

```
2 import lxml.html
3 import json
4 import urllib
5
6 index = 'http://en.wikipedia.org/w/a
7
8 print 'Scraping ' + index + '...'
9 raw_json = scraperwiki.scrape(index)
10 python_json = json.loads(raw_json)
11 html = python_json['parse']['text']
12
13
14
15
16
17
18
19
20
```

post-social methods?

issues in live social research

Noortje Marres & Esther Weltevrede

DMI Summer School 2012

Documentation RUN

Console Data Sources Chat

Starting run ...

Scraping http://en.wikipedia.org/w/api.php?act

Finished: 0.711 seconds elapsed

runfinished

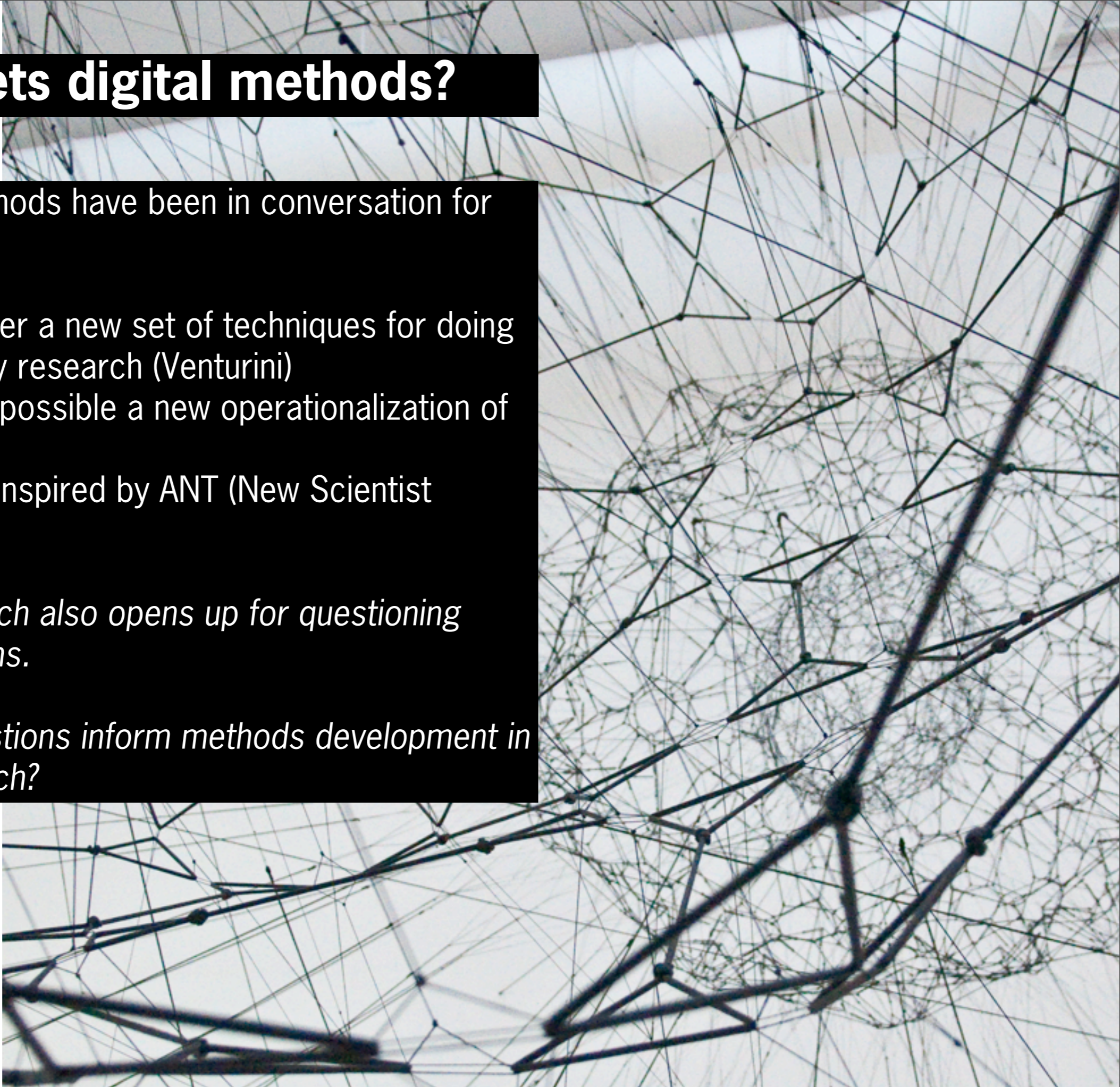
ANT meets digital methods?

ANT and digital methods have been in conversation for some time:

- Digital methods offer a new set of techniques for doing actor-network theory research (Venturini)
- Digitization makes possible a new operationalization of ANT (Latour, 2010)
- Digital methods is inspired by ANT (New Scientist review of Rogers)

Digital social research also opens up for questioning some key ANT claims.

How can these questions inform methods development in digital social research?



