverty practices by poverty categories (frequencies) – 1917.

	Deserving	Destitute	Distressed	Fallen	Homeless	Indigent	Misfortune	Needy	Poor	Worthy	Stranger	Total
Give$	1	9	2	0	0	0	1	7	19	2	0	41
Food	0	3	0	0	5	0	0	5	14	5	0	32
PaidWk	0	2	0	0	4	0	0	0	1	0	0	7
HomeWk	1	0	0	0	0	0	0	0	10	0	0	11
FindJob	0	3	0	0	6	0	0	2	7	4	0	22
Advise	0	2	1	0	2	0	0	1	1	0	3	10
Investg	0	8	2	0	2	0	0	1	7	1	0	21
JbTrain	0	2	1	0	1	0	0	0	5	0	0	9
Shelter	0	2	0	0	7	2	0	1	4	1	0	17
Asylum	0	20	0	6	3	31	0	5	12	18	0	95
Total	2	51	6	6	30	33	1	22	80	31	3	265

Semantic approach :
hand-coding of high-level categories appearing as rows and columns, then simple pattern matching from the New York City Charity Directories

socio-semantic categories

blockmodels and lattices

Mohr 1994 — "Soldiers, mothers, tramps and others: Discourse roles in the 1907 New York City charity directory"

Mohr & Duquenne 1997 — "The Duality of Culture and Practice: Poverty Relief in New York City, 1888-1917"

Table 2. Poverty practices by poverty categories (frequencies) – 1917.

	Deserving	Destitute	Distressed	Fallen	Homeless	Indigent	Misfortune	Needy	Poor	Worthy	Stranger	Total
Give$	1	9	2	0	0	0	1	7	19	2	0	41
Food	0	3	0	0	5	0	0	5	14	5	0	32
PaidWk	0	2	0	0	4	0	0	0	1	0	0	7
HomeWk	1	0	0	0	0	0	0	0	10	0	0	11
FindJob	0	3	0	0	6	0	0	2	7	4	0	22
Advise	0	2	1	0	2	0	0	1	1	0	3	10
Investg	0	8	2	0	2	0	0	1	7	1	0	21
JbTrain	0	2	1	0	1	0	0	0	5	0	0	9
Shelter	0	2	0	0	7	2	0	1	4	1	0	17
Asylum	0	20	0	6	3	31	0	5	12	18	0	95
Total	2	51	6	6	30	33	1	22	80	31	3	265

Semantic approach :
hand-coding of high-level categories appearing as rows and columns, then simple pattern matching from the New York City Charity Directories

Socio-semantic approach :
(1) block-modeling of (relief x organization type) X (target x gender)
(2) lattice on (relief x target)

Mohr 1994 — "Soldiers, mothers, tramps and others: Discourse roles in the 1907 New York City charity directory"

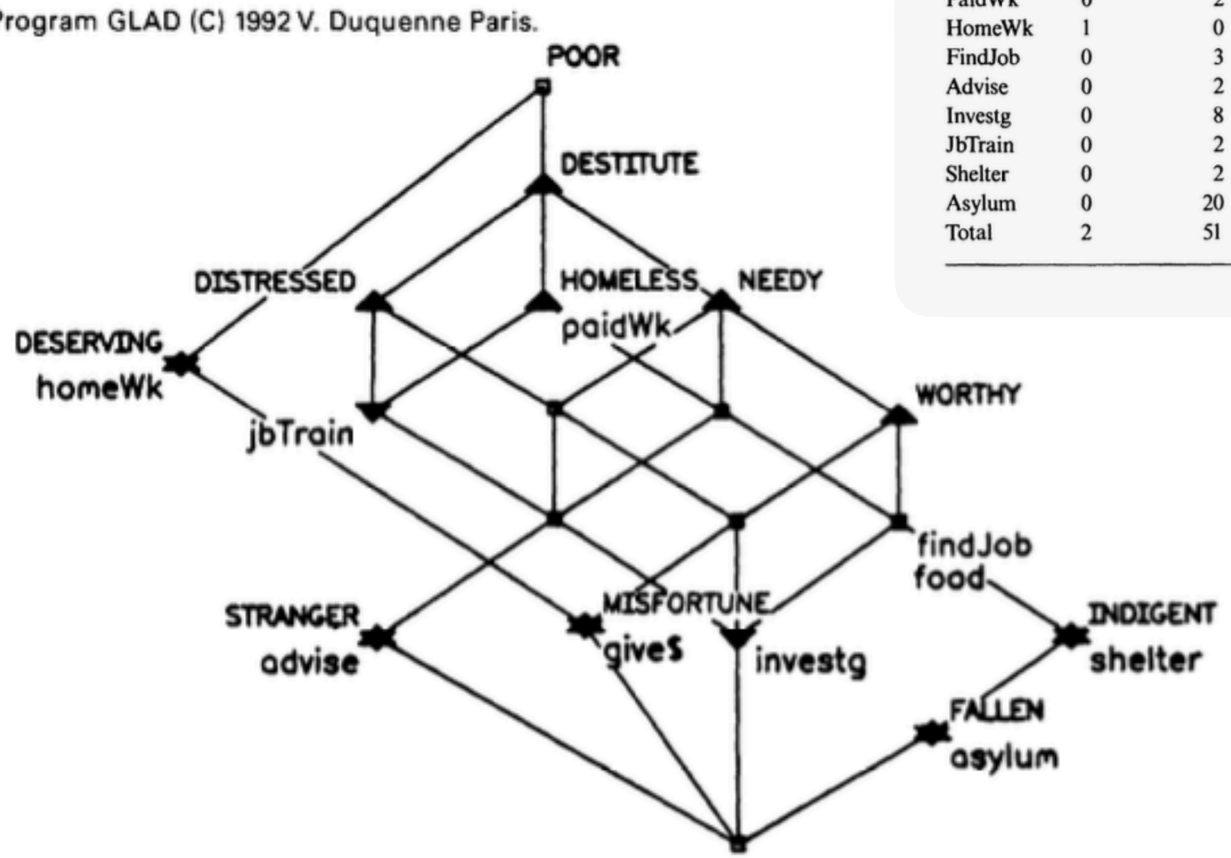
Mohr & Duquenne 1997 — "The Duality of Culture and Practice: Poverty Relief in New York City, 1888-1917"

socio-semantic categories

blockmodels and lattices

Table 2. Poverty practices by poverty categories (frequencies) – 1917.

	Deserving	Destitute	Distressed	Fallen	Homeless	Indigent	Misfortune	Needy	Poor	Worthy	Stranger	Total
Give$	1	9	2	0	0	0	1	7	19	2	0	41
Food	0	3	0	0	5	0	0	5	14	5	0	32
PaidWk	0	2	0	0	4	0	0	0	1	0	0	7
HomeWk	1	0	0	0	0	0	0	0	10	0	0	11
FindJob	0	3	0	0	6	0	0	2	7	4	0	22
Advise	0	2	1	0	2	0	0	1	1	0	3	10
Investg	0	8	2	0	2	0	0	1	7	1	0	21
JbTrain	0	2	1	0	1	0	0	0	5	0	0	9
Shelter	0	2	0	0	7	2	0	1	4	1	0	17
Asylum	0	20	0	6	3	31	0	5	12	18	0	95
Total	2	51	6	6	30	33	1	22	80	31	3	265



Semantic approach :
hand-coding of high-level categories appearing as rows and columns, then simple pattern matching from the New York City Charity Directories

Socio-semantic approach :
(1) block-modeling of (relief x organization type) X (target x gender)
(2) lattice on (relief x target)

socio-semantic categories

towards the meso-level: groups and clusters

socio-semantic clusters

- social clusters are typically deemed to correspond, at least implicitly, to underlying semantic boundaries
  – often a validation criterion

socio-semantic categories

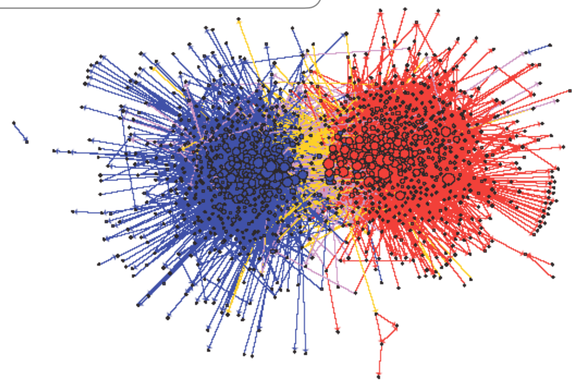
towards the meso-level: groups and clusters

socio-semantic clusters

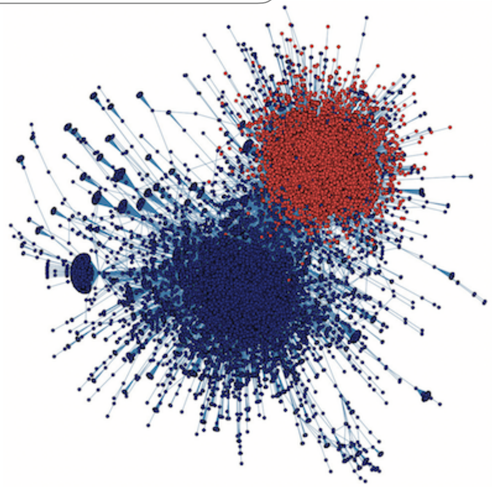
- social clusters are typically deemed to correspond, at least implicitly, to underlying semantic boundaries
  – often a validation criterion

- (1) appraising the socio-semantic cohesiveness of clusters
  are social clusters semantic clusters,
  are semantic clusters social clusters?

Adamic & Glance, 2004



Conover et al. 2011



socio-semantic categories

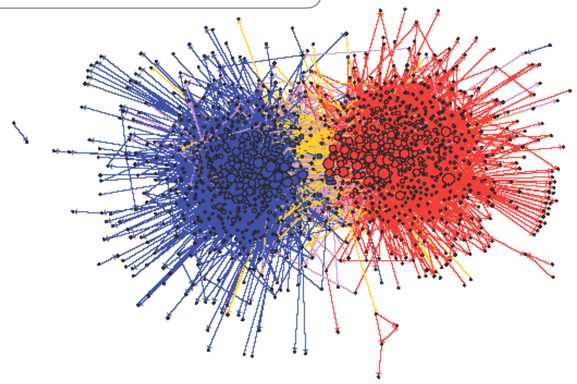
towards the meso-level: groups and clusters

socio-semantic clusters

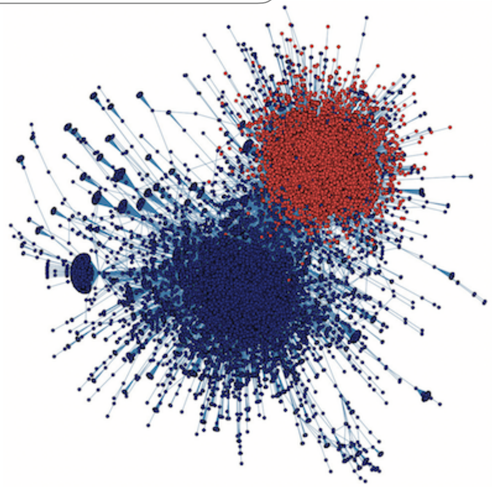
- social clusters are typically deemed to correspond, at least implicitly, to underlying semantic boundaries
  – often a validation criterion

- (1) appraising the socio-semantic cohesiveness of clusters
  are social clusters semantic clusters,
  are semantic clusters social clusters?

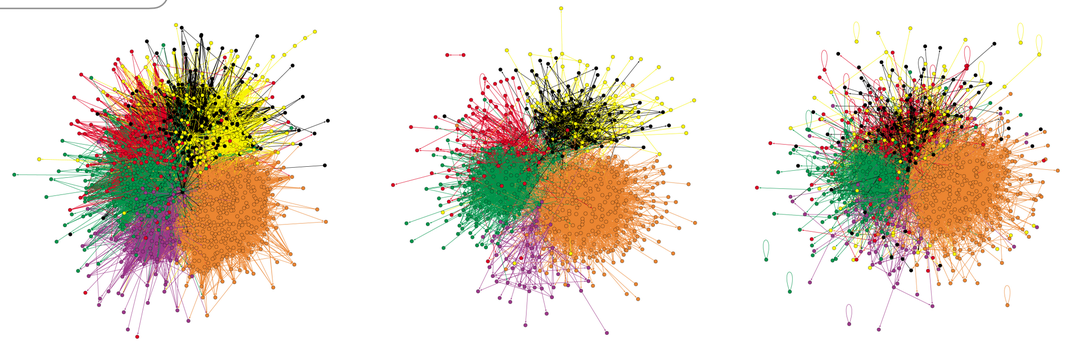
Adamic & Glance, 2004



Conover et al. 2011



Lietz et al. 2014



(a) Following (H=0.83) (b) Retweeting (H=0.90) (c) Mentioning (H=0.79)

socio-semantic categories

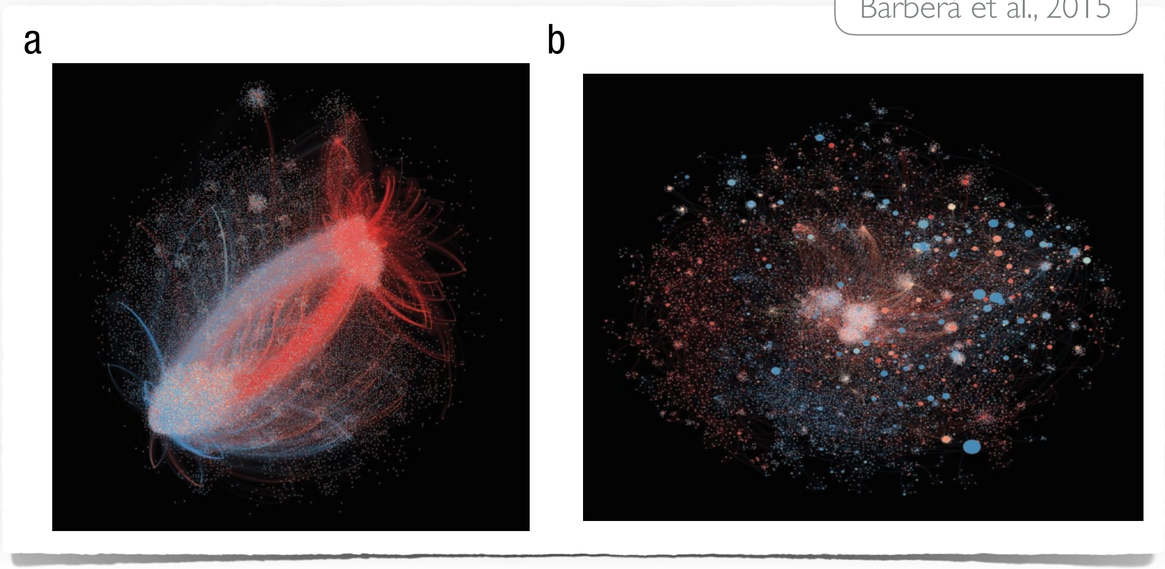
towards the meso-level: groups and clusters

socio-semantic clusters

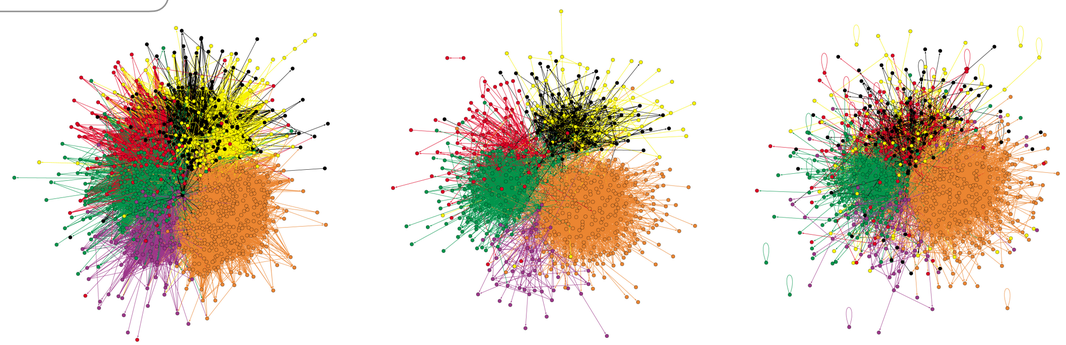
- social clusters are typically deemed to correspond, at least implicitly, to underlying semantic boundaries – often a validation criterion

- (1) appraising the socio-semantic cohesiveness of clusters
  are social clusters semantic clusters,
  are semantic clusters social clusters?

