HOW OPEN SCIENCE ELEVATES COLLABORATION IN COVID-19 RESEARCH - THE ROAD TO OPEN SCIENCE

Emilie HERMANS
National Open Access Desk *OpenAIRE*, Ghent University Library
Inge VAN NIEUWERBURGH
Co-ordinator scholarly communication, Ghent University Library

■ The current COVID-19 pandemic represents a significant threat that requires speedy knowledge exchange and transparency. It has undoubtedly demonstrated just how critical openly available scientific information can be for society. Where the realization of a fully open and transparent research practice with accessible and interoperable research outputs is still under construction, existing Open Science initiatives, policies and infrastructures now show effective ways to accelerate COVID-19 research. In this article we discuss how Open Science has become a crucial mindset when addressing current research practices, how Open Science principles show their strength in times of crisis and how existing infrastructures offer the possibility to be deployed as aggregators of open COVID-19 research.

■ De huidige COVID-19-pandemie is een bedreiging die snelle kennisuitwisseling en -transparantie noodzakelijk maakt. Deze tijd toont onweerlegbaar aan hoe doorslaggevend vrij beschikbare wetenschappelijke informatie kan zijn voor de maatschappij. Hoewel de realisatie van een volledig open en transparante onderzoekspraktijk, met toegankelijke en uitwisselbare onderzoeksresultaten, nog verder moet worden ontwikkeld, geven beleid en infrastructuur nu al effectieve manieren aan om COVID-19-onderzoek te versnellen. In dit artikel bespreken we hoe Open Science een cruciale manier van denken is geworden in de huidige onderzoekspraktijk, hoe Open Science-principes hun sterkte aantonen in tijden van crisis en hoe bestaande infrastructuur kan worden ingezet als verzamelaars van open COVID-19-onderzoek.

■ La pandémie actuelle de COVID-19 est une menace qui exige un échange rapide de connaissances et de la transparence. Cette période montre de manière incontestable combien l'information scientifique librement disponible peut être décisive pour la société. Bien que la réalisation d'une pratique de recherche totalement ouverte et transparente, avec des résultats de recherche accessibles et échangeables, doive encore être développée, la politique et l'infrastructure signalent déjà des moyens efficaces d'accélérer la recherche COVID-19. Dans cet article, nous examinons comment la Science Ouverte est devenue un mode de pensée crucial dans la pratique actuelle de la recherche, comment les principes de la Science Ouverte démontrent leur force en temps de crise, et comment les infrastructures existantes peuvent être déployées comme collecteurs pour une recherche COVID-19 ouverte.

We live in a paradoxical information world. To say that we are overwhelmed with information is a substantial understatement today. Everyone, from private citizens to public decision makers, regularly encounters an overabundance of information and must implement strategies and filters to select or deflect incoming information (Walgrave & Dejaeghere, 2017). Barriers are frequently established to restrict the deluge of information. In contrast, advocates of Open Science have been promoting broader access to scholarly information, which is often paywalled, though paid for by public funds, since the dawn of the internet. The possibilities of a worldwide platform induced the idea of a free flow of scientific information with the aim to make discovery, collaboration and broad dissemination easier. Combined with frustrations over an expensive and often slow publishing system, the view gained support that Open Access (OA), and more broadly Open Science, is the way forward to an alternative way of doing research.

FOSTER[1] defines Open Science as: '*... the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data.*' Open Science is about increased accessibility, transparency, accountability, reproducibility and trustworthiness of research. It is based on the principles of inclusion, fairness, equity, and sharing, and ultimately seeks to change the way research is done, who is involved and how it is valued. It aims to increasingly open up research to participation, validation and (re)use for the world to benefit.

Open Science encompasses a variety of practices, usually including areas like open access to publications, open research data, open source software/tools, open workflows, citizen science, open educational resources, and alternative methods for research evaluation including open peer review (Pontika et al., 2015).

The goal of full Open Access to publications is yet to be fulfilled, however progress has been made. It is estimated that the global proportion of openly accessible scholarly literature is at about 28% and is growing with the highest percentage of OA (45%) in the most recent year analysed, 2015 (Piwowar H, Priem J et al. 2018). This still leaves a massive amount of publications closed to the majority of

readers, particularly those working in poorer countries, where institutions struggle to afford costly academic journal subscription fees.