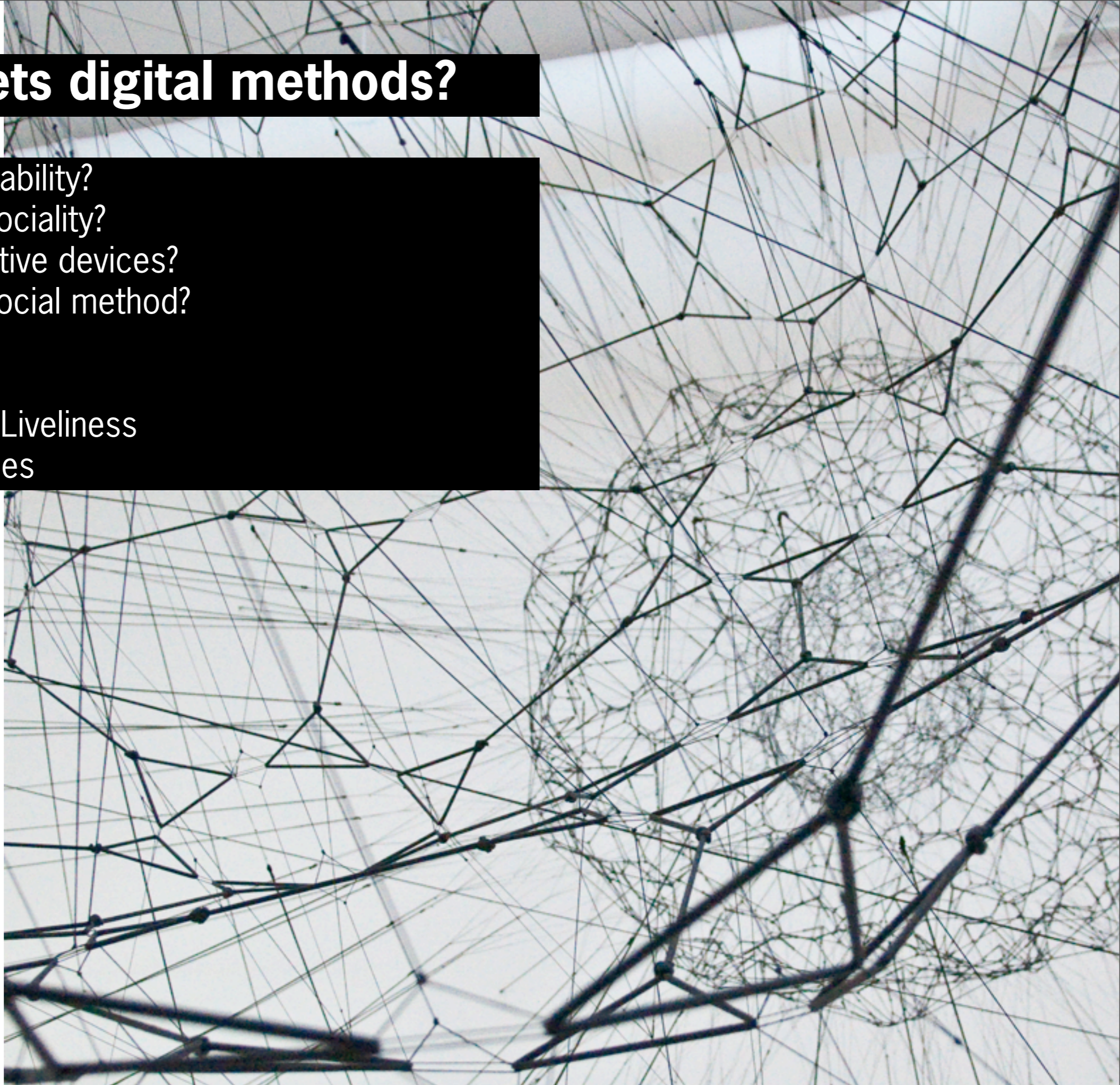
ANT meets digital methods?

Encounter 1: Traceability?
Encounter 2: Post-sociality?
Encounter 3: Formative devices?
Encounter 4: Post-social method?

Live Research

Case 1: Liveness & Liveliness

Case 2: Issue Lifelines



Encounter 1: Traceability?

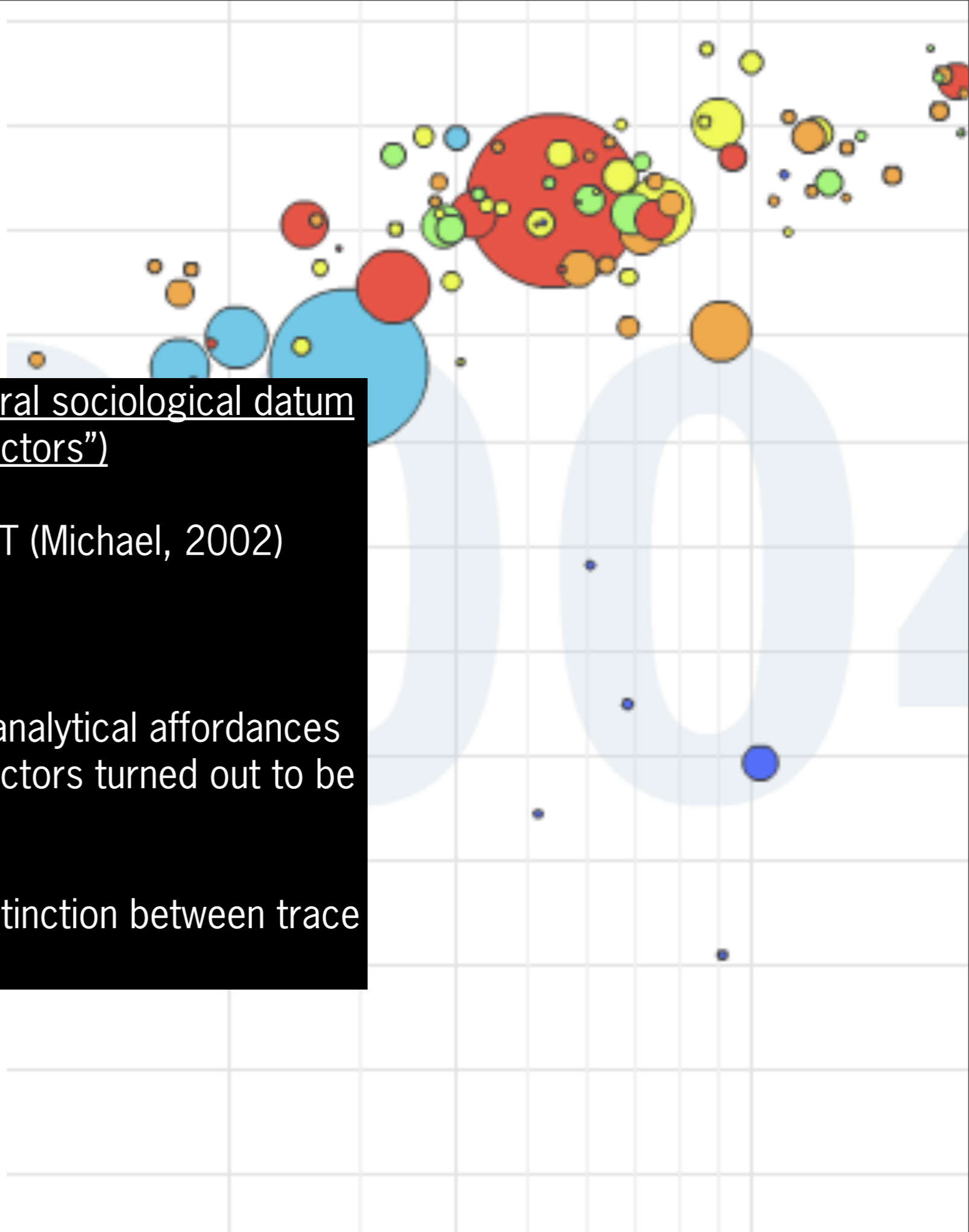
ANT proposed the 'trace' as the central sociological datum to 'replace' the variable ("follow the actors")

The critique of traceability in post-ANT (Michael, 2002)

...and the revenge of the Web:

tracing and tracking among the key analytical affordances of online networked media; here all actors turned out to be following the actors.

