The Replication Crisis and IP Law: A Novel Policy Tool for Open Science

Or Cohen-Sasson* & Ofer Tur-Sinai**

In recent years, the scientific community has faced a considerable problem—the replication crisis. Replication is the process of verifying scientific findings by repeating a published study. It is considered a cornerstone of the scientific enterprise, contributing to the credibility of research findings. Over the past two decades, however, replication has become increasingly difficult; in fact, in some disciplines the non-replicability rate is over 50%. A major factor accounting for this is diminished access to research materials required for replication (replication materials). This problem is particularly acute in computational studies, where the code, software documentation, datasets, and other information are often not shared.

In this Article, we address the replication crisis from the perspective of intellectual property (IP) law. Our goal is twofold: first, to investigate the extent to which IP law plays a part in impeding access to replication materials; and second, to explore potential solutions that could minimize this detrimental effect. One branch of IP law that has been identified by scholars as having a potential adverse effect on the ability to conduct replication studies is copyright law. This Article, however, shows that the impact of copyright law is likely minor, whereas other IP regimes—patent and trade

---

* Or Cohen-Sasson, Research Fellow, Canon Foundation in Europe; Ph.D. Candidate, Zvi Meitar Center for Advanced Legal Studies, Faculty of Law, Tel Aviv University.

** Ofer Tur-Sinai, Associate Professor of Law, Ono Academic College; Adjunct Professor, Hebrew University of Jerusalem. For helpful comments, suggestions, and discussions, the authors are grateful to Clark Asay, Eran Bareket, Michael Birnhack, Kathy Bowrey, Oren Bracha, Jorge Contreras, Carys Craig, Severine Dussolier, Niva Elkin-Koren, Sean Flynn, Jeanne Fromer, Janet Freilich, Gabriel Galvez-Behar, Yarden Haltovsky, Timothy Holbrook, Ariel Katz, Asa Kling, Lisa Larrimore Ouellette, Peter Mennel, Gideon Parchomovsky, Jason Rantanen, Jerome Reichman, Josh Sarnoff, Arul George Scaria, Emily Schaffer, Jacob Sherkow, Michal Shur-Ofry, and Brenda Simon, as well as the participants at PatCon 10 (Boston College), workshop of the Global Expert Network on Copyright User Rights (American University), workshop of the International Society for the History and Theory of Intellectual Property (University of Gothenburg), and the 2022 Annual Conference of European Policy for Intellectual Property (EPIP) (Cambridge University). The authors gratefully acknowledge support from the American University Program on Information Justice and Intellectual Property and from the Institute for Humane Studies (grant no. IHS016557).

secret law—have a greater impact in this domain. We find that a major reason for scientists to avoid sharing replication materials is the fear that doing so will compromise their ability to secure patent and trade secret protection.

As a solution, this Article proposes the Conditional-Access-Agreement (CAA)—a novel policy tool that establishes a private and controlled channel of communication between authors and replicators. Authors would be able to provide access to replication materials on demand, through this channel, for the exclusive purpose of conducting replication studies. The CAA mechanism provides a win-win solution: facilitating access to replication materials without jeopardizing scientists' chances of obtaining IP protection.

INTRODUCTION

In recent decades, science has faced a considerable problem—*the replication crisis*.[1] "Replication" is the process of verifying scientific findings by repeating a previous study.[2] It enables scientists to substantiate or refute existing theories and has long been considered a core principle of the scientific method.[3] Replication also increases public trust in science, an essential function in itself.[4]

The term "replication crisis" refers to the difficulty of replicating published studies.[5] One of the factors accounting for the recent replicability crisis is restricted access to research-related materials.[6] A scientific paper alone often does not contain all of the information necessary to replicate the study's results, and original researchers (*authors*) frequently do not share such information with the scientific community.[7]

---

1. Tobias Wingen et al., *No Replication, No Trust? How Low Replicability Influences Trust in Psychology*, 11 SOC. PSYCH. & PERSONALITY SCI. 454 (2020); ADVISORY COMMITTEE TO THE NATIONAL SCIENCE FOUNDATION, SOCIAL, BEHAVIORAL, AND ECONOMIC SCIENCES PERSPECTIVES ON ROBUST AND RELIABLE SCIENCE 21–23 (2015); NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE, REPRODUCIBILITY AND REPLICABILITY IN SCIENCE 71–104 (2019) [hereinafter NASEM REPORT].

2. NASEM Report, *supra* note 1, at 43.

3. *See* KARL POPPER, THE LOGIC OF SCIENTIFIC DISCOVERY 22–24, 66 (2002) ("Only by such repetitions can we convince ourselves that we are not dealing with a mere isolated 'coincidence', but with events which, on account of their regularity and reproducibility, are in principle inter-subjectively testable."); JOHN STADDON, SCIENTIFIC METHOD: HOW SCIENCE WORKS, FAILS TO WORK, AND PRETENDS TO WORK 5–6, 27–28, 45–49 (2018); Daniel J. Simons, *The Value of Direct Replication*, 9 PERSPS. ON PSYCH. SCI. 76, 76–77 (2014).

