e-ISSN 1574-180X

An International Journal on Grey Literature

Volume 17, Number 3, Autumn 2021

‘TRACKING THE INS AND OUTS OF GREY LITERATURE’

www.textrelease.com Grey Literature Network Service www.greynet.org

The Grey Journal

An International Journal on Grey Literature

COLOPHON

About TGJ

The Grey Journal is a flagship journal for the international grey literature community. It crosses continents, disciplines, and sectors both public and private.
The Grey Journal not only deals with the topic of grey literature but is itself a document type classified as grey literature. It is akin to other grey serial publications, such as conference proceedings, reports, working papers, etc.

The Grey Journal is geared to Colleges and Schools of Library and Information Studies, as well as, information professionals, who produce, publish, process, manage, disseminate, and use grey literature e.g. researchers, editors, librarians, documentalists, archivists, journalists, intermediaries, etc.

CIP
The Grey Journal (TGJ) : An international journal on grey literature / Dominic Farace (Journal Editor); Jerry Frantzen (Technical Editor) ; GreyNet International, Grey Literature Network Service. - Amsterdam: TextRelease, Volume 17, Number 3, Autumn 2021 – TIB (DE), CVTISR (SK), EBSCO (USA), ISTI CNR (IT), KISTI (KR), NIS IAEA (UN), NTK (CZ), and the University of Florida (USA) are Corporate Authors and Associate Members of GreyNet International.
e-ISSN 1574-180X (PDF)

Subscription Rate:
€240 institutional

Contact Address:
Back Issues, Document Delivery, Advertising, and Subscriptions:

TextRelease
Javastraat 194-HS
1095 CP Amsterdam
Netherlands
T +31 (0) 20 331.2420
info@textrelease.com
http://www.textrelease.com/glpublications.html

About GreyNet

The Grey Literature Network Services was established in order to facilitate dialog, research, and communication between persons and organizations in the field of grey literature. GreyNet further seeks to identify and distribute information on and about grey literature in networked environments. Its main activities include the International Conference Series on Grey Literature, the creation and maintenance of web-based resources, a moderated Listserv, and The Grey Journal. GreyNet is also engaged in the development of distance learning courses for graduate and post-graduate students, as well as workshops and seminars for practitioners.

Full-Text License Agreement
In 2004, TextRelease entered into an electronic licensing relationship with EBSCO Publishing, the world's most prolific aggregator of full text journals, magazines and other sources. The full text of articles in The Grey Journal (TGJ) can be found in *Library, Information Science & Technology Abstracts* (LISTA) full-text database.

Contents

'Tracking the Ins and Outs of Grey Literature'

EDITOR'S NOTE

SPECIAL JOURNAL ISSUES ON GREY LITERATURE IN 2022

Since 2005, The Grey Journal is the flagship journal for the international grey literature community, which in itself comprises multiple and diverse communities of practice in this field of information. Each carries out research and publishes results in a host of document types such as preprints, conference papers, policy statements, etc. Most often they are open access compliant.

The Grey Journal now in its 17th volume has on occasion included special issues alongside its regularly published Spring, Summer, and Autumn issues. Starting with the 18th volume, The Grey Journal under the editorial management of GreyNet International actively welcomes special issues submitted by organizations and institutions involved in the field of grey literature.

The benefits of publishing their work in The Grey Journal are significant and warrant the recognition given to articles in established and licensed journals that are abstracted and indexed by professional information services. These are guarantees to the authors and researchers that their work will gain greater access, citation, and statistical feedback.

Further, when enriched metadata alongside digital persistent identifiers such as the DOI, ORCiD, ROR, and Funder ID are added to their work, they acquire a ranking status afforded publish journal articles. For more information, contact journal@greynet.org.

Dominic Farace,
Journal Editor

Fake Science: Legal Implications in the Creation and Use of Fake Scientific Data Published as Grey Literature and Disseminated through social media*

Tomas A. Lipinski, School of Information Studies, University of Wisconsin and
Kathrine A. Henderson, LAC Group, United States

Abstract

In this six-part paper, the authors first define fake science as a concept and identify at a high level the problems and consequences of fake science dissemination especially where fake science is published as grey literature and/or disseminated across social media platforms. In addition, they identify factors contributing to the creation of fake science from the "the replication crisis" in scientific research to the impact of technologies such as Artificial Intelligence. Part 2 moves into the United States Legal Landscape and considers US policy around fake science and related issues illustrated through a detailed discussion of applicable statutes and case law. Specifically, the authors discuss ISP immunity under 47 U.S.C. § 230 and the Constitutional implications of the United States v. Alverez, 132 S. Ct. 2537 (2012) and the decision and the applications of Central Hudson Gas & Electric Corp. v. Public Service Commission of New York, 447 U.S. 557 (1980). There will also be consideration of fake grey data as commercial speech or as a deceptive trade practice. Part 3 addresses the European Legal Landscape through a discussion of applicable laws and legal precedents in a similar manner to part 2. Part 4, Comparisons of the United States and European Legal Landscapes looks at the similarities and differences between the United States and Europe in addressing their shared concerns over the creation, use and dissemination of fake scientific information. Part 5, Prevention and Deterrence considers measures and actions which help to reduce the creation of fake science or that mitigate the problems it creates. These measures and actions are presented and incorporated into the fake science lifecycle presented in Part 1, Problem Definition. In Part 6, the authors make recommendations including technology driven solutions designed to ferret out fake science and in turn reducing the serious problems fake science presents. Recommendations include Facebook and other social media AI tools; manually flagging fake data; and the creation of truth seeking algorithms.

Part I Problem Definition

The term fake science is a spin-off of the phenomenon called "fake news" which while "fake" is not news at all and so too it goes for fake science. Some definitions are in order:

Research misconduct is fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results. Fabrication is making up data or results and subsequently recording or reporting these. Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented. (Office of Research Integrity, n.d.) The authors contend that research misconduct may result in "fake science".

Fake science is part of the larger universe of fake news which is shared over social media. Fake news has much in common with research misconduct. Fake news may be completely fabricated or may contain some truth, but it lacks verifiable facts or sources. Fake news may also include verifiable facts; however, the language used is inflammatory, relevant details have been left out and/or the information is from a single point of view designed to evoke a specific response from the audience. (Desai, Mooney, & Oehrli, 2020) Other definitions say fake news is is "concocted from unsourced, unverified, often made-up information and then masterfully manipulated to pass as real and credible journalism. "(Andorfer, 2018, p. 1413) Further, the aim of fake news is to "intentionally deceive those who read it." (Andorfer, 2018, p. 1413)

* First published in the GL2020 Conference Proceedings, February 2021. https://doi.org/10.26069/greynet-2021-000.468-gg

Fake science dissemination, just like fake news dissemination, is the practice of spreading misinformation or disinformation.

Fake science may appear in peer-reviewed journals, popular press articles, or it may appear in grey literature. Grey literature which herein is defined as "any recorded, referable and sustainable data or information resource of current or future value, made publicly available without a traditional peer-review process" (Slavic, 2018).

The peer review process in simplest terms is an evaluation of one's scientific, academic or professional work by others in the same field--hence the peer-reviewed journal. Publication in peer reviewed journals is critical to career advancement in the academy. Peer review is not without its critics; nonetheless, it is well-entrenched, and therefore everyone is striving to be published at the required pace to achieve career success. Publish or perish as they say. Complicating an already pressure-filled publication model is the replication or reproducibility crisis under discussion over roughly the last decade. This crisis, according to Stanford's Encyclopedia of Philsophy is, collectively:

- the absence of published replication studies;
- widespread inability to reproduced studies previously published;
- evidence of publication bias;
- high prevalence of "questionable research practices" which inflate false positives; and
- the lack of transparency and completeness in the reporting of methods, data and analysis in scientific publication. (Fidler & Wilcox, 2018)

Some commentators recognize how the replication crisis created a space for improvement while also recognizing that any replication experiment is a challenge in and of itself. Particularly important to notions of what is fake science, Collins holds that where a conflict of results arises, scientists tend to fraction into two groups, each holding opposing interpretations of the results where such groups are "determined" and the "controversy runs deep" and importantly, the dispute between the groups cannot be resolved via further experimentation. (as cited in Fidler & Wilcox, 2018) In such cases, Collins claims that non-epistemic factors, the career, social and cognitive interests of the scientists, their reputation and that of their institutions, will partly determine which interpretation becomes the lasting view. (as cited in Fidler & Wilcox, 2018)

There are several well-recognized examples of fake science' that have been disseminated across social media and continue to influence public behavior long-after these flaws if not outright falsification of data has been exposed. Three exemplars:

Andrew Wakefield published an article in 1998 that linked autism to the measles, mumps and rubella vaccine. His study was a catalyst that led thousands of parents to stop vaccinating their children: a collective decision that left a resurgence of the measles in its aftermath. This happened despite Britain's General Medical Council ruling that "*the children that Wakefield studied were carefully selected and some of Wakefield's research was funded by lawyers acting for parents who were involved in lawsuits against vaccine manufacturers*." and a belated retraction of Wakefield's article by the Lancet in 2010. (Eggertson, 2010, p. E199) The antivaccination trend continues to be a cause célèbre. Advocacy groups and parents including celebrities continue to believe in Wakefield's work and continue to share misinformation across social media. In addition, there is a prevailing conspiracy theory that vaccine manufacturers are hiding the truth about the connection between the MMR vaccine and autism. "*This consipiracy is fuelled by parents' understandable longing to know the cause of their child's autism*", says Margaret Spoelstra, executive director of Autism Ontario despite the fact that no large study has replicated Wakefield's findings. "*We know that autism has a genetic cause and that there are environmental factors that we don't understand yet*,"

Spoelstra says. "*There's enormous pressure in the field to come up with those answers*." (Eggertson, 2010, p. E200)

Judy Mikovitz studied human retroviruses in the late 1980s and 1990s at the National Cancer Insitute before moving on to private research at the Whittemore Peterson Institute. (Neil & Campbell, 2020, p. 546). In the early 2000s, she gained recognition when she found evidence of a particular virus, XMRV, in the blood of patients who were suffering from chronic fatigue syndrome. There was a great deal of excitement around this discovery and many scientists began work to replicate her findings. (Neil & Campbell, 2020, p 546). Endeavors to replicate Mikovitz' work led to the conclusion that the blood samples used in the questionable study had been contanimated by the XMRV virus rather than the infection being present in the majority of the chronic fatigue research subjects. (Neil & Campbell, 2020, p 547) In addition, there was strong evidence that the data had been deliberated falsified. The entire incident was mired in scanda as well as another conspiracy theory manifesting. To wit, the scientific establishment is suppressing Mikowitz' findings. (Neil & Campbell, 2020, p 547) In the end, Science, the journal that published Mikowitz' study, took "*the highly unusual step of retracting the whole article without Mikovits or the Whittemore Peterson Institute's agreement because it became clear that the study was fraudulent and scientifically invalid*." (Neil & Campbell, 2020, p 548)

More recently, Mikovitz is responsible for a documentary about the COVID-19 pandemic which has been blocked by social media sites like Facebook including YouTube where "Plandemic" first premeried. YouTube quickly withdrew the documentary citing violation of its misinformation policies; however, contemporaneous reporting through news media outlets indicate that at least a million viewers had watched the first part of the documentary before the video was taken down. Plandemic II or "Plandemic: Indoctrination" expands claims that the COVID-19 pandemic was a "planned event". Although certain local media across the United States were said to be airing the documentary, this did not happen. Plandemic: Indoctrination is currently available only on two social media sites. (Spencer, McDonald, & Fichera, 2020)

"The story of broken windows is a story of our fascination with easy fixes and seductive theories. Once an idea like that takes hold, it's nearly impossible to get the genie back in the bottle." (Shankar et al., 2016) In 1982, long before social media made its debut, Criminologists George L. Kelling and James Q. Wilson proposed a theory linking disorder and incivility with more serious crime in an article published in The Atlantic entitled, Broken Windows The police and neighborhood safety. In the article, the two proposed a theory linking disorder and incivility with more serious crime using broken windows as a metaphpor. (McKee, 2018). The thinking was, if neighborhoods showed signs of neglect and petty crime was abundant, this would signal that the neighborhood was uncared for and furthermore, that it would behoove local law enforcement to address these smaller problems. To do so would result in a contraction of more serious crime while also empowering neighborhoods (Shankar et al., 2016). Both law enforcement and the public were enthusiastic about this approach, and especially relevant here, the science moved out of the academy and into popular press. (Shankar et al., 2016) This theory would be embraced in 1993 by the newly elected Mayor of New York City, Rudy Guiliani, who had run on a tough on crime platform. Guilliani and his Chief of Police, William Bratton applied the theory to New York City's policing practice and by 2001 had become one of Guiliani's crowning achievements. (Shankar et al., 2016) Meantime, Kelling and Wilson continued their work publishing another study in 2001 which provided additional support that the crime theory was working. Subsequent re-analyis has found flaws in this work. Even the most promising study by political scientist Wesley Skogan, recommended that the broken window study results be interpreted with caution. (McKee, 2018) However, Columbia Law Professor Bernard Harcourt found that the link between

neighborhood disorder and crime, namely purse snatching, assault, rape, and burglary vanished when poverty, neighborhood stability, and race were statistically controlled. (McKee, 2018) In addition, Hartcourt is particularly concerned about the theory in that it fostered zero-tolerance policies which are biased against disadvantaged segments of society. (McKee, 2018). Over time, Kelling himself thought it might be a good idea to move away from the theory, *"It's to the point now where I wonder if we should back away from the metaphor of broken windows. We didn't know how powerful it was going to be. It simplified, it was easy to communicate, a lot of people got it as a result of the metaphor. It was attractive for a long time. But as you know, metaphors can wear out and become stale."* (Shankar et al., 2016) The theory remains popular despite evidence that says at best the theory may have a modest impact on crime. One reason may be that we have no clear explanation of what reduced violent crime in the 1990s. (Shankar et al., 2016)

These examples could easily be construed as sensationalism and to some extent this is a fair point. After all, this is the nature of fake science and broadly, fake news. The impact on the public especially on vulnerable groups is one of the reasons why this is a problem that needs exploration and study.

As a starting point, let's consider the depth and breadth of retractions as well as the time frames from original publication of scientific studies to published retractions described below. It is easy to imagine what could happen in the interim, the time in which the science is believed to be credible and is shared often unquestioningly tthrough social media.

As of August 31, 2020, The Retraction Watch Database includes over 20,000 retractions. These are primarily peer-reviewed journal articles and the numbers themselve should give everyone pause. Some specifics related to our examples.

- Retractions based on contamination of cell lines/tissues, 60 instances
- Retractions based on concerns/issues on data, the first 600 instances were displayed when limiting using this filter
- Retractions based on concerns about results, the first 600 instances were displayed when limiting using this filter

The ten most highly cited studies as of May 2019 found in the database have been cited hundreds of times and in some cases more than a thousand times. These studies continue to be cited after the retraction is made although the nature of these citations, reporting out on the study flaws for example, is unknown and would require future study. Based on authors' analysis, time frames between original publication date and retraction date vary widely at the low end 2 years and at the high end 17 years with an average of 8.6 years from publication date to retraction date.

Social media may prolong the lifecycle of marred, hasty or subsequent misinterpretation of scientific data. It may also expand the reach of fake science and related conspiracy theories. The extent to which fake science may create real harm is of interest to law makers globally.

Part 2 United States Legal Landscape

Primarily due to the influence of the Free Speech clause of the First Amendment of the U.S. Constitution the legal options for regulating fake science or disinformation is limited if perhaps non-existent. In addition, Congress provided broad immunity for online service provider giving little legal incentive to control questionable but otherwise lawful content on platforms. In enacting 47 U.S.C. § 230 (Protection for private blocking and screening of offensive material) as part of the Telecommunications Act of 1996, Congress stated that "[i]t is the policy of the United States to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation." Under section 230 an online service is immune from civil liability for content other post on its

platform or service. "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." 47 U.S.C. § 230(b)(2). The definition of a service provider or "Interactive computer service" covers "any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions." 47 U.S.C. § 230(f)(2). There is no immunity for content that infringes intellectual property. 47 U.S.C. § 230(e)(2) ("Nothing in this section shall be construed to limit or expand any law pertaining to intellectual property."). There are protections for service providers relating to copyright infringement, but those protections are found elsewhere. In order to secure the protection of the safe harbor a service provider must act "expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity" 17 U.S.C. § 512(c)(1)(C). However, section 230 contains no such requirement. As a result, there is no legal incentive to a service provider facing claims of other tort (injury) harms to undertake any remedial steps.

There are considerations to reform section 230 using several different strategies. One proposed Senate bill is the Online Freedom and Viewpoint Diversity Act. S. 4534, 116th Cong., 2d Sess., Online Freedom and Viewpoint Diversity Act (09/08/2020). This proposed bill would expand the range of blocking that can be done in subsection (c)(2) which currently reads: "No provider or user of an interactive computer service shall be held liable on account of (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected." 47 U.S.C. § 230(c)(2)(A). It would accomplish this expansion by striking ''considers to be'' and inserting ''has an objectively reasonable belief is'' and adding to the list of content appropriately blocked to include ''promoting self-harm, promoting terrorism, or un lawful.'' Section 2 of proposed S. 4553. Hopefully, this change would encourage service providers to police their platforms for such content knowing that if harmed resulted from missed content the immunity would still apply. Section 2(1)(B)(i) of the proposed Online Freedom and Viewpoint Diversity Act would also amend the definition information content provider in current subsection (f)(3) of Section 230 which reads as follows: "The term 'information content provider' means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service." 47 U.S.C. § 230(f)(3). The proposed Online Freedom and Viewpoint Diversity Act would designate the current language of subsection (f)(3) into a new paragraph (f)(3)(A), then add a new paragraph (B): "(B) RESPONSIBILITY.—For purposes of subparagraph (A), being responsible in whole or in part for the creation or development of information (i) includes any instance in which a person or entity editorializes or affirmatively and substantively modifies the content of another person or entity; and (ii) does not include a change to the format, layout, or basic appearance of the content of another person or entity.'' Section 2(2)(B) and (A) of the proposed Online Freedom and Viewpoint Diversity Act. If a service provider labeled content posted on its platform as suspect it would be immune from liability for making that designation. This too, might encourage service provider to more actively police the content it makes available.

Another Senate bill is the Platform Accountability Consumer Transparency (PACT) Act. The PACT Act would constitute an elaborate revision of current Section 230 by adopting a mechanism by which a user of an interactive computer service or a third party could provide "notice of illegal content or illegal activity on the inter active computer service that substantially complies with the requirements under [new] paragraph (3)(B)(ii)… 47 U.S.C. 230(c)…[if so] the provider shall remove the content or stop the activity within 24 hours of

receiving that notice, subject to reasonable exceptions based on concerns about the legitimacy of the notice." S. 4066, 116th Cong., 2d Sess., Platform Accountability Consumer Transparency (PACT) Act, Section 6(a) (06/24/2020). The revised Section 230 under the PACT Act would resemble the take-down provisions in 17 U.S.C. § 512. The PACT Act defines illegal activity as "activity conducted by an information content provider that has been determined by a Federal or State court to violate Federal criminal or 18 civil law" and illegal content as "information provided by an information content provider that has been determined by a Federal or State court to violate Federal criminal or civil law or State defamation law." Section 6(b) of the proposed PACT Act, adding new subsection (f)(5) and (6).

On October 20, 2020 H.R. 8636, 116th Congress, 2d Session, Protecting Americans from Dangerous Algorithms Act was introduced. The bill is sponsored Representatives Malinowski and Eshoo. The bill was introduced in the wake of the militia instigated violence and shootings including two homicides in Kenosha, Wisconsin during a Black Lives Matter protest on the night of August 25, 2020. The militia communicated using Facebook. The bill amends Section 230 by removing immunity for a platform if its algorithm is used to amplify or recommend content directly relevant to a claim involving interference with civil rights (42 U.S.C. 1985); neglect to prevent interference with civil rights (42 U.S.C. 1986); and in cases involving acts of international terrorism (18 U.S.C. 2333) when "the claim involves a case in which the interactive computer service used an algorithm, model, or other computational process to rank, order, promote, recommend, amplify, or similarly alter the delivery or display of information (including any text, image, audio, or video post, page, group, account, or affiliation) provided to a user of the service if the information is directly relevant to the claim." Section 2, 2020 H.R. 8636, 116th Congress, 2d Session, Protecting Americans from Dangerous Algorithms Act, adding subsection (c)(3) to Section 230, entitled "Algorithmic Amplification." On September 23, 2020, Facebook as named as a party in a lawsuit regarding the wrongful death of one of the victims. A lawsuit is pending against Facebook. *Gittings, et al. v. Mathewson, et al.*, Civil Case No. 2:20-cv-1483 (E.D. Wis). "Social media is in Congress's cross hairs for its ability to spread content promoting conspiracy theories, propaganda and misinformation about the election and the COVID-19 pandemic. Companies including Facebook Inc. and Twitter Inc. rely on Section 230 to protect themselves from lawsuits relating to inflammatory material posted on their platforms." Henry Kenyon, Eshoo bill holds social media responsible for harm caused by algorithms Congressional Quarterly Roll Call, Data Privacy Briefing, no pagination in Westlaw (October 21, 2020).

The U.S. Supreme recently denied a Writ of Certiorari in *Enigma Software Group USA, LLC v. Malwarebytes, Inc.*, 946 F.3d 1040 (9th Cir. 2019). This would have provided the U.S. Supreme Court to review and possible limit the scope of the immunity provided in Section 230. *Enigma Software Group USA, LLC v. Malwarebytes, Inc.*, 946 F.3d 1040 (9th Cir. 2019), cert. denied 2020 WL 6037214 (October 13, 2020). Justice Thomas provided a memorandum opinion in denying certiorari. "Paring back the sweeping immunity courts have read into § 230 would not necessarily render defendants liable for online misconduct. It simply would give plaintiffs a chance to raise their claims in the first place. Plaintiffs still must prove the merits of their cases, and some claims will undoubtedly fail. Moreover, States and the Federal Government are free to update their liability laws to make them more appropriate for an Internet-driven society… Without the benefit of briefing on the merits, we need not decide today the correct interpretation of § 230. But in an appropriate case, it behooves us to do so." *Malwarebytes, Inc. v. Enigma Software Group USA, LLC*, 2020 WL 6037214, p. 4 (October 13, 2020). The provision of the statute in question in the cases was subsection (c)(2): "No provider or user of an interactive computer service shall be held liable on account of (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user

considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected." 47 U.S.C. § 230(c)(2). The case involved two software companies. "Malwarebytes and Enigma have been direct competitors since 2008, the year of Malwarebytes's inception. In their first eight years as competitors, neither Enigma nor Malwarebytes flagged the other's software as threatening or unwanted. In late 2016, however, Malwarebytes revised its PUP-detection criteria to include any program that, according to Malwarebytes, users did not seem to like. After the revision, Malwarebytes's software immediately began flagging Enigma's most popular programs—RegHunter and SpyHunter—as PUPs. Thereafter, anytime a user with Malwarebytes's software tried to download those Enigma programs, the user was alerted of a security risk and, according to Enigma's complaint, the download was prohibited, *i.e.* Malwarebytes "quarantined" the programs." *Enigma Software Group USA, LLC v. Malwarebytes, Inc.*, 946 F.3d 1040, 1047-1048 (9th Cir. 2019), cert. denied 2020 WL 6037214 (October 13, 2020). The appellate court concluded that because the blocking was done for an ill-motive the immunity did not apply: "we hold that § 230 does not provide immunity for blocking a competitor's program for anticompetitive reasons, and because Enigma has specifically alleged that the blocking here was anticompetitive, Enigma's claims survive the motion to dismiss. We therefore reverse the dismissal of Enigma's state-law claims and we remand for further proceedings." *Enigma Software Group USA, LLC v. Malwarebytes, Inc.*, 946 F.3d 1040, 1052 (9th Cir. 2019), cert. denied 2020 WL 6037214 (October 13, 2020). In discussing the exception in Section 230 for intellectual property claims, the Ninth Circuit observed that "Enigma's Lanham Act claim derives from the statute's false advertising provision. Enigma alleges that Malwarebytes mischaracterized Enigma's most popular software programs in order to divert Enigma's customers to Malwarebytes. These allegations do not relate to or involve trademark rights or any other intellectual property rights. Thus, Enigma's false advertising claim is not a claim "pertaining to intellectual property law" within the meaning of § 230(e)(2). The district court correctly concluded that the intellectual property exception to immunity does not encompass Enigma's Lanham Act claim." *Enigma Software Group USA, LLC v. Malwarebytes, Inc.*, 946 F.3d 1040, 1053-1054 (9th Cir. 2019), cert. denied 2020 WL 6037214 (October 13, 2020).

Looking beyond Section 230, it is likely that a law designed to regulate or prohibit false information would meet with constitutional challenge. This has occurred. The Supreme Court invalidated the federal Stolen Valor Act of 2005 [18 U.S.C. § 704(b)], which criminalized falsely representing oneself as having been awarded military medals or decorations. As less restrictive means are available to achieve the Congress' goal of truth in military medal records the statute failed to pass the strict scrutiny test applied to content based restrictions on speech. "A Government-created database could list Congressional Medal of Honor recipients. Were a database accessible through the Internet, it would be easy to verify and expose false claims." *United States v. Alvarez*, 567 U.S. 709, 729 (2012). The Court reiterated its belief less speech in not better speech, i.e., the Marketplace of Ideas concept. "The lack of a causal link between the Government's stated interest and the Act… The Government has not shown, and cannot show, why counterspeech would not suffice to achieve its interest. The facts of this case indicate that the dynamics of free speech, of counterspeech, of refutation, can overcome the lie." Id. at 726. Justice Kennedy, writing for the plurality commented that the "respondent's statements anything but contemptible, his right to make those statements is protected by the Constitution's guarantee of freedom of speech and expression." Id. at 730.

It could be argued that some fake new, especially when the claimed purports to be based on science is so outrageous that legal harm is found. Such claim would be based on the tort theory of Intentional infliction of emotional distress (IIED): claims must be "so outrageous in

character, and so extreme in degree, as to go beyond all possible bounds of decency, and to be regarded as atrocious, and utterly intolerable in a civilized community." *See* Restatement (Second) of Torts § 46 cmt. d (Am. Law Inst. 1965). The "actual malice" standard applicable to defamation cases was equally applicable to IIED claims brought by public figure. *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46 (1988). However, even here the influence of the Frist Amendment is present. In *Snyder v. Phelps*, 562 U.S. 443 (2011) a case involving the peaceful picketing of members of the military who were claimed to be homosexual. "Given that Westboro's speech was at a public place on a matter of public concern, that speech is entitled to "special protection" under the First Amendment. Such speech cannot be restricted simply because it is upsetting or arouses contempt" Id. at 458. Considering some disinformation surrounds co-called conspiracy theories the comment of Chief Justice Roberts is telling: "Because we find that the First Amendment bars Snyder from recovery for intentional infliction of emotional distress or intrusion upon seclusion—the alleged unlawful activity Westboro conspired to accomplish—we must likewise hold that Snyder cannot recover for civil conspiracy based on those torts." Id. at 460. Finally, the argument could be made that in some instances the disinformation is so pervasive that it creates a captive audience so the ability of the government to regulate would be greater. "As a general matter, we have applied the captive audience doctrine only sparingly to protect unwilling listeners from protected speech" Id. at 459. This *Snyder* Court also rejected this argument. Id.

As a result, the legal regulation of disinformation is near futile in the U.S. Voluntary, self-regulation by service providers is the only option. Likewise, in the European Union (EU), this is path chosen as most viable as well, coupled with nonintervention government strategies.

Part 3 European Legal Landscape

The approach to addressing the problems raised by fake news and fake science or disinformation, i.e., the term of preference in the European Union, is markedly different than the United States. In the EU, it is more common to have the private sector adopting a model of co-regulation and partnership with civil authority to combat a societal problem. In the EU, the 2017 and 2018 flurry of activity was designed in anticipation of the then upcoming European Parliament elections in May of 2019.

In 2017 a Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda, adopted by Special Rapporteurs appointed by international organisations [sic] identified the major concerns with disinformation. As this content is designed to mislead it "interfere[s] with the public's right to know and the right of individuals to seek and receive, as well as to impart, information and ideas of all kinds, regardless of frontiers, protected under international legal guarantees of the rights to freedom of expression and to hold opinions." Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda, adopted by Special Rapporteurs on Freedom of Expression and Access to Information, 1 (March 3, 2017), paraphrasing Article 11, European Union Charter of Fundamental Rights of the European Union and Article 19, Universal Declaration of Human Rights (10 Dec. 1948), U.N.G.A. Res. 217 A (III) (1948).

An initial scoping of strategies was presented in the Final report of the High Level Expert Group [HLEG] on Fake News and Online Disinformation: A Multi-Dimensional Approach to Disinformation (March 12, 2018). There are five "pillars" of action. Enhance the transparency of online news involving an adequate and privacy-compliant sharing of data about the systems that enable their circulation online, promote media and information literacy to counter disinformation and help users navigate the digital media environment, develop tools for empowering users and journalists to tackle disinformation fostering a positive engagement with fast-evolving information technologies, safeguard the diversity and sustainability of the

European news media ecosystem, and promote continued research on the impact of disinformation in Europe to evaluate the measures taken by different actors and constantly adjust the necessary responses. Id. at 5-6 and 35. The first three rely on cooperation between the private and public sectors.

The HLEG focused its comments on disinformation, which it defined as "as false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit." Id. at p. 10 It adopted the term "disinformation because "fake news" alone is "inadequate to capture" the complexity of the problem and because the term fake news is used by some politicians to "dismiss" or discredit information with which they disagree and so it "undermine[s] independent news media. Id. In a telling comment made is the context of the EU but so true of the climate in the U.S. the HLEG observed: "some problems of disinformation are animated by citizens individually or collectively sharing false and misleading content and that highly polarized societies with low levels of trust provide a fertile ground for the production and circulation of ideologically motivated disinformation." Id. at 11, citing, Weeks, B. E. (2015) Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. Journal of Communication. 65 (4), 699-719. The best strategy includes multi-stakeholder collaborations with minimal regulation without "politically dictated privatization of the policing and censorship" of what is deemed unacceptable." Id. at 20. One strategy proposed is aimed at "'diluting' disinformation through increased transparency and enhanced visibility and findability of trusted news content." Id. at p. 29. This comments echoes Justice Roberts opinion in the *Alvarez* decision. The answer to "bad" speech is more "good" speech so that the marketplace of idea decides. "Research suggests that detailed counter-messages and alternative narratives are often more effective than corrections in countering disinformation." Id. at p. 29, n. 50. This again is a direct parallel to the marketplace of ideas concept behind the Free Speech clause in the U.S. There is a clear concern by the HLEG to avoid governmental control of digital media. Id. at p. 30. As such, "the best responses are likely to be those driven by multi-stakeholder collaborations. Regulatory responses may quickly become inadequate to tackle a multi-faceted problem such as disinformation, whose nature and characteristics are bound to change fast with the evolution of technologies and digital behaviour [sic] patterns." Id. at 31.

The next month European Commission issues its essential framework regarding disinformation in the EU. European Commission Communication 'Tackling online disinformation': a European approach, COM(2018) 236 Final (April 26, 2018). Disinformation is "verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm. Public harm comprises threats to democratic political and policy-making processes as well as public goods such as the protection of EU citizens' health, the environment or security." Id.at pp. 3-4. It does not include "reporting errors, satire and parody, or clearly identified partisan news and commentary." Id.

The Communication identifies three causes for the rise in disinformation. "Economic insecurity, rising extremism and cultural shifts" offering a "breeding ground for disinformation." Id. at p. 4. The rise of platforms underscores a "media sector undergoing profound transformation." Id. Finally, "social networking technologies are manipulated to spread disinformation through a series of sequential steps: (i) creation; (ii) amplification through social and other online media; and (iii) dissemination by users." Id. at p. 5. Amplification occurs through algorithms designed to maximize the platform's busines model where advertising model is "click-based, which rewards sensational and viral content" along with bots that "artificially amplify the spread of disinformation." Id.

The Commission Communication highlight several harms from disinformation: the erosion of "trust in institutions and in digital and traditional media," a negative impact on "democracies

by hampering the ability of citizens to take informed decisions…[and] supports radical and extremist ideas and activities. It impairs freedom of expression." Id. at 1. The obligation of the state is to "refrain from interference and censorship and to ensure a favourable [sic] environment for inclusive and pluralistic public debate." Id. Blame was targeted at online platforms logically as most disinformation originates, resides and is amplified in that environment. Further, regulation at this information funnel point is more efficient than addressing the individual perpetrators of the disinformation. "These platforms ["particularly social media, video-sharing services and search engines"] have so far failed to act proportionately, falling short of the challenge posed by disinformation and the manipulative use of platforms' infrastructures." Id. at 2. Some platforms have taken limited initiatives to redress the spread of online disinformation, but only in a small number of countries and leaving out many users. The Commission believes there should be several strategies: "improve transparency regarding the origin of the origin and the way it is produced, sponsored, disseminated and targeted…to reveal possible attempts to manipulate opinion…promote diversity of information, foster credibility of information by providing an indication of its trustworthiness [] with the help of trusted flaggers…improving traceability, fashion inclusive solutions…awareness-raising, more media literacy, broad stakeholder involvement and the cooperation of public authorities, online platforms, advertisers, trusted flaggers, journalists and media groups. Id. at p. 6.

The Commission called for development of an EU-wide Code of Practice on Disinformation to which platforms and advertisers should commit. The Code would have numerous objectives: Improving scrutiny of advertisement placements & restrict targeting options for political advertising, ensuring transparency about sponsored content, including political and issue-based advertising, closing down fake accounts, indicators of trustworthiness, improving findability of trustworthy content, clear marking systems and rules for bots, empower uses by facilitating content discovery & tools for reporting disinformation, ensuring safeguards-by-design against disinformation (e.g. algorithm prioritization), providing trusted fact-checking organisations and academia with access to platform data. European Commission Communication, 'Tackling online disinformation': a European approach, p. 7-8 (April 26, 2018). See also, EU Code of Practice on Disinformation, pp. 3-4 (September 26, 2018) (discussing same).

In May the Commission reported on the implementation of the Communication. European Commission Report, Implementation of the Communication "Tackling online disinformation: a European Approach" COM(2018) 794 final (May 5, 2018). The Report from the Commission reiterated its initial aims to develop "a self-regulatory code of practice on disinformation for online platforms and the advertising industry in order to increase transparency and better protect users; the creation of an independent European network of fact-checkers to establish common working methods, exchange best practices and achieve the broadest possible coverage across the EU; the promotion of voluntary online identification systems to improve the traceability and identification of suppliers of information; and the use of the EU research and innovation programme [sic] (Horizon 2020) to mobilise [sic] new technologies, such as artificial intelligence, block chain and cognitive algorithms." Id. at p.1. Commenting on the planned Code of Practice on Disinformation, the 15 "commitments" of the Code are organized into five fields: scrutiny of ad placements, political advertising and issue-based advertising, integrity of services, empowering consumers and empowering the research community." Id. The first two fields reflect the concern with the upcoming 2019 EU elections, The report concluded that "the actions outlined in the Communication have been accomplished or launched during 2018. Online platforms and the advertising industry have agreed on a Code of Practice to increase online transparency and protect consumers, with a particular view to the European elections in 2019. A network of fact checkers is being created that will strengthen

capabilities to detect and debunk false narratives... Stakeholders should [] benefit from a number of research and innovation tools to identify and tackle disinformation… Awareness has increased across… sustained efforts, at EU and national level, to raise the level of media literacy and empower users, especially the younger generations, and improve critical thinking." Id. at 12.