In addition to open access to publications, open data, another Open Science theme, has gained momentum. To be able to assess the quality of a publication, an insight in the data and the methodology of the research is necessary. Moreover, research data are a source of information on their own and can be reused across research projects and disciplines. To trust data to be re-usable, open data wasn't enough, though. While the emphasis on open was elementary in the first few years, the focus has shifted to Findable, Accessible, Interoperable en Re-usable (FAIR) principles for data (Wilkinson et all, 2016). FAIR was introduced to highlight the main characteristics of carefully managed research data. In the end all open data should be FAIR, though not all FAIR data is necessarily open, following the principle "*As open as possible, as closed as necessary*". Some data should be protected, e.g. for privacy reasons. But still closed data should be FAIR as well, even if they are only accessible for the happy few.

With open access to publications and open data, the aim of a free flow of research output approaches. However, it also highlights that a lot of research information is scattered over different resources. An infrastructure linking and enriching this information, making research comprehensible, findable and contextualized, eases the situation. The Open Access Infrastructure for Research in Europe (*OpenAIRE*) is doing just that for years now. The infrastructure gathers publications, data, project information and research information from a variety of disciplines and sources across Europe and beyond, and adds services on top, making it a useful resource, for example to find COVID-19 information.

The implications of Open Science: lessons learned from a previous epidemic

The necessity of a direct and open way of communicating scholarly outcomes is highlighted in times of crisis, as seen with the 2013–2016 Ebola outbreak in West Africa and now again in the COVID-19 global pandemic.

The Ebola outbreak was both an example of how crucial access to information is and how Open Science can fuel fast track solutions. An illustrative example of how crucial information is missed because of lack of access, was highlighted by the New York Times.

They found that the original detection of Ebola in Liberia was underestimated because of the incorrect "conventional wisdom" that Ebola was not present in that part of Africa:

"*The conventional wisdom among public health authorities is that the Ebola virus, which killed at least 10,000 people in Liberia, Sierra Leone and Guinea, was a new phenomenon, not seen in West Africa before 2013. (The one exception was an anomalous case in Ivory Coast in 1994, when a Swiss primatologist was infected after performing an autopsy on a chimpanzee.)* (NY Times, 2015)."

But it turned out that Ebola had been detected in Liberia before. Papers behind paywalls noted significant evidence of antibodies to the Ebola virus in Liberia and in other nearby nations. Studies published in the eighties stated that Ebola antibody prevalence rates suggested the possibility of what some call "sanctuary sites," or persistent, if latent, Ebola infection in humans.

Of course, that doesn't mean that if these articles would have been openly available the Ebola outbreak could have been avoided, though it would have alerted health workers faster. However, the information had still to be found, accessed and passed down to Liberian doctors and health officials in a world that still suffers from great inequality when it comes to resources, access and health. The NY Times article concludes with a call for equity in the broadest sense. And this includes us questioning why, even today, downloading one of the papers would cost $45.

Luckily the outbreak also proved that international collaboration in health and science can be a major success. Months within the largest Ebola outbreak in history, an international group of researchers sequenced three viral genomes, sampled from patients in Guinea (Baize, S. et al., 2014). The data were made publicly available that same month. This rapid release of genomic data sparked collaboration from experts from diverse disciplines, allowing for a better understanding of the pathways of infection, epidemiologic spread, categorisation of disease, and a more effective and humane prevention, treatment and care (Yozwiak, N. et al, 2015). Scientific journals tried to facilitate the flow of data and analyses by opening up rules around publishing pre-printed data, removing paywalls, and increasing expectations regarding shared attribution (Abramowitz, S. et al, 2019).

The experiences with the Ebola virus attributed to the approach of the COVID-19 virus. Many publishers and service providers have stepped up to provide broader access to related research. Previous Open Science efforts now serve in effective ways to accelerate COVID-19 research. However, not all pitfalls could be avoided. As highlighted by Lonni Besançon et al.

(Lonni Besançon et al., 2020) more caution on the use of preprints by popular media is recommended. Preprints play an increasingly important role in the fast circulation of research, since many are openly available. Researchers know these preprints haven't been peer reviewed yet and as such have to be treated with caution, which they do. However, more and more journalists and the broader public have access as well, not always realising these preprints are part of the discussion in research. Lonni Besançon et al. highlight the issue that "*many journalists may not be trained to understand and navigate the complex academic publication system, and some journalist may be seeking sensationalist news headlines*" (Lonni Besançon et al., 2020). Findings from preprints should be communicated with particular caution, which Robbe De Graeve and Peter Brems also address in an article on VRT NWS "*Let op met pre-prints*" (Robbe De Graeve & Peter Brems, 2020).