Barberá et al., 2015



Lietz et al. 2014



(a) Following (H=0.83) (b) Retweeting (H=0.90) (c) Mentioning (H=0.79)

socio-semantic categories

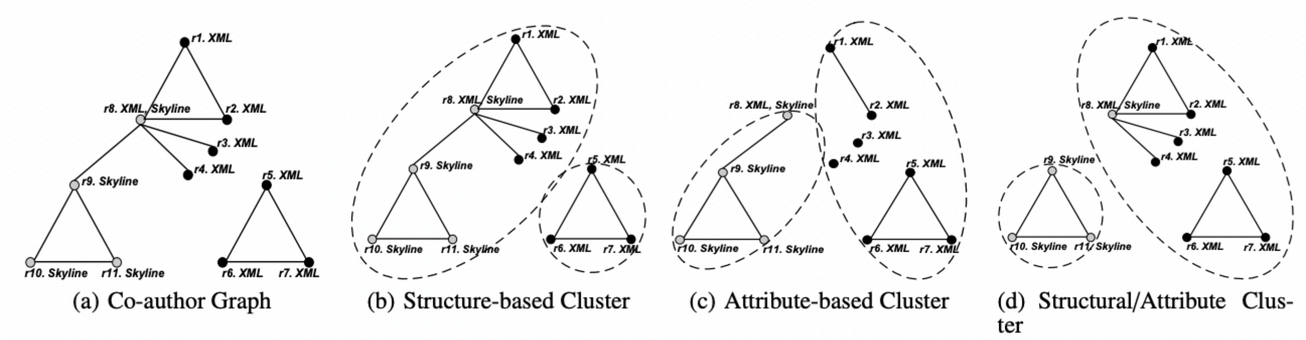
towards the meso-level: groups and clusters

Zhou et al. 2009

socio-semantic clusters

- social clusters are typically deemed to correspond, at least implicitly, to underlying semantic boundaries – often a validation criterion

- (1) appraising the socio-semantic cohesiveness of clusters
  are social clusters semantic clusters, are semantic clusters social clusters?
- (2) assuming the socio-semantic cohesiveness of clusters
  uncovering socio-semantic clusters better (as bipartite clusters, or as social clusters enriched by semantics)



(a) Co-author Graph (b) Structure-based Cluster (c) Attribute-based Cluster (d) Structural/Attribute Cluster

socio-semantic categories

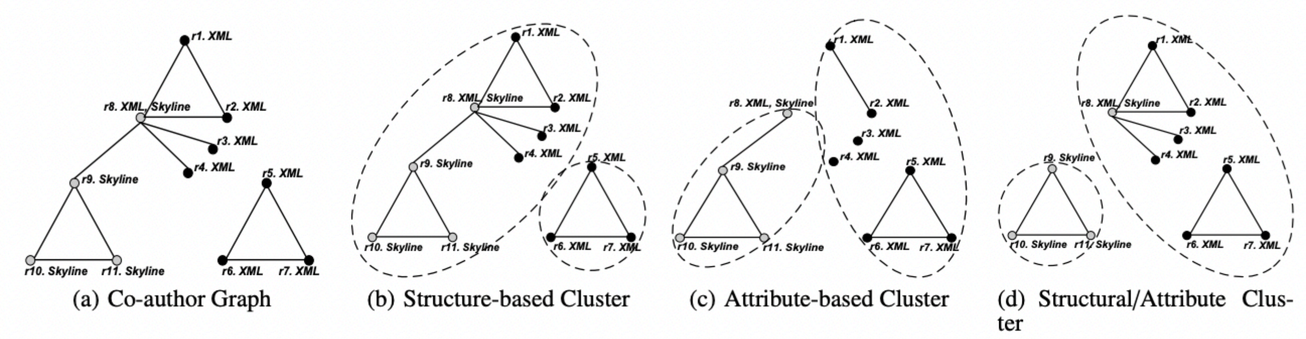
towards the meso-level: groups and clusters

Zhou et al. 2009

socio-semantic clusters

- social clusters are typically deemed to correspond, at least implicitly, to underlying semantic boundaries – often a validation criterion

- (1) appraising the socio-semantic cohesiveness of clusters
  are social clusters semantic clusters, are semantic clusters social clusters?
- (2) assuming the socio-semantic cohesiveness of clusters
  uncovering socio-semantic clusters better (as bipartite clusters, or as social clusters enriched by semantics)



(a) Co-author Graph (b) Structure-based Cluster (c) Attribute-based Cluster (d) Structural/Attribute Cluster

Bouveyron, Latouche, Zreik, 2018

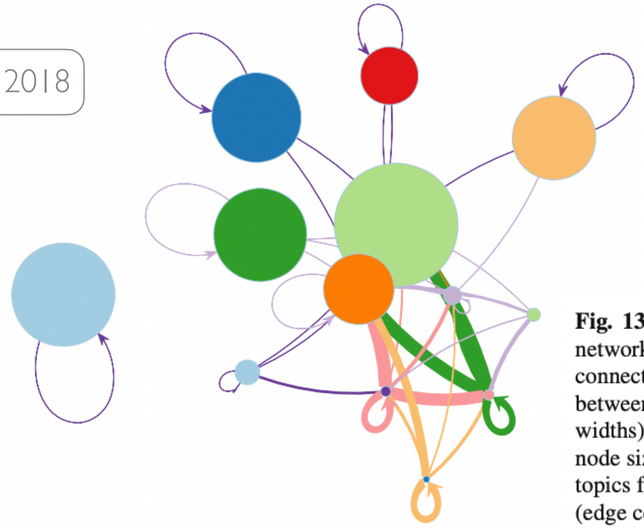


Fig. 13 Nips co-authorship network: summary of connection probabilities between groups (π, edge widths), group proportions (ρ, node sizes), and most probable topics for group interactions (edge colors)

configuration of socio-semantic clusters

overlaps and lacks thereof

Yang, 2011

"Community detection:
Topological vs. topical"

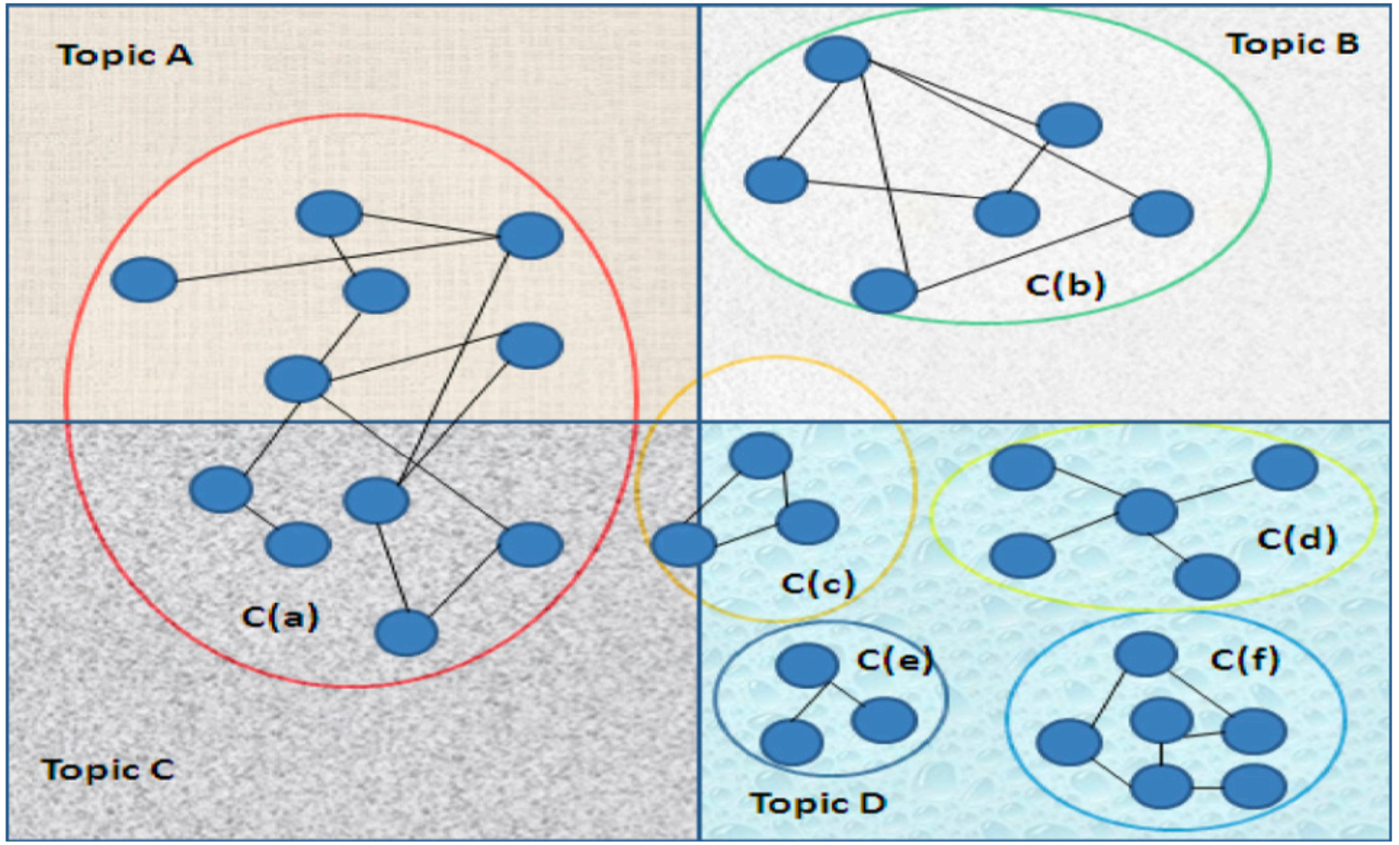
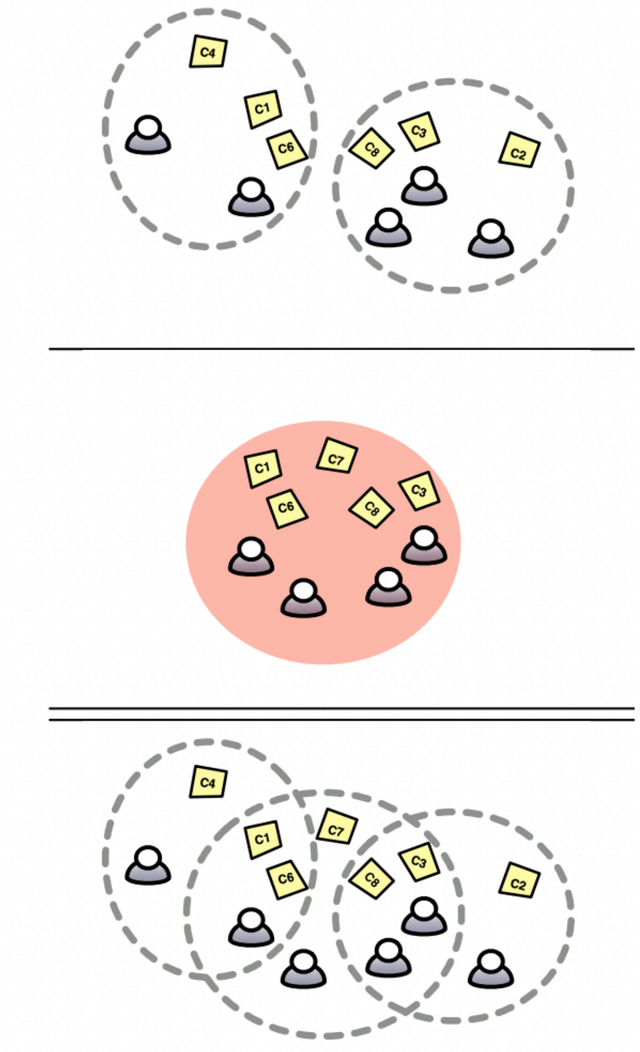


Fig. 1. Communities and topics.

socio-semantic categories

meso-level: hyperedges and hypergraphs



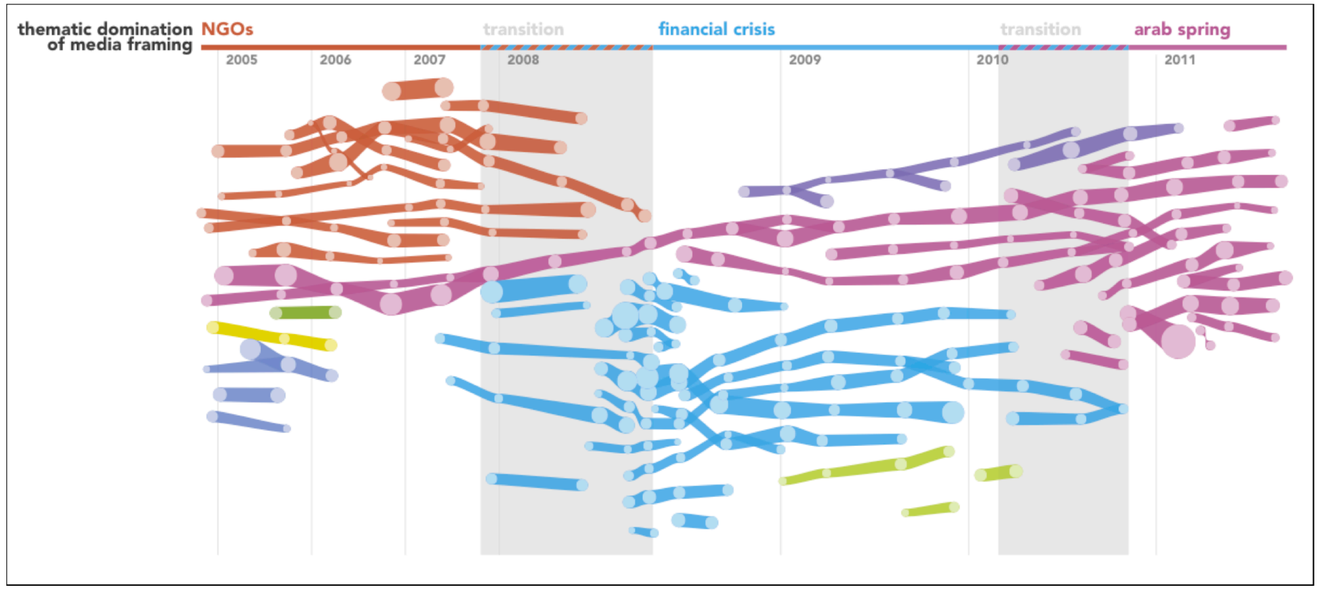
socio-semantic teams

- knowledge production teams
  meso-level: team formation not a sum of individual rationalities
- hypergraphs are a natural modeling framework
  socio-semantic hypergraphs are not reducible to social hypergraphs or bipartite graphs
- introduced a socio-semantic hypergraph model
  estimating *hypergraphic preferential attachment*
- strong socio-semantic preference for:
  - groups with very high or low proportion of 'experts'
  - repetition
    yet no relationship
    between semantic and social originality

Taramasco et al, 2010

socio-semantic categories

macro-level: phylogenies & lattices



Chavalarias, Cointet, Cornilleau, Duong, Mogoutov, Roth, Savy, Villard 2011

socio-semantic cluster dynamics

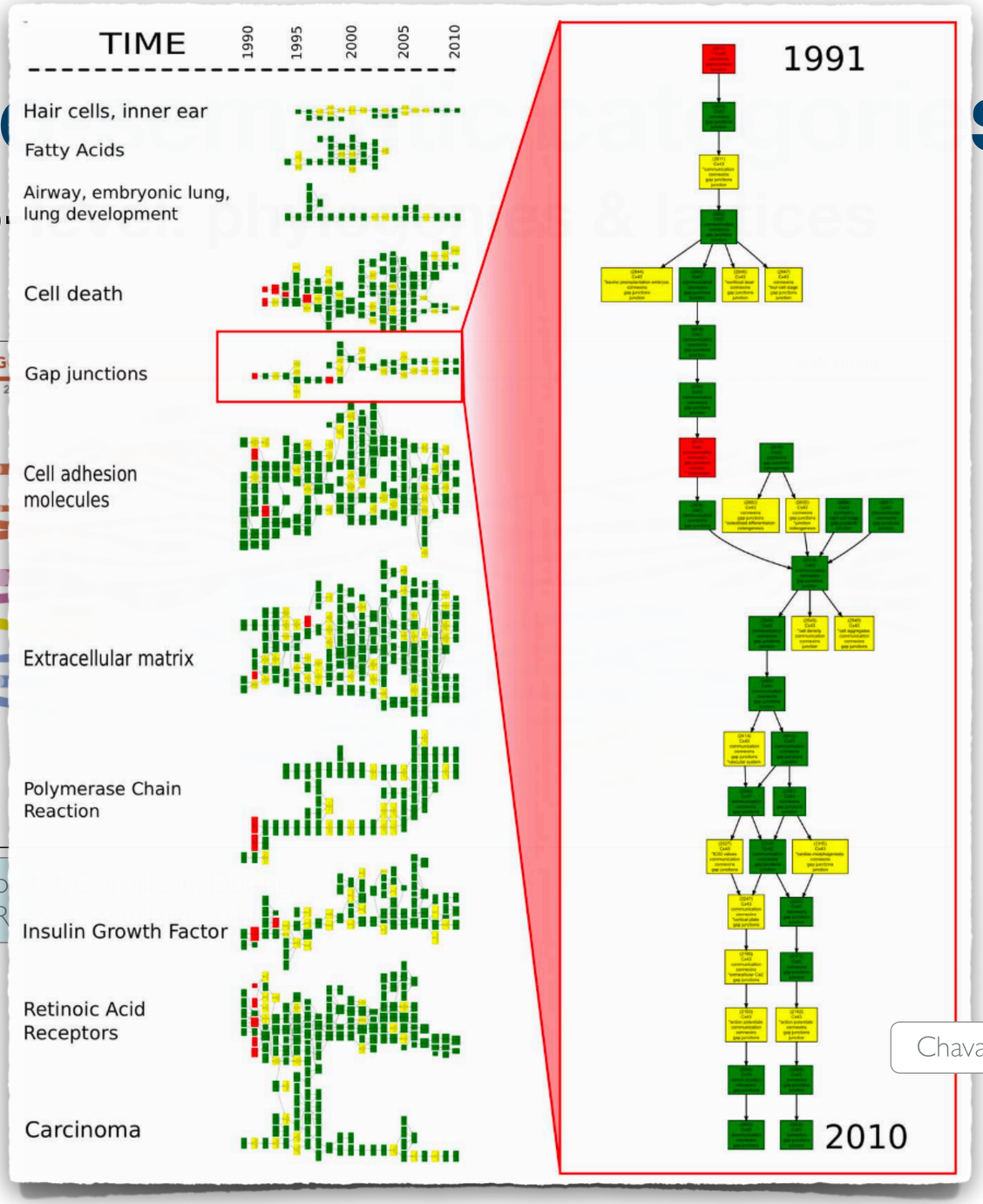
- clusters dynamics
  network snapshot matching
  clusters in temporal networks
  stream graphs and link streams

- isomorphic for social and semantic clusters
  yet characterizing phylogenies of intrinsically socio-semantic structures remains a rather open field

socio

macro

Chavalarias, Co
Mogoutov, R



Chavalarias & Cointet, 2013

socio-semantic cluster dynamics

- clusters dynamics
  network snapshot matching
  clusters in temporal networks
  stream graphs and link streams