The actual Code of Practice on Disinformation reiterated its purposes in the form of a series of "Commitments" in five areas: scrutiny of ad placements (transparency), political advertising and issue-based advertising (transparency), integrity of services (misuse of bots), empowering consumers (indicators of trustworthiness of content sources, media ownership and verified identity), and empowering the research community (platform data). European Commission, EU Code of Practice on Disinformation (September 26, 2018).

Motivated by upcoming "2019 European Parliament elections and more than 50 presidential, national or local/regional elections being held in Member States by 2020, it is urgent to step up efforts to secure free and fair democratic processes" the Commission issued its action plan in December of 2018. European Commission Joint Communication, Action Plan against Disinformation at p. 2 (JOIN(2018) 36 final (May 12, 2018), updated December 5, 2018. The Plan reiterated the harms of disinformation: "Disinformation undermines the trust of citizens in democracy and democratic institutions. Disinformation also contributes to the polarisation [sic] of public views and interferes in the democratic decision-making processes." Id. at 11. The Plan stressed the key role "played by civil society and the private sector (notably social media platforms) in tackling the problem of disinformation." Id. The Plan identified four pillars: improving the capabilities of Union institutions to detect, analyse [sic] and expose disinformation, strengthening coordinated and joint responses to disinformation, mobilising [sic] private sector to tackle disinformation and raising awareness and improving societal resilience. Id. at pp. 5-11. One of the coordinated and joint responses is the establishment of a Rapid Alert System "to provide alerts on disinformation campaigns in real-time." Id. at p.7. This is one tool that can "foster an open, democratic debate free from manipulation, including in the context of the forthcoming European elections." Id. Cooperation of the private sector is crucial to success: "Online platforms, advertisers and the advertising industry have a crucial role to play in tackling the disinformation problem, as its scale is directly related to the platforms' ability to amplify, target and spread disinformation messages of malicious actors." Id. at p. 8. The Action Plan observed the past failures of the private sector to "act appropriately to tackle the problem." Id. A key concept in all the Commission communication, plans, etc. is to increase what it phrases as "societal resilience." This reflects a view that most effective strategies meet the problem of disinformation from the bottom up. Id. at p. 9. This is accomplished through numerous strategies to increase media, information and data literacy: "the objective is for the Union and its neighbourhood [sic] to become more resilient against disinformation. This requires continuous and sustained efforts to support education and media literacy, journalism, fact-checkers, researchers, and the civil society as a whole." Id. at p. 12.

In June of 2019, the Commission issued a Report in the aftermath of Parliamentary elections. European Commission Joint Communication, Report on the Implementation of the Action Plan against Disinformation JOIN(2019) 12 Final (June 14, 2019). "The aim of disinformation is to distract and divide, to plant seeds of doubt by distorting and falsifying facts, thus confusing people and weakening their faith in institutions and established political processes." Id. at 1, n. 1. To this end the Report observed that "available evidence has not allowed [sic] to identify a distinct cross-border disinformation campaign from external sources specifically targeting the European elections. However, the evidence collected revealed a continued and sustained disinformation activity by Russian sources aiming to suppress turnout and influence voter

preferences... a consistent trend of malicious actors using disinformation to promote extreme views and polarise local debates, including through unfounded attacks on the EU." Id. at p. 3 (footnote emitted). As online platform data was not made available in sufficient and significant quantities "a conclusive assessment of the scope and impact of disinformation campaigns will take time and require a concerted effort by civil society, academia, public actors and online platforms." Id. at p.3.

The Report indicated that online platforms have taken measures to increase the integrity of services each offers. For example, from January to May leading up to the elections "Google reported to have globally removed more than 3.39 million Youtube channels and 8,600 channels for violations against its spam and impersonation policies. Facebook disabled 2.19 billion fake accounts in the first quarter of 2019 and acted specifically against 1,574 non-EU-based and 168 EU-based pages, groups and accounts engaged in inauthentic behaviour targeting EU Member States. Twitter challenged almost 77 million spam-like or fake accounts globally." Id. at p. 4, n. 11. In addition, the European Media Literacy Week included over 320 events in the EU during the week of March 18, 2019.

Source: European Commission Joint Communication, Report on the Implementation of the Action Plan against Disinformation JOIN(2019) 12 Final at p. 2 (June 14, 2019).

The report concluded that "preliminary analysis shows that it contributed to expose disinformation attempts and to preserve the integrity of the elections, while protecting freedom of expression. The highest turnout in the past twenty years (50.97 %) reflects the interest of the citizens for the Union and its importance for their lives." Id. at p. 9 (footnote eomitted). While disinforation is an evovling threat requiring "continuous research to update our policy toolboox… the objective remains the same: dividing our society and undermining the trust of citizens in democratic processes and institutions." Id.

A year later the detailed European Commission, Study for the "Assessment of the Implementation of the Code of Practice on Disinformation" Final Report (May 2020) was released. Commenting on the effectiveness of the "five pillars" the Final Reports made the following observations. Regarding the scrutiny of ad placements: "The Code has not effectively incentivised [sic] the platforms to provide data that is detailed enough to be of use in assessing the effectiveness of their existing policies with regards to scrutiny of ad placements." Id. at p. 39. There should also be "clearer definitions" and "[m]inimum data reporting requirements"

for the platforms. Id. at pp. 39 and 40, respectively. As the impetus for the Commission's work in the area was the 2019 European elections the Final Report concluded that stakeholders felt "that the Code, as it stands, places a disproportionate emphasis on political or issue-based advertising and should be updated to recognise [sic] that disinformation is also highly prevalent in terms of "organic" content by individual users, not just via advertising." Id. at p. 44 (regarding the second pillar: political and issue-based advertising). The stakeholders consulted also believed that there is still a "lack of data regarding this pillar that would allow independent verification of the information provided by the platforms." Id. at 46. Though platforms have ools and policies in place, "the focus is predominantly on preventing the creation of fake accounts/users and on detecting and deleting them. The status of such accounts/users (active, dormant, inactive etc.) and the reach of their activities is less known and reported on." Id. at p. 49.

Regarding the empowerment of consumers, the signatory platforms have developed a variety of tools: "Facebook has a context button which appears alongside links shared on its News Feeds…
Google has features such as Breaking News and Top News to ensure the prominence of authoritative content…Twitter has the 'verified accounts' function which means that celebrities, journalists, news organisations [sic] and politicians have verified accounts on Twitter (signalled [sic] by a blue badge with a white 'v' next to their name)… Mozilla included… a rollout of enhanced security features in the default setting of Firefox which highlights the quality of websites and provides other information about the website relevant to empower consumers…
Microsoft has a partnership with NewsGuard, which reviews online news sites across a series of nine journalistic integrity criteria." Id. at pp. 51 and 52 (footnotes omitted). The Final Report noted that these efforts "should be further developed and more widely implemented. This can then act as a minimum standard for all platforms to live up to further down the line… However, this does not automatically lead to consumer empowerment as the possibilities for this are not always know, and sometimes not even desired, by the (majority of the) consumers." Id. at 54.

The Final Report recounted a number of initiatives and collaboration the platforms are undertaking to support research on disinformation: "Twitter for example disclosed a significant archive of state-backed information operations on Twitter in October 2018… Mozilla Foundation which launched joint campaigns on transparency involving 71 researchers and 37 civil society organisations.102 Facebook reported that in April 2018 it launched a partnership with Social Science One (SS1), a group of 83 academic researchers, to share data with the academic research community while maintaining stringent privacy protections…. Microsoft also implemented partnering programmes [sic] with researchers (TAP), research institutions (Princeton University, Oxford Internet Institute) and with industry, including the participation of Bing News in the Trust Project… In collaboration with the International Fact Checking Network, Google News Lab launched FactCheck EU in March 2019 to provide fact checks from 19 organizations from 10 countries in 11 languages… Google also introduced new tools for researchers and the fact checking community: a ' Fact Check Explorer', which allows for exploration of Fact Checking journalism, and the 'Fact Check Markup Tool', which allows fact checkers to easily mark their own articles as fact-checks'" Id. at 55-56 (footnotes omitted). However, the Final Report noted that of the five pillars this is least developed. Id. at 56. Both the platforms and the research community criticized the other.
"Researchers noted that there is limited engagement with the research community and that the tools set up by platforms are still too weak, not transparent enough and not really user

friendly…According to platforms, the requests from researchers may sometimes be unclear or unrealistic, which makes the collaboration more difficult." pp. at 59-60.

The Final Report also noted at the member state level several countries have launched awareness and literacy campaigns, often in conjunction with trusted source or fact-checker portal. Fewer countries have actual legislative measures in place do the controversy that restrictive measures could "jeopardise [sic] democracy and censor the press." Id. at p. 78. Both France and Lithuania have measures that allow television channels to block disinformation (France: block television channels where "false information likely to affect the sincerity of the ballot" and Lithuania: television and radio stations blocking of "disinformation and information which is slanderous and offensive to a person or degrades human dignity and honour [sic]." Id. at pp. 78 and 79 (footnotes omitted). Initiatives in other member states including Hungary, Ireland and Italy have been discussed, or proposed or pending or recently enacted. Another approach taken in the Czech Republic is to target cyber-security and terrorist threats. Id. at pp. 80-81. Such measures would be broad enough to include disinformation if that was the form or context of the threat.

The most recent law is the European Commission Communication: the EU's fight against COVID-19 Disinformation released on June 10, 2020. As might be figured given the times in which we live the topic was scientific and medical disinformation related to COVID-19. The Communication addressed the issue of global "infodemic" and "flood of information about the virus, often false or inaccurate and spread quickly over social media." European Commission Joint Communication, Tackling COVID-19 Disinformation - Getting the Facts Right, JOIN/2020/8 Final, p. 1 (June 10, 2020). The Communication pointed to various forms of disinformation including "dangerous hoaxes and misleading healthcare information" conspiracy theories, hate speech regarding ethnic or religious group responsible for spread of the virus, consumer fraud (e.g., selling "miracle" cures), cybercrime (e.g., hacking/phishing using COVID-19 links), foreign actors (Russia and China are named in specific) influencing operation and disinformation campaign. Id. at p. 3. As with previous EU Commission communications there is an emphasis on platform cooperation (pp. 8-9) especially in supporting the work of fact-checkers and researchers (pp. 9-10) and the role of a free and independent media in reporting and outing sources of disinformation related to COVID-19 (pp. 10-13).

Overall, the EU is ahead of the US in identifying and defining disinformation, in raising awareness regarding its incidence and effects and in constructing collaborative multi-pronged plans and approaches for countering disinformation.

Part 4 Comparisons of the United States and European Legal Landscapes

Both the US and EU value freedom of expression, thus the EU has not a adopted a regulatory model but focused on cooperation between the private and public sector. In the US, due the First Amendment concerns the approach should be the same. Thus both the US and EU have the path of self-regulation and private/public collaboration as a framework to respond to the problem of disinformation. However, the policy attitudes in the US regarding the regulation of the private sector regarding information in general leave the US unprepared to mobilize the widespread and inclusive alliances necessary with private sector and civil society stakeholders. This is not to claim it is not occurring in the US especially in light of 2020 elections and its aftermath and the continuing pandemic, but it is not coordinated nor is there a unified federal plan or response. Rather platform and provider responses remain reactive, inconsistent and intermittent.

Part 5 Prevention & Deterrence

The practice of retracting fabricated data and results and leveraging legal remedies to help reduce the spread of fake science through social media comes late in the game. There are opportunities at earlier points along the continuum including prevention and deterrence of publishing fabricated or false research and results.

The United States Federal Government provides tremendous amounts of research funding and has a vested interested in preventing research misconduct where ever it has granted out monies. Within the Department of Health and Human Services, it has established an Office of Research Integrity [ORI]. The ORI is focused on ensuring the integrity of the research itself and disciplines those who do not adhere to standards.

Specifically, the ORI oversees and directs Public Health Service [PHS] research integrity on behalf of the Secretary of Health and Human Services. Per the ORI website, "*PHS provides nearly $38 billion for health research and development, primarily in the biomedical and behavioral sciences, through its extramural and intramural programs. Extramural funding supports research institutions outside of tthe federal government including medical shools, universities, hospitals and other reserch organizations with intramural funding supporting federal government agencies.* The ORI "*is responsible for ensuring the integrity of this research."* Operationally this translates to detection, investigation, and enforcement. The ORI also provides technical assistance to institutions and offers programming designed to promulgate responsible conduct. The agency is also empowered to discipline scientists who have committed research misconduct. (Office of Research Integrity, n.d.)

The ORI publishes research misconduct case summaries which include any imposed restrictions or sanctions against respondents. Once restrictions are lifted, the case summary is removed from the list. However, these actions are also recorded permanently in the Federal Registry. Six case summaries have been posted in 2020 as of September of that year. The publication date reflects the year in which investigations were completed. Respondents voluntarily enter into disciplinary agreements with the ORI; however, the agreements are not necessarily an admission of wrong doing. Authors' observations on 2020 investigations.

- Research misconduct is not concentrated in any particular discipline or setting.
- Fabricated and/or falsified data and results have been published in both grey literature and scholarly journals.
    - Typically most, if not all, of the Respondent's journal articles are rectracted; in some instances a correction is found to be sufficient
- Other misconduct such as plagiarism may have also occurred.
- A number of consequences negatively effect both the respondent and the institution.
    - A 2-3 panel committee must supervise the Respondent's research for a period of time. For the 2020 cases, the least amount of supervised time was 2 years and at the most 4 years except in the case of voluntary exclusion.
    - The committee must provide a report to the ORI certifying the research conducted by the Respondent meets standards.
    - The Respondent may be unable to fulfill hiis or her research obligations to the institution because he or she may not supervise, advise, and/or otherwise participate in grant funded research during the supervised period.
    - Some respondent's voluntarily exclude themselves from government or sub government contracts for 10 years.
    - In one cases,the Respondents's doctoral degree was revoked by the granting institution.

These consequences are career damaging if not career ending for the Respondent, but also put the institution in a difficult situation because many programs are heavily dependent on grant monies.

Part 6 Recommendations

Finding a path to reducing the spread of fake science is elusive and solutions have limitations. Culturally, socially, and emotionally, we humans like to have answers that resonate with our values, our hopes that calm our fears. People turn to conspiracies when there is an insufficient explanation. All of which is not to say that we should not intervene where possible to curtail the distribution of disinformation/misinformation at various points--teaching even the youngest of students to think critically, improving science literacy in government and amongst the public, retracting problematic studies and as a final recourse, imposing legal remedies.

Moving the needle will require a multifaceted strategy such as the five action-oriented pillars recommended in Fake News and Online Disinformation: A Multi-Dimensional Approach to Disinformation introduced in Part 3. The approach calls for among other things:

- Transparency around how online news is circulated
- Promotion of media and information literacy as a countermeasure and to help users navigate digital media
- Development of tools that would empower both users and journalist to tackle misinformation
- Fostering positive engagement in the rapid evolution of information technologies
- Finding ways to safeguard and sustain diversity in journalism and continued research

Multidimensional strategies require effort on the part of the social media industry, the journalism industry, education at all levels, and for users to take responsibilities for what they post and share.

Consider Andorfer's suggestion on three potential approaches including legal remedies already discussed herein. The other two are human judgement and purely technological solutions.

Human judgement places at least some responsibility on social media users and platform providers to identify and report posts that are dubious. In addition, encouraging users to moderate their own behaviors around sharing and reposting. (Andorfer, 2018, pp. 1413) There are several third party vendors that avail themselves to social media users who are willing to factcheck including the well known website, Snopes.com. At the platform level, Facebook for example has developed a flagging system wherein users can report problematic posts, which the company turns over to third parties to verify the accuracy of the information. (Andorfer, 2018, pp. 1414)

Facebook's human judgement solutions and others like it are imperfect. Among other things, it is increasingly difficult for even the savviest users, to identify fake news, especially deep fakes. Other issues include human error and bias.

Pure technological solutions utilizing artificial intelligence are also possible as these are increasingly sophisticated and capable of analyzing speech. Here too though there is the issue of bias and we may see the attitudes of corporations reflected, although not intentionally which could lead to censoring by social media companies like Facebook, Twitter and others.

Fake science has its own unique contributing factors stemming from the replication crisis, the limitations of the peer review process, the time it takes to discover falsification and fabrication, and long delays in retractions providing opportunity for misapplication and conspiracies to arise. These require multidimensional solutions to reduce systemic issues and to better communicate results to the general public.

References

Andorfer, A. (2018), Spreading Like Wildfire: Solutions for Abating the Fake News Problem on Social Media via Technology Controls and Government Regulation -- In Hastings Law Journal, 69(5), June 3, pp. 1410-1431. ISSN 0017-8322.

Desai, S., H. Mooney, & Oehrli, J. (2020), "Fake News," Lies and Propaganda: How to Sort Fact from Fiction - In: University of Michigan Library Research Guides.

Eggerston, L. (2010), Lancet retracts 12-year-old article linking autism to MMR vaccines -- In Canadian Medical Association Journal, 182(4), March 9, pp. E199-E200. 0820-3946 (print) 1488-2329 (web).

Fidler, F, & Wilcox, J. (2018, Winter), Reproducibility of Scientific Results - In: *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.).

Kenyon, H. (October 21, 2020), Eshoo bill holds social media responsible for harm caused by algorithms, *Congressional Quarterly Roll Call*, *Data Privacy Briefing*, no pagination in Westlaw.

McKee, A.(2018). Broken Windows Theory -- Encylopaedia Britannica, December 14.

Neil, S. & Campbell, E. (2020), Fake Science: XMRV, COVID-19, and the Toxic Legacy of Dr. Judy Mikovits - In: AIDS Research and Human Retroviruses, 36(7), July 7, pp. 545-549. DOI: 10.1089/aid.2020.0095

Office of Reserch Integrity, (n.d.), Official Website. United States Department of Health and Human Services. https://ori.hhs.gov/

Restatement (Second) of Torts (Am. Law Inst. 1965).

Retraction Watch (n.d.) Top 10 most highly cited retracted papers - In: Retraction Watch Leader Board. Accessed August 31, 2020.

Savić, D. (2018), When is grey too grey. pp. 19-24. orcid.org/0000-0003-1123-9693

Shankar, V., Benderev, C., Boyle, T. Klahr, R., Penman, M. Schmidt, J. (2016). How A Theory Of Crime And Policing Was Born, And Went Terribly Wrong. November 1, Hidden Brain, National Public Radio.

Spencer, S., McDonald,J. & Fichera, A. (2020). New 'Plandemic' Video Peddles Misinformation, Conspiracies. August 8, Factcheck.org.

Weeks, B. E. (2015) Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. Journal of Communication. 65 (4), 699-719.

Statutes

17 U.S.C. § 512 (Limitations on liability relating to material online).

47 U.S.C. § 230 (Protection for private blocking and screening of offensive material).

Bills

H.R. 8636, 116th Cong., 2d Sess., Protecting Americans from Dangerous Algorithms Act (10/20/2020).

S. 4066, 116th Cong., 2d Sess., Platform Accountability Consumer Transparency (PACT) Act (06/24/2020).

S. 4534, 116th Cong., 2d Sess., Online Freedom and Viewpoint Diversity Act (09/08/2020).

Cases

Enigma Software Group USA, LLC v. Malwarebytes, Inc., 946 F.3d 1040 (9th Cir. 2019), cert. denied 2020 WL 6037214 (October 13, 2020)

Gittings, et al. v. Mathewson, et al., Civil Case No. 2:20-cv-1483 (E.D. Wis) (09/20/2020).

Hustler Magazine, Inc. v. Falwell, 485 U.S. 46 (1988).

Snyder v. Phelps, 562 U.S. 443 (2011).

United States v. Alvarez, 567 U.S. 709, 729 (2012).

Documents

European Commission, Communication 'Tackling online disinformation': a European approach, COM(2018) 236 Final (April 26, 2018).

European Commission, Report, Implementation of the Communication "Tackling online disinformation: a European Approach" COM(2018) 794 final (May 5, 2018).

European Commission, Joint Communication, Action Plan against Disinformation at p. 2 (JOIN(2018) 36 final (May 12, 2018), updated December 5, 2018.

European Commission, EU Code of Practice on Disinformation (September 26, 2018).

European Commission, Joint Communication, Report on the Implementation of the Action Plan against Disinformation JOIN(2019) 12 Final (June 14, 2019).

European Commission, Study for the "Assessment of the Implementation of the Code of Practice on Disinformation" Final Report (May 2020).

European Commission, Joint Communication, Tackling COVID-19 Disinformation - Getting the Facts Right, JOIN/2020/8 Final, p. 1 (June 10, 2020).

Final report of the High Level Expert Group [HLEG] on Fake News and Online Disinformation: A Multi-Dimensional Approach to Disinformation (March 12, 2018).

Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda, adopted by Special Rapporteurs on Freedom of Expression and Access to Information, 1 (March 3, 2017).

Korea Institute of Science and Technology Information (KISTI)

English version - http://en.kisti.re.kr/

*** Vision**

World-class information research institute creating values for customers

*** Main functions**

*** Management and service of Korean R&D reports**

KISTI exclusively manages, preserves, and serves Korean R&D reports for citizens and government officials. It provides Korean R&D reports and their information with National science & Technology Information Service (NTIS) and National Discovery for Science Leaders (NDSL).

***Contact information**

KISTI email address: hcpark@kisti.re.kr

Headquarters: Tel : +82-42-869-1004, 1234 Fax: +82-42-869-0969

The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape*

Nicholas Fraser, Leibniz Information Centre for Economics, Germany
Liam Brierley, Department of Health Data Science; University of Liverpool, United Kingdom
Gautam Dey, MRC Lab for Molecular Cell Biology, United Kingdom
Jessica K. Polka, ASAPbio – Accelerating Science and Publication in Biology, United States
Máté Pálfy, The Company of Biologists, United Kingdom
Federico Nanni, The Alan Turing Institute, United Kingdom
Jonathon Alexis Coates, Hughes Hall College; University of Cambridge, United Kingdom

Abstract:
The world continues to face a life-threatening viral pandemic. The virus underlying the Coronavirus Disease 2019 (COVID-19), Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), has caused over 98 million confirmed cases and 2.2 million deaths since January 2020. Although the most recent respiratory viral pandemic swept the globe only a decade ago, the way science operates and responds to current events has experienced a cultural shift in the interim. The scientific community has responded rapidly to the COVID-19 pandemic, releasing over 125,000 COVID-19–related scientific articles within 10 months of the first confirmed case, of which more than 30,000 were hosted by preprint servers. We focused our analysis on bioRxiv and medRxiv, 2 growing preprint servers for biomedical research, investigating the attributes of COVID-19 preprints, their access and usage rates, as well as characteristics of their propagation on online platforms. Our data provide evidence for increased scientific and public engagement with preprints related to COVID-19 (COVID-19 preprints are accessed more, cited more, and shared more on various online platforms than non-COVID-19 preprints), as well as changes in the use of preprints by journalists and policymakers. We also find evidence for changes in preprinting and publishing behaviour: COVID-19 preprints are shorter and reviewed faster. Our results highlight the unprecedented role of preprints and preprint servers in the dissemination of COVID-19 science and the impact of the pandemic on the scientific communication landscape.

Introduction

Since January 2020, the world has been gripped by the Coronavirus Disease 2019 (COVID-19) outbreak, which has escalated to pandemic status, and caused over 98 million cases and 2.1 million deaths (43 million cases and 1.1 million deaths within 10 months of the first reported case)[1–3]. The causative pathogen was rapidly identified as a novel virus within the family Coronaviridae and was named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV 2)[4]. Although multiple coronaviruses are ubiquitous among humans and cause only mild disease, epidemics of newly emerging coronaviruses were previously observed in SARS in 2002[5] and Middle East Respiratory Syndrome (MERS) in 2012[6]. The unprecedented extent and rate of spread of COVID-19 has created a critical global health emergency, and academic communities have raced to respond through research developments.

New scholarly research has traditionally been communicated via published journal articles or conference presentations. The traditional journal publishing process involves the submission of manuscripts by authors to an individual journal, which then organises peer review, the process in which other scientists ("peers") are invited to scrutinise the manuscript and determine its suitability for publication. Authors often conduct additional experiments or analyses to address the reviewers' concerns in 1 or more revisions. Even after this lengthy process is concluded, almost half of submissions are rejected and require resubmission to a different journal[7]. The entire publishing timeline from submission to acceptance is estimated to take approximately 6

* First published in PLOS Biology, https://doi.org/10.1371/journal.pbio.3000959 April 2, 2021.

months in the life sciences[8,9]; the median time between the date a preprint is posted and the date on which the first DOI of a journal article is registered is 166 days in the life sciences[8].

Preprints are publicly accessible scholarly manuscripts that have not yet been certified by peer review and have been used in some disciplines, such as physics, for communicating scientific results for over 30 years[10]. In 2013, 2 new preprint initiatives for the biological sciences launched: PeerJ Preprints, from the publisher PeerJ, and bioRxiv, from Cold Spring Harbor Laboratory (CSHL). The latter established partnerships with journals that enabled simultaneous preprint posting at the time of submission[11]. More recently, CSHL, in collaboration with Yale and BMJ, launched medRxiv, a preprint server for the medical sciences[12]. Preprint platforms serving the life sciences have subsequently flourished, and preprints submissions continue to grow year on year; two-thirds of these preprints are eventually published in peer-reviewed journals[8].

While funders and institutions explicitly encouraged prepublication data sharing in the context of the recent Zika and Ebola virus disease outbreaks[13], usage of preprints remained modest through these epidemic[14]. The COVID-19 crisis represents the first time that preprints have been widely used outside of specific communities to communicate during an epidemic.

We assessed the role of preprints in the communication of COVID-19 research in the first 10 months of the pandemic, between January 1 and October 31, 2020. We found that preprint servers hosted almost 25% of COVID-19–related science, that these COVID-19 preprints were being accessed and downloaded in far greater volume than other preprints on the same servers, and that these were widely shared across multiple online platforms. Moreover, we determined that COVID-19 preprints are shorter and are published in journals with a shorter delay following posting than their non-COVID-19 counterparts. Taken together, our data demonstrate the importance of rapidly and openly sharing science in the context of a global pandemic and the essential role of preprints in this endeavour.

Results

COVID-19 preprints were posted early in the pandemic and represent a significant proportion of the COVID-19 literature

The COVID-19 pandemic has rapidly spread across the globe, from 3 patients in the city of Wuhan on the December 27, 2019 to over 46.1 million confirmed cases worldwide by the end of October 2020 (Fig 1A). The scientific community responded rapidly as soon as COVID-19 emerged as a serious threat, with publications appearing within weeks of the first reported cases (Fig 1B). By the end of April 2020, over 19,000 scientific publications had appeared, published both in scientific journals (12,679; approximately 65%) and on preprint servers (6,710; approximately 35%) (Fig 1B)—in some cases, preprints had already been published in journals during this time period and thus contribute to the counts of both sources. Over the following months, the total number of COVID-19–related publications increased approximately linearly, although the proportion of these which were preprints fell: By the end of October, over 125,000 publications on COVID-19 had appeared (30,260 preprints; approximately 25%). Given an output of approximately 5 million journal articles and preprints in the entirety of 2020 (according to data from Dimensions; https://dimensions.ai), the publication response to COVID-19 represented >2.5% of outputs during our analysis period. In comparison to other recent outbreaks of global significance caused by emerging RNA viruses, the preprint response to COVID-19 has been much larger; 10,232 COVID-19–related preprints were posted to bioRxiv and medRxiv in the first 10 months of the pandemic; in comparison, only 78 Zika virus–related and 10 Ebola virus–related preprints were posted to bioRxiv during the entire duration of the respective Zika virus epidemic (2015 to 2016) and Western African Ebola virus epidemic (2014 to 2016) (S1A Fig). This surge in COVID-19 preprints is not explained by general increases in preprint server usage; considering counts of outbreak-related and non-outbreak–related preprints for each outbreak (COVID-19, Ebola or Zika virus), preprint type was significantly associated with outbreak (chi-squared, $\chi 2 = 2559.2$, $p < 0.001$), with the proportion of outbreak-related preprints being greatest for COVID-19.

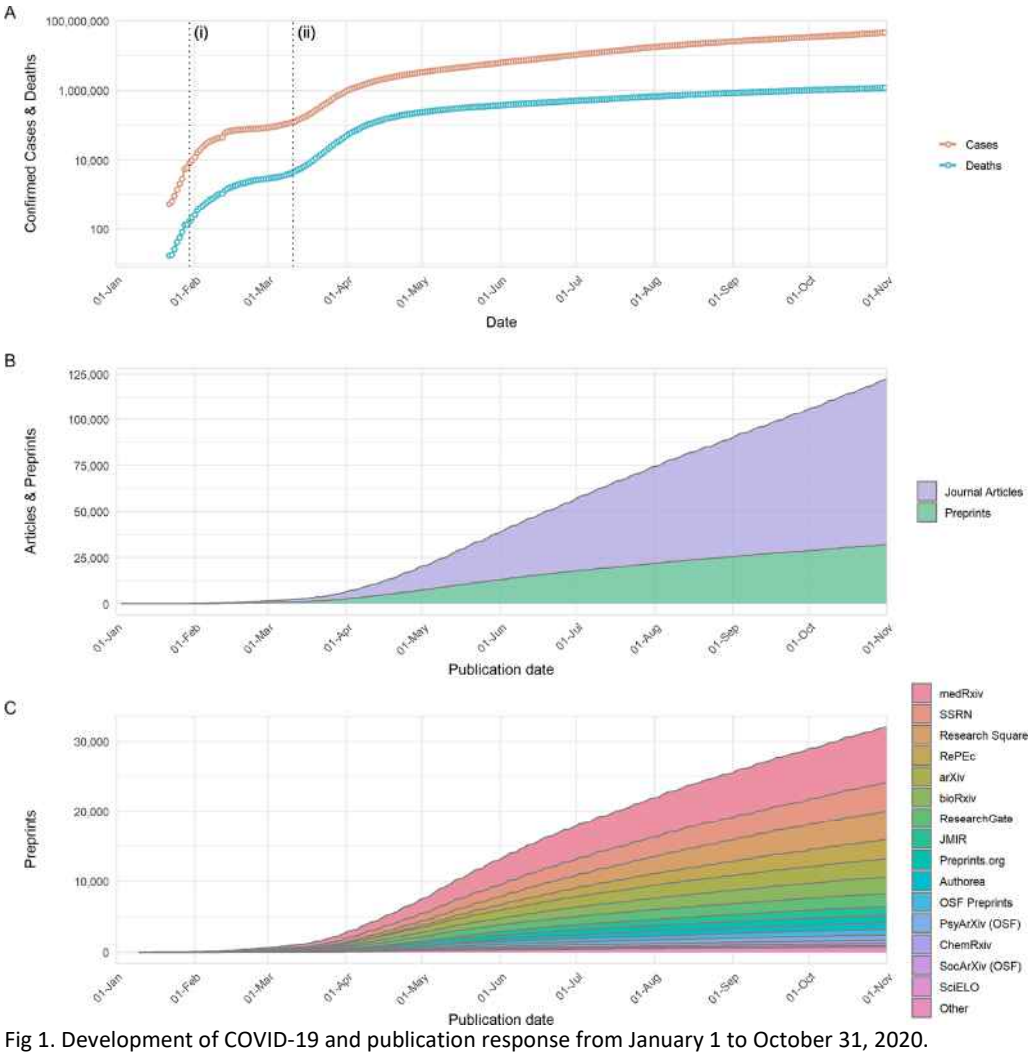


Fig 1. Development of COVID-19 and publication response from January 1 to October 31, 2020.

(A) Number of COVID-19 confirmed cases and reported deaths. Data are sourced from https://github.com/datasets/covid-19/, based on case and death data aggregated by the Johns Hopkins University Center for Systems Science and Engineering (https://systems.jhu.edu/). Vertical lines labelled (i) and (ii) refer to the date on which the WHO declared COVID-19 outbreak a Public Health Emergency of International Concern, and the date on which the WHO declared the COVID-19 outbreak to be a pandemic, respectively.

(B) Cumulative growth of journal articles and preprints containing COVID-19–related search terms.

(C) Cumulative growth of preprints containing COVID-19–related search terms, categorised by individual preprint servers. Journal article data in (B) are based upon data extracted from Dimensions (https://www.dimensions.ai; see Methods section for further details), and preprint data in (B) and (C) are based upon data gathered by Fraser and Kramer (2020). The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019; WHO, World Health Organization.

https://doi.org/10.1371/journal.pbio.3000959.g001

The 30,260 manuscripts posted as preprints were hosted on a range of preprint servers covering diverse subject areas not limited to biomedical research (Fig 1C, data from [15]). It is important to note that this number includes preprints that may have been posted on multiple preprint servers simultaneously; however, by considering only preprints with unique titles (case insensitive), it appears that this only applies to a small proportion of preprint records (<5%). The total number is preprints is nevertheless likely an underestimation of the true volume of preprints posted, as a number of preprint servers and other repositories (e.g., institutional repositories) that could be expected to host COVID-19 research are not included[15]. Despite being one of the newest preprint servers, medRxiv hosted the largest number of preprints (7,882); the next largest were SSRN (4,180), Research Square (4,089), RePEc (2,774), arXiv (2,592), bioRxiv (2,328), JMIR (1,218), and Preprints.org (1,020); all other preprint servers were found to host <1,000 preprints (Fig 1C).

One of the most frequently cited benefits of preprints is that they allow free access to research findings[16], while a large proportion of journal articles often remain behind subscription paywalls. In response to the pandemic, a number of journal publishers began to alter their open-access policies in relation to COVID-19 manuscripts. One such change was to make COVID-19 literature temporarily open access (at least for the duration of the pandemic), with over 80,000 papers in our dataset being open access (S1B Fig).

Attributes of COVID-19 preprints posted between January and October 2020

To explore the attributes of COVID-19 preprints in greater detail, we focused our following investigation on two of the most popular preprint servers in the biomedical sciences: bioRxiv and medRxiv. We compared attributes of COVID-19–related preprints posted within our analysis period between January 1 and October 31, 2020 against non-COVID-19–related preprints posted in the same time frame. In total, 44,503 preprints were deposited to bioRxiv and medRxiv in this period, of which the majority (34,271, 77.0%) were non-COVID-19–related preprints (Fig 2A, S1 Table). During the early phase of the pandemic, the posted monthly volumes of non-COVID-19 preprints was relatively constant, while the monthly volume of COVID-19 preprints increased, peaking at 1,967 in May, and subsequently decreased month by month. These patterns persisted when the 2 preprint servers were considered independently (S2A Fig). Moreover, COVID-19 preprints have represented the majority of preprints posted to medRxiv each month after February 2020.

The increase in the rate of preprint posting poses challenges for their timely screening. A minor but detectable difference was observed between screening time for COVID-19 and non-COVID-19 preprints (Fig 2B), although this difference appeared to vary with server (2-way ANOVA, interaction term; $F_{1,83333} = 19.22$, $p < 0.001$). Specifically, screening was marginally slower for COVID-19 preprints than for non-COVID-19 preprints deposited to medRxiv (mean difference = 0.16 days; Tukey honest significant difference [HSD] test, $p < 0.001$), but not to bioRxiv ($p = 0.981$). The slower screening time for COVID-19 preprints was a result of more of these preprints being hosted on medRxiv, which had slightly longer screening times overall; bioRxiv screened preprints approximately 2 days quicker than medRxiv independent of COVID-19 status (both $p < 0.001$; S2B Fig, S1 Table).