4. *See* Tobias Wingen et al., *supra* note 1; Christopher J. Ferguson, *"Everybody Knows Psychology Is Not a Real Science": Public Perceptions of Psychology and How We Can Improve Our Relationship with Policymakers, the Scientific Community, and the General Public*, 70 AM. PSYCH. 527, 530–31 (2015); Michal Bialek, *Replications Can Cause Distorted Belief in Scientific Progress*, 41 BEHAV. & BRAIN SCIS. e122 (2018).

5. Ferguson, *supra* note 4, at 528–29.

6. *See Replication Crisis in Psychology: Overview, Causes & Examples*, STUDY.COM (Sept. 20, 2022), https://study.com/academy/lesson/replication-crisis-psychology-overview-causes-examples.html [https://perma.cc/T9LS-BRE6].

7. *See* Matthew Hutson, *Artificial Intelligence Faces Reproducibility Crisis*, 359 SCI. 725 (2018); Zeeya Merali, *Computational Science…Error*, 467 NATURE 775 (2010); Daniel Stockemer et al., *Data Access, Transparency, and Replication: New Insights from the Political Behavior Literature*, 51 POL. SCI. & POL. 799 (2018); R. Michael Alvarez et al., *Research Replication: Practical Considerations*, 51 POL. SCI. & POL. 422 (2018); Christine L. Borgman, *The Conundrum of Sharing Research Data*, 63 J. AM. SOC'Y FOR INFO. SCI. & TECH, 1059 (2012); Jelte Wicherts et al., *The Poor Availability of Psychological Research Data for Reanalysis*, 61 AM. PSYCH. 726 (2006) ("Seventy-three percent of the authors did not share their data."); Gary King, *Replication, Replication*, 28 POL. SCI. & POL. 444, 446 (1995).

Insufficient access to the materials needed for replication is a particularly acute problem in *computational studies*,[8] i.e., studies that harness computing capabilities to investigate complex phenomena in multiple disciplines, such as biology, medicine, engineering, and social science.[9] Replicating computational studies requires access to complete and accurate information regarding various elements used by the author.[10] Depending on the nature of the original study and the type of the validating study,[11] such elements may include the underlying data, code, software documentation, and other components (*replication materials*).[12] Unfortunately, access to these elements

8. *See* Victoria Stodden et al., *An Empirical Analysis of Journal Policy Effectiveness for Computational Reproducibility*, 115 PNAS 2584, 2588 (2018) [hereinafter Stodden et al., *Journal Policy*] (finding that replication materials could not be obtained for more than 55% of the sample studies, and successfully reproducing only 26% of all sample studies' results); Randall J. LeVeque et al., *Reproducible Research for Scientific Computing: Tools and Strategies for Changing the Culture*, 14 COMPUTING SCI. & ENG'G 13 (2012); Victoria Stodden et al., *Enabling the Verification of Computational Results: An Empirical Evaluation of Computational Reproducibility*, P-RECS'18: PROC. FIRST INT'L WORKSHOP ON PRAC. REPRODUCIBLE EVALUATION COMPUT. SYS. (2018) [hereinafter Stodden et al., *Verification of Computational Results*] (finding that 67% of the sample studies could not be easily replicated, and finding that "the main barriers to computational reproducibility are inadequate documentation of code, data, and workflow information (70.9%), missing code function and setting information, and missing licensing information (75%)"); Merali, *supra* note 7, at 775; Hutson, *supra* note 7, at 725–26 (finding that only 6% of the sample articles included code, circa 30% included test data, and about 54% included pseudo-code); Gregorio Robles, *Replicating MSR: A Study of the Potential Replicability of Papers Published in the Mining Software Repositories Proceedings*, 7TH IEEE WORKING CONF. ON MINING SOFTWARE REPOSITORIES 171, 174–77 (2010) (finding that processed dataset, code, and tools are rarely shared, even when authors explicitly state they built such tools).

9. *Computational Science*, NATURE PORTFOLIO, https://www.nature.com/subjects/computational-science [https://perma.cc/3DAM-74TX]; *see* Denis Noble, *The Rise of Computational Biology*, 3 NATURE REV. MOLECULAR CELL BIOLOGY 459 (2002); *see also* Norman P. Hummon & Thomas J. Fararo, *The Emergence of Computational Sociology*, 20 J. MATHEMATICAL SOCIO. 79 (1995).