Or do digital media reposition the distinction between trace and variable?



Encounter 2: Post-sociality?

ANT proposed to move beyond “the social”

to trace the formation and transformation of dynamic, heterogeneous assemblages (associations of humans and non-humans)

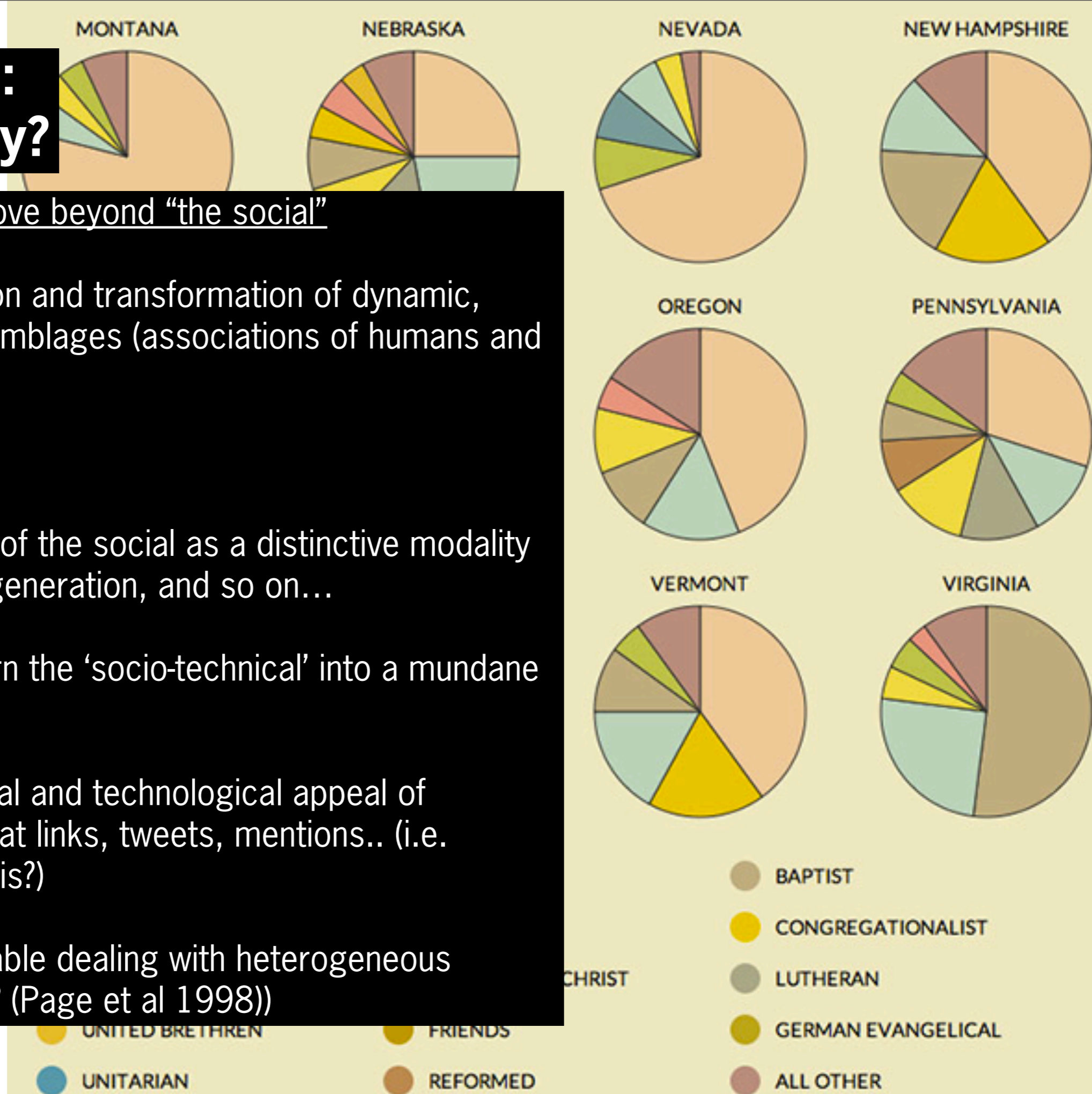
After the Web...:

1. the re-affirmation of the social as a distinctive modality of interaction, data-generation, and so on...

(Or did Facebook turn the ‘socio-technical’ into a mundane term?)

2. the methodological and technological appeal of reduction: look only at links, tweets, mentions.. (i.e. heterogeneity in crisis?)

(Or does the link enable dealing with heterogeneous information formats? (Page et al 1998))



Encounter 3: Formative Devices?

ANT attaches special importance to formats:

The progressive re-formatting of claims: From 'scientist Y suggests that X' to 'many believe X' to 'X' (see Latour Science in Action)

Digital devices like Google similarly rely on modifiers (links, mentions) to extract order from a 'panoply of elements' (Didier, 2009/2010),

After the Web:

ANT claims about formatting applied the metaphor of info-processing to knowledge – this doesn't quite work when knowledge becomes information processing?

Digital methods seek to *deploy* formatting effects for purposes of social research



Live Research

What is post-social method:

'Tracing the formation and transformation of heterogeneous assemblies.'

This project is newly relevant in the current online context:

- the rise of the real-time web (special interest in content dynamics)
- medium-specificity: using the formats of the mediums for research (to detect heterogeneous formation)

Liveness & Liveliness

Work in progress method development:

Instead of the most current, we are interested in the most active issues and terms

Instead of 'the most relevant result', we are interested in issue variation (over time)

(from search to research)

Tracing the variable composition of the object time:

Detecting liveliness?

