Twitter rape threats and the **D**iscourse **o**f **O**nline **M**isogyny (DOOM)

Mark McGlashan @Mark_McGlashan

Claire Hardaker @DrClaireH (in absentia)
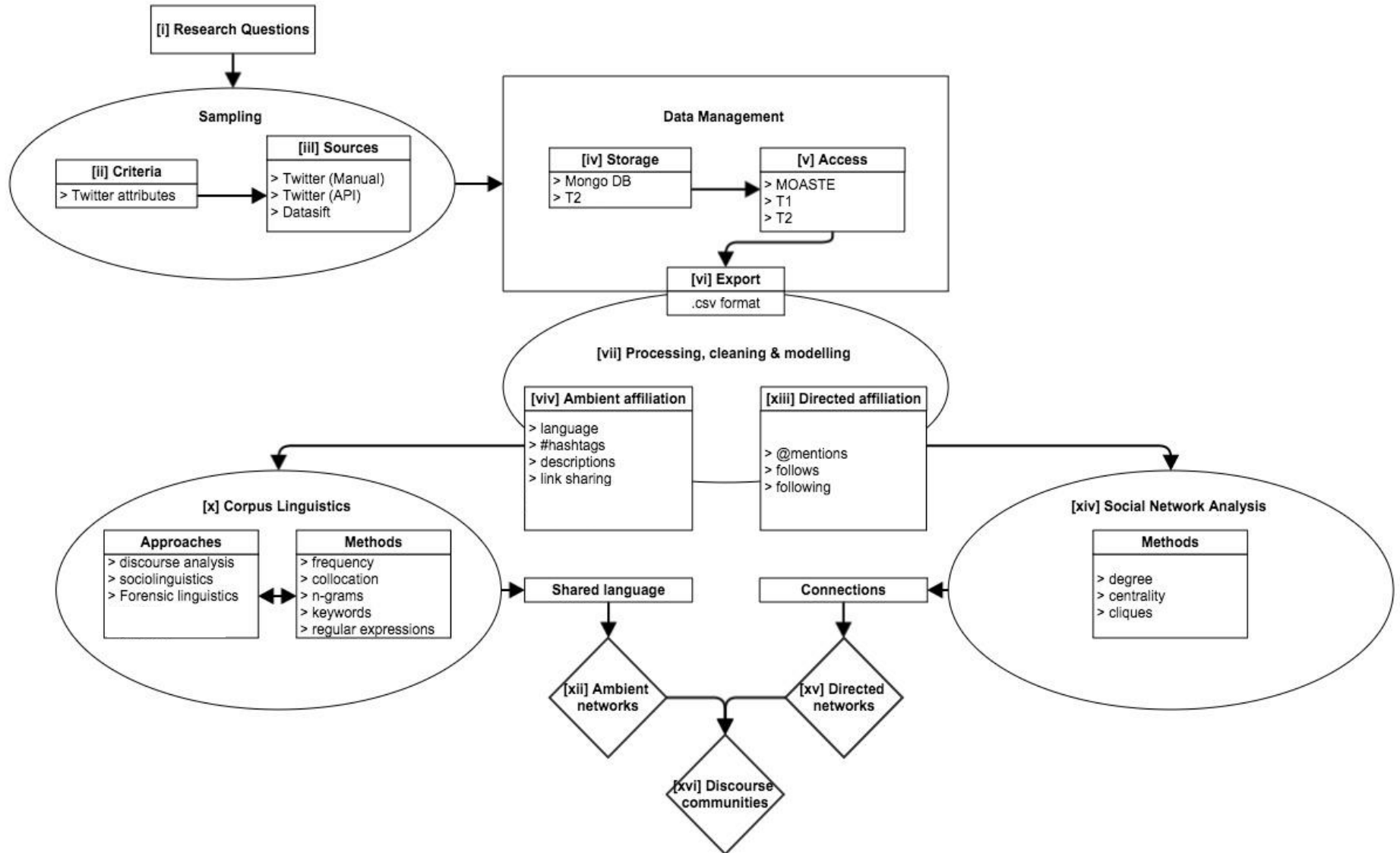
- **What are we investigating?** Hate speech, sexist abuse, violent forms of misogyny online

- **Why?** Investigative bodies etc. pressured to act against increasingly violent online misogyny, but remarkable lack of research

- The aims of DOOM – understand online misogyny, address public concerns, and inform/advise policy, practice, legislation

http://**cass.lancs.ac.uk**

CASS
Corpus Approaches
to Social Science

- **Why** this corpus and event?
  - Known case of misogynistic online abuse
  - Known target
  - Legal proceedings
    - John Nimmo, Isabella Sorley, Peter Nunn
- **How** did we gather it?
  - Seed corpus and DataSift
- How did we analyse it?
  - Corpus linguistics (CL) AntConc & R
  - Social network analysis (SNA) Gephi & R
  - Discourse analysis (DA)

CoCoA Framework - Corpus-assisted Community Analysis

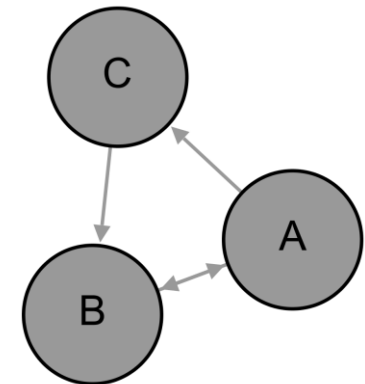Big thanks to Steve Wattam for the dev on this
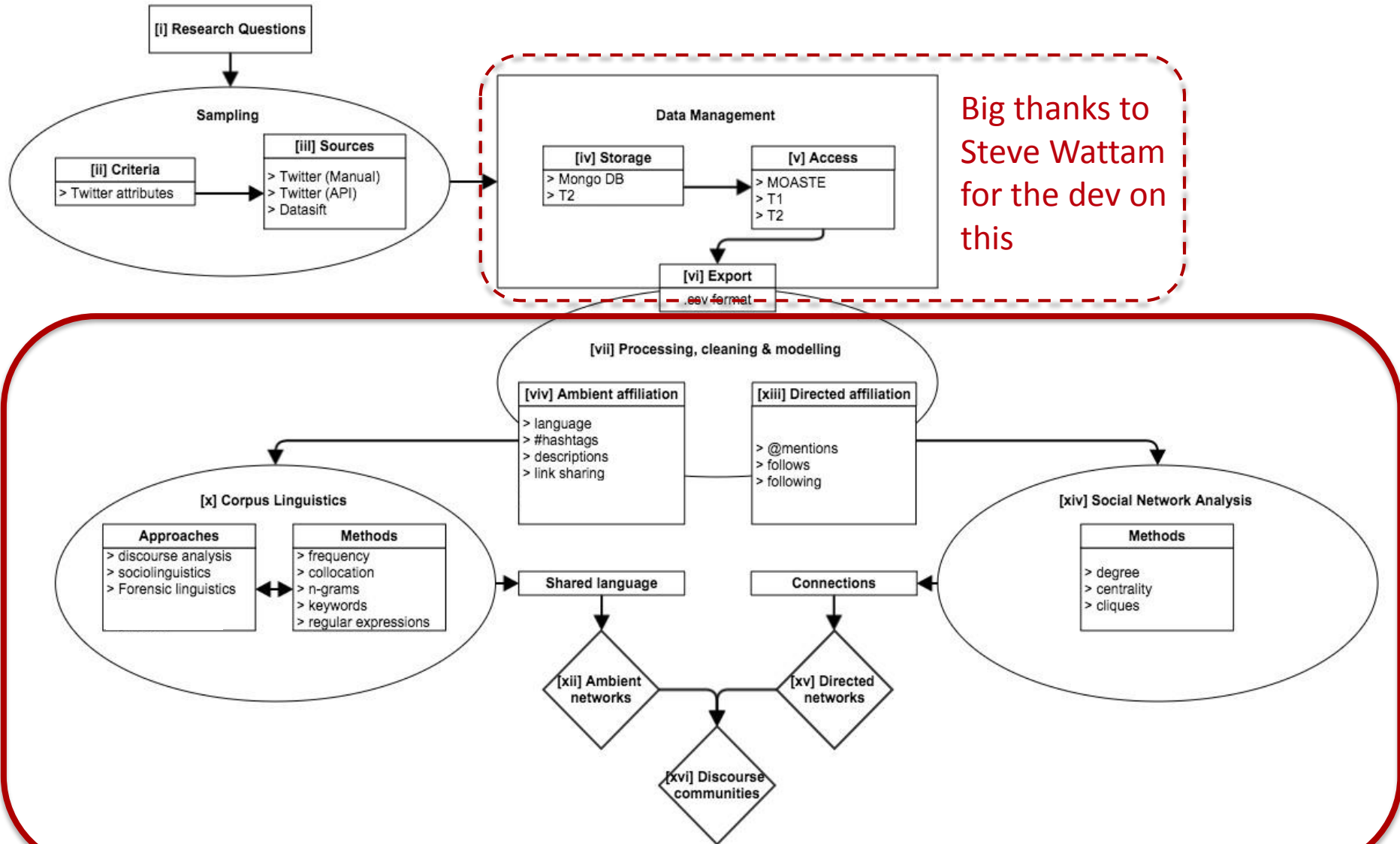
Direct affiliation