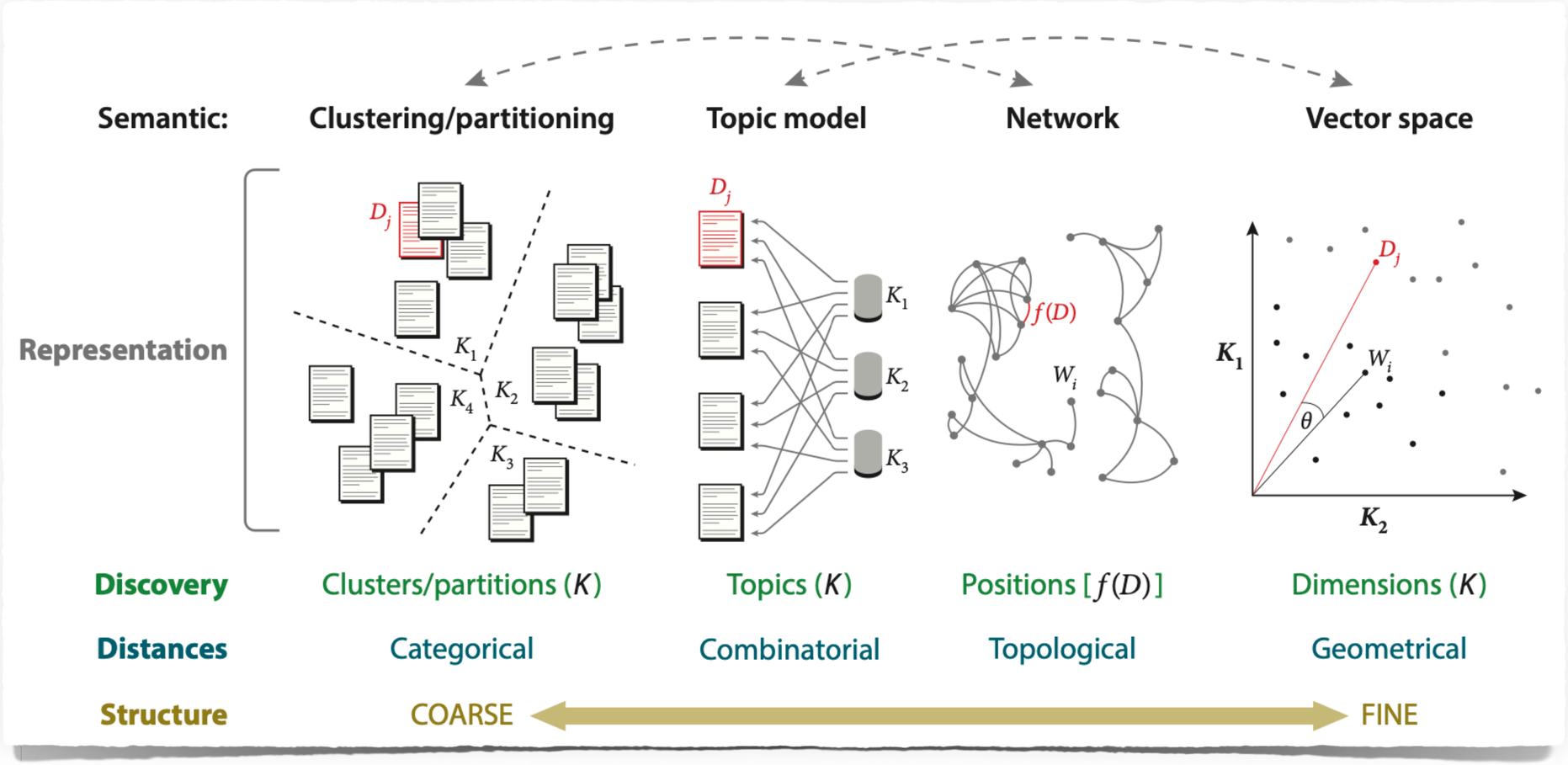
- isomorphic for social and semantic clusters
  yet characterizing phylogenies of intrinsically socio-semantic structures remains a rather open field

# of clouds and bags

a.k.a. what social scientists are often working with

clouds and bags

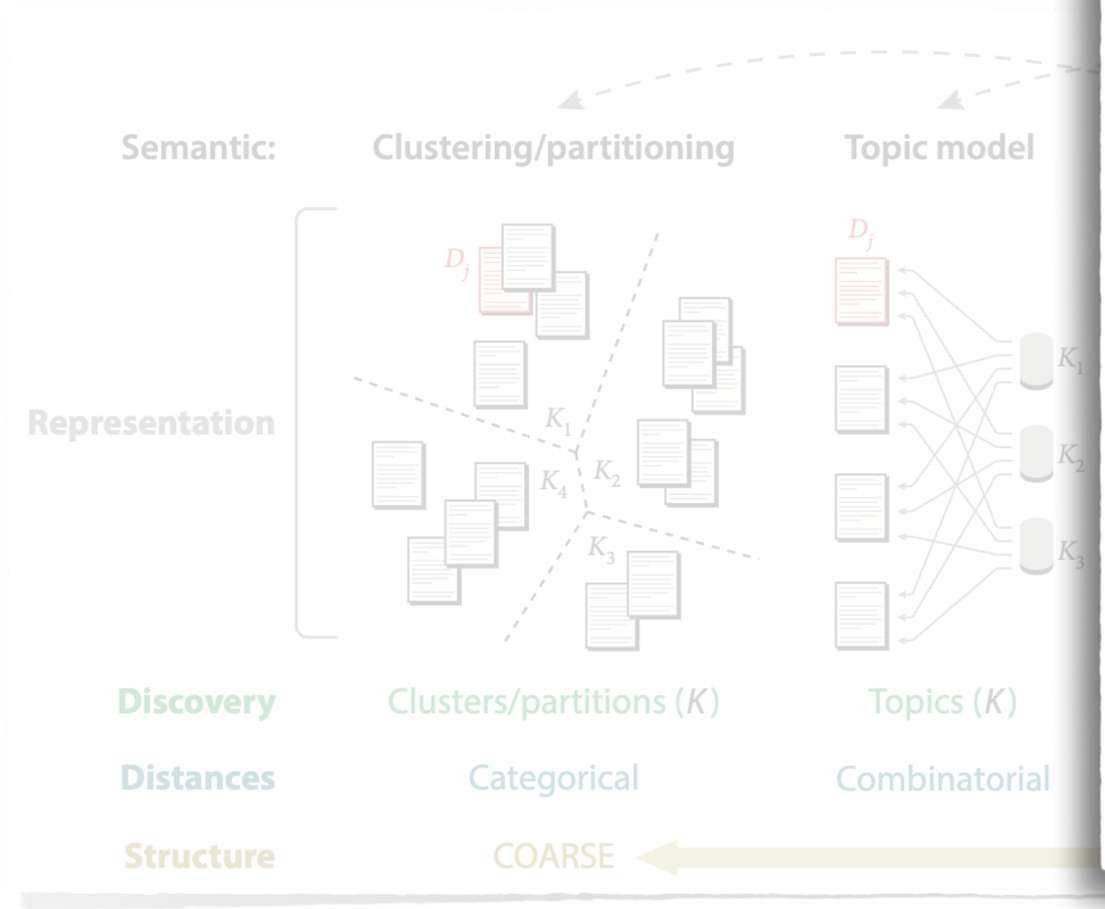
the big four times two

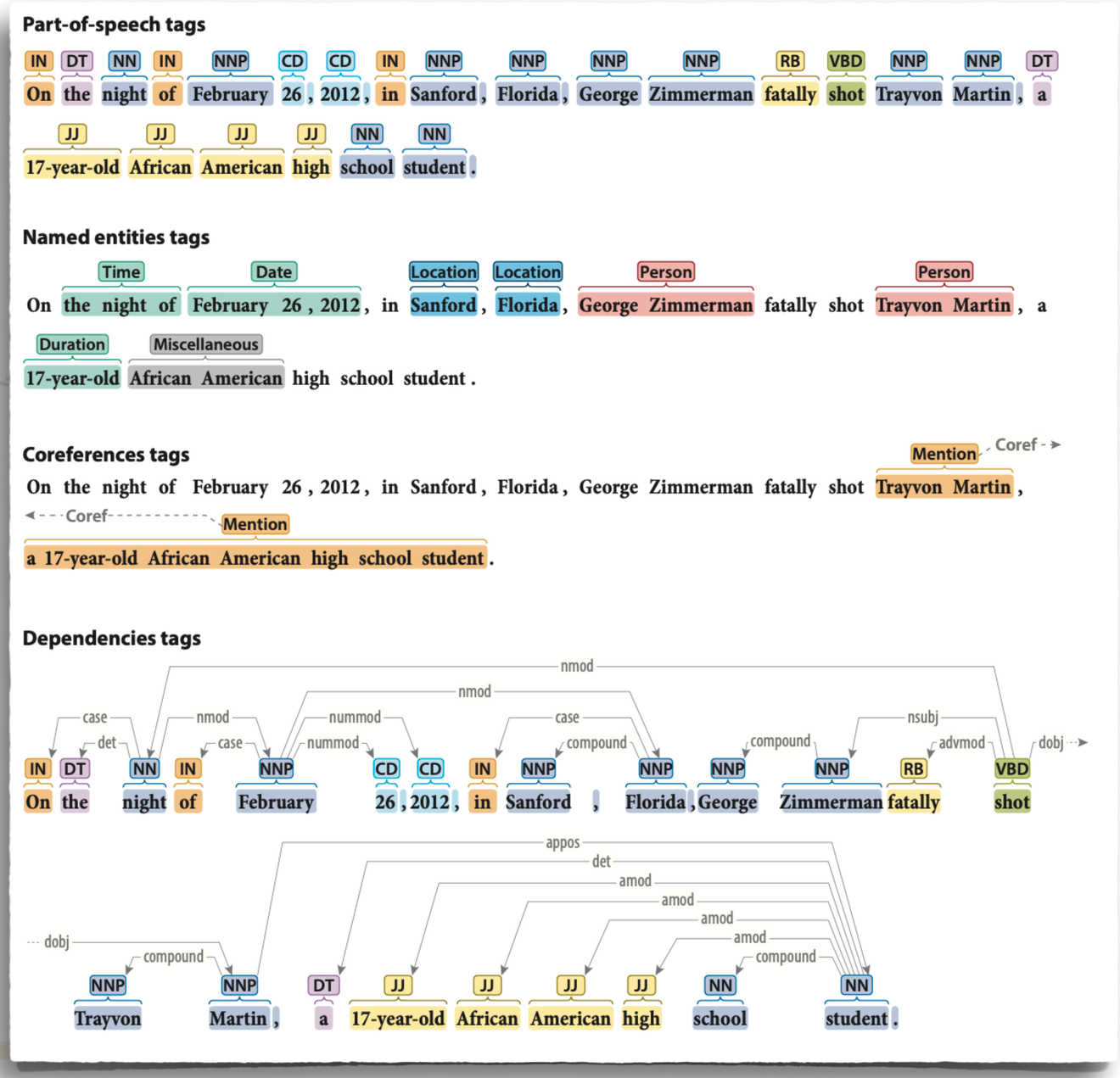


Evans and Aceves, 2016

clouds and bags

the big four times two

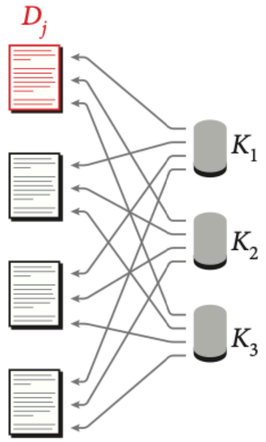




Evans and Aceves, 2016

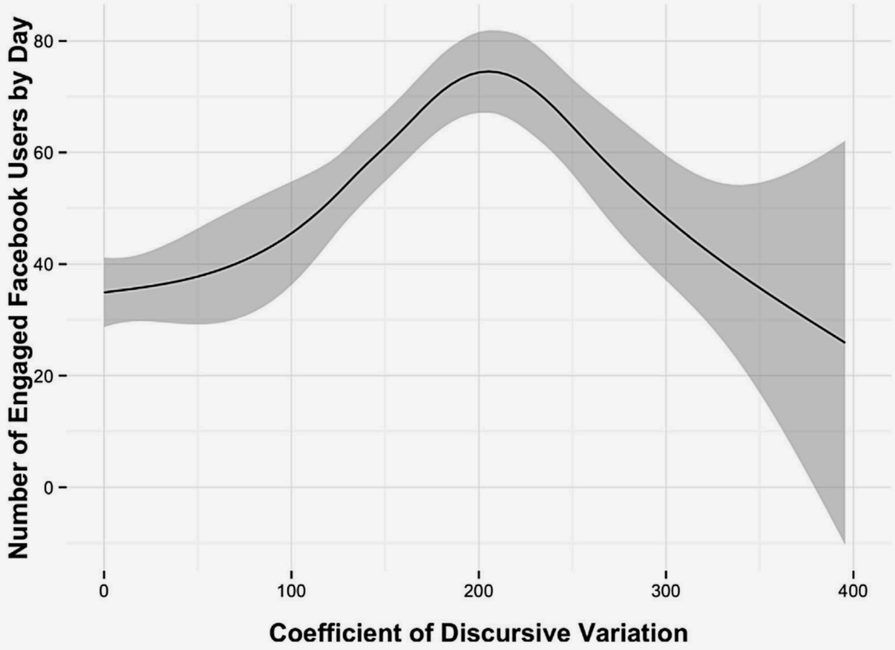
clouds and bags

topic models



Topics (K)

Combinatorial



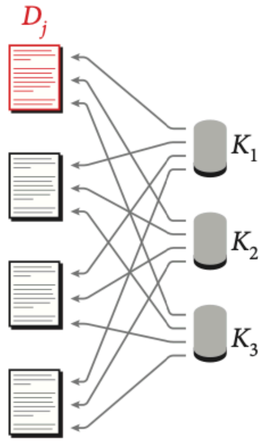
Bail 2016

"Cultural carrying capacity: Organ donation advocacy, discursive framing, and social media engagement"

- posts from 42 organ donor organizations observed over 18 months,
- structural topic models,
- then diversity

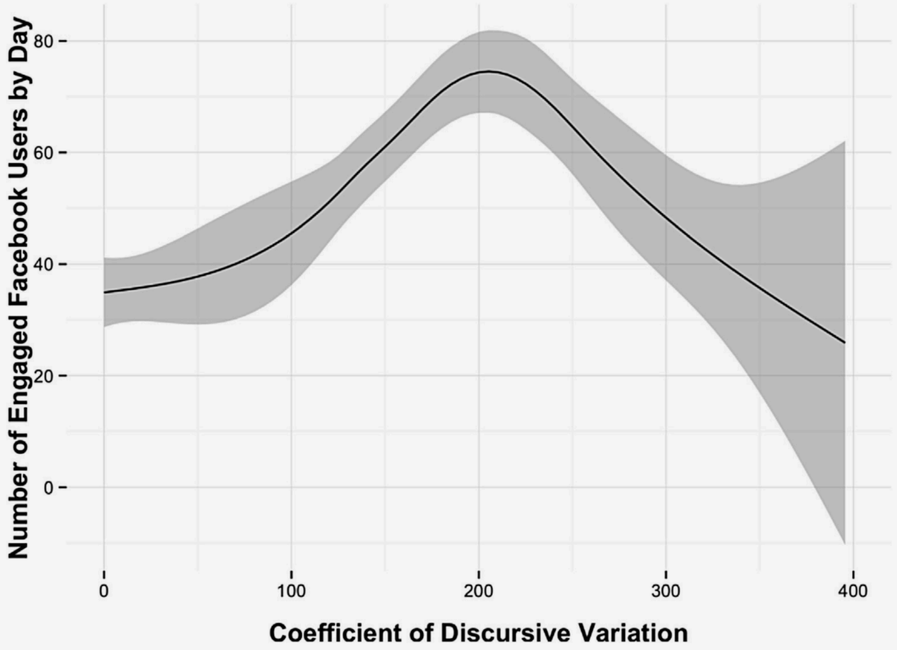
clouds and bags

topic models



Topics (K)

Combinatorial

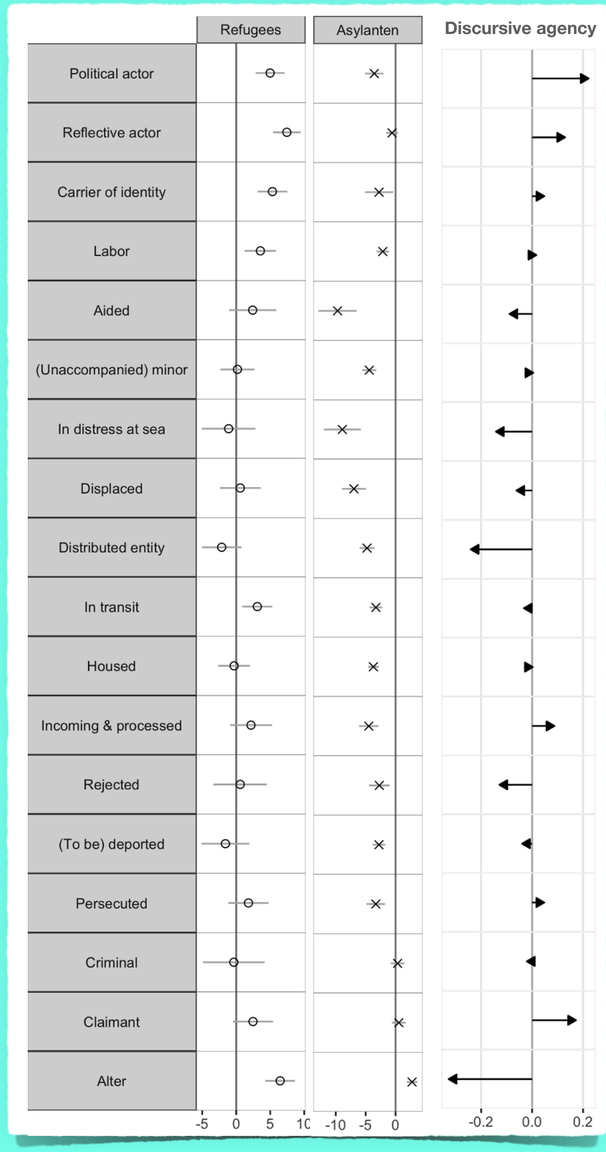


Bail 2016

"Cultural carrying capacity: Organ donation advocacy, discursive framing, and social media engagement"

- posts from 42 organ donor organizations observed over 18 months,
- structural topic models,
- then diversity

Motif class	Frequency	Example sentence	Motif
Agent	23 290	Berlin welcomes refugees.	A_Berlin
Treatment	89 837	Berlin welcomes refugees.	t_welcome
Action	131 804	Refugees enter Germany.	a_enter
Patient	20 161	Refugees enter Germany.	P_Germany
Characterization	107 280	Syrian refugees are welcome.	be_Syrian be_welcome
Attribution	54 495	Refugees' motivation was high.	N_motivation



Stuhler 2021

"What's in a category? A new approach to Discourse Role Analysis"