ANT meets digital methods?

2 pilot studies

Dataset

(0.18 seconds)

Google data 'crisis'

Co-word machine input: top 100 titles.

Duration: 1 January 2012 - 15 February 2012

Objective: show co-word variation over time

Web

Crisis

ty for single homeless people. We are dedicated to ending
g life-changing services and campaigning for ...
What we do - Crisis Skylight London

[n help](#) · [Crisis](#)

ou can help

buying a **Crisis** Kit, you will provide the items that could turn
a homeless persons life around for good. Buy a **Crisis** Kit today ...

Books

More

Show search tools

[Volunteer](#) · [How you can help](#) · [Crisis](#)

[www.crisis.org.uk](#) > [How you can help](#)

Volunteer. We could not change so many lives without the hard work and commitment
of our volunteers. Over 10000 people give their time, energy and skills to ...

[Crisis Hidden Gigs](#) · [Art and music](#) · [Crisis](#)

[www.crisis.org.uk](#) > ... > [Art and music](#) > [Crisis Hidden Gigs](#)

Crisis held two Hidden Gigs with Coldplay in December 2011. The events, one in
Newcastle and the other in Liverpool, raised £250000 for **Crisis**. Watch this ...

[News for crisis](#)



[Central bank group: Crisis fixes still needed](#)

[Fox News](#) - 1 hour ago

Governments, banks and households struggling with too much debt are
dragging down the world's economy and more needs to be done to ...

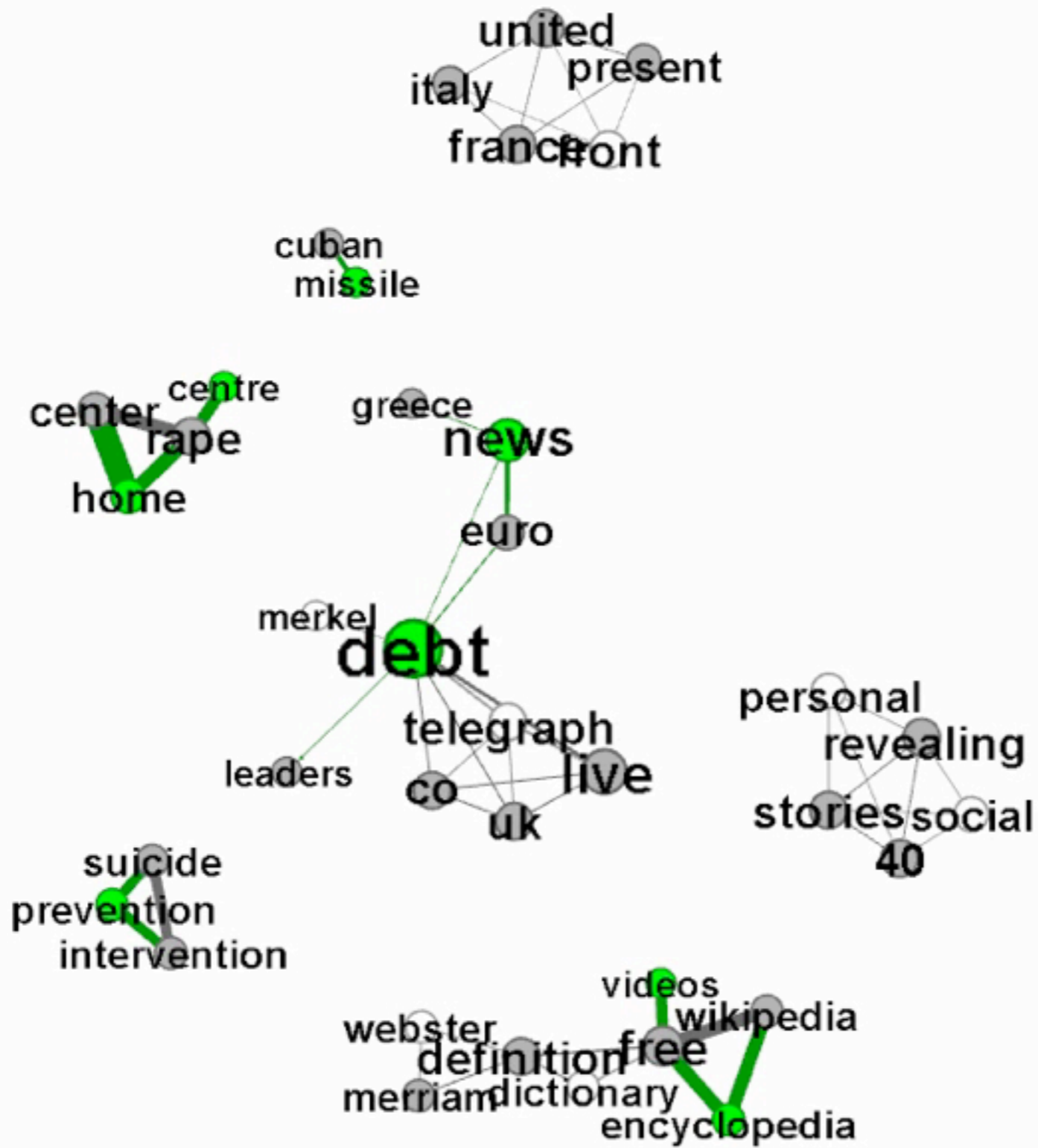
New York Ti...

[Greek crisis: An odyssey seen through ancient myth](#)

[San Francisco Chronicle](#) - 3 hours ago

[Ashraf faces hot summer of crisis](#)

[Financial Times](#) - 1 hour ago



A- A- Arial-BoldMT, 32

2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Issue Lifelines

Instead of Twitter 'trending topics', we are interested in using Twitter to detect issue variation over time

Question: How lively is "climate action" on Twitter?

Objective: Use the co-word machine as Issue Biographer.