- Connections with 'directionality'
- 'directed graphs'[1] representing relationships ('edges') between actors ('nodes')
  - Asymmetric (e.g. I follow you, you don't follow me)
  - Symmetric (e.g. we mention each other in our tweets)

[3] cf. Scott, J. (2013). *Social Network Analysis*. 3rd Ed. London: Sage.

EDL News @edlnews · Jan 19
Do Nigel Farage's jack booted blue rinse army want to legalise chuffing on a huge joint? #ukip #kipperstoners pic.twitter.com/mQKkczlBht

↩   ⟲ 1   ★   •••

The Voice Of Reason @TheVoiceOfRea13 · Jan 19
@edlnews oh fuck off u sad little lefty cunt

↩   ⟲   ★   •••

EDL News @edlnews · Jan 19
@TheVoiceOfRea13 First kipper tantrum of the day.

↩   ⟲   ★   •••

The Voice Of Reason @TheVoiceOfRea13 · Jan 19
@edlnews love to know wat sort of reward you think your goin to receive for opposing anything the majority wants u sad cunt

↩   ⟲   ★   •••

scott cheadle @schead · Jan 19
@TheVoiceOfRea13 @edlnews Ah, the voice of reason. The voice of UKIP. The voice of an illiterate oink unable to string a sentence together.

↩   ⟲ 1   ★ 1   •••

The Voice Of Reason @TheVoiceOfRea13 · Jan 19
@schead @edlnews eat shit do gooding prick u haven't got a fuckin brain in your head beleive ul be rewarded for bein a lefty cunt
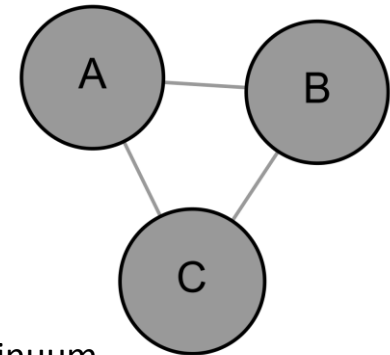
↩   ⟲ 1   ★   •••

EDL News @edlnews · Jan 19
@TheVoiceOfRea13 @schead superb kipper meltdown. first class! lol

Ambient affiliation[1,2]

- "to commune with others without necessarily engaging in direct conversational exchanges"[2]

- Shared behaviours/characteristics

- undirected graphs[3]

[1] Zappavigna, M. (2012) *Discourse of Twitter and social media.* London: Continuum.

[2]Zappavigna, M. (2013) Enacting identity in microblogging through ambient affiliation. *Discourse & Communication.* 8(2). pp. 209-228.

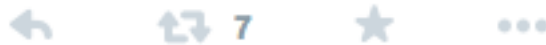[3] cf. Scott, J. (2013). *Social Network Analysis*. 3rd Ed. London: Sage.

CASS
Corpus Approaches
to Social Science

BernieBurnette @BB5163 · Jan 9
#UK: Manchester Muslims demanding laws to ban criticism of their paedophile "prophet". #UKIP #CharlieHebdo #EDL #DDL #PVV #Bluehand #TCOT

↩   ♺ 7   ★   •••

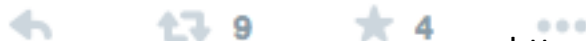**James Bond** @EnglandsTerrier · 3h
How its made.. #Bluehand pic.twitter.com/iDAD6CTNR2

Andy V @avjestica · Jan 8
Can anyone tell me why the hell we have Muslim Council of Great Britain... or what the hell they do (other than claim benefits!!)
#bluehand

↩   ♺ 9   ★ 4   •••

http://**cass.lancs.ac.uk**

- Twitter data -> ambient affiliation
  - Corpora of tweet text
  - Corpora of descriptions text
  - Frequency lists
    - Word|N-gram|Hashtag|link shares|etc.
- Twitter data -> direct affiliation
  - Interaction 'edgelists'
    - Mentions
    - Retweets
    - Follower/friends relationships
- BONUS: Twitter data -> Timeseries
    - tweets|retweets|mentions|link shares|etc. over time

Any tweets which mention or originate from

@CCriadoPerez

between 25/06/2013 and 25/09/2014

CCP Twitter Activity
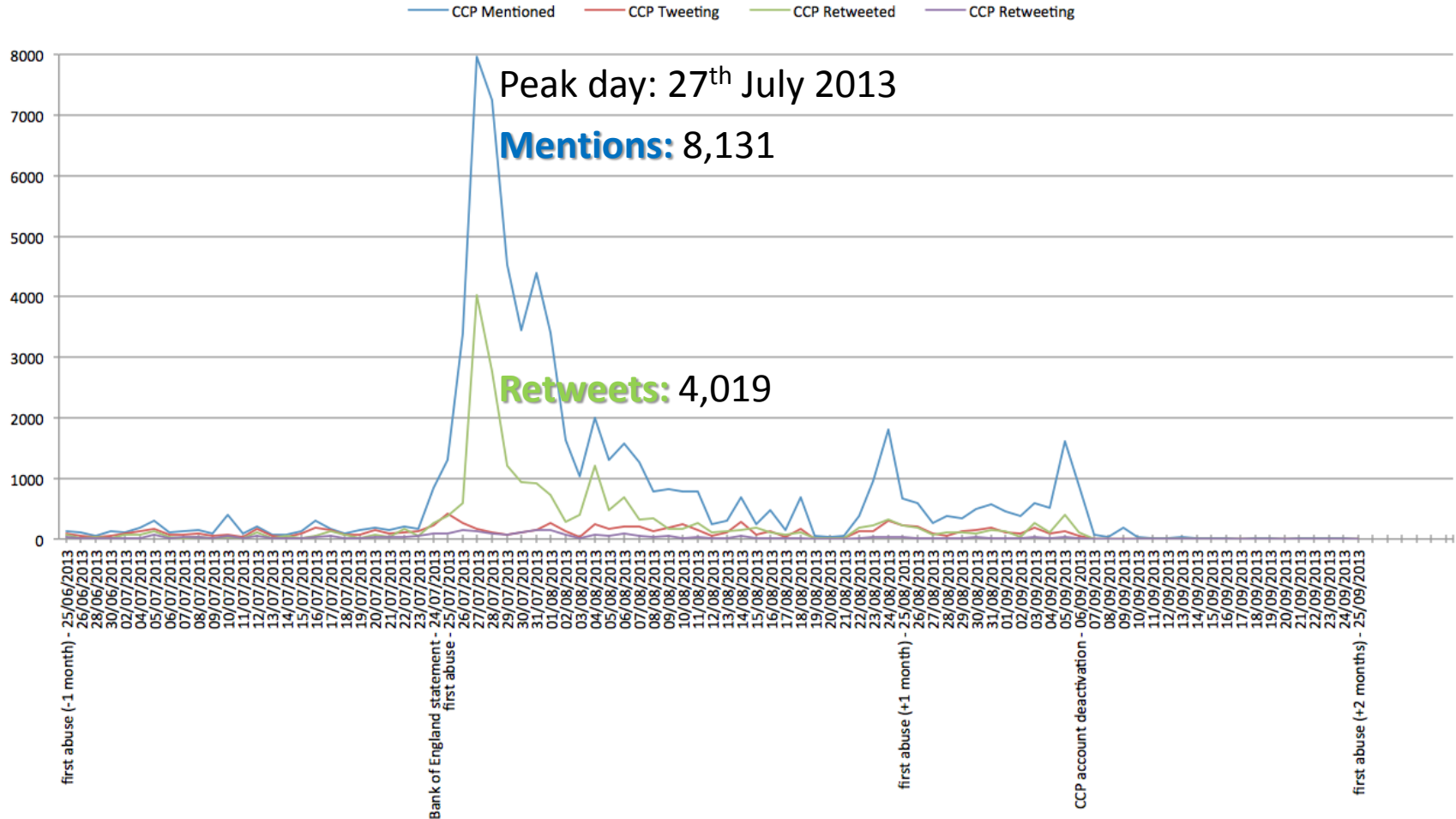
Peak day: 27th July 2013

Mentions: 8,131

Retweets: 4,019

http://**cass.lancs.ac.uk**

1. **How are rape threats constructed?**

    –   What language is used **as part of** misogynistic abuse online?

