Complexity beyond pure reason?

Emergence through the lens of coincidence

Markus Luczak-Roesch | @mluczak

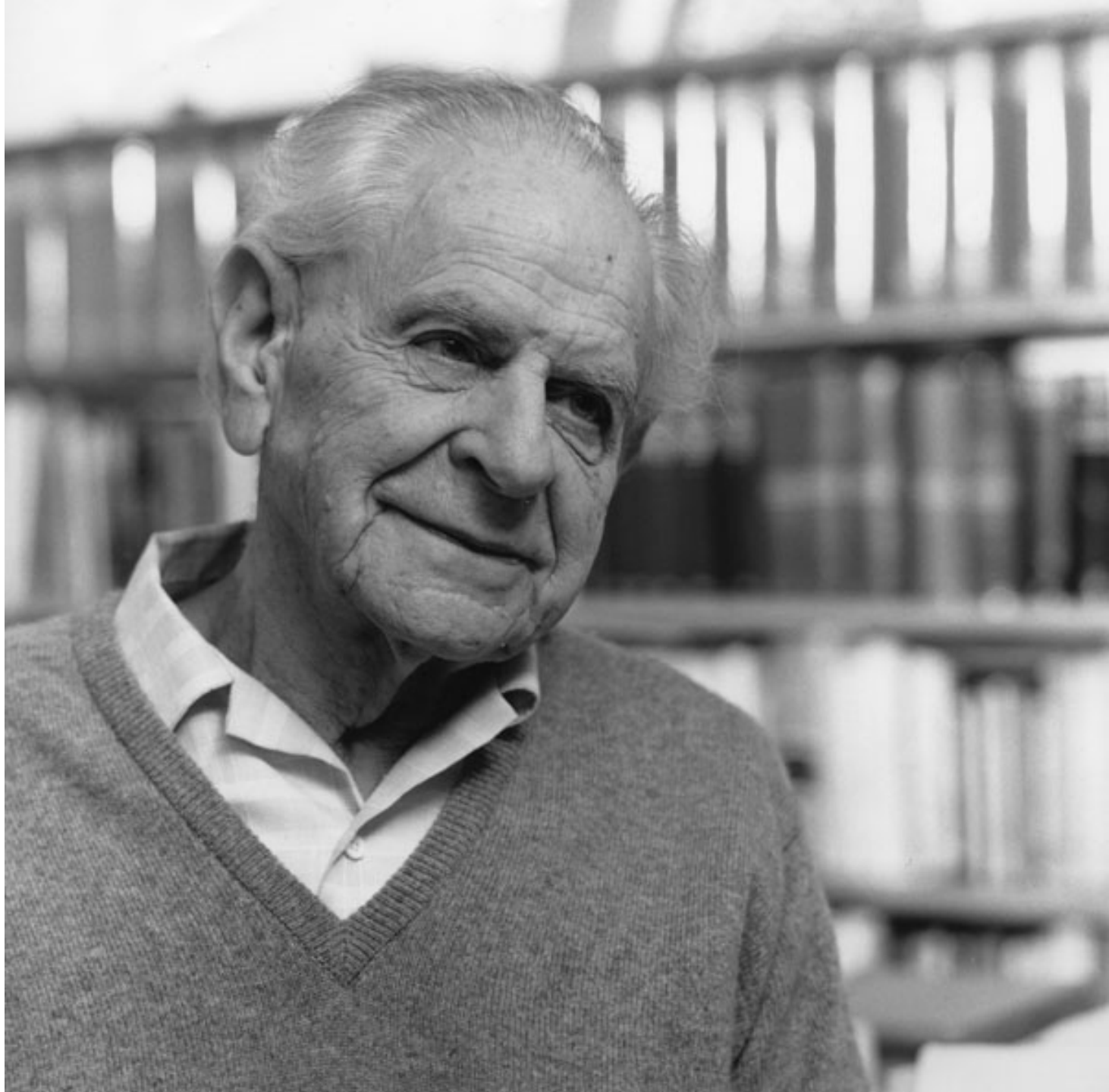
Victoria University of Wellington
School of Information Management

A story of three worlds[1]

World 1
physical objects and events
(our ability to make noises)

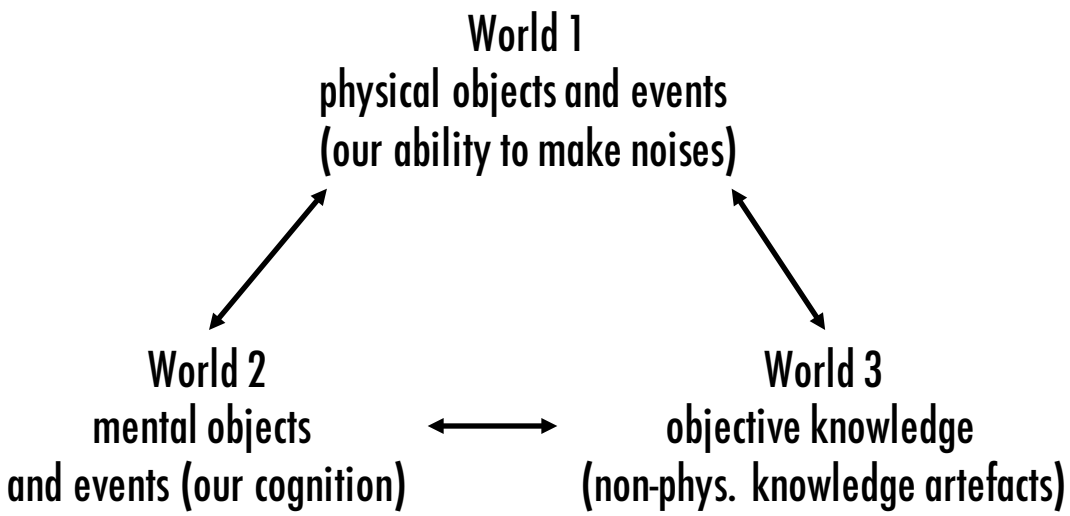
World 2
mental objects
and events (our cognition)

World 3
objective knowledge
(non-phys. knowledge artefacts)



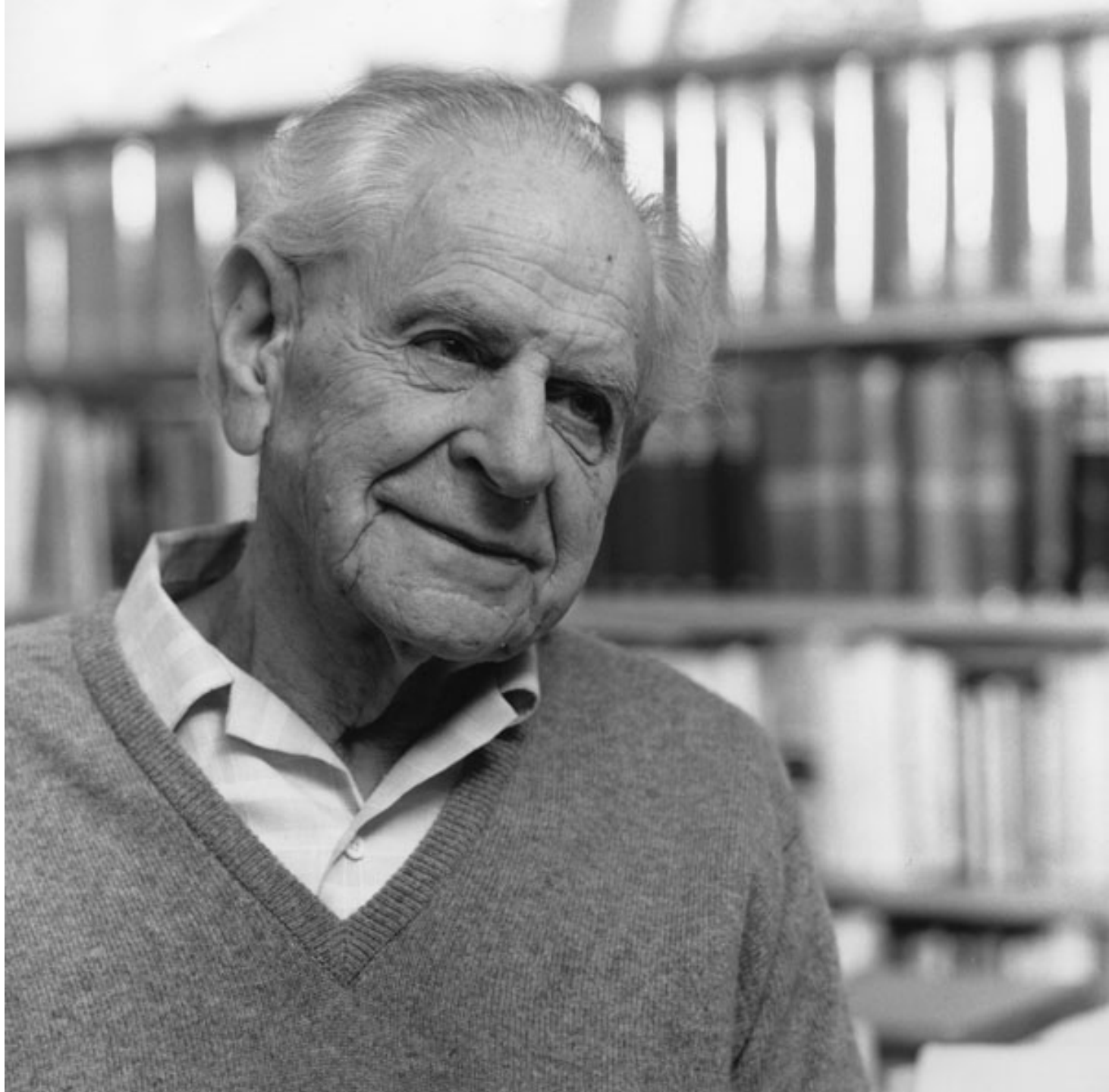
[1] Popper, K., 2013. Knowledge and the Body-Mind Problem: In defence of interaction. Routledge. (original lecture in 1969)

A story of three worlds

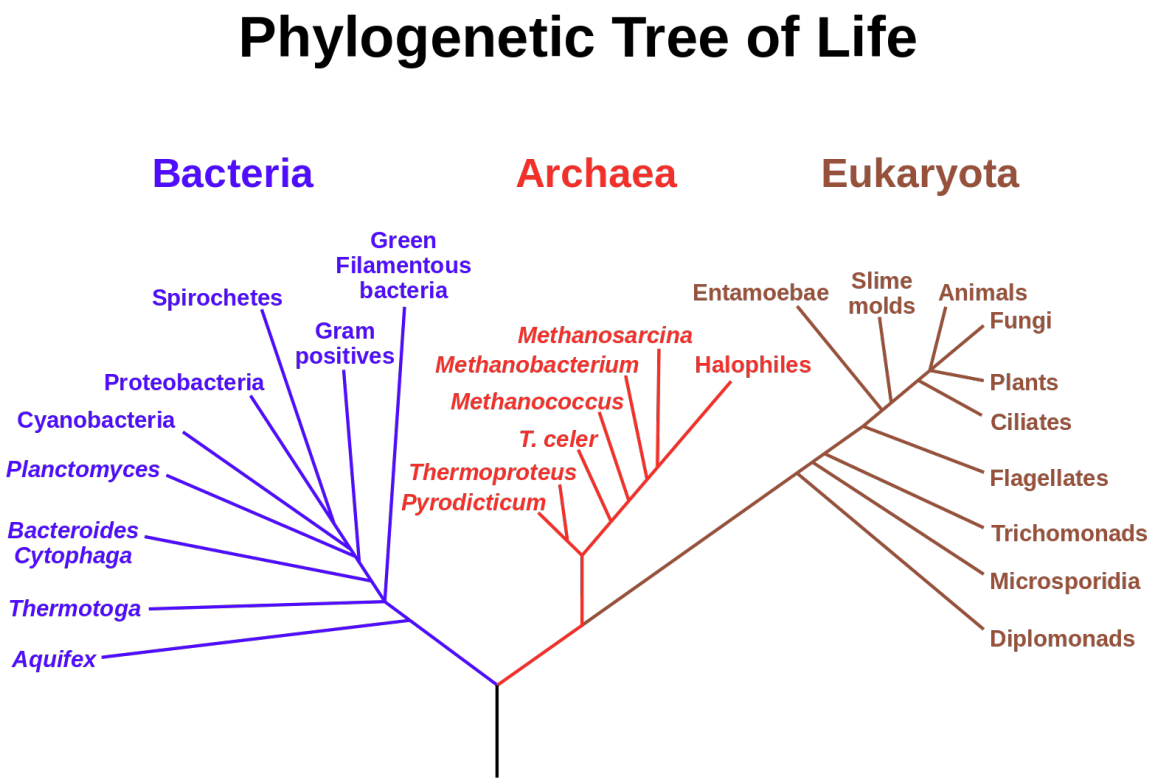


"a transcendental method in Kant's [3] sense of attempting to understand the conditions of knowledge itself" [2]

[2] Luczak-Roesch M, Tinati R, O'Hara K. (2017) What an entangled Web we weave: An information-centric approach to socio-technical systems. PeerJ Preprints 5:e2789v1 https://doi.org/10.7287/peerj.preprints.2789v1
[3] Kant, I., Critique of pure reason. Translated by Norman Kemp Smith. London Macmillan 532 1934.



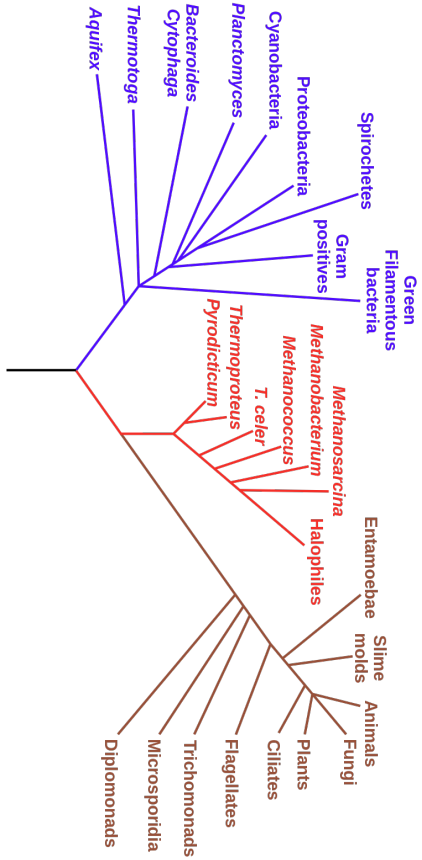
Evolutionary emergence



Public domain image from https://en.wikipedia.org/wiki/File:Phylogenetic_tree.svg

Evolutionary emergence

- Popper suggests the view that evolution means "ascending" in the genetic tree misses the point
  - what's optimal under the conditions at one point may become suboptimal when the conditions change
  - instead of "ascending into higher forms" it is "increased variety" that should be regarded



Coincidence

- events that are temporally related
  but have no observed causal relationship

- C.G. Jung coined the term "synchronicity" for cases of "acausal but meaningful coincidences"
  - underpins his theory of the collective unconscious
  - Obscure theory of parapsychology or fundamental to our understanding of the mind?

Jung, C.G., 1952. Synchronicity: An Acausal Connecting Principle. In Vol. 8. of the Collected Works of CG Jung. Princeton University Press, 2010.

Diaconis and Mosteller on coincidence[9]

- directions for a general theory of coincidences are
  - hidden cause
  - ~~psychology~~
  - Multiple Endpoints and the Cost of "Close"
  - The Law of Truly Large Numbers

- but they emphasized
  - **"[…] we are handicapped by lack of empirical work. We do not have a notion of how many coincidences occur per unit of time […]"**

[9] Diaconis, P. and Mosteller, F., 2006. Methods for studying coincidences. In Selected Papers of Frederick Mosteller (pp. 605-622). Springer, New York, NY.

Some contemporary problems

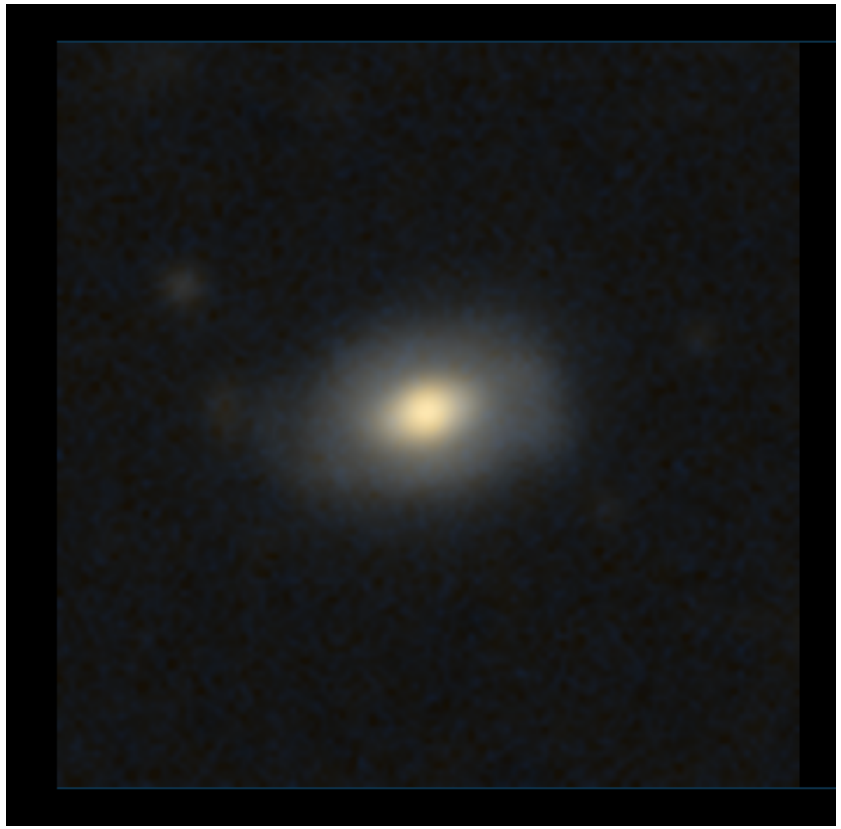
What is a Social Machine and how does it compute?

