Supplementary Material – Social Media News Communities: Gatekeeping, Coverage, and Statement Bias

This document contains supplementary materials to the paper "Social Media News Communities: Gatekeeping, Coverage, and Statement Bias."

1 Data processing

Table 1 includes the list of data sources used in the paper, as described in Section 3. For each data source, we have included its name, Twitter accounts, the number of articles we collected, and the size of its community, along with the Country/Genre.

Table 1: News sources including name, Twitter accounts, number of articles, size of community, country where the media is based and genre.

Name	Twitter accounts	Arts.	Comm.	Country/Genre	Name	Twitter accounts	Arts.	Comm.	Country/Genre
ABC	@ABC	565	478	USA/Gen.	Houston Chron.	@HoustonChron	326	40	USA/Gen.
ABC	@ABCWorldNews	140	69	USA/Gen.	Hindustan Times	@htTweets	583	118	India/Gen.
Atlanta J.-C.	@ajc	166	53	USA/Gen.	Huffington Post	@HuffingtonPost	2,545	4,406	USA/Gen.
Al Jazeera	@AJELive	34	189	Qatar/Gen.	Indian Express	@IndianExpress	122	57	India/Gen.
Al Jazeera	@AJEnglish	667	645	Qatar/Gen.	LA Times	@latimes	1,151	391	India/Gen.
Al Arabiya	@AlArabiya	284	151	Dubai/Gen.	Mathrubhumi	@mathrubhumi	63	7	USA/Gen.
AP	@AP	454	649	USA/Gen.	Metafilter	@metafilter	92	6	India/Gen.
Arizona Rep.	@arizonarepublic	124	35	USA/Gen.	Metro	@MetroUK	485	75	USA/Gen.
The Australian	@australian	127	10	Australia/Gen.	Miami Herald	@MiamiHerald	285	194	UK/Gen.
Bankrate	@Bankrate	228	18	USA/Bus.	National Post	@nationalpost	1,240	184	USA/Gen.
BBC	@BBCBreaking	44	606	UK/Gen.	Navbharat Times	@NavbharatTimes	479	10	Canada/Gen.
BBC	@BBCBus.	160	1,045	UK/Bus.	NBC	@NBCNews	638	324	USA/Gen.
BBC	@BBCNews	735	3,512	UK/Gen.	News Austrlia	@newscomauHQ	395	77	Australia/Gen.
BBC	@BBCWorld	363	2,150	UK/Gen.	Newsmax	@Newsmax_Media	320	49	USA/Gen.
bdnews24	@bdnews24com	724	146	Bangladesh/Gen.	New Jersey	@njdotcom	253	68	USA/Gen.
Bloomberg	@BloombergNews	822	724	USA/Bus.	National Review	@NRO	174	87	USA/Gen.
Breitbart News	@BreitbartNews	748	1,111	USA/Gen.	NY Post	@nypost	1,075	270	USA/Gen.
Business Std.	@bsindia	234	9	India/Bus.	NY Times	@nytimes	801	1,638	USA/Gen.
CBC	@CBCNews	246	148	Canada/Gen.	Philly.com	@phillydotcom	268	22	USA/Gen.
CBS	@CBSNews	599	435	USA/Gen.	Reuters	@Reuters	824	775	UK/Gen.
Chicago Trib.	@chicagotribune	224	117	USA/Gen.	Russia Today	@RT_com	881	1,203	Rusia/Gen.
CNN	@cnni	575	235	USA/Gen.	Sky News	@SkyNews	410	340	UK/Gen.
CNN	@CNNMoney	380	90	USA/Bus.	Sydney Morn. Hrld.	@smh	439	171	Australia/Gen.
Chr. Sci. Monit.	@csmonitor	685	6	USA/Gen.	Sunday Times	@Suntimes	356	80	USA/Gen.
Daily Finance	@daily_finance	4,531	56	USA/Bus.	The Age	@theage	558	119	USA/Gen.
Dallas News	@dallasnews	233	45	USA/Gen.	The Atlantic	@TheAtlantic	584	346	USA/Gen.
Denver Post	@denverpost	760	96	USA/Gen.	The Economist	@TheEconomist	153	116	UK/Gen.
Digital Spy	@digitalspy	540	112	UK/Ent.	Herald Sun	@theheraldsun	1,170	20	Australia/Gen.
DNA	@dna	376	25	India/Gen.	TIME	@TIME	455	614	USA/Gen.
Deutsche Welle	@dw_english	521	16	Germany/Gen.	Times of India	@timesofindia	3,753	404	India/Gen.
Economic Times	@EconomicTimes	1,782	69	India/Bus.	Toronto Star	@TorontoStar	717	157	Canada/Gen.
EuroNews	@euronews	170	31	France/Gen.	UPI	@UPI	1,165	98	USA/Gen.
Examiner	@examinercom	95	28	USA/Ent.	USA Today	@USATODAY	408	196	USA/Gen.
Forbes	@Forbes	2,062	1,484	USA/Bus.	U.S News	@usnews	797	65	USA/Gen.
FOX	@FoxNews	599	1,156	USA/Gen.	Voices of America	@VOA_News	520	81	USA/Gen.
France 24	@France24_en	249	101	France/Gen.	Washington Post	@washingtonpost	534	482	USA/Gen.
Global Mail	@globeandmail	2,348	194	Canada/Gen.	Washington Times	@washtimes	488	184	USA/Gen.
Guardian	@guardian	650	1,089	UK/Gen.	WND News	@worldnetdaily	287	268	USA/Gen.
Guardian World	@guardianworld	945	749	UK/Gen.	Wall St. Journ.	@WSJeurope	4	312	Europe/Gen.

