

Supplementary material to:

Orduña-Malea, E. & Bautista-Puig, N. (in press). Research assessment under debate: disentangling the interest around the DORA Declaration on Twitter. *Scientometrics*.
<https://doi.org/10.1007/s11192-023-04872-6>

Table of contents

Appendix A

Co-authorship network	3
-----------------------------	---

Appendix B

Sentiment Analysis tests	4
--------------------------------	---

Appendix C

Distribution of languages.....	5
--------------------------------	---

Appendix D

Twitter users mentioning DORA-related hashtags.....	6
---	---

Appendix E

Network of Twitter users.....	7
-------------------------------	---

Appendix F

Clusters of the co-occurrence network of terms.....	9
---	---

Appendix G

Co-occurrence network of hashtags	13
---	----

Appendix H

Individual tube lay-outs	15
--------------------------------	----

Appendix I

Average sentiment analysis per month.....	16
---	----

Appendix J

Sentiment analysis broken by impact for each hashtag category	18
---	----

Appendix K

Output-level impact broken by hashtag category	21
--	----

Appendix A

Co-authorship network

According to Dimensions, there are 744 publications, including "sfdora.org" in the full text. Fig.A1 includes a co-authorship network that identifies the leading institutions that have discussed or mentioned the topic in the literature.

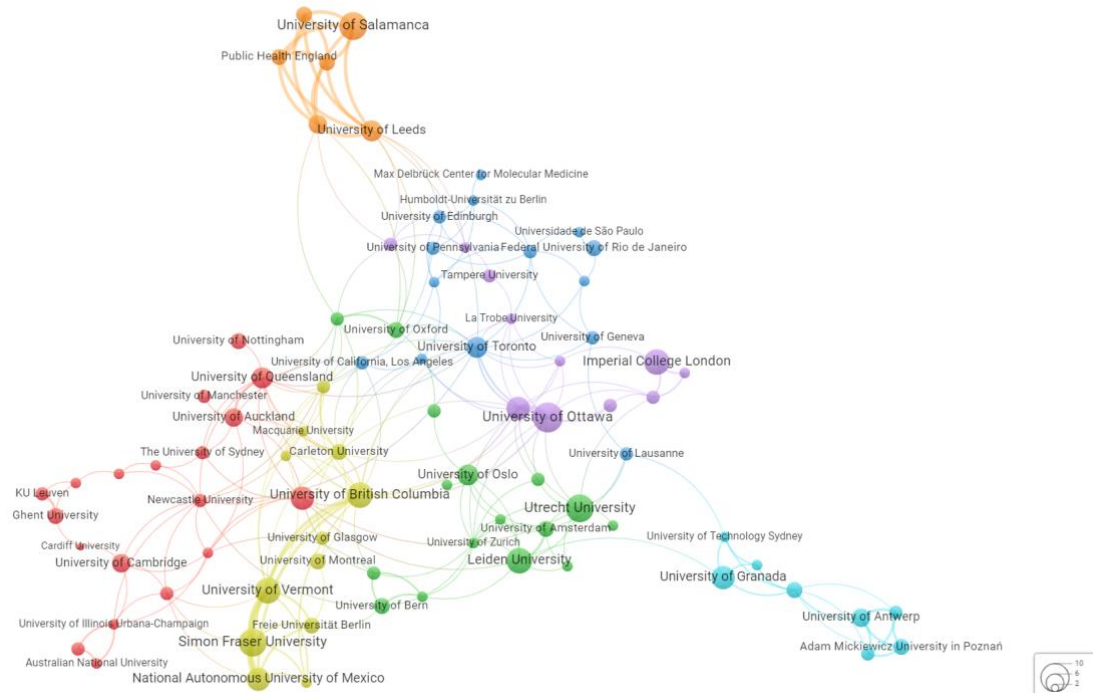


Fig.A1. Co-authorship network at the institution level for publications, including "sfdora.org"

Source: Dimensions. Map created with VOSviewer (<https://www.vosviewer.com>).

Note: this map excludes those publications not mentioning DORA's website. Therefore, it should be used only for illustrative purposes.

Appendix B

Sentiment analysis tests

At the initial stage, AFINN (2,477 terms) from the *tidytext* library, which ranges between -5 (negative sentiment) and +5 (positive sentiment), was used. Although it performed well (12,545 tweets classified 76% of tweets in English), it does not include a Spanish dictionary (36.87% of the dataset was not classified). Later, the *NRC Valence, Arousal, and Dominance (NRC-VAD) Lexicon* (with 19,971 terms in different languages) ranging from 0 to 1 (valence represents the positivity and negativity dimension) was used (Mohammad, 2018). However, this method was not able to classify 7% of tweets.

The SentiStrength¹ tool was also used (Thelwall et al., 2010). This application estimates the strength of positive and negative sentiment in short texts assigning values ranging from -1 to -5 (negative) and 1 to 5 (positive). It has been widely used to analyze Twitter sentiment (e.g., Thelwall, Buckley & Paltoglou, 2011). Our dataset was fully classified using this method. Subsequently, the results were normalized on a scale of -1 to 1 and compared with those generated by CorTexT. We found that 15% of positive tweets yielded the same values with both tools, while 9% of negative tweets showed the same value (neutral tweets with a value of 0 were excluded). However, the differences between the two tools were minimal for positive tweets (an average difference of 0.07 points) and negative tweets (0.14).