    –   What language is used **in relation to** it?

2. **Who is talking about rape threats?**

    –   Who is talking about (or making) threats online?

    –   Do they form networks?

    –   If so, how?

- Extracted top 20 (or fewer) most frequent words (less stopwords) in tweets for each day (92 days)
  - Finds most freq. talk **per day**
- Aggregated 92 lists of top 20 (or fewer) words into single frequency list
  - List of most freq. words during entire period that accounts for **daily variation**
- Final list = **456.** we looked in more detail at the top twenty…

CASS
Corpus Approaches
to Social Science

Rank	Word	Freq	Rank	Word	Freq
1	Just	4688	11	Can	2694
2	Abuse	4037	12	Rape	2589
3	Twitter	3971	13	Get	2549
4	Don't (dont)	3614	14	Good	2234
5	Women	3388	15	Men	1896
6	People	3208	16	You're (youre)	1763
7	Threats	2996	17	Will	1399
8	Like	2963	18	Support	1277
9	I'm (im)	2830	19	One	981
10	Think	2812	20	Well	937

Overall corpus: top twenty words

http://**cass.lancs.ac.uk**

CASS
Corpus Approaches
to Social Science

Rank	Word	Freq	Rank	Word	Freq
1	Just	4688	11	Can	2694
2	**Abuse**	**4037**	12	**Rape**	**2589**
3	Twitter	3971	13	Get	2549
4			14		2234
5			15		1896
6	People	3208	16	You're	1763
7	**Threats**	**2996**	17	Will	1399
8	Like	2963	18	Support	1277
9	I'm	2830	19	One	981
10	Think	2812	20	Well	937

appalling abuse, disgusting abuse, frenetic abuse
abuse of @caroladec...

**Rape threats
(one of the most common phrases overall)**

abuse button, abuse on twitter

Overall corpus: top twenty words

http://**cass.lancs.ac.uk**

Defining risky users

- Low risk (147 accounts)
  - General insults and abuse
  - Sexualised, misogynistic remarks
  - Affiliation with high risk users
- High risk (61 accounts)
  - Intent to menace (i.e. to cause fear, threat of harm)
  - (Sexually) aggressive/threatening
  - Incitement to suicide
  - (Repeated) harassment

- Overall proportion of abusive tweets
  - UPPER estimate: 1,957 tweets, from 208 accounts
    - From **low and high** risk users
    - 2.5:100 tweets abusive
  - LOWER estimate: 705 tweets, from 61 accounts
    - From **high** risk users
    - 0.9:100 tweets abusive

CASS
Corpus Approaches
to Social Science

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	Cunt	189.2	12	Pussy	73.7
3	Lol	173.4	13	Raping	73.3
4	Bitch	161.4	14	Penis	68.3
5	Raep	158.6	15	Fucking	60.1
6	Jews	119.6	16	Your	59.1
7	Faggot	104.3	17	Ass	56.5
8	Nigger	92.8	18	Gay	51.1
9	Me	83.8	19	Cake	49.8
10	I	80.6	20	Cock	49.7

High and low risk: top twenty **key**words

http://**cass.lancs.ac.uk**

CASS
Corpus Approaches
to Social Science

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	Cunt	189.2	12	Pussy	73.7
3	Lol	173.4	13	Raping	73.3
4	Bitch	161.4	14	Penis	68.3
5	Raep	158.6	15	Fucking	60.1
6	Jews	119.6	16	Your	59.1
7	Faggot	104.3	17	Ass	56.5
8	Nigger	92.8	18	Gay	51.1
9	Me	83.8	19	Cake	49.8
10	I	80.6	20	Cock	49.7

Rape

http://**cass.lancs.ac.uk**

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	Cunt	189.2	12	Pussy	73.7
3	Lol	173.4	13	Raping	73.3
4	Bitch	161.4	14	Penis	68.3
5	Raep	158.6	15	Fucking	60.1
6	Jews	119.6	16		59.1
7	Faggot	104.3	17		56.5
8	Nigger	92.8	18	Gay	51.1
9	Me	83.8	19	Cake	49.8
10	I	80.6	20	Cock	49.7

**Misspelling =
important!
'raep train'**

Rape

CASS
Corpus Approaches
to Social Science

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	**Cunt**	**189.2**	12	Pussy	73.7
3	Lol	173.4	13	Raping	73.3
4	**Bitch**	**161.4**	14	Penis	68.3
5	Raep	158.6	15	Fucking	60.1
6	Jews	119.6	16	Your	59.1
7	Faggot	104.3	17	Ass	56.5
8	Nigger	92.8	18	Gay	51.1
9	Me	83.8	19	Cake	49.8
10	I	80.6	20	Cock	49.7

Misogyny

http://**cass.lancs.ac.uk**

CASS
Corpus Approaches
to Social Science

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	Cunt	189.2	12	Pussy	73.7
3	Lol	173.4	13	Raping	73.3
4	Bitch	161.4	14	Penis	68.3
5	Raep	158.6	15	Fucking	60.1
6	Jews	119.6	16	Your	59.1
7	**Faggot**	**104.3**	17	Ass	56.5
8	Nigger	92.8	**18**	**Gay**	**51.1**
9	Me	83.8	19	Cake	49.8
10	I	80.6	20	Cock	49.7

Homophobia

http://**cass.lancs.ac.uk**

CASS
Corpus Approaches
to Social Science

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	Cunt	189.2	12	Pussy	73.7
3	Lol	173.4	13	Raping	73.3
4	Bitch	161.4	14	Penis	68.3
5	Raep	158.6	15	Fucking	60.1
6	**Jews**	**119.6**	16	Your	59.1
7	Faggot	104.3	17	Ass	56.5
8	**Nigger**	**92.8**	18	Gay	51.1
9	Me	83.8	19	Cake	49.8
10	I	80.6	20	Cock	49.7

Racism/anti-Semitism

http://**cass.lancs.ac.uk**

Rank	Word	Keyness	Rank	Word	Keyness
1	rape	270.5	10	Loool	75.7
2	Cunt	189.2	**12**	**Pussy**	**73.7**
3	Lol	173.4	13	Raping	73.3
4	Bitch	161.4	**14**	**Penis**	**68.3**
5	Raep	158.6	**15**	**Fucking**	**60.1**
6	Jews	119.6	16	Your	59.1
7	Faggot	104.3	**17**	**Ass**	**56.5**
8	Nigger	92.8	18	Gay	51.1
9	Me	83.8	19	Cake	49.8
10	I	80.6	**20**	**Cock**	**49.7**

Genitalia/anatomy

http://**cass.lancs.ac.uk**

1. **How are rape threats constructed?**

    – What language is used **as part of** misogynistic abuse online?

    – What language is used **in relation to** it?