Examples of stories found using the method in Section 3.2 are listed on Table 2.

Table 2: Examples of stories found in our dataset.

Story (date 2013)	Arts.	Sources	Tweets	Comms.
Former UK PM Margaret Thatcher dies (Apr 9th)	309	67	49K	51
Tensions in the Korean peninsula (Apr 11th)	402	62	53K	48
Presidental elections in Venezuela (Apr 13th)	18	16	1.2K	52

The correlation of $r^2 = 0.74$ described in Section 3.3 is supported by Figure 1.
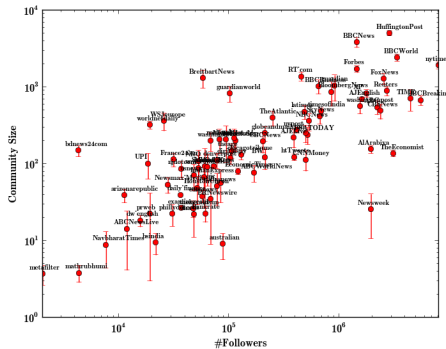
Figure 1: Number of followers of the corporate account of each news source versus size of the (active) community of each source. These quantities are correlated ($r^2 = 0.74$), but communities are 3 orders of magnitude smaller than the number of followers.

2 Selection biases

The correlations described in Section 4.1 are a sub-set of those found in Table 3.

Table 3: Correlation between quantities of articles and stories and community sizes. The table shows the Person correlation between the logarithm of the metrics.

	MA	MS	TF	TS	TP	Median
MA. Articles in media	-					454
MS. Stories in media	0.83	-				98
TF. Followers of media	0.36	0.49	-			105 K
TS. Stories in comm.	0.29	0.48	0.80	-		263
TP. People in comm.	0.23	0.38	0.74	0.73	-	117