Preprint servers offer authors the opportunity to post updated versions of a preprint, enabling them to incorporate feedback, correct mistakes, or add additional data and analysis. The majority of preprints existed as only a single version for both COVID-19 and non-COVID-19 works, with very few preprints existing in more than 2 versions (Fig 2C). This may somewhat reflect the relatively short time span of our analysis period. Although distributions were similar, COVID-19 preprints appeared to have a slightly greater number of versions, 1 [IQR 1] versus 1 [IQR 0]; Mann–Whitney test, $p < 0.001$). The choice of preprint server did not appear to impact on the number of versions (S2C Fig, S1 Table).

bioRxiv and medRxiv allow authors to select from a number of different Creative Commons license types when depositing their work: CC0 (No Rights Reserved), CC-BY (Attribution), CC BY-NC (Attribution, Noncommercial), CC-BY-ND (Attribution, No Derivatives), and CC-BY-NC-ND (Attribution, Noncommercial, No Derivatives). Authors may also select to post their work without a license (i.e., All Rights Reserved) that allows text and data mining. A previous analysis has found that bioRxiv authors tend to post preprints under the more restrictive license types[17], although there appears to be some confusion among authors as to the precise implications of each license type[18]. License choice was significantly associated with preprint category (chi-squared, $\chi^2 = 336.0$, df = 5, $p < 0.001$); authors of COVID-19 preprints were more likely to choose the more restrictive CC-BY-NC-ND or CC-BY-ND than those of non-COVID-19 preprints and less likely to choose CC-BY (Fig 2D). Again, the choice of preprint server did not appear to impact on the type of license selected by the authors (S2D Fig).

Given the novelty of the COVID-19 research field and rapid speed at which preprints are being posted, we hypothesised that researchers may be posting preprints in a less mature state, or based on a smaller literature base than for non-COVID preprints. To investigate this, we

compared the word counts and reference counts of COVID-19 preprints and non-COVID-19 preprints from bioRxiv (at the time of data extraction, HTML full texts from which word and reference counts were derived were not available for medRxiv) (Fig 2E). We found that COVID-19 preprints are on average 32% shorter in length than non-COVID-19 preprints (median, 3,965 [IQR 2,433] versus 5,427 [IQR 2,790]; Mann–Whitney test, $p < 0.001$) (S1 Table). Although the length of preprints gradually increased over the analysis period, COVID-19 preprints remained shorter than non-COVID-19 preprints with a similar difference in word count, even when adjusted for factors such as authorship team size and bioRxiv subject categorisation (S1 Model, S2 and S3 Tables). COVID-19 preprints also contain fewer references than non-COVID-19 preprints (Fig 2F), although not fewer than expected relative to overall preprint length, as little difference was detec ted in reference:word count ratios (median, 1:103 versus 1:101; $p = 0.052$). As word counts increased over time, the reference counts per preprint also steadily increased.

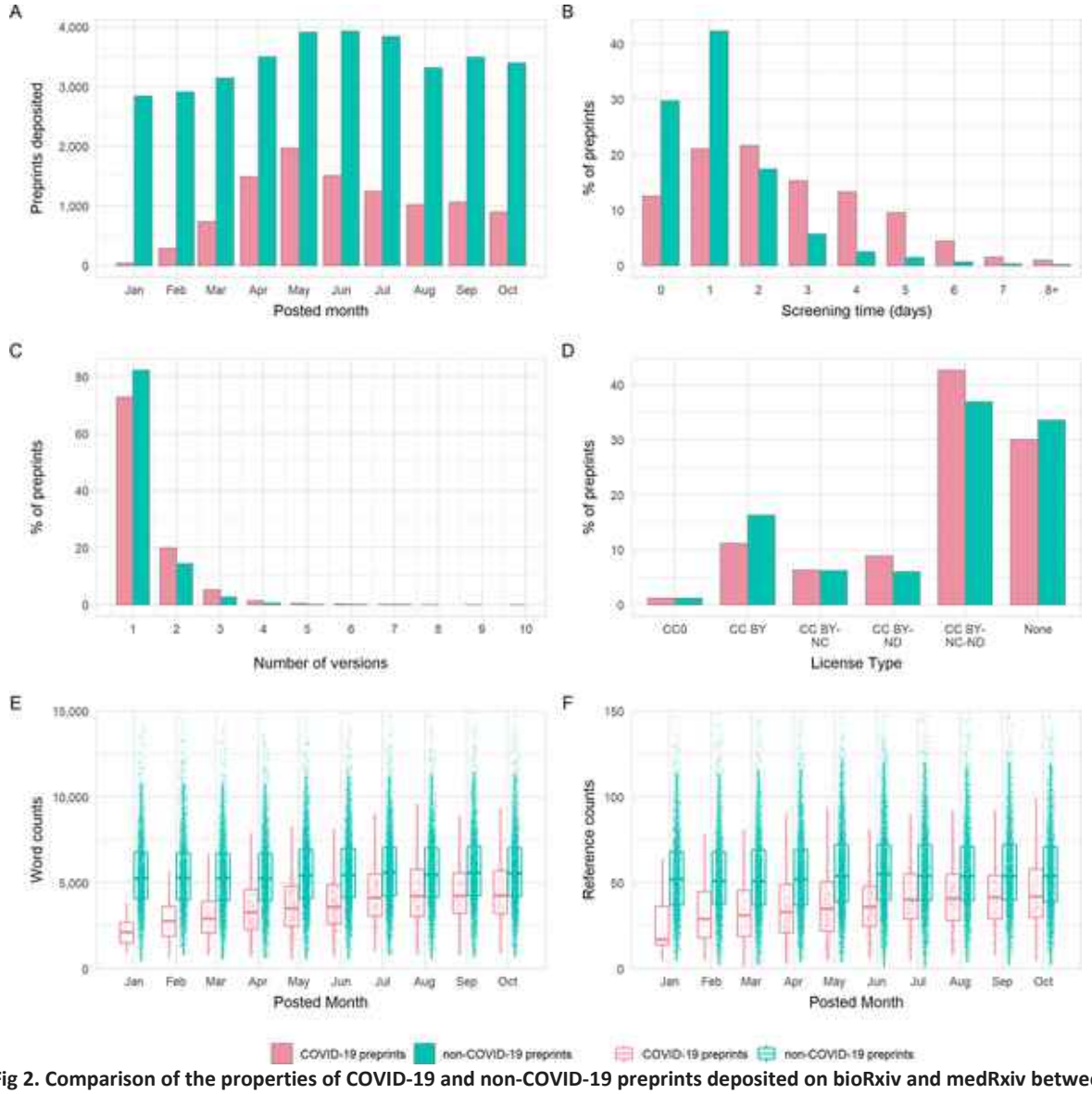


Fig 2. Comparison of the properties of COVID-19 and non-COVID-19 preprints deposited on bioRxiv and medRxiv between January 1 and October 31, 2020.

(A) Number of new preprints deposited per month.

(B) Preprint screening time in days.

(C) License type chosen by authors.

(D) Number of versions per preprint.

(E) Boxplot of preprint word counts, binned by posting month.

(F) Boxplot of preprint reference counts, binned by posting month. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019. https://doi.org/10.1371/journal.pbio.3000959.g002

Scientists turned to preprints for the first time to share COVID-19 science

The number of authors per preprint may give an additional indication as to the amount of work, resources used, and the extent of collaboration in a manuscript. Although little difference was seen in number of authors between preprint servers (S1 Table), COVID-19 preprints had a marginally higher number of authors than non-COVID-19 preprints on average (median, 7 [IQR 8] versus 6 [IQR 5]; $p < 0.001$), due to the greater likelihood of large (11+) authorship team sizes (Fig 3A). However, single-author preprints were approximately 2.6 times more common for COVID-19 (6.1% of preprints) than non-COVID-19 preprints (2.3% of preprints) (Fig 3A).

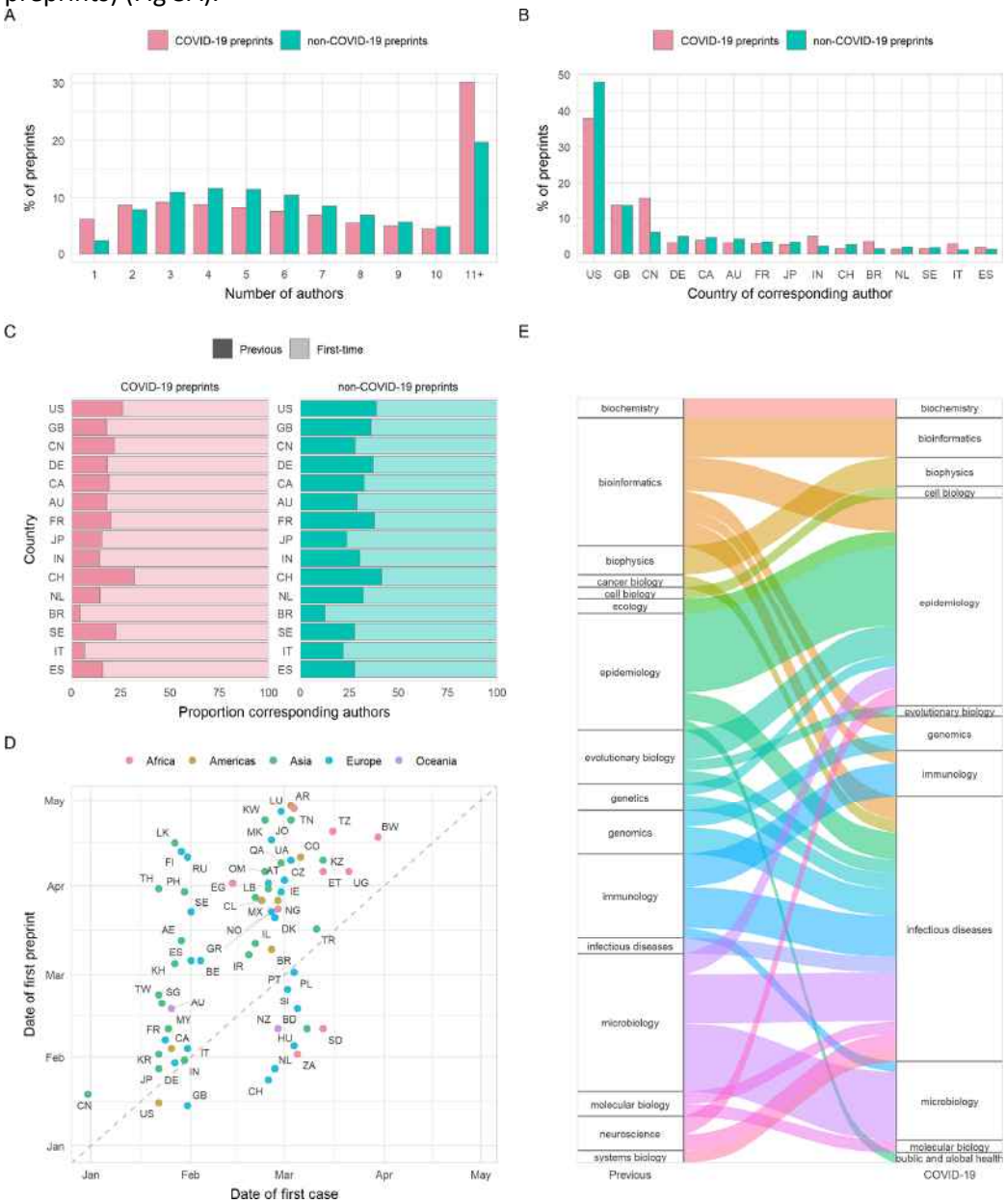


Fig 3. Properties of authors of COVID-19 and non-COVID-19 preprints deposited on bioRxiv and medRxiv between January 1 and October 31, 2020.

(A) Proportion of preprints with *N* authors. (B) Proportion of preprints deposited by country of corresponding author (top 15 countries by total preprint volume are shown). (C) Proportions of COVID-19 and non-COVID-19 corresponding authors from each of the top 15 countries shown in (B) that had previously posted a preprint (darker bar) or were posting a preprint for the first time (lighter bar). (D) Correlation between date of the first preprint originating from a country (according to the affiliation of the corresponding author) and the date of the first confirmed case from the same country for COVID-19 preprints. (E) Change in bioRxiv/medRxiv preprint posting category for COVID-19 preprint authors compared to their previous preprint (COVID-19 or non-COVID-19), for category combinations with n > = 5 authors. For all panels containing country information, labels refer to ISO 3166 character codes. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019. https://doi.org/10.1371/journal.pbio.3000959.g003

The largest proportion of preprints in our dataset were from corresponding authors in the United States, followed by significant proportions from the United Kingdom and China (Fig 3B). It is notable that China is overrepresented in terms of COVID-19 preprints compared to its non-COVID-19 preprint output: 39% of preprints from Chinese corresponding authors were COVID-19 related, compared to 16.5% of the US output and 20.1% of the UK output. We also found a significant association for corresponding authors between preprint type (COVID-19 or non-COVID-19) and whether this was the author's first bioRxiv or medRxiv preprint (chi-squared, $\chi^2 = 840.4$, df = 1, $p < 0.001$). Among COVID-19 corresponding authors, 85% were posting a preprint for the first time, compared to 69% of non-COVID-19 corresponding authors in the same period. To further understand which authors have been drawn to begin using preprints since the pandemic began, we stratified these groups by country (S4 Table) and found significant associations for the US, UK, Germany, India (Bonferroni adjusted $p < 0.001$), France, Canada, Italy ($p < 0.01$), and China ($p < 0.05$). In all cases, a higher proportion were posting a preprint for the first time among COVID-19 corresponding authors than non-COVID-19 corresponding authors. Moreover, we found that most countries posted their first COVID-19 preprint close to the time of their first confirmed COVID-19 case (Fig 3D), with weak positive correlation considering calendar days of both events (Spearman rank; $\rho = 0.54$, $p < 0.001$). Countries posting a COVID-19 preprint in advance of their first confirmed case were mostly higher-income countries (e.g., US, UK, New Zealand, and Switzerland). COVID-19 preprints were deposited from over 100 countries, highlighting the global response to the pandemic.

There has been much discussion regarding the appropriateness of researchers switching to COVID-19 research from other fields[19]. To quantify whether this phenomenon was detectable within the preprint literature, we compared the bioRxiv or medRxiv category of each COVID-19 preprint to the most recent previous non-COVID-19 preprint (if any) from the same corresponding author. Most corresponding authors were not drastically changing fields, with category differences generally spanning reasonably related areas. For example, some authors that previously posted preprints in evolutionary biology have posted COVID-19 preprints in microbiology (Fig 3E). This suggests that—at least within the life sciences—principal investigators are utilising their labs' skills and resources in an expected manner in their contributions to COVID-19 research.

COVID-19 preprints were published quicker than non-COVID-19 preprints

Critics have previously raised concerns that by forgoing the traditional peer-review process, preprint servers could be flooded by poor-quality research[20,21]. Nonetheless, earlier analyses have shown that a large proportion of preprints (approximately 70%) in the biomedical sciences are eventually published in peer-reviewed scientific journals[8]. We assessed differences in publication outcomes for COVID-19 versus non-COVID-19 preprints during our analysis period, which may be partially related to differences in preprint quality. Published status (published/unpublished) was significantly associated with preprint type (chi-squared, $\chi^2 = 186.2$, df = 1, $p < 0.001$); within our time frame, 21.1% of COVID-19 preprints were published in total by the end of October, compared to 15.4% of non-COVID preprints. As expected, greater proportions published were seen among preprints posted earlier, with over 40% of COVID-19 preprints submitted in January published by the end of October and less than 10% for those published in August or later (Fig 4A). Published COVID-19 preprints were distributed across many journals, with clinical or multidisciplinary journals tending to publish the most COVID-19 preprints (Fig 4B). To determine how publishers were prioritising COVID-19 research, we compared the time from preprint posting to publication in a journal. The time interval from posting to subsequent publication was significantly reduced for COVID-19 preprints by a difference in medians of 48 days compared to non-COVID-19 preprints posted in the same time period (68 days [IQR 69] versus 116 days [IQR 90]; Mann–Whitney test, $p < 0.001$). This did not appear to be driven by any temporal changes in publishing practices, as the distribution of publication times for non-COVID-19 preprints was similar to our control time frame of January to December 2019 (Fig 4C). This acceleration additionally varied between publishers (2-way ANOVA, interaction term preprint type*publisher; $F_{9,5273} = 6.58$, $p < 0.001$)

and was greatest for the American Association for the Advancement of Science (AAAS) at an average difference of 102 days (Tukey HSD; $p < 0.001$) (Fig 4D).

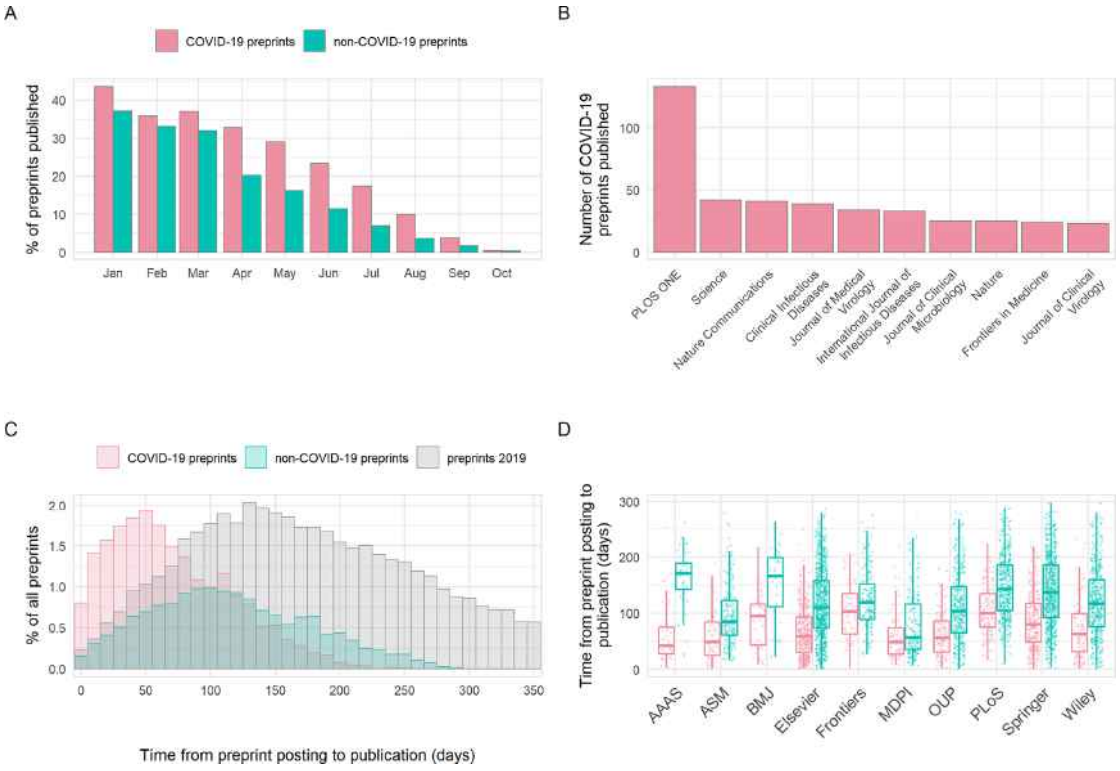


Fig 4. Publication outcomes of COVID-19 and non-COVID-19 preprints deposited on bioRxiv and medRxiv between January 1 and October 31, 2020.

(A) Percentage of COVID-19 versus non-COVID-19 preprints published in peer-reviewed journals, by preprint posting month. (B) Destination journals for COVID-19 preprints that were published within our analysis period. Shown are the top 10 journals by publication volume. (C) Distribution of the number of days between posting a preprint and subsequent journal publication for COVID-19 preprints (red), non-COVID-19 preprints posted during the same period (January to October 2020) (green), and non-COVID-19 preprints posted between January and December 2019 (grey). (D) Time from posting on bioRxiv or medRxiv to publication categorised by publisher. Shown are the top 10 publishers by publication volume. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019.

https://doi.org/10.1371/journal.pbio.3000959.g004

Extensive access of preprint servers for COVID-19 research

At the start of our time window, COVID-19 preprints received abstract views at a rate over 18 times that of non-COVID-19 preprints (Fig 5A) (time-adjusted negative binomial regression; rate ratio = 18.2, z = 125.0, $p < 0.001$) and downloads at a rate of almost 30 times (Fig 5B) (rate ratio = 27.1, z = 124.2, $p < 0.001$). Preprints posted later displayed lower usage rates, in part due to the reduced length of time they were online and able to accrue views and downloads. However, decreases in both views and downloads by posting date was stronger for COVID-19 preprints versus non-COVID-19 preprints (preprint type*calendar day interaction terms, both $p < 0.001$); each additional calendar month in posting date resulted in an estimated 24.3%/7.4% reduction in rate of views and an estimated 28.5%/12.0% reduction in rate of downloads for COVID-19/non-COVID-19 preprints, respectively. Similar trends of decrease were observed when restricting view and download data to the first respective month of each preprint, with highest rates of usage for those posted in January (S3A and S3B Fig). The disparity between COVID-19 and non-COVID-19 preprints suggests that either COVID-19 preprints continued to slowly accumulate total usage well beyond their first month online (Fig 5) and/or they received a more diluted share of relative initial interest as larger volumes of preprints (and publications) were available by later months (Fig 1B).

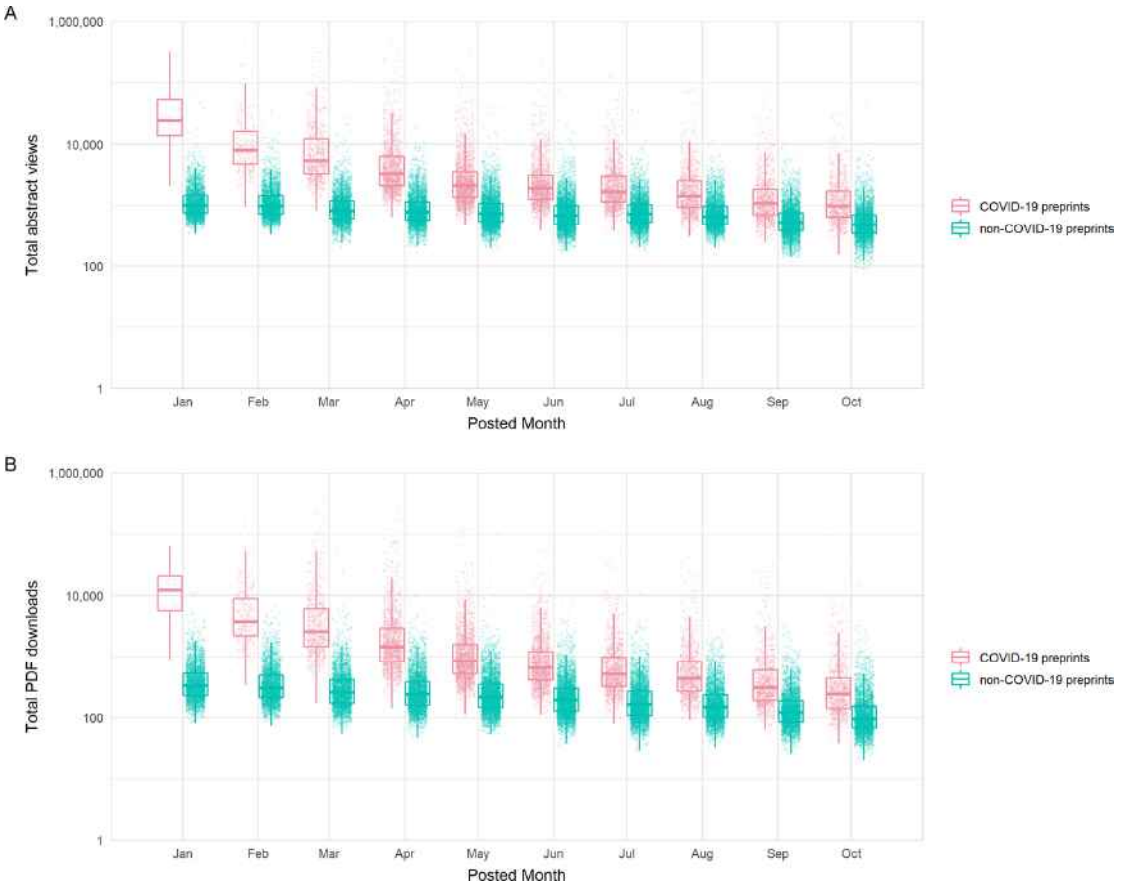


Fig 5. Access statistics for COVID-19 and non-COVID-19 preprints posted on bioRxiv and medRxiv.
(A) Boxplots of abstract views, binned by preprint posting month. (B) Boxplots of PDF downloads, binned by preprint posting month. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019. https://doi.org/10.1371/journal.pbio.3000959.g005

To confirm that usage of COVID-19 and non-COVID-19 preprints was not an artefact of differing preprint server reliance during the pandemic, we compared usage rates during the pandemic period with those from the previous year (January to December 2019), as a non-pandemic control period. Beyond the expected effect of fewer views/downloads of preprints that have been uploaded for a shorter time, the usage data did not differ from that prior to the pandemic (S3C and S3D Fig).

Secondly, we investigated usage across additional preprint servers (data kindly provided by each of the server operators). We found that COVID-19 preprints were consistently downloaded more than non-COVID-19 preprints during our time frame, regardless of which preprint server hosted the manuscript (S3E Fig), although the gap in downloads varied between server (2-way ANOVA, interaction term; $F_{3,89990} = 126.6$, $p < 0.001$). Server usage differences were more pronounced for COVID-19 preprints; multiple post hoc comparisons confirmed that bioRxiv and medRxiv received significantly higher usage per COVID-19 preprint than all other servers for which data were available (Tukey HSD; all p values < 0.001). However, for non-COVID-19 preprints, the only observed pairwise differences between servers indicated greater bioRxiv and medRxiv usage than Research Square (Tukey HSD; $p < 0.001$). This suggests that specific attention has been given disproportionately to bioRxiv and medRxiv as repositories for COVID-19 research.

COVID-19 preprints were shared and cited more widely than non-COVID-19 preprints

We quantified the citation and online sharing behaviour of COVID-19 preprints using citation count data from Dimensions (https://dimensions.ai) and counts of various altmetric indicators using data from Altmetric (https://altmetric.com) (Fig 6; further details on data sources in Methods section). In terms of citations, we found higher proportions overall of COVID-19

preprints that received at least a single citation (57.9%) than non-COVID-19 preprints (21.5%) during our study period of January 1 to October 31, 2020, although the citation coverage expectedly decreased for both groups for newer posted preprints (Fig 6A). COVID-19 preprints also have greater total citation counts than non-COVID-19 preprints (time-adjusted negative binomial regression; rate ratio = 13.7, z = 116.3, $p < 0.001$). The highest cited COVID-19 preprint had 652 citations, with the 10th most cited COVID-19 preprint receiving 277 citations (Table 1); many of the highest cited preprints focussed on the viral cell receptor, angiotensin converting enzyme 2 (ACE2), or the epidemiology of COVID-19.

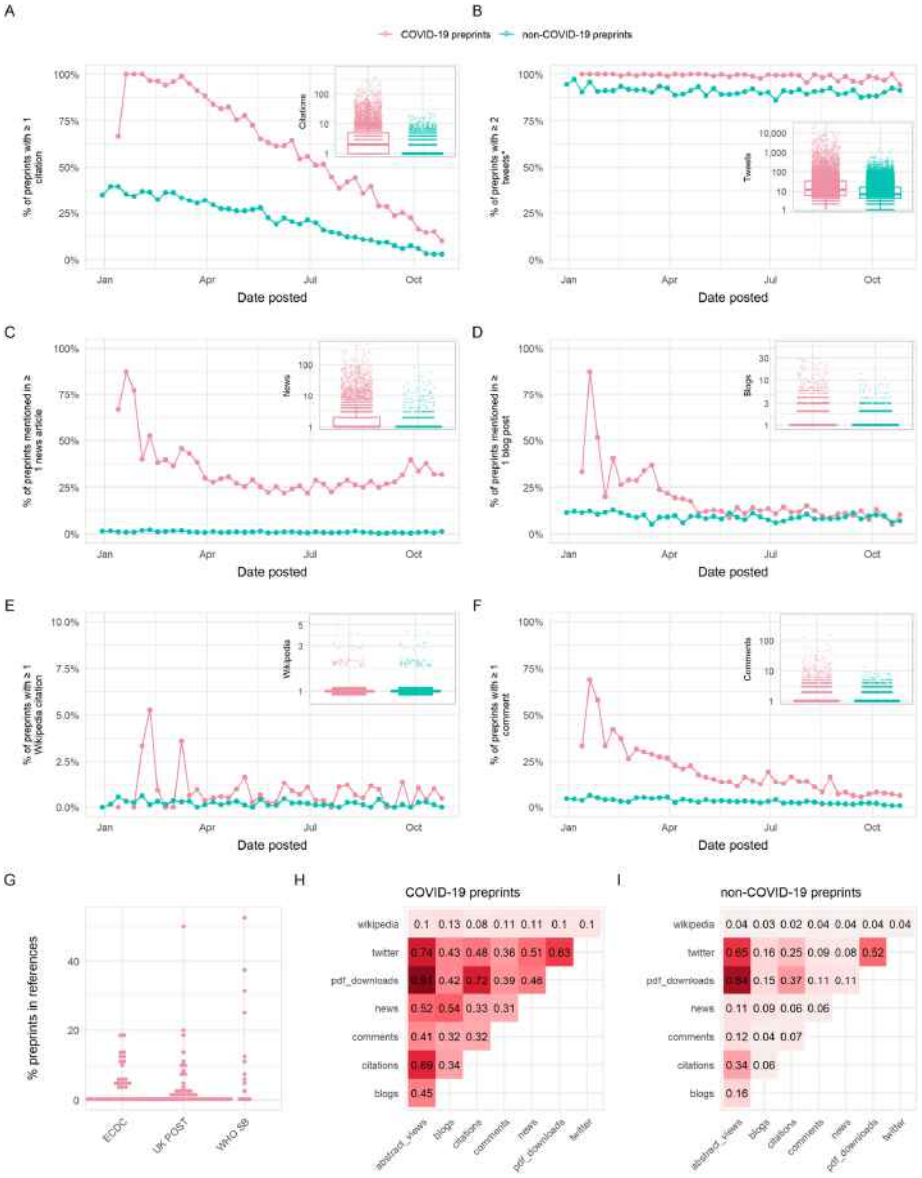


Fig 6. Usage of COVID-19 and non-COVID-19 preprints posted on bioRxiv and medRxiv between January 1 and October 31, 2020.

Panels (A)–(F) show the proportion of preprints receiving at least 1 citation or mention in a given source, with the exception of panel (B) which shows the proportion of preprints receiving at least 2 tweets (to account for the fact that each preprint is tweeted once automatically by the official bioRxiv/medRxiv Twitter accounts). The inset in each panel shows a boxplot comparing citations/mentions for all COVID-19 and non-COVID-19 preprints posted within our analysis period. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility. Data are plotted on a log-scale with +1 added to each count for visualisation. (G) Proportion of preprints included in reference lists of policy documents from 3 sources: the ECDC, UK POST, and WHO SB. (H) Spearman correlation matrix between indicators shown in panels (A)–(F), as well as abstract views and PDF downloads for COVID-19 preprints. (I) Spearman correlation matrix between indicators shown in panels (A)–(F), in addition to abstract views and PDF downloads for non-COVID-19 preprints. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019; ECDC, European Centre for Disease Prevention and Control; UK POST, United Kingdom Parliamentary Office of Science and Technology; WHO SB, World Health Organization Scientific Briefs.

https://doi.org/10.1371/journal.pbio.3000959.g006

Table 1. Top 10 cited COVID-19 preprints.

Rank	Source	DOI	Title	Posted date	Citations
1	medRxiv	10.1101/2020.02.06.20020974	Clinical characteristics of 2019 novel coronavirus infection in China	February 9, 2020	652
2	bioRxiv	10.1101/2020.02.07.937862	Severe acute respiratory syndrome-related coronavirus—The species and its viruses, a statement of the Coronavirus Study Group	February 11, 2020	513
3	medRxiv	10.1101/2020.01.23.20018549	Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions	January 24, 2020	361
4	bioRxiv	10.1101/2020.01.26.919985	Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov	January 26, 2020	359
5	medRxiv	10.1101/2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	March 30, 2020	358
6	bioRxiv	10.1101/2020.01.31.929042	The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells	January 31, 2020	313
7	bioRxiv	10.1101/2020.02.03.931766	Specific ACE2 Expression in Cholangiocytes May Cause Liver Damage After 2019-nCoV Infection	February 4, 2020	304
8	bioRxiv	10.1101/2020.01.22.914952	Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin	January 23, 2020	302
9	bioRxiv	10.1101/2020.01.30.927806	The digestive system is a potential route of 2019-nCov infection: a bioinformatics analysis based on single-cell transcriptomes	January 31, 2020	285
10	medRxiv	10.1101/2020.02.11.20021493	Laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections	February 12, 2020	277

COVID-19, Coronavirus Disease 2019.

https://doi.org/10.1371/journal.pbio.3000959.t001

Sharing of preprints on Twitter may provide an indicator of the exposure of wider public audiences to preprints. COVID-19 preprints received greater Twitter coverage (98.9% received >1 tweet) than non-COVID-19 preprints (90.7%) (note that the threshold for Twitter coverage was set at 1 rather than 0, to account for automated tweets by the official bioRxiv and medRxiv Twitter accounts) and were tweeted at an overall greater rate than non-COVID-19 preprints (rate ratio = 7.6, z = 135.7, $p < 0.001$) (Fig 6B). The most tweeted non-COVID-19 preprint received 1,656 tweets, whereas 8 of the top 10 tweeted COVID-19 preprints were tweeted over 10,500 times each (Table 2). Many of the top 10 tweeted COVID-19 preprints were related to transmission, reinfection, or seroprevalence. The most tweeted COVID-19 preprint (26,763 tweets) was a study investigating antibody seroprevalence in California[22]. The fourth most tweeted COVID-19 preprint was a widely criticised (and later withdrawn) study linking the SARS-CoV-2 spike protein to HIV-1 glycoproteins[23].

Table 2. Top 10 tweeted COVID-19 preprints.

Rank	Source	DOI	Title	Posted date	Tweets	News articles	Blogs
1	medRxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	April 17, 2020	26,763	434	55
2	medRxiv	10.1101/2020.04.04.20053058	Indoor transmission of SARS-CoV-2	April 7, 2020	21,831	187	34
3	medRxiv	10.1101/2020.07.15.20151852	Effect of Hydroxychloroquine in Hospitalized Patients with COVID-19: Preliminary results from a multi-centre, randomized, controlled trial.	July 15, 2020	17,534	83	5
4	bioRxiv	10.1101/2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	January 31, 2020	16,608	102	25
5	medRxiv	10.1101/2020.05.19.20105999	SARS-CoV-2 RNA concentrations in primary municipal sewage sludge as a leading indicator of COVID-19 outbreak dynamics	May 22, 2020	16,582	63	8
6	medRxiv	10.1101/2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	March 30, 2020	14,614	106	18
7	medRxiv	10.1101/2020.10.14.20212555	Multi-organ impairment in low-risk individuals with long COVID	October 16, 2020	12,871	34	6
8	medRxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	March 10, 2020	12,484	354	29
9	medRxiv	10.1101/2020.08.03.20167395	Viable SARS-CoV-2 in the air of a hospital room with COVID-19 patients	August 4, 2020	11,770	121	8
10	medRxiv	10.1101/2020.03.30.20048165	Association of BCG vaccination policy with prevalence and mortality of COVID-19	April 6, 2020	10,701	7	0

COVID-19, Coronavirus Disease 2019.

https://doi.org/10.1371/journal.pbio.3000959.t002

To better understand the discussion topics associated with highly tweeted preprints, we analysed the hashtags used in original tweets (i.e., excluding retweets) mentioning the top 100 most tweeted COVID-19 preprints (S4A Fig). In total, we collected 30,213 original tweets containing 11,789 hashtags; we filtered these hashtags for those occurring more than 5 times and removed a selection of generic or overused hashtags directly referring to the virus (e.g., "#coronavirus" and "#covid-19"), leaving a final set of 2,981 unique hashtags. While many of the top-used hashtags were direct, neutral references to the disease outbreak such as "#coronavirusoutbreak" and "#wuhan," we also found a large proportion of politicised tweets using hashtags associated with conspirational ideologies (e.g., "#qanon," "#wwg1wga," an

abbreviation of "Where We Go One, We Go All" a tag commonly used by QAnon supporters), xenophobia (e.g., "#chinazi"), or US-specific right-wing populism (e.g., "#maga"). Other hashtags also referred to topics directly associated with controversial preprints, e.g., "#hydroxychloroquine" and "#hiv," both of which were major controversial topics associated with several of the top 10 most tweeted preprints.