"Real life is and must be full of all kinds of social constraint – the very processes from which society arises. Computers can help if we use them to create **abstract social machines on the Web: processes in which the people do the creative work and the machine does the administration**."

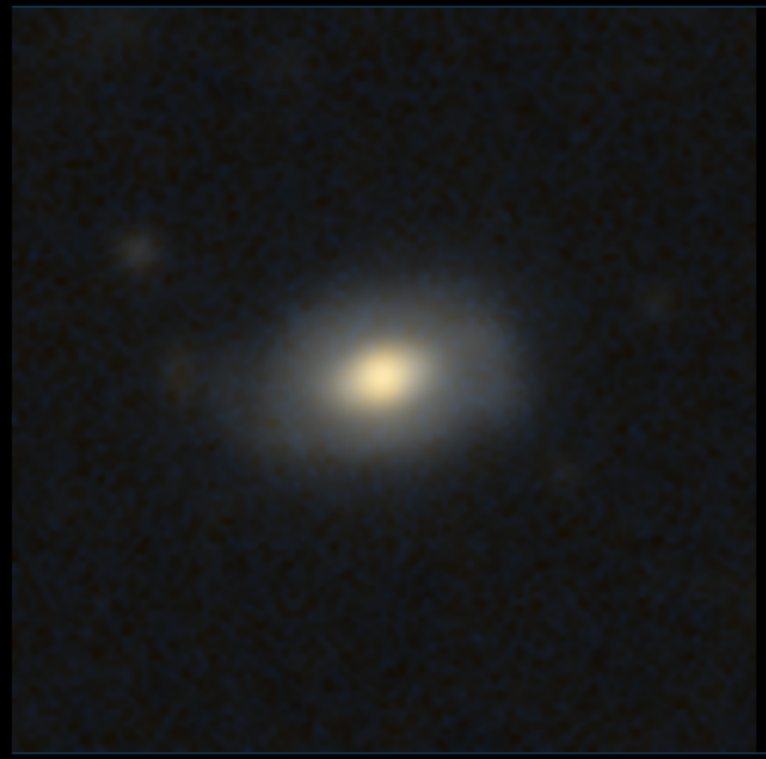
Berners-Lee, Tim; Mark Fischetti (1999). Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its inventor. Britain: Orion Business. ISBN 0-7528-2090-7.

ARCSAT and SDSS telescope buildings at the Apache Point Observatory in Sunspot, USA

Classifying galaxies is a human task



CLASSIFY STORY SCIENCE GALAXY ZOO DISCUSS PROFILE LANGUAGE



Classify

GAMA Favourite Invert

Examples Restart

Note: Please always classify the galaxy in the centre of the image.

SHAPE

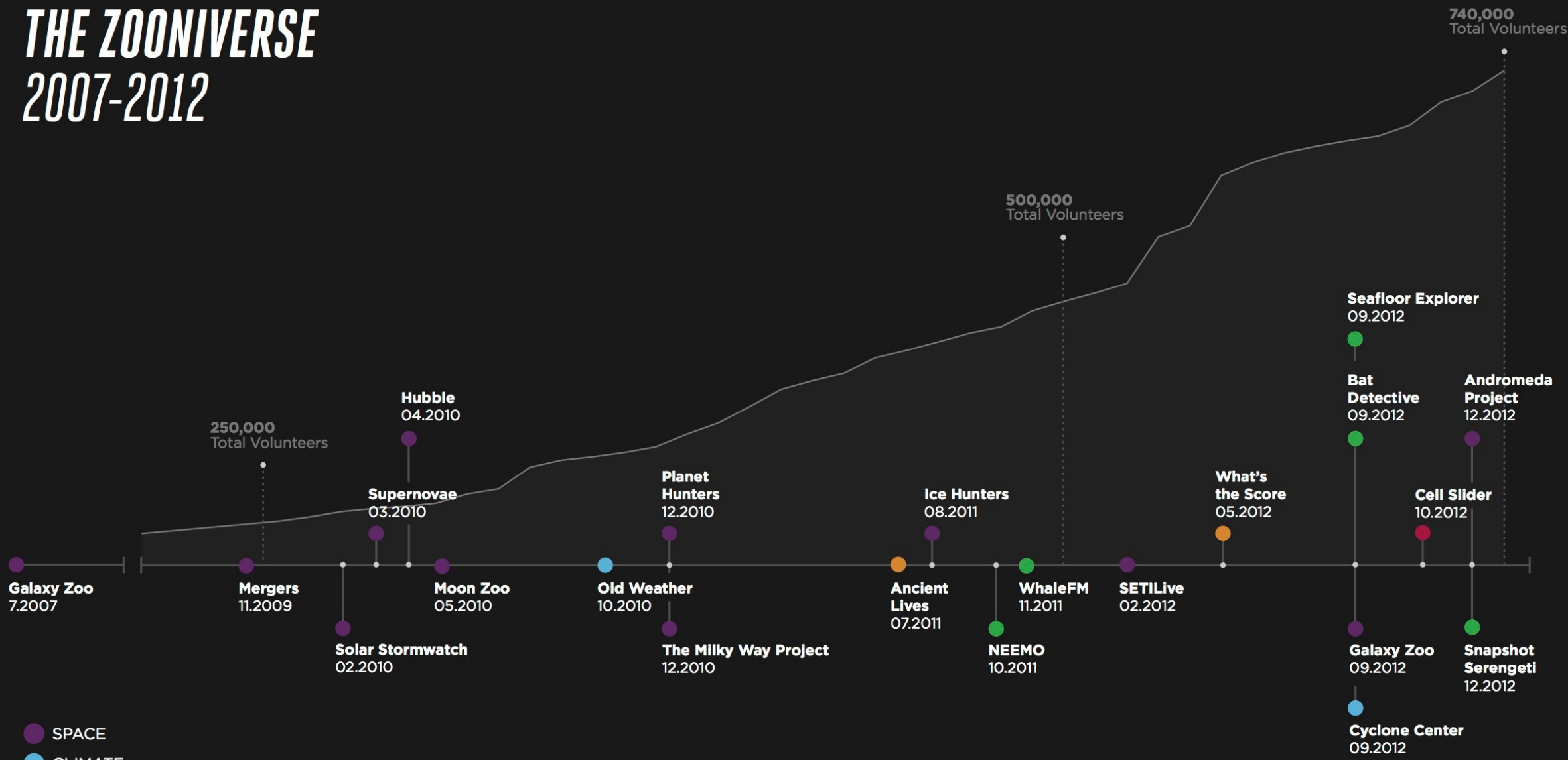
Is the galaxy simply smooth and rounded, or does it have features?

Smooth

Features

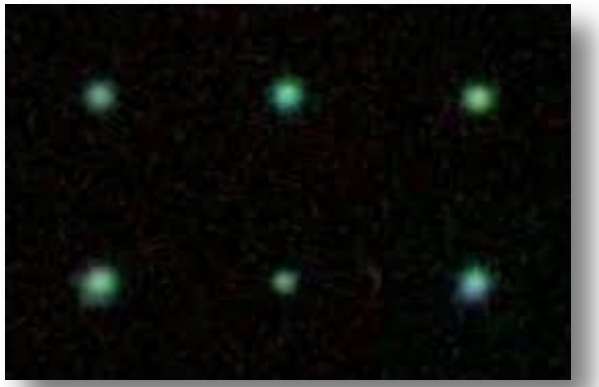
Star or artifact

THE ZOONIVERSE
2007-2012



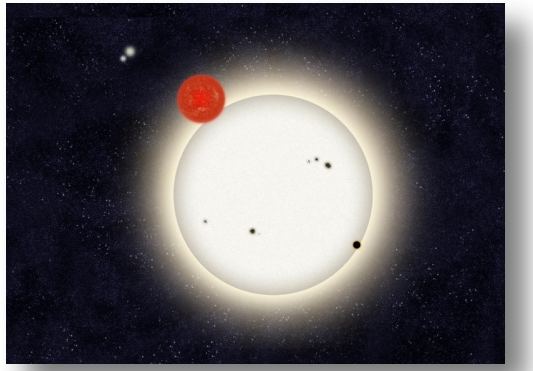
Source: http://zooniverse.org

Hanny's Voorwerp
Galaxy Zoo [2007]



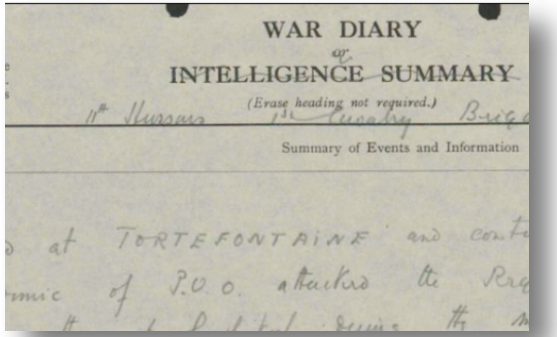
Green Pea Galaxies
Galaxy Zoo [2007]

Yellow Balls
Milky Way [2009]



Circumbinary Planet Ph1b
Planet Hunter [2012]

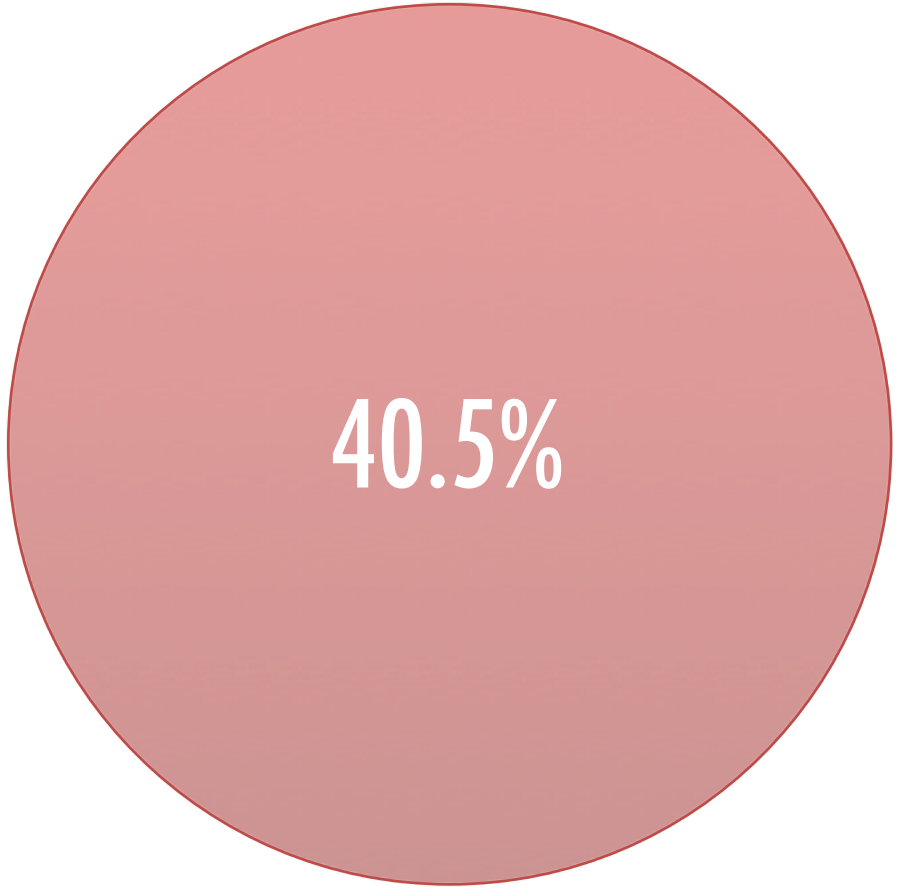
Convict Worm
Seafloor Explorer [2012]

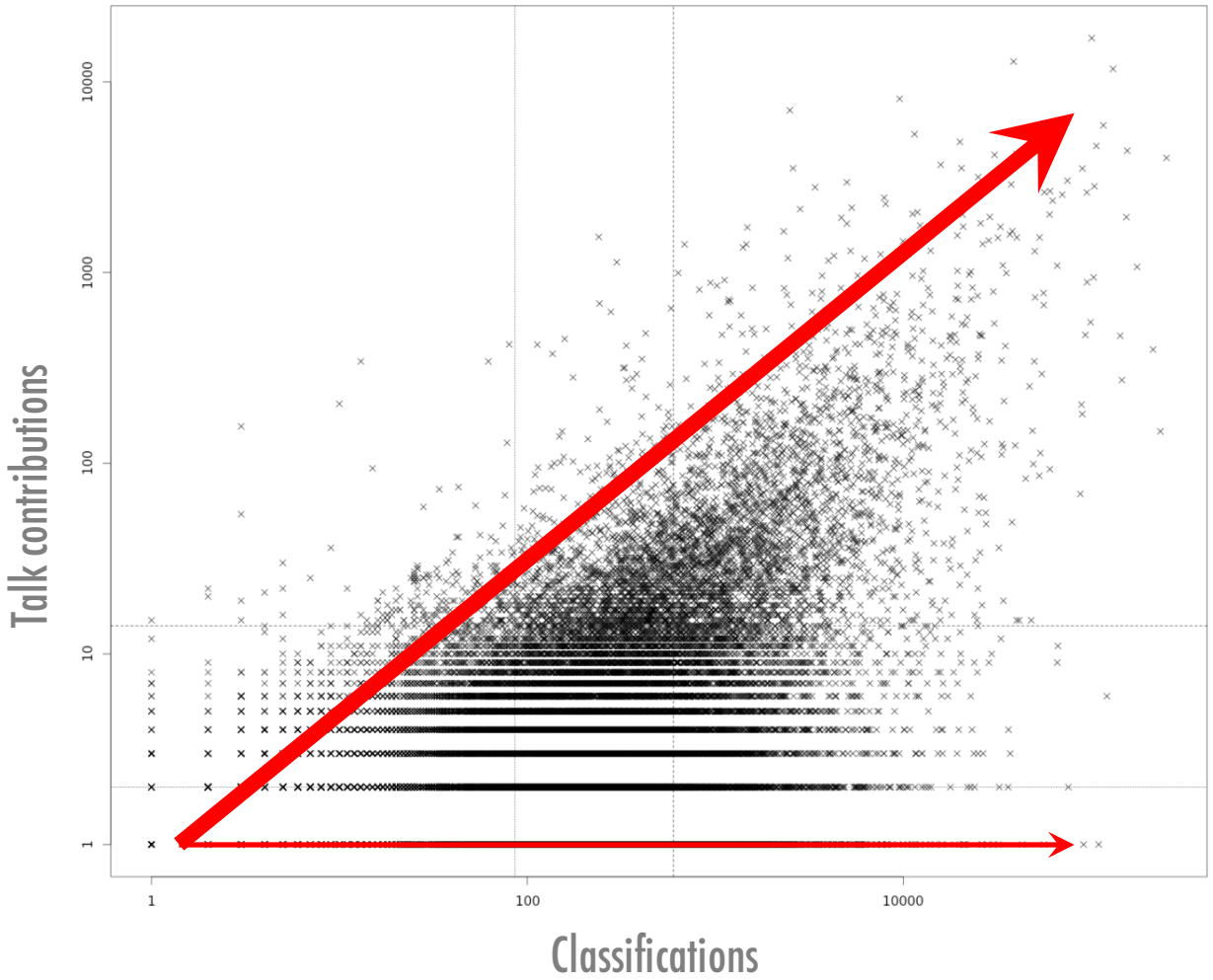


Spanish Flu
Operation War Diaries [2014]

Source: http://zooniverse.org

Task and talk participation





Community-level linguistic change

Project	initial 10%	most recent 10%
PH	Stable domain specific vocabulary	day, transit, **httparchive**. . . , possibl, star, **kid**, dip, look, planet, like
SF	Emerging domain specific vocabulary	corallinealga, anemon, object, hermitcrab, bryozoan, stalkedtun, shrimp, left, cerianthid, sanddollar
NN	Stable problem/error reporting	like, field, record, date, name, can, click, look, get, label

Emergent coordination

Q8-2 #transit at 18.6d

by user1 6 hours ago

Q8-2 #eclipse atat 18.6.

by user2 3 days ago

. . .

Q3-1 #transit #plain #gap

by user3 13 days ago

. . .

Q11-1 #transitingplanet #glitch

by user4 17 days ago

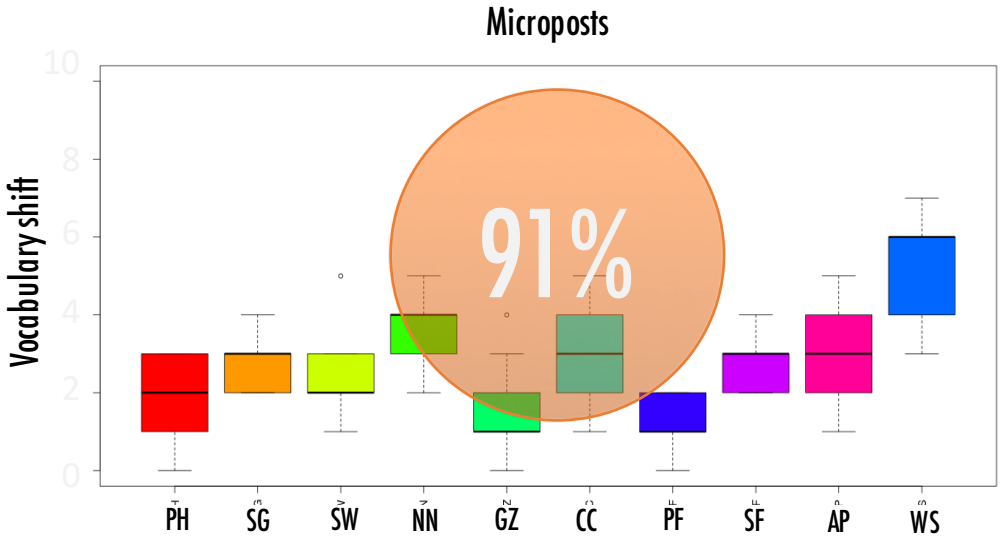
. . .