References

- Mohammad, S. (2018, July). *Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words*. In Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers) (pp. 174-184).
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558. <https://doi.org/10.1002/asi.21416>
- Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62(2), 406-418. <https://doi.org/10.1002/asi.21462>

¹ <http://sentistrength.wlv.ac.uk>

Appendix C

Distribution of languages

Table C1. Distribution of languages used in the tweets published about DORA

Language	Complete dataset		User mention dataset		Hashtag mention dataset		DORA dataset	
	No. Tweets	%	No. Tweets	%	No. Tweets	%	No. Tweets	%
English	17,201	83.0	11,024	79.3	718	85.0	5,459	91.3
Spanish	1,349	6.5	1,129	8.1	36	4.3	184	3.1
Dutch	354	1.7	326	2.3	3	0.4	25	0.4
French	361	1.7	269	1.9	48	5.7	44	0.7
German	242	1.2	225	1.6	0	0.0	17	0.3
Undetermined	561	2.7	419	3.0	10	1.2	132	2.2
Others	649	3.1	503	3.6	30	3.6	116	1.9
Total tweets	20,717	100	13,895	67	845	4	5,977	29

Note: only languages with >1% in the complete dataset are displayed.

Appendix D

Twitter users mentioning DORA-related hashtags

Table D1. Twitter users who include DORA-related hashtags most frequently

User	Tweets	Type	Role	Genre	All tweets	Followers
Alfonso Martínez Arias	84	Personal	Researcher	Male	12,000	7,616
American Society for Cell Biology	52	Institutional	Society	N/A	12,800	29,161
Stephen Curry	34	Personal	Researcher	Male	79,100	21,792
Lupicinio Iñiguez-Rueda	20	Personal	Researcher	Male	77,100	4,403
Christopher Jackson	17	Personal	Researcher	Male	111,900	36,681
Anna Hatch	16	Personal	Practitioner	Female	6,981	1,368
SPARC	14	Institutional	Coalition	N/A	13,300	17,884
The Winnower	11	Institutional	Publisher	N/A	15,400	4,872
Jim Woodgett	10	Personal	Researcher	Male	73,800	10,134
Stephen Royle	8	Personal	Researcher	Male	16,000	5,827
Redalyc	8	Institutional	Database	N/A	25,500	33,922

Note: total tweets and followers counts as of 6 June 2023. These tweets do not include mentions to DORA account.

Appendix E

Network of Twitter users

Fig.E1 includes the network of users related to the DORA Declaration, where node **D** represents the DORA dataset, and node **C** (community) means the user-mention dataset. Twitter users (the remaining nodes) are linked when they receive a mention from tweets included in the DORA and community datasets, respectively. As we can observe, the overlap of users among those who have been co-mentioned with DORA or mentioned directly by DORA is limited to a few users.

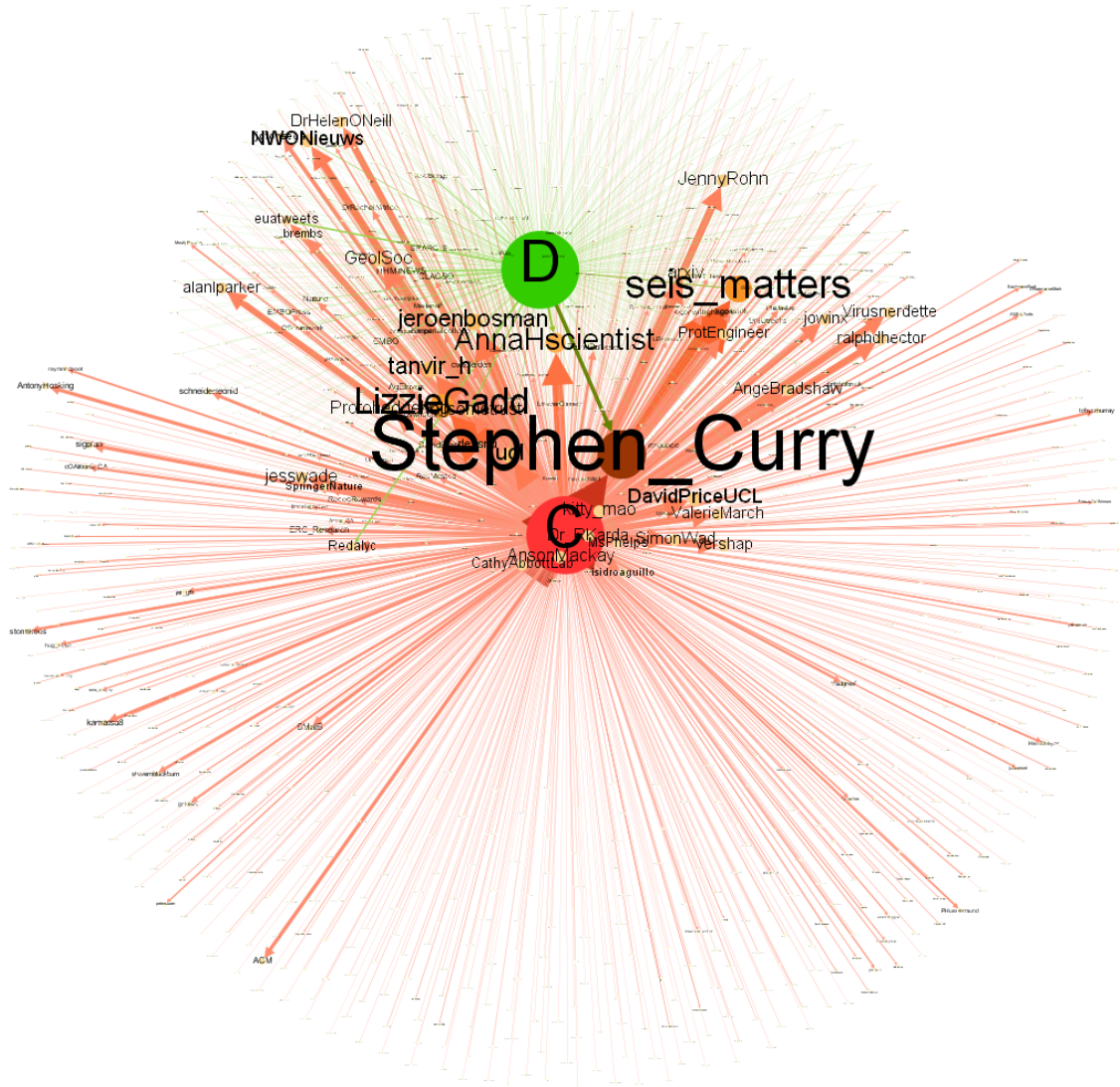


Fig.E1 Network of Twitter users being mentioned from the DORA dataset (node D) or the user.mention dataset (node C)

Note: Only those authors attaining at least five mentions are included for clarity.