« German news coverage of refugees (2010-2020),

which employs a set of distinct discourse roles such as

***refugee as claimant of welfare benefits*,**

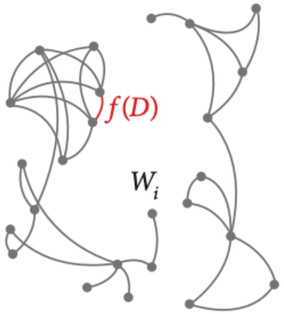
refugee in distress at sea,

and

***refugee as a criminal.* »**

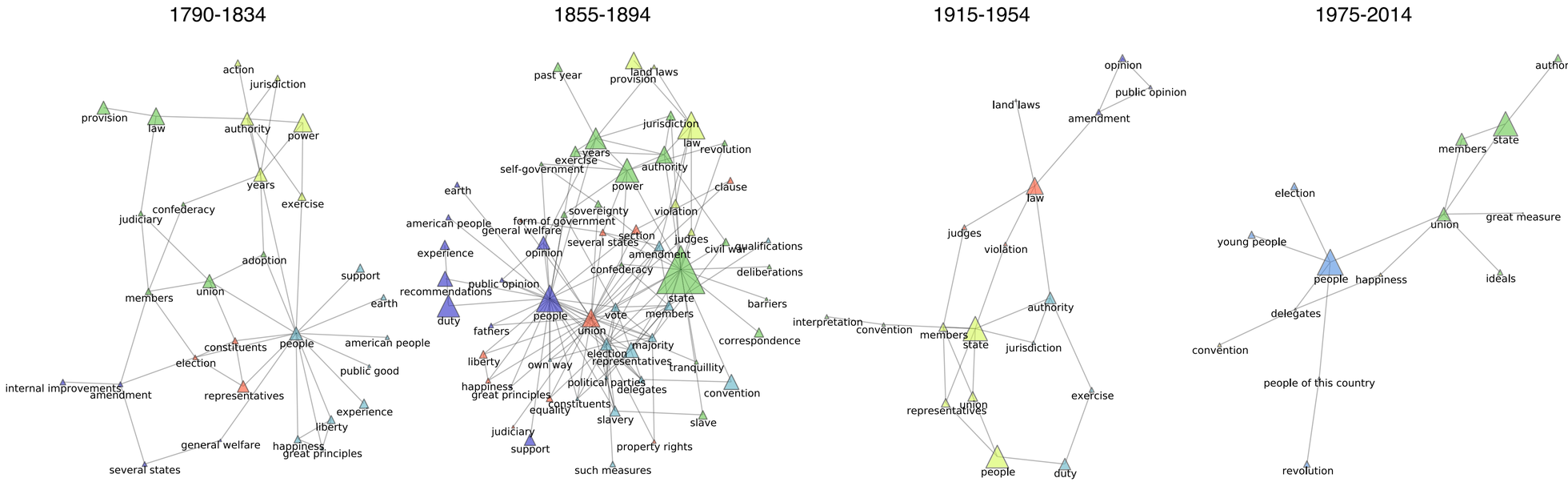
clouds and bags

networks and clusters



Positions $[f(D)]$

Topological

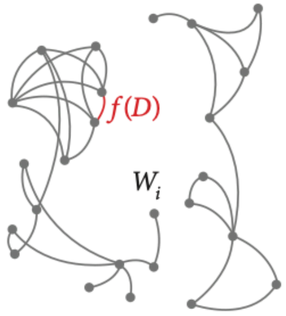


Rule, Cointet, Bearman, 2015

"Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014"

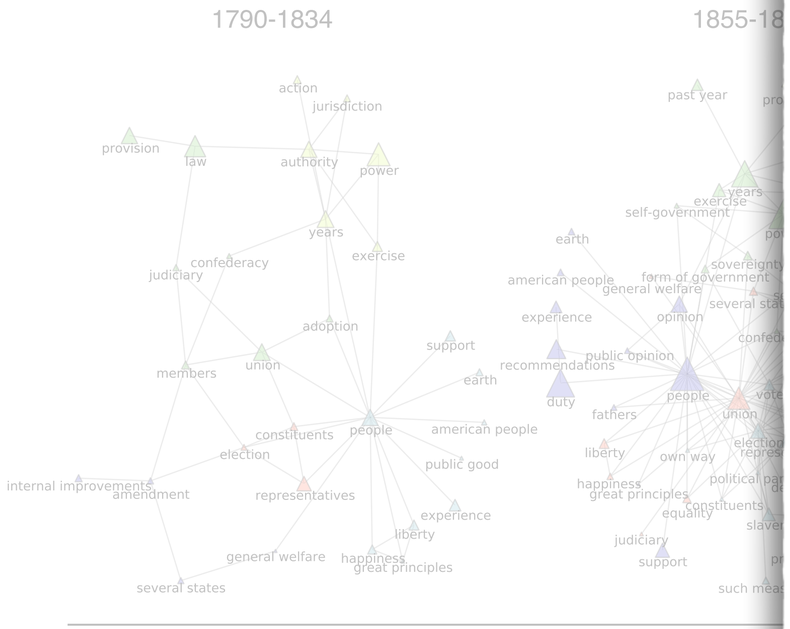
clouds and bags

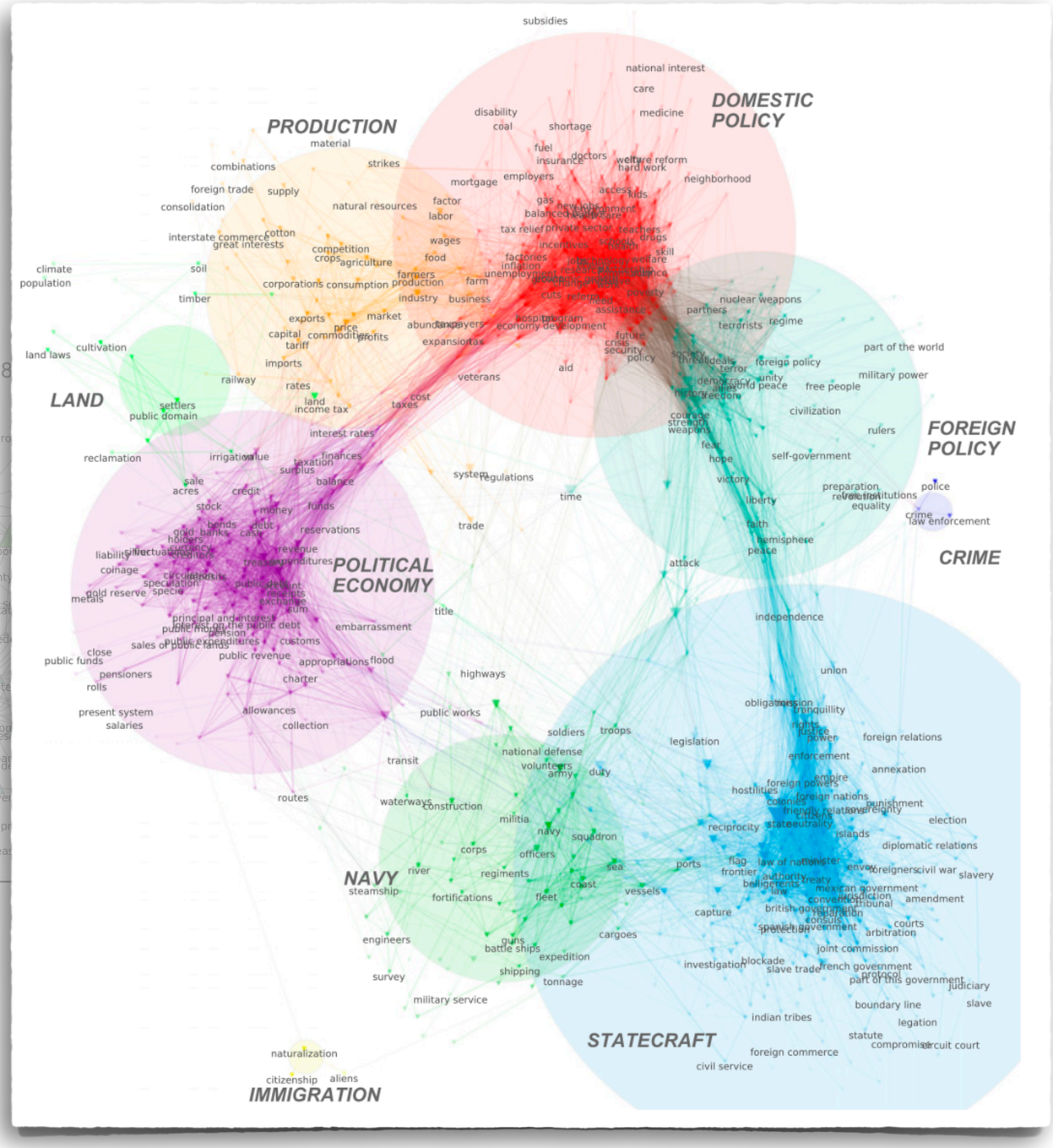
networks and clusters



Positions $[f(D)]$

Topological



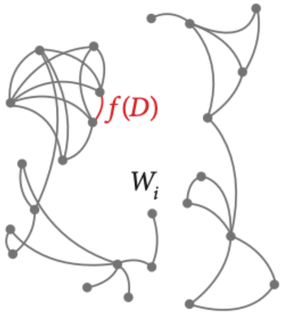


Rule, Cointet, Bearman, 2015

"Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014"

clouds, bags and tags

networks and clusters



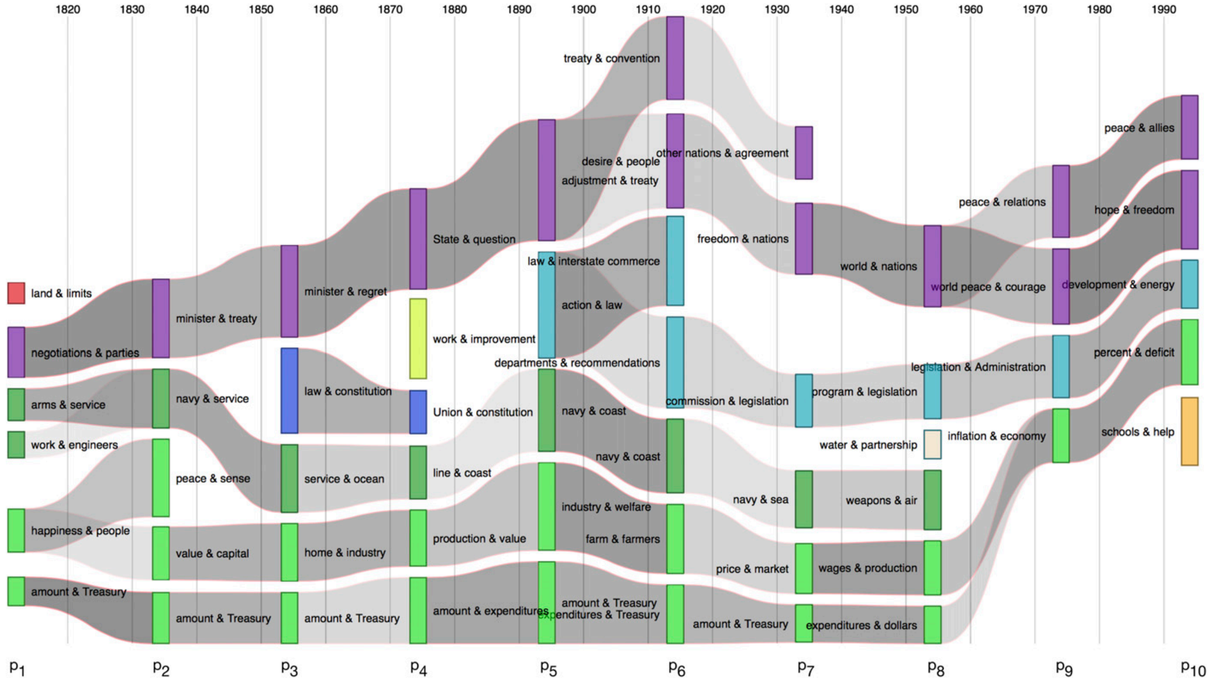
Positions $[f(D)]$

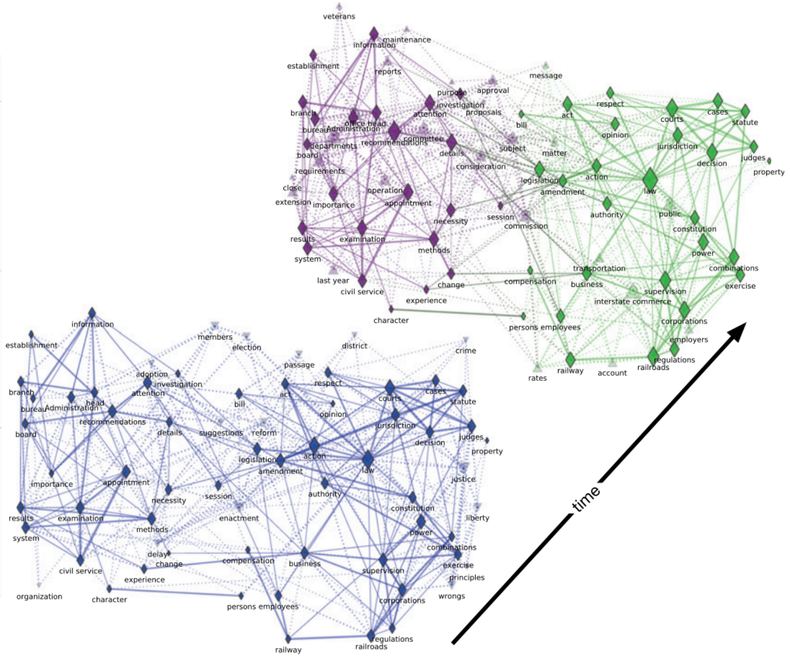
Topological

Rule, Cointet, Bearman, 2015

"Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014"

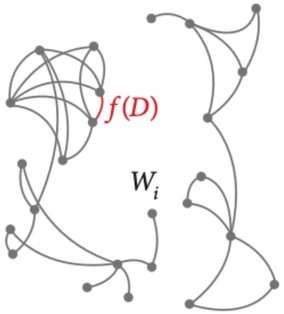
	action & law (1875-1914)	departments and recommendations (1895-1934)	law & interstate commerce (1895-1934)
action & law (1875-1914)	liberty; justice; passage; adoption; principles; election; reform; crime; members; enactment; district; wrongs; suggestions; delay; organization	results; session; board; appointment; details; recommendations; methods; bureau; importance; attention; civil service; experience; character; Administration; system; examination; change; head; investigation; branch; establishment; necessity; information	jurisdiction; judges; railroads; decision; exercise; amendment; bill; authority; power; business; act; action; respect; legislation; corporations; courts; compensation; property; employees; regulations; persons; law; cases; railway; constitution; statute; supervision; combinations; opinion
departments and recommendations (1895-1934)	results; session; board; appointment; details; recommendations; methods; bureau; importance; attention; civil service; experience; character; Administration; system; examination; change; head; investigation; branch; establishment; necessity; information	office; committee; approval; close; last year; operation; subject; requirements; departments; commission; maintenance; veterans; purpose; consideration; extension; proposals; reports	
law & interstate commerce (1895-1934)	jurisdiction; judges; railroads; decision; exercise; amendment; bill; authority; power; business; act; action; respect; legislation; corporations; courts; compensation; property; employees; regulations; persons; law; cases; railway; constitution; statute; supervision; combinations; opinion		transportation; interstate commerce; rates; public; employers; message; account; matter





clouds, bags and tags

networks and interactions



Positions [$f(D)$]

Topological

Conover, Ratkiewicz, Francisco, Gonçalves, Flammini, Menczer, 2011

"Political Polarization on Twitter"

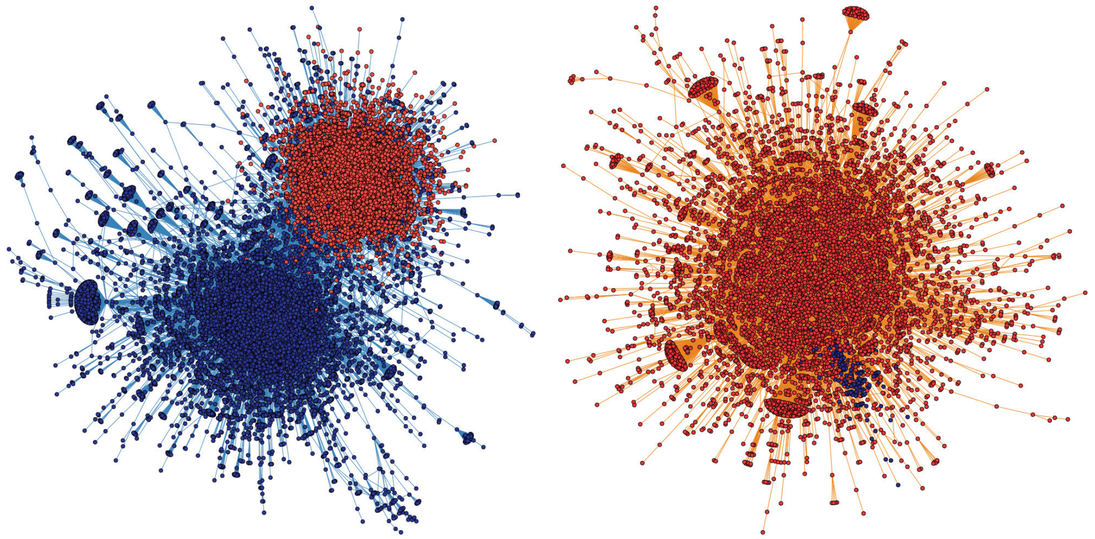


Figure 1: The political retweet (left) and mention (right) networks, laid out using a force-directed algorithm. Node colors reflect cluster assignments (see § 3.1). Community structure is evident in the retweet network, but less so in the mention network. We show in § 3.3 that in the retweet network, the red cluster A is made of 93% right-leaning users, while the blue cluster B is made of 80% left-leaning users.

Table 5: Ratios between observed and expected number of links between users of different political alignments in the mention and retweet networks.

