



HAL
open science

Using Social Media Bots to Keep up with a Vibrant Research Field: The Example of @MOF_papers

François-Xavier Coudert

► **To cite this version:**

François-Xavier Coudert. Using Social Media Bots to Keep up with a Vibrant Research Field: The Example of @MOF_papers. *Chemistry of Materials*, 2023, 35 (7), pp.2657-2660. 10.1021/acs.chemmater.3c00383 . hal-04065672

HAL Id: hal-04065672

<https://hal.science/hal-04065672>

Submitted on 12 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Using social media bots to keep up with a vibrant research field: the example of @MOF_papers

François-Xavier Coudert^{a,*}

^a Chimie ParisTech, PSL Research University, CNRS,
Institut de Recherche de Chimie, Paris, 75005 Paris, France

* fx.coudert@chimieparistech.psl.eu ; <https://www.coudert.name>

Introduction

The number of academic papers published worldwide is a steadily increasing function of time. The cause for this is multifactorial and is due in part to the strong investment and rapid growth in scientific research from China¹ and a number of emerging countries.² It can also be ascribed in part to the pressure on individual researchers to publish more and more, as performance metrics are used in too many hiring, promotion, and award committees.³ This increase occurs across the board and impacts all fields of research, although in some very novel areas the progression is even faster than average. Therefore, it can be overwhelming for researchers to stay informed about the latest developments and trends in their field, and avoid missing published papers that would be key to their own research. This is particularly true of students and early-career researchers (graduate students and postdocs in particular), who do not necessarily have the benefit of a long experience in the field and a good network in place, as well as those who enter a new field or work in highly multidisciplinary areas.

To help researchers keep up with the rapidly increasing literature available online, several solutions can be used. Some are software-based and employ either proprietary publication databases and their alert capabilities (the two best-known examples are probably Web of Science⁴ and Scopus⁵) or free online services that offer saved searches and inbox alerts (Google Scholar⁶, for example). Another way is to regularly perform exhaustive literature reviews or to read published reviews, but both solutions necessarily create a lag between the publication of a paper and its discovery. Crowdsourcing is also a common way to stay on top of the literature, through informal discussions with colleagues — whether they occur in the physical world at journal clubs and conferences, or online on social platforms.

Most of us spend a nontrivial part of our time online on social media, and platforms such as Twitter have shown a clear power in accelerating the dissemination of information to a large readership.⁷ In the academic world, it has impacted both the research enterprise and higher education.⁸ For these reasons, I created in April 2014 a tool for keeping track of the new papers published in the field of metal–organic frameworks (MOFs), and related materials. @MOF_papers was born and started

tweeting on 25 April 2014, originally without any software: it used a combination of free online services, based on Yahoo! Pipes, to filter the journal feeds and post the new papers to Twitter. In June 2018, after the closing of the previous platform, I replaced it by a Python program I had written, PapersBot, that did the same job and could be used by others as well. Here, I present the working principles of PapersBot and present some of the insight I have gained from operating the @MOF_papers bot.

Design of the software

The design of the PapersBot software is relatively simple (see Figure 1): it pulls information about the latest published articles in a series of journals entered in its database, filters it according to specific keywords, and then pushes the content to social media like Twitter and, most recently, the federated social network Mastodon. While most journals and publishers have ways to set up personalized alerts on their website, all of them also share the information about latest published articles in a simple standardized format: web-based RSS feeds. These are simple files with XML syntax and consistent content across journals, allowing to identify papers and their metadata — without having to parse the details of the webpages of each journal.

From all the RSS feeds of individual journals, the content used is merged into a single feed and filtered according to keywords or regular expressions. The PapersBot code is flexible for this filtering, and for the @MOF_papers bot I chose a simple keyword-based filter, matching on both paper title and abstracts: the bot will select papers featuring the terms *metal-organic framework*, *covalent organic framework*, *imidazolate framework*, *porous coordination polymer*, and *framework material*. All new papers corresponding to these criteria are then tweeted: the tweet includes the paper title, the publisher link, and an image if one was present in the original RSS feed. Papers that have been posted are marked in the bot's database through their DOI, to avoid potential duplication — for example, items can be posted again by the publisher when they are assigned page numbers. The monitoring is performed automatically once every hour, based on the GitHub Actions platform. The same bot can monitor feeds and post to multiple social media platforms, with the recent introduction of a Mastodon feed at https://mstdn.science/@MOF_papers in November 2022, in addition to the original Twitter feed at https://twitter.com/MOF_papers.

Over the years, various additional features have been tested. In particular, the PapersBot code can recognize known journals and tag the corresponding journal Twitter account when the paper is posted. This was well received by both users and journal community managers, but led to the @MOF_papers bot being suspended for spamming in August 2018: the feature has been disabled since then.