2. **Who is talking about rape threats?**

    – Who is talking about (or making) threats online?
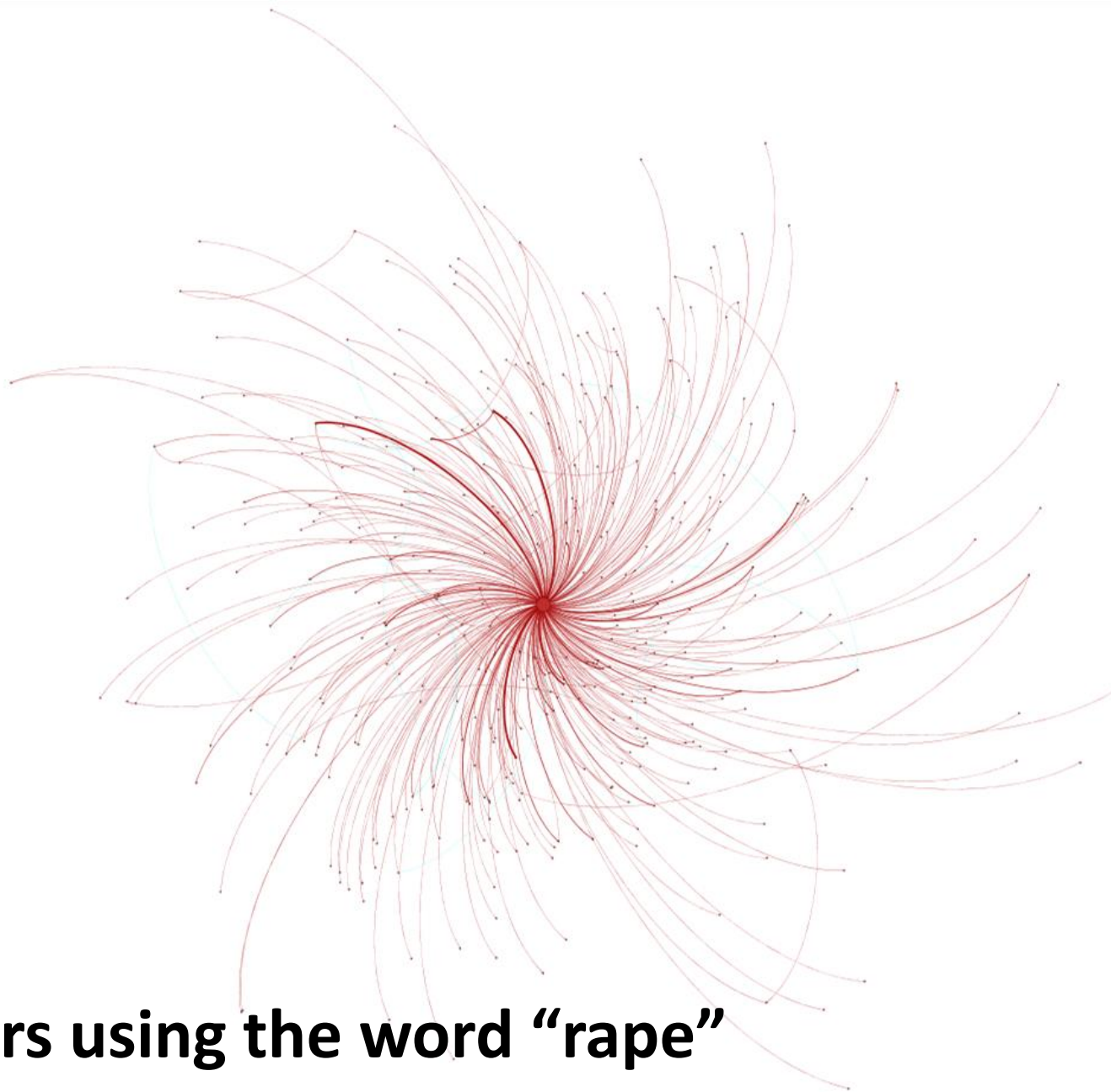
    – Do they form networks?