In the TCE list with P=31.2017, R=9.026 R_Earth.

by user2 17 days ago

Q5-1 #eclipse at 20.8, KID 9772531

by user2 17 days ago

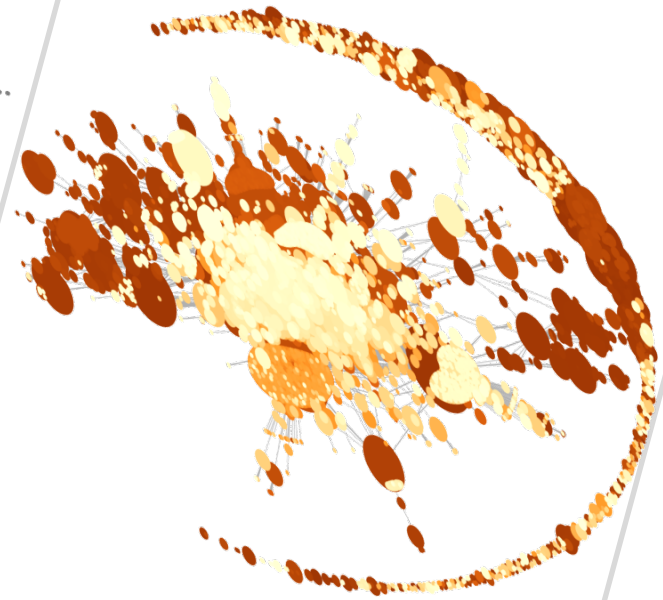


[3] Luczak-Roesch, M., Tinati, R., Simperl, E., Van Kleek, M., Shadbolt, N., & Simpson, R. (2014). Why won't aliens talk to us? Content and community dynamics in online citizen science. Proceedings of the Eighth AAAI Conference on Weblogs and Social Media, {ICWSM} 2014, Ann Arbor, Michigan, USA, June 1-4, 2014.

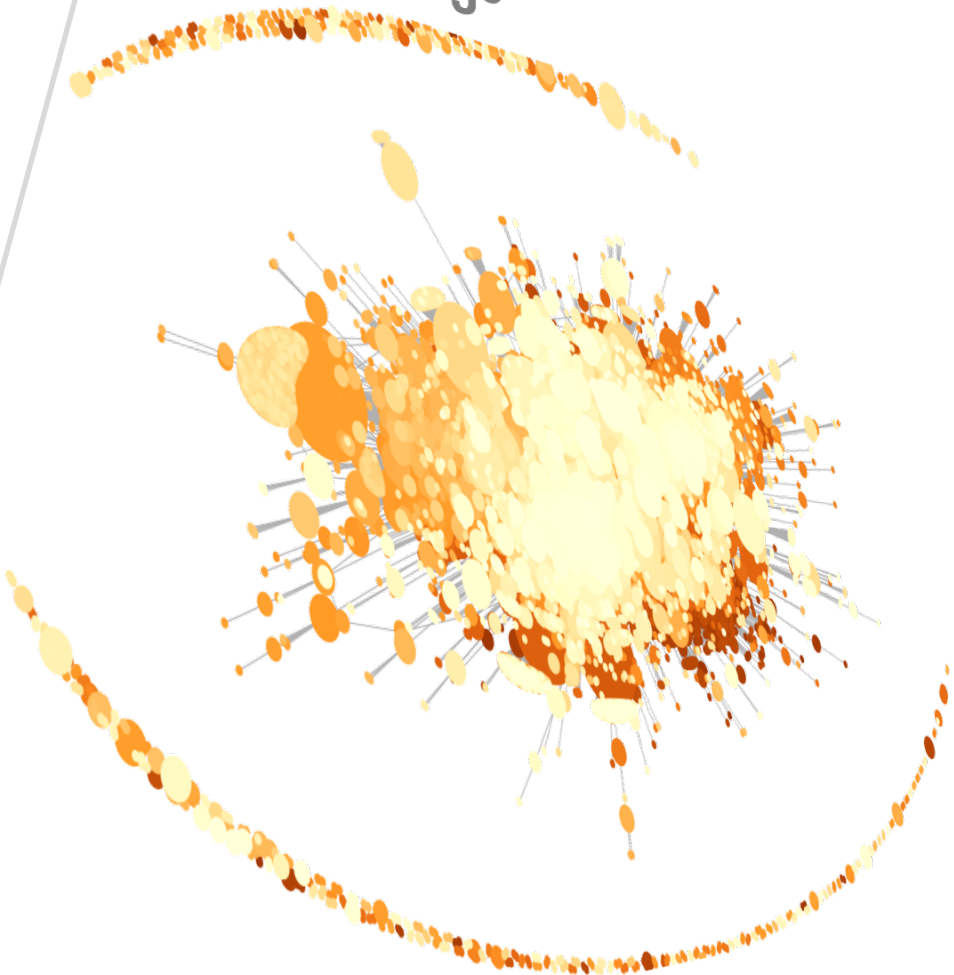
Networks within and out of the Zooniverse

Threads, classified subjects, incoming and outgoing links,...

Mentions

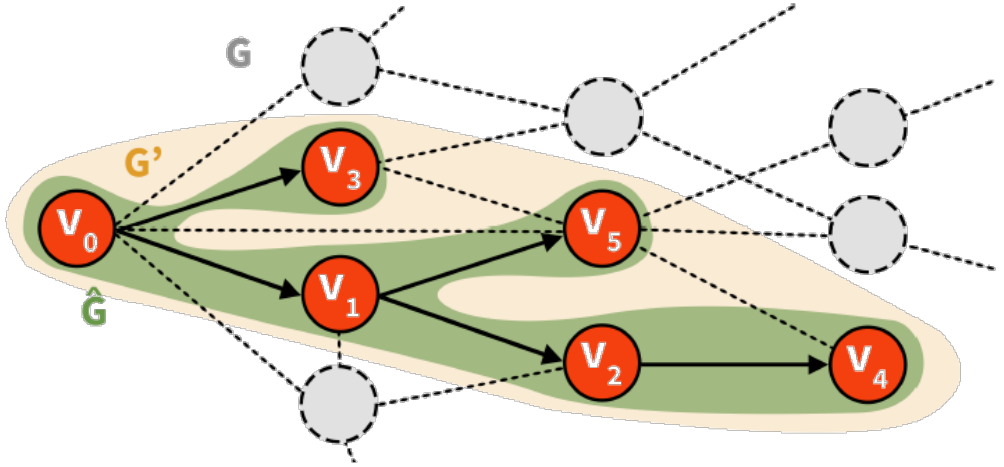


Tags

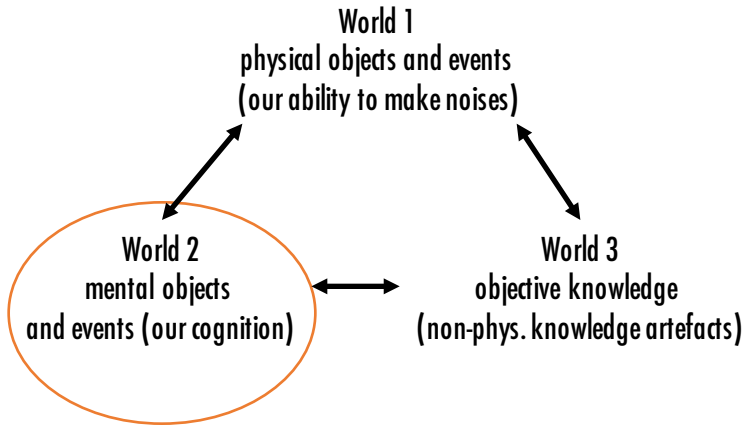


Boundaries of context-rich methods

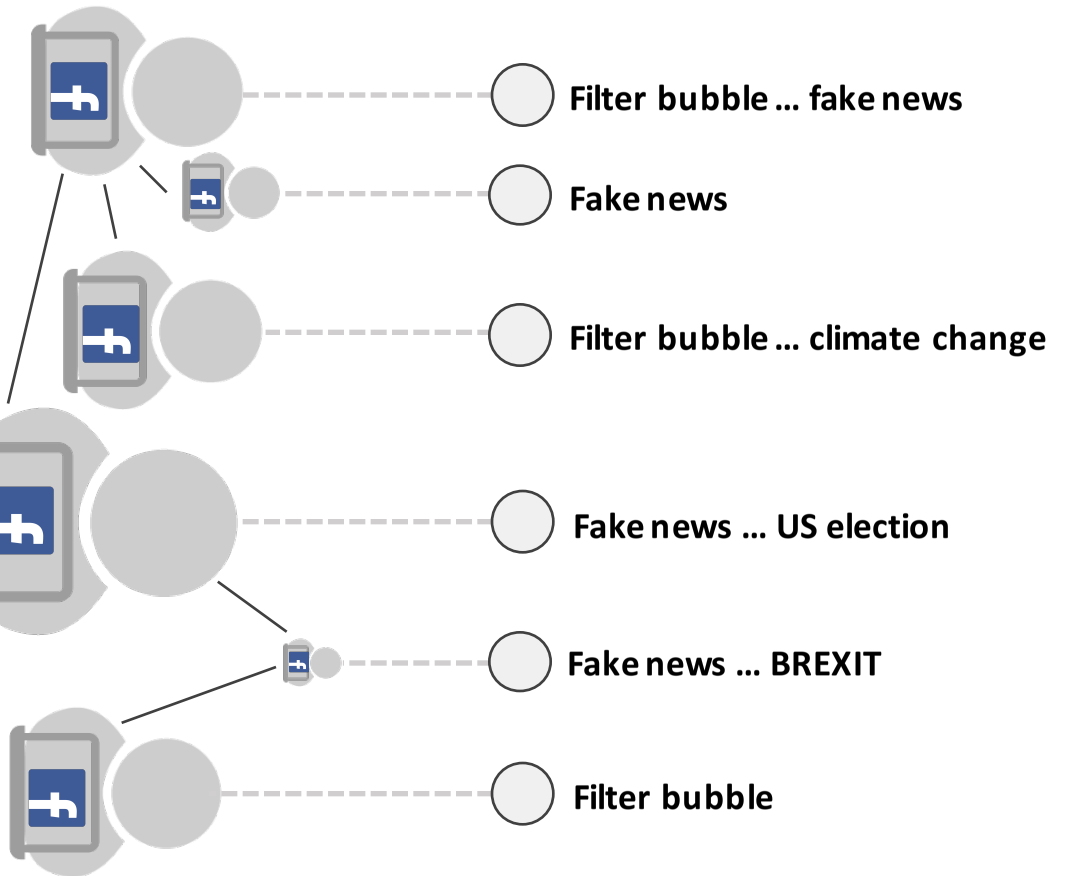
"An **informational cascade** occurs when it is optimal for an individual, having observed the actions of those ahead him, to follow the behavior of the preceding individual without regard to his own information." [4]

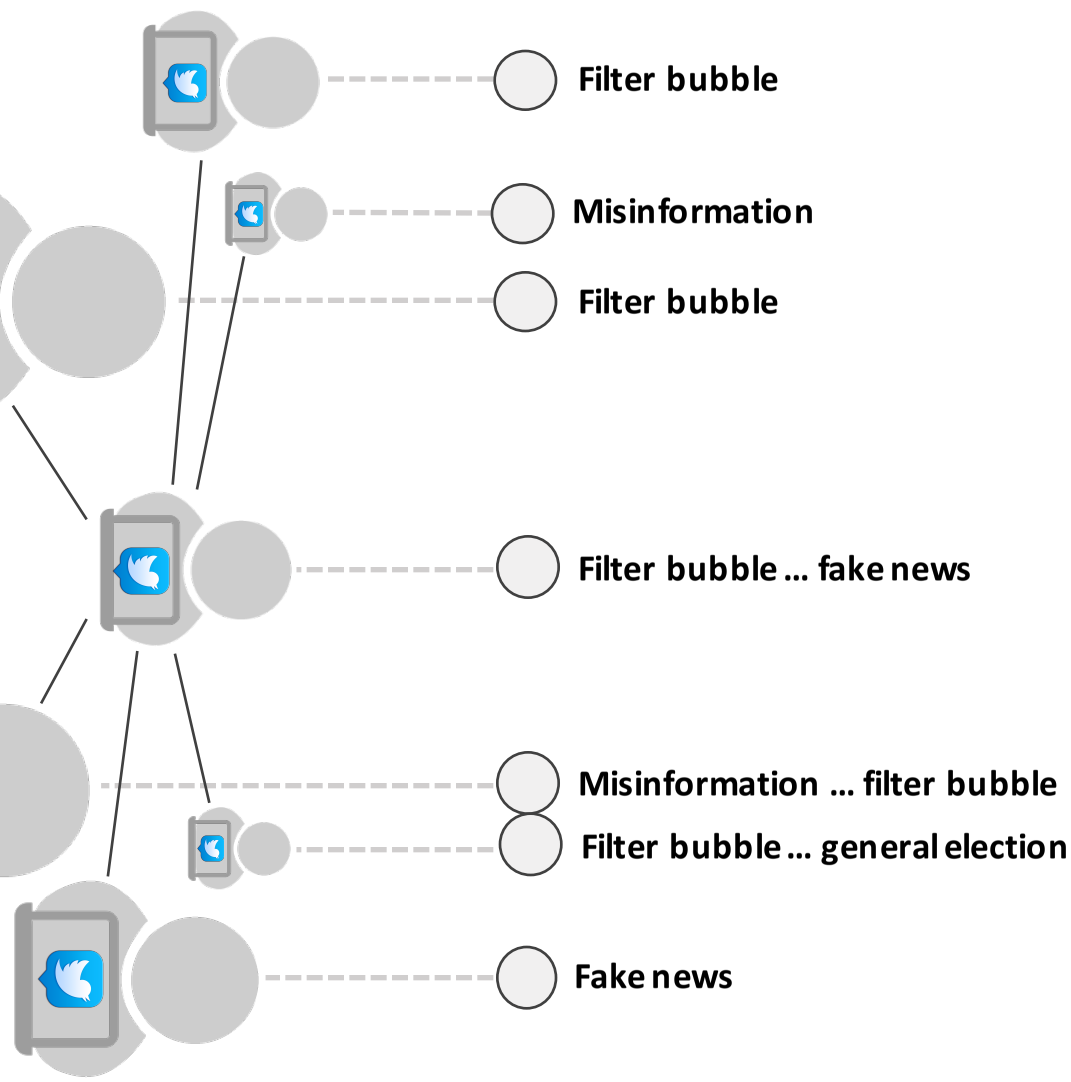


[5]



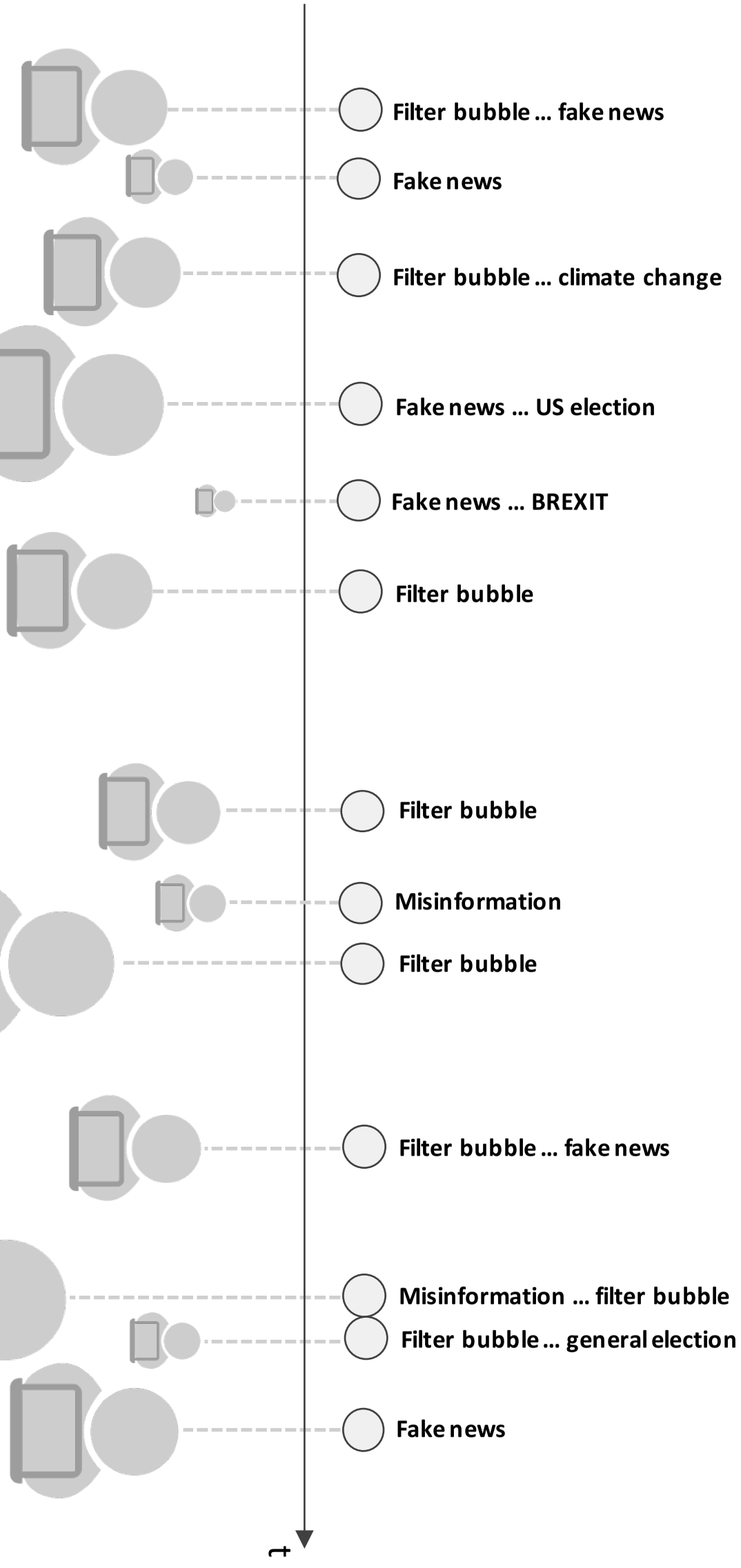
[4] Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. "A theory of fads, fashion, custom, and cultural change as informational cascades." Journal of political Economy (1992): 992-1026.
[5] Cheng, Justin, et al. "Can cascades be predicted?." Proceedings of the 23rd international conference on World wide web. International World Wide Web Conferences Steering Committee, 2014.