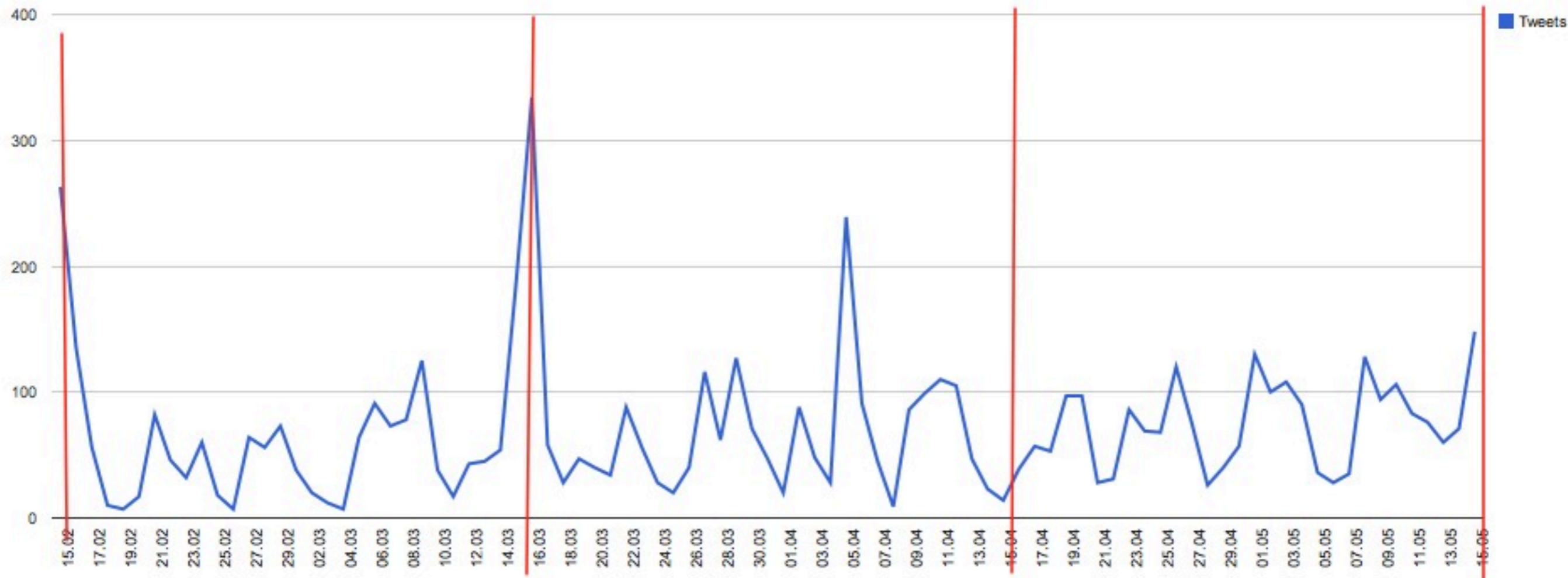
Dataset

Twitter data “Climate Action”.

Co-word machine input: Tweets.

Focus on three intervals: 15Feb-14Mar, 15Mar-14Apr, 15Apr-14May.

Objective: profile the co-word relations of key hashtags.



Machine settings for the lifeline tracker

Keyword profiling.

Determines changes in associational profile per interval.

Identifies degree of connectivity & change.

New words connecting, words disappearing & current connections.

Handwritten notes on a whiteboard:

- 1501a
- Outliers weirdos? innovation
- tired Slogans
- everyday language
- $\frac{C_{ij}^2}{C_i C_j}$
- 7h

ANAL: KEYWORD

	FREQ	DEGREE	AP VAK	WORDS IN	WORDS OUT	AP KW
SLICE 1	20	70				WORD 1, WORD 2
SLICE 2	70	3	0.9	WORD 1, 2	WORD X ...	WORD 1, 2
SLICE 3	8	4	0.8	WORD 1, 2	WORD X ...	WORD 1, 2
...						

Diagram:

```

    graph TD
      INPUT --> KW[KEYWORDS]
      KW --> CM[COOL MACHINE]
      CM --> OUT[OUTPUT]
  
```

Vertical calculations on the right:

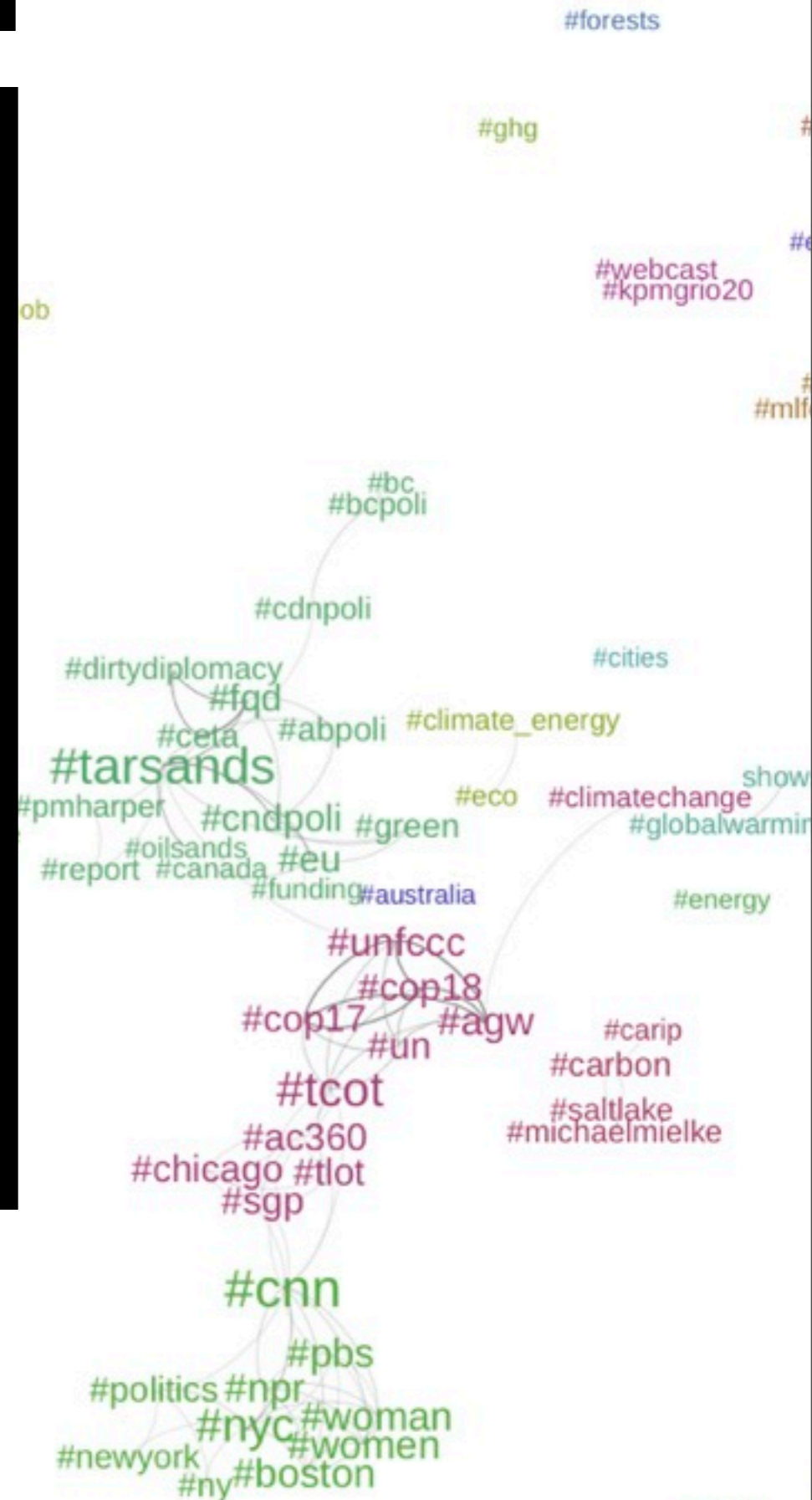
$$\begin{array}{r} 25 \\ 450 \\ 9 \\ \hline 75 \end{array}$$