    – If so, how?
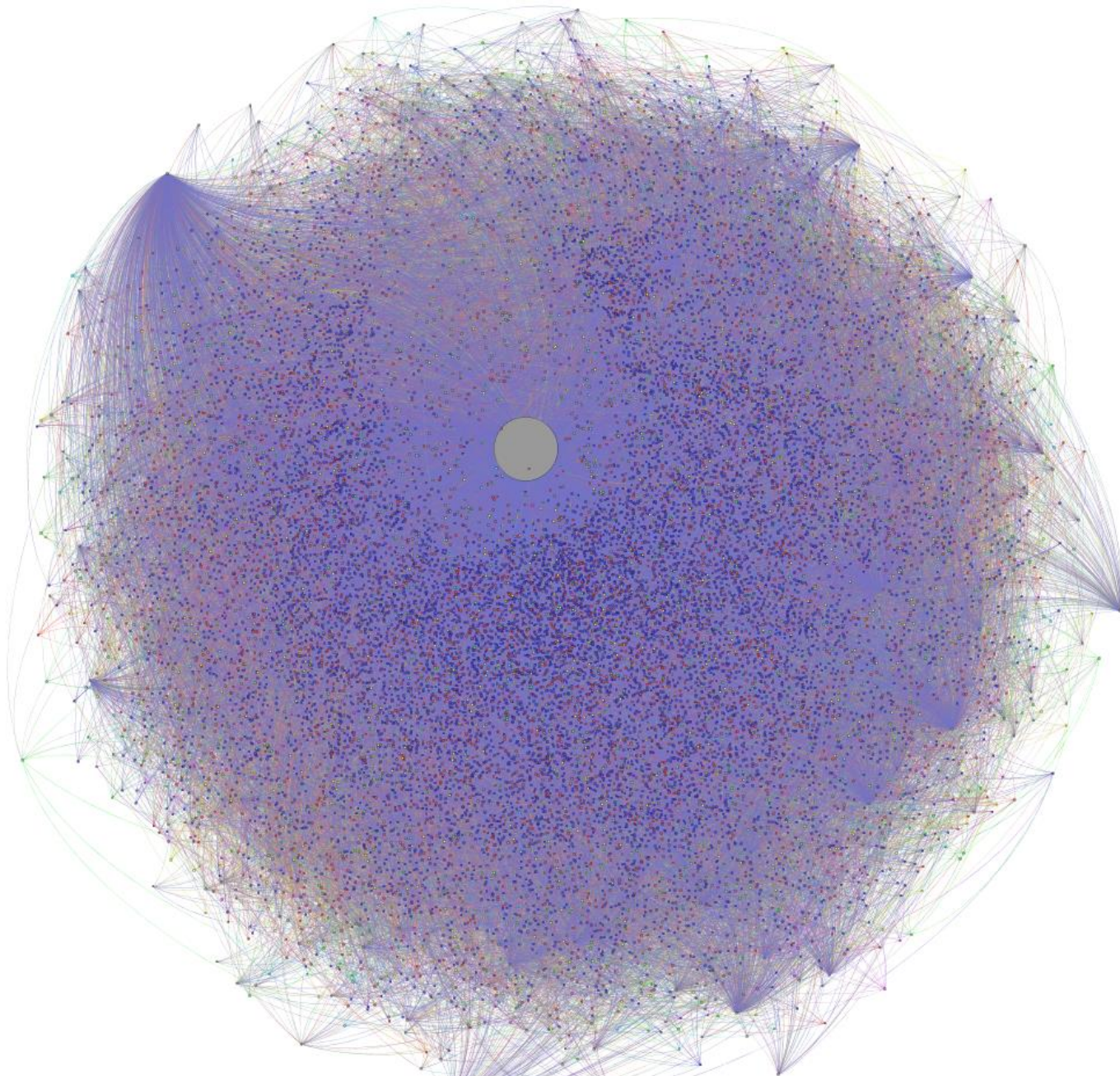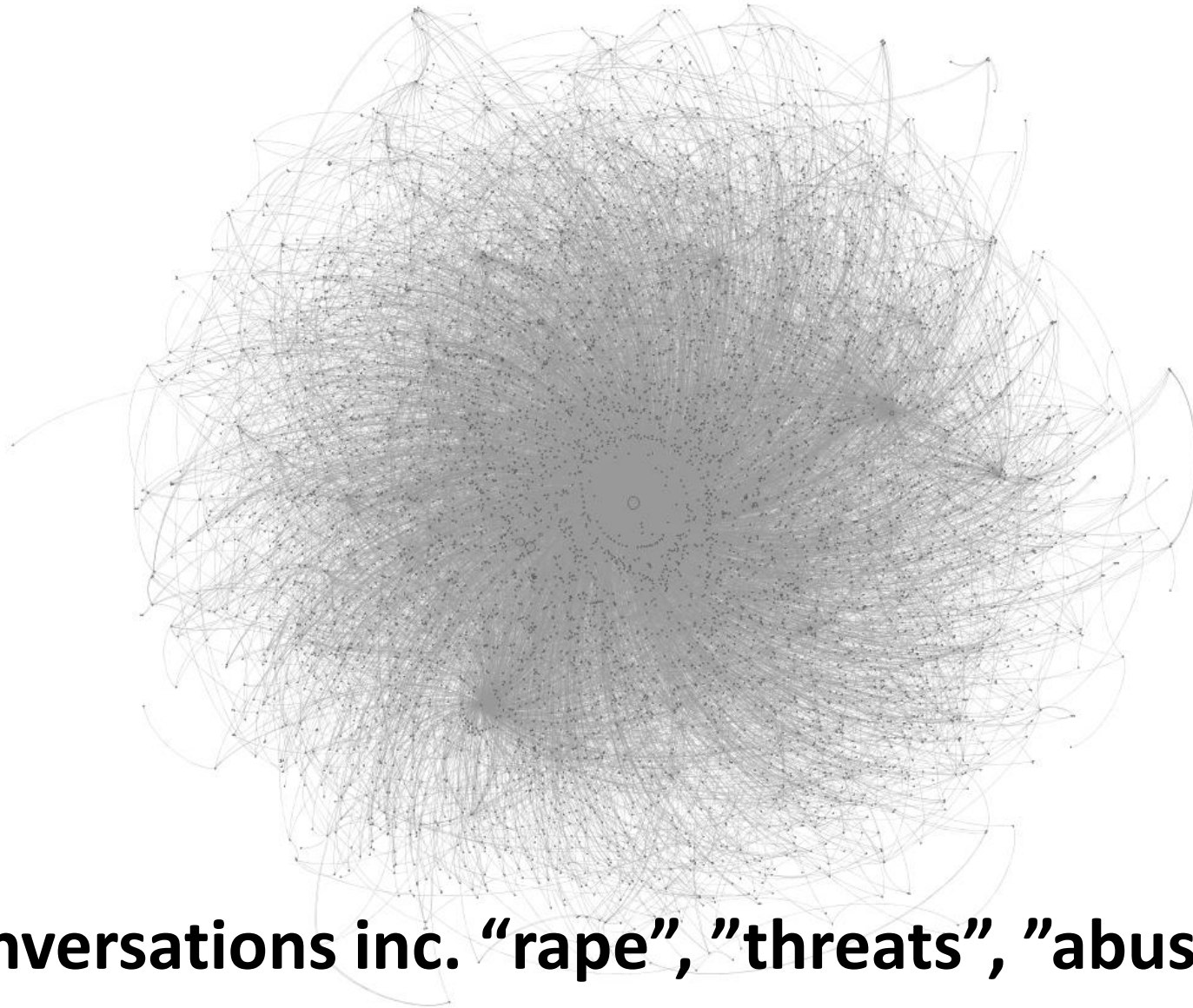
Conversations inc. "rape", "threats", "abuse"
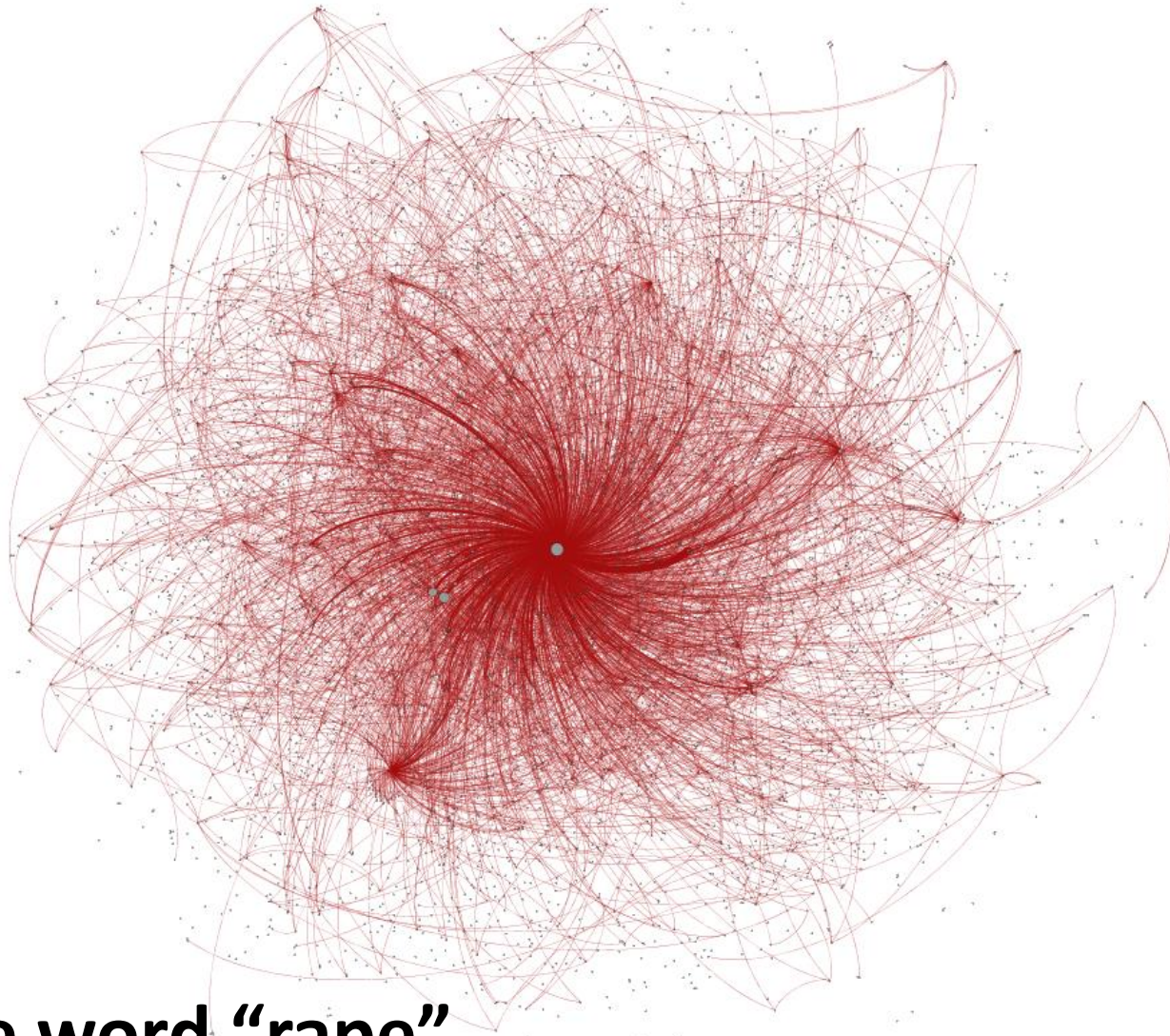
Just the word "rape"

Who is talking about rape threats?