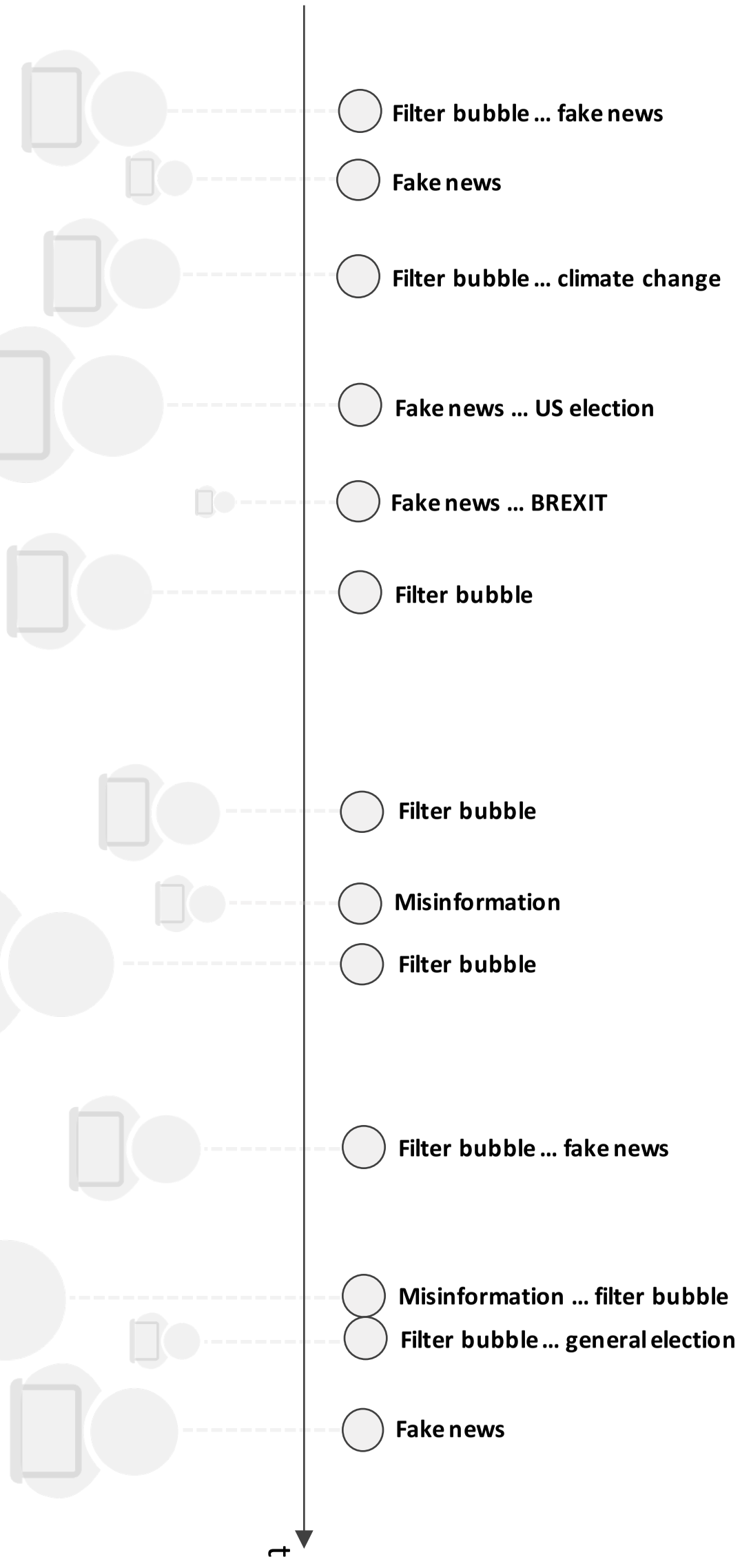

Filter bubble ... fake news
Fake news
Filter bubble ... climate change
Fake news ... US election
Fake news ... BREXIT
Filter bubble

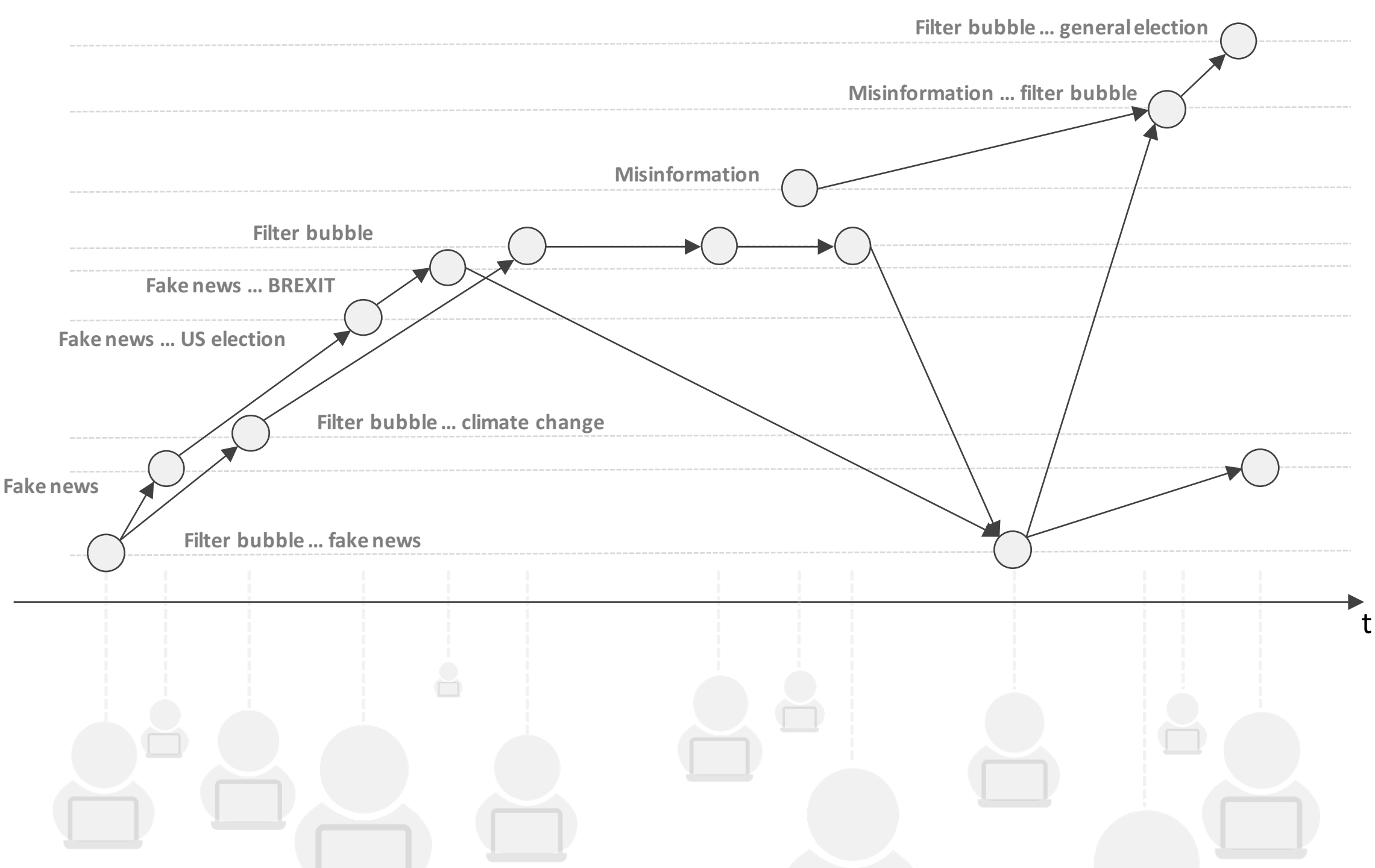

Filter bubble
Misinformation
Filter bubble
Filter bubble ... fake news
Misinformation ... filter bubble
Filter bubble ... general election
Fake news

We can observe situations when online communication does not happen along explicit social ties (especially in critical situations when the time to make decisions is limited). **Instead of talking explicitly with each other people are broadcasting about the same event or topic.**

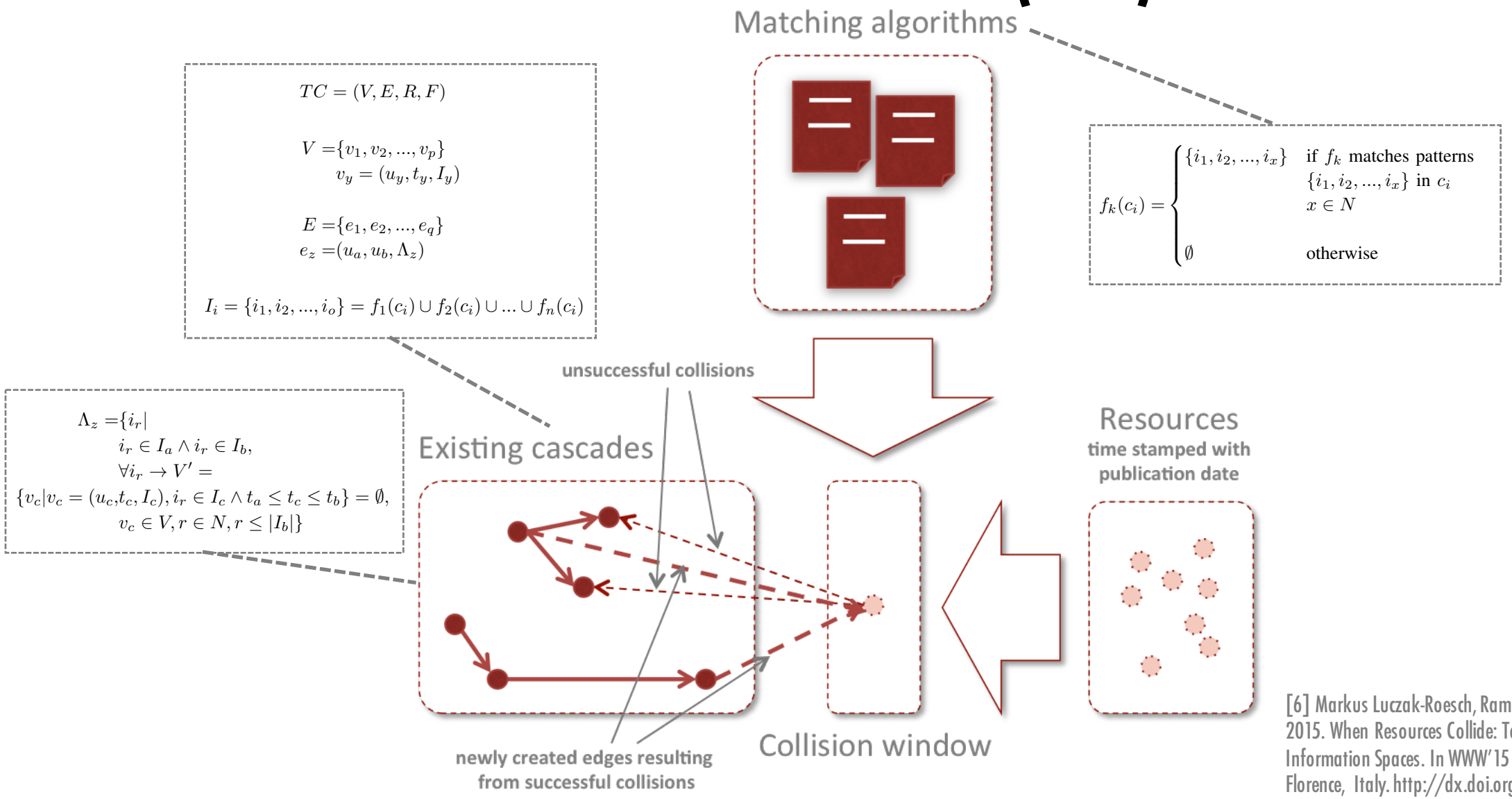
Source: United Nations Development Programme, https://goo.gl/Z1uXdV, CC BY-NC-ND 2.0


Filter bubble ... fake news
Fake news
Filter bubble ... climate change
Fake news ... US election
Fake news ... BREXIT
Filter bubble
Filter bubble
Misinformation
Filter bubble
Filter bubble ... fake news
Misinformation ... filter bubble
Filter bubble ... general election
Fake news
t


Filter bubble ... fake news
Fake news
Filter bubble ... climate change
Fake news ... US election
Fake news ... BREXIT
Filter bubble
Filter bubble
Misinformation
Filter bubble
Filter bubble ... fake news
Misinformation ... filter bubble
Filter bubble ... general election
Fake news
t

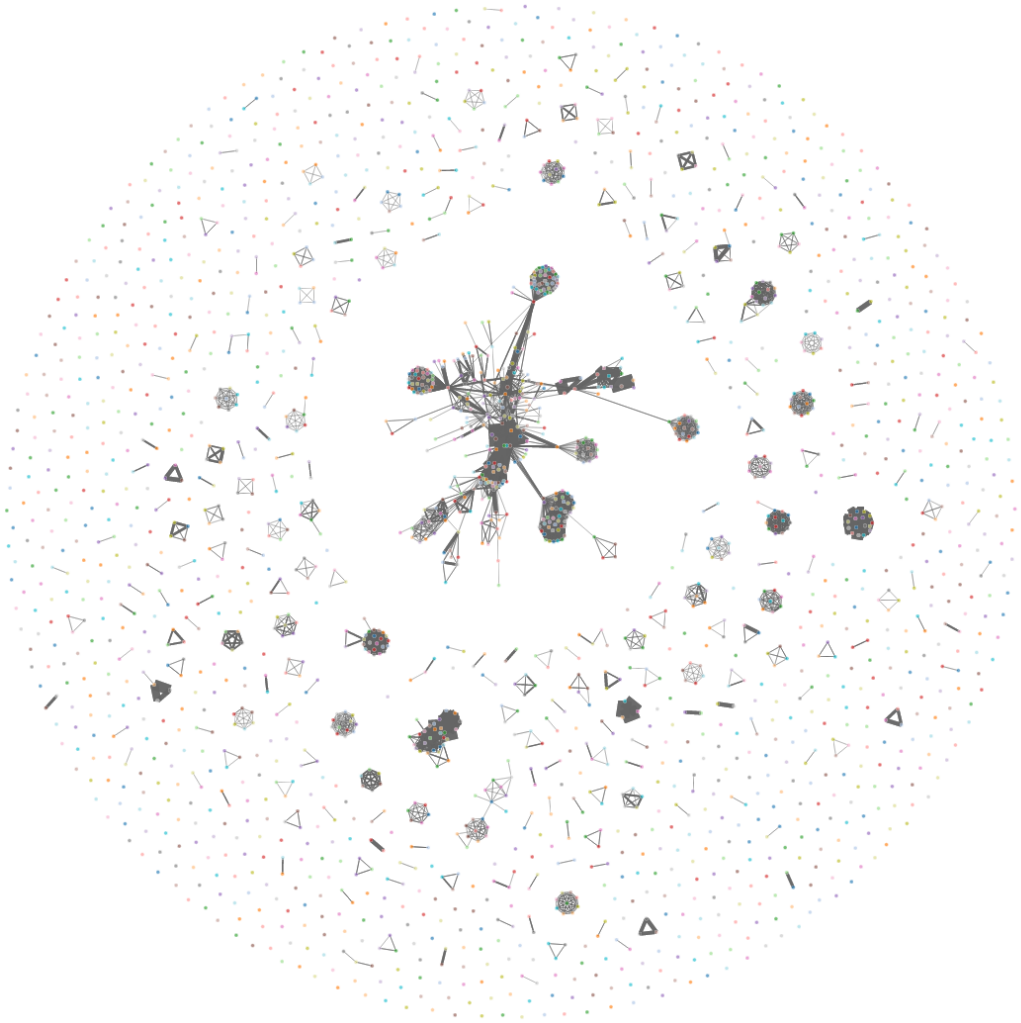

Filter bubble ... general election
Misinformation ... filter bubble
Misinformation
Filter bubble
Fake news ... BREXIT
Fake news ... US election
Filter bubble ... climate change
Fake news
Filter bubble ... fake news
t

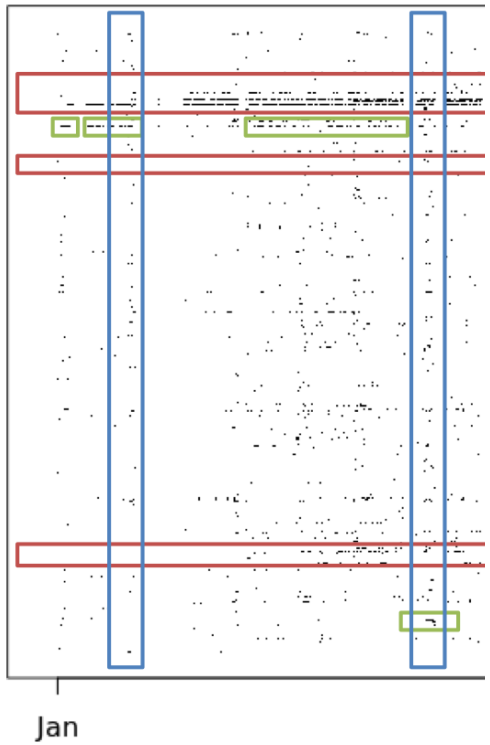
Transcendental Information Cascades (TICs)



[6] Markus Luczak-Roesch, Ramine Tinati, and Nigel Shadbolt. 2015. When Resources Collide: Towards a Theory of Coincidence in Information Spaces. In WWW'15 Companion, May 18–22, 2015, Florence, Italy. http://dx.doi.org/10.1145/2740908.2743973