The fraction of exclusive stories per each data source is depicted in Figure 2

The distribution of prominence of stories, as described in Section 4.2, is depicted for a set of example news media in Figure 3.
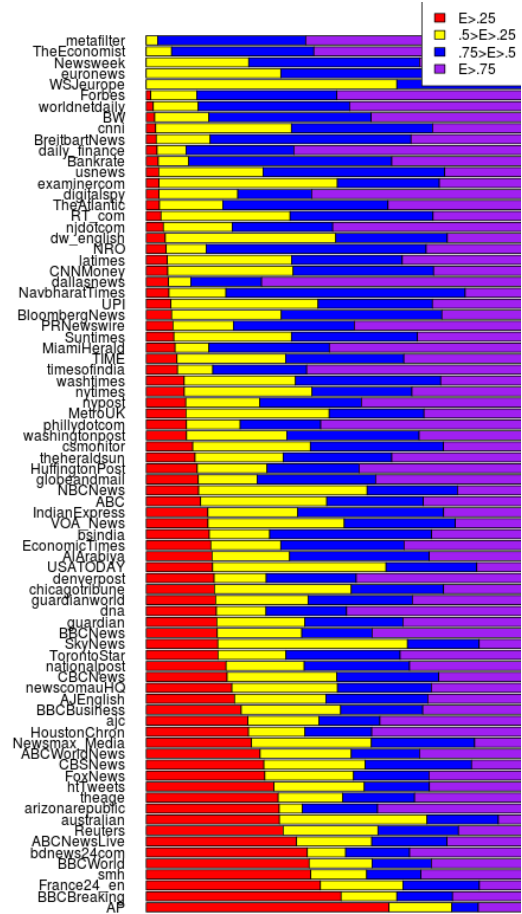

Figure 2: Exclusivity of stories. The fraction of exclusive articles varies widely across online news sources. $E > 0.75$ are articles that are basically unique to one source, while $E < 0.25$ appear in two or more sources with minor differences.
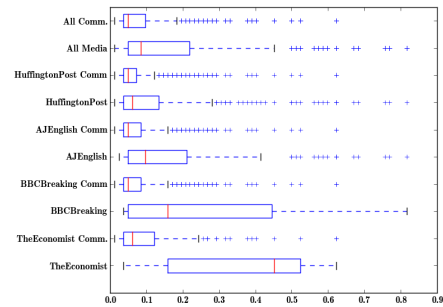

Figure 3: Prominence of stories. Social media news communities share more often niche content, and in general do not share stories having high prominence. This is in sharp contrast with online news media sources, which prefer stories having in general higher prominence than the social media ones.

3   Coverage bias

Table 4 presents more correlations between selection and coverage biases, as described in Section 5.1.

Table 4: Correlation between selection and coverage biases as presented by news media sources. Correlations above 0.8 between different biases are shown in bold-face.

	SS	SP	CS	CP	CT	CT'
SS. Selection bias by stories	1	-	-	-	-	-
SP. Selection bias by people	0.66	1	-	-	-	-
CS. Coverage bias by story words	**0.81**	0.63	1	-	-	-
CP. Coverage bias by people mentions	0.68	**0.94**	0.68	1	-	-
CT. Coverage bias by tweets (all)	0.65	0.53	**0.84**	0.60	1	-
CT'. Coverage bias by tweets (community)	0.30	0.22	0.40	0.28	0.37	1

4   Statement bias

Figure 4 supplements observations on Section 6 with anecdotal observations about the death of Margaret Thatcher on April 8th, 2013: social media users were described as "dancing on the grave" of the former UK Prime Minister, while traditional news media was much more circumspect.[1]

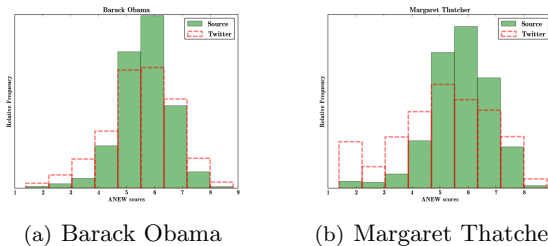(a) Barack Obama   (b) Margaret Thatcher

Figure 4: Distribution of valence scores in mentions in media sources and social media communities, for two politicians in our sample. Mentions in social media exhibit a wider range of expression and a tendency towards more negative sentiments.

---

[1] http://www.independent.co.uk/voices/comment/margaret-thatchers-death-newspapers-pay-respect-while-social-media-dances-on-her-grave-8565679.html