	Mention		Retweet	
	→ Left	→ Right	→ Left	→ Right
Left	1.23	0.68	1.70	0.05
Right	0.77	1.31	0.03	2.32

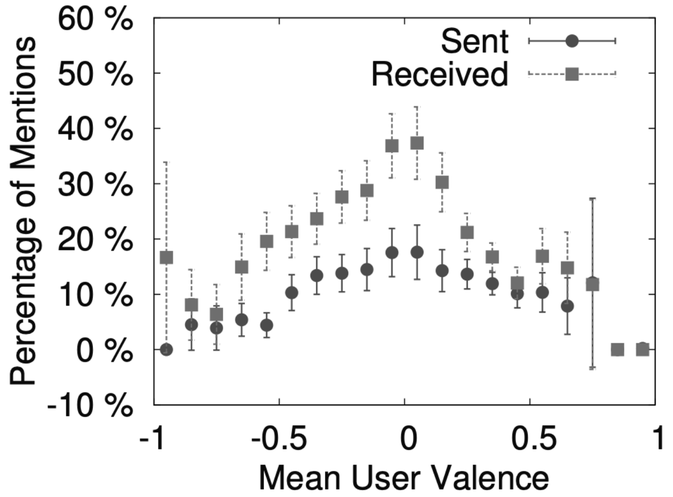
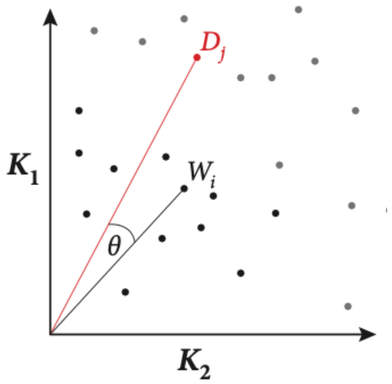


Figure 3: Proportion of mentions a user sends and receives to and from ideologically-opposed users relative to her valence. Points represent binned averages. Error bars denote 95% confidence intervals.

clouds and bags

embeddings

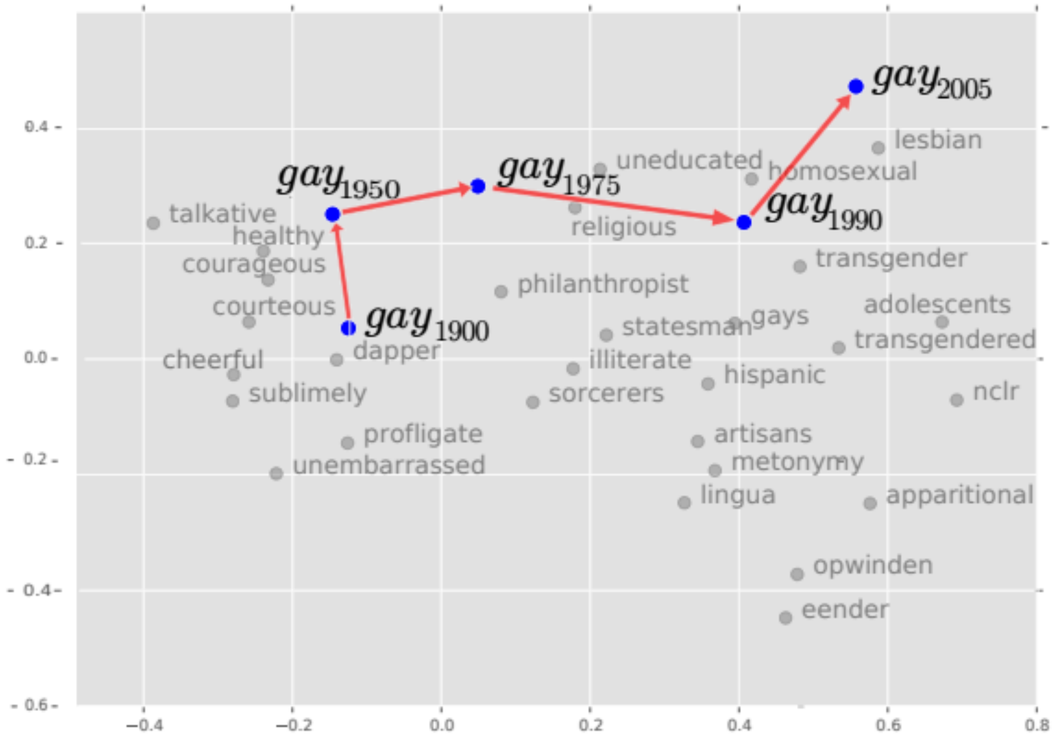


Dimensions (K)

Geometrical

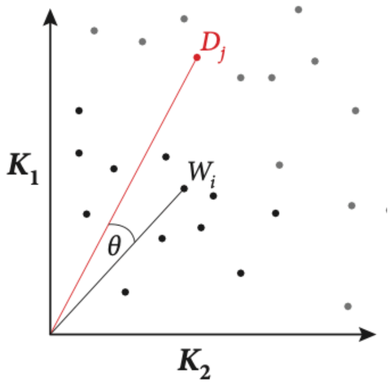
Kulkarni, Perozzi, Al-Rfou, Skiena, 2015

"Statistically Significant Detection of Linguistic Change"



clouds and bags

embeddings

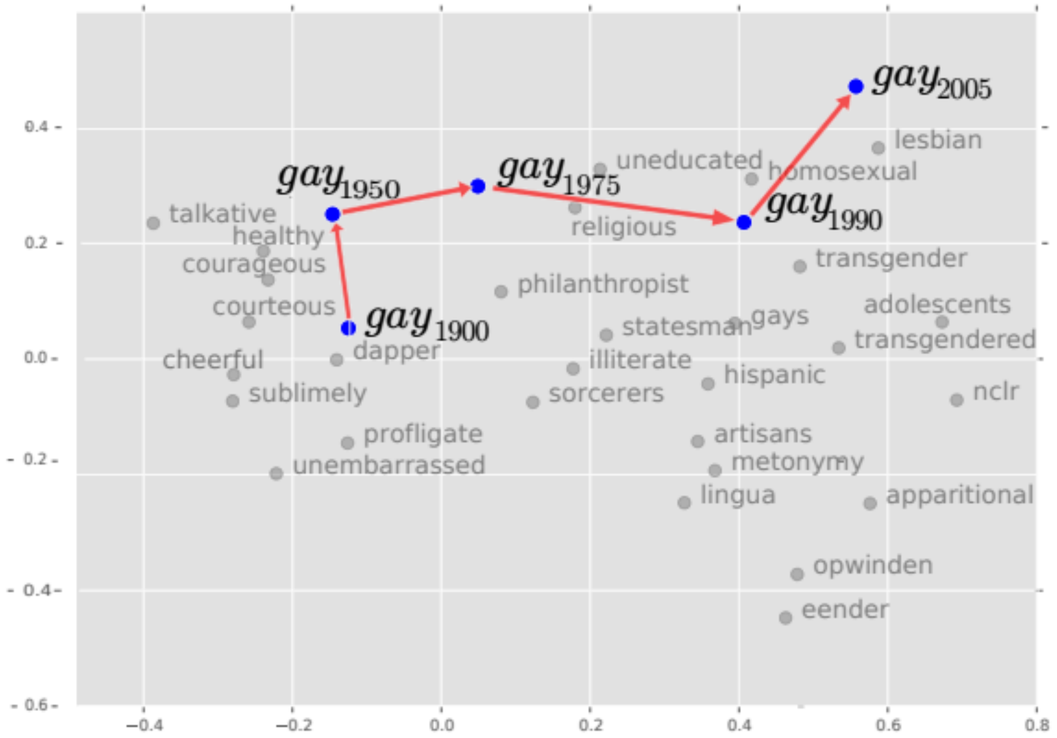


Dimensions (K)

Geometrical

Kulkarni, Perozzi, Al-Rfou, Skiena, 2015

"Statistically Significant Detection of Linguistic Change"



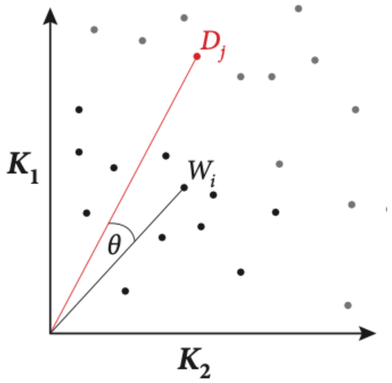
Kutuzov et al, 2018

"Diachronic word embeddings and semantic shifts: a survey"

Further: quest for laws of semantic change e.g., laws of
- prototypicality (semantic shift faster for prototypical words)
- conformity (semantic shift slower for more frequent words)
- innovation (semantic shift faster for polysemous words)

clouds and bags

embeddings

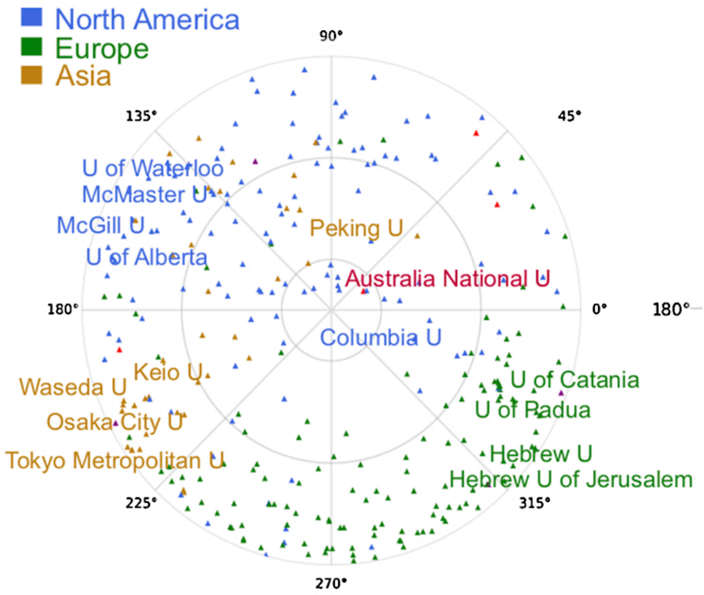


Dimensions (K)

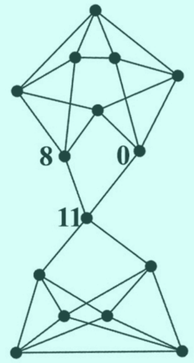
Geometrical

Li, Wu, Evans, 2020

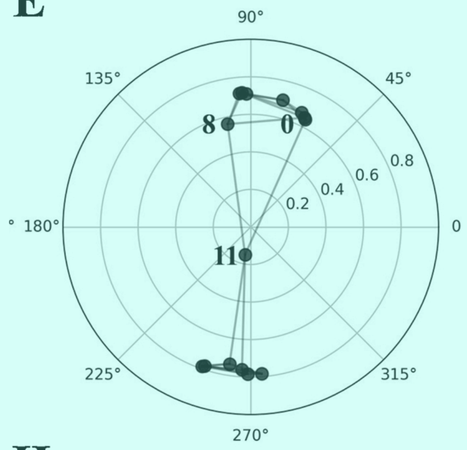
"Social centralization and semantic collapse: Hyperbolic embeddings of networks and text"



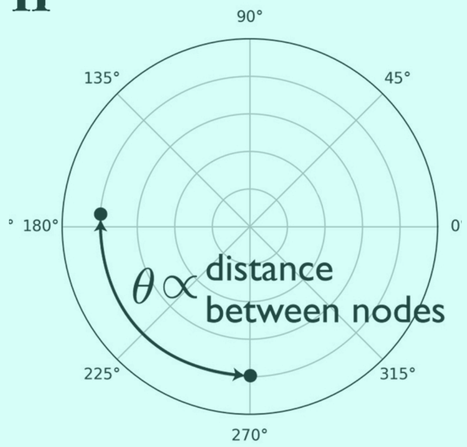
B



E

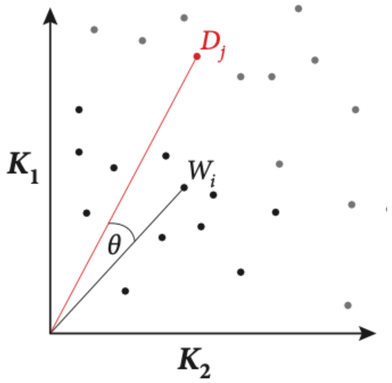


H



clouds and bags

embeddings

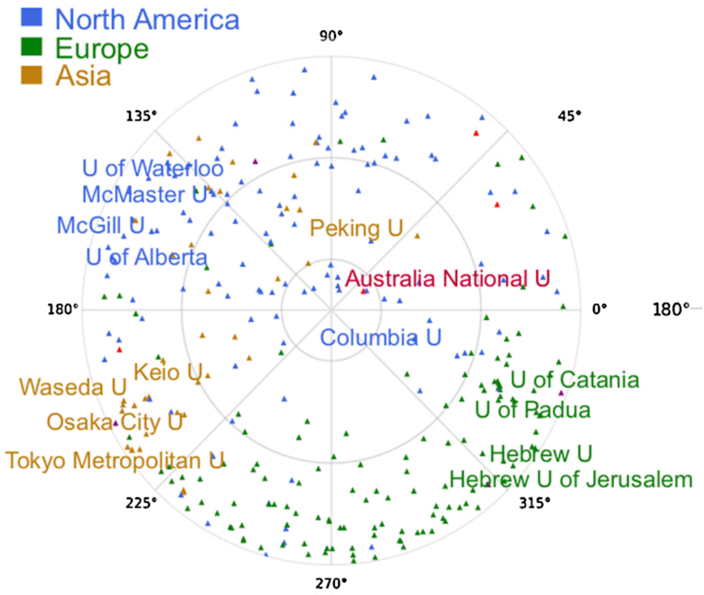


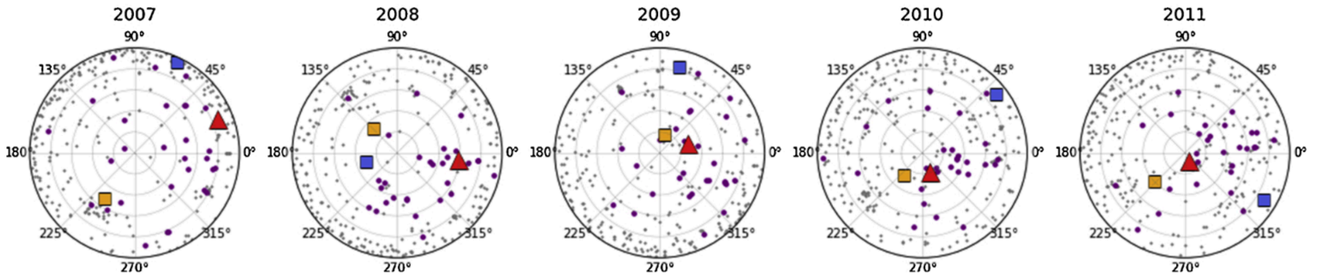
Dimensions (K)

Geometrical

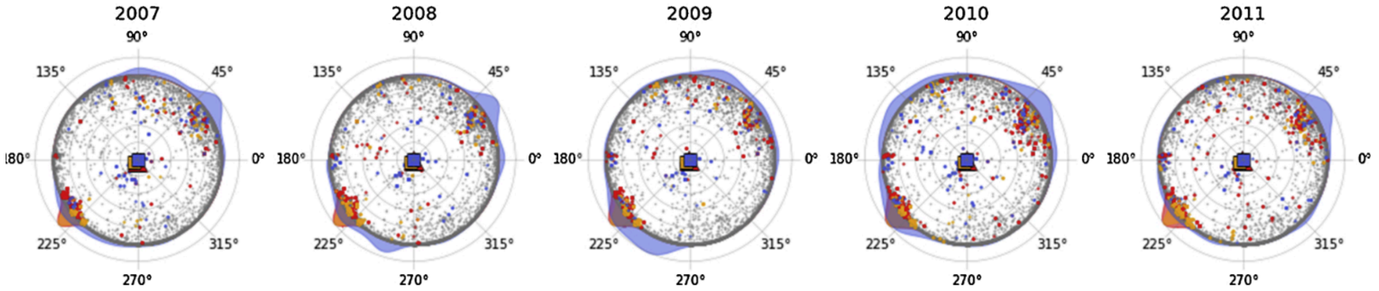
Li, Wu, Evans, 2020

"Social centralization and semantic collapse: Hyperbolic embeddings of networks and text"

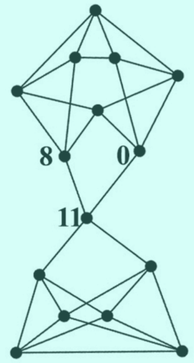




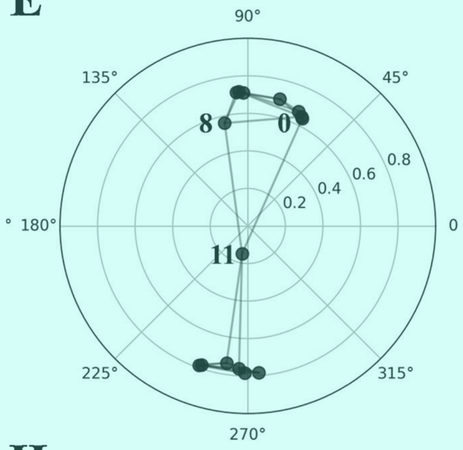
Movement of institutions in **collaboration space (above)** and **semantic space (below)**



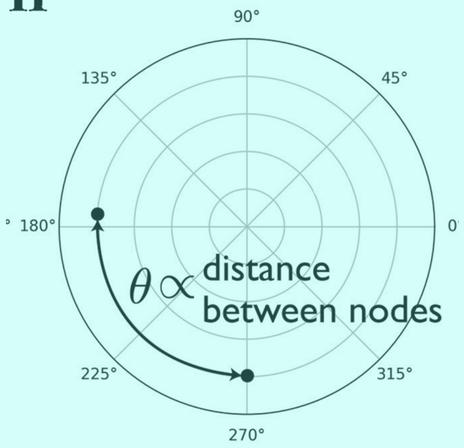
B



E



H



the sentence level

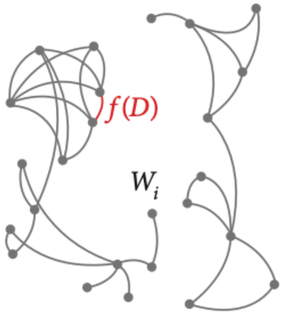
a.k.a. what social scientists are often looking forward to

the sentence level

network categorization

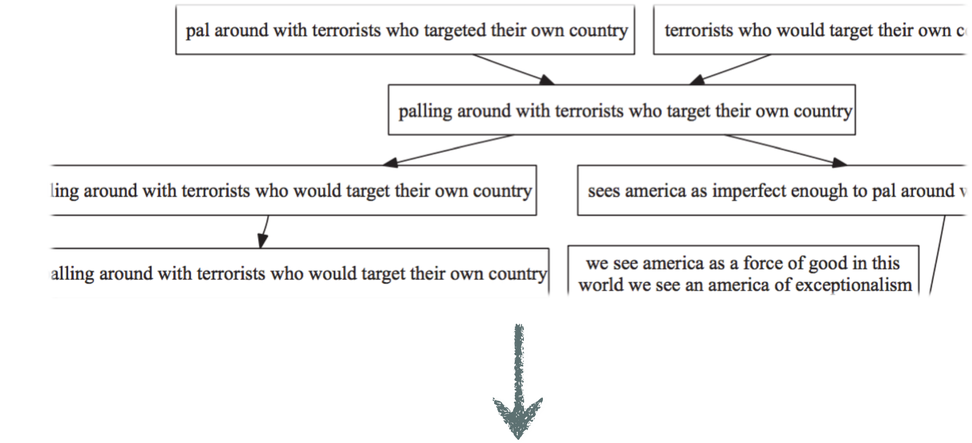
Leskovec, Backstrom, Kleinberg, 2009

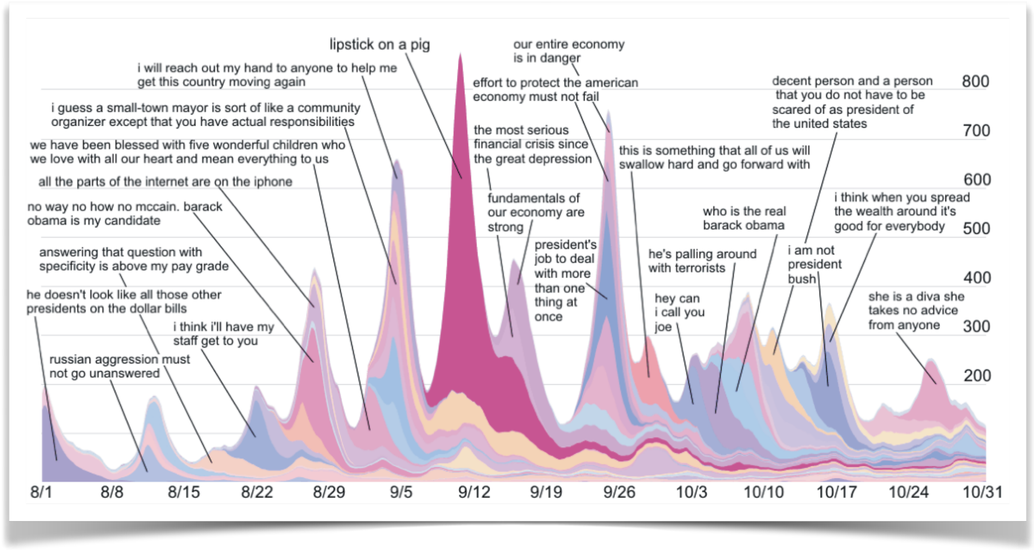
"Meme-tracking and the dynamics of the news cycle"