10. *See* David Donoho et al., *Reproducible Research in Computational Harmonic Analysis*, 11 COMPUTING SCI. & ENG'G 8, 9 (2009); Xiaoli Chen et al., *Open is Not Enough*, 15 NATURE PHYSICS 113, 113 (2019); Harold Thimbleby, *Explaining Code for Publication*, 33 SOFTWARE, PRAC. & EXPERIENCE 975, 977–79 (2003); Hutson, *supra* note 7, at 725–26; B. D. McCullough et al., *Lessons from the JMCB Archive*, 38 J. MONEY, CREDIT & BANKING 1093 (2006); Steve M. Easterbrook, *Open Code for Open Science?*, 7 NATURE GEOSCIENCE 779, 780 (2014).

11. For a discussion regarding the two main types of validating studies—reproductions and replications—see *infra* notes 32–34 and accompanying text.

12. *See supra* note 10. Particularly, *see* Chen et al., *supra* note 10; Easterbrook, *supra* note 10; Donoho et al., *supra* note 10 (explaining that conducting replication studies requires the full computational environment that produced the original results).

is far from guaranteed, which makes some of the studies difficult or impossible to replicate.[13]

This Article addresses the replication crisis in computational studies from the perspective of intellectual property (IP) law. Our goal in this Article is twofold: first, to investigate the extent to which IP law and policy play a part in impeding access to replication materials; and second, to explore possible solutions that could minimize this detrimental effect.

To appreciate the impact of IP law on replicability, one must first study the factors that account for the wide phenomenon of restricted access to replication materials. As we elaborate below, commercial and proprietary concerns often play a major role, alongside other considerations, in incentivizing researchers (or the organizations they are affiliated with) to avoid sharing materials they are not required to disclose. The desire to secure IP protection for various aspects arising out of the research may naturally strengthen this incentive.[14]

One branch of IP law relevant to the discussion of IP and replicability is copyright law. Various components of a computational study can be protected by copyright.[15] Nevertheless, our discussion below demonstrates that copyright's role in hindering replicability is not as significant as it may seem. In short, to the extent replication materials are not publicly accessible, they are not susceptible to copying, making copyright law largely irrelevant; while if the materials are accessible and the only concern is copyright liability, existing policy levers (e.g., fair use [16]) could alleviate such concerns to a large extent.

By contrast, patent law and trade secret law may have a significant impact on access to replication materials. Patent law can protect inventions that are developed in the course of scientific research. As a condition for patent protection, however, the invention must be novel and non-obvious, i.e., it must not have been publicly disclosed and must be sufficiently inventive.[17] Accordingly, a risk averse researcher (or affiliated organization) that publishes a scientific study would be disinclined to share supplementary

---

13. *See supra* note 8; Babatunde Kazeem Olorisade, Pearl Brerton, & Peter Andras, Reproducibility *of Studies on Text Mining for Citation Screening in Systematic Reviews: Evaluation and Checklist*, 73 J. BIOMEDICAL INFORMATICS 1, 1-13 (2017), https://www.sciencedirect.com/science/article/pii/S1532046417301661 [https://perma.cc/UW4X-DJBV]; Heidi Seibold et al., *A Computational Reproducibility Study of PLOS ONE Articles Featuring Longitudinal Data Analyses,* 17 PLOS One 5 (2021) https://doi.org/10.1371/journal.pone.0269047 [https://perma.cc/ZHM6-G5XB].

14. *See infra* Part III.

15. *See infra* notes 125–128 and accompanying text.

16. *See* 17 U.S.C. § 107.

17. 35 U.S.C. §§ 102–103.

information (including replication materials) to maximize chances of obtaining patent protection.

Researchers may also wish to claim trade secret protection for replication materials as an additional or alternative measure to guard against misappropriation. Notably, all replication materials could be protected as trade secrets, including algorithms, code, and data. Yet, to qualify for protection, the information must not be in the public domain, and the rightful holder of the information must take reasonable measures to keep it secret.[18] Thus, again, the law may strengthen researchers' incentive to avoid disclosing information in order to secure a competitive advantage in the market.

Our analysis of patent and trade secret law reveals a disturbing phenomenon: These legal regimes encourage the concealment of replication materials. Given the high societal interest in replications, a natural inclination of scholars and policy makers could be to consider limiting the scope of IP rights.[19] We argue, however, that narrowing IP protection is not imperative to enable replicability. This Article proposes a solution that rejects a binary view—either IP rights or replicability—and suggests an inclusive approach, maintaining that IP rights and replicability can coexist.

Concretely, we propose a new policy tool: the *Conditional Access Agreement* (CAA). When submitting a paper for publication, an author would execute a CAA with the journal, pledging to provide—upon a replicator's request—full access to replication materials. The CAA would specify that any replicator seeking access to the materials must sign a *Non-Disclosure Agreement* (NDA) to obtain them. The NDA would prohibit disclosure or use of the information for any purpose other than replication. We consider below various mechanisms to encourage the implementation of this policy.