Hashtag profiling over time intervals

#tarsands
#eu
#cdnpoli
#agw
#green
#fqd
#cndpoli
#politics
#unfccc
#ceta
#health
#flooding

#jobs
#san
#intern
#job
#cop18
#cop17
#climatechange
#energy
#globalwarming
#environment
#policy
#losangeles
#nonprofit

#wind #nuclear
#solar
#coal #germany
#coalwatch



Associational profiling

	A	B	C	D	E	F		
1	key	time	degree	similarity	associational profile		in	out
2	#environment	0	2		#politics (1), #climate (1)			
3	#environment	1	5	0	#ows (1), #clean (1), #climatechange (1), #fish (2), #wind (#wind (2), #fish (2), #climatechange (2), #clean (2), #climate (2), #po		
4	#environment	2	10	0	#esg (1), #innovation (1), #energy (1), #sustain (1), #carbo	#amherstma (2), #sustainability (2), #climate (2), #g	#wind (2), #fish (
5	#tarsands	0	14		#fqd (6), #ceta (3), #climate (8), #pmharper (2), #report (2), #unfccc (1), #law (1), #cdnpoli (2), #abpoli (3), #cndpoli (7), #oilsan			
6	#tarsands	1	5	0.028171808	#ngp (1), #unfccc (1), #agw (1), #science (1), #climatechan	#climatechange (2), #science (2), #agw (2), #ngp (2)	#cndpoli (2), #ab	
7	#job	0	1		#missouri (1)			
8	#job	1	6	0	#ca (3), #san (4), #california (3), #jobs (4), #intern (4), #act	#intern (2), #action (2), #jobs (2), #california (2), #s	#missouri (2)	
9	#energy	0	1		#climate (1)			
10	#energy	1	6	0.40824829	#earthconservation (1), #data (1), #eco (1), #politicians (1)	#clean (2), #politicians (2), #eco (2), #earthconserv	#climate (0)	
11	#energy	2	9	0	#esg (1), #environment (1), #innovation (1), #sustain	, #globalwarming (2), #climatechange (2), #argentina	#climate (2), #cl	
12	#green	0	8		#cdnpoli (1), #climate (4), #agw (1), #bcpoli (1), #cndpoli (3), #eu (3), #tarsands (3), #fail (1)			
13	#green	1	3	0	#footprint (1), #mobile (1), #ict (1)	#ict (2), #mobile (2), #footprint (2)	#eu (2), #tarsand	
14	#green	2	19	0	#policy (1), #minority (1), #action (1), #blog (1), #carbonta	#governance (2), #econdev (2), #regions (2), #globa	#ict (2), #mobile	
15	#agw	0	5		#cdnpoli (1), #climate (1), #green (1), #bcpoli (1), #climatechange (1)			
16	#agw	1	15	0.076980036	#cop18 (12), #cop17 (12), #unfccc (13), #un (5), #tcot (4),	#carbon (2), #vic (2), #ngp (2), #tarsands (2), #scien	#bcpoli (2), #cdn	
17	#globalwarming	0	2		#climate (2), shows (2)			
18	#globalwarming	1	4	0.534522484	#climate (1), shows (1), #eco (1), #climatechange (2)	#climatechange (2), #eco (2), shows (-0.666666666	shows (0.666666	
19	#globalwarming	2	6	0.507092553	#climatechange (4), #oil (3), #energy (3), #argentina (3), #	#action (2), #argentina (2), #energy (2), #oil (2), #cl	#eco (2), shows	
20	#unfccc	0	6		#law (1), #cdnpoli (1), #tarsands (1), #pmharper (1), #climate (1), #report (1)			
21	#unfccc	1	10	0.018202892	#cop18 (12), #cop17 (12), #agw (13), #un (5), #tcot (4), #e	#ngp (2), #science (2), #climatechange (2), #eu (2),	#climate (2), #pr	
22	#unfccc	2	3	0	#climate (4), #newzealand (4), #australia (4)	#australia (2), #newzealand (2), #climate (2)	#ngp (2), #tarsan	
23	#climatechange	0	1		#agw (1)			
24	#climatechange	1	23	0.169030851	#climate (1), #congress (1), #health (1), #flooding (1), #clir	#ngp (2), #futurewewant (2), #economics (2), #oilsa	#agw (0)	
25	#climatechange	2	12	0.230021853	#stupid (1), #riverrally (1), #carbontax (1), #green (1), #cit	#sydney (2), #rpcd (2), #oil (2), #energy (2), #stupid	#unfccc (2), #ngp	
26	#cop18	1	6		#cop17 (12), #unfccc (12), #agw (12), #un (5), #tcot (4), #eu (1)			
27	#cop17	1	6		#cop18 (12), #unfccc (12), #agw (12), #un (5), #tcot (4), #eu (1)			
28	#health	1	7		#climate (434), #flooding (433), #cleanairact (10), #congress (3), #climatechange (1), #ows (1), #eco (1)			
29	#health	2	6	0.998238141	#climate (28), #flooding (27), #cleanairact (2), #ncds (1), #	#security (2), #pogoblog (2), #ncds (2), #cleanairact	#ows (2), #eco (2)	
30	#flooding	1	6		#climate (433), #health (433), #cleanairact (10), #congress (3), #climatechange (1), #ows (1)			
31	#flooding	2	4	0.998995767	#climate (27), #health (27), #cleanairact (2), #pogoblog (1)	#pogoblog (2), #cleanairact (-1.3333333333333333), #	#ows (2), #clima	
32	#san	1	6		#ca (3), #california (3), #jobs (4), #intern (4), #job (4), #action (4)			
33	#jobs	1	7		#ca (3), #san (4), #california (3), #intern (4), #job (4), #action (4), #intjobs (1)			
34	#intern	1	6		#ca (3), #san (4), #california (3), #jobs (4), #job (4), #action (4)			
35	#policy	2	4		#minority (1), #action (1), #blog (1), #green (1)			