Appendix F

Clusters of the co-occurrence network of terms

ID	Term	Weight	cluster id	cluster label
1	hong kong principles	49	0	open science & research assessment
2	use metrics research	99	0	open science & research assessment
4	research assessment	626	0	open science & research assessment
9	humanities social sciences	20	0	open science & research assessment
12	open access	39	0	open science & research assessment
17	imperative scientific output	30	0	open science & research assessment
24	blog post	55	0	open science & research assessment
31	impact factor	139	0	open science & research assessment
32	hiring promotion tenure	72	0	open science & research assessment
33	hiring promotion	254	0	open science & research assessment
39	research integrity	39	0	open science & research assessment
40	research evaluation	332	0	open science & research assessment
48	open science	112	0	open science & research assessment
50	research impact	52	0	open science & research assessment
51	review promotion tenure	44	0	open science & research assessment
55	research metrics	34	0	open science & research assessment
58	leiden manifesto	82	0	open science & research assessment
60	use journal	71	0	open science & research assessment
63	practice research	35	0	open science & research assessment
66	sfdora blog post	2	0	open science & research assessment
77	scientific content paper	67	0	open science & research assessment
80	career researchers	49	0	open science & research assessment
82	reward system	20	0	open science & research assessment
84	research outputs	138	0	open science & research assessment
85	use metrics	15	0	open science & research assessment
86	research institutions	156	0	open science & research assessment
87	sfdora community interview	38	0	open science & research assessment
88	evaluation system	38	0	open science & research assessment
91	funding agencies institutions	31	0	open science & research assessment
94	dora declaration	55	0	open science & research assessment
99	research funders	72	0	open science & research assessment
100	promotion tenure decisions	89	0	open science & research assessment
3	research articles	116	1	Outputs scientific research
6	ways output scientific	120	1	Outputs scientific research
7	research evaluated funding	133	1	Outputs scientific research
10	outputs scientific research	223	1	Outputs scientific research
28	research assessment reform	33	1	Outputs scientific research
59	agencies academic institutions	124	1	Outputs scientific research
61	data reagents software	94	1	Outputs scientific research
89	evaluated funding agencies	133	1	Outputs scientific research
97	software intellectual property	90	1	Outputs scientific research
5	recherche intra signe	66	2	INRA signe déclaration
79	intra signe d'@claration	66	2	INRA signe déclaration
81	évaluation recherche intra	66	2	INRA signe déclaration
8	best practice	162	3	INRA signe déclaration
36	challenge research assessment	131	3	INRA signe déclaration
47	influence specific research	143	3	INRA signe déclaration
65	value influence specific	143	3	INRA signe déclaration

74	specific research outputs	143	3	INRA signe déclaration
93	research assessment practices	183	3	INRA signe déclaration
95	assessment practices	37	3	INRA signe déclaration
11	case study	185	4	Academics career assessment & innovation
13	assessment policies practices	36	4	Academics career assessment & innovation
14	career assessment reform	87	4	Academics career assessment & innovation
16	reform stories innovation	83	4	Academics career assessment & innovation
26	space rubric	31	4	Academics career assessment & innovation
29	research assessment policies	33	4	Academics career assessment & innovation
37	academic assessment	75	4	Academics career assessment & innovation
49	career assessment	60	4	Academics career assessment & innovation
52	assessment reform stories	87	4	Academics career assessment & innovation
57	policy practice	53	4	Academics career assessment & innovation
70	authorship practices	2	4	Academics career assessment & innovation
75	dora idea action	25	4	Academics career assessment & innovation
76	stories innovation change	123	4	Academics career assessment & innovation
78	researchassessment reform	60	4	Academics career assessment & innovation
15	francisco declaration research	599	5	San Francisco Declaration
25	san francisco declaration	609	5	San Francisco Declaration
54	declaration research assessment	746	5	San Francisco Declaration
56	signs san francisco	106	5	San Francisco Declaration
90	research assessment dora	291	5	San Francisco Declaration
27	impact research outputs	161	6	Impact Metrics
43	value impact research	161	6	Impact Metrics
44	impact influence policy	164	6	Impact Metrics
46	research impact influence	157	6	Impact Metrics
62	range impact measures	194	6	Impact Metrics
68	influence policy practice	192	6	Impact Metrics
69	indicators research impact	162	6	Impact Metrics
18	scientific content	126	7	Scientific content & committees making decisions
19	publication metrics	173	7	Scientific content & committees making decisions
35	promotion tenure	109	7	Scientific content & committees making decisions
41	decisions funding hiring	130	7	Scientific content & committees making decisions
64	committees making decisions	96	7	Scientific content & committees making decisions
20	works journal impact	240	8	Journal Impact Factor

21	declaration addiction harmreduction	411	8	Journal Impact Factor
23	impact factor jif	428	8	Journal Impact Factor
45	dora declaration addiction	411	8	Journal Impact Factor
53	journal impact factors	824	8	Journal Impact Factor
71	factor jif addiction	411	8	Journal Impact Factor
72	dose dora declaration	411	8	Journal Impact Factor
73	harmreduction jif letsabandonthejif	411	8	Journal Impact Factor
92	jif addiction good	411	8	Journal Impact Factor
96	addiction good dose	411	8	Journal Impact Factor
98	addiction harmreduction jif	411	8	Journal Impact Factor
22	measure quality individual	592	9	Individual assessment & quality
30	scientists contributions hiring	480	9	Individual assessment & quality
34	promotion funding decisions	673	9	Individual assessment & quality
38	contributions hiring promotion	490	9	Individual assessment & quality
42	surrogate measure quality	576	9	Individual assessment & quality
67	quality individual research	608	9	Individual assessment & quality
83	individual research articles	565	9	Individual assessment & quality

Appendix G

Co-occurrence of hashtags

A co-occurrence network of hashtags was also performed to find general thematic clusters by using *CorTextT's* network mapping tool.