As well as featuring heavily on social media, COVID-19 research has also pervaded print and online news media. In terms of coverage, 28.7% of COVID-19 preprints were featured in at least a single news article, compared to 1.0% of non-COVID-19 preprints (Fig 6C), and were used overall in news articles at a rate almost 100 times that of non-COVID-19 preprints (rate ratio = 92.8, z = 83.3, $p < 0.001$). The top non-COVID-19 preprint was reported in 113 news articles, whereas the top COVID-19 preprints were reported in over 400 news articles (Table 3). Similarly, COVID-19 preprints were also used more in blogs (coverage COVID-19/non-COVID-19 preprints = 14.3%/9.1%, rate ratio = 3.73, z = 37.3, $p < 0.001$) and Wikipedia articles (coverage COVID-19/non-COVID-19 preprints = 0.7%/0.2%, rate ratio = 4.47, z = 7.893, $p < 0.001$) at significantly greater rates than non-COVID-19 preprints (Fig 6D and 6E, Table 4). We noted that several of the most widely disseminated preprints that we classified as being non-COVID-19 related featured topics nonetheless relevant to generalised infectious disease research, such as human respiratory physiology and personal protective equipment.

Table 3. Top 10 COVID-19 preprints covered by news organisations.

Rank	Source	DOI	Title	Posted date	Tweets	News articles	Blogs
1	bioRxiv	10.1101/ 2020.04.29.069054	Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2	April 30, 2020	6,848	449	29
2	medRxiv	10.1101/ 2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	April 17, 2020	26,763	434	55
3	medRxiv	10.1101/ 2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	April 21, 2020	10,385	411	27
4	medRxiv	10.1101/ 2020.10.15.20209817	Repurposed antiviral drugs for COVID-19; interim WHO SOLIDARITY trial results	October 15, 2020	8,569	396	25
5	medRxiv	10.1101/ 2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	March 10, 2020	12,484	354	29
6	medRxiv	10.1101/ 2020.05.15.20103655	Differential Effects of Intervention Timing on COVID-19 Spread in the United States	May 20, 2020	1,831	295	16
7	medRxiv	10.1101/ 2020.07.09.20148429	Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection	July 11, 2020	2,167	281	27
8	medRxiv	10.1101/ 2020.08.12.20169359	Effect of Convalescent Plasma on Mortality among Hospitalized Patients with COVID-19: Initial Three-Month Experience	August 12, 2020	2,746	264	26
9	medRxiv	10.1101/ 2020.06.22.20137273	Effect of Dexamethasone in Hospitalized Patients with COVID-19: Preliminary Report	June 22, 2020	5,698	246	26
10	medRxiv	10.1101/ 2020.03.11.20031096	Relationship between the ABO Blood Group and the COVID-19 Susceptibility	March 16, 2020	4,055	245	23

COVID-19, Coronavirus Disease 2019.

https://doi.org/10.1371/journal.pbio.3000959.t003

Table 4. Top 10 commented on COVID-19 preprints.

Rank	Source	DOI	Title	Posted date	Comments count
1	medRxiv	10.1101/ 2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	April 17, 2020	582
2	medRxiv	10.1101/ 2020.03.24.20042937	Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study	March 28, 2020	149
3	bioRxiv	10.1101/ 2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	January 31, 2020	129
4	medRxiv	10.1101/ 2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	April 21, 2020	129
5	bioRxiv	10.1101/ 2020.04.29.069054	Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2	April 30, 2020	75
6	medRxiv	10.1101/ 2020.03.11.20031096	Relationship between the ABO Blood Group and the COVID-19 Susceptibility	March 16, 2020	72
7	medRxiv	10.1101/ 2020.03.27.20043752	Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months	March 30, 2020	61
8	medRxiv	10.1101/ 2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	March 30, 2020	58
9	medRxiv	10.1101/ 2020.04.16.20067835	Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs	April 22, 2020	56
10	medRxiv	10.1101/ 2020.04.05.20054361	Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters	April 8, 2020	53

COVID-19, Coronavirus Disease 2019.

https://doi.org/10.1371/journal.pbio.3000959.t004

A potential benefit of preprints is that they allow authors to receive an incorporate feedback from the wider community prior to journal publication. To investigate feedback and engagement with preprints, we quantified the number of comments received by preprints

directly via the commenting system on the bioRxiv and medRxiv platforms. We found that non-COVID-19 preprints were commented upon less frequently compared to COVID-19 preprints (coverage COVID-19/non-COVID-19 preprints = 15.9%/3.1%, time-adjusted negative binomial regression; rate ratio = 11.0, z = 46.5, $p < 0.001$) (Fig 6F); the most commented non-COVID-19 preprint received only 68 comments, whereas the most commented COVID-19 preprint had over 580 comments (Table 5). One preprint, which had 129 comments, was retracted within 3 days of being posted following intense public scrutiny (Table 4, doi: 10.1101/2020.01.30.927871). As the pandemic has progressed, fewer preprints were commented upon. Collectively, these data suggest that the most discussed or controversial COVID-19 preprints are rapidly and publicly scrutinised, with commenting systems being used for direct feedback and discussion of preprints.

Table 5. Top 10 most blogged COVID-19 preprints.

Rank	Source	DOI	Title	Posted date	Tweets	News articles	Blogs
1	medRxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	April 17, 2020	26,763	434	55
2	medRxiv	10.1101/2020.04.04.20053058	Indoor transmission of SARS-CoV-2	April 7, 2020	21,831	187	34
3	bioRxiv	10.1101/2020.04.29.069054	Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2	April 30, 2020	6,848	449	29
4	medRxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	March 10, 2020	12,484	354	29
5	medRxiv	10.1101/2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	April 21, 2020	10,385	411	27
6	medRxiv	10.1101/2020.07.09.20148429	Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection	July 11, 2020	2,167	281	27
7	medRxiv	10.1101/2020.06.22.20137273	Effect of Dexamethasone in Hospitalized Patients with COVID-19: Preliminary Report	June 22, 2020	5,698	246	26
8	medRxiv	10.1101/2020.08.12.20169359	Effect of Convalescent Plasma on Mortality among Hospitalized Patients with COVID-19: Initial Three-Month Experience	August 12, 2020	2,746	264	26
9	bioRxiv	10.1101/2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	January 31, 2020	16,608	102	25
10	bioRxiv	10.1101/2020.03.30.015347	Susceptibility of ferrets, cats, dogs, and different domestic animals to SARS-coronavirus-2	March 31, 2020	4,168	209	25
11	medRxiv	10.1101/2020.10.15.20209817	Repurposed antiviral drugs for COVID-19; interim WHO SOLIDARITY trial results	October 15, 2020	8,569	396	25

COVID-19, Coronavirus Disease 2019.

https://doi.org/10.1371/journal.pbio.3000959.t005

Within a set of 81 COVID-19 policy documents (which were manually retrieved from the European Centre for Disease Prevention and Control (ECDC), United Kingdom Parliamentary Office of Science and Technology (UK POST), and World Health Organization Scientific Briefs (WHO SB)), 52 documents cited preprints (Fig 6G). However, these citations occurred at a relatively low frequency, typically constituting less than 20% of the total citations in these 52 documents. Among 255 instances of citation to a preprint, medRxiv was the dominant server cited ($n = 209$, 82%), with bioRxiv receiving a small number of citations ($n = 21$) and 5 other servers receiving ≤10 citations each (arXiv, OSF, preprints.org, Research Square, and SSRN). In comparison, only 16 instances of citations to preprints were observed among 38 manually collected non-COVID-19 policy documents from the same sources.

To understand how different usage and sharing indicators may represent the behaviour of different user groups, we calculated the Spearman correlation between the indicators presented above (citations, tweets, news articles, blog mentions, Wikipedia citations, and comment counts) as well as with abstract views and download counts as previously presented (Fig 6H and 6I). Overall, we found stronger correlations between all indicators for COVID-19 preprints compared to non-COVID-19 preprints. For COVID-19 preprints, we found expectedly strong correlation between abstract views and PDF downloads (Spearman $\rho = 0.91$, $p < 0.001$), weak to moderate correlation between the numbers of citations and Twitter shares (Spearman $\rho = 0.48$, $p < 0.001$), and the numbers of citations and news articles (Spearman $\rho = 0.33$, $p < 0.001$) suggesting that the preprints cited extensively within the scientific literature did not necessarily correlate with those that were mostly shared by the wider public on online platforms. There was a slightly stronger correlation between COVID-19 preprints that were most blogged and those receiving the most attention in the news (Spearman $\rho = 0.54$, $p < 0.001$) and moderate correlation between COVID-19 preprints that were most tweeted and those receiving the most attention in the news (Spearman $\rho = 0.51$, $p < 0.001$), suggesting

similarity between preprints shared on social media and in news media. Finally, there was a weak correlation between the number of tweets and number of comments received by COVID-19 preprints (Spearman ρ = 0.36, $p < 0.001$). Taking the top 10 COVID-19 preprints by each indicator, there was substantial overlap between all indicators except citations (S4B Fig).

In summary, our data reveal that COVID-19 preprints received a significant amount of attention from scientists, news organizations, the general public, and policy-making bodies, representing a departure for how preprints are normally shared (considering observed patterns for non-COVID-19 preprints).

Discussion

The usage of preprint servers within the biological sciences has been rising since the inception of bioRxiv and other platforms[10,25]. The urgent threat of a global pandemic has catapulted the use of preprint servers as a means of quickly disseminating scientific findings into the public sphere, supported by funding bodies encouraging preprinting for COVID-19 research[26,27]. Our results show that preprints have been widely adopted for the dissemination and communication of COVID-19 research, and in turn, the pandemic has greatly impacted the preprint and science publishing landscape[28].

Changing attitudes and acceptance within the life sciences to preprint servers may be one reason why COVID-19 research is being shared more readily as preprints compared to previous epidemics. In addition, the need to rapidly communicate findings prior to a lengthy review process might be responsible for this observation (Fig 3). A recent study involving qualitative interviews of multiple research stakeholders found "early and rapid dissemination" to be among the most often cited benefits of preprints[16]. These findings were echoed in a survey of approximately 4,200 bioRxiv users[10] and are underscored by the 6-month median lag between posting of a preprint and subsequent journal publication[8,16]. Such timelines for disseminating findings are clearly incompatible with the lightning-quick progression of a pandemic. An analysis of publication timelines for 14 medical journals has shown that some publishers have taken steps to accelerate their publishing processes for COVID-19 research, reducing the time for the peer-review stage (submission to acceptance) on average by 45 days and the editing stage (acceptance to publication) by 14 day[29], yet this still falls some way short of the approximately 1 to 3 days screening time for bioRxiv and medRxiv preprints (Fig 2B). This advantage may influence the dynamics of preprint uptake: As researchers in a given field begin to preprint, their colleagues may feel pressure to also preprint in order to avoid being scooped. Further studies on understanding the motivations behind posting preprints, for example, through quantitative and qualitative author surveys, may help funders and other stakeholders that support the usage of preprints to address some of the social barriers for their uptake[30].

One of the primary concerns among authors around posting preprints is premature media coverage[16,31]. Many preprint servers created highly visible collections of COVID-19 work, potentially amplifying its visibility. From mid-March 2020, bioRxiv and medRxiv included a banner to explain that preprints should not be regarded as conclusive and not reported on in the news media as established information[32]. Despite this warning message, COVID-19 preprints have received unprecedented coverage on online media platforms (Fig 6). Indeed, even before this warning message was posted, preprints were receiving significant amounts of attention. Twitter has been a particularly notable outlet for communication of preprints, a finding echoed by a recent study on the spread of the wider (i.e., not limited to preprints) COVID-19 research field on Twitter, which found that COVID-19 research was being widely disseminated and driven largely by academic Twitter users[33,34]. Nonetheless, the relatively weak correlation found between citations and other indicators of online sharing (Fig 6H) suggests that the interests of scientists versus the broader public differ significantly: Of the articles in the top 10 most shared on Twitter, in news articles or on blogs, only one is ranked among the top 10 most cited articles (S4B Fig). Hashtags associated with individual, highly tweeted preprints reveal some emergent themes that suggest communication of certain preprints can also extend well beyond scientific audiences (S4A Fig) [34]. These range from good public health practice ("#washyourhands") to right-wing philosophies (#chinalies), conspiracy

theories ("#fakenews" and "#endthelockdown"), and xenophobia ("#chinazi"). Many of the negative hashtags have been perpetuated by public figures such as the former President of America and the right-wing media[35,36]. Following President Trump's diagnosis of COVID-19, one investigation found a wave of anti-Asian sentiment and conspiracy theories across Twitter[37]. This type of misinformation is common to new diseases, and social media platforms have recently released a statement outlining their plans to combat this issue[38]. An even greater adoption of open science principles has recently been suggested as one method to counter the misuse of preprints and peer-reviewed articles[24]; this remains an increasingly important discourse.

The fact that news outlets are reporting extensively on COVID-19 preprints (Fig 6C and 6D) represents a marked change in journalistic practice: Pre-pandemic, bioRxiv preprints received very little coverage in comparison to journal articles[25]. This cultural shift provides an unprecedented opportunity to bridge the scientific and media communities to create a consensus on the reporting of preprints[21,39]. Another marked change was observed in the use of preprints in policy documents (Fig 6G). Preprints were remarkably underrepresented in non-COVID-19 policy documents yet present, albeit at relatively low levels, in COVID-19 policy documents. In a larger dataset, two of the top 10 journals which are being cited in policy documents were found to be preprint servers (medRxiv and SSRN in fifth and eighth position, respectively)[40]. This suggests that preprints are being used to directly influence policymakers and decision-making. We only investigated a limited set of policy documents, largely restricted to Europe; whether this extends more globally remains to be explored[41]. In the near future, we aim to examine the use of preprints in policy in more detail to address these questions.

As most COVID-19-preprints were not yet published, concerns regarding quality will persist[20]. This is partially addressed by prominent scientists using social media platforms such as Twitter to publicly share concerns about poor-quality COVID-19 preprints or to amplify high-quality preprints[42]. The use of Twitter to "peer-review" preprints provides additional public scrutiny of manuscripts that can complement the more opaque and slower traditional peer-review process. In addition to Twitter, the comments section of preprint servers can be used as a public forum for discussion and review. However, an analysis of all bioRxiv comments from September 2019 found a very limited number of peer-review style comments[43]. Despite increased publicity for established preprint review services (such as PREreview[44,45]), there has been a limited use of these platforms[46]. However, independent preprint review projects have arisen whereby reviews are posted in the comments section of preprint servers or hosted on independent websites[47,48]. These more formal projects partly account for the increased commenting on the most high-profile COVID-19 preprints (Fig 4). Although these new review platforms partially combat poor-quality preprints, it is clear that there is a dire need to better understand the general quality and trustworthiness of preprints compared to peer-reviewed articles. Recent studies have suggested that the quality of reporting in preprints differs little from their later peer-reviewed articles[49], and we ourselves are currently undertaking a more detailed analysis. However, the problem of poor-quality science is not unique to preprints and ultimately, a multipronged approach is required to solve some of these issues. For example, scientists must engage more responsibly with journalists and the public, in addition to upholding high standards when sharing research. More significant consequences for academic misconduct and the swift removal of problematic articles will be essential in aiding this. Moreover, the politicisation of public health research has become a polarising issue, and more must be done to combat this; scientific advice should be objective and supported by robust evidence. Media outlets and politicians should not use falsehoods or poor-quality science to further a personal agenda. Thirdly, transparency within the scientific process is essential in improving the understanding of its internal dynamics and providing accountability.

Our data demonstrate the indispensable role that preprints, and preprint servers, are playing during a global pandemic. By communicating science through preprints, we are sharing research at a faster rate and with greater transparency than allowed by the current journal

infrastructure. Furthermore, we provide evidence for important future discussions around scientific publishing and the use of preprint servers.

Methods

Preprint metadata for bioRxiv and medRxiv

We retrieved basic preprint metadata (DOIs, titles, abstracts, author names, corresponding author name and institution, dates, versions, licenses, categories, and published article links) for bioRxiv and medRxiv preprints via the bioRxiv Application Programming Interface (API; https://api.biorxiv.org). The API accepts a "server" parameter to enable retrieval of records for both bioRxiv and medRxiv. We initially collected metadata for all preprints posted from the time of the server's launch, corresponding to November 2013 for bioRxiv and June 2019 for medRxiv, until the end of our analysis period on October 31, 2020 (*N* = 114,214). Preprint metadata, and metadata related to their linked published articles, were collected in the first week of December 2020. Note that where multiple preprint versions existed, we included only the earliest version and recorded the total number of following revisions. Preprints were classified as "COVID-19 preprints" or "non-COVID-19 preprints" on the basis of the following terms contained within their titles or abstracts (case insensitive): "coronavirus," "covid-19," "sars-cov," "ncov-2019," "2019-ncov," "hcov-19," "sars-2." For comparison of preprint behaviour between the COVID-19 outbreak and previous viral epidemics, namely Western Africa Ebola virus and Zika virus (S1 Fig), the same procedure was applied using the keywords "ebola" or "zebov" and "zika" or "zikv," respectively.

For a subset of preprints posted between September 1, 2019 and April 30, 2020 (*N* = 25,883), we enhanced the basic preprint metadata with data from a number of other sources, as outlined below. Note that this time period was chosen to encapsulate a 10-month analysis period from January 1 to October 31, 2020, in which we make comparative analysis between COVID-19 and non-COVID-19–related preprints, (*N* = 44,503), as well as the preceding year from January 1 to December 31, 2019 (*N* = 30,094), to use as a pre-COVID-19 control group. Of the preprints contained in the 10-month analysis period, 10,232 (23.0%) contained COVID-19–related keywords in their titles or abstracts.

For all preprints contained in the subset, disambiguated author affiliation and country data for corresponding authors were retrieved by querying raw affiliation strings against the Research Organisation Registry (ROR) API (https://github.com/ror-community/ror-api). The API provides a service for matching affiliation strings against institutions contained in the registry, on the basis of multiple matching types (named "phrase," "common terms," "fuzzy," "heuristics," and "acronyms"). The service returns a list of potential matched institutions and their country, as well as the matching type used, a confidence score with values between 0 and 1, and a binary "chosen" indicator relating to the most confidently matched institution. A small number (approximately 500) of raw affiliation strings returned from the bioRxiv API were truncated at 160 characters; for these records, we conducted web scraping using the rvest package for R[50] to retrieve the full affiliation strings of corresponding authors from the bioRxiv public web pages, prior to matching. For the purposes of our study, we aimed for higher precision than recall, and thus only included matched institutions where the API returned a confidence score of 1. A manual check of a sample of returned results also suggested higher precision for results returned using the "phrase" matching type, and thus we only retained results using this matching type. In a final step, we applied manual corrections to the country information for a small subset of records where false positives would be most likely to influence our results by (a) iteratively examining the chronologically first preprint associated with each country following affiliation matching and applying manual rules to correct mismatched institutions until no further errors were detected (*n* = 8 institutions); and (b) examining the top 50 most common raw affiliation strings and applying manual rules to correct any mismatched or unmatched institutions (*n* = 2 institutions). In total, we matched 54,289 preprints to a country (72.8%); for COVID-19 preprints alone, 6,692 preprints (65.4%) were matched to a country.

Note that a similar, albeit more sophisticated method of matching bioRxiv affiliation information with the ROR API service was recently documented by Abdill and colleagues[51].

Word counts and reference counts for each preprint were also added to the basic preprint metadata via scraping of the bioRxiv public web pages (medRxiv currently does not display full HTML texts, and so calculating word and reference counts was limited to bioRxiv preprints). Web scraping was conducted using the rvest package for R[50]. Word counts refer to words contained only in the main body text, after removing the abstract, figure captions, table captions, acknowledgements, and references. In a small number of cases, word counts could not be retrieved because no full text existed; this occurs as we targeted only the first version of a preprint, but in cases where a second version was uploaded very shortly (i.e., within a few days) after the first version, the full-text article was generated only for the second version. Word and reference counts were retrieved for 61,397 of 61,866 bioRxiv preprints (99.2%); for COVID-19 preprints alone, word and reference counts were retrieved for 2,314 of 2,333 preprints (99.2%). Word counts ranged from 408 to 49,064 words, while reference counts ranged from 1 to 566 references.

Our basic preprint metadata retrieved from the bioRxiv API also contained DOI links to published versions (i.e., a peer-reviewed journal article) of preprints, where available. In total, 22,151 records in our preprint subset (29.7%) contained links to published articles, although of COVID-19 preprints, only 2,164 preprints contained such links (21.1%). It should be noted that COVID-19 articles are heavily weighted towards the most recent months of the dataset and have thus had less time to progress through the journal publication process. Links to published articles are likely an underestimate of the total proportion of articles that have been subsequently published in journals—both as a result of the delay between articles being published in a journal and being detected by bioRxiv and bioRxiv missing some links to published articles when, e.g., titles change significantly between the preprint and published version[25]. Published article metadata (titles, abstracts, publication dates, journal, and publisher name) were retrieved by querying each DOI against the Crossref API (https://api.crossref.org), using the rcrossref package for R[52]. With respect to publication dates, we use the Crossref "created" field which represent the date on which metadata was first deposited and has been suggested as a good proxy of the first online availability of an article[53,54]. When calculating delay from preprint posting to publication dates, erroneous negative values (i.e., preprints posted after published versions) were ignored. We also retrieved data regarding the open access status of each article by querying each DOI against the Unpaywall API (https://unpaywall.org/products/api) via the roadoi package for R [55].

Usage, altmetrics, and citation data

For investigating the rates at which preprints are used, shared, and cited, we collected detailed usage, altmetrics, and citation data for all bioRxiv and medRxiv preprints posted between January 1, 2019 and October 31, 2020 (i.e., for every preprint where we collected detailed metadata, as described in the previous section). All usage, altmetrics, and citation data were collected in the first week of December 2020.

Usage data (abstract views and PDF downloads) were scraped from the bioRxiv and medRxiv public web pages using the rvest package for R [50]. bioRxiv and medRxiv web pages display abstract views and PDF downloads on a calendar month basis; for subsequent analysis (e.g., Fig 4), these were summed to generate total abstract views and downloads since the time of preprint posting. In total, usage data were recorded for 74,461 preprints (99.8%)—a small number were not recorded, possibly due to server issues during the web scraping process. Note that bioRxiv web pages also display counts of full-text views, although we did not include these data in our final analysis. This was partially to ensure consistency with medRxiv, which currently does not provide display full HTML texts, and partially due to ambiguities in the timeline of full-text publishing—the full text of a preprint is added several days after the preprint is first available, but the exact delay appears to vary from preprint to preprint. We also compared rates of PDF downloads for bioRxiv and medRxiv preprints with other preprint

servers (SSRN and Research Square) (S3C Fig)—these data were provided directly by representatives of each of the respective preprint servers.

Counts of multiple altmetric indicators (mentions in tweets, blogs, and news articles) were retrieved via Altmetric (https://www.altmetric.com), a service that monitors and aggregates mentions to scientific articles on various online platforms. Altmetric provide a free API (https://api.altmetric.com) against which we queried each preprint DOI in our analysis set. Importantly, Altmetric only contains records where an article has been mentioned in at least one of the sources tracked; thus, if our query returned an invalid response, we recorded counts for all indicators as 0. Coverage of each indicator (i.e., the proportion of preprints receiving at least a single mention in a particular source) for preprints were 99.3%, 10.3%, 7.4%, and 0.33 for mentions in tweets, blogs news, and Wikipedia articles, respectively. The high coverage on Twitter is likely driven, at least in part, by automated tweeting of preprints by the official bioRxiv and medRxiv Twitter accounts. For COVID-19 preprints, coverage was found to be 99.99%, 14.3%, 28.7%, and 0.76% for mentions in tweets, blogs, news, and Wikipedia articles, respectively.

To quantitatively capture how high-usage preprints were being received by Twitter users, we retrieved all tweets linking to the top 10 most-tweeted preprints. Tweet IDs were retrieved via the Altmetric API service and then queried against the Twitter API using the rtweet package[56] for R, to retrieve full tweet content.

Citations counts for each preprint were retrieved from the scholarly indexing database Dimensions (https://dimensions.ai). An advantage of using Dimensions in comparison to more traditional citation databases (e.g., Scopus, Web of Science) is that Dimensions also includes preprints from several sources within their database (including from bioRxiv and medRxiv), as well as their respective citation counts. When a preprint was not found, we recorded its citation counts as 0. Of all preprints, 13,298 (29.9%) recorded at least a single citation in Dimensions. For COVID-19 preprints, 5,294 preprints (57.9%) recorded at least a single citation.

Comments

bioRxiv and medRxiv html pages feature a Disqus (https://disqus.com) comment platform to allow readers to post text comments. Comment counts for each bioRxiv and medRxiv preprint were retrieved via the Disqus API service (https://disqus.com/api/docs/). Where multiple preprint versions existed, comments were aggregated over all versions. Text content of comments for COVID-19 preprints were provided directly by the bioRxiv development team.

Screening time for bioRxiv and medRxiv

To calculate screening time, we followed the method outlined by Steve Royle[57]. In short, we calculate the screening time as the difference in days between the preprint posting date and the date stamp of submission approval contained within bioRxiv and medRxiv DOIs (only available for preprints posted after December 11, 2019). bioRxiv and medRxiv preprints were filtered to preprints posted between January 1 and October 31, 2020, accounting for the first version of a posted preprint.

Policy documents

To describe the level of reliance upon preprints in policy documents, a set of policy documents were manually collected from the following institutional sources: the ECDC (including rapid reviews and technical reports), UK POST, and WHO SB (*n* = 81 COVID-19–related policies, *n* = 38 non-COVID-19–related policies). COVID-19 policy documents were selected from January 1, 2020 to October 31, 2020. Due to the limited number of non-COVID-19 policy documents from the same time period, these documents were selected dating back to September 2018. Reference lists of each policy document were then text mined and manually verified to calculate the proportion of references that were preprints.

Journal article data

To compare posting rates of COVID-19 preprints against publication rates of articles published in scientific journals (Fig 1B), we extracted a dataset of COVID-19 journal articles from Dimensions (https://www.dimensions.ai), via the Dimensions Analytics API service. Journal articles were extracted based on presence of the following terms (case insensitive) in their titles or abstracts: "coronavirus," "covid-19," "sars-cov," "ncov-2019," "2019-ncov," "hcov-19," and "sars-2." Data were extracted in the first week of December 2020 and covered the period January 1, 2020 to October 31, 2020. To ensure consistency of publication dates with our dataset of preprints, journal articles extracted from Dimensions were matched with records in Crossref on the basis of their DOIs (via the Crossref API using the rcrossref package for R[52]), and the Crossref "created" field was used as the publication date. The open access status of each article (S1B Fig) was subsequently determined by querying each DOI against the Unpaywall API via the roadoi package for R [55].

Statistical analyses

Preprint counts were compared across categories (e.g., COVID-19 or non-COVID-19) using chi-squared tests. Quantitative preprint metrics (e.g., word count and comment count) were compared across categories using Mann–Whitney tests and correlated with other quantitative metrics using Spearman rank tests for univariate comparisons.

For time-variant metrics (e.g., views, downloads, which may be expected to vary with length of preprint availability), we analysed the difference between COVID-19 and non-COVID-19 preprints using generalised linear regression models with calendar days since January 1, 2020 as an additional covariate and negative binomially distributed errors. This allowed estimates of time-adjusted rate ratios comparing COVID-19 and non-COVID-19 preprint metrics. Negative binomial regressions were constructed using the function "glm.nb" in R package MASS[58]. For multivariate categorical comparisons of preprint metrics (e.g., screening time between preprint type and preprint server or publication delay between preprint type and publisher for top 10 publishers), we constructed 2-way factorial ANOVAs, testing for interactions between both category variables in all cases. Pairwise post hoc comparisons of interest were tested using Tukey HSD while correcting for multiple testing, using function "glht" while setting multiple comparisons to "Tukey" in R package multcomp [53].

Parameters and limitations of this study

We acknowledge a number of limitations in our study. Firstly, to assign a preprint as COVID-19 or not, we used keyword matching to titles/abstracts on the preprint version at the time of our data extraction. This means we may have captured some early preprints, posted before the pandemic that had been subtly revised to include a keyword relating to COVID-19. Our data collection period was a tightly defined window (January to October 2020) which may impact upon the altmetric and usage data we collected as those preprints posted at the end of October would have had less time to accrue these metrics.

Supporting information

S1 Fig. Preprints represent a higher proportion of the pandemic-related literature for COVID-19 than previous pandemics, and most articles are open access.

(A) Total number of preprints posted on bioRxiv and medRxiv during multiple epidemics: Western Africa Ebola virus, Zika virus, and COVID-19. The number of preprints posted that were related to the epidemic and the number that were posted but not related to the epidemic in the same time period are shown. Periods of data collection for Western Africa Ebola virus (January 24, 2014 to June 9, 2016) and Zika virus (March 2, 2015 to November 18, 2016) correspond to the periods between the first official medical report and WHO end of Public Health Emergency of International Concern declaration. The period of data collection for COVID-19 refers to the analysis period used in this study, January 1, 2020 to October 31, 2020. (B) Comparison of COVID-19 journal article accessibility (open versus closed access) according to data provided by Unpaywall (https://unpaywall.org). The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019; WHO, World Health Organization. https://doi.org/10.1371/journal.pbio.3000959.s001
(TIFF)

S2 Fig. Properties of COVID-19 and non-COVID-19 preprints categorised by preprint server.
(A) Number of new preprints posted to bioRxiv versus medRxiv per month. (B) Preprint screening time in days for bioRxiv versus medRxiv. (C) Number of preprint versions posted to bioRxiv versus medRxiv. (D) License type chosen by authors for bioRxiv versus medRxiv. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019.
https://doi.org/10.1371/journal.pbio.3000959.s002
(TIFF)

S3 Fig. Additional access statistics for bioRxiv and medRxiv preprints.
(A) Boxplots of abstracts views received by COVID-19 and non-COVID-19 preprints in the same calendar month in which they were posted, binned by preprint posting month. (B) Boxplots of PDF downloads received by COVID-19 and non-COVID-19 preprints in the same calendar month in which they were posted, binned by preprint posting month. (C) Boxplots of total abstract views for non-COVID preprints between January 2019 and October 2020, binned by preprint posting month (D) Boxplots of total PDF downloads for for non-COVID preprints between January 2019 and October 2020, binned by preprint posting month. (E) Comparison of PDF downloads for COVID-19 and non-COVID-19 preprints across multiple preprint servers. Red shaded areas in (C) and (D) represent our analysis time period, concurrent with the COVID-19 pandemic. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019.
https://doi.org/10.1371/journal.pbio.3000959.s003
(TIFF)

S4 Fig. Additional COVID-19 preprint usage data.
(A) Wordcloud of hashtags for the 100 most tweeted COVID-19 preprints. The size of the word reflects the hashtag frequency (larger = more frequent). Only hashtags used in at least 5 original tweets (excluding retweets) were included. Some common terms relating directly to COVID-19 were removed for visualisation ("covid19," "coronavirus," "ncov2019," "covid," "covid2019," "sarscov2," "2019ncov," "hcov19," "19," "novelcoronavirus," "corona," "coronaovirus," "coronarovirus," and "coronarvirus"). (B) Euler diagram showing overlap between the 10 most tweeted COVID-19 preprints, the 10 most covered COVID-19 preprints in the news, the 10 most blogged about preprints, the 10 most commented-upon preprints, and the 10 most cited COVID-19 preprints. The data underlying this figure may be found in https://github.com/preprinting-a-pandemic/pandemic_preprints and https://zenodo.org/record/4587214#.YEN22Hmnx9A. COVID-19, Coronavirus Disease 2019.
https://doi.org/10.1371/journal.pbio.3000959.s004
(TIFF)

S1 Table. Descriptive statistics for COVID-19 and non-COVID-19 preprints broken down by server.
COVID-19, Coronavirus Disease 2019.
https://doi.org/10.1371/journal.pbio.3000959.s005
(XLSX)

S2 Table. Outputs from mixed-effects regression predicting word count using all bioRxiv preprints.
https://doi.org/10.1371/journal.pbio.3000959.s006
(XLSX)

S3 Table. Outputs from mixed-effects regression predicting word count using only published bioRxiv preprints.
https://doi.org/10.1371/journal.pbio.3000959.s007
(XLSX)

S4 Table. Statistics for first time or previous posting of preprints by senior authors based on country.
https://doi.org/10.1371/journal.pbio.3000959.s008
(XLSX)

S1 Model. Mixed-effects regression models to investigate alternative factors on length of preprints.
https://doi.org/10.1371/journal.pbio.3000959.s009
(DOCX)

Acknowledgments

The authors would like to thank Ted Roeder, John Inglis, and Richard Sever from bioRxiv and medRxiv for providing information relating to comments on Coronavirus Disease 2019 (COVID-19) preprints. We would also like to thank Martyn Rittman (preprints.org), Shirley Decker-Lucke (SSRN), and Michele Avissar-Whiting (Research Square) for kindly providing usage data. Further thanks to Helena Brown and Sarah Bunn for conversations regarding media usage and government policy.

References

1.WHO. COVID-19 Weekly Epidemiological Update-11. 2020 Oct. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/weekly-epi-update-11.pdf.