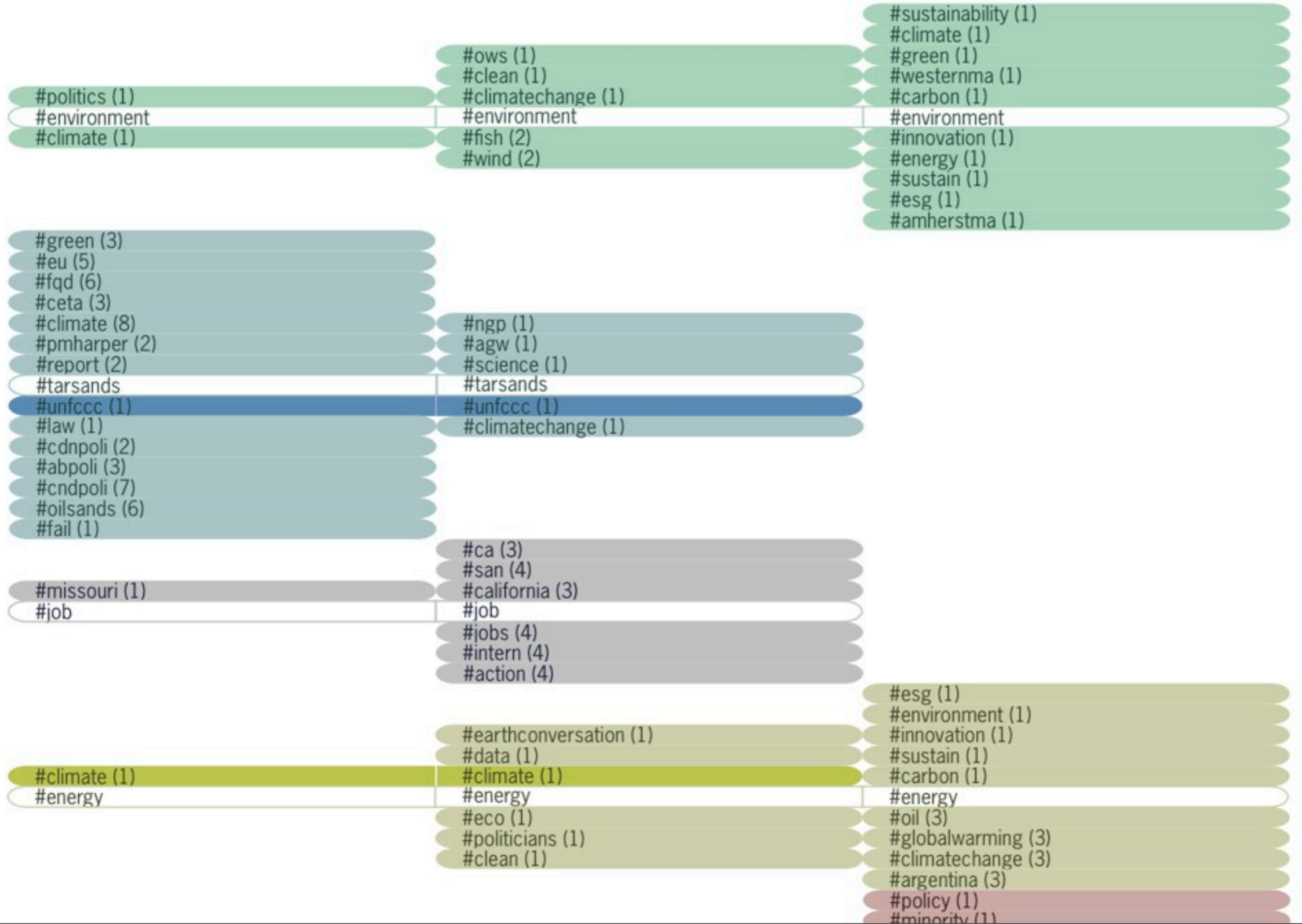
Climate Action Associational Profiles

15-02-2012

15-03-2012

15-04-2012

15-05-2012



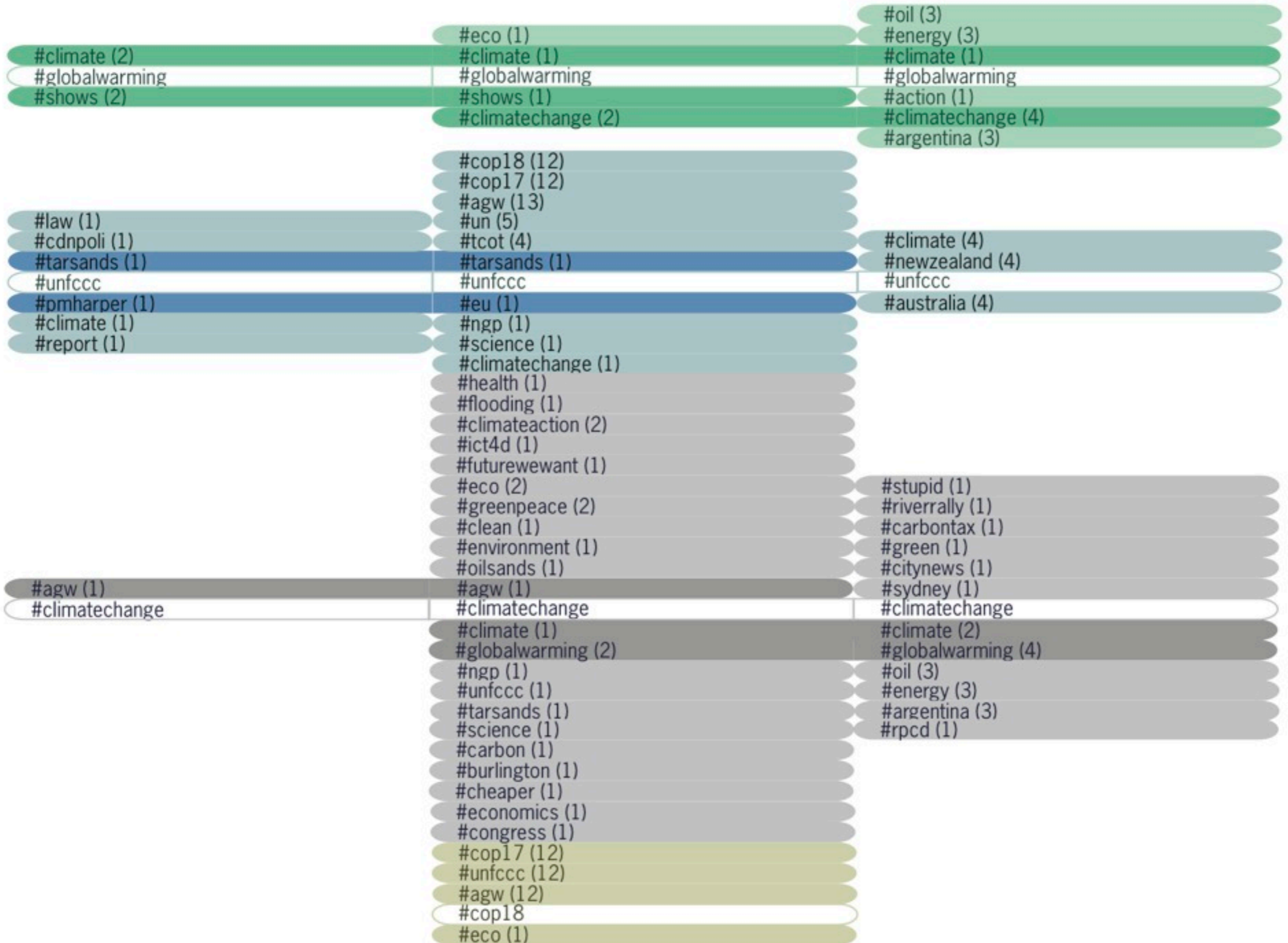
Climate Action Associational Profiles

15-02-2012

15-03-2012

15-04-2012

15-05-2012



hashtag actor profiling #green

Interval I

socialopine.com (2)
reuters.com (1) secure2.edf.org (1) thegreenmiles.blogspot.com (1) theithacajournal.com (1)

publicserviceeurope.com (2) redgreenandblue.org (2)
canadians.org (1) EarthAdapt.com (1) ericsson.com (1) greenallianceblog.org.uk (1) jobs.greenbiz.com (1)

motherjones.com (2)
ping.fm (1) treehugger.com (1) unep.org (1)

Interval II

motherjones.com (2)
EarthAdapt.com (1) ericsson.com (1) greenallianceblog.org.uk (1) jobs.greenbiz.com (1)

redgreenandblue.org (2) socialopine.com (2)
theithacajournal.com (1) treehugger.com (1)

Interval III

guardian.co.uk (5)
govtoday.co.uk (1) huffingtonpost.com (1) isustainableearth.com (1) jobs.greenbiz.com (1)

news.xydo.com (1) portlandtribune.com (1) straight.com (1) studiocity.patch.com (1) taketochange.com (2) thehill.com (1) tibidy.com (1) uanews.org (2) mb.com.ph (2)

beautiful-planet.info (1) broadbandcommission.org (1) cityofevanston.org (2) climateactionprogramme.org (1) climatesummit.org.au (1)

2degreesnetwork.com (1) EarthAdapt.com (1)

Course

The Link
 The Website
 The Engine
 The Spheres
 The Webs
 Post-demographics
 Networked Content

Digital Methods

Summer 2012
 About
 FAQ
 Course
 Tools
 Projects by Theme
 Research Protocols
 Summer School
 Winter School
 Papers and Publications
 Blog

Co-word Lifelines

tbc

Members

Noortje, Carolin...

Introduction

Co- word analysis is a method of content analysis that makes textual data amendable to network analysis. Co- word analysis works by detecting word pairs or 'co-occurrences': words that occur in each others's vicinity (word distance of say 3, 4 or 5 words), and proceeds by weighing these word relations (assigning them values in terms of relative strength and quantity of connections). Combining these measures, co-word analysis in principle makes it possible to detect the emergence of new significant themes, so called 'pockets of innovation' and, more generally, changes in thematic associations over time (Callon et al, 1983; Danowski, 2009).

Co-word analysis has been included in a variety of digital research and online visualisation tools (eg [Infomous](#); [Wordij](#)). The form of analysis gained special relevance in the context of digital research and allows to address the phenomenon of 'happening content': the dynamic nature of online information ecologies as well as issues which raises the question of how to detect, valuate and display significant variation in and of data. Fresh terms may arise, new reports are released, and events occur that can shift the relations among actors and issues, and so the question arises how best to capture, analyse and render such change. In this sense, co-word analysis allows to study 'happening content': it offers a way to detect fluctuations in networked content, as an alternative, or complement, to more established 'methods of the medium', such as inlink count or the measurement of word frequency,





Thank you.