Similarly, the analysis of hashtags can provide information on the topics of discussion (Fig.G1). Nine topics have been identified through the co-occurrence of hashtags: "Open & events"; "Open science"; "Open & journals"; "REF" (Research Excellence Framework); "Research assessment and impact factor"; "Recognition and rewards"; "Academic chatter & phdchat"; "eLearning and assessment"; "sustainableopenaccess and journal impact factor".

More specific topics have been highlighted, such as events, discussions about the REF exercise in the United Kingdom (separated from research assessment), or tweets related to science-related Twitter accounts (e.g., Academic Chatter™). Moreover, we can observe the relevance of open science and its relation to different topics (i.e., events, sustainability).

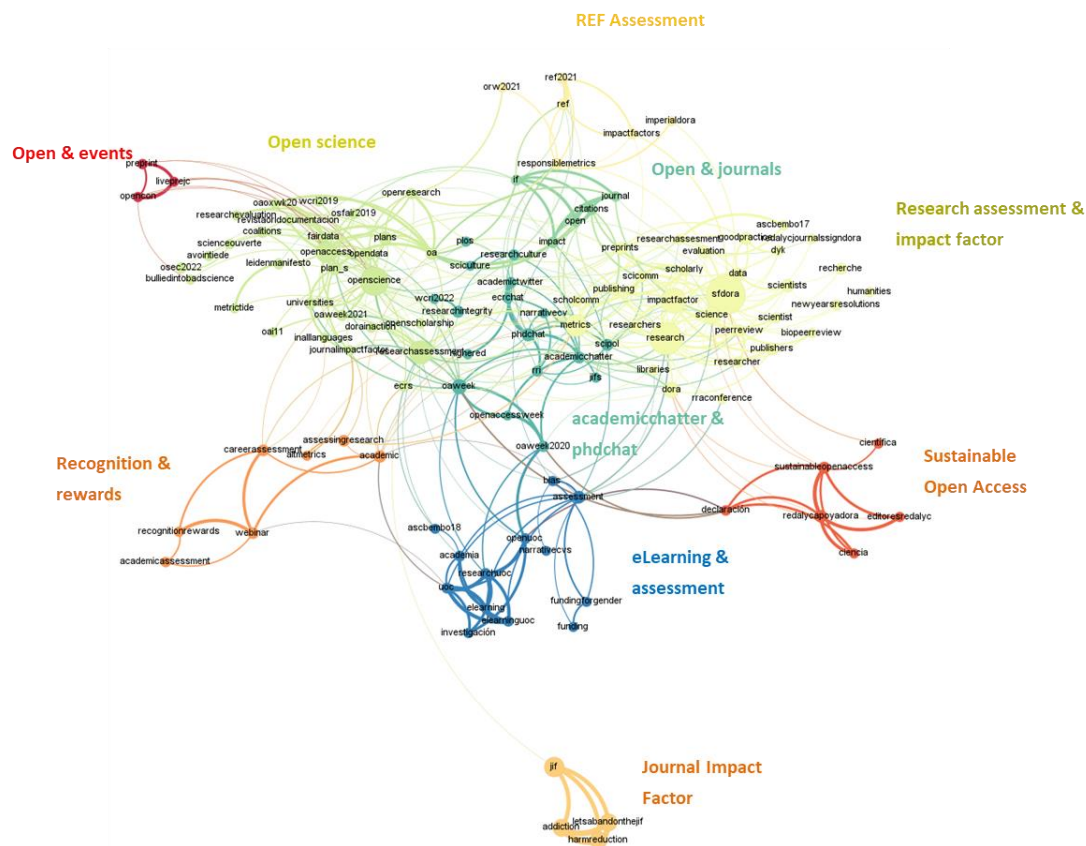


Fig.G1 Topics of discussion about DORA identified from a co-occurrence network of hashtags. Note: hashtags extracted and depurated with CorTextT (300 keywords). Map created with Gephi 0.10.1. Available at: <https://documents.cortext.net/lib/mapexplorer/explorerjs.html?file=https://assets.cortext.net/docs/fffc8783a05cc1ed16feb7ad2799d4e4#>