2.Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020;382:727–33. pmid:31978945
View Article PubMed/NCBI Google Scholar

3.WHO. Coronavirus Disease (COVID-19) Weekly Epidemiological Update—24. 2021 Jan. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20210127_weekly_epi_update_24.pdf

4.Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020;5:536–44. pmid:32123347
View Article PubMed/NCBI Google Scholar

5.Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. In: http://dx.doi.org/10.1056/NEJMoa030781 [Internet]. Massachusetts Medical Society; 7 Oct 2009 [cited 13 May 2020]. https://doi.org/10.1056/NEJMoa030781 pmid:12690092

6.Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. N Engl J Med. 2012;367:1814–20. pmid:23075143
View Article PubMed/NCBI Google Scholar

7.Wallach JD, Egilman AC, Gopal AD, Swami N, Krumholz HM, Ross JS. Biomedical journal speed and efficiency: a cross-sectional pilot survey of author experiences. Res Integr Peer Rev. 2018;3:1. pmid:29451557
View Article PubMed/NCBI Google Scholar

8.Abdill RJ, Blekhman R. Tracking the popularity and outcomes of all bioRxiv preprints. Pewsey E, Rodgers P, Greene CS. Elife. 2019;8:e45133. pmid:31017570
View Article PubMed/NCBI Google Scholar

9.Björk B-C, Solomon D. The publishing delay in scholarly peer-reviewed journals. J Informet. 2013;7:914–23.
View Article Google Scholar

10. Sever R, Roeder T, Hindle S, Sussman L, Black K-J, Argentine J, et al. bioRxiv: the preprint server for biology. bioRxiv. 2019;833400.
View Article Google Scholar

11. Kaiser J, 2014, Am 12:00. BioRxiv at 1 year: A promising start. In: Science | AAAS [Internet]. 11 Nov 2014 [cited 13 May 2020]. Available from: https://www.sciencemag.org/news/2014/11/biorxiv-1-year-promising-start

12. Rawlinson C, Bloom T. New preprint server for medical research. BMJ. 2019;365. pmid:31167753
View Article PubMed/NCBI Google Scholar

13. Wellcome Trust. Sharing data during Zika and other global health emergencies | Wellcome. In: Wellcomeacuk [Internet]. 10 Feb 2016 [cited 13 May 2020]. Available from: https://wellcome.ac.uk/news/sharing-data-during-zika-and-other-global-health-emergencies

14. Johansson MA, Reich NG, Meyers LA, Preprints LM. An underutilized mechanism to accelerate outbreak science. PLoS Med. 2018;15:e1002549. pmid:29614073
View Article PubMed/NCBI Google Scholar

15. Fraser N, Kramer B. covid19_preprints. 2020.
View Article Google Scholar

16. Chiarelli A, Johnson R, Pinfield S, Preprints RE. Scholarly Communication: An Exploratory Qualitative Study of Adoption, Practices, Drivers and Barriers. F1000Res. 2019;8:971. pmid:32055396
View Article PubMed/NCBI Google Scholar

17. Himmelstein D. The licensing of bioRxiv preprints. Satoshi Village. 2016 [cited 19 May 2020]. Available from: https://blog.dhimmel.com/biorxiv-licenses/
View Article Google Scholar

18.ASAPbio. asapbio/licensing. ASAPbio; 2018. Available from: https://github.com/asapbio/licensing

19. Gog JR. How you can help with COVID-19 modelling. Nat Rev Phys. 2020; 1–2.
View Article Google Scholar

20. Bagdasarian N, Cross GB, Fisher D. Rapid publications risk the integrity of science in the era of COVID-19. BMC Med. 2020;18:192. pmid:32586327
View Article PubMed/NCBI Google Scholar

21. Sheldon T. Preprints could promote confusion and distortion. Nature. 2018;559:445–6. pmid:30042547
View Article PubMed/NCBI Google Scholar

22. Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. medRxiv. 2020; 2020.04.14.20062463.
View Article Google Scholar

23. Pradhan P, Pandey AK, Mishra A, Gupta P, Tripathi PK, Menon MB, et al. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag. bioRxiv. 2020; 2020.01.30.927871. pmid:32511314
View Article PubMed/NCBI Google Scholar

24. Besançon L, Peiffer-Smadja N, Segalas C, Jiang H, Masuzzo P, Smout C, et al. Open Science Saves Lives: Lessons from the COVID-19 Pandemic. bioRxiv. 2020; 2020.08.13.249847.
View Article Google Scholar

25. Fraser N, Momeni F, Mayr P, Peters I. The relationship between bioRxiv preprints, citations and altmetrics. Quant Sci Stud. 2020:1–21.
View Article Google Scholar

26. Wellcome Trust. Coronavirus (COVID-19): sharing research data | Wellcome. 31 Jan 2020 [cited 21 May 2020]. Available from: https://wellcome.ac.uk/coronavirus-covid-19/open-data

27. Wellcome Trust. Publishers make coronavirus (COVID-19) content freely available and reusable | Wellcome. 16 Mar 2020 [cited 21 May 2020]. Available from: https://wellcome.ac.uk/press-release/publishers-make-coronavirus-covid-19-content-freely-available-and-reusable

28. Ioannidis JPA, Salholz-Hillel M, Boyack KW, Baas J. The rapid, massive infection of the scientific literature and authors by COVID-19. bioRxiv. 2020; 2020.12.15.422900.
View Article Google Scholar

29. Horbach SPJM. Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. Quant Sci Stud. 2020;1:1056–67.
View Article Google Scholar

30. Penfold NC, Polka JK. Technical and social issues influencing the adoption of preprints in the life sciences. PLoS Genet. 2020;16:e1008565. pmid:32310942
View Article PubMed/NCBI Google Scholar

31.ASAPbio. Preprint authors optimistic about benefits: preliminary results from the #bioPreprints2020 survey. In: ASAPbio [Internet]. 27 Jul 2020 [cited 1 Feb 2021]. Available from: https://asapbio.org/biopreprints2020-survey-initial-results

32. Inglis J. We've just put an additional, cautionary note about the use of preprints on every @biorxivpreprint https://t.co/08eSXL4dDi. In: Twitter [Internet]. 1 Feb 2020 [cited 22 May 2020]. Available from: https://twitter.com/johnringlis/status/1223598414493077505

33. Fang Z, Costas R. Tracking the Twitter attention around the research efforts on the COVID-19 pandemic. ArXiv200605783 Cs. 2020 [cited 16 Sep 2020]. Available from: http://arxiv.org/abs/2006.05783

34. Carlson J, Harris K. Quantifying and contextualizing the impact of bioRxiv preprints through automated social media audience segmentation. PLoS Biol. 2020;18:e3000860. pmid:32960891
View Article PubMed/NCBI Google Scholar

35. Yaqub U. Tweeting During the Covid-19 Pandemic: Sentiment Analysis of Twitter Messages by President Trump. Digit Gov Res Pract. 2021;2:1–7.
View Article Google Scholar

36. Gruzd A, Mai P. Going viral: How a single tweet spawned a COVID-19 conspiracy theory on Twitter. Big Data Soc. 2020;7:2053951720938405.
View Article Google Scholar

37. Anti-Defamation League. At the Extremes: The 2020 Election and American Extremism | Part 3. In: At the Extremes: The 2020 Election and American Extremism | Part 3 [Internet]. 10 Aug 2020 [cited 27 Jan 2021]. Available from: https://www.adl.org/blog/at-the-extremes-the-2020-election-and-american-extremism-part-3

38. Lally C, Christie L. COVID-19 misinformation UK Parliam POST. 2020 [cited 21 May 2020]. Available from: https://post.parliament.uk/analysis/covid-19-misinformation/, https://post.parliament.uk/analysis/covid-19-misinformation/

39. Fleerackers A, Riedlinger M, Moorhead L, Ahmed R, Alperin JP. Communicating Scientific Uncertainty in an Age of COVID-19: An Investigation into the Use of Preprints by Digital Media Outlets. Health Commun. 2021;0:1–13. pmid:33390033
View Article PubMed/NCBI Google Scholar

40. Adie E. COVID-19-policy dataset. 2020.
View Article Google Scholar

41. Yin Y, Gao J, Jones BF, Wang D. Coevolution of policy and science during the pandemic. Science. 2021;371:128–30. pmid:33414211
View Article PubMed/NCBI Google Scholar

42. Markus A, Oransky I, Retraction Watch. Eye for Manipulation: A Profile of Elisabeth Bik. In: The Scientist Magazine® [Internet]. 7 May 2019 [cited 21 May 2020]. Available from: https://www.the-scientist.com/news-opinion/eye-for-manipulation—a-profile-of-elisabeth-bik-65839

43. Malički M, Costello J, Alperin JP, Maggio LA. From amazing work to I beg to differ—analysis of bioRxiv preprints that received one public comment till September 2019. bioRxiv. 2020; 2020.10.14.340083.
View Article Google Scholar

44.OASPA. COVID-19 Publishers Open Letter of Intent—Rapid Review. In: OASPA [Internet]. 27 May 2020 [cited 13 May 2020]. Available from: https://oaspa.org/covid-19-publishers-open-letter-of-intent-rapid-review/

45. Johansson MA, Saderi D. Open peer-review platform for COVID-19 preprints. Nature. 2020;579:29–9. pmid:32127711
View Article PubMed/NCBI Google Scholar

46. Brierley L. The role of research preprints in the academic response to the COVID-19 epidemic. 2020.
View Article Google Scholar

47. Vabret N, Samstein R, Fernandez N, Merad M. Advancing scientific knowledge in times of pandemics. Nat Rev Immunol. 2020:1–1. pmid:31792373
View Article PubMed/NCBI Google Scholar

48.MIT Press. The MIT Press and UC Berkeley launch Rapid Reviews: COVID-19. In: MIT News | Massachusetts Institute of Technology [Internet]. 29 Jun 2020 [cited 13 Sep 2020]. Available from: https://news.mit.edu/2020/mit-press-and-uc-berkeley-launch-rapid-reviews-covid-19-0629

49. Carneiro CFD, Queiroz VGS, Moulin TC, Carvalho CAM, Haas CB, Rayêe D, et al. Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. Res Integr Peer Rev. 2020;5:16. pmid:33292815
View Article PubMed/NCBI Google Scholar

50. Wickham H, RStudio. rvest: Easily Harvest (Scrape) Web Pages. 2019. Available from: https://CRAN.R-project.org/package=rvest

51. Abdill RJ, Adamowicz EM, Blekhman R. International authorship and collaboration in bioRxiv preprints. bioRxiv. 2020; 2020.04.25.060756. pmid:32716295
View Article PubMed/NCBI Google Scholar

52. Chamberlain S, Zhu H, Jahn N, Boettiger C, Ram K. rcrossref: Client for Various "CrossRef" "APIs" 2020. Available from: https://CRAN.R-project.org/package=rcrossref

53. Fang Z, Costas R. Studying the accumulation velocity of altmetric data tracked by Altmetric.com. Scientometrics. 2020;123: 1077–1101.
View Article Google Scholar

54. Haustein S, Bowman TD, Costas R. When is an article actually published? An analysis of online availability, publication, and indexation dates. ArXiv150500796 Cs. 2015 [cited 22 Jan 2021]. Available from: http://arxiv.org/abs/1505.00796

55. Jahn N, rOpenSci TS roadoi: Find Free Versions of Scholarly Publications via Unpaywall. 2019. Available from: https://CRAN.R-project.org/package=roadoi

56. Kearney M. rtweet: Collecting and analyzing Twitter data. J Open Source Softw. 2019;4:1829.
View Article Google Scholar

57. Royle Steve. Screenager: screening times at bioRxiv. In: quantixed [Internet]. 30 Mar 2020 [cited 22 May 2020]. Available from: https://quantixed.org/2020/03/30/screenager-screening-times-at-biorxiv/

58. Venables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer-Verlag; 2002. https://doi.org/10.1007/978-0-387-21706-2

PsycEXTRA®

EXPAND BEHAVIORAL SCIENCE RESEARCH WITH THE PREMIER RESOURCE FOR GRAY LITERATURE

Devoted to curating and indexing hard-to-find content from authoritative sources, PsycEXTRA® allows researchers to go beyond traditional peer-reviewed research materials. This high-quality and relevant database combines bibliographic records with unique full-text materials, focusing on the latest conference presentations and papers, newsletters, reports, patient-oriented factsheets and brochures, magazines, monographs, and standards and guidelines relevant to the needs of students, faculty, and clinicians alike.




Explore original, cutting-edge, high-quality research

- Updated biweekly
- More than 300,000 records and growing
- No coverage overlap with PsycINFO®, creating an ideal companion database
- Full text for more than 70% of records
- Ongoing updates of select professional literature from multiple state and government entities, associations, foundations, and more
- Average of 25,000 records or more added each year

Available via EBSCO

Grey literature is a necessary facet in a critical approach to gambling research*

David G. Baxter, Fiona Nicoll, and **Murat Akçayir,**
Deptartment of Political Science, University of Alberta, Canada

Abstract:
Commercial gambling has seen massive global expansion in the past 25 years. It is a huge industry selling a risky form of entertainment: problem gambling is the only non-substance addiction recognized in the DSM-5, affecting an average of 2.3% of people in jurisdictions where prevalence data are available. Gambling also harms people who gamble below the clinical threshold of "problem gambling", as well as the friends, families and communities of people who gamble. Gambling harm is disproportionally felt by racialized peoples and people of lower socioeconomic status. As such, researchers and governments are increasingly viewing gambling as a public health issue.

Gambling research is published in both the primary and grey literature, and the integrity of gambling research is a topic of increasingly heated debate. Bibliometric reviews have found that gambling research is heavily focused on the psychological and biological characteristics of people with problem gambling, with less emphasis on the gambling products themselves and how they are provided. While the gambling grey literature is recognized as valuable by the gambling research community, it has not yet been systematically assessed.

In this paper we present the grey literature analysis portion of a pilot project to use a big data approach to produce a mapping review of gambling research from five nations: Australia, Canada, New Zealand, United Kingdom, and United States. For primary research publications on gambling, we performed systematized searches on the Scopus and Web of Science databases. For gambling grey literature, we retrieved all grey literature documents in the GREO International Gambling Research Evidence Centre. For the period of 2014-2018, the grey literature search yielded 360 reports, compared to 1292 articles in the primary literature search. The proportion of grey literature greatly varied by country, ranging from <10% in USA to nearly 50% in New Zealand. Content analysis revealed that the problems investigated in gambling grey literature are very different from the published literature: the top problems in the published literature were young gamblers and online gambling, whereas the top problems in the grey literature were the prevalence of problem gambling and the health and well-being of the public. This demonstrates that the grey literature is a vital piece of the puzzle to understanding this public health issue.

Background

Gambling as a public health issue

Commercial gambling is expanding in many countries, and is increasingly viewed as a public health issue (Korn & Shaffer, 1999; Wardle, Reith, Langham, & Rogers, 2019). Although global gambling revenues are difficult to accurately measure, a 2015 report estimated the global gambling market to be worth US$423 billion, and would grow to US$635 billion by 2022 (Morgan Stanley 2015, as cited in Cassidy, 2020). Commercial gambling has been called a “dangerous commodity” similar to tobacco or alcohol, and presents addiction issues and other social harms that are disproportionately borne by vulnerable populations (Markham & Young, 2015). Currently, problem gambling is the only non-substance addiction recognized by the American Psychiatric Association’s classification of mental disorders, and one of only two recognized by the WHO’s International Classification of Diseases (American Psychiatric

* First published in the GL2020 Conference Proceedings, February 2021. https://doi.org/10.26069/greynet-2021-000.469-gg

Association, 2013; World Health Organization, 2018). Problem gambling prevalence rates average 2.3% worldwide in countries where data are available (Williams, Volberg, & Stevens, 2012). Problem gambling is more prevalent in poor and racialized communities (Abbott, Volberg, Bellringer, & Reith, 2004), and in North America is most prevalent in Indigenous populations (Williams, Stevens, & Nixon, 2011).

Many others are harmed by gambling besides those with problem gambling: on average, one person's problem gambling negatively impacts six others who are close to them (Goodwin, Browne, Rockloff, & Rose, 2017), and there are many more people who gamble and experience a low or moderate amount of gambling-related harm, but do not reach the threshold for a clinical diagnosis of problem gambling (Browne et al., 2016).

Why a critical approach to gambling research?

When governments legalize or expand gambling, they often simultaneously create gambling research and treatment programmes which are funded from a levy based on a percentage of gambling revenues (GREO, 2020). Although there are multiple potential conflicts of interest with doing research funded by directly by the gambling industry (Kim, Dobson, & Hodgins, 2016), there are also concerns when research is funded indirectly from gambling revenues via a levy. In this model, the research funding body and the researchers themselves both necessarily benefit from gambling revenues growing/remaining high (Adams & Rossen, 2012), which may discourage research that challenges the status quo of gambling operation (Livingstone et al., 2018).

There have been a small number of empirical investigations into these issues. The 2013 "Fair Game" report interviewed 109 gambling researcher stakeholders and found that when gambling research funds were mediated through specialized agencies the researchers were pressured to research politically "safe" topics such as problem gambling instead of broader issues about how gambling is provided (Cassidy, Loussouarn, & Pisac, 2013). This narrow focus on psychological and psychiatric perspectives on individual people with problem gambling was replicated in two bibliometric mapping reviews of thousands of gambling research articles, (Akçayir, Nicoll, Baxter, & Palmer, Forthcoming 2021; Baxter, Hilbrecht, & Wheaton, 2019). Only two quantitative empirical reviews of gambling research have specifically investigated the effect of gambling industry funding on research bias. These studies found no significant differences between industry and non-industry funded research, but these studies were themselves directly funded by the gambling industry and their findings should be treated with caution (Ladouceur, Shaffer, Blaszczynski, & Shaffer, 2019; Shaffer et al., 2019).

Gambling's grey literature

The above empirical reviews all share a significant limitation: they only review research published in journal articles and exclude the significant body of gambling research published as grey literature. Many of the aforementioned government-funded gambling research programmes primarily publish research reports as grey literature. For example, the Ontario Problem Gambling Research Centre (OPGRC), funded by the Government of Ontario, operated from 2000 to 2013 and was the largest single funder of gambling research at the time, awarding an average of CAD$2.23 million per year; some OPGRC studies were published as academic journal articles but all research was also published as peer-reviewed grey reports (OPGRC, 2013). Many other types of organizations publish gambling research as grey literature, including non-governmental organizations, think tanks, and gambling industry trade organizations (GREO, n.d.).

Despite the wealth of grey literature evidence on gambling topics, it has not been considered in debates about the conduct on gambling research, and it is often not included in

knowledge syntheses on gambling topics. It is an established best practice to include grey literature in health research knowledge syntheses such as systematic reviews (Hopewell, McDonald, Clarke, & Egger, 2007), because grey documents often cover different topics and are more in-depth than their traditional counterparts, and can be less subject to publication bias toward statistically significant results, amongst other potential benefits (Bonato, 2018). Although guidelines have been developed for systematically searching public health grey literature (Godin, Stapleton, Kirkpatrick, Hanning, & Leatherdale, 2015), a recent high-profile systematic review on gambling interventions excluded grey literature by claiming that it is called 'grey' because it is not peer-reviewed and therefore of questionable scientific robustness and reliability (Ladouceur, Shaffer, Blaszczynski, & Shaffer, 2017). We argue that, on the contrary, grey literature can be of equal or greater scientific rigour than a peer-reviewed journal article. This is because the accepted definition of "grey literature" is concerned with document formats, publishing bodies, method of publication, and meeting a *minimum* level of quality. (Schöpfel, 2010). In contrast, peer-reviewed articles can fall short on several quality indices, as recent debates about the reliability and replicability of experiments, data collection and analyses of important and influential publications attest.(Baker, 2016; Heathers, 2020; Jackson, 2020)

In order to address this significant gap in gambling research debates, this study presents a pilot study which we believe is the first investigation of the qualities of gambling's grey literature research publications in comparison to the primary literature on gambling. We aim to map the recent primary and grey literature on gambling in five countries: Australia, Canada, New Zealand, the United Kingdom, and the United States. These five countries were selected because they share similar recent histories of gambling liberalization, and their academic and grey documents are primarily published in English. We address the following research questions:

1. What proportion of gambling research is published as grey literature in the five study countries?
2. What issues are most often investigated in gambling research? Do the issues differ between primary and grey research publications?
3. Which academic disciplines study gambling? Do disciplines differ between primary and grey research publications?

Methods

The present study employs a mapping review methodology. A mapping review is a type of knowledge synthesis that aims to describe and categorize information within a body of knowledge of known scope (Grant & Booth, 2009).

Literature Search and Inclusion/Exclusion Criteria

The primary literature searches were performed on the Scopus and Web of Science databases, which are the two largest multidisciplinary indexes of peer-reviewed journal articles. The search methods and inclusion/exclusion criteria for producing this dataset are described elsewhere (Akçayir, Nicoll, Baxter, & Palmer, 2021), but in summary a broad keyword search for "gambl*" was used to maximize recall of relevant articles. Inclusion criteria required articles to be empirical studies published in academic journals between 2014-2018, where gambling is the central topic of investigation and the first author is from one of the five target countries.

The grey literature sample was retrieved from the Gambling Research Exchange (GREO) International Gambling Research Evidence Centre. The GREO Evidence Centre's collection policy is international in scope and prioritizes grey literature (Baxter & Hilbrecht, 2020), and

it's grey literature collection is strongest in the five countries of this study. Guided by Adams et al.'s framework of evaluating grey literature tiers based on source expertise and outlet control (Adams, Smart, & Huff, 2017), our search was limited to the document type categories "White Papers" and "Reports", with the exclusion of Reports subtypes that are not research documents (i.e., Legislation, Annual Report, Policy Document). As the scope of the GREO Evidence Centre is gambling only, no keyword limits were imposed on the search results. Similar to the primary literature search, results were limited to the time period of 2014-2018, and required to be published by an organization or government from one of the five target countries. All searches were conducted in 2019 with the last search performed on 7 August 2019.

Data Coding and Analysis

All primary articles and grey reports were coded for descriptive information to ascertain year of publication and study location. For primary literature, the location was assigned to the country of the first author listed on the study. For grey literature, location was assigned based on the location of the publisher/producing body.

To identify and code the main "issues" investigated in each study, an inductive content analysis was performed, following the methodology of Akçayir and Akçayir (2017). For each study, the abstract, keywords, and background/introduction sections were read to find purpose statements and research questions identifying the issue of investigation. This information from each study was input into the qualitative data analysis program Atlas.ti 7 and coded. The coded issues were then grouped into categories based on their similarities and assigned to the most descriptive wording that was used.

For disciplinary analysis, primary literature was assigned a discipline based on the self-description of the journal in which the article is published. For grey literature, as the publications do not have a disciplinary journal, documents were assigned a discipline based on the affiliation of the first author of the document, if such information is available.

Results

The primary literature dataset included 1292 empirical journal articles, while the grey literature search resulted in 360 reports, primarily from national and subnational government sources. A summary of the distribution of primary and grey documents by country is summarized in Table 1.

Table 1: distribution and proportion of primary and grey literature publications, 2014-2018.

Country	Primary Literature	Grey Literature	% Grey Literature
Australia	325	80	19.8%
Canada	307	156	33.7%
New Zealand	32	31	49.2%
United Kingdom	198	48	19.5%
United States	430	45	9.5%
Total	**1292**	**360**	**21.8%**

The proportion of gambling research publications that were grey literature varied considerably between countries, ranging from under 10% in the United States to nearly 50% in New Zealand. This suggests that the different countries in the study may have different publication norms for gambling research, or broader research generally.

Top "Issues" of Investigation

Our content analysis revealed many issues that are studied in gambling research, but for the purposes of this analysis only the top ten issues for each literature type are presented. These results are presented in Table 2.

Table 2: The top ten issues investigated in primary and grey literature research publications on gambling, 2014-2018.

Primary Literature (N=1292)			**Grey Literature (N=360)**		
Issue	**n**	**% of Total**	**Issue**	**n**	**% of Total**
Young adult gamblers	111	8.59%	Health problem/well-being	31	8.61%
Online gambling	110	8.51%	Prevalence	30	8.33%
Slot machines/EGMs	86	6.66%	Slot machines/EGMs	24	6.67%
Children & Adolescents	60	4.64%	Responsible gambling	20	5.56%
Treatment	56	4.33%	Treatment	20	5.56%
Advertising	47	3.64%	Online gambling	17	4.72%
Characteristics	47	3.64%	Sports betting	17	4.72%
Gambling motivations	46	3.56%	Gambling harms	15	4.17%
Sports betting	45	3.48%	Harm minimization	15	4.17%
Impulsivity	43	3.33%	Assessment	12	3.33%

The content analysis reveals some similarities, as well as some important differences. In both samples we see an interest in the same specific forms of gambling (i.e., "Online gambling", "Slot machines/EGMs", and "Sports Betting") as well treatment for problem gambling.

The popular issues unique to the primary literature include interest in certain types of individuals gambling (i.e., "Young adult gamblers", "Children & Adolescents") and what about them makes them gamble (i.e., "Characteristics", "Gambling Motivations", and "Impulsivity"), as well as the effects of advertising. These results largely replicate the heavy emphasis on the psychology of individual gamblers found in previous mapping reviews of gambling research.

In the grey literature sample, we see more interest in the whole population (i.e., "Prevalence" and "Assessment") and concern for gambling being harmful (i.e., "Health problem/well-being", "Gambling harms", and "Harm minimization"). Finally, we also see the theme "Responsible gambling".

The concept of responsible gambling is based on the "Reno Model of Responsible Gambling" (Blaszczynski, Ladouceur, & Shaffer, 2004), which has guided much of gambling policy development in the five study countries since its publication in 2004. The Reno Model referred to all gambling stakeholders, including governments, gambling providers, health services, and community and consumer groups share the responsibility of creating gambling policies that minimize gambling-related harms, however it has been criticized more recently as having evolved to shift the responsibility of gambling harms onto individual gamblers who don't "gamble responsibly" (Hancock & Smith, 2017). Indeed, a recent scoping review on responsible gambling found that by far the most predominant theme was "responsible gambling tools and interventions", such as a pop-up message on a slot machine suggesting you take a break, or the ability to ban oneself from the casino (Reynolds, Kairouz, Ilacqua, & French, 2020). These types of gambling harm minimization measures focus the responsibility of preventing gambling harm on the person doing the gambling, not on the gambling provider.

Disciplines studying gambling

The results for disciplinary approaches of gambling research are presented in Figure 1 (primary literature) and Figure 2 (grey literature). The data must be compared with caution as they were collected differently, and each have their own limitations. each primary literature journal article was automatically assigned the self-described discipline of the journal in which it was published. The issue is that many journals describe themselves as "interdisciplinary" although the individual articles published within the journal are usually within a single discipline. As a result, nearly one-quarter of articles were assigned to the uninformative category "Interdisciplinary". Conversely, although the grey literature documents were assigned a discipline based on the first author's professional or academic department affiliation, over one-third of items either had an institutional author or the first author's discipline could not be ascertained and were therefore coded as "Not Applicable".

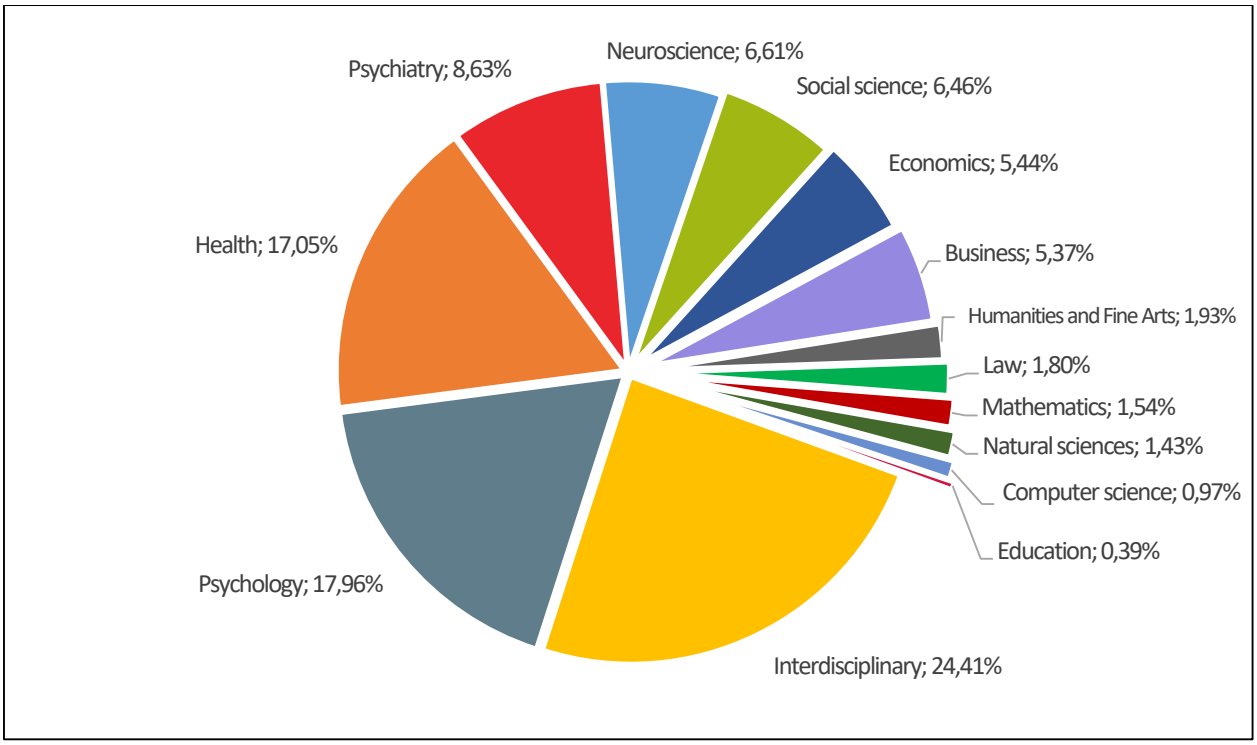


Figure 1: Academic disciplines of primary literature research publications on gambling, according to journal self-description (N=1292)

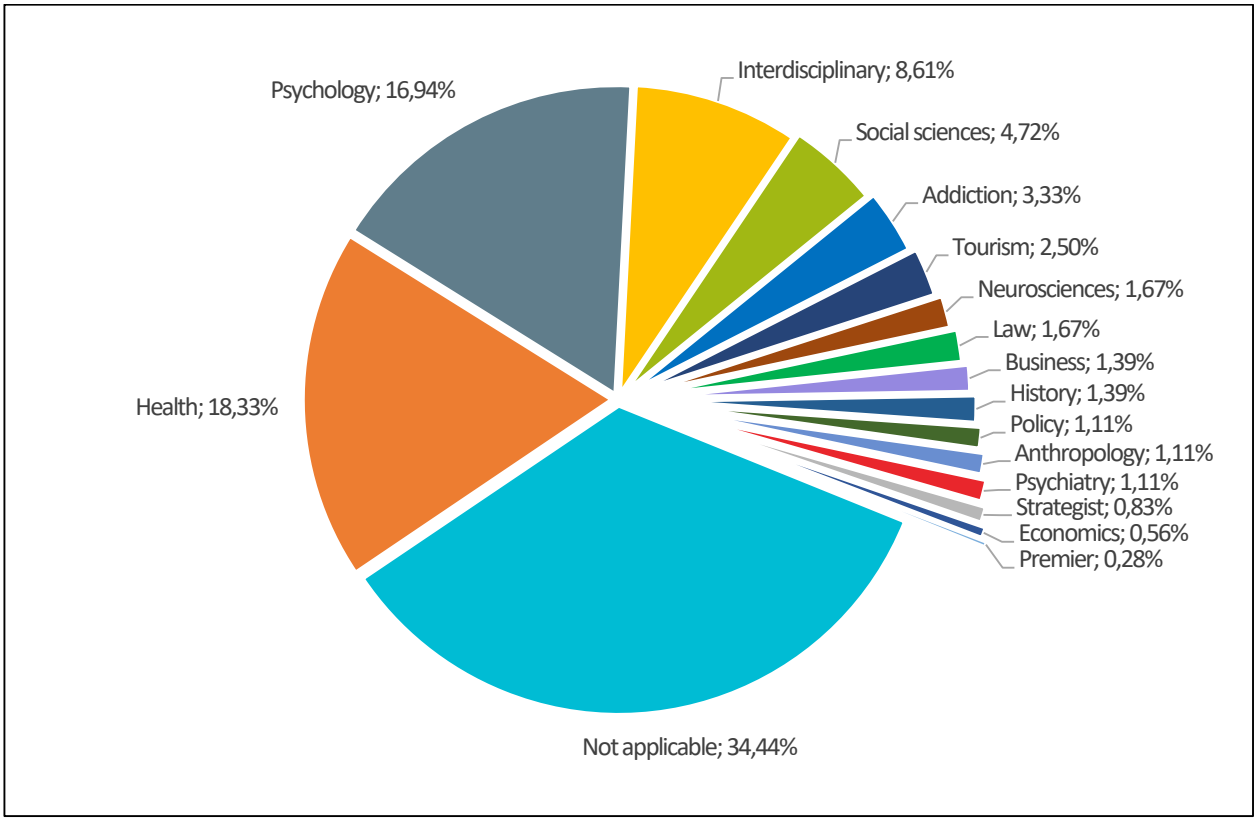


Figure 2: Academic disciplines of grey literature research publications on gambling, according to first-author affiliation (N=360).

Despite these limitations, there are some interesting parallels. Where data were available, they grey literature was more diverse in the total number of disciplinary approaches, but when we focus on the major categories, they are very similar. In the primary literature we expect psychology, health, psychiatry, and neuroscience to dominate as was found in previous research. For the grey literature, on the other hand, our previous analysis found that the grey literature tended to investigate broader societal and population issues, so we might expect less to see less psychological approaches and more social sciences. However, we in the grey literature sample we see psychology is still much larger than the social sciences and other disciplines.

Discussion

This study represents the first broad-scale descriptive and thematic analysis comparing the primary and grey literatures of gambling research. We found that grey literature represents a significant portion of gambling research produced: an average of over 20% of all research documents across the five countries in the sample. The proportion of grey literature varied a lot between the five countries, with the greatest outlier being New Zealand where 49% of gambling research documents were grey publications. New Zealand is also unique in the sample as it is the only country where the gambling legislation requires the government to take public health approach to gambling programmes, including government-funded gambling research (Government of New Zealand, 2003, s.317). This unique model may be specific to gambling or reflect broader cultural differences with how health research or research in general is produced and published. In either case, a stronger culture of publishing gambling research as grey literature may positively affect the quality of the grey research itself: If grey research reports are well-regarded in New Zealand and academics are therefore more incentivized to produce them, the research may be more robust.

The findings of our content analysis are important as they demonstrate the unique value of gambling's grey literature. In the primary literature we found a focus on individual gamblers and their psychological characteristics, as has been reported in previous mapping reviews. On the other hand, the grey literature tended to focus on broader social and population-level issues, as well as minimizing gambling harm. This is not surprising because the majority of grey literature gambling research is funded by government bodies, and governments are accountable to their whole population and are responsible for the health and well-being of the people. We however note that gambling grey literature research in our study may still represent the states' interests in a neoliberal "responsible gambling" approach to gambling policy, whereby the responsibility to prevent gambling harm is downloaded onto individual citizens who gamble.

The quality of gambling grey literature has been questioned but has yet to be assessed. Regardless of this, our finding that gambling's grey literature has some overlap and some differences from the primary literature justifies its inclusion in systematic and scoping reviews as this will yield more comprehensive knowledge syntheses on gambling. Additionally, by including gambling grey literature in systematic reviews, the quality of the studies will be assessed by subject matter experts performing the reviews, thus shedding light on the actual relative quality of gambling's grey research.

Our preliminary disciplinary analysis did not find dramatic differences in the disciplinary approaches between primary and grey gambling research, but taken together with the results of the content analysis, the similarity is interesting. As expected, the primary literature focuses on individual gamblers and their behaviour and thus is dominated by the disciplines of psychology, health, psychiatry, and neuroscience. On the other hand, the grey literature is focused on broader population-level health and social issues, so we might expect disciplines such as health, social sciences and economics to dominate. Although "Health" is the largest discipline in the grey literature sample, "Psychology" is a close second, and the proportions of both social sciences and economics are lower than in the primary literature sample. One possible explanation is that because most of the primary academic research is psychological, most of the researchers who become recognized "gambling experts" are psychologists, who are then called upon to do the (primarily non-psychological) gambling research for governments and other grey literature producing bodies. As the field of gambling studies is already criticized for being overly focused on perspectives of psychology and related sciences, this would be another argument to create more space for social science and humanities research into gambling in academic settings.