Existing infrastructures accelerate COVID-19 research: OpenAIRE COVID-19 Gateway and Zenodo COVID-19 community

The Covid-19 outbreak and previous health crises articulate the importance of rapid Open Access to current research to ensure that the latest critical data and research is accessible to those who most need it without restriction. It is thus crucial that scientific responses are based on international collaboration that brings together the best minds and available data from different countries for the benefit of all. Today COVID-19 emphasised the fact that global challenges require global solutions. The necessity for a collaborative effort that is efficient and fast, requires information to be accessible widely and openly, a challenge Open Science practices and infrastructure are prepared for. Useful models and infrastructure for responsible sharing of research outcomes have been developed by the broader Open Science community. Let's take a closer look at the response of two of them: *OpenAIRE* and *Zenodo*.

OpenAIRE is a European endeavour aiming at facilitating Open Science. It is a European community and service driven framework providing a support network for enabling Open Science. A human network of supporting agencies for Open Science in 34 European countries assists in promoting Open Science in the broad sense whether on a practical or policy level.

Besides this human network, services are provided based on the infrastructural pillar of *OpenAIRE*. With the overarching goal of providing an infrastructure linking and enriching research results, the *OpenAIRE* research Graph[2] establishes an open and sustainable scholarly communication infrastructure responsible for the overall management, analysis, enriching, provision, monitoring and cross-linking of all research outcomes. This is achieved by linking information stored in repositories and by many other content providers for scholarly publications and data, connected to project information. Ultimately all this information is presented in one place, establishing an open and sustainable scholarly communication infrastructure where projects are linked to their publications, datasets, funder information and much more. It provides a huge advantage for e.g. institutions and funders as they can now discover all research output per project, funder or data provider, but also for research projects to have an overview of the research outcomes and their accessibility. With several checks and balances in place, it provides a rich source of information as a trusted service provider.

With the need to scan and collect trusted research output on COVID-19 quickly, *OpenAIRE* created the COVID-19 gateway in collaboration with the European Commission and other key players ELIXIR, EMBL and RDA. This COVID-19 Open Research Gateway[3] provides a single entry point to COVID-19 resources, working closely with (European) disciplinary research infrastructures. The effort leverages upon *OpenAIRE*'s long-held experience in gathering and inferring diverse research subjects with specialised topics. It will be a fast collaborative exercise in identifying valuable resources from a vast range of scholarly information, filtering and aggregating COVID-19 related records and providing the necessary links to funders and institutions. As a result, anyone can discover and access contextualised research output specific to COVID-19, for free, from a single access point. The *OpenAIRE* COVID-19 Gateway collects many EU and global initiatives. One of the initiatives it is linked to is the *Zenodo* COVID-19 community[4].

Zenodo is a free repository for storing and sharing publications, data, software and other research artefacts. It is co-funded by the European Commission and hosted by CERN. It provides the wider scientific community with the option of storing its data in a non-commercial environment and making it freely available to society at large. *Zenodo* also provides the feature to create a community: a specific place to collect and accept or reject uploads submitted. Hence the researcher or research group can create a space dedicated for their research output, workshop or project. On top of that, by assigning DOIs and promoting open access, everything is citable and discoverable.

Already in March 2020, *OpenAIRE* and *Zenodo* created a specific community to collect all research results that could be relevant for the scientific community worldwide working on COVID-19 and SARS-CoV-2. It is a curated community, which means a team of

experts nominated by *OpenAIRE*, reviews records. Each time a *Zenodo* user wants to add a record to the community, an e-mail is sent to the curators who will decide whether the record is relevant to the COVID-19 community and whether to include the record or not. A data curator is also assigned to scout for further COVID-19 uploads outside the community and coordinate the COVID-19 curation efforts on *Zenodo* with those of other teams worldwide.

The *OpenAIRE* COVID-19 gateway reveals the power of openly accessible, interconnected research information. It shows how previous Open Science efforts contribute to effective ways to accelerate COVID-19 research. It is also a powerful motivator to invest in infrastructure, technical and human, to be able to offer such useful services that deliver open access research results for all and in all disciplines.