Appendix H

Individual tube lay-outs

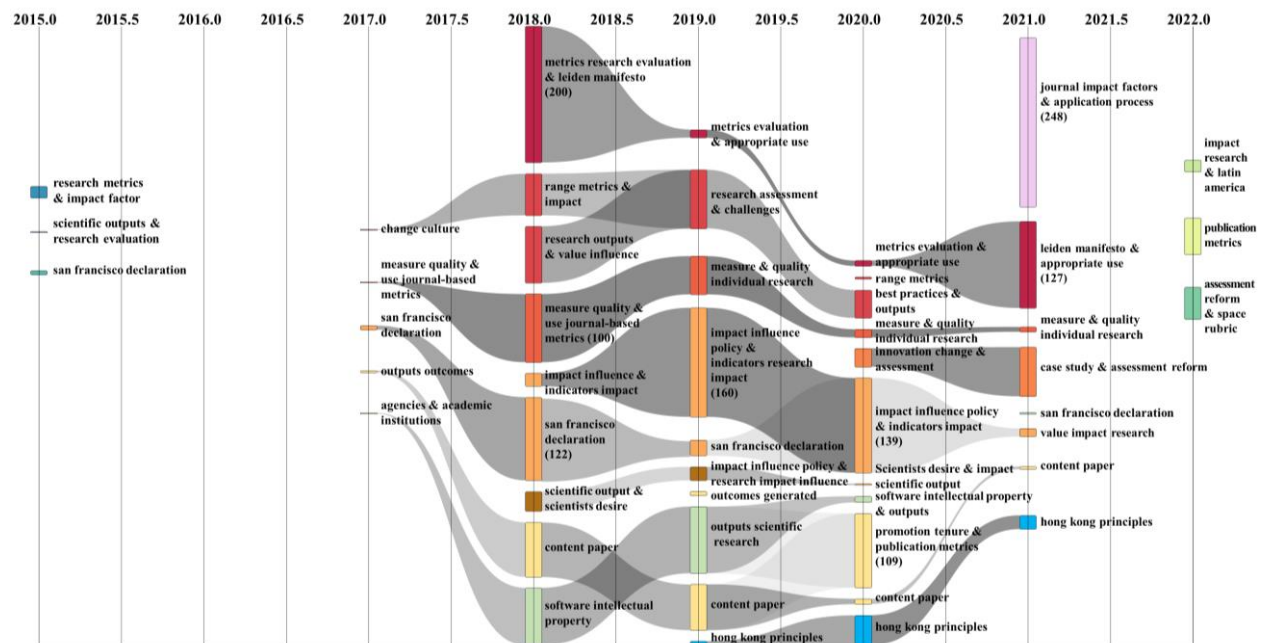


Figure H.1 Sankey diagram of the evolution of topics over time by the DORA account

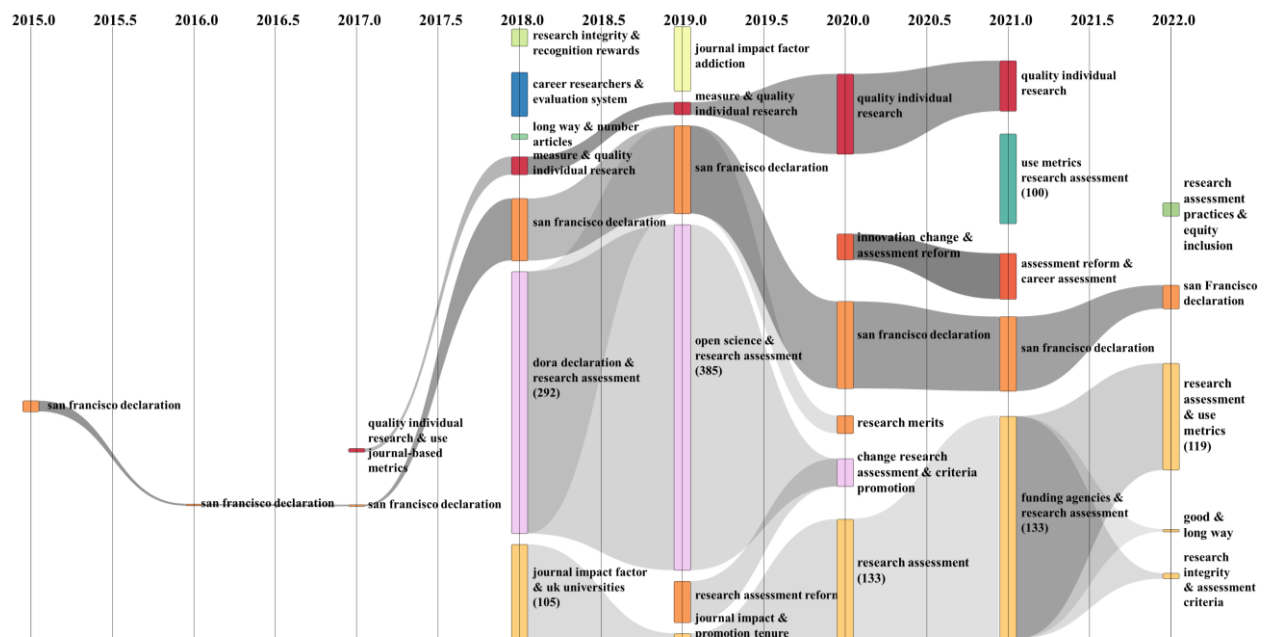


Figure H.2 Sankey diagram of the evolution of topics over time by the user-mention dataset

Note: Topics with more than 100 documents were indicated in brackets

Appendix I

Average sentiment analysis per month

month	General	DORA	Community	Hashtag to DORA
2015-04	1.500		1.500	
2015-05	0.700		0.700	6.000
2015-06	0.286		0.286	4.000
2015-07	0.375		0.375	0
2015-08	1.444		1.444	
2015-09	0.267		0.267	0
2015-10	-0.727		-0.727	3.000
2015-11	-0.421		-0.421	0.500
2015-12	1.333		1.333	1
2016-01	-0.667		-0.667	1.500
2016-02	0.250		0.250	2.429
2016-03	1.400		1.400	3.000
2016-04	2.400		2.400	4.000
2016-05	0		0	0
2016-06	0.833		0.833	2.000
2016-07	0.250		0.250	
2016-08	0		0	
2016-09	0.500		0.500	1
2016-10	-2.000		-2.000	
2016-11	0		0	-2.000
2016-12	2.000		2.000	
2017-01	0		0	
2017-02	1.389		1.389	2.000
2017-03	-6.000		-6.000	0
2017-04	1.676		1.676	1.750
2017-05	1		1	
2017-06	0.579		0.579	0
2017-07	0.867		0.867	0
2017-08	0.750		0.750	
2017-09	0.750		0.750	0
2017-10	1.897		1.897	
2017-11	2.423		2.423	
2017-12	1.818		1.818	0.750
2018-01	0.904		0.904	
2018-02	0.977		0.977	0.938
2018-03	1.549		1.549	4.000
2018-04	1.408		1.408	1
2018-05	1.190		1.190	1
2018-06	1.365		1.365	1.857