Study Limitations and Future Work

This study is the first of its kind and its results should be taken as preliminary. In particular, the dataset is limited to a recent five-year window and the grey literature sample was only retrieved from one database. Although the GREO Evidence Centre database's topic scope matches this project perfectly and is a strong collection in the five study countries, the database is operated by a Canadian organization and thus Canadian grey publications may be more fully represented in this dataset than other nations; in particular the United States may be underrepresented as it is a large country with gambling research and policy publications spread across fifty states. The disciplinary data is also limited as a large number of documents were coded as "Interdisciplinary" and "Not Applicable". Another analysis of the primary literature sample found that "Interdisciplinary" articles mirrored the dominant disciplines of psychology, health, psychiatry, and neuroscience (Akçayir et al., Forthcoming 2021), but more complete data coding must be done for the grey literature for further comparison.

Our forthcoming work aims to address these limitations. We have expanded a timeframe to 1996-2018 to allow for longitudinal analysis of trends, and the grey literature search has been thoroughly expanded to a full systematic search following current guidelines (Godin et al., 2015). This future work will continue to use a big data meta-analytic approach explore the characteristics of gambling grey literature in other areas of critical interest, such as publication types and funding sources. In keeping with the principles of Open Science, the complete dataset of the project including data from this pilot study will be published to a data repository in 2021.

References

Abbott, M. W., Volberg, R., Bellringer, M. E., & Reith, G. (2004). *A review of research on aspects of problem gambling*. London: Responsibility in Gambling Trust. Retrieved from https://niphmhr.aut.ac.nz/__data/assets/pdf_file/0004/7537/summary_of_auckland_report.pdf

Adams, P. J., & Rossen, F. (2012). A tale of missed opportunities: Pursuit of a public health approach to gambling in New Zealand. *Addiction, 107*(6), 1051-1056. https://doi.org/https://doi.org/10.1111/j.1360-0443.2012.03800.x

Adams, R. J., Smart, P., & Huff, A. S. (2017). Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies. *International Journal of Management Reviews, 19*(4), 432-454. https://doi.org/https://doi.org/10.1111/ijmr.12102

Akçayir, M., & Akçayir, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*(C), 1-11. https://doi.org/10.1016/j.edurev.2016.11.002

Akçayir, M., Nicoll, F., Baxter, D. G., & Palmer, Z. S. (2021). Whose responsibility is it to prevent or reduce gambling harm? A mapping review of current empirical research. *International Journal of Mental Health and Addiction*. https://doi.org/https://doi.org/10.1007/s11469-020-00459-x

Akçayir, M., Nicoll, F., Baxter, D. G., & Palmer, Z. S. (Forthcoming 2021). Patterns of influence and academic collaboration in gambling research: A co-citation analysis. *Critical Gambling Studies*.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Washington, DC: American Psychiatric Association.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 533*, 452-454. https://doi.org/https://doi.org/10.1038/533452a

Baxter, D. G., & Hilbrecht, M. (2020). From "Grey Literature" to "Specialized Resources": Rethinking terminology to enhance grey literature access and use. *The Grey Journal, 16*(1), 49-56.

Baxter, D. G., Hilbrecht, M., & Wheaton, C. T. J. (2019). A mapping review of research on gambling harm in three regulatory environments. *Harm Reduction Journal, 16*, 19. https://doi.org/10.1186/s12954-018-0265-3

Blaszczynski, A., Ladouceur, R., & Shaffer, H. J. (2004). A science-based framework for responsible gambling: the Reno Model. *Journal of Gambling Studies, 20*(3), 301-317. https://doi.org/10.1023/B:JOGS.0000040281.49444.e2

Bonato, S. (2018). *Searching the Grey Literature: A Handbook for Searching Reports, Working Papers, and Other Unpublished Research*. New York: Rowman and Littlefield.

Browne, M., Langham, E., Rawat, V., Greer, N., Li, E., Rose, J., . . . Best, T. (2016). *Assessing gambling-related harm in Victoria: A public health perspective*. Melbourne: Victorian Responsible Gambling Foundation. Retrieved from https://www.responsiblegambling.vic.gov.au/__data/assets/pdf_file/0007/28465/Browne_assessing_gambling-related_harm_in_Vic_Apr_2016-REPLACEMENT2.pdf

Cassidy, R. (2020). *Vicious Games: Capitalism and Gambling*. London: Pluto Press.

Cassidy, R., Loussouarn, C., & Pisac, A. (2013). *Fair Game: Producing Gambling Research*. London: Goldsmiths, University of London. Retrieved from https://prism.ucalgary.ca/bitstream/handle/1880/50242/Fair_Game_Web_Final.pdf?sequence=1&isAllowed=y

Godin, K., Stapleton, J., Kirkpatrick, S. I., Hanning, R. M., & Leatherdale, S. T. (2015). Applying systematic review search methods to the grey literature: A case study examining guidelines for school-based breakfast programs in Canada. *Systematic Reviews, 4*(1), 138. https://doi.org/10.1186/s13643-015-0125-0

Goodwin, B. C., Browne, M., Rockloff, M., & Rose, J. (2017). A typical problem gambler affects six others. *International Gambling Studies, 17*(2), 276-289. https://doi.org/10.1080/14459795.2017.1331252

Government of New Zealand. (2003). *Gambling Act 2003* Retrieved from https://www.legislation.govt.nz/act/public/2003/0051/latest/DLM207497.html

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91-108. https://doi.org/https://doi.org/10.1111/j.1471-1842.2009.00848.x

GREO. (2020). *Systems of Funding for Gambling Research*. Brief report prepared for the Gambling Commission, Birmingham, UK.: Retrieved from https://doi.org/10.33684/2020.003

GREO. (n.d.). Specialized Resources. Retrieved from https://www.greo.ca/en/greo-resource/specialized-resources.aspx

Hancock, L., & Smith, G. (2017). Critiquing the Reno Model I-IV international influence on regulators and governments (2004–2015)— the distorted reality of "responsible gambling". *International Journal of Mental Health and Addiction, 15*(6), 1151-1176. https://doi.org/10.1007/s11469-017-9746-y

Heathers, J. (2020). The Lancet has made one of the biggest retractions in modern history. How could this happen? *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2020/jun/05/lancet-had-to-do-one-of-the-biggest-retractions-in-modern-history-how-could-this-happen

Hopewell, S., McDonald, S., Clarke, M. J., & Egger, M. (2007). Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews, 2*. https://doi.org/10.1002/14651858.MR000010.pub3

Jackson, G. (2020). 'Disgusting' study rating attractiveness of women with endometriosis retracted by medical journal. *The Guardian*. Retrieved from https://www.theguardian.com/society/2020/aug/05/disgusting-study-rating-attractiveness-of-women-with-endometriosis-retracted-by-medical-journal

Kim, H. S., Dobson, K. S., & Hodgins, D. C. (2016). Funding of gambling research: Ethical issues, potential benefit and guidelines. *Journal of Gambling Issues, 32*. https://doi.org/https://doi.org/10.4309/jgi.2016.32.7

Korn, D. A., & Shaffer, H. J. (1999). Gambling and the health of the public: Adopting a public health perspective. *Journal of Gambling Studies, 15*(4), 289-365. https://doi.org/10.1023/A:1023005115932

Ladouceur, R., Shaffer, P., Blaszczynski, A., & Shaffer, H. J. (2017). Responsible gambling: A synthesis of the empirical evidence. *Addiction Research & Theory, 25*(3), 225-235. https://doi.org/10.1080/16066359.2016.1245294

Ladouceur, R., Shaffer, P., Blaszczynski, A., & Shaffer, H. J. (2019). Responsible gambling research and industry funding biases. *Journal of Gambling Studies, 35*(2), 725-730. https://doi.org/10.1007/s10899-018-9792-9

Livingstone, C., Adams, P., Cassidy, R., Markham, F., Reith, G., Rintoul, A., . . . Young, M. (2018). On gambling research, social science and the consequences of commercial gambling. *International Gambling Studies, 18*(1), 56-68. https://doi.org/10.1080/14459795.2017.1377748

Markham, F., & Young, M. (2015). "Big Gambling": The rise of the global industry-state gambling complex. *Addiction Research & Theory, 23*(1), 1-4. https://doi.org/10.3109/16066359.2014.929118

OPGRC. (2013). *Building knowledge: Ontario Problem Gambling Research Centre annual report 2012/13*. Guelph, CA: Ontario Problem Gambling Research Centre. Retrieved from https://www.greo.ca/en/about-us/resources/Documents/2012-13-annual-report.pdf

Reynolds, J., Kairouz, S., Ilacqua, S., & French, M. (2020). Responsible gambling: A scoping review. *Critical Gambling Studies, 1*(1), 23-39. https://doi.org/10.29173/cgs42

Schöpfel, J. (2010). *Towards a Prague Definition of Grey Literature*. Paper presented at the Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues., Prague, Czech Republic. https://archivesic.ccsd.cnrs.fr/sic_00581570/document

Shaffer, P. M., Ladouceur, R., Williams, P. M., Wiley, R. C., Blaszczynski, A., & Shaffer, H. J. (2019). Gambling research and funding biases. *Journal of Gambling Studies, 35*(3), 875-886. https://doi.org/10.1007/s10899-019-09875-8

Wardle, H., Reith, G., Langham, E., & Rogers, R. D. (2019). Gambling and public health: We need policy action to prevent harm. *BMJ, 365*, 1807. https://doi.org/10.1136/bmj.l1807

Williams, R. J., Stevens, R. M. G., & Nixon, G. (2011). Gambling and problem gambling in North American Aboriginal People. In Y. D. Belanger (Ed.), *First Nations Gaming in Canada* (pp. 166-194). Winnepeg: University of Manitoba Press.

Williams, R. J., Volberg, R., & Stevens, R. (2012). *The Population Prevalence of Problem Gambling: Methodological Influences, Standardized Rates, Jurisdictional Differences, and Worldwide Trends*. Guelph, CA: Ontario Problem Gambling Research Centre. Retrieved from https://www.greo.ca/Modules/EvidenceCentre/Details/population-prevalence-problem-gambling-methodological-influences-standardized-rates-0

World Health Organization. (2018). *International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision)*. Geneva: World Health Organization. Retrieved from https://icd.who.int/en

Grey Literature Resources generate and drive Awareness to a Circular Economy: An Explorative Research Project*

Dominic Farace and **Jerry Frantzen**, GreyNet International

Abstract:
In September 2019, the first seminar of its kind dealt with grey literature and a circular economy i.e., an economic system aimed at eliminating waste and the continual use of resources. The information compiled in advance of that seminar and the interest shown by the participants provide the lead-up to this study.
Grey literature resources are a significant part of the information industry and like other industries in a circular economy such as textile, construction, and logistics, the role and value of these resources must be understood and demonstrated. This study looks at how grey literature resources are a vehicle for other industries in a circular economy, and at the same time how they themselves are part of an industry, which drives a circular economy.
This study first sets out to gain insight into the opinions of GreyNet's community of practice with regard to a circular economy and to determine if there is consensus. This is carried out via an online community-based survey. The study will further look at the way in which grey literature resources can be seen as a vehicle for other industries in a circular economy. This part of the study is carried out via a search of the literature based on a sample of a number of industries. Together, these results will allow us to explore good practices in generating societal awareness to a circular economy and in doing so, drive awareness to the value of grey literature resources.
As with any explorative research, the outcome has yet to be fully defined. Nevertheless, it is expected that the findings of the survey will indicate a sufficient level of consensus within GreyNet's community of practice. And, the search of the literature will demonstrate that grey literature can be seen as a vehicle for other industries in driving awareness to a circular economy. Nevertheless, differences in the volume of references to grey literature may occur depending on the type of industry included in the sample.

Introduction

Almost two years ago at a social event, the conversation turned to circular economy. A term I had heard in passing, but which I chose to read up on later that evening. Half way through the wiki article, the association of terms used in grey literature became apparent. When I looked at the references, what also became apparent were the number of references from the grey literature. Last year Jerry Frantzen, my co-author and I presented a one-of-a-kind workshop on grey literature and circular economy[1]. The takeaway from that event was the lead up to this research project.

Before proceeding further, the definition of circular economy used in this study, which is shared in great part by both the Ellen MacArthur Foundation[2] and Wikipedia reads as follows: "Circular economy is an economic system aimed at eliminating waste and the continual use of resources. Circular systems employ among other means the reuse and sharing of resources. This regenerative approach is in contrast to the traditional linear economy, which has a 'take, make, dispose' model of production.[3]"

Project Goals and Method

The goal of this research project was to gain insight into the opinions of GreyNet's community of practice[4] with regard to circular economy and to determine if there is consensus. Further, the goal was to explore the way in which grey literature resources can be seen as a vehicle for other industries in a circular economy. In so doing, it would also corroborate grey literature's role as a driver in the information industry. The method of approach in this study was twofold:

* First published in the GL2020 Conference Proceedings, February 2021. https://doi.org/10.26069/greynet-2021-000.456-gg

first to construct a questionnaire and distribute it online among GreyNet's community of practice. And secondly, to carry out a web search of a sample of industries engaged in circular economy.

Survey Questionnaire

In formulating the ten questions that comprise the instrument used in this study, a short list of terms and concepts were drawn from the workshop on Circular Economy and Grey Literature held in the Summer of 2019. Each of the terms and concepts selected would have a particular connotation and use in both the fields of grey literature and circular economy. From the list that was compiled, nine of the terms and concepts were selected and further elaborated. This aided in formulating the nine close-ended survey questions. The tenth question was open ended. The final edited version of the online questionnaire[5] was then entered in SurveyMonkey and a link was generated for online distribution.

Survey Population and Respondents

The survey was online accessible for a four-week period in March of 2020. GreyNet's Distribution List, its Social Media[6] (namely Facebook and LinkedIn) as well as the GreyGuide[7], GreyNet's Web Access Portal and Repository comprised the population used in the study.

Survey Population	Survey Respondents	% Questions Answered
GreyNet's Distribution List, Social Media, and Web Access Portal	72	93.75%

This resulted in 72 survey respondents that answered a near 94% of the survey questions. While the population of the survey was not strictly controlled, all of the survey recipients are known to have some affiliation with grey literature. It can be noted that one-third of the respondents provided their contact details solicited in the final question of the survey.

	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
Q1	45	19	7	0	0
Q2	40	26	4	1	0
Q3	25	28	13	6	0
Q4	18	27	23	4	0
Q5	26	31	13	2	0
Q6	26	29	9	8	0
Q7	44	21	5	0	0
Q8	5	25	32	7	2
Q9	28	35	7	1	1

Table 1: Overview Survey Results

A clear majority of responses were in agreement on seven of the close-ended survey questions dealing with Q1 open access, Q2 loss of information, Q3 information overload and underuse, Q5 persistent identifiers, Q6 knowledge transfer, Q7 accessibility from production through to publication, and Q9 public awareness. A marked uncertainty was found with regard to two of the close-ended survey questions dealing with Q4 a dated mindset and Q8 reusability instead of replicability of research data.

Web Search Procedure

In the second method of approach used in our study, five industries identified with Circular Economy were selected for web-searching – namely: Agriculture, Utilities, Logistics, Textile, and the Automotive industry. Google[8] and Google Scholar[9] were the two web search engines used in the study. It is widely held that the majority of content in Google Search is primarily not controlled by commercial publishing, while Google Scholar's content is. The search terms that were used both standalone and in combination with one another include: "circular economy", "grey literature", and the five industries mentioned above. The results of the web

search were then ranked by their number of hits. For comparative purposes, the web searches first carried out in June 2020 were again repeated in October 2020.

GOOGLE *(Primarily not controlled by Commercial Publishing)* ***June 10, 2020***

"circular economy" 7.610.000	"circular economy" **and** "A.B.C.D.E."	Ranking	"circular economy" **and** "A.B.C.D.E." **and** "grey literature"	Ranking	"grey literature" 1.520.000
Industries:					
A. Agriculture	4.220.000	1	9.250	1	
B. Utilities	362.000	5	4.380	3	
C. Logistics	1.660.000	2	5.120	2	
D. Textile	546.000	4	2.100	5	
E. Automotive	718.000	3	4.090	4	

GOOGLE *(Primarily not controlled by Commercial Publishing)* ***Oct. 5, 2020***

"circular economy" 9.370.000	"circular economy" **and** "A.B.C.D.E."	Ranking	"circular economy" **and** "A.B.C.D.E." **and** "grey literature"	Ranking	"grey literature" 1.700.000
Industries:					
A. Agriculture	8.480.000	1	15.600	1	
B. Utilities	420.000	5	3.790	5	
C. Logistics	3.580.000	2	11.400	2	
D. Textile	643.000	4	8.070	4	
E. Automotive	1.310.000	3	8.990	3	

Table 2: Google Search Results

In June of 2020, a Google Search of the term “circular economy” produced over 7.6 million hits. The same search in October 2020 increased to over 9.3 million hits. In June of 2020, a Google Search of the term “grey literature” produced over 1.5 million hits. The same search in October 2020 increased to 1.7 million hits. When the term “circular economy” was combined separately with each search term corresponding to the five industries in our study, results indicated that Agriculture and Logistics had the highest number of hits respectively. And, the number of hits for each of those two industries more than doubled between June and October of 2020.

When the search term “grey literature” was then further added to the Google Search string results show that {“circular economy” + “agriculture” + “grey literature”} yielded 9250 hits in June and 15600 hits in October. An increase of 6350 hits. The search {“circular economy” + “logistics” + “grey literature”} yielded 5120 hits in June and 11400 hits in October. An increase of 6280 hits.

GOOGLE SCHOLAR *(Primarily controlled by Commercial Publishing)* ***June 10, 2020***

"circular economy" 88.000	"circular economy" and "A.B.C.D.E."	Ranking	"circular economy" and "A.B.C.D.E." and "grey literature"	Ranking	"grey literature" 144.000
Industries:					
A. Agriculture	13.900	2	392	1	
B. Utilities	5.360	5	107	5	
C. Logistics	18.100	1	274	2	
D. Textile	11.400	3	125	4	
E. Automotive	9.120	4	162	3	

GOOGLE SCHOLAR *(Primarily controlled by Commercial Publishing)* ***Oct. 5, 2020***

"circular economy" 101.000	"circular economy" and "A.B.C.D.E."	Ranking	"circular economy" and "A.B.C.D.E." and "grey literature"	Ranking	"grey literature" 162.000
Industries:					
A. Agriculture	32.800	1	481	1	
B. Utilities	5.860	5	120	5	
C. Logistics	19.700	2	347	2	
D. Textile	12.400	3	148	4	
E. Automotive	10.100	4	189	3	

Table 3: Google Scholar Search Results

The results of our search in Google Scholar are described here in much the same way as with the Google search. In June of this year, a search in Google Scholar for the term "circular economy" produced 88,000 hits. The same search in October increased to over 100,000 hits. In June, a search in Google Scholar of the term "grey literature" produced 144,000 hits. The same search in October 2020 increased to 162,000 hits. When the term "circular economy" was combined separately with each search term corresponding to the five industries in our study, results indicated that Agriculture and Logistics also had the highest number of hits. However, the ranking in June, where Logistics was first and Agriculture was second turned the other way around in October. While both of the two industries showed an increase, the increase for "Agriculture" + "Circular Economy" was 58%, while the increase for "Logistics" + "Circular Economy" was 8%.

When the search term "grey literature" was then further added to the Google Scholar search string results show that {"circular economy" + "agriculture" + "grey literature"} yielded 392 hits in June and 481 hits in October. An increase of 89 hits. The search {"circular economy" + "logistics" + "grey literature"} yielded 274 hits in June and 347 hits in October. An increase of 73 hits.

Some Concluding Remarks

Results of the survey show a significant level of consensus in GreyNet's community of practice. However, there is an uncertainty regarding the FAIR data principle[10] of reusability juxtaposed to replicability, which requires further consideration. Gathered from comments pertaining to Question 4, the expression 'dated mindset'[11] was not immediately understood.

The Google and Google Scholar searches provide clear indicative results; however, other search engines and databases are needed to corroborate findings. Variance was demonstrated among the five industries in the study with regard to circular economy, and this variance remained when the term grey literature was added to the search string. Google Search

accounted for an average of 7069 hits across the five industries in the study, while Google Scholar accounted for an average of only 235 hits.

In order to better assess the role of grey literature, it is recommended that search by type of publication and/or publishing body also be included. In close, being a part of the information industry, grey literature can be viewed as both a driver and vehicle for other industries in a circular economy.

Project Spinoff

The term circular economy did not appear in any of the survey questions. However, the terms and concepts used in formulating the survey questions are found in both the fields of grey literature and circular economy. Twenty-four of the 72 survey respondents, who chose to provide their contact details were later asked if the survey and survey data[12] had potential for reuse; and if so, how would they envision this in relation to another area of study? Comments from three of the recipients follow: 1. to understand the socio-economic consequences of the FAIR approach, 2. in repurposing data from COVID-19 tests and observations into research guides, literature searches, and informational webinars, and 3. to engage the research community in reflecting and articulating its needs for open sources.

References

[1] http://www.greynet.org/greyforumseries/circulareconomy.html

[2] https://www.ellenmacarthurfoundation.org/circular-economy/what-is-the-circular-economy

[3] https://en.wikipedia.org/wiki/Circular_economy

[4] http://www.greynet.org/home.html

[5] Survey questionnaire https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:161400/tab/2

[6] http://www.greynet.org/home/socialmedia.html

[7] http://greyguide.isti.cnr.it/

[8] https://www.google.com/

[9] https://scholar.google.com/

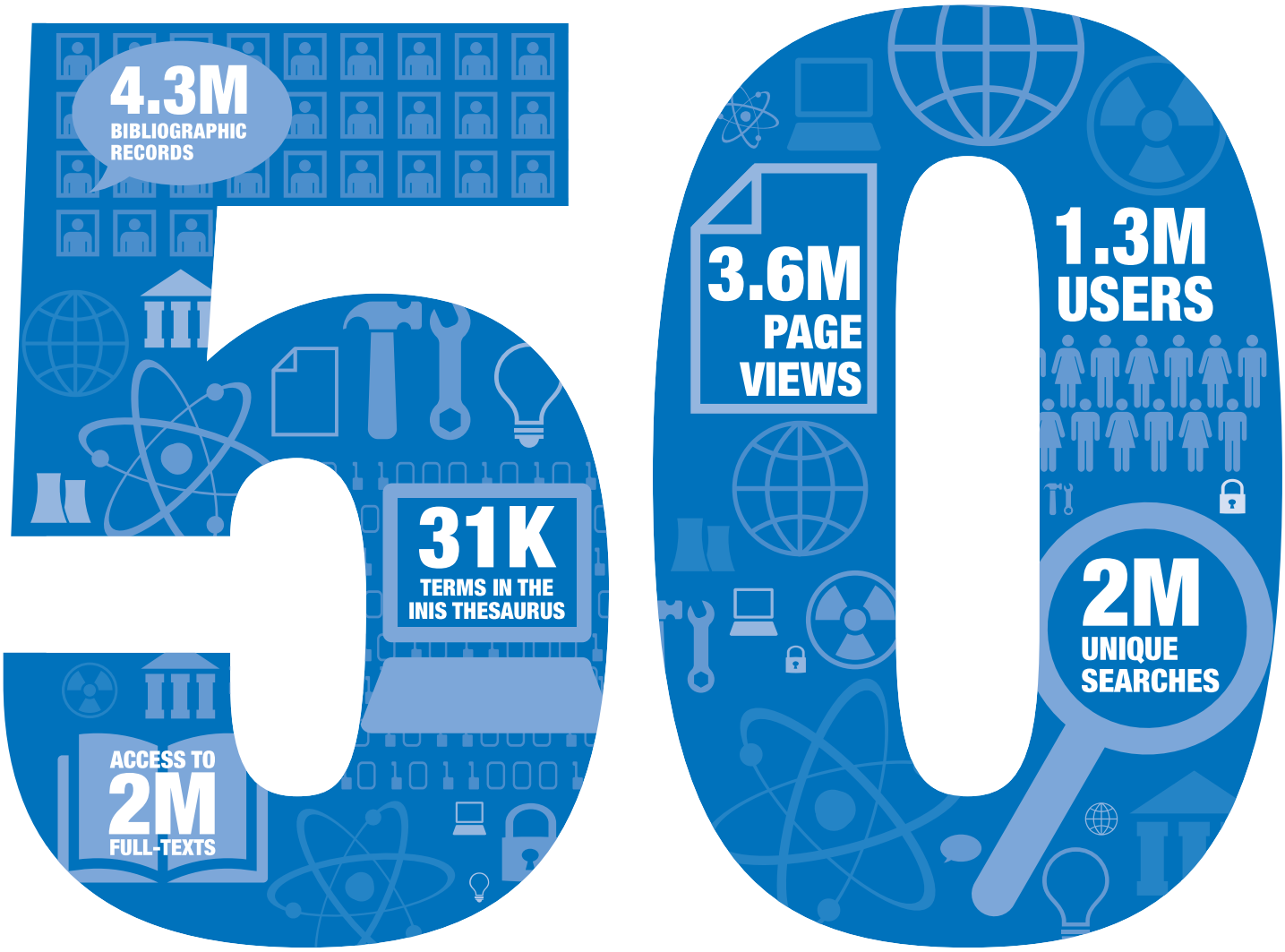
[10] https://www.go-fair.org/fair-principles/

[11] The dated mindset of the 20th Century is an obstacle in the field of grey literature.

[12] Survey data described in a data paperfile:///C:/Users/GreyNet/Downloads/2020-RGL01-005%20(2).pdf

50 YEARS OF INIS

THE WORLD'S TRUSTED NUCLEAR REPOSITORY



Looking for nuclear information?

Want to preserve your nuclear information?

INIS CAN HELP!

The International Nuclear Information System (INIS) was established in 1970 "to foster the exchange of scientific and technical information on peaceful uses of atomic energy".

132 countries and 14 international organizations contribute their national nuclear literature, making it the world's leading open access repository for nuclear science and technology literature.

Explore INIS and find a wealth of information on physics, radiation, climate change, health, etc. **Preserve** your nuclear information by storing it in our trusted repository.

IAEA International Nuclear Information System INIS

INIS.Feedback@iaea.org

iaea.org/inis

20-10469E

International Nuclear Information System (INIS) 50 Years of Successful Contribution to Nuclear Science and Society*

Dobrica Savić, Nuclear Information Section, IAEA

Abstract:
INIS was established in May 1970, as a mechanism to provide access to a comprehensive collection of references to the world's nuclear literature. It has grown from a modest 25-member endeavour to a unique global information resource with a membership of 132 countries. INIS maintains a repository of over 4.4 million bibliographic records, of which 2 million are full text. In 2020, 1.7 million unique visitors made over 2.5 million searches, viewing 4 million web pages.
This paper discusses how INIS operates, the role of its members, the importance of international cooperation, its contribution to nuclear science, its information sharing goals, and the benefits to society of open access to nuclear information.
Keywords: INIS; nuclear information

Introduction

The onset of the cold war in 1947 ushered in an era of fear and uncertainty in nuclear technology. President Dwight D. Eisenhower's *Atoms for Peace* speech to the UN General Assembly in 1953 spurred on the founding of the International Atomic Energy Agency (IAEA) in 1957. The Statute of the IAEA recognizes the need to "…foster the exchange of scientific and technical information on peaceful uses of atomic energy". Thus, with the IAEA Board of Governors approval, INIS was established in May 1970 as a mechanism to provide access to a comprehensive collection of references to the world's nuclear literature.

INIS has grown from a modest 25-member endeavour to a unique global information resource with a membership of 132 countries. It maintains a repository of over 4.4 million bibliographic records, of which 2 million are full text. In 2020, 1.7 million unique visitors made over 2.5 million searches, viewing 4 million web pages.

The peaceful use of atomic energy brings numerous benefits to society. Nuclear technology is used almost everywhere on a daily basis, particularly in the areas of health, environment, water, industrial applications, food and agriculture. INIS offers access to a multitude of documents, reports, articles, and other papers related to science and nuclear technology — a veritable treasure trove for scientists, researchers, government administrators, students and many others. The unique subject area and the sheer volume of information offered by the INIS repository represents a major resource of nuclear information, technological developments and scientific discoveries.

While it is not possible to include all of INIS' achievements throughout its 50 years, this paper will concentrate on the creation of INIS, how it functions, its goals, and its current key role as a global resource of nuclear information.

What is INIS

INIS, as part of the IAEA, is one of the world's largest and most comprehensive repositories of published literature in the field of nuclear science and technology. Operating under special membership arrangements, INIS, currently comprising 132 countries, is a collaborative effort. INIS Liaison Officers (ILOs) are designated by their government authorities and are responsible for collecting their national literature and preparing input to the INIS repository, disseminating information contained in INIS products, and promoting INIS within their national boundaries. Preservation and dissemination are centralized within the INIS Secretariat in Vienna.

The INIS repository contains bibliographic references and full-text documents of conventional and non-conventional (grey) literature, including scientific and technical reports, conference proceedings, patents and theses.

* First published in the GL2020 Conference Proceedings, February 2021.

INIS subject scope covers all areas of IAEA activities, including 50 related categories. The highest areas of input can be seen below (Fig. 1).

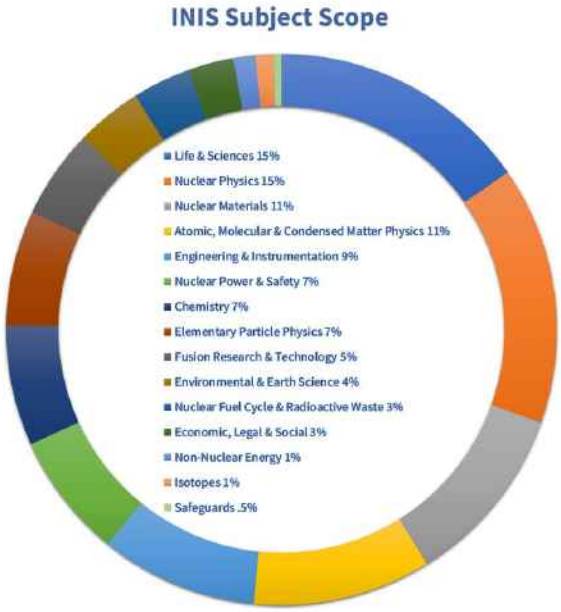


Figure 1: INIS main subject areas

INIS goals

INIS' mandate under the IAEA is to:

- foster the exchange of scientific and technical information on the peaceful use of nuclear science and technology
- collect, process, preserve and disseminate nuclear information
- increase awareness in Member States of the importance of maintaining efficient and effective systems for managing nuclear information resources
- assist Member States with capacity building and training
- provide information services and support to the IAEA and its Member States

INIS also maintains a multilingual thesaurus in Arabic, Chinese, English, French, German, Japanese, Russian and Spanish, providing translations of thousands of technical terms that help navigate and search the collection.

INIS activities are organized in cooperation with its Member States. In addition to regular operational contact with ILOs, consultative meetings take place biennially to discuss policy issues and the overall direction of INIS. The main activities of INIS include:

- Information collection
  - Collect and process bibliographic metadata and full texts of nuclear literature published in IAEA Member States
- Information preservation
  - Electronically preserve non-conventional or 'grey' literature, such as nuclear-related documents, policy and technical reports, and other full-text publications from Member States and international organizations
- Information sharing
  - Ensure free access of the INIS collection to Internet users around the world
- Nuclear knowledge organization
  - Create and maintain the INIS Thesaurus as a major tool for describing nuclear information and knowledge in a structured form
- Capacity building
  - Assist INIS members in improving their effectiveness in nuclear information management

Creation of INIS

With an initial 25 members, the collaborative effort to collect nuclear literature in 1970 was modest — 3950 records were entered into the database in the first year. A seemingly small step which marked a significant beginning for the leading global nuclear information system that INIS would become.

Never had such a geographically and linguistically diverse group of nations cooperated to offer, from a central repository, free, easy to find, and trusted information to scientists, researchers, information specialists, students, government officials, and other users. Initially, the inputting process was tedious and required a lot of manpower. Member States would mail paper documents to the IAEA headquarters in Vienna, where they would be photographed and converted to microfiche. Afterwards, INIS staff would check the incoming information, combine it into a single computer-readable file and distribute it to Member States as machine-readable tapes and semi-monthly abstracting journals.

INIS today

Thanks to the continued cooperation with its ever-growing number of global members and the implementation of innovations in technology throughout the years, INIS has seen dynamic growth in the number of records input to its repository —currently 4.4 million records. Over 100,000 new records are added each year. Improvements in technology include digitalization, the deployment of modern databases and search engines, automated classification, artificial intelligence and machine learning, and harvesting. Figure 2 shows the evolution of records input by country over the last 50 years.

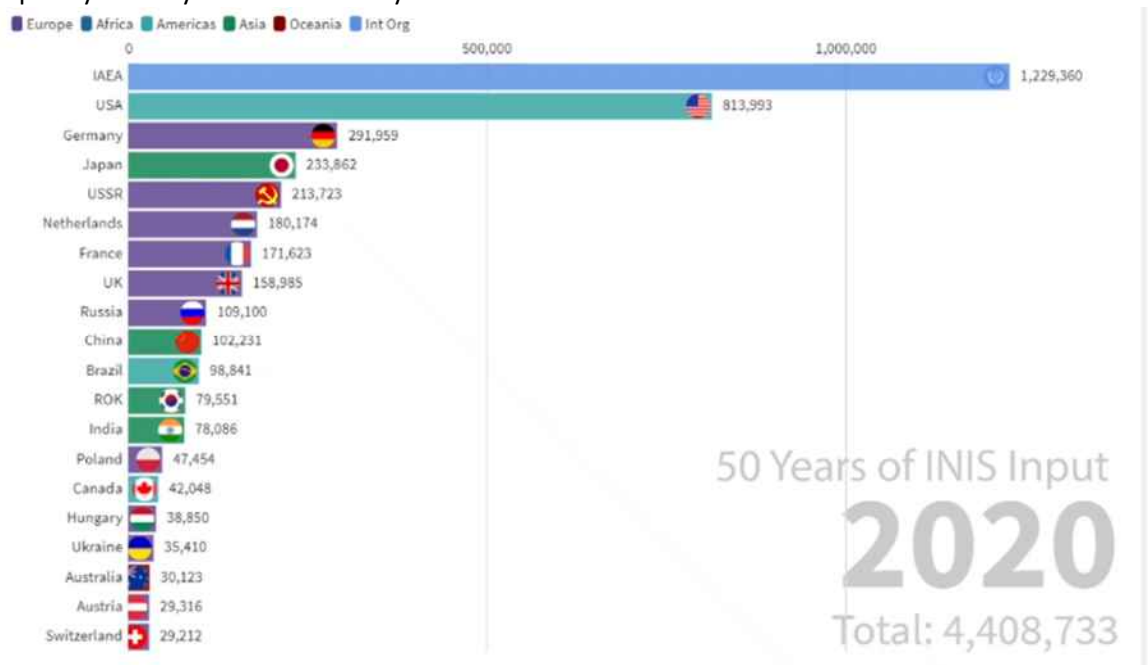


Figure 2: 50 years of INIS input

One of INIS' greatest assets is its collection of more than 2 million full text documents. These can be downloaded directly from INIS servers or through URL or DOI links provided as a part of the INIS bibliographic record.

Another important characteristic of INIS is its open and free accessibility and availability. Statistics show that the repository has been accessed from every corner of the world (Table 1). Its target audience includes researchers, students, government officials, journalists, and the general public.