Positions $[f(D)]$

Topological





the sentence level

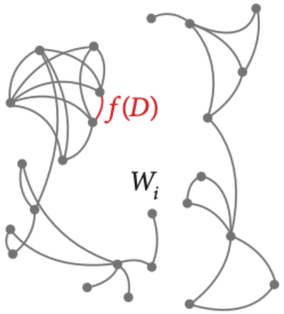
network categorization

Leskovec, Backstrom, Kleinberg, 2009

"Meme-tracking and the dynamics of the news cycle"

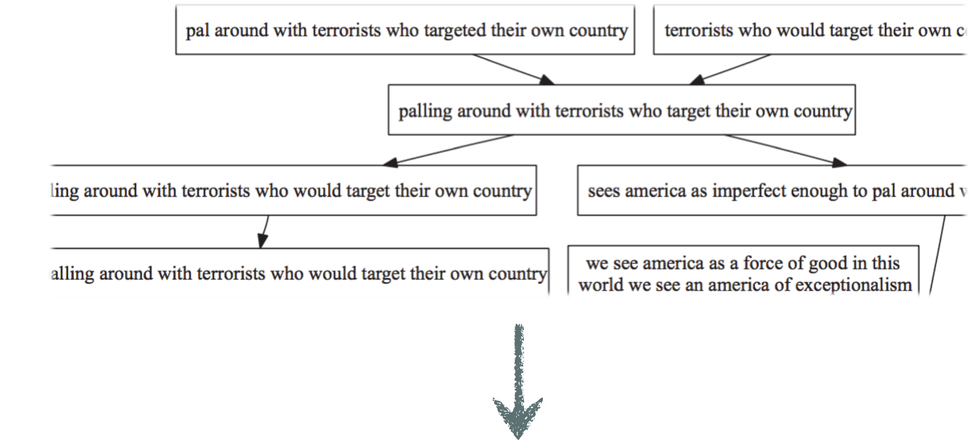
Simmons, Adamic, Adar, 2011

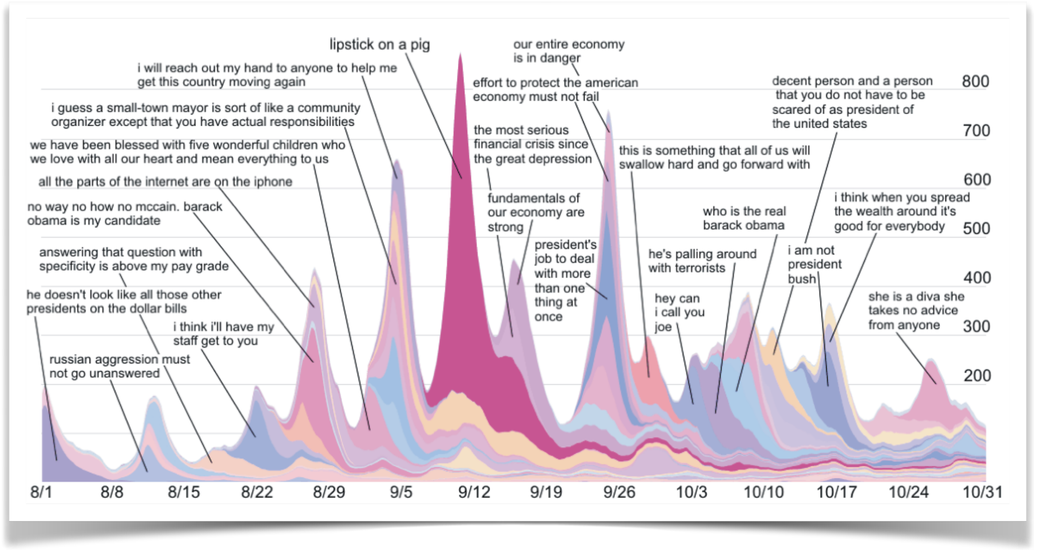
"Memes Online: Extracted, Subtracted, Injected, and Recollected"



Positions $[f(D)]$

Topological





abn.com

"i find that governor sarah palin abused her power by violating alaska statute 39 52 110 a of the alaska executive branch ethics act"

demconwatchdog.com

"i find that governor sarah palin abused her power by violating alaska statute 39"

lzydata.blogspot.com

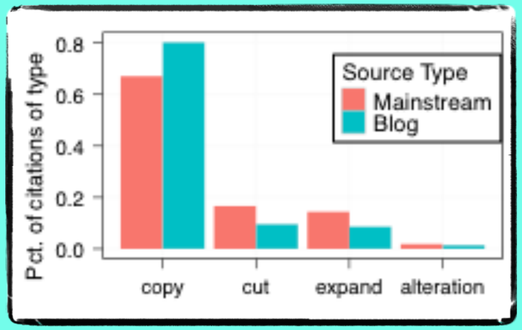
"abused her power"

thenation.com

"i find that governor palin abused her power by violating alaska statute 39 52 110 a of the alaska executive branch ethics act *alaska statute 39 52 110 a provides 'the legislature affirms that each public officer holds office as a public trust…"*

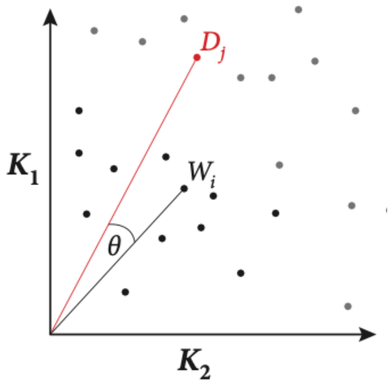
blogs.abcnews.com

"i find that governor sarah palin abused her power by violating alaska statute 39 52 110 a of the alaska *executive branch ethics act compliance with the code of ethics is not optional*"



the sentence level

embedding-based categorization



Dimensions (K)

Geometrical

Angelov, 2020 "Top2Vec: Distributed representation of topics"

Tonneau, Avouac, Roth "Short text Topic-Facet Modeling"

1. Topic modeling

Short text
(e.g. claims, social media posts)

Sentence **embedding**
(SBERT)

Dimensionality **reduction**
(UMAP)

Cluster **detection**
(HDBSCAN)

healthcare

ecology

policy reforms

2. Facet modeling

« We need to invest in green agriculture to curb climate change »

ecology

« Renewable energy can help fight climate change »

Re-applying **topic modeling**

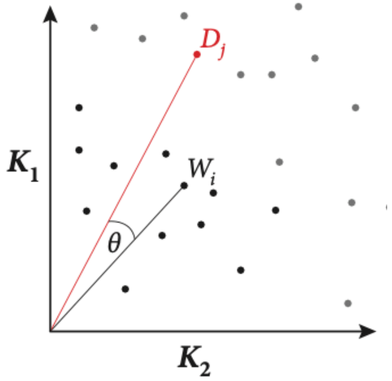
Facet clusters

energy

agriculture

the sentence level

embedding-based categorization



Dimensions (K)

Geometrical

Angelov, 2020 — "Top2Vec: Distributed representation of topics"

Tonneau, Avouac, Roth — "Short text Topic-Facet Modeling"

1. Topic modeling

Short text
(e.g. claims, social media posts)

Sentence **embedding**
(SBERT)

Dimensionality **reduction**
(UMAP)

Cluster **detection**
(HDBSCAN)

healthcare

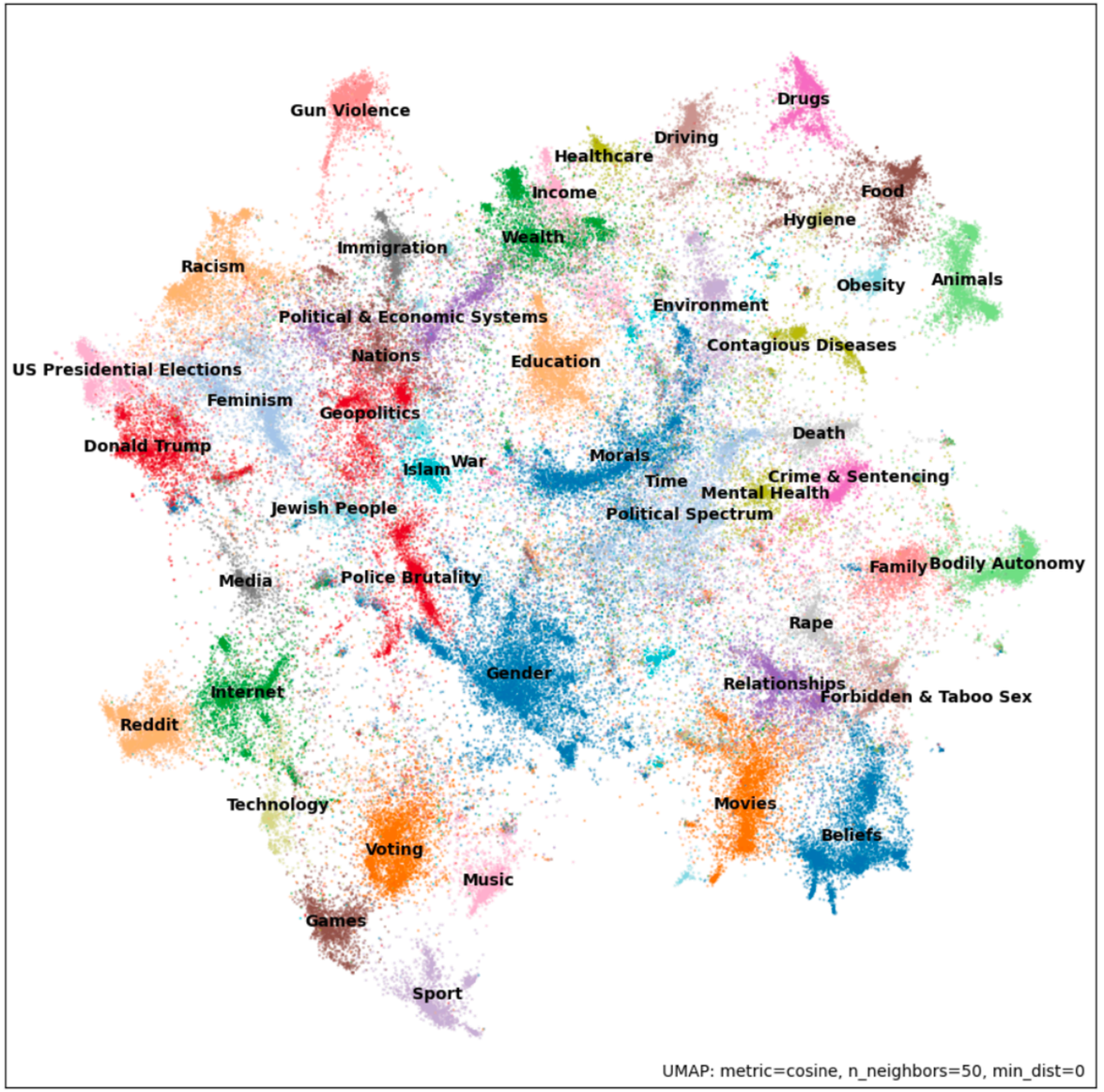
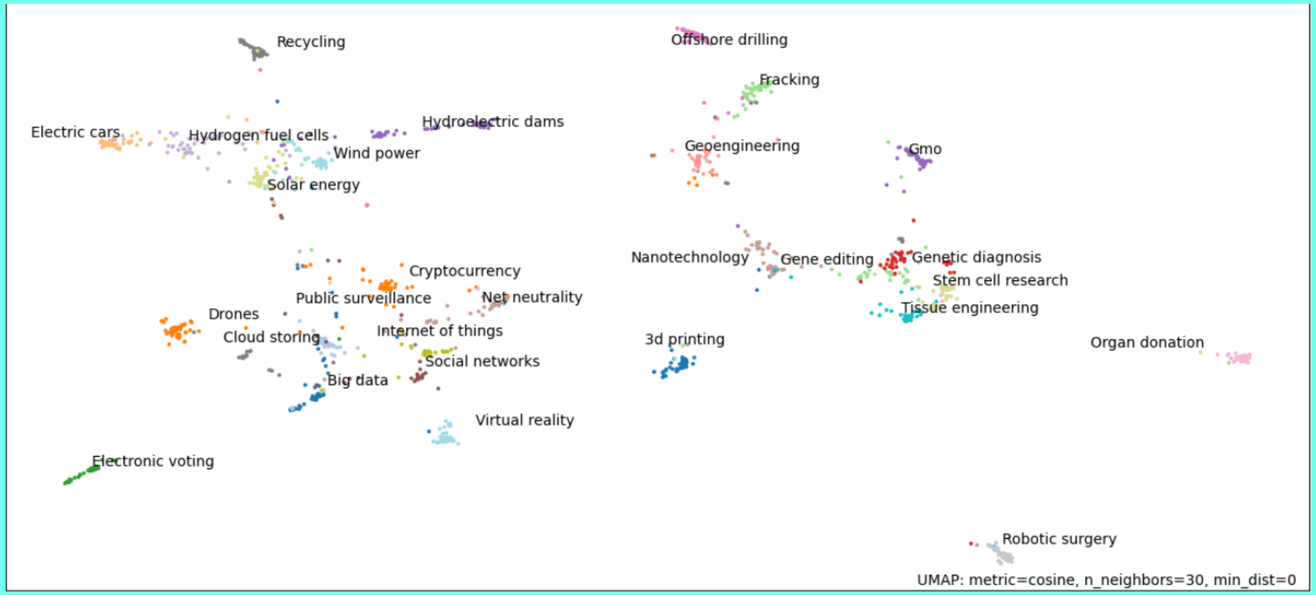


Figure 3: 2D projection of the topics predicted by the Topic2Facet model for the Reddit CMV Corpus



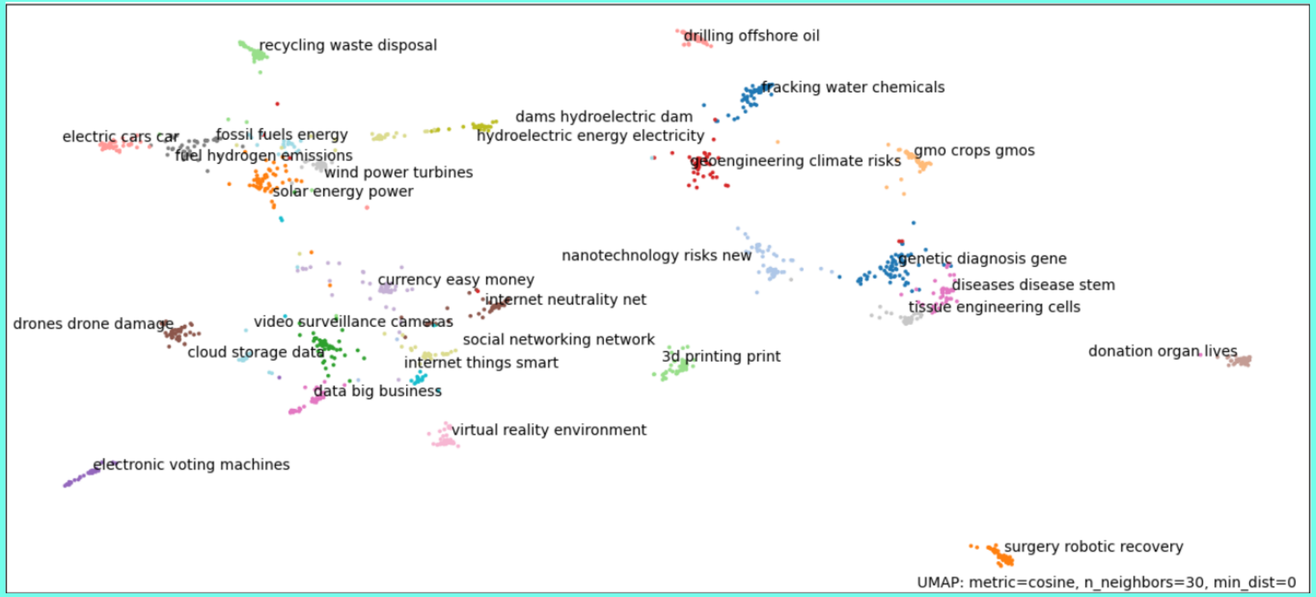


Figure 2: Comparison of gold (top) and predicted (bottom) topics on the UKP ASPECT corpus