2018-07	1.302		1.302	1.071
2018-08	1.611		1.611	2.364
2018-09	1.508		1.508	1.091
2018-10	1.257	1.148	1.403	2.250
2018-11	0.945	0.891	1.323	2.400
2018-12	1.533	1.593	1.333	1.571
2019-01	1.441	1.429	1.457	1
2019-02	1.300	1.333	1.200	2.143
2019-03	1.086	1.071	1.127	-0.600
2019-04	1.525	1.563	1.429	1.385
2019-05	1.364	1.374	1.338	1.667
2019-06	1.648	1.776	0.803	1.600
2019-07	1.489	1.638	1.045	3.667
2019-08	1.370	1.575	1.029	0
2019-09	1.136	1.062	1.433	3.000
2019-10	1.279	1.332	0.885	1.167
2019-11	1.125	1.027	1.541	1.714
2019-12	1.165	1.112	1.313	
2020-01	1.355	1.500	1	0.857
2020-02	1.707	1.926	0.984	3.500
2020-03	1.160	1.245	1	2.250
2020-04	1.158	1.059	1.244	8.000
2020-05	1.454	1.651	1.014	0.500
2020-06	0.891	0	0.896	2.000
2020-07	0.877		0.877	1.500
2020-08	0.684		0.684	0.600
2020-09	1.344		1.344	
2020-10	0.835		0.835	1.375
2020-11	1.012		1.012	4.000
2020-12	0.763		0.763	1.200
2021-01	1.145		1.145	2.667
2021-02	1.115		1.115	1.333
2021-03	1.171		1.171	-1
2021-04	0.965		0.965	
2021-05	1.141		1.141	1
2021-06	1.171		1.171	1
2021-07	1.398		1.398	-0.667
2021-08	0.996		0.996	
2021-09	1.364		1.364	
2021-10	1.090		1.090	3.000

2021-11	1.348	1.472	1.278	1
2021-12	1.122	1.165	1.019	
2022-01	1.232	1.351	1	
2022-02	1.071		1.071	1
2022-03	1.119		1.119	
2022-04	1.558	1.846	1.485	1
2022-05	1.410		1.410	1.667
Average	0.954	1.326	0.918	

Appendix J

Sentiment analysis broken by impact for each hashtag category

The average sentiment polarity among categorization and impact metrics (likes, retweets, replies, and quotes) is summarized in the Figure J.1.

Breaking down the sentiment data by impact metric, we can observe that negative sentiments are the ones with the highest impact in the case of likes counts for specific topics such as Institutions (-3 sentiment score; 36 likes on average), Community (-5; 21) and Gender (-2; 20). On the contrary, retweets stand out with positive sentiment (10) in Peer Review (30 retweets on average) but are scored negatively when it comes to Gender (-2 sentiment score; 17 retweets), Career (-1; 16.5) and Ethics topics (-1; 16.3). Otherwise, although replies are scarce, they attain a higher impact regarding Data Source (4 sentiment score; 2 replies) and DORA support (3; 2). Similarly, quote counts show a higher influence on negative perceptions, especially for Community (-5 sentiment score; 3 quotes), Institutions (-3; 2), or Metrics (-6; 2).

Some topics present a higher polarity, suggesting that different opinions (either positive or negative) are embraced, such as the Peer review category impact in terms of retweets (one tweet with a sentiment score of -5 with 12 retweets on average, and other tweets with a sentiment score of +10 with 30 retweets on average), the Metrics category in terms of quotes (sentiment score of -6 with two quotes, and sentiment score of +6 with 1.4 quotes on average), and the Location category in terms of replies (sentiment score of -5 with one reply on average, and sentiment score of +3 with 1.8 replies on average). The Open category is also noteworthy in terms of quotes (sentiment score of -8 with one quote and sentiment score of +7 with 0.7 quotes on average). These results denote each metric provides supplementary information on output-level impact and sentiment, which differs among topics.