LAST 10 YEARS	
1. United States	11. Australia
2. India	12. Russia
3. United Kingdom	13. Italy
4. Germany	14. Pakistan
5. Japan	15. Indonesia
6. Canada	16. Turkey
7. Korea, Republic of	17. Malaysia
8. Brazil	18. Spain
9. France	19. Philippines
10. Iran	20. Egypt

Table 1: Top 20 countries of users in last 10 years

INIS has a proven record as an excellent tool for the preservation of nuclear information. IAEA Member States have been able to recover from INIS valuable information that has been lost or damaged. One of these success stories is the Yerevan Physics Institute (YerPhi) in Armenia, whose physical collection had been damaged during storage. Thankfully, the information had been digitally preserved in the INIS repository making it possible to not only recover the information, but to set up a dedicated website linking to the documents in INIS.

Challenges and Opportunities

The challenges that INIS faces today can also be opportunities in its role as a key player in the world of information sharing and preservation.

The biggest challenge INIS faces are from Google and Google Scholar. The Google challenge is two-fold. The first challenge is external, as seen through user statistics — the number of users coming directly to INIS to search for what they need. It is generally assumed that everything can be found by searching Google and therefore not worth further time and effort searching elsewhere. The second challenge is internal — within our organizations — with a reluctance to invest financially in internal information management operations since it seems that everything is already available on the Internet, such as Google.

INIS has benefited from Google Scholar by having all of its documents indexed and made available through their search engine. Still, it should not be forgotten that without INIS having spent 50 years collecting this information, it would not now be widely available and easily accessible. It is important to recognize that documents need to be prepared and input online by someone in order for search engines to find them.

Another challenge arises from the current economic situation, which impacts INIS Member States, and, by extension, the IAEA budget. Despite acknowledging and praising the importance of information, it becomes a vulnerable target in budget cuts. Economic circumstances negatively impact the readiness of Member States to invest in collecting their national information resources and making them available to INIS. This directly impacts the level of funds available to INIS to maintain its repository.

The final challenge is unprecedented, unpredictable, quickly evolving and usually very expensive technological innovations. Digital transformation requires substantial hardware and software updates, changes in established work procedures and methodologies, and upgrades to evolving workforce skills — which necessitates substantial training and re-training.

Wherever challenges are encountered, the opportunities that come with them should also be examined. The greatest opportunity for INIS lies with the trust its members hold in this joint venture, proven throughout 50 years, millions of records, and its millions of users. Its huge collection of bibliographic and full-text records is a remarkable asset that needs to be maintained, reused, repackaged, and repurposed in multiple ways to offer new possibilities and opportunities.

INIS' popularity with its user base is also an immense asset. In 2020 alone, more than 1.7 million unique users visited INIS and performed 2.5 million searches, opened 4 million pages

and downloaded almost 200,000 documents. INIS' relationship with its user base needs to be nourished by offering high quality, relevant, reliable, and trustworthy information.

The INIS Thesaurus, with over 31,000 terms, offers another opportunity to continue its transformation into modern taxonomies and ontologies. These are regarded as the main building blocks for web-based semantic applications and the use of artificial intelligence.

Conclusion

INIS celebrates its 50th anniversary recognizing its many achievements, and with high hopes and expectations. As in the previous half century, INIS is set to play a key role as a global resource of nuclear information in the coming decades.

Bibliography

39th Consultative Meeting of INIS Liaison Officers. IAEA 2018.

From Paper to Digital: IAEA's INIS Marks 50 Years of Nuclear Information Sharing. Retrieved November 2020 from *https://bit.ly/3ulCdhl*

Hakopov, Zaven, Mironov, Dmitry, Savić, Dobrica, & Svetashova, Yulia (Sep 2018). Automated KOS-based Subject Indexing in INIS. CEUR Workshop Proceedings, 17-28.

Huxlin, M. (1998). What is INIS?. ANZ Nuclear Medicine, 29(2), 18.

INIS Progress and Activity Report 2019. IAEA 2020.

International Atomic Energy Agency, Department of Nuclear Energy, INIS and Nuclear Knowledge Management Section, Vienna (Austria) (2011). Nuclear Information and Knowledge, No 11, June 2011 (INIS-XA--11R0582). International Atomic Energy Agency (IAEA): IAEA.

International Atomic Energy Agency, Department of Nuclear Energy, Nuclear Information Section, Vienna (Austria) (2015). Nuclear Information Newsletter No 17, October 2015 45th INIS Anniversary Newsletter (INIS-XA--15M4022). International Atomic Energy Agency (IAEA)

International Nuclear Information System (INIS). Retrieved November 2020 from *https://www.iaea.org/resources/databases/inis*

International Nuclear Information System (INIS) - The First Forty Years 1970-2010. Retrieved November 2020 from *https://bit.ly/3aE2KPd*

International Nuclear Information System (INIS) in Review 2019. IAEA 2019.

Prabhu, Anjali H., & Gupta, Rajiv (2020). Insights into International Nuclear Information System (INIS): scientometric dimensions for 2010-2019 (BARC--2020/E/008). India.

Savić, Dobrica, & St-Pierre, Germain (2013). Digital Preservation at International Nuclear Information System (INIS). 15 international conference on grey literature - The grey audit: A field assessment in grey literature, International Atomic Energy Agency (IAEA): IAEA.

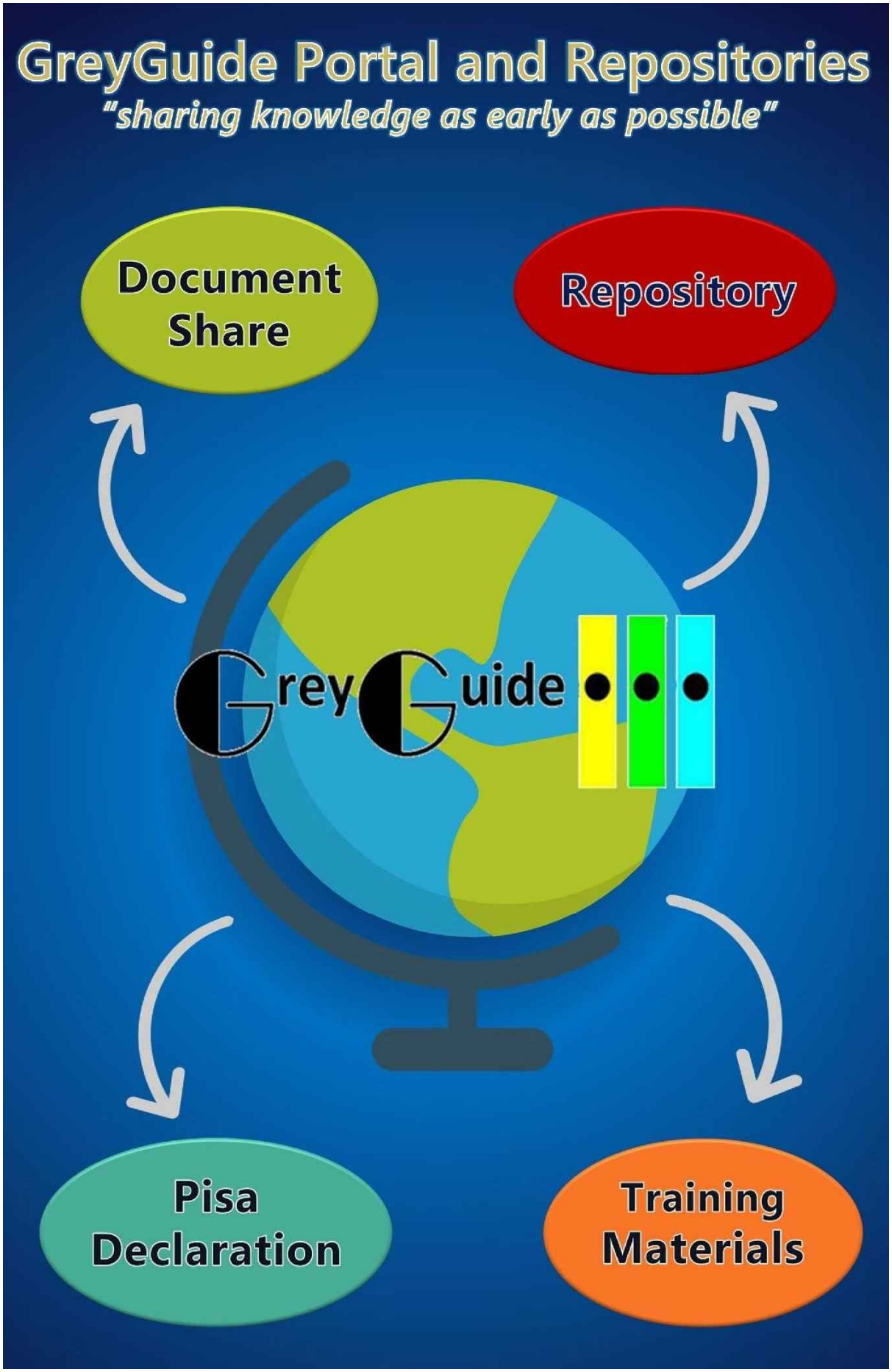
Savić, Dobrica (2013). Nuclear Information Democratization. Online Searcher, 37(6), 30-32,49-51.

Savic, Dobrica (2014). Using Google Search Appliance (GSA) to search digital library collections: A Case Study of the INIS Collection Search. JLISIt (Firenze Online), 5(2), 61-83. doi:104403/jlisit-10071

Savić, Dobrica (2016). INIS: Nuclear Grey Literature Repository. Grey Journal (Online), 7-14.

Savić, Dobrica (2016). Public Interest in Accessing the INIS Collection. GL-conference Series Conference Proceedings, 11.

Zheludev, I.S., & Groenewegen, H.W. (1978). INIS: The International Nuclear Information System. IAEA Bulletin, 20(4), 7-17.


GreyGuide Portal and Repositories
"sharing knowledge as early as possible"
Document Share
Repository
GreyGuide
Pisa Declaration
Training Materials

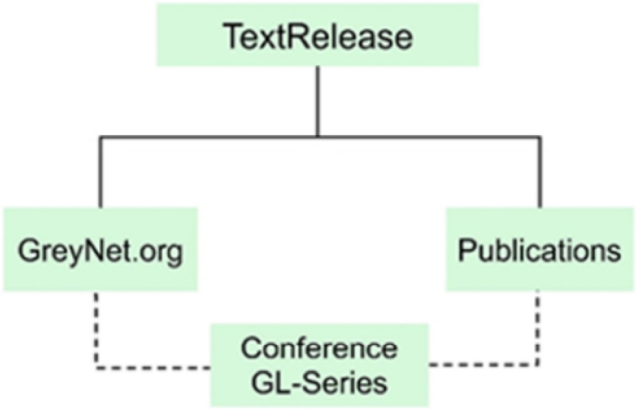
Report on Business, A linked, wiki-like edition. – Autumn 2021

GreyNet International

e-ISSN 2542-4572

This business report is intended to provide a comprehensive, up to date overview of the workings of TextRelease and its International Grey Literature Network Service. TextRelease is an independent information and publishing bureau specializing in the field of grey literature and networked information.

GreyNet was founded in 1992 and re-launched in 2003 under the leadership and direction of TextRelease. Since its re-launch, it has developed into an international network capable of serving various sectors of government, academics, business and industry as well as subject based communities producing, processing, distributing, and preserving grey literature. GreyNet's mission is dedicated to research, publication, open access, education, and public awareness to grey literature, which requires now more than ever an infrastructure commensurate to its real potential.



In order to accommodate further growth, enhance its existing resources, and rise to challenges required for innovation, GreyNet powered by TextRelease seeks to outsource the International Conference Series on Grey Literature. GreyNet will continue as corporate author to the serial publications published by TextRelease such as the GL Conference Proceedings, the Program Books, and The Grey Journal. Likewise, all partnerships, licensed agreements, collaborations with service providers, and contracts undertaken by TextRelease on behalf of GreyNet will remain in effect. The work by GreyNet's Resource Policy Committee, Social Media Publishing Committee, and Education and Training Committee will likewise remain ongoing. GreyNet will further maintain representation on the Conference Program Committee.

This year's edition of GreyNet's Business Report incorporates alongside Uniform Resource Locators (URLs) a number of persistent identifiers (PIDs), namely ORCiD, ROR ID, DOI, OpenDOAR, and Crossref Funder ID. This will enable the reader to easily navigate to primary sources and resources linked to a given text. More on these and other developments are found in this wiki-like edition of the Business Report.

Dominic Farace, Director

Autumn 2021

Leadership in the field Grey Literature

Dominic Farace,
Director 1992 - ...

Jerry Frantzen,
Tech. Assistant 1996 - ...

Dr. Dominic Farace is the Founder/Director of GreyNet International, Grey Literature Network Service, since 1992. In the above capacity, he is author, journal editor, guest lecturer, and curriculum developer in the field of grey literature. Farace is also a recipient of the Golden Candle Award in 2000 for his work in the field of grey literature as well as the Victorine van Schaick Prize in 2008 for The Grey Journal, International Journal on Grey Literature.

Company Information

TextRelease	Established April 1, 2003
Address:	Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
Phone:	+31 (0) 20 331 2420
Email:	info@textrelease.com
Legal Status:	Independent, Sole Proprietorship
Commercial Banking:	Rabobank Amsterdam, IBAN: NL70 RABO 0313 5853 42
Credit Card Recipient:	Visa Card, MasterCard, and American Express
Tax Office (NL):	VAT No. NL001525626B80
Registered (NL):	Amsterdam Chamber of Commerce No. 34188522
Registered (USA):	DUNS No. 408545049
GreyNet International	ISNI, International Standard Name Identifier 0000 0001 1508 0451
GreyNet International	ROR ID, Research Organization Registry

GreyNet was formerly powered by

TransAtlantic[DF] GreyNet	1992-1998, Amsterdam Chamber of Commerce No. 229159
MCB/UP GreyNet	1998-2000, Amsterdam Chamber of Commerce No. 33229159

Notice GreyNet is provided by TextRelease (DUNS Number 408545049), which is a private company based in Amsterdam. TextRelease is a sole proprietorship owned and operated by Dominic Farace. TextRelease has no agents or other signatories.

GreyNet Associate and Institutional Members

TIB Hannover Germany

Associate Member 2017 - …

EBSCO Publishing, United States

Associate Member 2005 - …

ISTI-CNR, Italy

Crossref Funder ID

Associate Member 2012 - …

CVTISR, Slovak Republic

Associate Member 2012 - …

University of Florida, United States

Crossref Funder ID

Associate Member 2018 - …

KISTI, South Korea

Crossref Funder ID

Associate Member 2013 - …

NTK, Czech Republic

Associate Member 2010 - …

Nuclear Information Section, IAEA

Crossref Funder ID

Associate Member 2013 - …

GreyNet provides two levels of organizational membership: **Associate** and **Institutional**. The Associate Members together constitute GreyNet's corporate authorship, which is tasked with capturing various types of information including factual, bibliographic, biographic, documentary and full-text along with accompanying metadata, research data, and post-publication data issuing from the annual Conference Series on Grey Literature as well as related workshops in this field of information. This body of textual and non-textual content is then further compiled, edited, and published in a number of serial publications.

Partnerships and Agreements

Over the years, GreyNet powered by TextRelease has entered into a number of agreements with other businesses and organizations for the advancement of grey literature. Among these include:

FEDGrey Working Group

Inist

Repository

Portal and Repository

TIB AV-Portal	German National Library for Science and Technology, 2018 - …
WorldWideScience.org	World Wide Science Alliance, 2016 - …
OpenAIRE	Open Access Infrastructure for Research in Europe, 2015 - …
ISTI-CNR	GreyGuide Repository and Web Access Portal, 2013 - …
DANS Data Archive	GreyNet Research Data entry and open access, 2012 - …
FedGrey Working Group, LoC	FEDLINK-GreyNet Memorandum of Understanding, 2011 - …
ICSTI	GreyNet Partnership, 2009 - …
Inist-CNRS	OpenGrey Repository, [formerly OpenSIGLE] 2007 - 2017
EBSCO / LISTA-FT	License Agreement for The Grey Journal, 2005- ...

Other Publishing Agreements

CSA/PAIS Int.	A&I Service for 'The Grey Journal', 2006 - …
LISA Index	A&I Service for 'The Grey Journal', 2006 - …
SCOPUS / Elsevier	A&I Service for 'The Grey Journal', 2009 - …
Thomson Reuters	A&I Service for 'The Grey Journal', 2009 - …
Clarivate – Web of Science	A&I Service for the GL Conference Proceedings, 2009 - …
Curran's Scopus | Compendex	A&I Service for the GL Conference Proceedings, 2004 - …
Cabell's Directories	TGJ Journal Listing, 2013 - …
EBSCO	TGJ Subscription Agent, 2005 - …
Emerald	Rights to former GreyNet content 1994-2000 via open access, 2009
De Gruyter/Saur	Monograph on Grey Literature, 2010

GreyNet International – Committee's and Chairmen 2021

CPC	SMP
GL2020 Conference Program Committee *Chair* ***Dr. Plato L. Smith, USA*** **International Conference Series on Grey Literature**	**Social Media Publishing Committee** *Chair* ***Dr. Joachim Schöpfel, France*** **Serials, Non-Serials, and Social Media**
RPC	**E&T**
Resource Policy Committee *Chair* ***Dr. Dobrica Savić, Austria*** **Repositories, Portals, and Web-based Resources**	**Education and Training Committee** *Chair* ***David Baxter, Canada*** **Library and Information Studies**

In providing an infrastructure commensurate to its mission and core activities, GreyNet is supported by four working committees. The four chairpersons form the GreyNet's Joint Chair, which meets annually. Together they advance GreyNet's manifold mission dedicated to research, publication, open access, education, and public awareness to grey literature.

TextRelease Products, Services, and Professional Activities

As Program and Conference Bureau, the main activities of TextRelease include the identification and negotiation with potential conference hosts for the annual conference venues and the identification and acquisition of conference sponsors and advertisers. The formation of Program Committees and Chairpersons along with the initial draft of conference themes incorporated in the announcements and call-for-papers. TextRelease maintains contact with authors throughout the lead-up to the conference providing them with guidelines for their content submissions and informing them of their status on the conference program. TextRelease is further responsible for the design, compilation, editing, and publication of conference materials and documents issuing from the Conference Series. Foremost among these are the conference site, newsletters, program books, and conference proceedings.

International Conference Series on Grey Literature, 1993-2021




The GL Conference Series began as a biannual event in the early '90s, however with its relaunch in 2003, the conference series rose to an annual event. The GL Conference Series maintains a transatlantic character alternating between Europe and North America. However, participants from recognized organizations worldwide actively sponsor and participate.

GL1 **Weinberg Report 2000**
Amsterdam RAI Congress Center, 13-15 December (NL) **1993**

GL2 **Grey Exploitations in the 21st Century**
Catholic University of America, Washington D.C., 2-3 November (USA) **1995**

GL3 **Perspectives on the Design and Transfer of STI**
EU Jean Monnet Building, Luxembourg, 13-14 November (LUX) **1997**

GL4 **New Frontiers in Grey Literature**
Kellogg Conference Center, Washington D.C., 4-5 October (USA) **1999**

GL5 **Grey Matters in the World of Networked Information**
KNAW Congress Center, Amsterdam, 4-5 December (NL) **2003**

GL6 **Work on Grey in Progress**
NYAM Conference Center, New York, 6-7 December (USA) **2004**

GL7 **Open Access to Grey Resources**
INIST-CNRS Congress Centre, Nancy, 5-6 December (FR) **2005**

GL8 **Harnessing the Power of Grey**
UNO Conference Center, New Orleans, 4-5 December (USA) **2006**

GL9 **Grey Foundations in Information Landscape**
Provincial House, Antwerp, 10-11 December (BE) **2007**

GL10 **Designing the Grey Grid for Information Society**
Amsterdam Science Park, 8-9 December (NL) **2008**

GL11 **The Grey Mosaic, Piecing It All Together**
Library of Congress, Washington D.C., 14-15 December (USA) **2009**

GL12 **Transparency in Grey Literature, Grey Tech Approaches …**
National Library of Technology, Prague, 6-7 December (CZ) **2010**

GL13 **The Grey Circuit, From Social Networking to Wealth Creation**
Library of Congress, Washington D.C., 5-6 December (USA) **2011**

GL14 **Tracking Innovation through Grey Literature**
National Research Council, Rome, 29-30 November (IT) **2012**

GL15 **The Grey Audit, A Field Assessment of Grey Literature**
CVTISR, Bratislava, Slovak Republic, 2-3 December (SK) **2013**

GL16 **Grey Literature Lobby, Engines and Requesters for Change**
Library of Congress, Washington D.C., 8-9 December (USA) **2014**

GL17 **A New Wave of Textual and Non-Textual Grey Literature** (NL) **2015**
KNAW Congress Center, Amsterdam, 1-2 December

GL18 **Leveraging Diversity in Grey Literature** (USA) **2016**
The New York Academy of Medicine, New York, 28-29 November

GL19 **Public Awareness and Access to Grey Literature** (Italy) **2017**
National Research Council of Italy, CNR Rome, 23-24 October

GL20 **Research Data Fuels and Sustains Grey Literature** (USA) **2018**
Loyola University New Orleans, Louisiana, 3-4 December

GL21 **Open Science Encompasses New Forms of Grey Literature** (Germany) **2019**
German National Library for Science and Technology, 22-23 October

GL2020 **Applications of Grey Literature for Science and Society** (Italy) **2020**
National Research Council of Italy, *19 November*

GL2021 **Digital Transformation of Grey Literature: Exploring Next Generation Grey**
OBA Forum Amsterdam, Netherlands, Forthcoming *6-7 December* (NL) **2021**

International Conference Series on Grey Literature

Conference Lifecycle and Workflow

Milestones and Deliverables

PRE-CONFERENCE	CONFERENCE	POST-CONFERENCE
Conference Announcement	Program Books	Author full-text papers
Call-for-Papers	Conference Presentations	Conference Proceedings
Conference Proposals	Conference Evaluation	Conference Papers
Author Memoranda	Non-Exclusive Rights Agreement	PowerPoint Slides
Author BIOs	Video Recording	Video Presentations
Call-for-Posters	Poster Prize	Conference Posters
Conference Site	Social Media	Research data
GreyNet Award Dinner	Pre-Conference Announcement	DOI Minting

Conference Call for Tenders 2022

GL Conference Hosts 1993-2021

CNR Rome	Italy
CUA Washington D.C.	United States
CVTISR Bratislava	Slovak Republic
EU CEC/DG XIII	Luxembourg
EWI Brussels	Belgium
FEDLINK, Library of Congress	United States
German National Library for Science and Technology	Germany
Inist-CNRS, Nancy	France
MCB University Press	United Kingdom
NTK, National Library of Technology	Czech Republic
NYAM, New York Academy of Medicine	United States

GL Conference Sponsors 1993-2021

ACM Digital Library	United States
British Library	United Kingdom
BIOSIS	United States
City of Amsterdam	Netherlands
City of Nancy	France
CRL	United States
DANS-KNAW	Netherlands
Deanet Media Company	Italy
DTIC	United States
EAGLE	Luxembourg
EBSCO Publishing	United States
EIPASS | European Informatics Passport	Italy
Elsevier	Netherlands
Endeavor Information Services	United States
EU Bookshop	Luxembourg
EU Cordis	Luxembourg
GESIS	Germany
ICSTI	France
IIa, Inc.	United States
ISTI-CNR	Italy
JST	Japan
KISTI	Korea
LLC	United States
NIS-IAEA	Austria
NTIS	United States
NLE	United States
NLM	United States
OSTI-DOE	United States
Springer Verlag	Germany
Swets	Netherlands
Swinburne University of Technology	Australia
Thomson Reuters	United States
TIB Hannover	Germany
University of Florida; George A. Smathers Libraries	United States
University of Ljubljana	Slovenia
UWM, University of Wisconsin, Madison	United States
Wiley	Italy

International Journal on Grey Literature

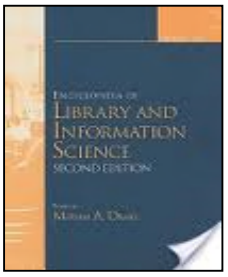
The Grey Journal 2005-2021

ISSN 1574-1796 e-ISSN 1574-180X

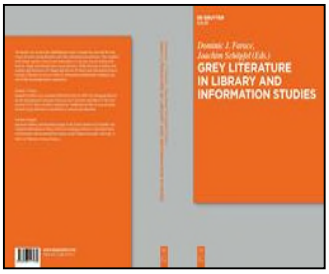
TGJ Volume	7, Number 3, Autumn	2011	Research and Education in Grey Literature
TGJ Volume	6, Number 1, Spring	2010	Government Alliance to Grey Literature
TGJ Volume	6, Number 2, Summer	2010	Shared Strategies for Grey Literature
TGJ Volume	6, Number 3, Autumn	2010	Research on Grey Literature in Europe
TGJ Volume	5, Number 1, Spring	2009	Paperless Initiatives for Grey Literature
TGJ Volume	5, Number 2, Summer	2009	Archaeology and Grey Literature
TGJ Volume	5, Number 3, Autumn	2009	Trusted Grey Sources and Resources
TGJ Volume	4, Number 1, Spring	2008	Praxis and Theory in Grey Literature
TGJ Volume	4, Number 2, Summer	2008	Access to Grey in a Web Environment
TGJ Volume	4, Number 3, Autumn	2008	Making Grey more Visible
TGJ Volume	3, Number 1, Spring	2007	Grey Standards in Transition and Use
TGJ Volume	3, Number 2, Summer	2007	Academic and Scholarly Grey
TGJ Volume	3, Number 3, Autumn	2007	Mapping Grey Resources
TGJ Volume	2, Number 1, Spring	2006	Grey Matters for OAI
TGJ Volume	2, Number 2, Summer	2006	Collections on a Grey Scale
TGJ Volume	2, Number 3, Autumn	2006	Using Grey to Sustain Innovation
TGJ Volume	1, Number 1, Spring	2005	Publish Grey or Perish
TGJ Volume	1, Number 2, Summer	2005	Repositories - Home2Grey
TGJ Volume	1, Number 3, Autumn	2005	Grey Areas in Education

Other Serial and Non-Serial Publications

PRQ



ELIS



De Gruyter | Saur

Publishing Research Quarterly - Spring Issues on Grey Literature 2004-2007 ISSN 1053-8801

Encyclopedia of Library and Information Sciences, 2017 (3rd Ed.) ISBN 978-0-8493-9712-7

De Gruyter - Grey Literature in Library and Information Studies, 2010 ISBN 978-3-598-11793-0

Other GreyNet Publications

Conference Proceedings on Grey Literature	2004 - 2021	ISSN 1386-2316
Conference Program Books on Grey Literature	2003 - 2019	ISSN 1386-2308
GreyNet Quarterly Newsletter	2009 - 2021	ISSN 1877-6035
GreyWorks Summer Workshop Series	2009 - 2019	ISSN 2211-1425
GreyForum Series, Onsite-Online Workshops	2013 - 2019	ISSN 2213-5735

Social Media and Networking

GreyNet LinkedIn Group https://www.linkedin.com/groups/3718857/

GreyNet Twitter Social Blogging https://twitter.com/GreyLitNet

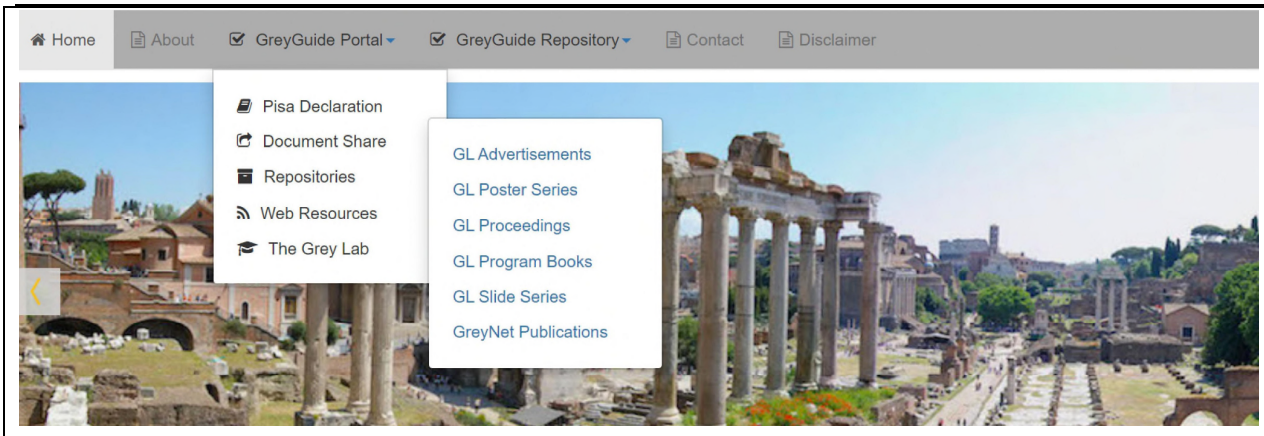
GreyNet International Facebook	https://www.facebook.com/greynetinternational
GreyNet RSS Feeds	http://www.textrelease.com/rss.xml
Wikipedia	https://en.wikipedia.org/wiki/Grey_Literature_Network_Service

GreyNet's Web Access Portal and Repository

GreyNet's Position on the information landscape is Redefined

The GreyGuide Repository and Portal to Good Practices and Resources in Grey Literature offers information professionals, practitioners and students shared common ground in the field of grey literature. In early 2014, GreyNet (content provider) reached agreement with ISTI-CNR (service provider) to embark on the GreyGuide Project. GreyNet collections and web content that were only accessible either via its website or allied conference site began migration to the GreyGuide Repository and Portal. This allowed for combined search, browse, and retrieval capability across collections - whereby standardized metadata and full-text can be harvested online. Alongside good practices in grey literature, GreyNet's in-house collections of persons and organizations in the field of grey literature, as well as research projects, conference proposals, conference papers, and other resources in grey literature are now openly accessible via the GreyGuide.

GreyGuide Portal



GreyGuide Repository

997 Records as of September 10, 2021

BIO	Who is in Grey Literature	**240** Metadata records with ORCiDs and ROR IDs
GLA	Conference Abstracts/Proposals	**190** Metadata records
GLP	Conference Papers	**443** Full-text metadata records with DOIs
RGL	Resources in Grey Literature	**124** Multidisciplinary and multiple document types

In 2020 the GreyGuide Repository was registered in OpenDOAR, the Directory of Open Access Repositories

Starting in 2020, the ROR ID (Research Organization Registry) was entered in the BIO metadata records

Starting in 2018, the ORCiD (a unique, persistent digital identifier assigned to an author-researcher) was included as a metadata field in the **BIO Collection** as well as in **TGJ, The Grey Journal**.

In 2018, GreyNet became a minting service for DOIs (a unique, digital object identifier that is assigned to content, which provides a persistent link to its location on the Internet). The DOI field is included in the **GLP and RGL Collections**.

GreyNet's Open Access Resources and Service Providers

GreyNet completed the second phase of an Enhanced Publications Project (EPP) in which research data accompanying conference preprints are entered in the DANS Data Archive and cross-linked to metadata records in the OpenGrey Repository. Since the project's completion in 2012, research data accompanying conference preprints in the GL-Series are requested as an integral part of the annual call-for-papers and is now embedded in GreyNet's workflow. GreyNet currently has over 48 published research datasets and 17 accompanying Data Papers in DANS.

The TIB AV-Portal is a web-based platform for quality-tested scientific videos. TIB and GreyNet International have a Media Partnership. Video recordings of GreyNet International events are hosted on the AV-Portal. The GreyNet International logo will be shown on the publisher's page in the AV-Portal and GreyNet International is acknowledged as publisher in the metadata of each video. Currently there are 62 published video presentations in GreyNet's collection.

Over 375 full-text metadata records originating in the GL Conference Series including their corresponding full-text papers, PowerPoint slides, abstracts, biographical notes, and post-publication commentaries are available via the OpenGrey Repository http://www.opengrey.eu. This web-based resource has become a significant reference tool for GreyNet's own referral service. In 2007, GreyNet (data provider) signed an agreement with INIST (service provider) in which all of its past and future collections of conference preprints would be openly accessible in the OpenGrey Repository. In 2009, GreyNet purchased back the rights to content from its first four conference proceedings (1994-2000), making these also openly accessible and thus guaranteeing that its collections would be comprehensive.

Agreement with OpenAIRE has been reached with the GreyGuide and OpenGrey, whereby GreyNet's Collections of Conference Preprints and published Conference Papers are now harvested and accessible in OpenAIRE via its Discovery services, which now accounts for 928 research outcomes.

WorldWideScience.org is a global science gateway comprised of national and international scientific databases and portals among which now include the GreyGuide, OpenAIRE, OpenGrey, TIB AV-Portal, and DANS via NARCIS. WorldWideScience.org accelerates scientific discovery and progress by providing one-stop searching of databases from around the world. Multilingual WorldWideScience.org provides real-time searching and translation of globally-dispersed multilingual scientific literature.

The Pisa Declaration on Policy Development for Grey Literature Resources

In April 2014, information professionals from prominent European academic and research centers met in Pisa to address issues dealing with the policy and management of grey literature resources. This meeting resulted in the formulation of the Pisa Declaration on Policy Development for Grey Literature Resources – a fifteen-point roadmap, which serves as a guide for organizations involved in the production, publication, access and use of grey literature well into the 21st Century. The Pisa Declaration much like the Budapest, Bethesda, and Berlin Declarations are neither inventions of the mind nor improvements in doing the same things better – instead, they offer a different approach in dealing with information and data to meet the market needs and requirements of today. Until now, the problem was the lack of cooperation and coordination between and among organizations dealing with grey literature. However, the time of going it alone is now ended. The Pisa Declaration marks the close of an era of *ad hoc* policy and decision making with regard to grey literature resources. Since its original publication in English, the Pisa Declaration has been endorsed by 150 information professionals from renowned organizations worldwide and has now been translated into 22 languages: Armenian, Arabic, Bulgarian, Croatian, Czech, Dutch, French, German, Greek, Hindi, Hungarian, Italian, Japanese, Korean, Macedonian, Russian, Serbian, Slovak, Spanish, Tagalog, Turkish, and Ukrainian.

Conference based Projects on Grey Literature, 1993-2000

Since the initial launch of the GL-Conference Series, GreyNet in collaboration with its pool of authors compiled and edited four information resources. And since its relaunch in 2003, GreyNet annually carries out research projects beneficial to stakeholders in the international grey literature community.