Applying TICs to the Wikipedia edit history

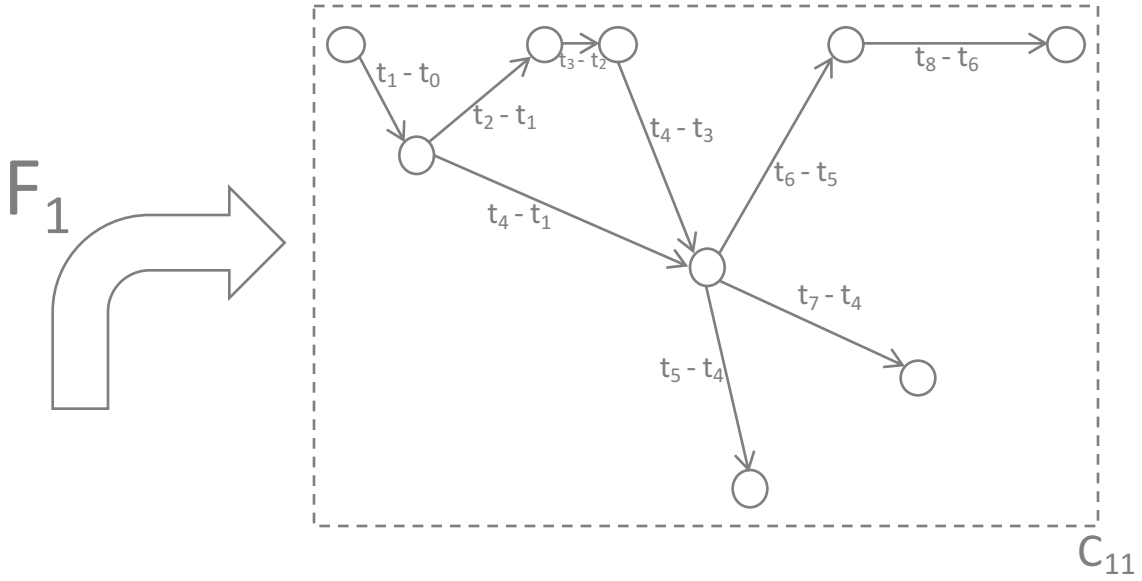


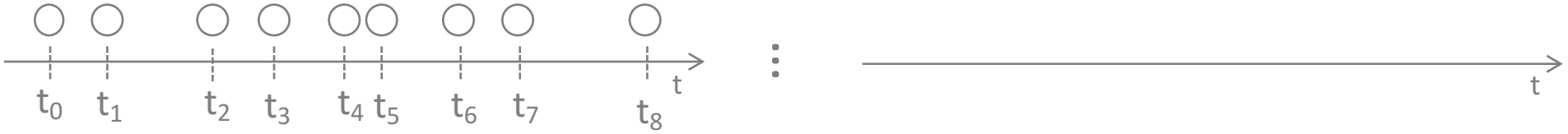


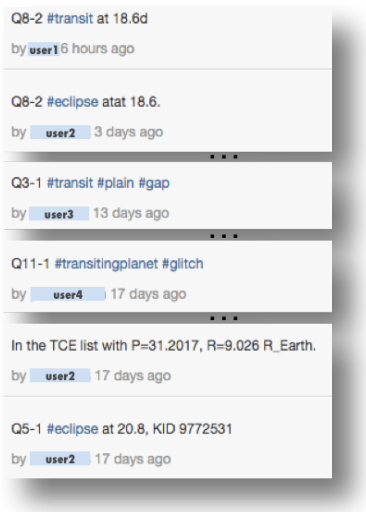
[7]

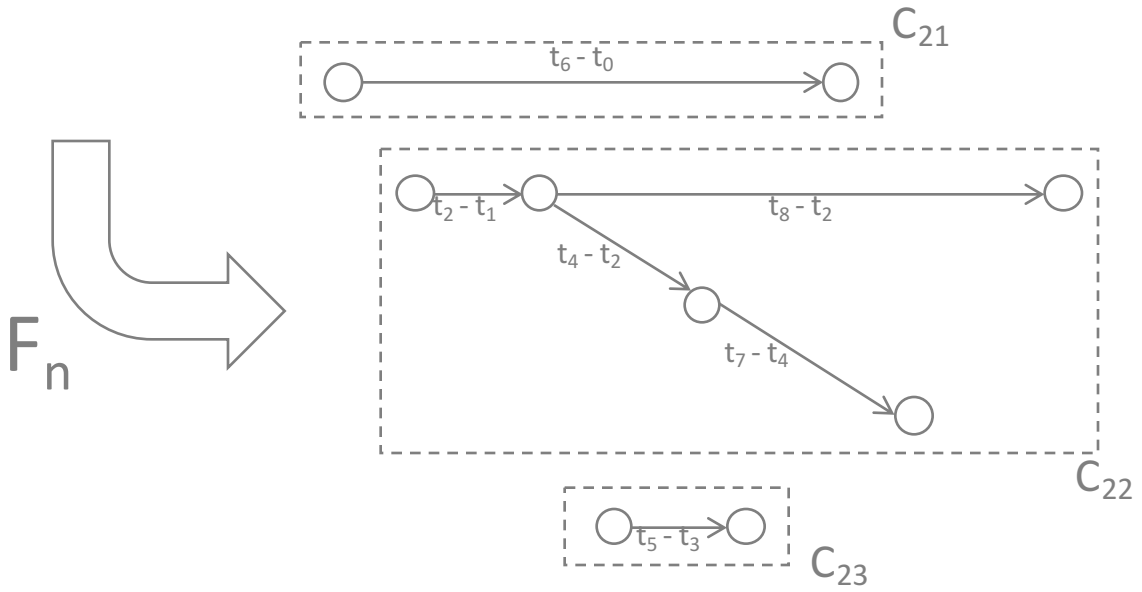
[7] Tinati, R., Luczak-Roesch, M. and Hall, W., 2016, April. Finding Structure in Wikipedia Edit Activity: An Information Cascade Approach. In Proceedings of the 25th International Conference Companion on World Wide Web (pp. 1007-1012). International World Wide Web Conferences Steering Committee.

There is more than one "reality"



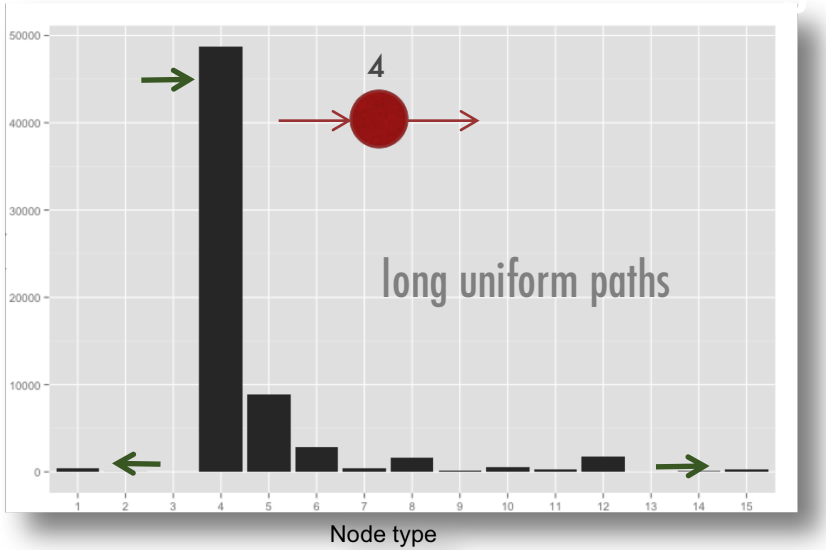




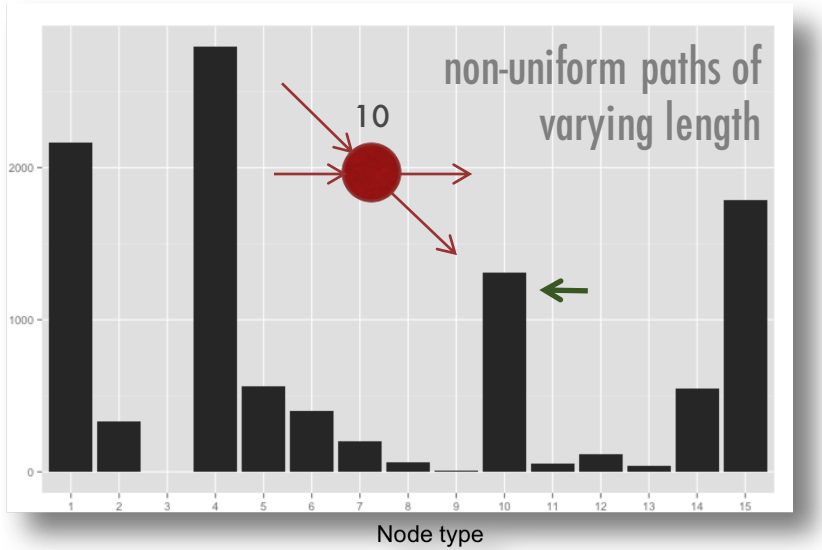


Analyzing low-level properties of the multiple states of a system that exist at the same time

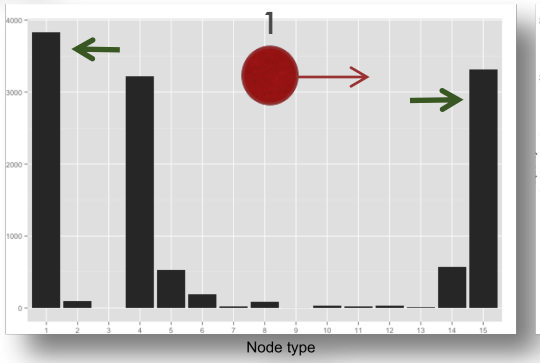
Tags

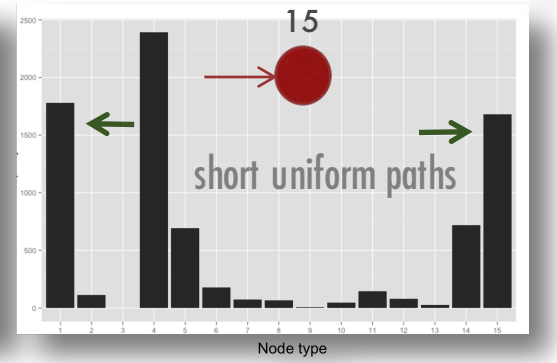


URIs



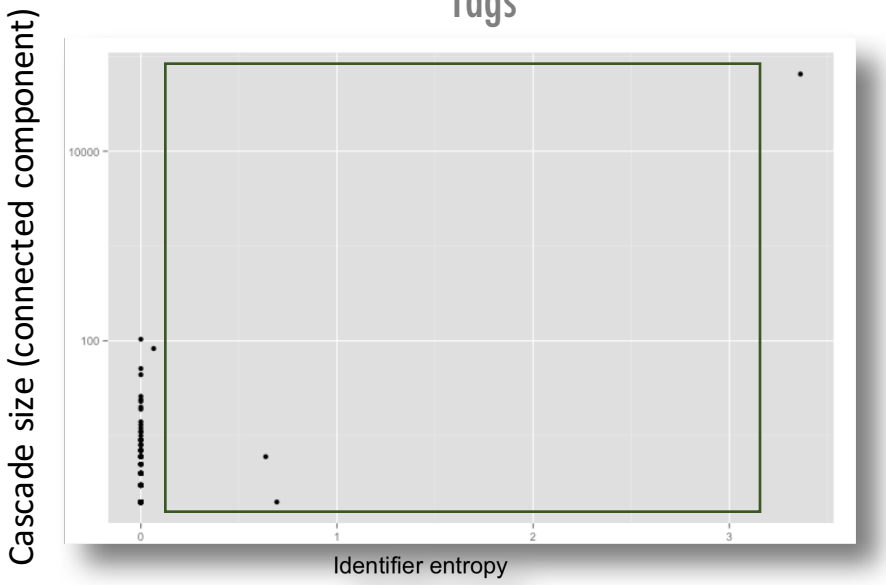
KID & APH



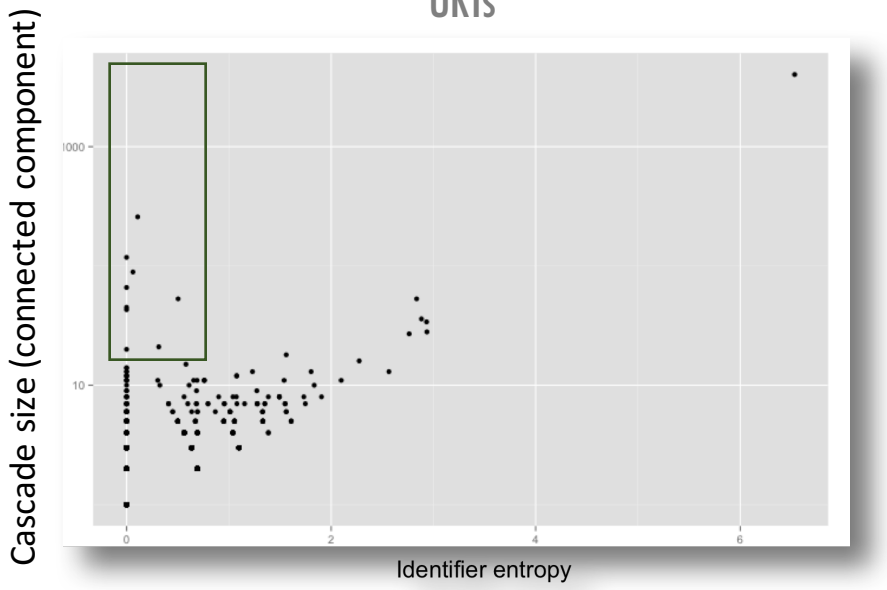


Analyzing low-level properties of the multiple states of a system that exist at the same time

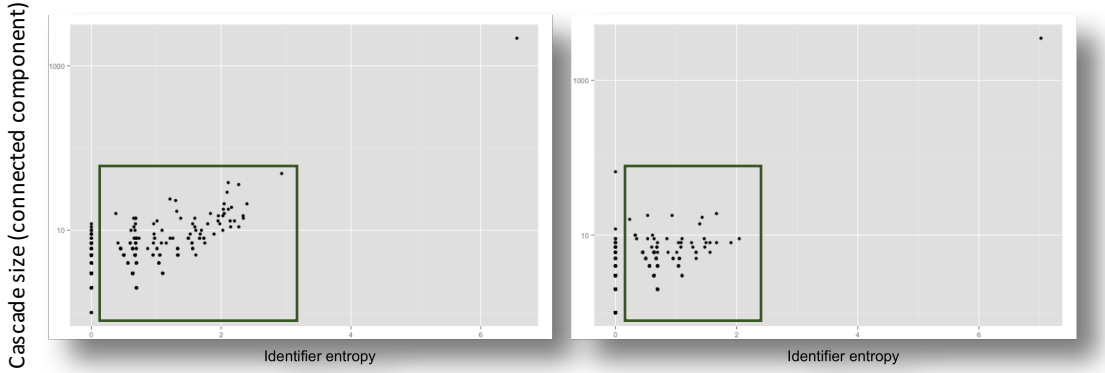
Tags



URIs



KID & APH

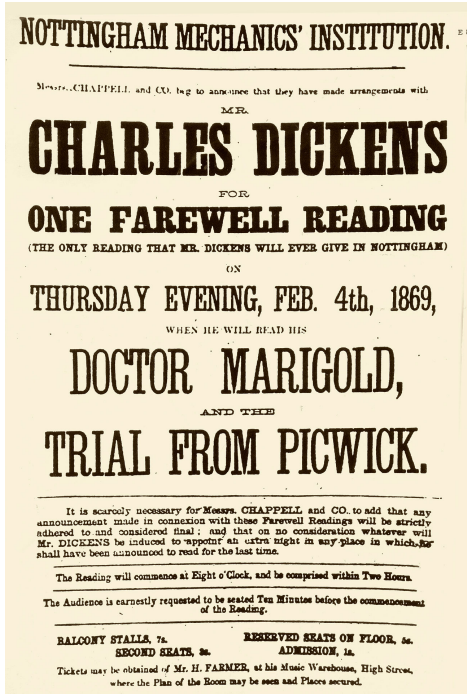


varying profiles randomness with growing cascade size

[8] Markus Luczak-Roesch, Ramine Tinati, Max van Kleek, and Nigel Shadbolt. 2015. From coincidence to purposeful flow? Properties of transcendental information cascades. In IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Paris, FR.

TICs applied to literary texts

From Dickens to Data Science



NOTTINGHAM MECHANICS' INSTITUTION.

Messrs. CHAPPELL and CO. beg to announce that they have made arrangements with

MR.

CHARLES DICKENS

FOR

ONE FAREWELL READING

(THE ONLY READING THAT MR. DICKENS WILL EVER GIVE IN NOTTINGHAM)

ON

THURSDAY EVENING, FEB. 4th, 1869,

WHEN HE WILL READ HIS

DOCTOR MARIGOLD,

AND THE

TRIAL FROM PICWICK.

It is scarcely necessary for Messrs. CHAPPELL and CO. to add that any announcement made in connexion with these Farewell Readings will be strictly adhered to and considered final; and that on no consideration whatever will Mr. DICKENS be induced to appear an extra night in any place in which he shall have been announced to read for the last time.

The Reading will commence at Eight o'Clock, and be comprised within Two Hours.