Emilie Hermans
Inge Van Nieuwerburgh
Ghent University, OpenAIRE
Sint-Hubertusstraat 8
9000 Gent
emilie.hermans@ugent.be
orcid.org/0000-0001-6594-7826
inge.vannieuwerburgh@ugent.be
orcid.org/0000-0002-7285-1598
https://lib.ugent.be
http://*openaire*.eu

October 2020

References

Walgrave, Stefaan; Dejaeghere, Yves. Surviving Information Overload: How Elite Politicians Select Information. *Governance* [online], 10 May 2016 (consulted on 23 July 2020), Vol. 30, no. 2, p. 229-244. <https://onlinelibrary.wiley.com/doi/abs/10.1111/gove.12209>

Piwowar, Heather; Priem, Jason; Larivière, Vincent; Alperin, Juan Pablo; Matthias, Lisa; Norlander, Bree; Farley, Ashley; West, Jevin; Haustein, Stefanie. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* [online], 13 February 2018 (consulted on 6 August 2020) 6:e4375. <https://doi.org/10.7717/peerj.4375>

Wilkinson, Mark; Dumontier, Michel; Aalbersberg, Ijsbrand; et al. The FAIR Guiding Principles for *scientific data* management and stewardship. Scientific Data [online], 15 March 2016 (consulted on 7 August 2020), Vol. 3, Article number: 160018. <https://doi.org/10.1038/sdata.2016.18>

Pontika, Nancy; Pearce, Knoth; Cancellieri, Matteo; Pearce, Samuel. Fostering Open Science to Research using a Taxonomy and an eLearning Portal. *i-KNOW '15: Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business* [online], October 2015 (consulted on 23 July 2020), no. 11, p. 1–8 <https://doi.org/110.1145/2809563.2809571>

Dahn, Bernice, Mussah, Vera; Nutt, Cameron. Yes, We Were Warned About Ebola. *The New York Times* [online], 7 April 2015 (consulted on 9 August 2020) <https://www.nytimes.com/2015/04/08/opinion/yes-we-were-warned-about-ebola.html?_r=0>

Baize, Sylvain; Pannetier, Delphine; Pharm, D.; Oestereich, Lisa; et al. Emergence of Zaire Ebola Virus Disease in Guinea. *The New England Journal of Medicine* [online], 9 October 2014 (consulted on 25 August 2020), Vol. 317, No. 15, p. 1418-1425 <https://doi.org/10.1056/NEJMoa1404505>

Abramowitz, Sharon; Giles-Vernick, Tamara; Webb, Jimm; Tappan, Jennifer; et al.
Data Sharing in Public Health Emergencies: Anthropological and Historical Perspectives on Data Sharing during the 2014-2016 Ebola Epidemic and the 2016 Yellow Fever Epidemic [online]. Wellcome Trust, 2018 (consulted on 14 July 2020). Report: Anthropology, p. 1-124. <https://www.glopid-r.org/wp-content/uploads/2019/07/data-sharing-in-public-health-emergencies-yellow-fever-and-ebola.pdf>

Besançon, Lonni; Peiffer-Smadja, Nathan; Segalas, Corentin; Jiang, Haiting; Masuzzo, Paola; Smout, Cooper; Deforet, Maxime; Leyrat, Clémence (2020). Open Science Saves Lives: Lessons from the COVID-19 Pandemic. *BioRxiv* [online], 18 August 2020 (consulted on 2 September), nr. 8, p. 1-19. <https://doi.org/10.1101/2020.08.13.249847>

Yozwiak, Nathan; Schaffner, Stephen F.; Sabeti, Pardis C. Data sharing: Make outbreak research open access. *Nature*, 26 February 2015, vol. 518, nr. 7540, p. 477-479. <https://doi.org/10.1038/518477a>

De Graeve, Robbe; Brems, Peter. Beschermen komkommers en roken écht tegen corona? 5 tips om je weg te vinden in het kluwen van 50.000 coronastudies. *VRT NWS* [online], 11 september 2020 (consulted on 2 October 2020). <https://www.vrt.be/vrtnws/nl/2020/09/09/beschermen-roken-en-komkommers-echt-tegen-corona-vijf-tips-om-g/>

Notes

1. The *FOSTER* portal is an e-learning platform that brings together the best training resources addressed to those who need to know more about Open Science, or need to develop strategies and skills for implementing Open Science practices in their daily workflows. <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition> (consulted on 2 October 2020)

2. *OpenAIRE* Research Graph [online] <https://graph.*openaire*.eu/> (consulted on 2 October 2020)

3. *OpenAIRE* for COVID-19 [online] <https://www.*openaire*.eu/*openaire*-activities-for-covid-19> (consulted on 2 October 2020)

4. Coronavirus Disease Research Community - COVID-19 [online] <https://*zenodo*.org/communities/covid-19> (consulted on 2 October 2020)