1993-2000 Bibliographic Database on the Topic of Grey Literature

1995-2000 International Guide to Persons and Organizations in Grey Literature

1997-2000 Glossary of Terms in Grey Literature

1999-2000 International Journal on Grey Literature (IJGL)

Conference based Research Projects on Grey Literature, 2003-2021

2003-2004 A Review of Four Information Professionals Their Work and Impact on the Field of GL

2004-2005 Citation Analysis and Grey Literature: Stakeholders in the Grey Circuit

2004-2005 Grey Literature Survey 2004: Research project tracking developments in the field of GL

2005-2006 Access to Grey Content: An Analysis of GL based on Citation and Survey Data

2006-2007	Knowledge Generation in the Field of GL: A Review of Conference-based Research
2007-2008	Grey Literature: A Pilot Course constructed and implemented via Distance Education
2008-2009	OpenSIGLE, Home to GreyNet's Research Community and its Grey Literature
2009-2010	OpenSIGLE - Crossroads for Libraries, Research and Educational Institutions in GL
2010-2011	Peering through the Review Process: Towards Transparency in Grey Literature
2011-2012	Linking full-text grey literature to underlying research and post-publication data
2013-2014	GreyGuide - Guide to Good Practice in Grey Literature
2014-2015	GreyGuide, GreyNet's web access portal and lobby for change in Grey Literature
2015-2016	Leveraging Grey Literature – Capitalizing on Value and the Return on Investment
2016-2017	Policy Development for GL Resources: An Assessment of the Pisa Declaration
2017-2018	Data Papers are Witness to Trusted Resources in Grey Literature: A Project Use Case
2018-2019	Open Data engages Citation and Reuse: A Follow-up Study on Enhanced Publication
2019-2020	AccessGrey: Securing Open Access to Grey Literature for Science and Society
2020-2021	Grey Literature Resources generate and drive Awareness to a Circular Economy
2021	Persistent Identifiers and Grey Literature: A PID Project and GreyNet Use Case

GreyNet Annual Award for Outstanding Achievement in Grey Literature

Since 1999, GreyNet has issued an award in recognition for an author(s) contribution to the field of grey literature. Nominations for the GreyNet Award are based on a number of criteria (1) Results from conference evaluation forms, (2) Publication of the authors' full-text paper in the official Conference Proceedings, (3) Selection and publication of the authors work as a journal article, thus exporting and distributing research results originating in the Conference Series, and (4) Prior history of the author and/or author's organization in the field of grey literature. The recipient may also be an information professional, whose significant contributions over the years has sustained the growth and development of GreyNet International.

GreyNet Awards 1999 – 2021

1999 Julia Gelfand, University of California Irvine, USA
Awarded for her work as journal editor on grey literature.

2000 Daniela Luzi, National Research Centre, Italy
Awarded for her methodological approaches to grey literature.

2004 Bertrum MacDonald, Dalhousie University, Canada
Awarded for his study on the marine science community and grey literature.

2005 Keith Jeffrey, CCLRC U.K. and Anne Asserson, University of Bergen, Norway
Awarded for their work on CRIS systems for grey literature.

2006 Marcus Banks, New York University School of Medicine, USA
Awarded for his work on open access to grey literature.

2007 Todd A. Chavez [et al.], University of South Florida, USA
Awarded for his research on the Karst community and grey literature.

2008 Anne Gentil-Beccot, CERN, Switzerland
Awarded for her research on the physics community and grey literature.

2009 Debbie Rabina, Pratt Institute's School of Information and Library Science, USA
Awarded for her research on copyright licenses and legal deposit of grey literature.

2010 Joachim Schöpfel, University of Lille, France
Award presented for his extensive research and publications on grey literature.

2010 Christiane Stock, Inist-CNRS, France
Award presented for her work and publication on the OpenSIGLE Repository

2011 Bonnie C. Carroll and June Crowe, Information International Associates, Inc. USA
Awards presented for their work on scientific data in grey literature and NGOs.

2012 Petra Pejšová, National Technical Library, Prague, Czech Republic
Award presented for her research and management skills in the field of grey literature

2013 Marcus Vaska, Knowledge Resource Service; University of Calgary, Canada
Award presented for his extended research in raising awareness of grey literature.

2014 Kiyoshi Ikeda, Intellectual Resources Department, Japan Atomic Energy Agency, JAEA
Award for his work in collecting, organizing, and disseminating data and information on nuclear accidents and in particular the Fukushima Daiichi Nuclear Power Station Accident.

2015 Marnix van Berchum, Data Archiving and Networked Services, DANS-KNAW, Netherlands
Award presented for the technical support as service provider for GreyNet's research data issuing from the International Conference Series on Grey Literature**.**

2015 Stefania Biagioni, Institute of Information Science and Technologies, ISTI-CNR, Italy
Award presented for the development and maintenance of the GreyGuide Repository and Web Access Portal for GreyNet's Collections and Resources**.**

2016 Blane K. Dessy, Library of Congress, Washington D.C., USA
Award presented in recognition for sustained support research in the field of grey literature.

2017 Dobrica Savić, Nuclear Information Section; International Atomic Energy Agency, UN
Award for his leading role in drafting, promoting, and assessing the 'Pisa Declaration on Policy Development for Grey Literature Resources'.

2018 Tomas A. Lipinski, University of Wisconsin - Milwaukee; School of Information Studies, USA
Award for his corpus of publications dealing with legal issues and information rights in the field of grey literature.

2019 Plato L. Smith, University of Florida, George A. Smathers Libraries, USA
Award for his work dealing with grey literature research data, its reuse, and preservation. As well as for his profound organizational support.

2020 Silvia Giannini, Institute of Information Science and Technologies, ISTI-CNR, Italy
Special recognition for her professional and decisive style of management of the Twenty-Second International Conference on Grey Literature during a year challenged by a pandemic

2021 Carlo Carlesi, Institute of Information Science and Technologies, ISTI-CNR, Italy
Ten years of service to the international grey literature community as technical developer and support of the GreyGuide – GreyNet's web-access Portal and Repository

2021 Jerry Frantzen, GreyNet International, Grey Literature Network Service, Netherlands
Twenty-five years of service to the international grey literature community as technical editor, project assistant, and the overall technical support of of GreyNet International

GreyNet Honorary Member Register

In 2017, the international Grey Literature Network Service celebrated its Twenty-fifth Anniversary and understands that this milestone was reached in great part through the financial and content contributions of its associate, institutional, and individual members. In recent years, a number of GreyNet Members have retired; however, they continue to share interest in the field of grey literature and remain in contact with GreyNet's range of information services and activities. As an organization, GreyNet chooses to honor these information professionals and seeks to offer a token of lasting thanks for their sustained support.

⁎ **Roberta Shaffer,** USA ⁎ **Herbert Gruttemeier,** France ⁎ **Keith Jeffery,** United Kingdom

⁎ **June Crowe,** USA ⁎ **Leonid Pavlov,** Russia ⁎ **Janie Kaplan,** USA ⁎ **Rosvita Vaska,** Canada

Educational and Training Initiatives in Grey Literature

Course modules and training are available not only to instructors and students at Schools and Colleges of Library and Information Science, but also to professionals and practitioners working to serve their knowledge-based communities. Through the years, GreyNet has taken an active role in curriculum development and instruction. Such initiatives include:

Workshops and Seminars

Introduction to Grey Literature, Supply and Demand Side Grey, Strategies and Future Trends, etc.

Distance Education Program 2007-2009

Undergraduate Accredited Course on Grey Literature via the University of New Orleans, USA

Guest Lecture Series 2004-2008

Lectures on Grey Literature for Masters Students at the University of Amsterdam, Netherlands

Presentations at Conference, Meetings, and Other Events

Invited Papers and Presentations on topics related to Grey Literature

GreyWorks - Summer Workshop Series on Grey Literature

ISSN 2211-1425

Benchmarks and Forecasts on Grey Literature, Washington DC, USA	May 5, 2009
Benchmarks and Forecasts on Grey Literature, Amsterdam, NL	July 2, 2009
Transparency Governs the Grey Landscape, Washington DC, USA	Aug.10,2010
Ten Strategies for Grey Literature, Amsterdam, NL	Sept.9, 2011
Strategic Mapping of Grey Literature, The Hague, NL	Aug.31,2012
Grey Boot Camp: Exposure to the Grey Landscape, Washington DC, USA	July 25,2013
Leveraging Grey Literature: A Training in Resource Assessment, DC, USA	Aug.11,2015
Leveraging Grey Resources: A Training Module, UvA, Amsterdam, NL	Sept.21,2016
Data Papers: An Open Access Tool for Citizen Science, OBA, Amsterdam, NL	Aug.16, 2018

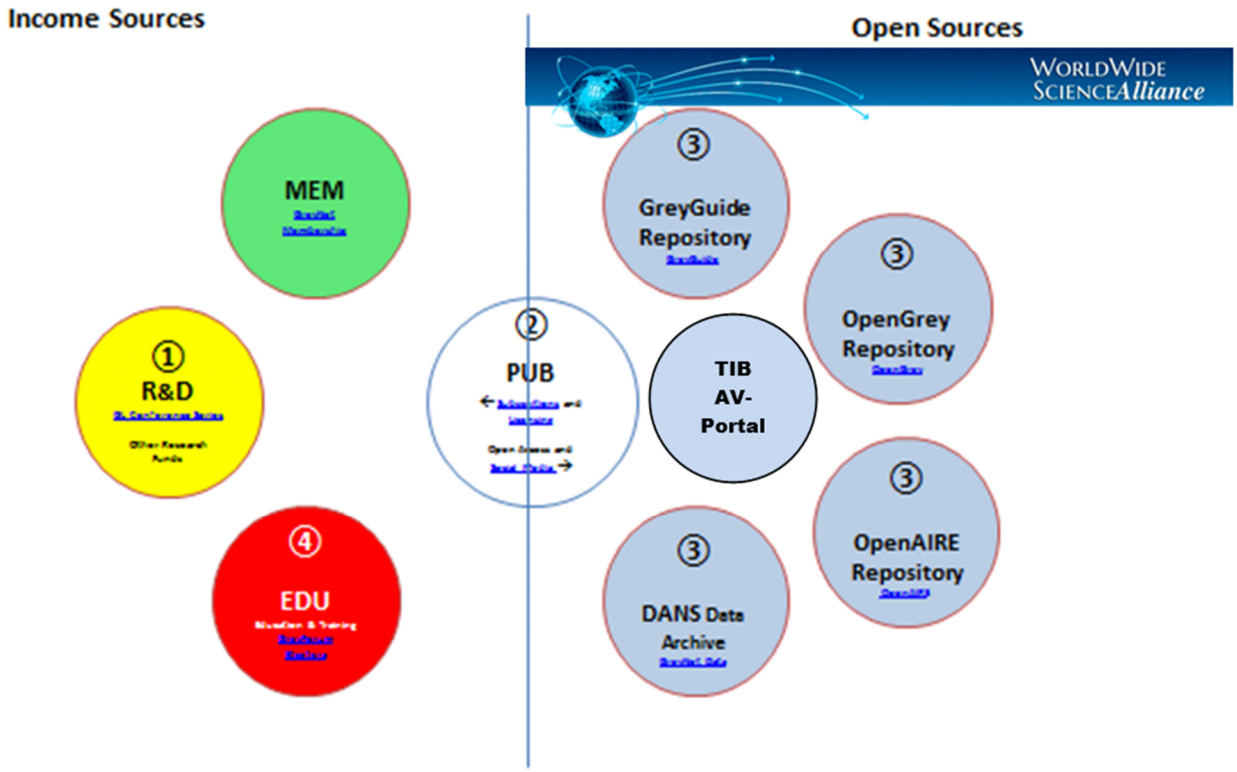
GreyForum - Onsite and Online Seminars and Workshops

"Where grey literature provides common ground for information professionals"

ISSN 2213-5735

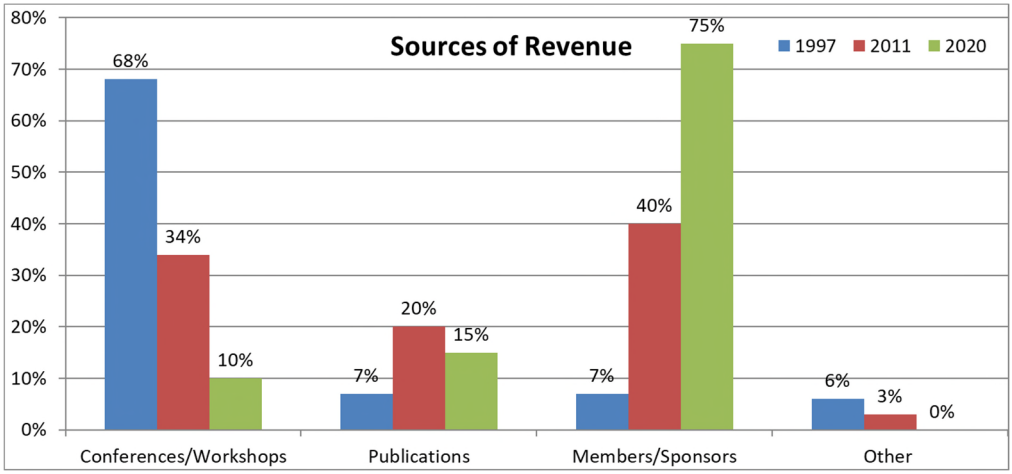
Grey Literature and Information Ethics, St. Hilda's College, Oxford, UK	Jan.31, 2013
Grey Literature and Information Rights, KNVI Conference, NL	Nov.14,2013
Grey Literature and Policy Development, CNR Pisa, Italy	April 7, 2014
Grey Literature and Digital Preservation, Pinto Huis, Amsterdam, NL	Nov.30,2015
Wikipedia for Grey Resources, Library of Congress, Washington DC, USA	May 2, 2016
Data Papers, Trusted Tool in Research & Data Sharing, University of Florida, USA	Mar.20, 2018
Open Data for Research and Enhanced Publication, CNR, Pisa, Italy	May.25, 2018
Data Papers, The Why and How, NTK Prague, Czech Republic	Oct. 23, 2018
Grey Literature and the Circular Economy, OBA, Amsterdam, NL	Sept, 5, 2019

GreyNet's Sustained mix of Income and Open Sources



Financial Indicators

Since its relaunch in 2003, GreyNet has increased and diversified its sources of revenue. Where it initially was dependent almost solely on participant fees issuing from the international conference series, it has since increased revenues issuing from membership fees, sponsorship fees, advertising, the sale of publications, journal licensing, royalties, as well as educational and training activities. In 2020, membership and sponsorship fees accounted for 75% of the revenues, conference and workshops fees accounted for 10%, and publication fees were 15%. Together, these three sources of revenue accounted for 100% of GreyNet's gross income.



As seen in the chart above, revenues issuing directly from conference fees have considerably decreased over the years, while the number of conference participants has not. This can be accounted for by the fact that GreyNet has acquired more organizational members as well as conference sponsors. Another factor that explains the decrease in conference revenues are the reductions now offered to authors, students, and employees of host and partner organizations.

Operating Costs

Direct and indirect costs associated with TextRelease and its Grey Literature Network Service include human resources, office maintenance, travel and lodging (related to site visits, conferences, workshops, committee meetings and other field work), conference materials, publishing costs, etc.

SWOT Analysis / Call for Tenders

Strengths

Leadership role, an established and recognized international information network, GreyNet as Brand Name, Corporate Memory, Partnerships and Agreements, Good Will, Commitment to Open Access, A near 30 years expertise in the field of grey literature, Sustained information resources, and a track record of before and after sales delivery and service, Established banking and fiscal control.

Weaknesses

ICT issues such as website functionality including graphic design, online ordering, use/user statistics and their analyses, underdeveloped marketing, sales, and promotional activities. Insufficient reserves for further investment. A sole proprietorship seen as an obstacle to potential membership. Insufficient presence at key national and international conferences where grey literature is either directly or indirectly related to the themes and programs.

Opportunities

Further role in training and education both onsite and online, as well as research and (digital) publication. Ability to upgrade, enhance, and leverage existing products and services. Increased potential for sponsorship. Better access to funding and grants. Streamlining the workflow through outsourcing. Further cooperation with the open access community as well as further cooperation with commercial publishers and licensing agents. More use of available social media and networks.

Threats

Continuity, inability to respond to growth leading to missed opportunities, infrastructure's command of the ever-changing information landscape, an uncertain financial climate.

the grey Journal

Victorine van Schaick Prize 2008

International Journal on Grey Literature

Subscription Order Form

For GreyNet Members free or reduced rates, see http://www.greynet.org/membership.html

TGJ Volume 18, 2022	*Type of Subscription:*	*Amount in Euros*	*Total*
THE GREY JOURNAL - PDF/ Email / PWP ISSN 1574-180X *Annual Subscription, including Electronic Handling*	☐ Institutional	€ 240	€

Customer information

Name:	
Organisation:	
Postal Address:	
City/Code/Country:	
E-mail Address:	

Check one of the boxes below for your Method of Payment:

❑ Direct transfer to TextRelease, Rabobank Amsterdam, Netherlands
BIC: RABONL2U **IBAN:** NL70 RABO 0313 5853 42, with reference to 'TGJ/The Grey Journal'

❑ MasterCard/Eurocard ❑ Visa card ❑ American Express

Card No. _____ Expiration Date: _____

Print the name that appears on the credit card, here _____

Signature: _____ **CVC II code:** _____ *(Last 3 digits on signature side of card)*

Place: _____ Date: _____

NOTE **Creditcard transactions authorized by Saferpay, Secure Paygate for VisaCard and MasterCard**

Correspondence Address:

TextRelease Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
Tel +31-(0) 20-331.2420 • info@textrelease.com • www.textrelease.com

Twenty-Third International Conference on Grey Literature

Digital Transformation of Grey Literature: Exploring Next Generation Grey

OBA Congress Center • 6-7 December 2021
Oosterdokstraat 110, Amsterdam, Netherlands

Day 1 Conference Program

08:30 **REGISTRATION DESK OPENS WITH COFFEE AND TEA SERVICE**

09:00 **OPENING SESSION**

▪ **Welcome Address**	▪ **Keynote Address**	▪ **Opening Address**
Martin Berendse CEO Amsterdam Public Library Netherlands	Dr. Gregory J. Gordon President SSRN, Elsevier United States	Dr. Donatella Castelli Head of InfraScience, CNR Italy

MODERATOR – PLATO L. SMITH, UNIVERSITY OF FLORIDA; GEORGE A. SMATHERS LIBRARIES, USA

11:00 SESSION 1 – DIGITAL TRANSFORMATION AND THE CHANGING INFORMATION LANDSCAPE

- **The Impact of Digital Transformation on the Sustainability of Grey Literature**
  Dobrica Savić, Nuclear Information Section, International Atomic Energy Agency, NIS-IAEA, United Nations
- **Deepfakes: A Digital Transformation Leads to Misinformation**
  Nika Nouri and Julia Gelfand, University of California, Irvine United States
- **Exploring Video Abstracts to share Science**
  Margret Plank, Leibniz Information Centre for Science and Technology, Germany
- **Grey Literature citation and inclusion rates in gambling review articles: Opportunities for improvement**
  David Baxter, Department of Political Science, University of Alberta, Canada
- **Persistent Identifiers and Grey Literature: A PID Project and GreyNet Use Case**
  Dominic Farace, GreyNet International; Stefania Biagioni and Carlo Carlesi, GreyGuide ISTI-CNR, Italy
  Chris Baars, DANS-KNAW, Netherlands

14:00 SESSION 2 – INSIGHTS AND ISSUES CHALLENGING GREY LITERATURE

- **Grey Literature in Open Repositories: New Insights and New Issues**
  Joachim Schöpfel and Eric Kergosien, University of Lille- GERiiCO, France
  Hélène Prost, CNRS - GERiiCO, France; Florence Thiault, University of Rennes 2, France
- **Digital Grey Soviet Science**
  Yuliya B. Balashova, Saint Petersburg State University, Russia
- **Zines as Nonbinary Objects and Questions of Privilege**
  Andrea Marshall, Centre for Media and Celebrity Studies, United States
- **The endless life of OA Journals from myth to reality: Survey on the present status of vanished OA journals in Iran and future prospect**
  Hamideh Memari Hanji and Saeideh Memari Hanji, National Library and Archives of Iran
- **Burning Grey: The Worldwide Influence of a Locally Published Grey Literature**
  Vince Ervin V. Palcullo, Donna May C. Rivera, Via Marie F. Dumenden, and Ma. Cynthia T. Peleña, Central Philippine University; Joy F. Geromiano and Daryl Superio, Southeast Asian Fisheries Development Center, Philippines

16:30 INTRODUCTION TO CONFERENCE POSTERS AND SPONSOR SHOWCASE

Lightening presentations in the main conference hall in advance of the Poster Session on Day Two

18:30 CONFERENCE DINNER AND GREYNET AWARD PRESENTATIONS

Grand Café Restaurant 1e Klas - Stationsplein 15, Platform 2B Central Station

TextRelease
GL2021 Program and Conference Bureau
Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
www.textrelease.com • conference@textrelease.com
Tel.+31-20-331.2420

Twenty-Third International Conference on Grey Literature

Digital Transformation of Grey Literature: Exploring Next Generation Grey

OBA Congress Center • 6-7 December 2021
Oosterdokstraat 110, Amsterdam, Netherlands

Day 2 Conference Program

09:00 **REGISTRATION DESK OPENS WITH COFFEE AND TEA SERVICE**

09:30 **POSTER SESSION AND SPONSOR SHOWCASE**

The Poster Session continues on the morning of Day Two, where the presenters meet with delegates and participants in an informal setting. Those presenting conference posters are eligible for the Poster Prize 2021 that will be awarded during the Closing Session. Posters will be judged on their innovative content, relevance to the conference topics, graphic design, and presentation.

MODERATOR – JUDITH C. RUSSELL, DEAN OF UNIVERSITY LIBRARIES, UNIVERSITY OF FLORIDA, USA

11:30 **SESSION 3 - ENSURING PUBLIC ACCESS TO RESEARCH DATA**

Ensuring Public Access to Research Data: Perspectives from Three Academic Research Libraries

Rick Anderson, University Librarian, Brigham Young University, United States
Michael Levine-Clark, Dean of Libraries, University of Denver, United States
Judith C. Russell, Dean of University Libraries, University of Florida, United States

In this session, three academic research librarians will provide an overview of CHORUS; outline some of the challenges faced by research institutions of different sizes in managing data; articulate the differing roles of libraries in this process at these universities; and describe how these institutions might make use of CHORUS or similar tools to more effectively manage data and link it to related publications - and also seek input from and potential partnerships among the GreyNet community to further our common objectives.

MODERATOR – JOACHIM SCHÖPFEL, UNIVERSITY OF LILLE, FRANCE

14:00 **SPECIAL PANEL SESSION – EXPLORING NEXT GENERATION GREY**

David Baxter, University of Alberta, Canada
Silvia Giannini and *Anna Molino,* Institute of Information Science and Technologies, ISTI-CNR, Italy
Tomas A. Lipinski, School of Information Studies, University of Wisconsin—Milwaukee, United States
Veronika Potočnik, National and University Library, Slovenia
Dobrica Savić, Nuclear Information Section, NIS-IAEA, United Nations

GL2021 offers the many and diverse communities of practice in the field of grey literature a unique opportunity to collaborate in addressing and defining the next phase in the digital transformation of grey literature. Together this can be accomplished by unlocking the potential next generation grey holds for information science and society. The panel members will issue prepared statements and actively solicit discussion. The statements by the panelists, recorded comments by the conference participants, and results of an online survey carried out prior to the conference will be compiled and further drafted in a conference paper and subsequent journal article entitled 'Amsterdam Manifest on Grey Literature: What the future holds for this field of information'. This document will mark the 30th Anniversary of the International Conference Series on Grey Literature, 1992-2022.

16:00 **CLOSING SESSION –** Report by the Conference Moderators, Presentation of the Poster Prize, Conference Handoff, and Farewell

16:30 **POST CONFERENCE TOUR OF THE AMSTERDAM MAIN PUBLIC LIBRARY**

Twenty-Third International Conference on Grey Literature

Digital Transformation of Grey Literature: Exploring Next Generation Grey

OBA Forum - Amsterdam, Netherlands
December 6-7, 2021

Call for Posters

Title of Poster:	
Author Name(s):	**Phone:**
Organization(s):	**Email:**
Postal Address:	**URL:**
Postal Code – City – Country:	

Guidelines

Persons who seek to present a poster during GL2021 are invited to submit an English abstract between 250-300 words. The abstract should describe the project, activity, information product or service. The abstract should likewise include a title, name(s) of the creator(s) and their full address information. Abstracts are an important source of information available prior to the conference that is accessible to conference delegates and the international grey literature community.

Due Date for Submission

Timely registration is a guarantee for your placement on the conference program. The conference venue is able to accommodate a limited number of physical posters that will be mounted on display panels allowing for optimal viewing. Abstracts in MSWord should be emailed to conference@textrelease.com on or before **November 15th 2021**. Those submitting poster abstracts will receive verification upon receipt accompanied by further guidelines for posters, submission of slides, Conference Registration, etc.

Poster Prize 2021

Those presenting conference posters are also eligible for the Poster Prize 2021 that will be awarded during the conference Closing Session. Posters will be judged by a panel of jurors on their innovative content, relevance to the conference topics, graphic design, and submitted abstract.

Related Conference Topics		
❑ Cite Grey	❑ Evidence based Grey	❑ Policy development
❑ Citizen Science	❑ Funding Grey	❑ Reuse of Data
❑ Communities of Practice	❑ Linking Grey Resources	❑ Vanished OA Journals
❑ Digital Grey	❑ Optics of Grey	❑ Other Related Topic:

TextRelease
GL2021 Program and Conference Bureau

Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
www.textrelease.com • conference@textrelease.com
Tel.+31-20-331.2420

Author Information

Akçayir, Murat **155**
Murat Akcayir is an Alberta Gambling Research Institute postdoctoral research fellow in meta-analysis at the University of Alberta. He received both his masters and PhD degree in Educational Technology. He was also a visiting researcher at the Faculty of Science and Technology in Athabasca University. His research interests include literature reviews, meta-analysis, emerging technologies and digital games.
ORCID_ID https://orcid.org/0000-0002-7607-9478

Baxter, David G. **155**
David Baxter is the Information Specialist at Gambling Research Exchange (GREO). He is responsible for the collection development of GREO's Evidence Centre, a digital library of information resources, research datasets, and other grey literature that help advance knowledge of gambling-related issues. David sits on GreyNet's LIS Education and Training Committee, and is interested in the role of grey literature as a source of evidence for gambling research and policy. David holds a Master of Information from University of Toronto and a Bachelor of Science Honours in Biology from McMaster University.
ROR_ID https://ror.org/047mprv74
ORCID_ID https://orcid.org/0000-0001-5235-6728
E-mail david@greo.ca

Farace, Dominic **165**
Dominic Farace is Head of GreyNet International and Director of TextRelease, an independent information bureau specializing in grey literature and networked information. He holds degrees in sociology from Creighton University (BA) and the University of New Orleans (MA). His doctoral dissertation in social sciences is from the University of Utrecht, The Netherlands, where he has lived and worked since 1976. After six years heading the Department of Documentary Information at the Royal Netherlands Academy of Arts and Sciences (SWIDOC/KNAW), Farace founded GreyNet, Grey Literature Network Service in 1992. He has since been responsible for the International Conference Series on Grey Literature (1993-2013). In this capacity, he also serves as Program and Conference Director as well as managing editor of the Conference Proceedings. He is editor of The Grey Journal and provides workshops and training in the field of grey literature.
ROR_ID https://ror.org/01pxfxj80
ORCID_ID https://orcid.org/0000-0003-2561-3631
Email: info@greynet.org

Fraser, Nicholas **131**
Nicholas Fraser is a postdoctoral researcher in the Web Science research group at ZBW - Leibniz Information Centre for Economics, Kiel, Germany.
ROR_ID https://ror.org/03a96gc12
ORCID_ID https://orcid.org/0000-0002-7582-6339

Frantzen, Jerry **165**
Jerry Frantzen graduated in 1999 from the Amsterdam University of Applied Sciences/Hogeschool van Amsterdam (HvA) in Library and Information Science. Frantzen is the technical editor of The Grey Journal (TGJ). And, since 1996, he is affiliated with GreyNet International, Grey Literature Network Service, as a freelance technical consultant.
ORCID_ID https://orcid.org/0000-0002-3405-7078
ROR_ID https://ror.org/047mprv74

Henderson, Kathrine A. **113**
Kathrine Andrews Henderson is research analyst with LAC Group. She is part of a unique team of "virtual" researchers who provide "Library as a Service" to major law firms and corporations. Prior to this Ms. Henderson was the research librarian for the Office of the Auditor General for the State of Arizona. Earlier in her library career, Henderson was as an academic librarian. She was the Instructional Programs Librarian at Thunderbird School of Global Management and served in other roles including time as the business librarian for Arizona State University's Fletcher Library. Kathrine has expertise in business and legal research, intellectual property, and information ethics and has used this expertise to contribute to her field. Recently, she published a chapter on Intellectual Property Ethics in Foundations of Information Ethics, John Burgess and Emily Knox, editors. Other works include co-authoring Case Studies in Library and Information Science Ethics with Elizabeth Buchanan. In January 2018, Henderson was appointed to the Information Outlook Advisory Council for the Special Library Association. In the past, she served as Co-Director of the International Society for Ethics and Information Technology (INSEIT) and as an editor for ACM's Computers & Society. Henderson holds a Masters Degree in Library and Information Science from the University of Wisconsin-Milwaukee and a Bachelor of Science in Management from Arizona State University. In 2017, The School of Information Studies at UWM honored Kathrine as one of 50 Distinguished Alumni as part of the school's 50th Anniversary celebration.
ROR_ID https://ror.org/014jgh757
Email: kat.henderson@libsource.com

Author Information *CONTINUED*

Lipinski, Tomas A. **113**

Professor Lipinski completed his Juris Doctor (J.D.) from Marquette University Law School, Milwaukee, Wisconsin, received the Master of Laws (LL.M.) from The John Marshall Law School, Chicago, Illinois, and the Ph.D. from the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Dr. Lipinski has worked in a variety of legal settings including the private, public and non-profit sectors. He is the author of numerous articles and book chapters; his monographs include, THE LIBRARY'S LEGAL ANSWER BOOK co-authored with Mary Minow (2003); the COPYRIGHT LAW IN THE DISTANCE EDUCATION CLASSROOM (2005), THE COMPLETE COPYRIGHT LIABILITY HANDBOOK FOR LIBRARIANS AND EDUCATORS (2006), and THE LIBRARIAN'S LEGAL COMPANION FOR LICENSING INFORMATION RESOURCES AND SERVICES (2012). Recent articles and chapters include, Click Here to Cloud: End User Issues in Cloud Computing Terms of Service Agreements, in CHALLENGES OF INFORMATION MANAGEMENT BEYOND THE CLOUD: 4TH INTERNATIONAL SYMPOSIUM ON INFORMATION MANAGEMENT IN A CHANGING WORLD, IMCW 2013 (Revised Selected Papers.), with Kathrine Henderson, Hate Speech: Legal and Philosophical Aspects, in THE HANDBOOK OF INTELLECTUAL FREEDOM CONCEPTS (2014), in 2013 with Andrea Copeland, Look before you License: The Use of Public Sharing Websites in building Patron Initiated Public Library Repositories, PRESERVATION, DIGITAL TECHNOLOGY & CULTURE and in 2012, Law vs. Ethics, Conflict and Contrast in Laws Affecting the Role of Libraries, Schools and other Information Intermediaries, JOURNAL OF INFORMATION ETHICS. He has been a visiting professor in summers at the University of Pretoria-School of Information Technology (Pretoria, South Africa) and at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Professor Lipinski was the first named member of the Global Law Faculty, Faculty of Law, University of Leuven (Katholieke Universiteit Leuven), Belgium, in Fall of 2006 where he continues to lecture annually at its Centers for Intellectual Property Rights and Interdisciplinary Center for Law and ICT. He is active in copyright education and policy-making, chairing the ACRL Copyright Discussion Group, a member of the ALA OITP Committee on Legislation Copyright Subcommittee, a member of the Copyright and Other Legal Matters Committee of IFLA and serves as an IFLA delegate to the World Intellectual Property Organization's Standing Committee on Copyright and Other Rights. In October of 2014 he returned to the University of Wisconsin—Milwaukee to serve as Professor and Dean of its i-School, the School of Information Studies.

ROR_ID https://ror.org/031q21x57

Email tlipinsk@uwm.edu

Nicoll, Fiona **155**

Fiona Nicoll is the author of Gambling in Everyday Life: Spaces, Moments and Products of Enjoyment and is based in the Political Science department at the University of Alberta where she holds an Alberta Gambling Research Institute (AGRI) Chair in gambling policy. She is a co-editor of Critical Gambling Studies and the author of numerous book chapters and articles on reconciliation and Indigenous sovereignty, critical race and whiteness studies, queer theory and critical theory and pedagogy in the neo-liberal university.

ROR_ID https://ror.org/030ykd126

Savić, Dobrica **171**

Dr Dobrica Savić is Head of the Nuclear Information Section (NIS) of the IAEA. He holds a Doctorate degree from Middlesex University in London, an MPhil degree in Library and Information Science from Loughborough University, UK, an MA in International Relations from the University of Belgrade, Serbia, as well as a Graduate Diploma in Public Administration, Concordia University, Montreal, Canada. He has extensive experience in the management and operations of web, library, information and knowledge management, as well as records management and archives services across various United Nations agencies, including UNV, UNESCO, World Bank, ICAO, and the IAEA. His main interests are digital transformation, creativity, innovation and use of information technology in library and information services.

ROR_ID https://ror.org/00gtfax65

ORCID_ID https://orcid.org/0000-0003-1123-9693

Contact: www.linkedin.com/in/dobricasavic

Notes for Contributors

Non-Exclusive Rights Agreement

- I/We (the Author/s) hereby provide TextRelease (the Publisher) non-exclusive rights in print, digital, and electronic formats of the manuscript. In so doing,
- I/We allow TextRelease to act on my/our behalf to publish and distribute said work in whole or part provided all republications bear notice of its initial publication.
- I/We hereby state that this manuscript, including any tables, diagrams, or photographs does not infringe existing copyright agreements; and, thus indemnifies TextRelease against any such breach.
- I/We confer these rights without monetary compensation and with the understanding that TextRelease acts on behalf of the author/s.

Submission Requirements

Manuscripts should not exceed 15 double-spaced typed pages. The size of the page can be either A-4 or 8½x11 inches. Allow 4cm or 1½ inch from the top of each page. Provide the title, author(s) and affiliation(s) followed by your abstract, suggested keywords, and a brief biographical note.
A printout or PDF of the full text of your manuscript should be forwarded to the office of TextRelease. A corresponding MS Word file should either accompany the printed copy or be sent as an attachment by email. Both text and graphics are required in black and white.

REFERENCE GUIDELINES

General

i. All manuscripts should contain references
ii. Standardization should be maintained among the references provided
iii. The more complete and accurate a reference, the more guarantee of an article's content and subsequent review.

Specific

iv. Endnotes are preferred and should be numbered
v. Hyperlinks need the accompanying name of resource; a simple URL is not acceptable
vi. If the citation is to a corporate author, the acronym takes precedence
vii. If the document type is known, it should be stated at the close of a citation.
viii. If a citation is revised and refers to an edited and/or abridged work, the original source should also be mentioned.

Examples

Youngen, G.W. (1998), Citation patterns to traditional and electronic preprints in the published literature. - In: College & Research Libraries, 59 (5) Sep 1998, pp. 448-456. - ISSN 0010-0870

Crowe, J., G. Hodge, and D. Redmond (2010), Grey Literature Repositories: Tools for NGOs involved in public health activities in developing countries. – In: Grey Literature in Library and Information Studies, Chapter 13, pp. 199-214. – ISBN 978-3-598-11793-0

DCMI, Dublin Core Metadata Initiative Home Page http://purl.oclc.org/metadata/dublin_core/

Review Process

The Journal Editor first reviews each manuscript submitted. If the content is suited for publication and the submission requirements and guidelines complete, then the manuscript is sent to one or more Associate Editors for further review and comment. If the manuscript was previously published and there is no copyright infringement, then the Journal Editor could direct the manuscript straight away to the Technical Editor.

Journal Publication and Article Deposit

Once the journal article has completed the review process, it is scheduled for publication in The Grey Journal. If the Author indicated on the signed Rights Agreement that a preprint of the article be made available in GreyNet's Archive, then browsing and document delivery are immediately provided. Otherwise, this functionality is only available after the article's formal publication in the journal.

Contents
'Tracking the Ins and Outs of Grey Literature'