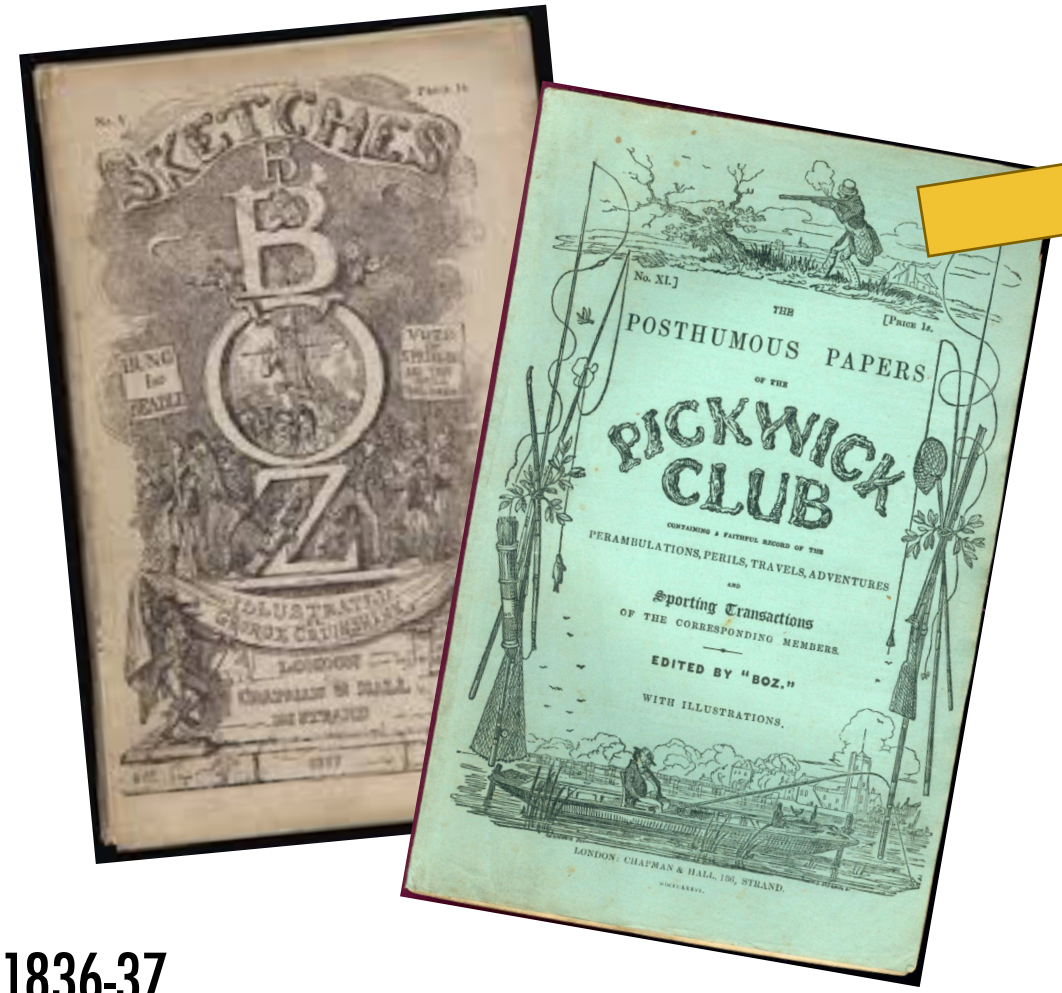
The Audience is earnestly requested to be seated Ten Minutes before the commencement of the Reading.

BALCONY STALLS, 7s. RESERVED SEATS ON FLOOR, 5s.
SECOND SEATS, 3s. ADMISSION, 1s.

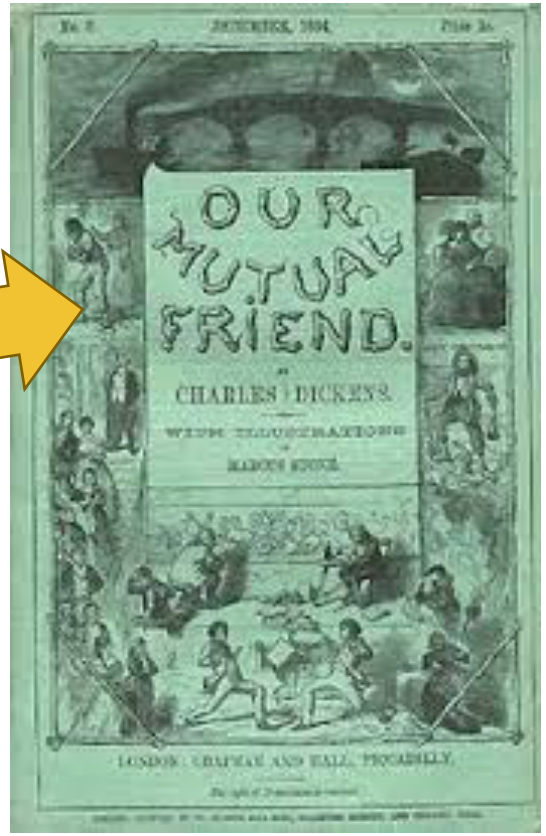
Tickets may be obtained of Mr. H. FARMER, at his Music Warehouse, High Street, where the Plan of the Room may be seen and Places secured.

All fourteen of Dickens's completed novels were published serially in weekly or monthly instalments.

From "Sketches" to Novels



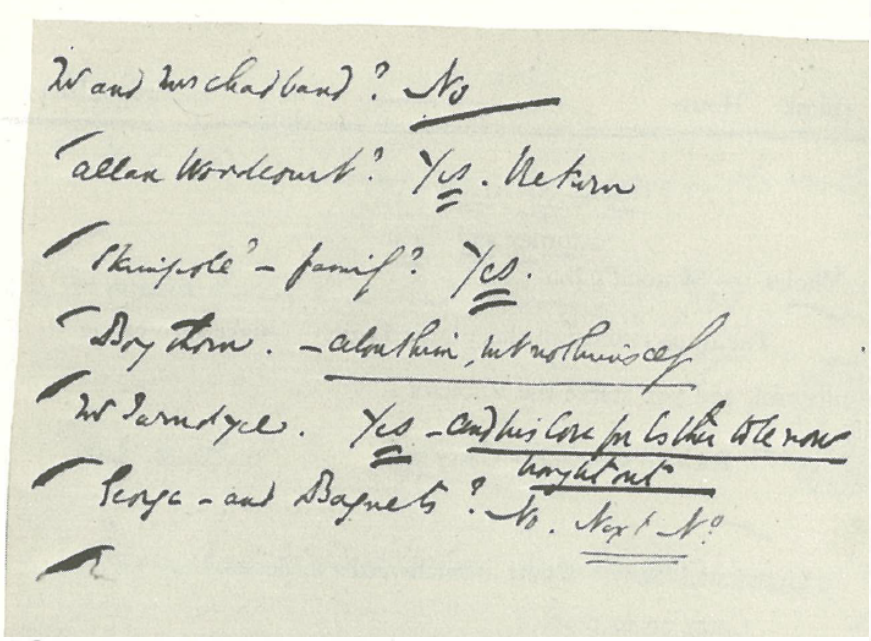
1836-37



1864-65

"I have endeavoured in the progress of this Tale, to resist the temptation of the current Monthly Number, and to keep a steadier eye upon the general purpose and design."
Preface to Martin Chuzzlewit (1844)

Managing Characters



Mr and Mrs Chadband? No

allan Woodcourt? Yes. Return

Skimpole? – family? Yes.

Boythorn. —About him, but not himself

Mr Jarndyce. Yes –And his love for Esther to be now brought out

George – and Bagnets? No. Next No.

Casby? Yes.

Flora, Mr F's Aunt? Yes

Pancks? Yes.

Miss Wade? No

Lagnier? No

Cavalletto? Carry through

The Meagleses? No

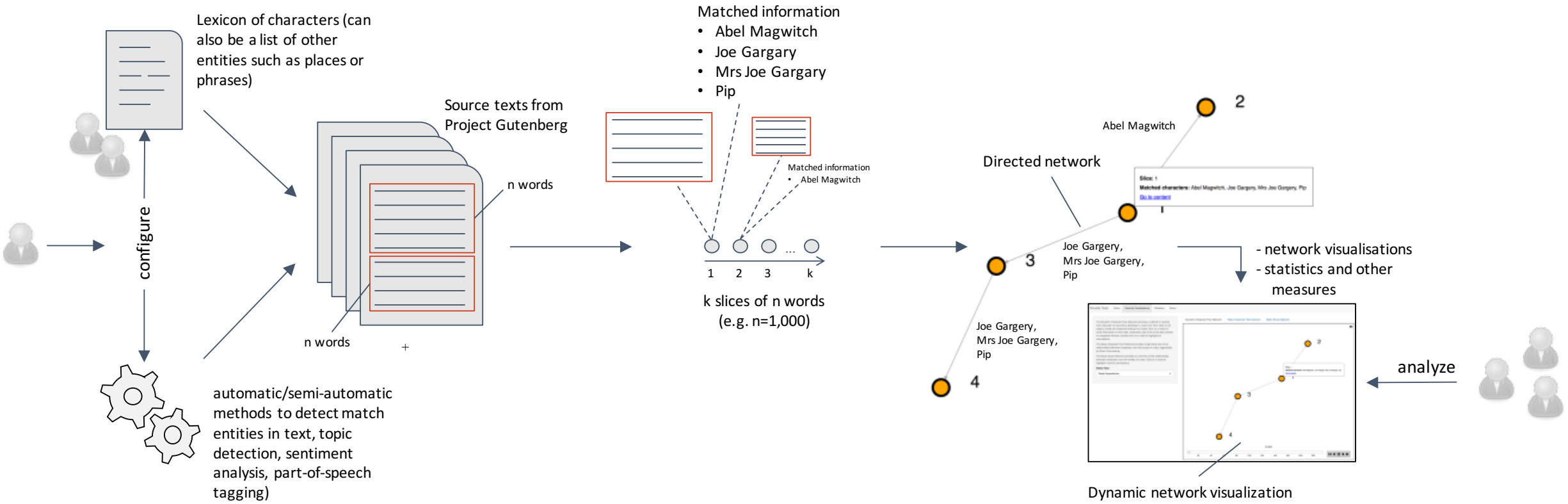
Pet and Gowran? No. (Next No.)

Daniel Doyce? Slightly

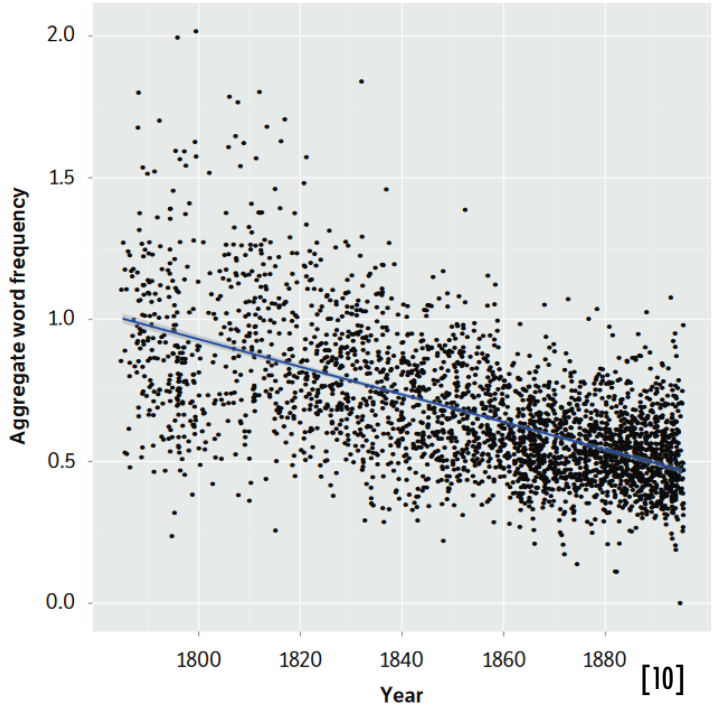
Plornish Family? No.

Dickens' Working Notes for His Novels. Edited by Harry Stone, U of Chicago Press, 1987.

Transcendental Information Cascades applied to English literature



"Not-So-Distant Reading"



[10]

Figure 8: Aggregate term frequencies of the abstract values fields combined in novels, 1785-1900.

Jockers, Matthew. Macroanalysis: Digital Methods & Literary History. 2013, p. 83.

Table 6.3. Classification accuracy and thirty-year generations

	1780 (%)	1790 (%)	1800 (%)	1810 (%)	1820 (%)	1830 (%)	1840 (%)	1850 (%)	1860 (%)	1870 (%)
1780s signal	92	8								
1790s signal	14	60	8	7	2	1	4		3	
1800s signal	14	16	39	13	5	2	5	8		
1810s signal	9	8	8	53	5	3	6	8		
1820s signal	9	1	7	29	32	11	7	2		1
1830s signal	11	4	10	7	7	31	13	4	11	
1840s signal			5	5	9	7	28	34	7	4
1850s signal					2		4	58	36	
1860s signal								16	71	14
1870s signal										100

Note: Columns represent the actual decades of composition, and rows represent the computed "signal" associated with the decade in the row label. Percentages in the cells show the proportion of attributions to the row signal in the column decade.

Distant Reading "tackles literary problems by scientific means: hypothesis-testing, computational modeling, quantitative analysis....understanding literature not by studying particular texts, but by aggregating and analyzing massive amounts of data."[11]

By working with large corpora of texts, "distant reading" methods explore how literary history works on a "macroscopic" level.

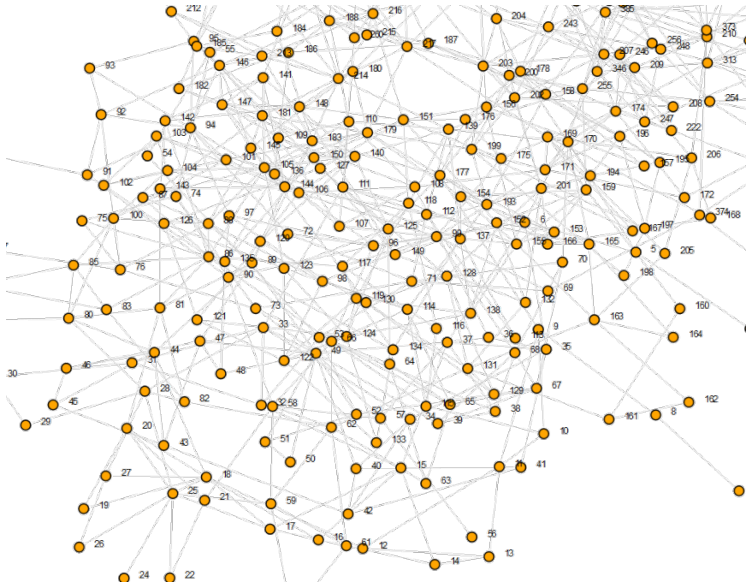
[10] Heuser, Ryan & Le-Khac, Long. A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method. Stanford Literary Lab. Pamphlet 4, May 2012.
[11] Shulz, Kathryn. "What is Distant Reading?" New York Times, 24 June 2011.

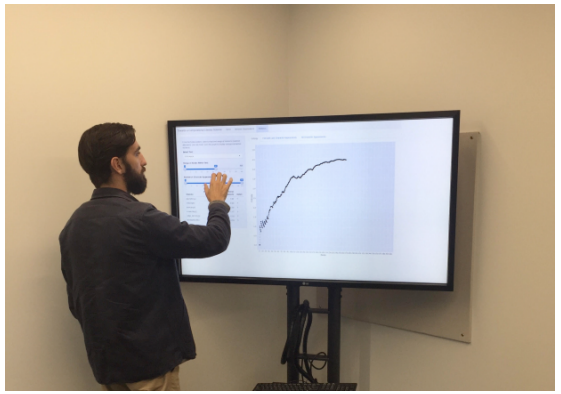
"Not-So-Distant Reading"

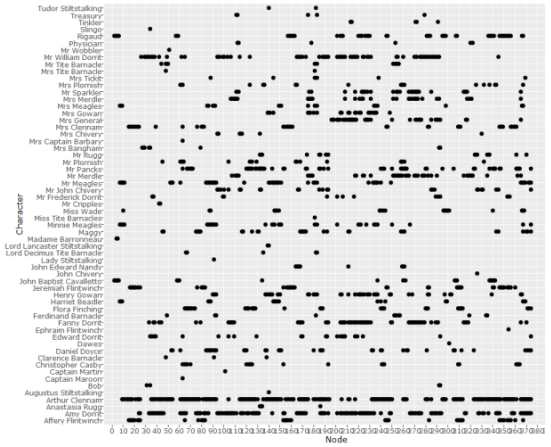
"The point, then, is that data-driven approaches are **not just doing the same thing better or at a larger scale**. They are **doing a different thing altogether**: interacting with the objects of the world. Traditional literary criticism, on the contrary, interacts with the past, with tradition. While one falsifies theories, the other develops from them. The figure of one is the data visualization. And the figure of the other is narrative."[12]

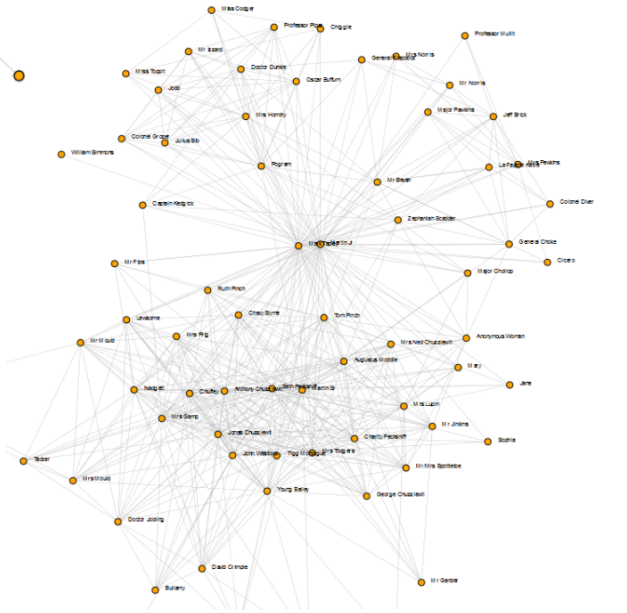
[12] Rosenthal, Jesse. "Narrative Against Data." Genre, vol. 50, no. 1, 2017, p. 9.