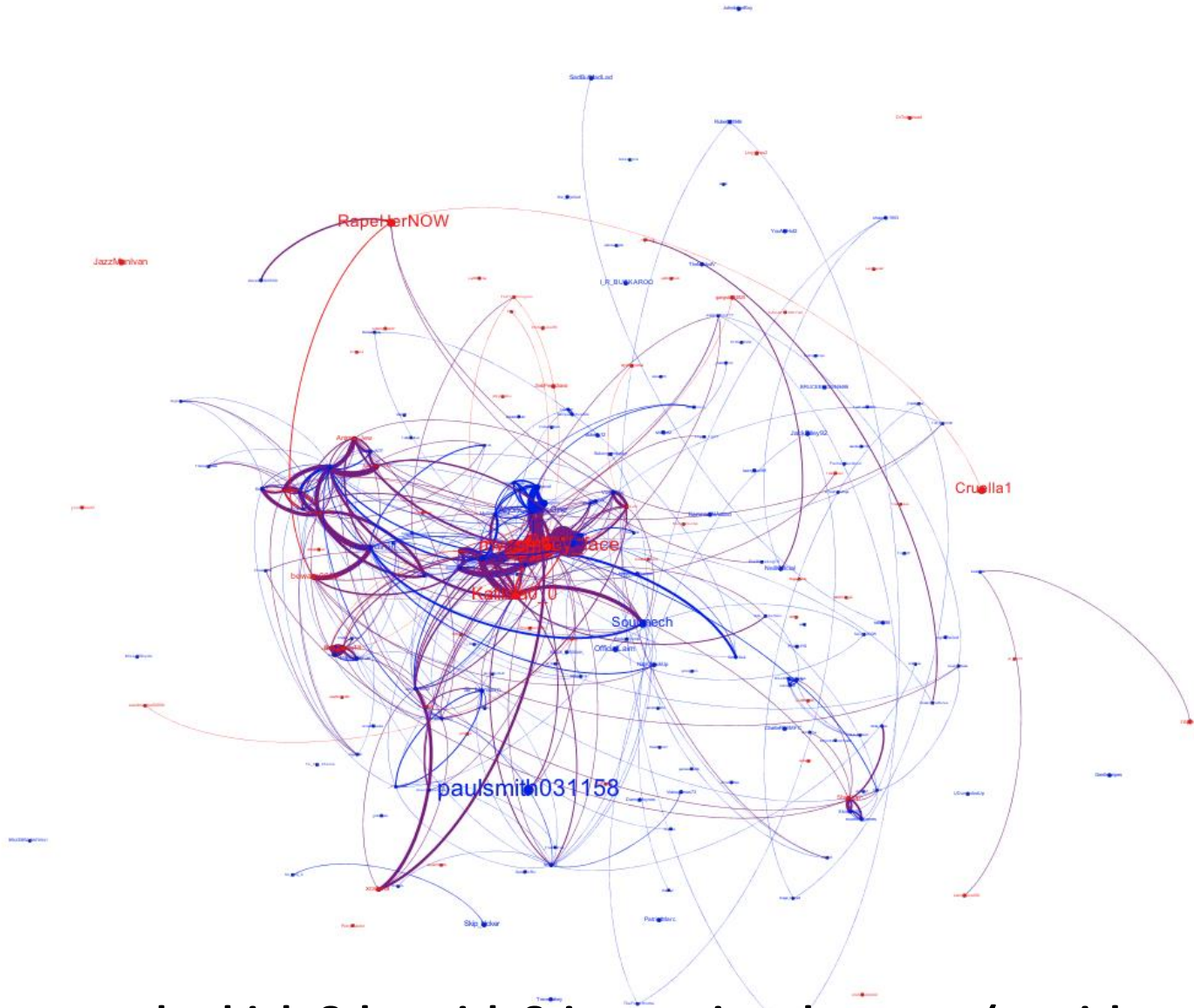|	Word in description	Frequency of occurrence in unique user descriptions (no risk)
Writer/media	Writer	1060
	Editor	431
	Journalist	354
	Books	351
	Blogger	335
	Author	316
	Media	369
Political/activism	Feminist	953
	Politics	598
Education	Student	505
	Teacher	286
	History	284
	Geek	337
Women (relational)	Women	379
	Mum	428
	Mother	339
	Wife	302