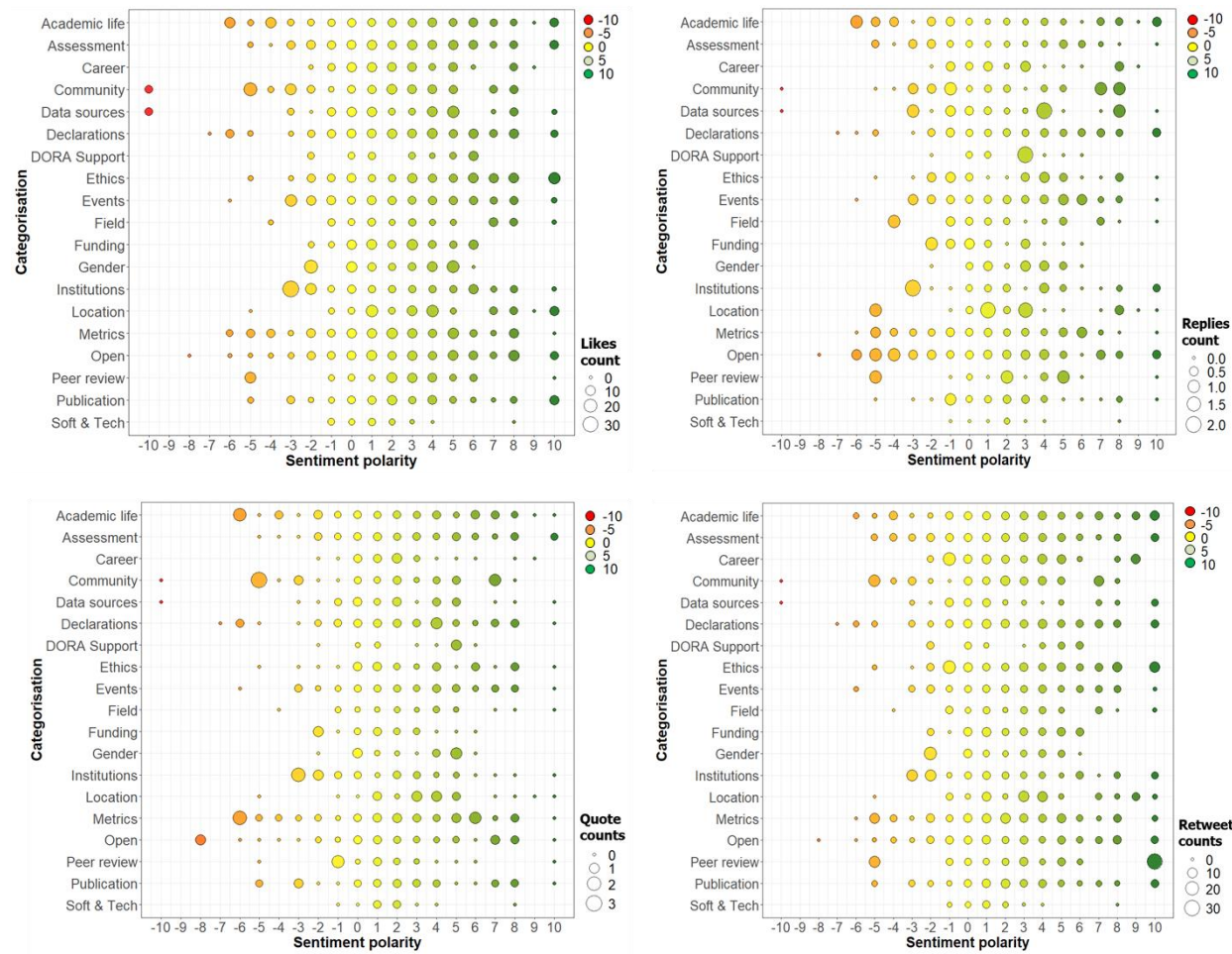


Fig.J.1 Average sentiment polarity among categorization and links (upper left), replies (upper right), quotes (lower left), and retweets (lower right)

Appendix K

Output-level impact broken by hashtag category

The impact of the tweets has been broken down according to the hashtag categories, observing remarkable differences according to the category, impact metric, and the origin of the tweet (Table K.1).

For example, both datasets have generated a similar number of tweets containing hashtags related to "Metrics." Of these, the tweets published by DORA achieve more likes and quotes, but the tweets posted by the Twitter community that mention DORA are more retweeted.

Tweets about "Assessment" and "Declarations" come mainly from the DORA dataset but are more replied to when published by the user-mention dataset. Tweets about "Scientific publication" are posted primarily by the user-mention dataset but are more retweeted when they come from the DORA dataset. Tweets about "Events" and "Open" achieve more relevance in the user-mention dataset.

Table K1 Publication and impact of tweets about DORA broke down by hashtag category

HASHTAG CATEGORY	TWEET COUNT			LIKE COUNT			REPLY COUNT			QUOTE COUNT			RETWEET COUNT		
	UMD	DD	ALL	UMD	DD	ALL	UMD	DD	ALL	UMD	DD	ALL	UMD	DD	ALL
Assessment	619	899	1,518	3,104	6,942	10,046	164	60	224	182	385	567	1,706	5,109	6,815
Declarations	380	880	1,260	2,688	6,206	8,894	171	102	273	247	391	638	1,557	5,672	7,229
Open	942	154	1,096	6,632	1,103	7,735	387	17	404	344	48	392	3,864	1,130	4,994
Academic life	492	412	904	2,477	2,580	5,057	107	45	152	150	201	351	1,431	2,725	4,156
Metrics	476	427	903	1,814	4,721	6,535	166	69	235	108	295	403	1,024	3,882	4,906
Events	707	160	867	3,985	577	4,562	245	10	255	213	48	261	2,275	747	3,022
Scientific publication	231	125	356	850	847	1,697	52	14	66	43	61	104	500	906	1,406
Institutions	223	39	262	1,172	112	1,284	54	1	55	66	8	74	615	249	864
Community	100	87	187	624	670	1,294	26	10	36	29	47	76	386	601	987
Field/Discipline/Subject	139	45	184	421	206	627	36	4	40	19	9	28	270	169	439
DORA Support	93	29	122	602	7	609	29	0	29	25	0	25	410	221	631
Gender, diversity, equality & inclusion	56	29	85	349	328	677	18	3	21	17	26	43	140	184	324
Position/Career/Award	44	34	78	236	268	504	11	6	17	15	15	30	197	246	443
Data sources	67	10	77	298	113	411	20	3	23	21	5	26	137	69	206
Ethics	64	13	77	436	53	489	21	0	21	22	4	26	378	65	443
Funding	50	23	73	402	217	619	28	2	30	12	10	22	168	192	360
Location	58	12	70	208	82	290	22	0	22	7	3	10	114	58	172
Peer review	31	31	62	162	117	279	9	2	11	2	8	10	78	168	246
Software & Technology	38	3	41	54	26	80	1	0	1	5	3	8	28	26	54
TOTAL	4,810	3,412	8,222	26,514	25,175	51,689	1,567	348	1,915	1,527	1,567	3,094	15,278	22,419	37,697

UMD: user-mention dataset; DD: DORA dataset.