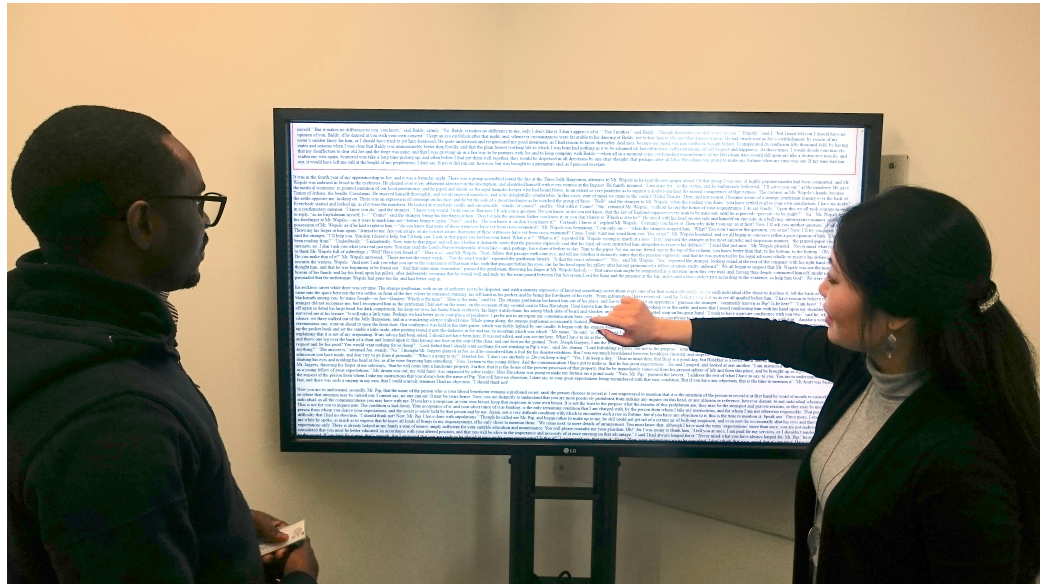
“Not-So-Distant Reading”





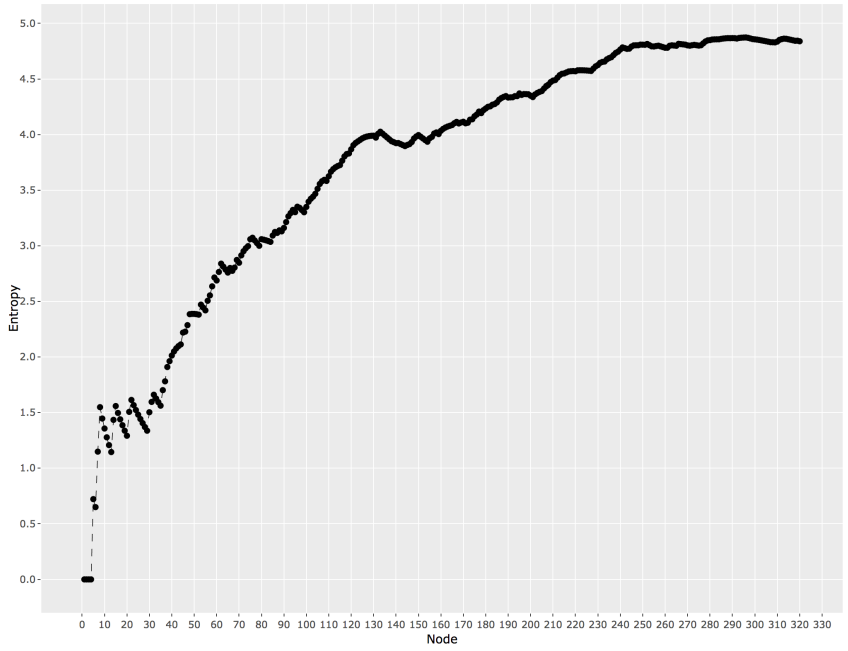




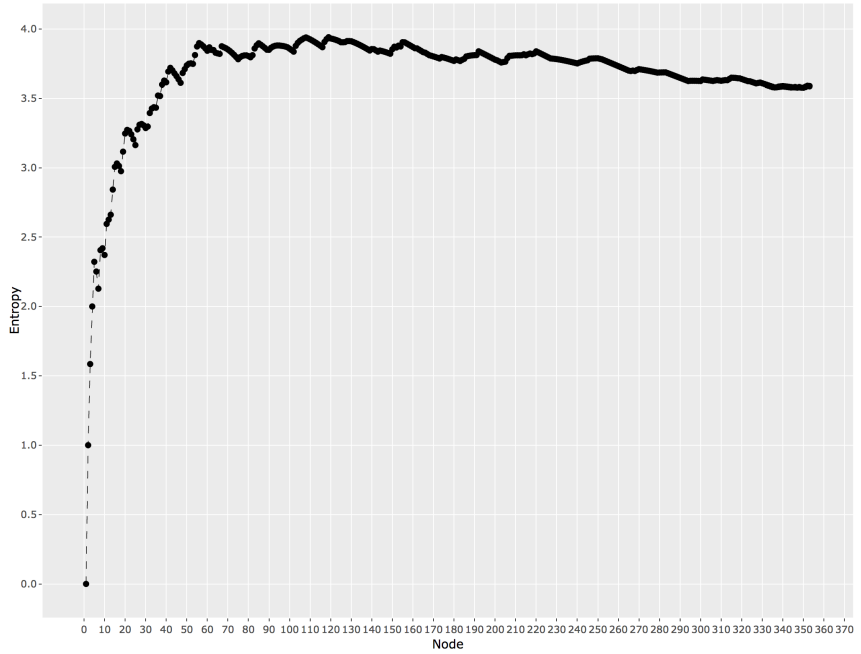


“Not-So-Distant Reading”

Pickwick Papers

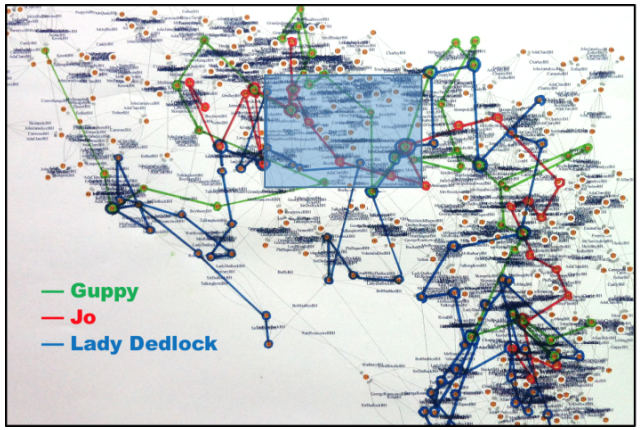


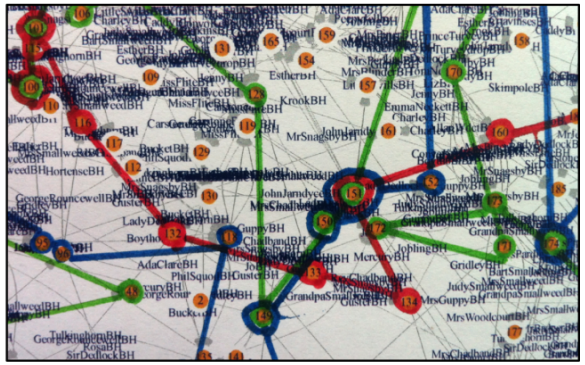
Our Mutual Friend

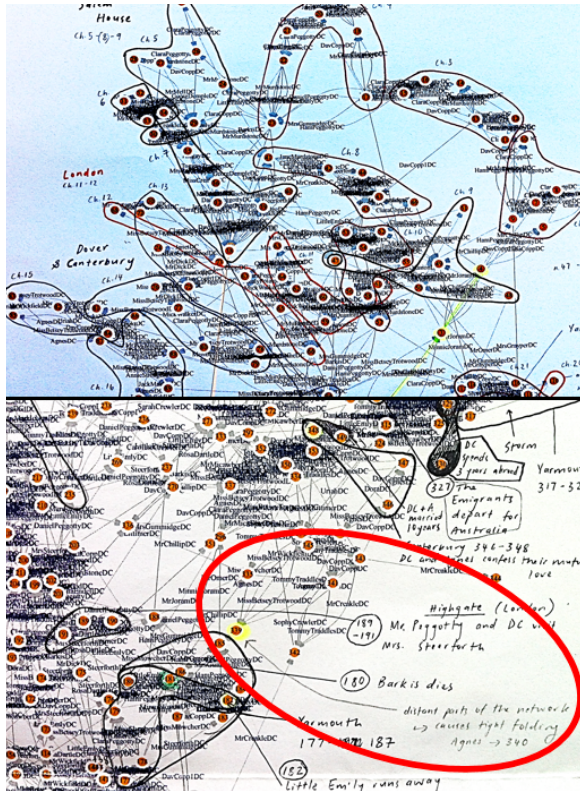


"Not-So-Distant Reading"

Although our method draws on foundational methodologies within digital humanities (hypothesis-testing, quantitative analysis), it explores the seeming contradiction between 'distant' and 'close' reading.







[13] Luczak-Roesch, M., Grener, A. and Fenton, E., 2018. Twenty Thousand Leagues Above the Book: An Interactive Visual Analytics Approach to Literature. In Proceedings of the International Conference on Supporting Group Work (GROUP), ACM. DOI: 10.1145/3148330.3154507.

[14] Luczak-Roesch, M., Grener, A. & Fenton, E. (2018). Not-so-distant reading: A dynamic network approach to literature. it - Information Technology, 60(1), pp. 29-40. Retrieved 1 Mar. 2018, from doi:10.1515/itit-2017-0023

Public demo

https://stia.shinyapps.io/tlit/

Further TIC applications

Discrete vs. continuous data

Q8-2 #transit at 18.6d

by user1 6 hours ago

Q8-2 #eclipse atat 18.6.

by user2 3 days ago

. . .

Q3-1 #transit #plain #gap

by user3 13 days ago

. . .

Q11-1 #transitingplanet #glitch

by user4 17 days ago

. . .

In the TCE list with P=31.2017, R=9.026 R_Earth.

by user2 17 days ago

Q5-1 #eclipse at 20.8, KID 9772531

by user2 17 days ago

Image source: https://en.wikipedia.org/wiki/Electroencephalography#/media/File:Spike-waves.png, CC BY-SA 2.0

EEG brain wave recordings

Image source: https://en.wikipedia.org/wiki/Electroencephalography#/media/File:Spike-waves.png, CC BY-SA 2.0

EEG brain wave recordings

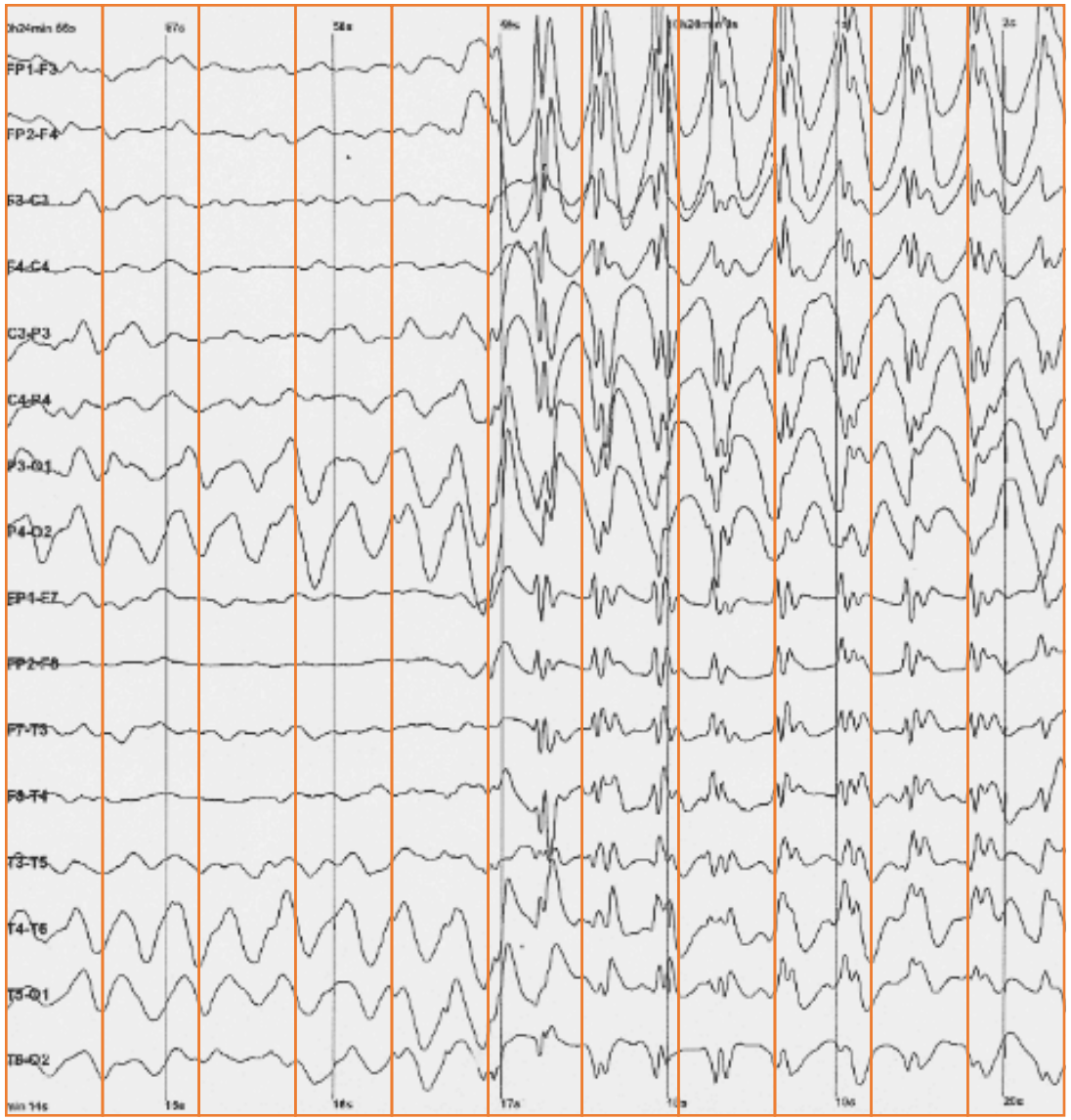
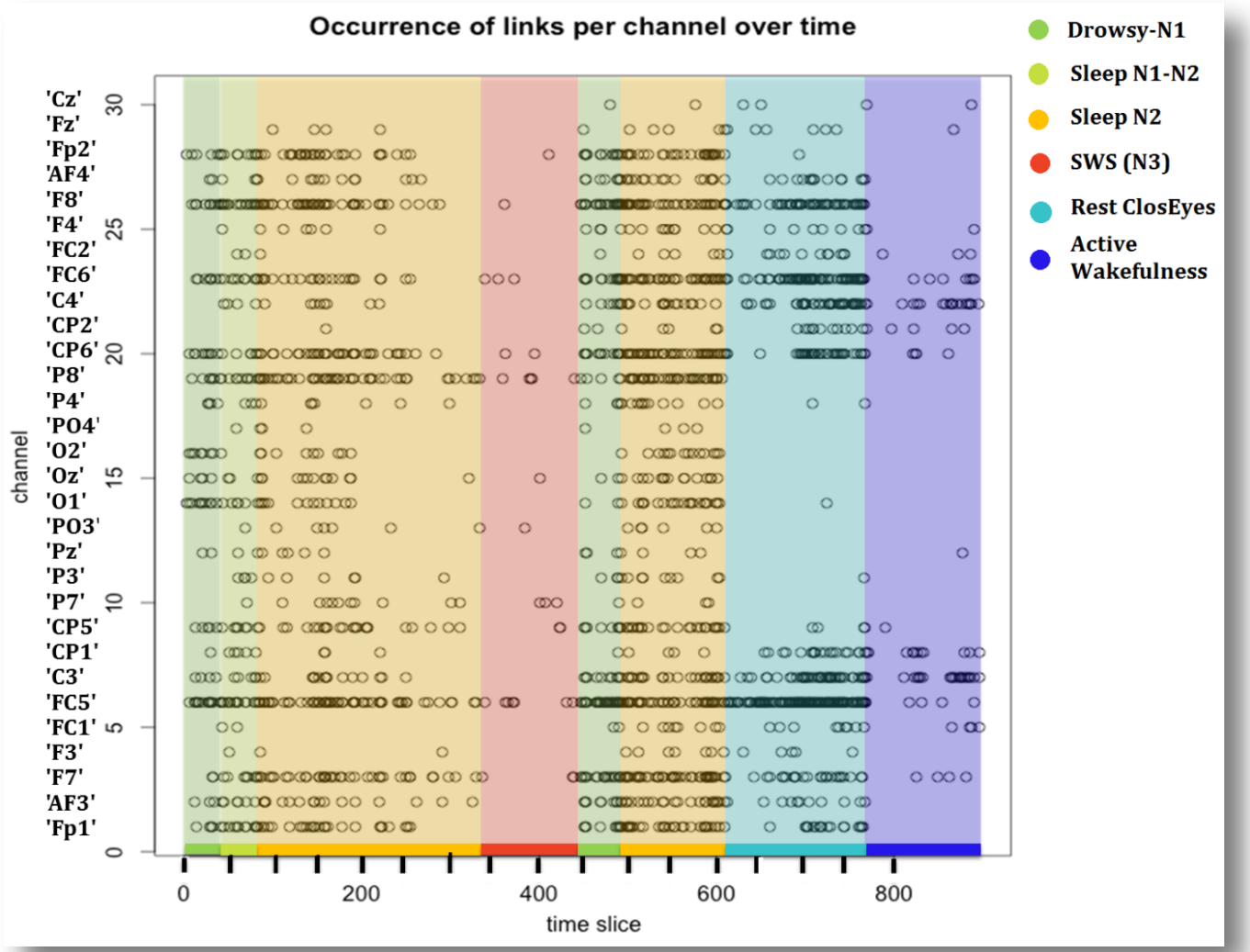


Image source: https://en.wikipedia.org/wiki/Electroencephalography#/media/File:Spike-waves.png, CC BY-SA 2.0

Linking based on similarity of spectral density (Euclidian distance)