Finally, the PapersBot code base is made freely available as open-source software (available at <https://github.com/fxcoudert/PapersBot>) and has been used by several other research groups to maintain and publish their own literature survey bots. Table 1 shows a list of such bots I am aware of, in chemistry, materials sciences and connected fields. We can see that the tools have a diversity of use: some of them have a broad topic, heavy traffic and many users, while others focus on more specialized topics. Because

of their sometimes related topic, there is some overlap between their followers but taking this into account, they account together for more than 37,000 unique users. This number highlights the importance that these paper bots have attained in our fields.

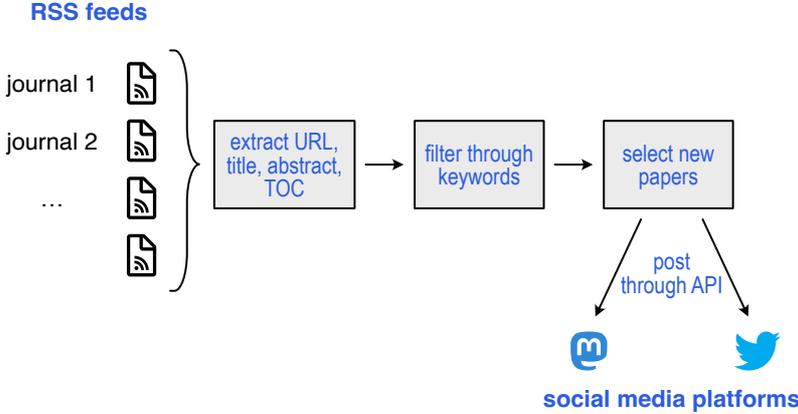


Figure 1. Overview of the PapersBot workflow.

Bot	Topic	Creation	Followers
@MOF_papers	MOFs, COFs, PCPs	April 2014	11,816
@cryoEM_Papers	cryogenic electron microscopy	August 2018	11,307
@BatteryPapers	battery materials	April 2020	4,173
@COF_papers	covalent-organic frameworks	February 2019	3,688
@ECat_papers	electrocatalysis	June 2018	3,160
@POP_Papers	porous organic polymers	December 2017	2,565
@POC_Papers	molecular/supramolecular cages, organic or coordination	April 2017	2,014
@CO2_papers	carbon dioxide	June 2018	1,673
@MechChemPapers	mechanochemistry	January 2020	1,553
@Zeolite_papers	zeolites	June 2018	978
@PharmSolidForms	pharmaceutical solid forms	November 2015	454
@mat_mol_ml	ML in chemistry and materials science	March 2019	355
@SpinStatePapers	spin-states in transition-metal complexes	January 2015	342
@methox_papers	methane oxidation	June 2018	206
@Flow_Papers	flow chemistry	September 2021	126
@GisansPapers	neutron reflectivity and GISANS	August 2018	62
@IT_papers	ice templating	February 2014	41
@ChemGardenPaper	chemobionics	October 2022	15

Table 1. List of paper bots in chemistry and materials sciences, not necessarily exhaustive, along with topic, date of creation and number of followers (as of 12 February 2023).

Users of the @MOF_papers bot

On 12 February 2023, the @MOF_papers bot had 11,816 followers. In Twitter terminology, a *follower* is a user who had subscribed to a given account, whose posts will then appear on the main chronological page (the *timeline*). We also have anecdotal evidence that some colleagues use the bot without subscribing (or even without being Twitter users themselves), simply by bookmarking its profile page in their browser.

Out of the 11,816 followers, we could analyze the geographical localization of 6,687 users. Not all users chose to indicate their location, and because location is a free text field, in some cases it could not be mapped to geographical coordinates by the Nominatim geocoder. Figure 2 presents a map with

the number of followers per country. We find that the bot has users from 125 countries, spanning all over the world — with notable exceptions in areas where scientific research is underrepresented, including Africa and central Asia. Most of the countries with many followers are countries with a large research activity and population (USA, China, India, etc.). However, there are some countries that show a large number of followers compared to their respective research output or number of researchers: Thailand, Mexico, Turkey, Brazil and Pakistan. From comments posted online by users from these countries, it appears clear that there is an added value for the bot in countries where access to expensive proprietary databases is rare because research funding is insufficient. In addition, I note that China is underrepresented in the geographical distribution (5.0% of followers, compared to e.g. 16.4% in the USA) due to Twitter is officially blocked in China, and access being possible only through a virtual private network (VPN). Recent papers are also distributed by other channels more accessible to Chinese researchers, including WeChat messaging groups ranging in size from single research group to multiple universities. Some groups including selection or editorialization of the content, as well as translation into Chinese.