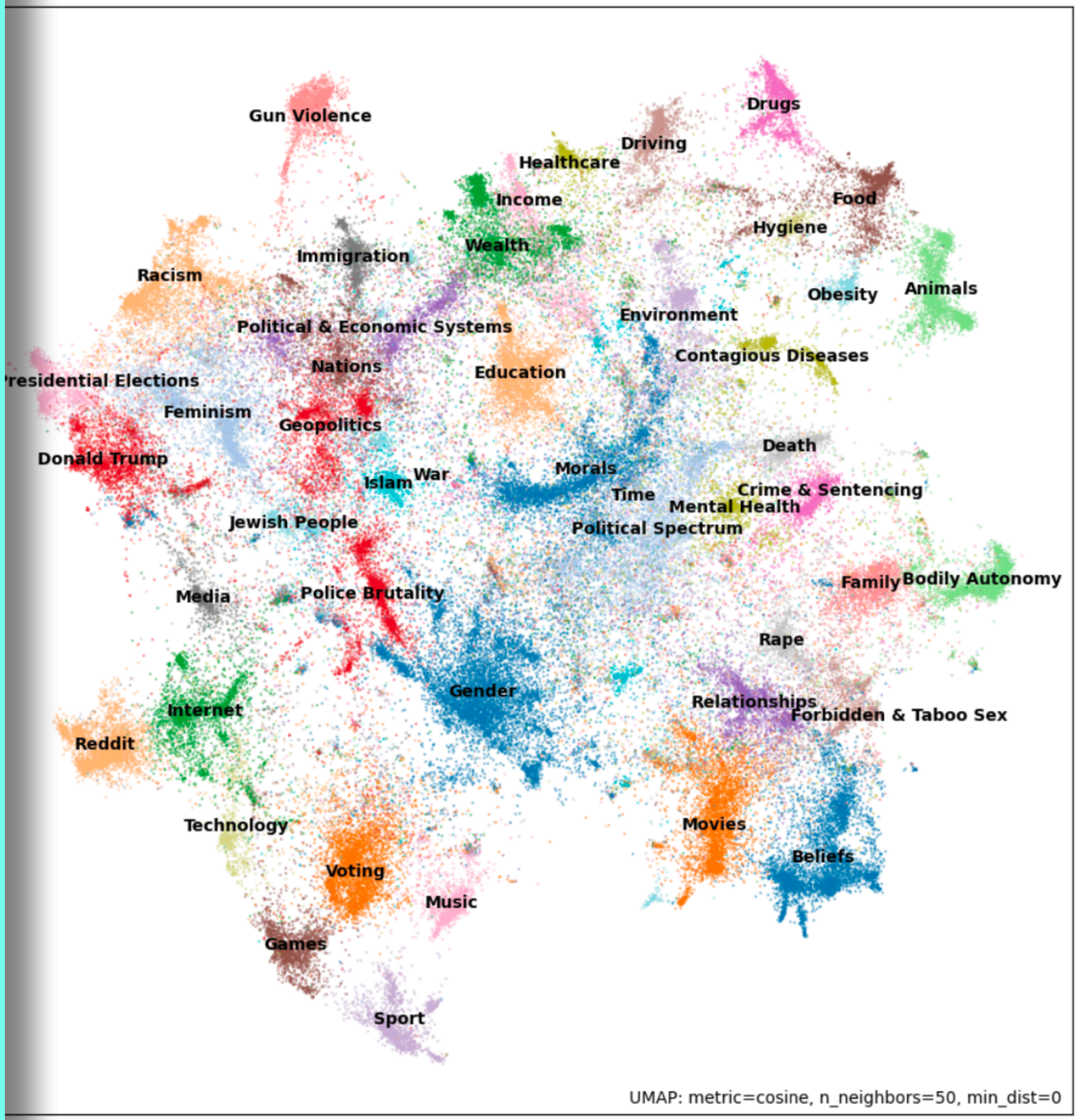
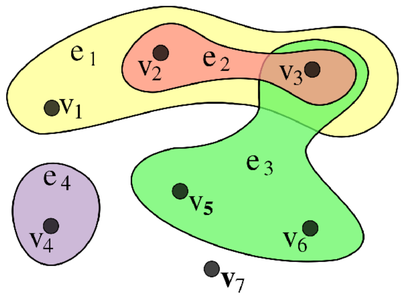


Figure 3: 2D projection of the topics predicted by the Topic2Facet model for the Reddit CMV Corpus

the sentence level

pattern definition and extraction



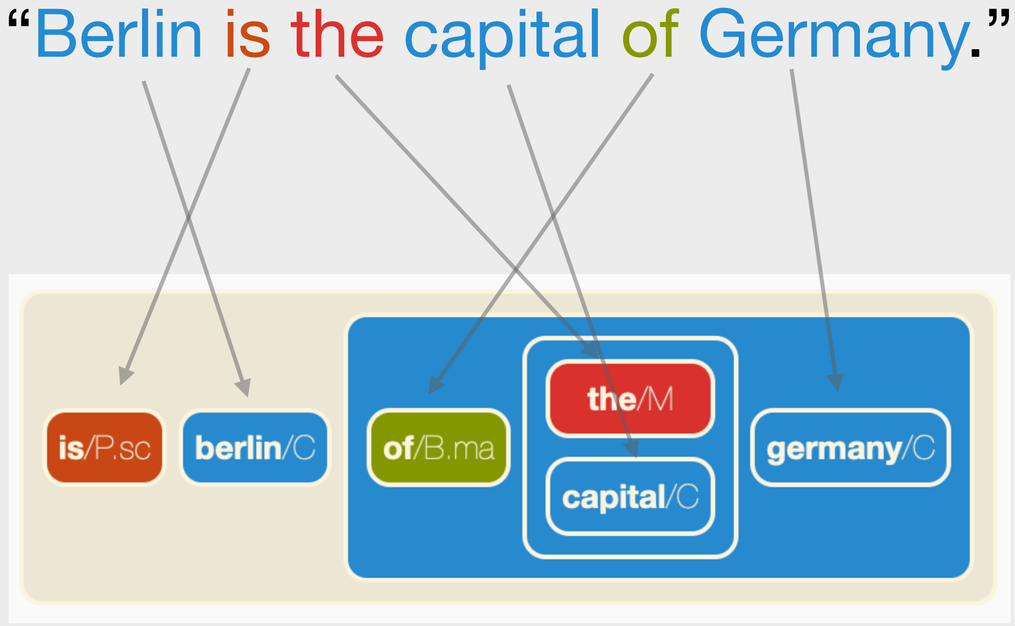
Semantic hypergraphs

Topological

Menezes & Roth, 2019

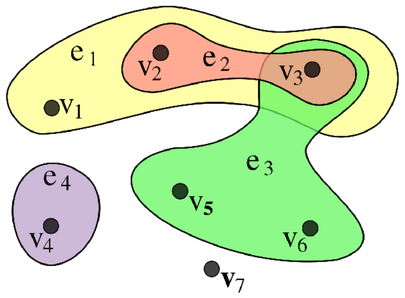
Bonaccorsi, Melluso, Chiarello, Fantoni, 2021a, 2021b