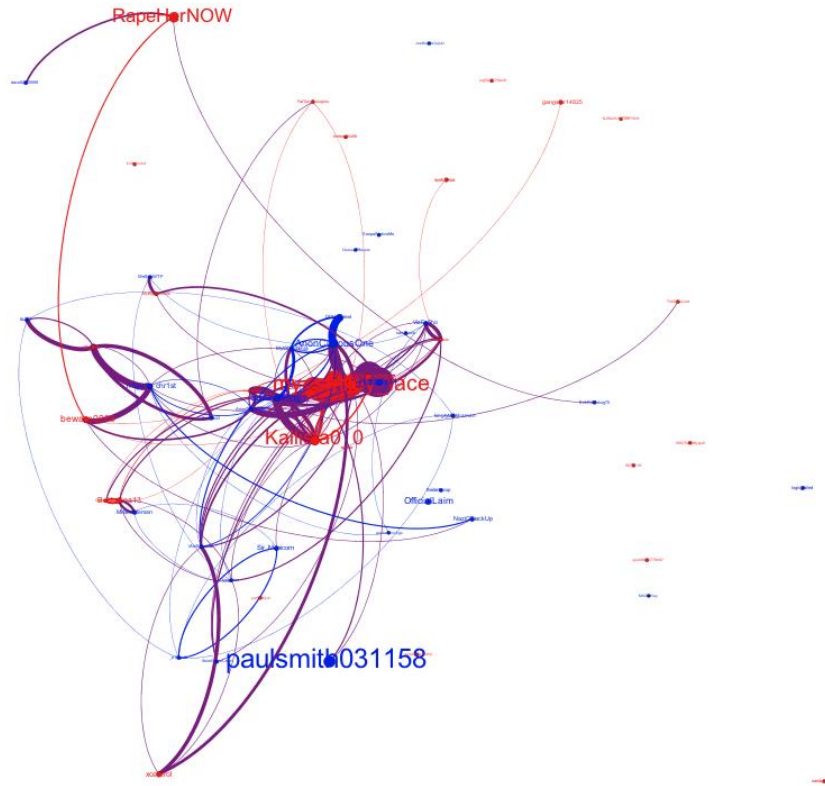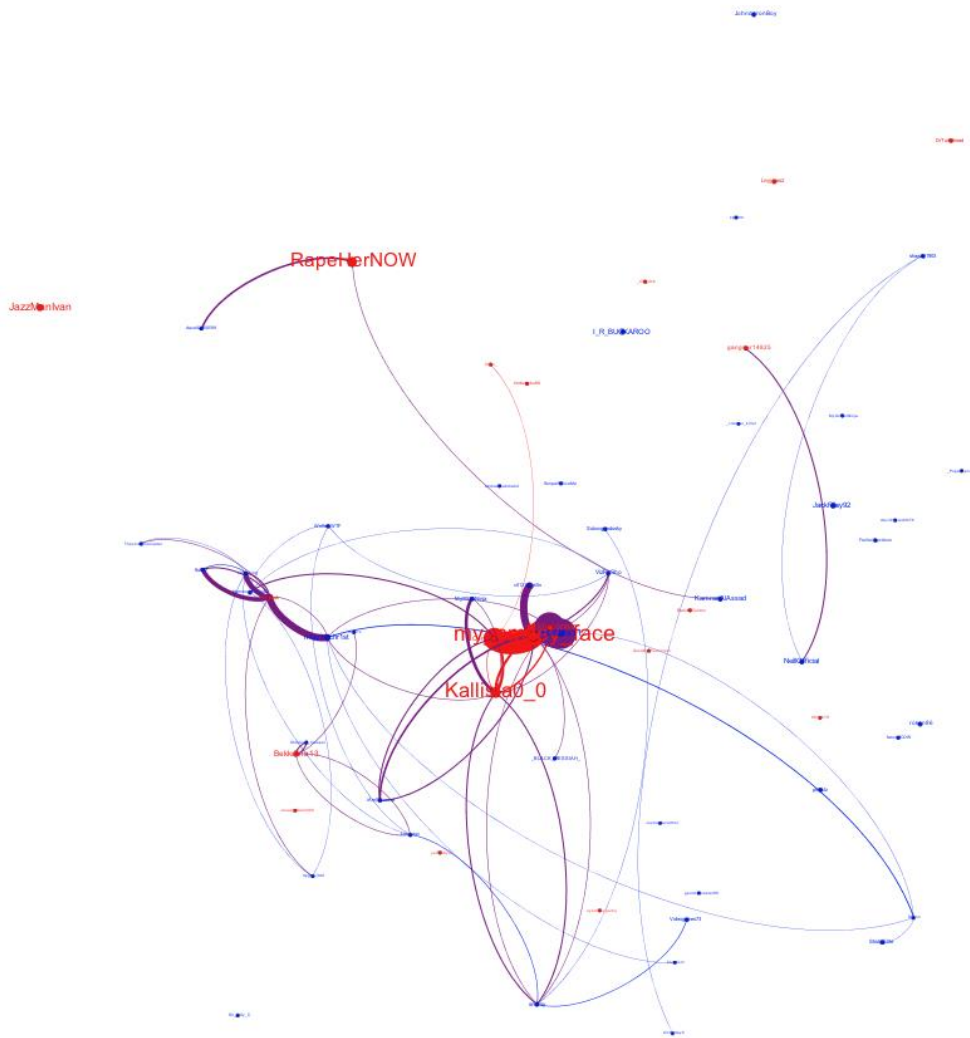**Writers using the word "rape"**
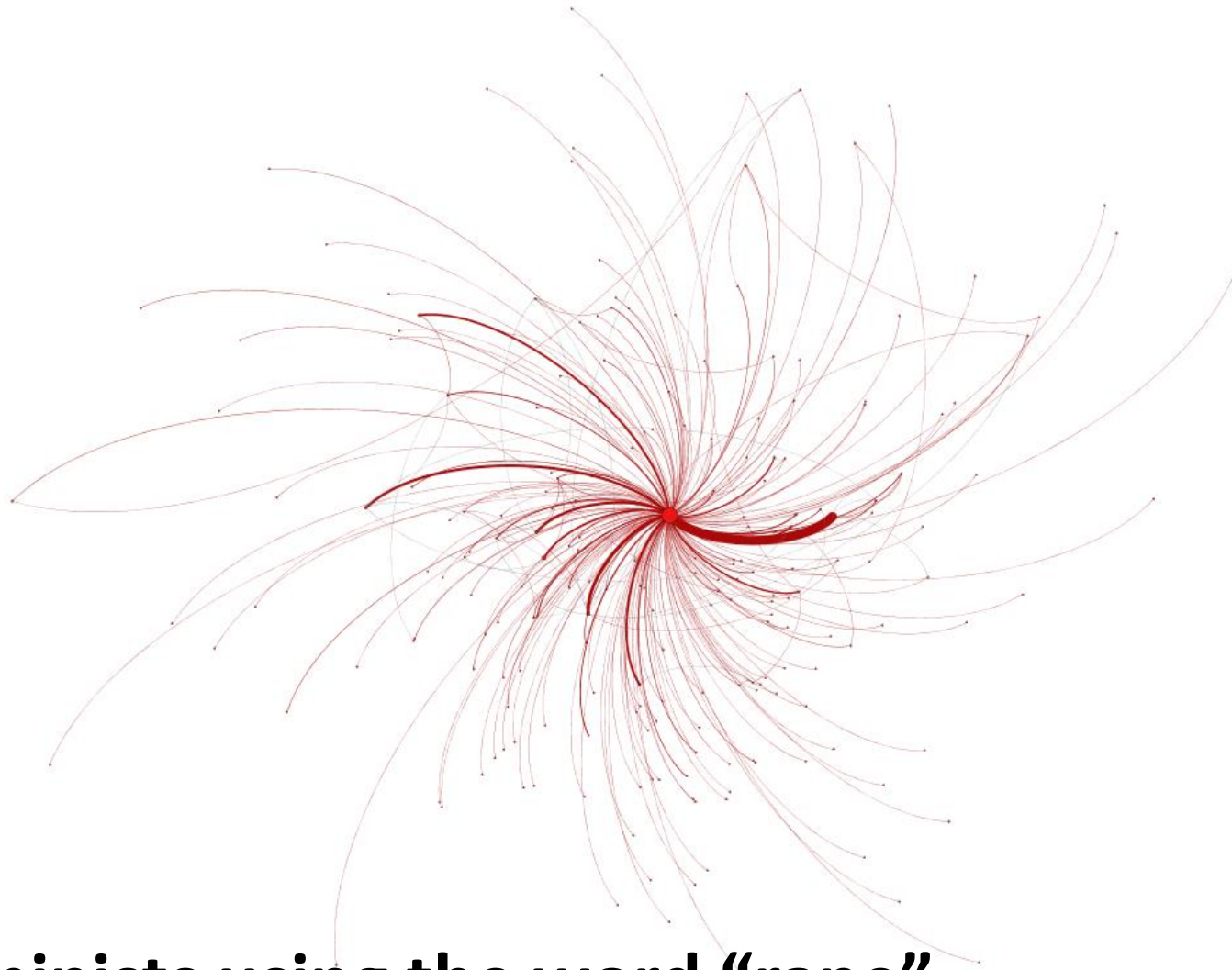
Feminists using the word "rape"
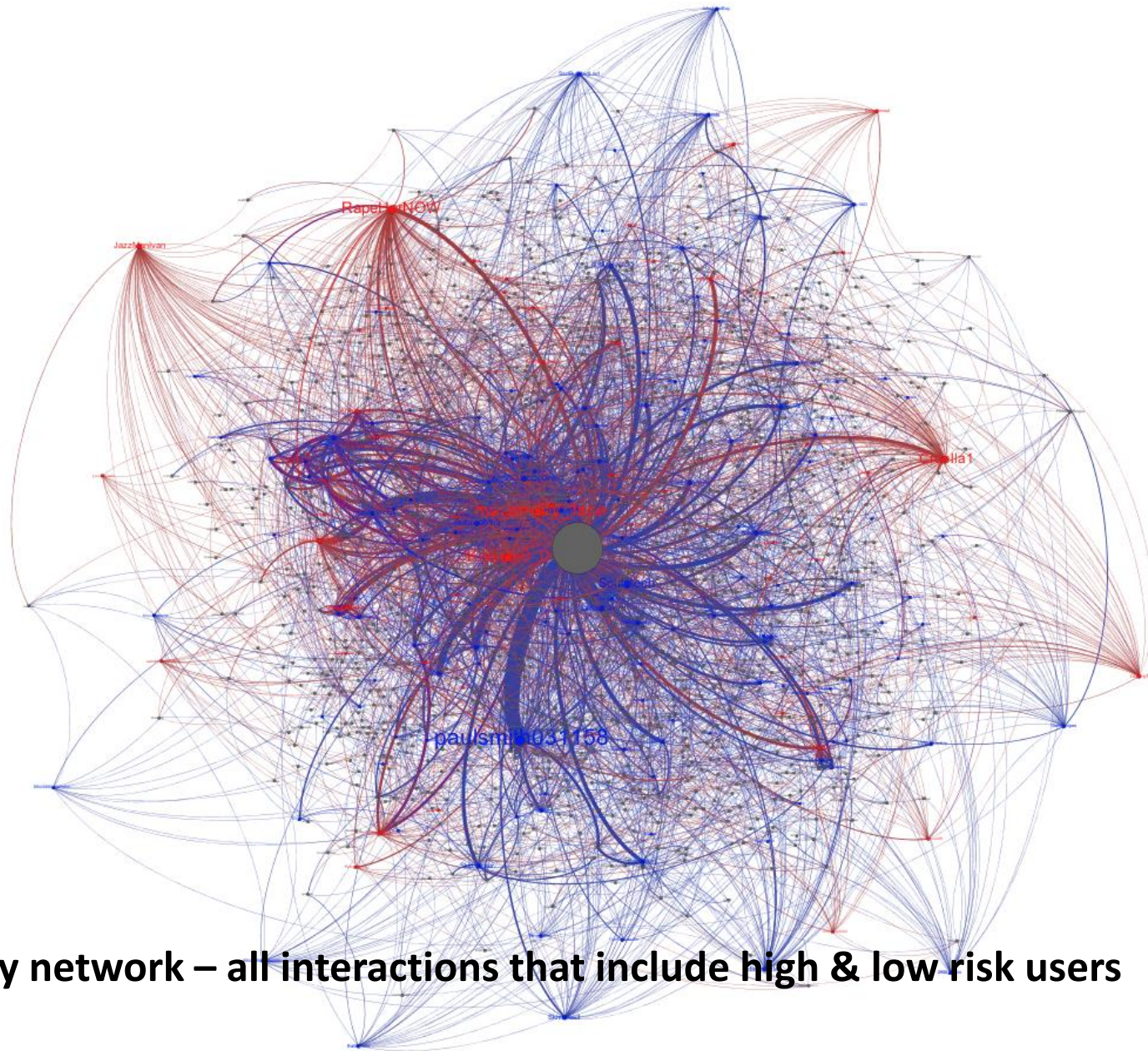
Risky network – all interactions that include high & low risk users