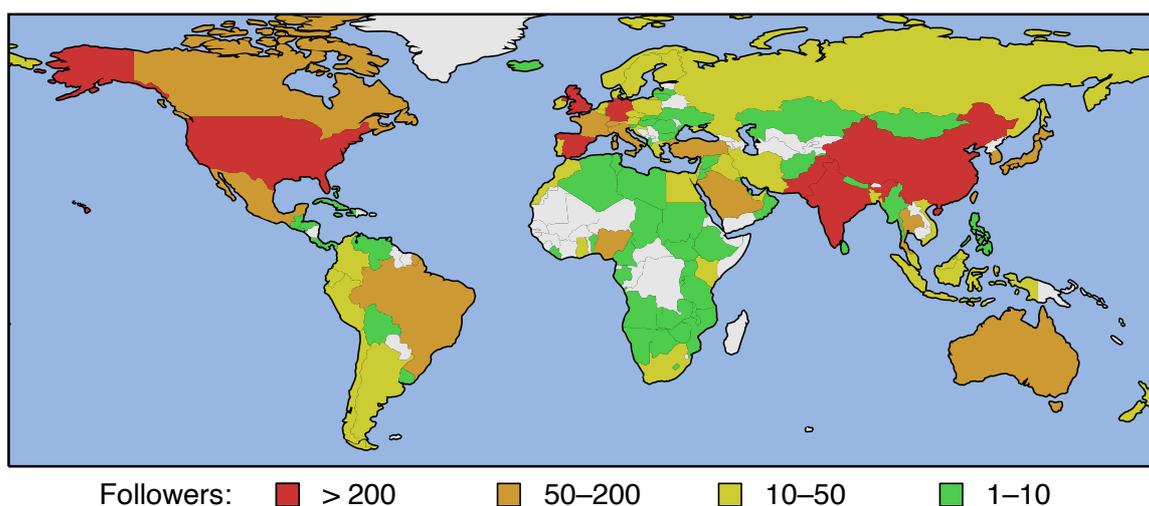


Figure 2. Map of @MOF_papers followers per country, when their profile location could be geocoded.

In order to better understand the user base of the bot, a simple survey was run through Twitter itself (as a series of poll questions) from 12 to 15 February 2023. It shows that followers are active in public higher education, with only 3.6% of respondents working in the private sector, and 1.2% being retired. The main category of users is graduate students, which represent 45.0% of respondents. There are relatively few undergrad students (3.9% of responses), probably because they are more focused in their literature review, and may not develop long-term workflows. Post-doctoral researchers represent

16.9 % of users, while 28.3% of respondents are academic staff (most of them having both research and teaching responsibilities). @MOF_papers therefore has a broad distribution of users across academia.

About their usage of the tool, a vast majority respondents cite only research (93%), while 7% say they use it for both research and teaching. More surprising to me is the distribution across device: 56% says they check the bot on their phone, 12% on a computer, and 32% on both phone and computer. This statistics shows the blurring of lines between professional and personal devices, where someone can check the latest papers when they have to wait for a bus, or similar situations. Finally, 82.0% of users who answered that question identified as male, 17.6% as female, and 0.4% as nonbinary or other. While this is a more unbalanced gender distribution than academic chemical research in the US, it may be due to a large number of users from countries where gender balance in higher education and research is far from parity.