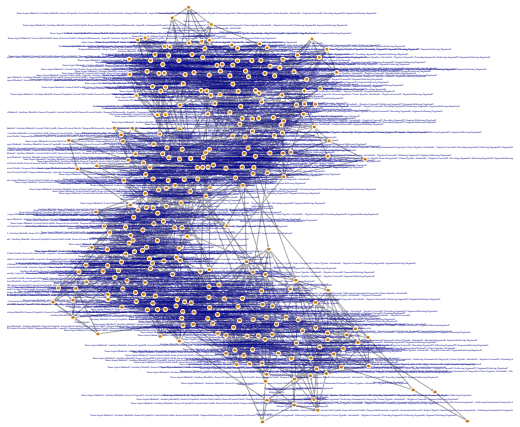
TICs applied to microlinguistic data

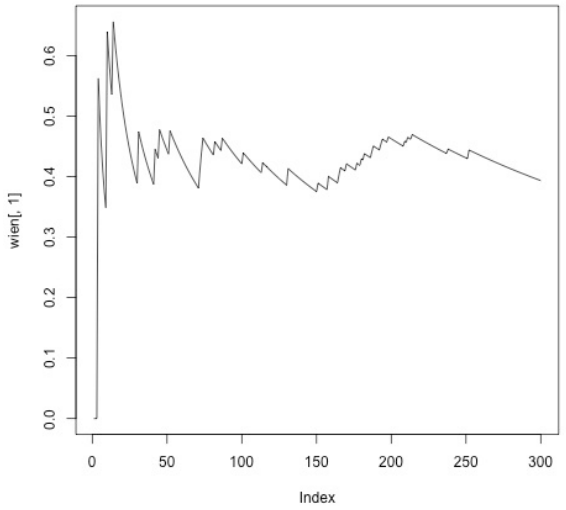
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
Tense-Aspect Marker	Auxiliary/Modal	Form of Copula	Lexical Verb	Lexical Verb Code	Form of Lexical Verb	Temporal Reference	Aspect	Grammatical Person	Subject Type	Sentence Type	Form of negation	Following Grammatical Category	Clause Type	Word Order in Clause	Adverbial	Negative Concord	Preceding Segment	Following Segment	Segment Following Negation	Community	Speaker Number	Speaker Code	Speaker Age	Number of people present in IV	Speaker sex	Urban sojourner?	Line Number	Context
/	/	/	have	h	b	n	c	2	s	d	/	/	w		-	/	/	/	/	H	1	1	87	4	F	N	11	when you HAVE children and you talk to them
/	/	/	say	<	b	n	h	2	s	d	/	/	r		-	/	V	0	/	H	1	1	87	4	F	N	11	and them don't hear what you SAY
/	/	/	talk	n	b	n	h	2	s	d	/	/	m		-	/	k	t	/	H	1	1	87	4	F	N	11	when you have children and you TALK to them
/	d	/	hear	s	b	n	h	6	o	n	n	/	w		-	/	/	/	h	H	1	1	87	4	F	N	11	and them don't HEAR what you say
d	z	/	give	f	b	n	c	6	o	d	/	/	t		-	/	/	/	/	H	1	1	87	4	F	N	12	them does GIVE you plenty trouble and plenty worry, s
/	m	/	go	g	b	n	h	2	s	n	t	/	m		-	/	/	/	g	H	1	1	87	4	F	N	13	these young people now you cannot GO with them now
/	/	0	talk	n	n	n	c	2	s	d	/	v	s		-	/	V	t	/	H	1	1	87	4	F	N	13	karn you Ø TALKING to young people
/	/	/	give	f	b	n	c	6	o	d	/	/	m		-	/	v	j	/	H	1	1	87	4	F	N	14	them GIVE you word so you can't go with them
/	/	/	have to	h	b	n	c	2	s	d	/	/	m		-	/	/	/	/	H	1	1	87	4	F	N	14	so you just HAVE to leave them
/	m	/	go	g	b	n	h	2	s	n	n	/	m		-	/	/	/	k	H	1	1	87	4	F	N	14	them give you word so you can't GO with them
/	/	/	bring	n	b	p	c	3	p	d	/	/	s		-	/	N	m	/	H	1	1	87	4	F	N	15	so karn I know my mother BRING me up with manners
/	/	/	know	k	b	p	c	1	s	d	/	/	m		-	/	V	m	/	H	1	1	87	4	F	N	15	so karn I KNOW my mother bring me up with manners
/	/	/	meet	n	b	n	h	2	s	d	/	/	p		-	/	t	V	/	H	1	1	87	4	F	N	16	"if you MEET anybody in the street, talk to them
/	/	/	say	<	b	n	p	3	s	d	/	/	m		-	/	V	0	/	H	1	1	87	4	F	N	16	she SAY, "
/	/	/	talk	n	b	n	h	2	s	m	/	/	a		-	/	k	t	/	H	1	1	87	4	F	N	16	"if you meet anybody in the street, TALK to them
/	/	0	do	d	n	n	c	6	o	d	/	v	p		-	/	m	d	/	H	1	1	87	4	F	N	16	"if them DOING anything wrong, talk to them
/	/	/	talk	n	b	n	h	6	o	d	/	/	a		-	/	k	t	/	H	1	1	87	4	F	N	17	"if them doing anything wrong, TALK to them
/	/	/	tell	n	b	p	h	6	o	d	/	/	r		-	/	l	m	/	H	1	1	87	4	F	N	17	and just like what them TELL me,
/	/	/	tell	n	b	p	h	6	p	d	/	/	m		-	/	l	m	/	H	1	1	87	4	F	N	17	my grandmother, my mother TELL me
/	s	0	be sick	b	0	n	c	3	p	d	/	a	s		-	/	V	s	/	H	1	1	87	4	F	N	18	you see me eye SICK you know
/	/	/	see	>	b	n	c	2	s	d	/	/	s		-	/	V	m	/	H	1	1	87	4	F	N	18	you SEE me eye sick you know
/	/	/	tell	n	b	n	h	1	s	d	/	/	m		-	/	l	D	/	H	1	1	87	4	F	N	18	I TELL them so
/	m	/	see	>	b	n	c	1	0	n	n	/	m		-	/	/	/	s	H	1	1	87	4	F	N	18	Can't SEE good out of me eye.
/	/	/	have	h	b	n	c	1	s	d	/	/	m		-	/	/	/	/	H	1	1	87	4	F	N	19	I HAVE a cataract.

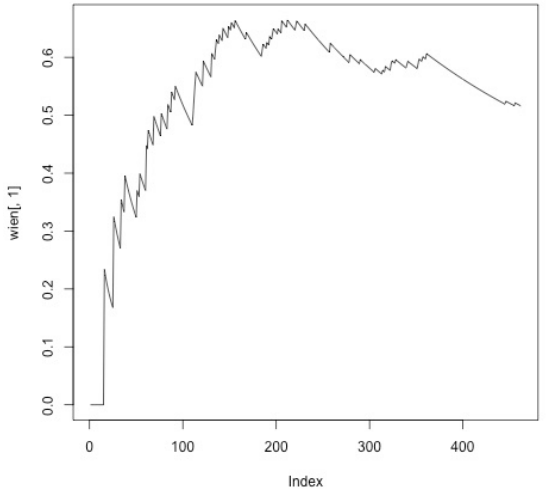
Walker, J.A. and Meyerhoff, M., 2006. Zero copula in the eastern Caribbean: Evidence from Bequia. American Speech, 81(2), pp.146-163.

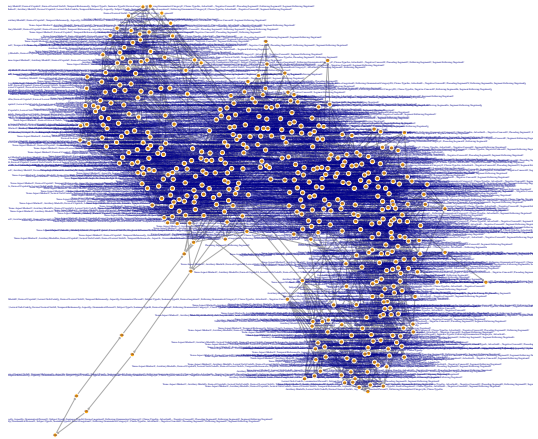
TICs applied to microlinguistic data

- bridging the micro and the macro level
- explain evolutionary dynamics of languages that could not by quantified so far









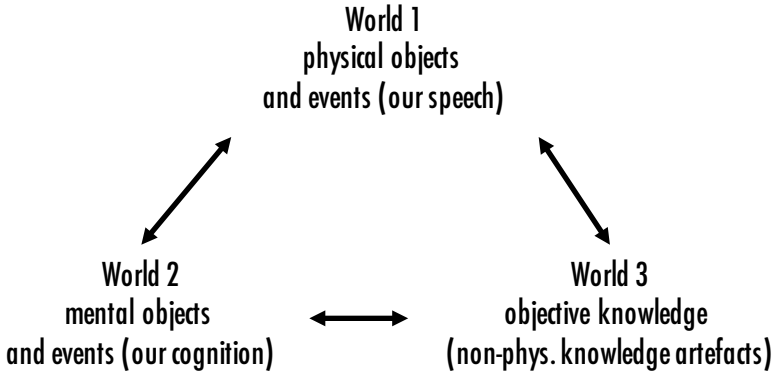
Tomasello, M., 2000. First steps toward a usage-based theory of language acquisition. Cognitive linguistics, 11(1/2), pp.61-82.
Tummers, J., Heylen, K. and Geeraerts, D., 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. Corpus Linguistics and Linguistic Theory, 1(2), pp.225-261.

Further application areas of TICs (ongoing projects)

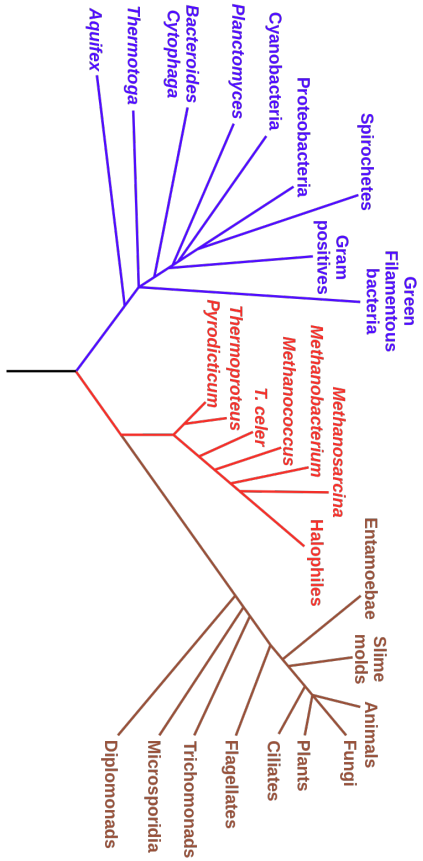
- Autobiographical memories from people with and without depression
- more English texts
- text in other languages than English

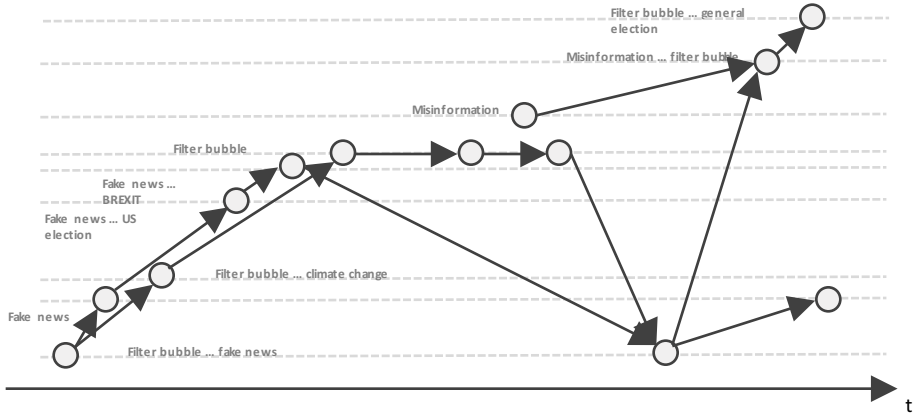
Back to the start

Evolutionary emergence and TICs

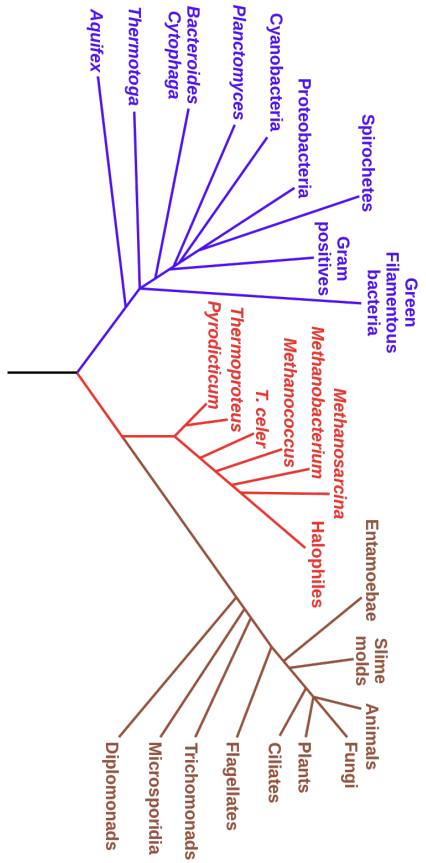


"a transcendental method in Kant's [3] sense of attempting to understand the conditions of knowledge itself"[2]

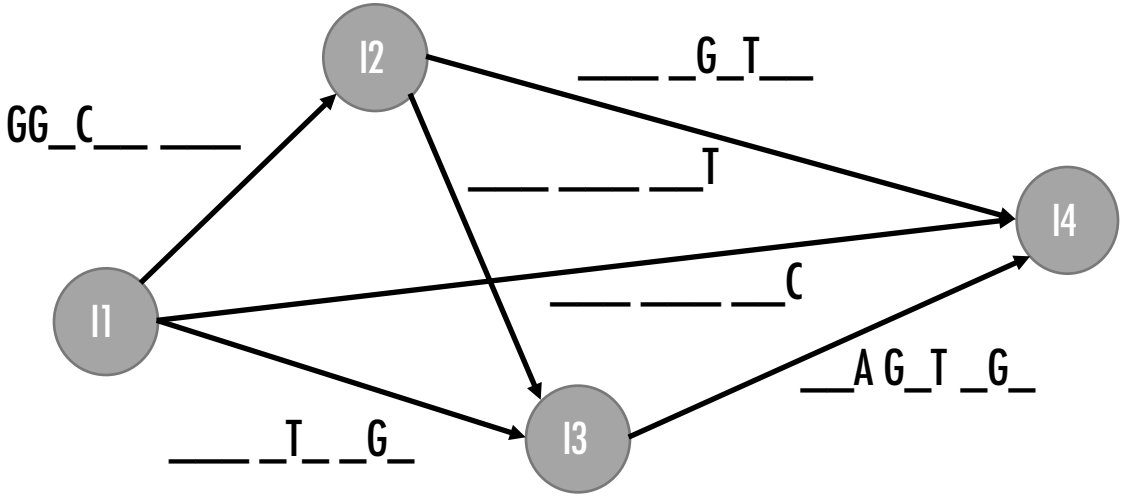


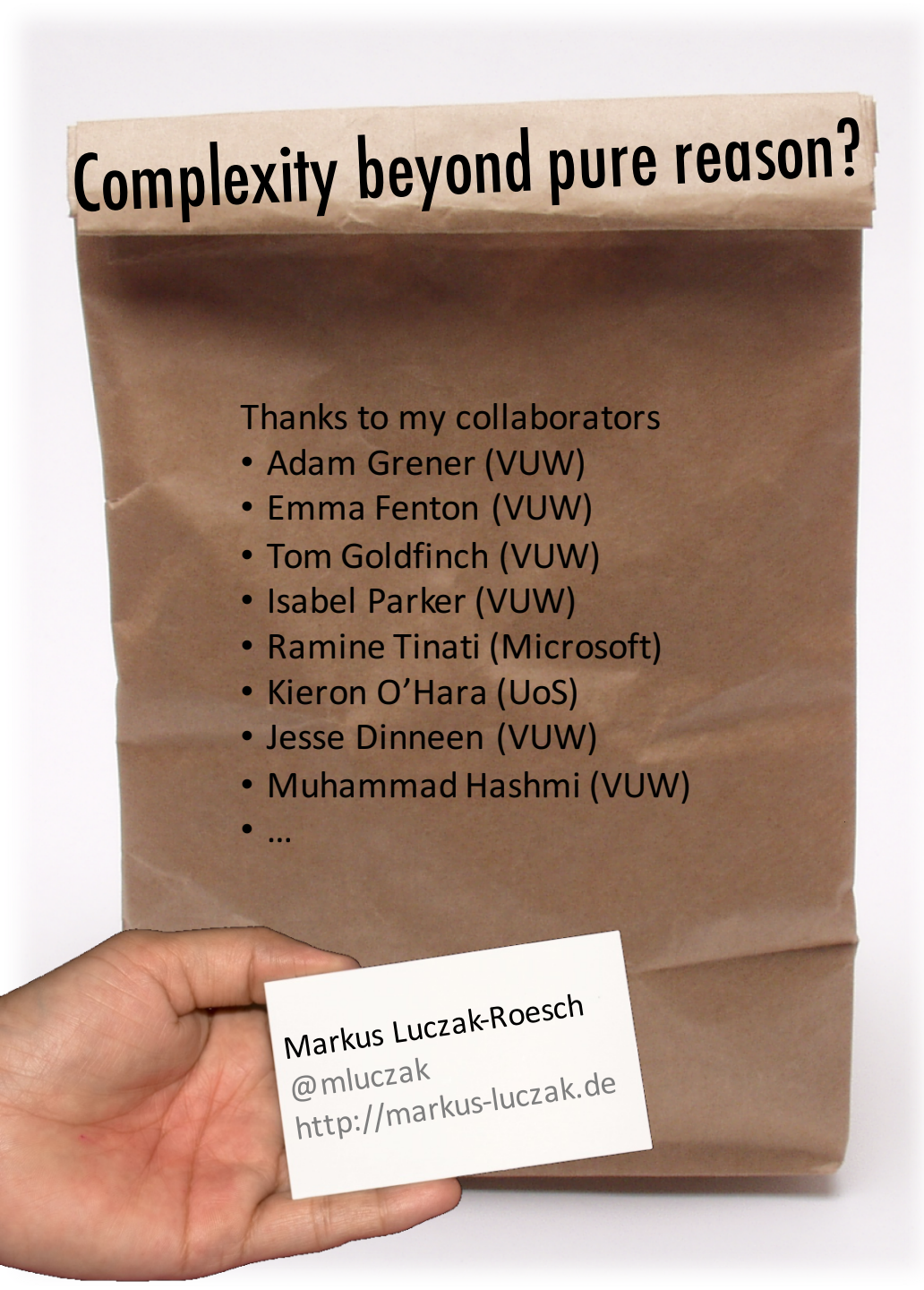


Evolutionary emergence and TICs



I1: {GGG CTA GGC}
I2: {GGT CGC TTT}
I3: {AAA GTT CGT}
I4: {CTA GGT TGC}





- The case for Transcendental Information Cascades
  - (some) philosophical foundations
  - contemporary applications

- Admitting my own dilemma
  - Are TICs a potential candidate model to study coincidences at scale?
  - Is it just good to know that there is this layer of complexity in almost any system or does it carry more general meaning for our understanding of emergence?