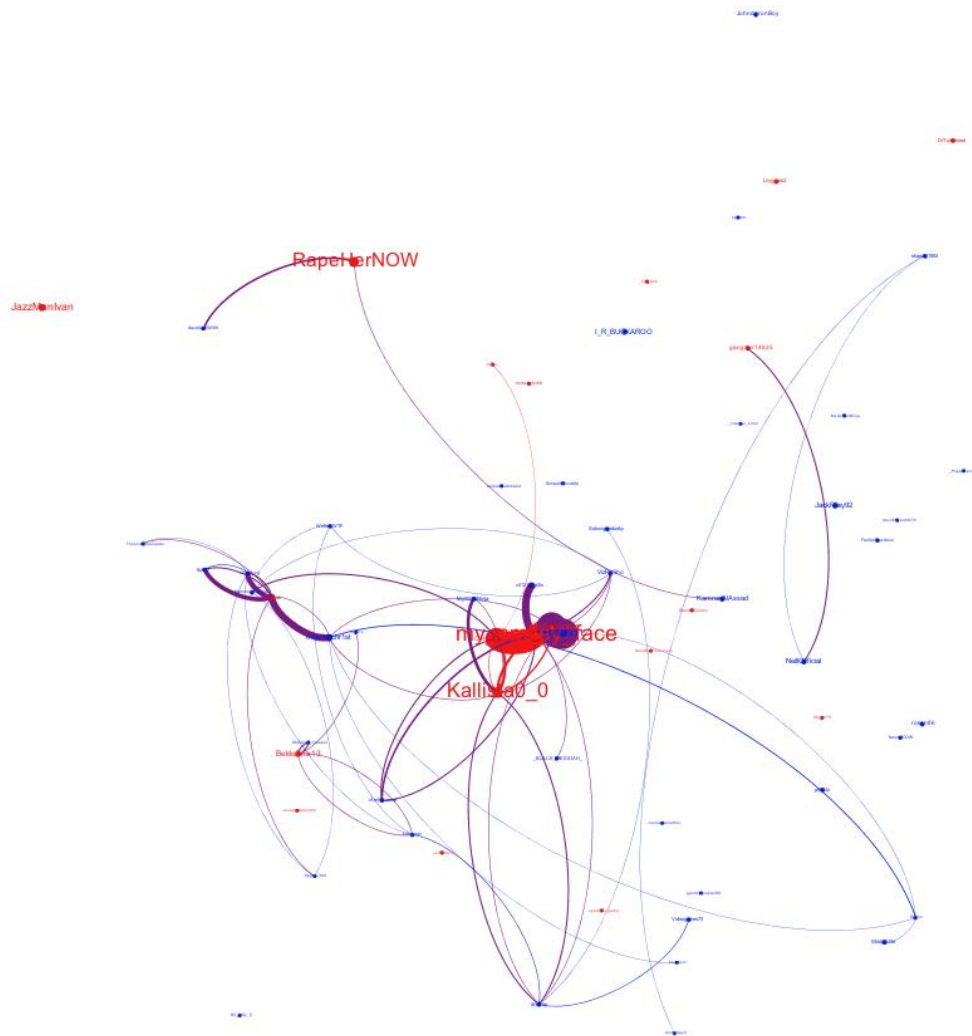**Risky network – high & low risk & interactions between (no-risk removed)**

http://**cass.lancs.ac.uk**

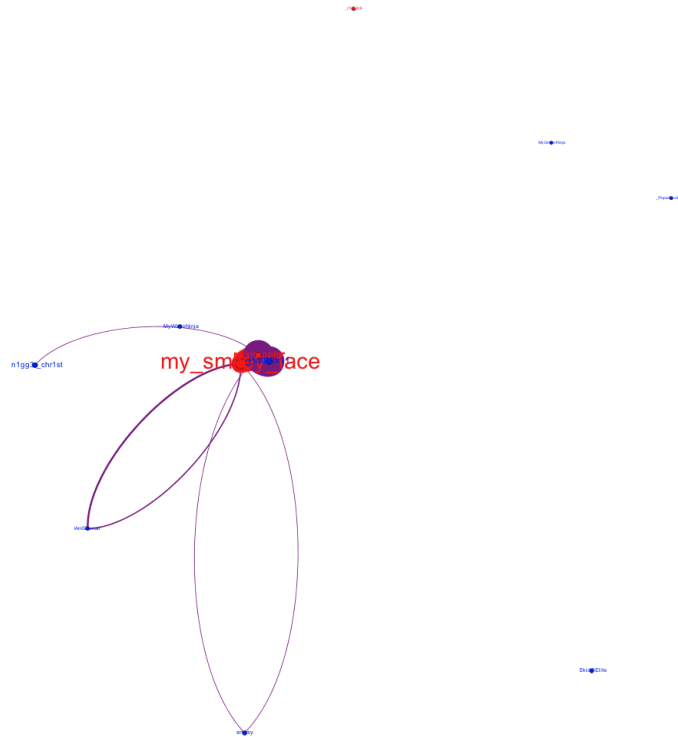**Talk in risky network: rape ("rape", "raep", "raping")**

Talk in risky network: misogyny ("cunt", "bitch")
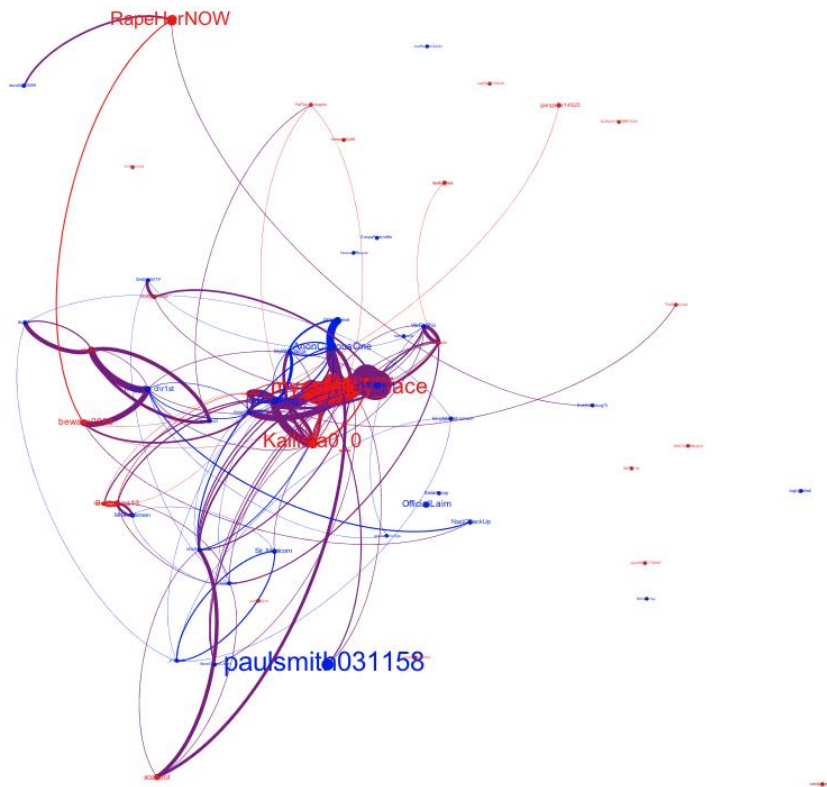
Who is talking about rape threats?

	Word in description	Frequency of occurrence in user descriptions
Group affiliation	**.*sec|.*sek** e.g. #uncletomsec #stfusec #cuntsec #idgafsec	494
	anon #anonymous #cananon #anonmus anonymous	67
	League #rustleleague	56
Racism	Niggas	42
	Nigger	11
Rape	#rapecr3w	101
	Raping	15
	Rape	7

http://**cass.lancs.ac.uk**

Talk in risky network: misogyny ("cunt", "bitch")

http://**cass.lancs.ac.uk**

Risky groups (*sec/*sek) & correlation w/risky talk: misogyny ("cunt", "bitch")

Talk in risky network: rape ("rape", "raep", "raping")

Risky groups (*sec/*sek) & correlation w/risky talk: rape ("rape", "raep", "raping")

1. How are rape threats constructed?

- Use of language relating to
  - Violence (e.g. abuse, rape)
  - Homophobia
  - Racism
  - "Faithism" (e.g. Anti-Semitism)
  - Genitalia and anatomy

2. Who is talking about rape threats?

- Non-risky users: writers, activists, etc.
  - (i.e. word analysis alone is not enough – too much "noise")
- Risky users: forming coherent networks using in-group markers
  - in bios, e.g. *sec/*sek
  - in language, e.g. "raep" (also "raep train", "conductor", etc.)
  - in who they affiliate with, mention, retweet, etc.

CASS
Corpus Approaches
to Social Science

- Future work:
  - Generating linguistic profiles for abusive online behaviour (e.g. hate speech, threats)
  - Methods for detecting abusive online behaviour
  - Identifying escalation of that behaviour (individual or group) – possibly threat/threat-level indication?
  - Investigating how risky communities (networks) develop and evolve