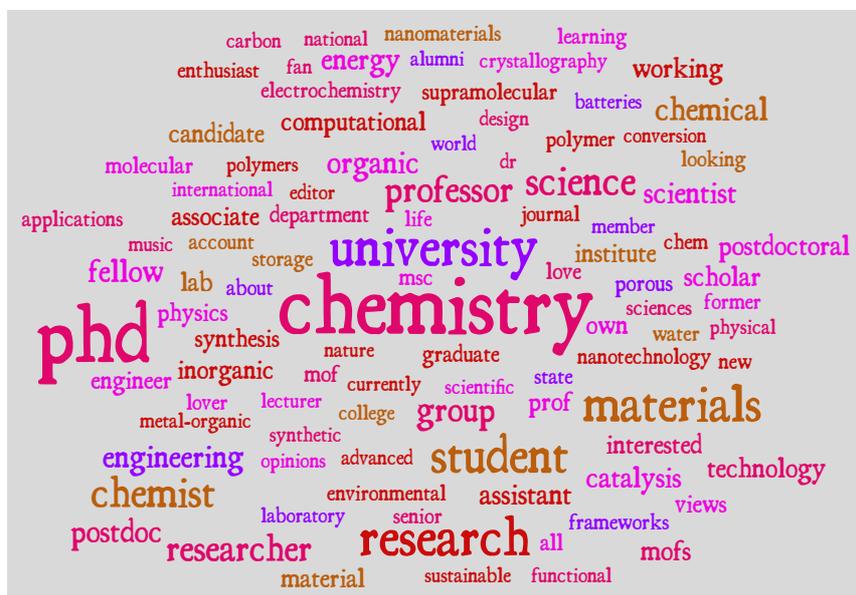


Figure 3. Most frequent words, and associated frequencies, featured in profiles of @MOF_papers followers. Analysis is restricted to the Latin alphabet, common words and country names have been excluded.

Another way to learn more about the users of the bot is to data-mine the information publicly available in their Twitter profiles. Figure 3 shows a word cloud of the most common occurrences. Beyond the very common words such as “chemistry” and “university”, we can see quite well the most common themes and applications of MOFs and nanoporous materials more generally: catalysis, energy, environment, batteries, sustainability, nanotechnology, carbon capture, etc. We also see the large diversity of subfields involved in our community, as well as from synthesis to engineering, from organic and inorganic to physical and computational chemistry.

Finally, Figure 4 shows the 24 most frequent emojis featured in user profiles (once national flags have been excluded). We find a lot of different symbols related to chemistry (test tube, microscope, alembic, both men and women figures wearing a lab coat) used by people to show their professional occupation, but we also find some more specific emojis related to subfields: high voltage and battery emojis are commonly used to indicate energy materials and batteries; DNA helix and Petri dish for biology; sun for solar energy and related materials; seedling and earth emoji are commonly used for environmental sciences; etc. We note that the most common emojis not science- or academic-related are quite universal: rainbow and rainbow flag, football, books, and coffee.



Figure 4. Most frequent emojis in followers’ profiles. National flags were excluded from the analysis.

Perspectives

The @MOF_papers bot is quite active, tweeting an average of 400 papers per month, and generating 500,000 “impressions” (unique tweet views) per month. More than just reading the content posted by the bot, the community has created other uses — not originally intended. Several groups and journals tag their new publications with the @MOF_papers handle, and others can use that as an “alert” to see content not only posted by the bot, but where it is tagged. This allows to communicate on papers that the bot misses — it happens, and I then try to tweet them manually — or on other community events, such as conferences. I have been solicited a few times by conference organizers to make the bot an “official” part of the conference (retweeting content, such as lectures announcements or online posters), but I have declined as I felt it would bias the coverage of the bot in some way.

Beyond its current activity, it is possible that there would be other uses for a “paper bot”. Right now, although the bot is quite successful, its main limitation is that it is entirely run by one person on free resources (the GitHub servers through their Education plan). Moreover, recent changes in the leadership and direction of Twitter towards monetization have threatened its existence. In this spirit, I opened in November 2022 a Mastodon account to have presence on at least two social media platforms, but so far the Mastodon account is far less popular, with 55 followers and very little interaction of users with the bot. It remains to see whether opening to more platforms (such as Instagram and Facebook)

would be useful to the community, or if Twitter remains in the future the common “watering hole” of our online chemistry and materials research community.

Acknowledgments

This work is supported by the Agence Nationale de la Recherche under project MESAMM (ANR-19-CE05-0031).

References

- ¹ Brainard, J.; Normile, D. China rises to first place in most cited papers, *Science*, 2022, <https://doi.org/10.1126/science.ade4585>
- ² Adams, J. The fourth age of research, *Nature*, 2013, 497, 557–560 <https://doi.org/10.1038/497557a>
- ³ Binswanger, M. (2014). “Excellence by Nonsense: The Competition for Publications in Modern Science.” In: Bartling, S., Friesike, S. (eds) *Opening Science*. Springer, Cham. https://doi.org/10.1007/978-3-319-00026-8_3
- ⁴ Web of Science, Clarivate. <https://clarivate.com/webofsciencegroup/solutions/web-of-science/>
- ⁵ Scopus, Elsevier. <https://www.scopus.com/>
- ⁶ Google Scholar. <https://scholar.google.com/>
- ⁷ Tapper, E. B.; Mirabella, R.; Walicki, J. J.; Banales, J. M. *J. Hepatology*, 2021, 75, 271–274 <https://doi.org/10.1016/j.jhep.2021.05.020>
- ⁸ Malik, A.; Heyman-Schrum, C.; Johri, A. Use of Twitter across educational settings: a review of the literature. *Int. J. Educ. Technol. High Educ.*, 2019, 16, 36 <https://doi.org/10.1186/s41239-019-0166-x>