Note: Normalizing by dataset size would generate less informative low values as the values are not elevated. However, this should be considered when contextualizing the results. The temporal evolution of hashtag categorization is available in Table S.3.

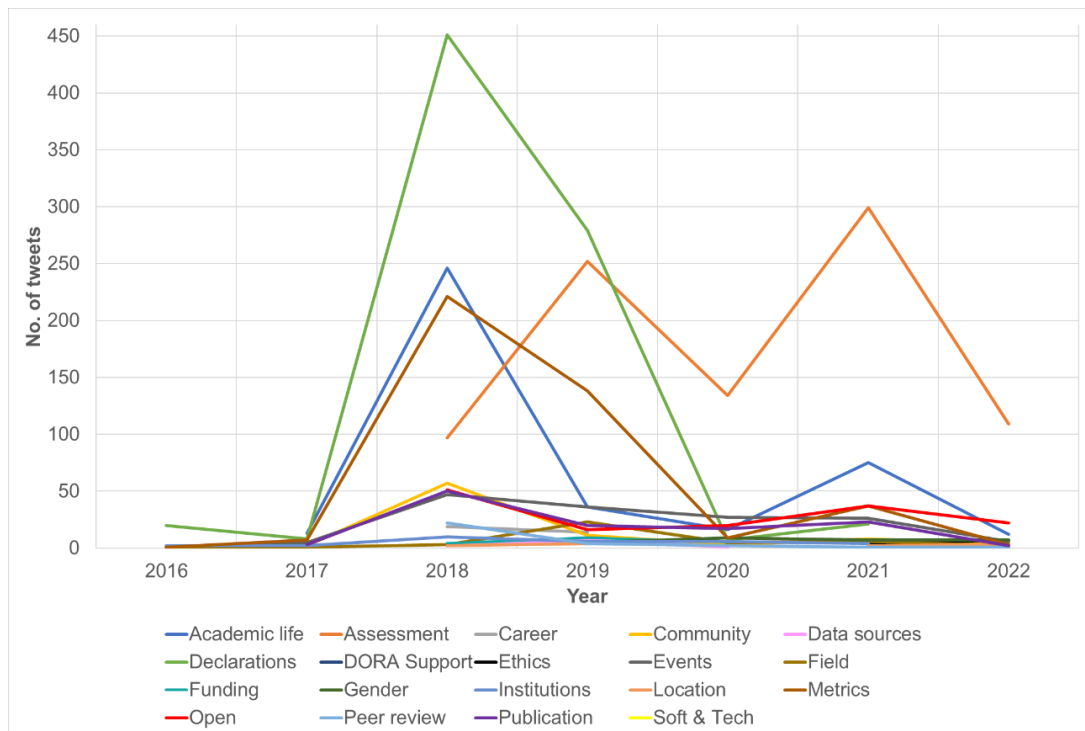


Figure K.1 Temporal evolution of hashtag categorisation (DORA dataset)

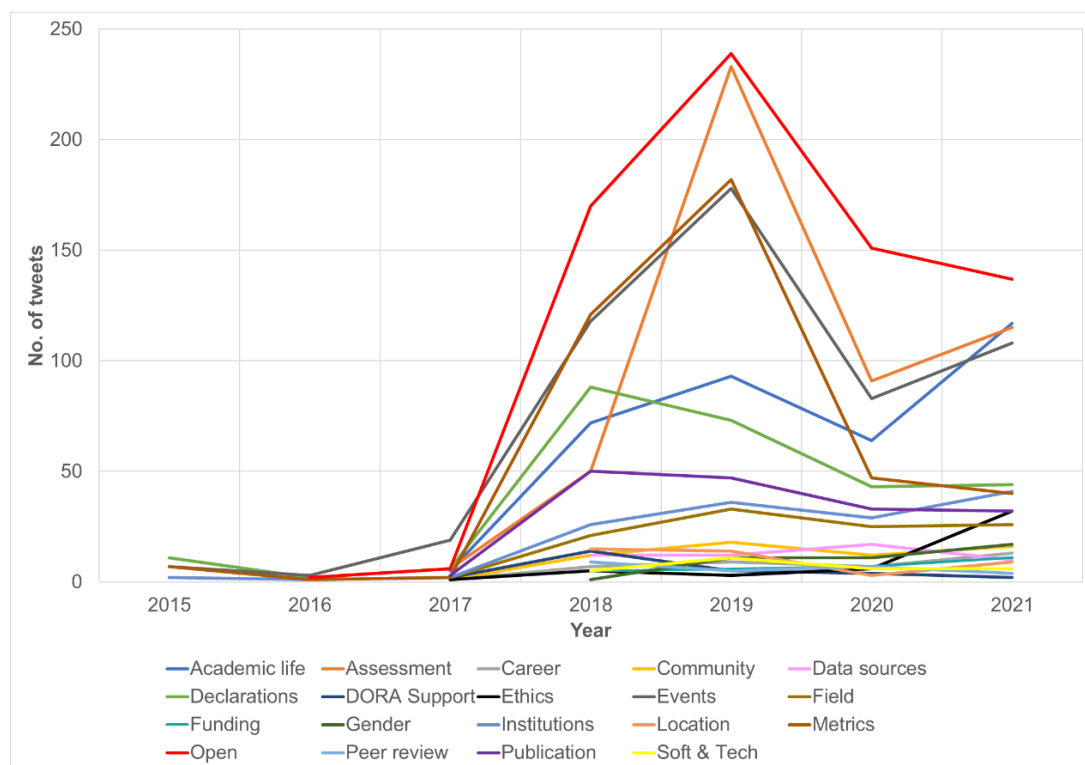


Figure K.2 Temporal evolution of hashtag categorisation (user-mention dataset)