Menezes, Pottier, Roth, 2023



the sentence level

pattern definition and extraction



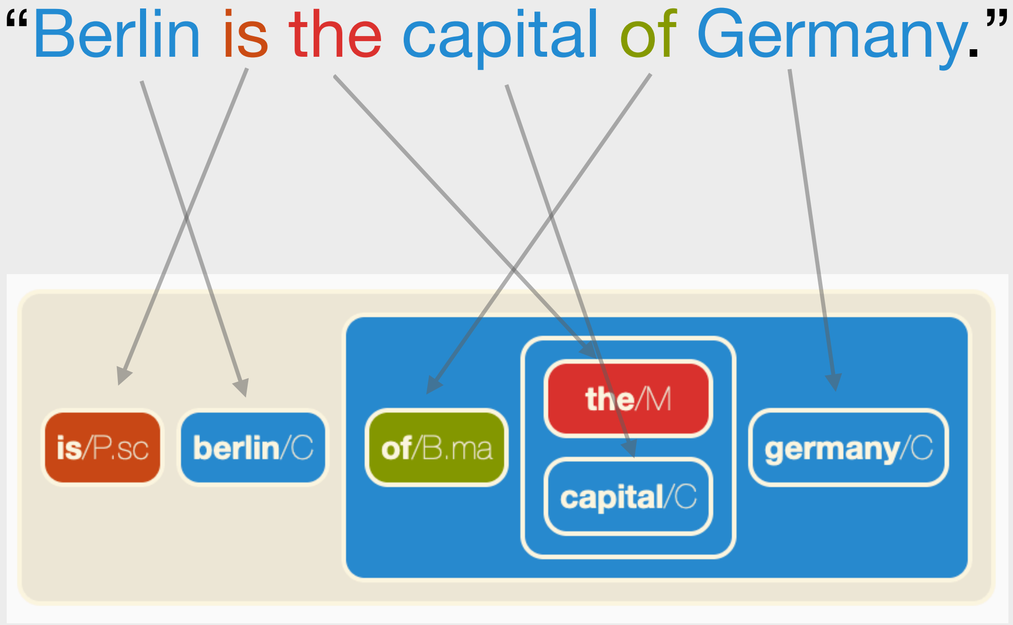
Semantic hypergraphs

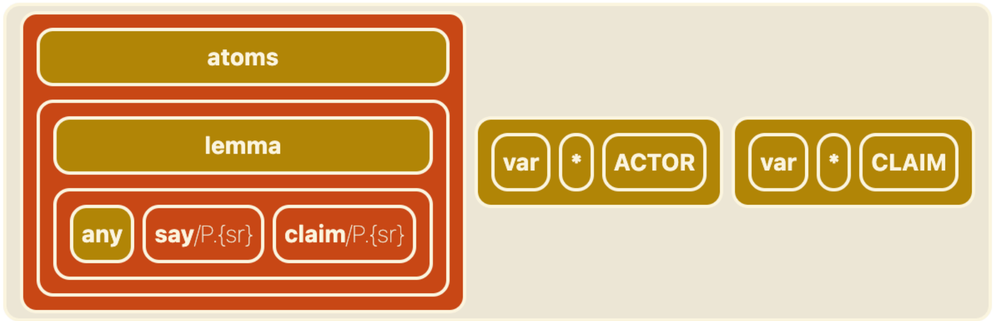
Topological

Menezes & Roth, 2019

Bonaccorsi, Melluso, Chiarello, Fantoni, 2021a, 2021b

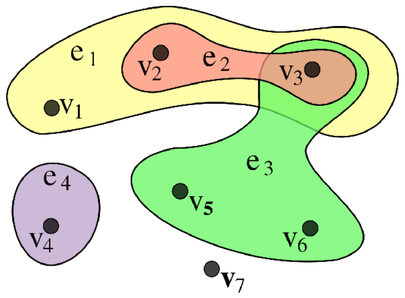
Menezes, Pottier, Roth, 2023





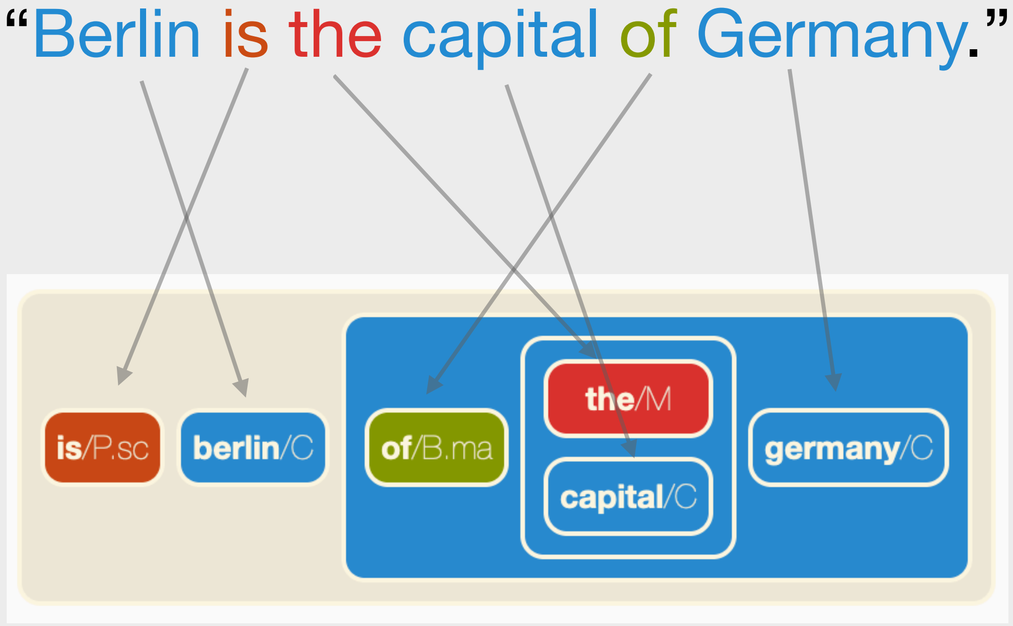
the sentence level

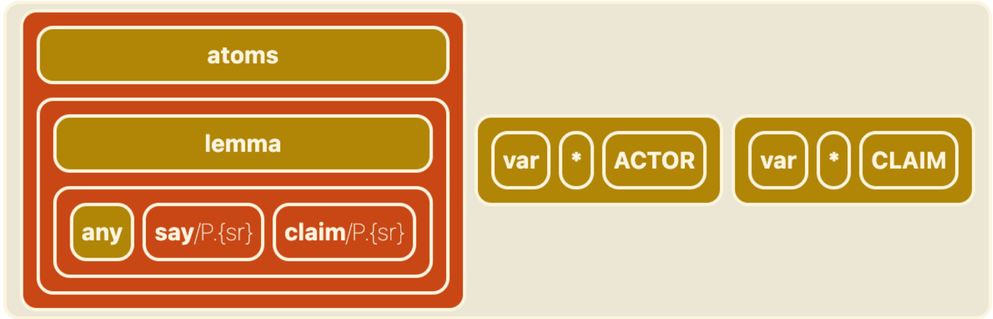
pattern definition and extraction

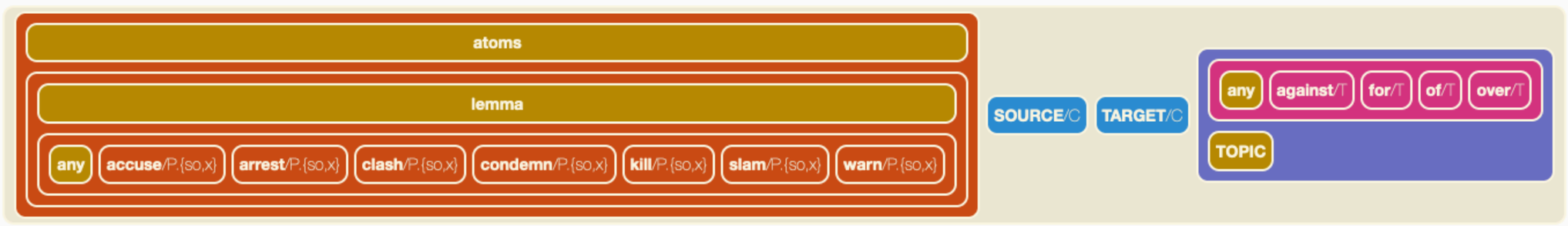


Semantic hypergraphs

Topological







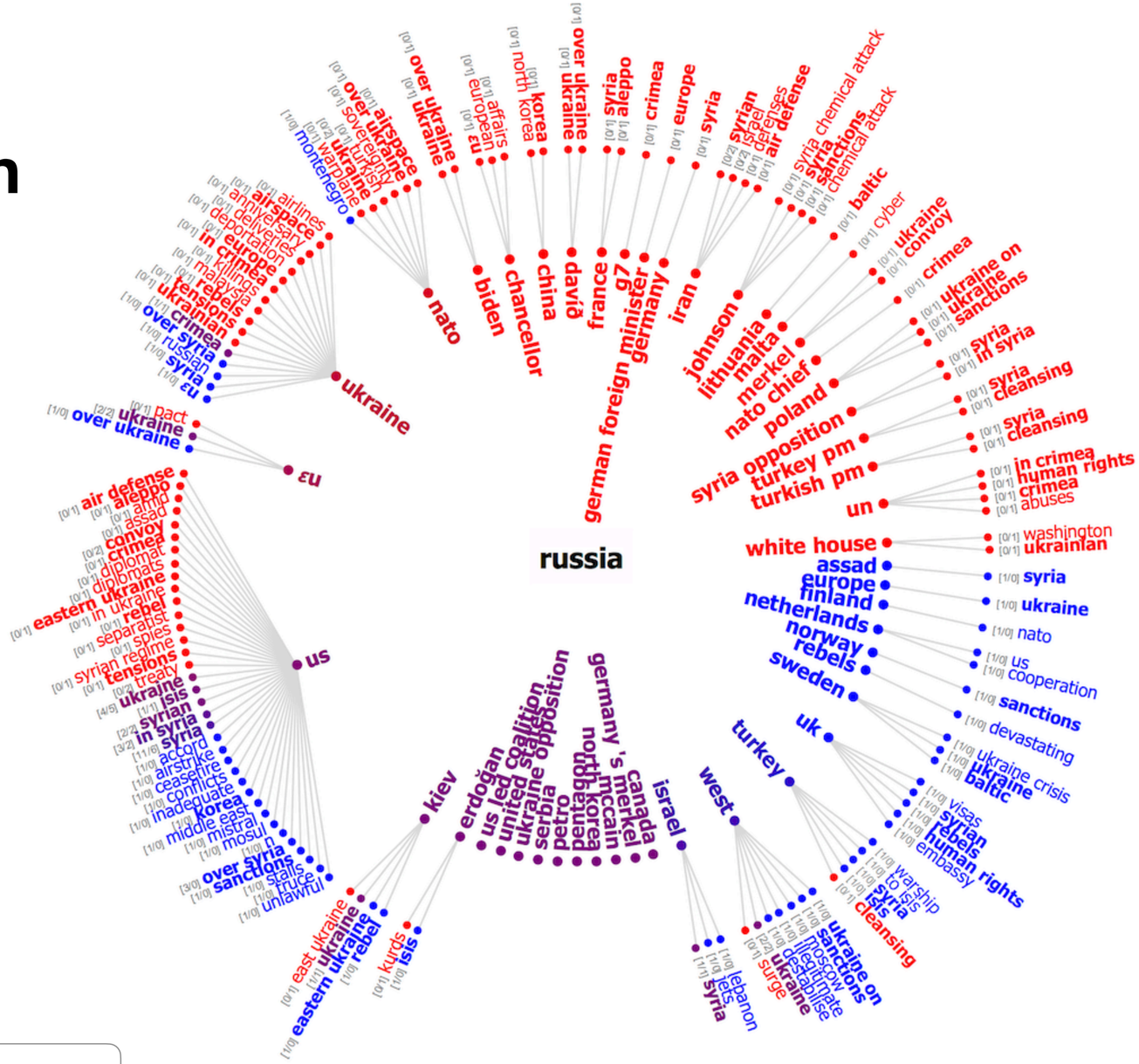
Menezes & Roth, 2019

Bonaccorsi, Melluso, Chiarello, Fantoni, 2021a, 2021b

Menezes, Pottier, Roth, 2023

the sentence level

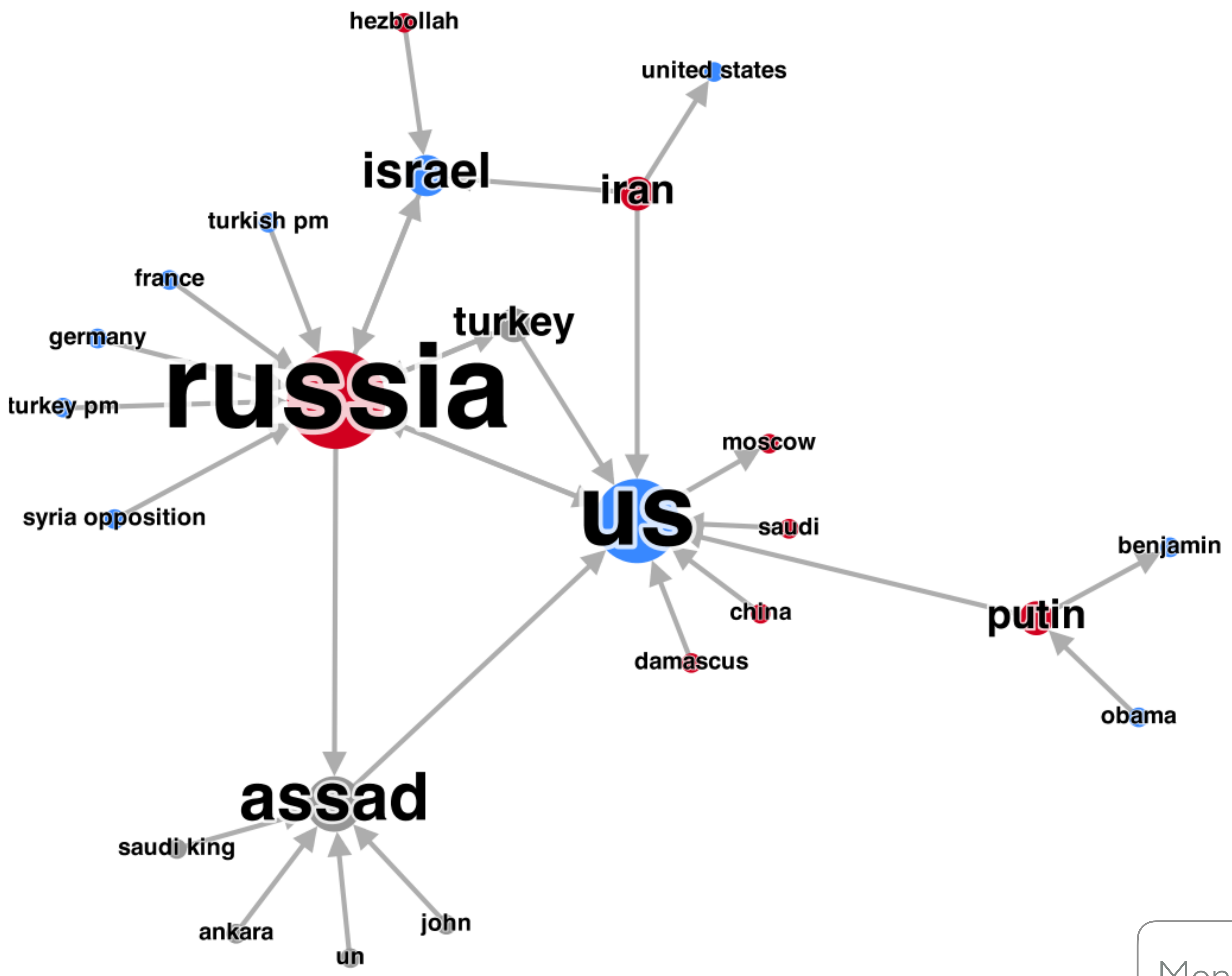
pattern definition and extraction

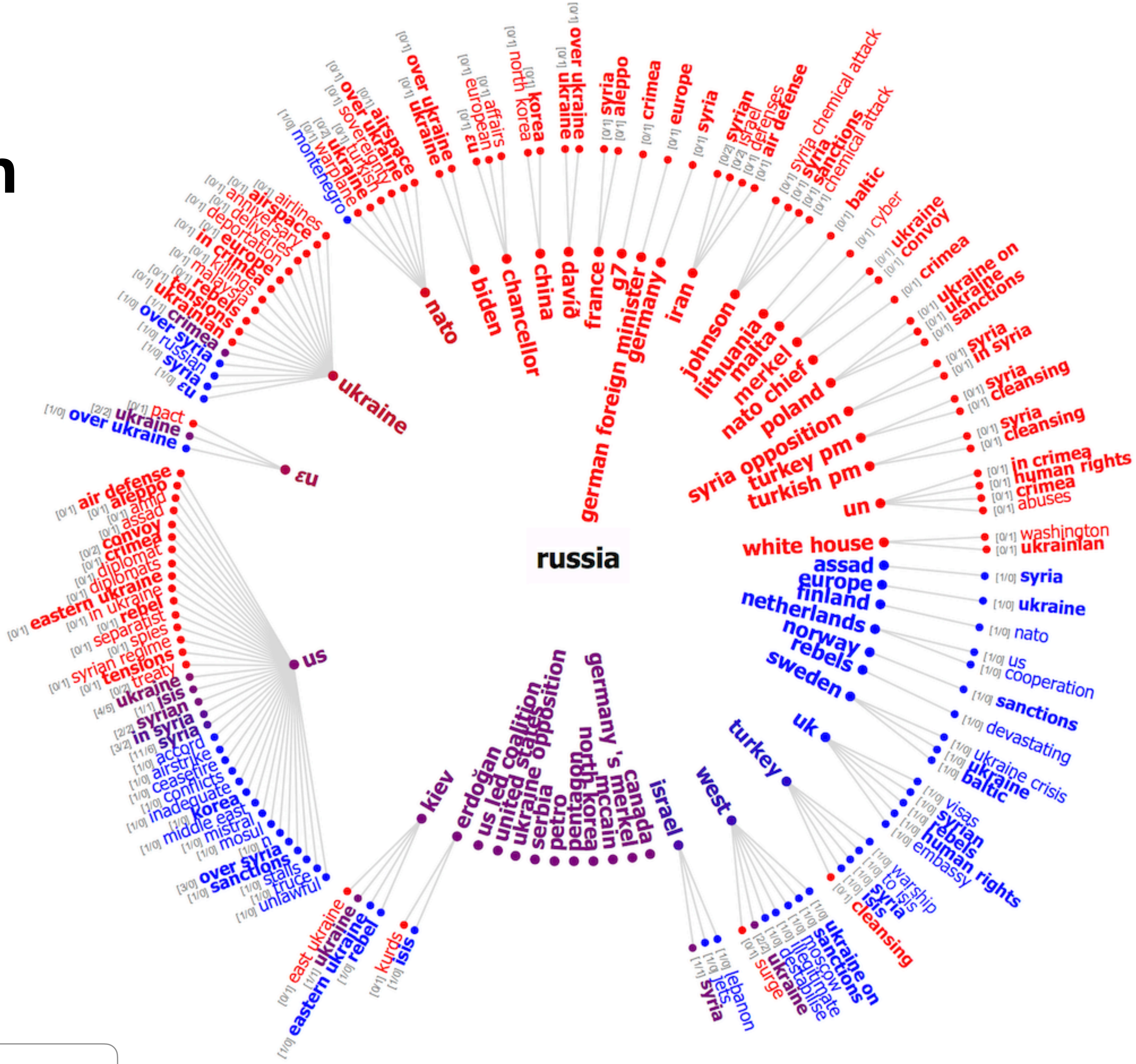


Menezes & Roth, 2019

the sentence level

pattern definition and extraction





Menezes & Roth, 2019

the sentence level

pattern definition and extraction

Set		Atoms
Claim predicates	*positive* (P^+)	show/P, indicate/P, confirm/P, support/P, suggest/P, reveal/P, provide/P, validate/P, exist/P, demonstrate/P, verify/P, imply/P, illustrate/P, find/P, point/P, exhibit/P, establish/P, obtain/P, hold/P, follow/P
	negative (P^-)	reject/P, challenge/P, fail/P
EKC concepts		kuznets/C, ekc/C, turning/C
Curve concepts		curve/C, shape/C, shaped/C
Negative modifiers		not/M, n't/M, no/M, little/M, poor/M
Result concepts	(R)	result/C, finding/C, test/C, evidence/C, support/C

predicate p based on	$\mathcal{N}(H)$	H contains an N-curve	*EKC validation claim type*
P^+	True	True	X
		False	*negative result*
	False	True	*negative result*
		False	*positive result*
P^-	True		X
	False	True	X
		False	*negative result*

Menezes, Pottier, Roth, 2023

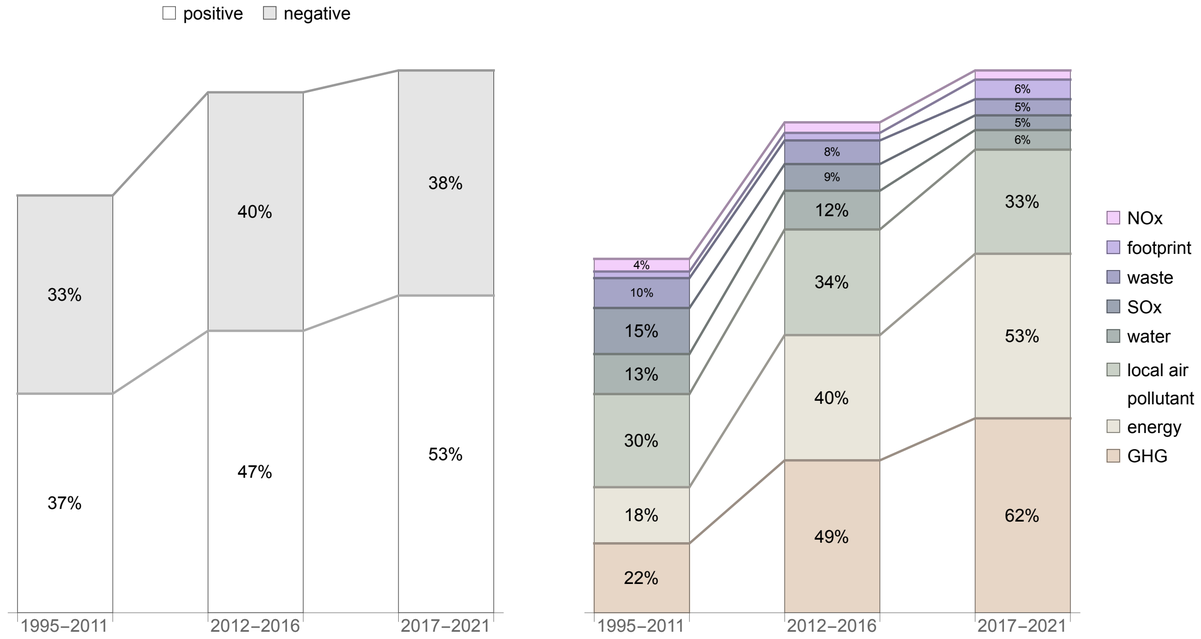


Figure 2: *Left:* estimated percentage of articles with a positive or a negative result (the bars do not add to 100% because articles can present no results, or both positive and negative results). *Right:* percentage of articles evoking a relationship on a given topic (sums over 100% as an article may address several topics, bar elements representing less than 4% were not labeled).

EKC validation claim type	Precision	Recall
positive result	.809	.847
negative result	.833	.366

the sentence level

pattern definition and extraction

Block	#authors	pos	neg	endogamy	year	#A	$\overline{\#A}$	k
A	4	.318	.076	.002	2007.7	7.50	0.44	4.5
B	366	.397	.380	.034	2012.4	3.22	0.76	5.1
C	8	.605	.455	.027	2015.9	7.75	1.35	11.1
D	392	.538	.380	.104	2018.1	4.27	1.26	8.9
E	64	.472	.366	.080	2016.6	6.13	1.00	15.6

Table 4: Various metrics per author block, including ratios of positive and negative EKC claims, endogamy measured as ratio of citations received from co-authors, mean year of publication, mean number of articles per author (#A) and further normalized per year ($\overline{\#A}$), and mean number of unique co-authors (k).

Menezes, Pottier, Roth, 2023

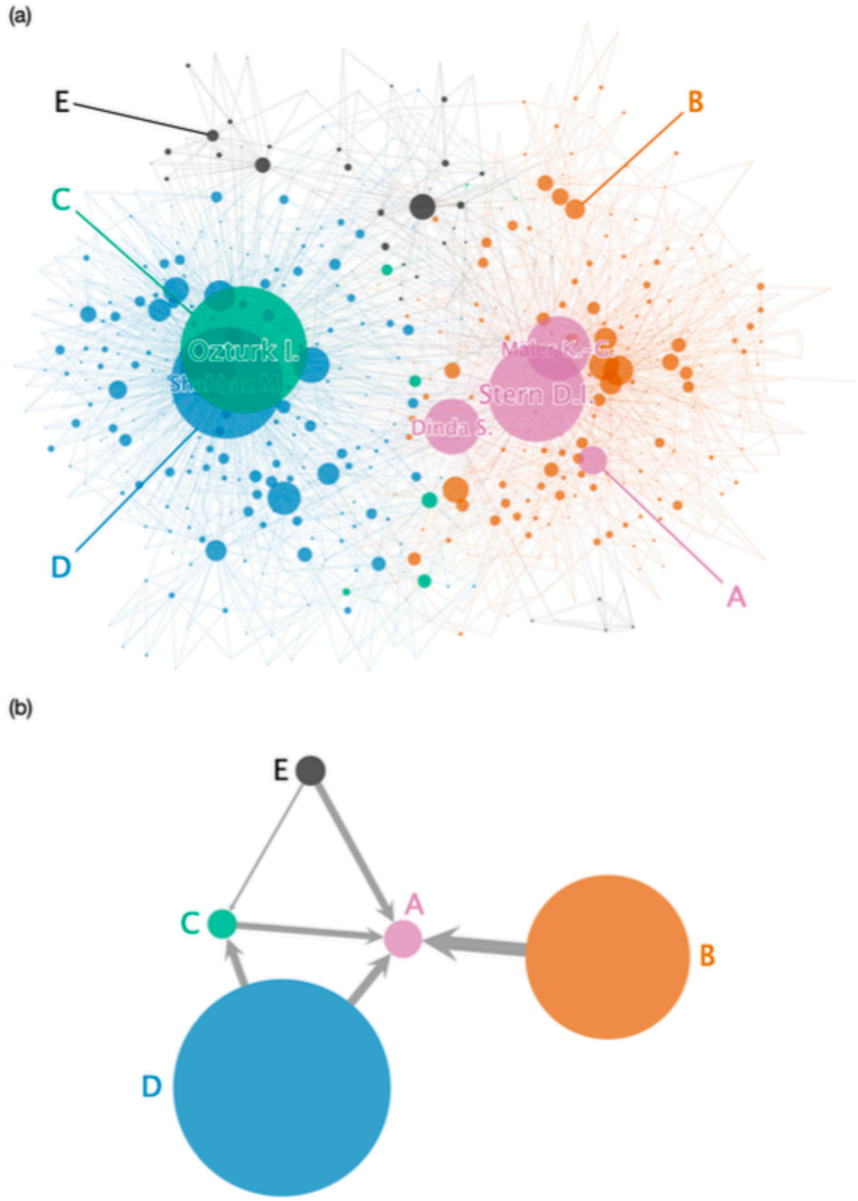


Figure 3: (a) Graph of citations between authors, colored by blocks determined by SBM. Node radius is proportional to in-degree. (b) Graph of author blocks. Edge thickness is proportional to the probability of connection from authors of one block to another. Node radius is proportional to the total number of citations (for the entire corpus) received by the authors in the block.

quite a bit descriptive so far

a.k.a. the gap between models and empirical observations

morphogenesis of socio-semantic clusters

ubiquitous homophily, a variety of outcomes

Flache, Mäs, Feliciani, Chattoe-Brown, Deffuant, Huet, Lorenz, 2017

Models of social influence:
Towards the next frontiers.

- assimilative social influence
  (interaction increases similarity)
- similarity bias
  (the other way around: similarity influences interaction)
- repulsive influence
  (usually combining assimilation with repulsion)

social influence may reinforce similarity

yet

social influence may reinforce divergence

morphogenesis of socio-semantic clusters

ubiquitous homophily, a variety of outcomes

Flache, Mäs, Feliciani, Chattoe-Brown, Deffuant, Huet, Lorenz, 2017

Models of social influence: Towards the next frontiers.

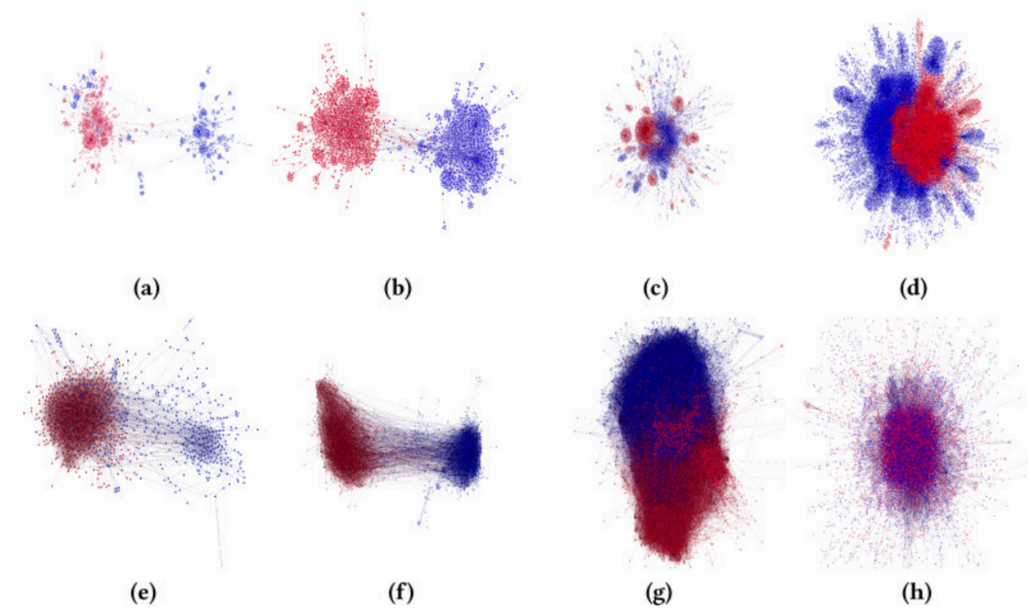
- assimilative social influence
  (interaction increases similarity)
- similarity bias
  (the other way around: similarity influences interaction)
- repulsive influence
  (usually combining assimilation with repulsion)

social influence may reinforce similarity

yet

social influence may reinforce divergence

- yet...



Garimella et al., 2016, inter alia