The CAA mechanism would establish private—not public—access to replication materials, conditioned upon and protected by an NDA. As the materials would not be released to the public domain, their disclosure would not negate patent or trade secret protection. The CAA would thus facilitate

---

18. *Trade Secrets*, WIPO, https://www.wipo.int/tradesecrets/en/ [https://perma.cc/97XY-SF4H].

19. *See, e.g.*, Victoria Stodden, *Enabling Reproducible Research: Licensing Scientific Innovation*, 13 INT'L J. COMM. L. & POL'Y 1, 22 (2009) [hereinafter Stodden, *Enabling Reproducible Research*] (suggesting that researchers will license replication materials in advance under permissive terms; for example, licensing code under the Apache 2.0 or the MIT license, which allow commercial use); Victoria Stodden, *Intellectual Property and Computational Science*, *in* OPENING SCIENCE: THE EVOLVING GUIDE ON HOW THE INTERNET IS CHANGING RESEARCH, COLLABORATION AND SCHOLARLY PUBLISHING 225, 231 (Bartling & Friesike eds., 2014) [hereinafter Stodden, *IP*] (suggesting to license copyright-protected data under the Creative Commons license, which allows, *inter alia*, commercial use); Steven Shavell, *Should Copyright of Academic Works Be Abolished?*, 2 J. LEGAL ANALYSIS 301 (2010) (proposing to abolish copyright in scientific publications, contending that it will generate Open Access).

replicability by granting limited, non-public access to replication materials on demand without jeopardizing the proprietary nature or the commercial potential of the study. While not a panacea for the replication crisis, implementing the proposed CAA policy may remove a major barrier to replicability.

Vast non-legal scholarship has investigated reasons for the replication crisis and has offered ways to mitigate it through organizational, cultural, and scientific measures.[20] Yet, there is sparse literature regarding the role that the legal system—particularly IP law—plays in this domain.[21] A few scholars have explored questions situated at the interface between IP law and replication studies. Of them, only a handful of legal scholars have touched upon the problem of access to materials needed to replicate scientific studies on which we focus in this Article.[22] A particularly notable contribution to the literature in this domain is a series of papers by Victoria Stodden, who has focused on the replication crisis in computational studies and investigated certain areas in which IP law may pose a barrier to access to replication materials.[23] The primary development Stodden advocates is to encourage more replications-friendly behavior by IP owners.[24] A different dimension of

---

20. *See, e.g.*, Jens B. Asendorpf et al., *Recommendations for Increasing Replicability in Psychology*, *in* METHODOLOGICAL ISSUES AND STRATEGIES IN CLINICAL RESEARCH 607 (Alan E. Kazdin ed., 2016); Michèle B Nuijten, *Practical Tools and Strategies for Researchers to Increase Replicability*, 61 DEVELOPMENTAL MED. & CHILD NEUROLOGY 535 (2019); Luke Plonsky, *Quantitative Considerations for Improving Replicability in CALL and Applied Linguistics*, 32 CALICO J. 232 (2015).

21. For notable exceptions, see Michael W. Carroll, *Sharing Research Data and Intellectual Property Law: A Primer*, 13 PLOS BIOLOGY e1002235 (2015); Michal Shur-Ofry, *Access-to-Error*, 34 CARDOZO ARTS & ENT. L.J. 357 (2016); Stodden, *IP*, *supra* note 19.

22. *See, e.g.*, Stodden, *IP*, *supra* note 19 (focusing primarily on access to code used in computational studies); Carroll, *supra* note 21 (focusing on open repositories of research data, where in the absence of clear permission, users may find themselves infringing intellectual property rights).

23. *See, e.g.*, Stodden, *IP*, *supra* note 19; Stodden, *Enabling Reproducible Research*, *supra* note 19; Victoria Stodden et al., *Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals*, 8 PLOS ONE e67111 (2013) [hereinafter Stodden et al., *Toward Reproducible Computational Research*]; Stodden et al., *Journal Policy*, *supra* note 8; LeVeque et al., *supra* note 8; Stodden et al., *Verification of Computational Results*, *supra* note 8; Victoria Stodden & Isabel Reich, *Software Patents as a Barrier to Scientific Transparency: An Unexpected Consequence of Bayh-Dole*, (presented at the "From Data to Solutions" IGERT Seminar, Columbia University, Nov. 13, 2012) (available at https://academiccommons.columbia.edu/doi/10.7916/D86Q25HP [https://perma.cc/K4KE-MEM3]).

24. *See, e.g.*, Stodden, *IP*, *supra* note 19 (recommending that authors publish data under Creative Commons licenses; suggesting that copyright in publications remain in the hands of authors, assuming they would allow open access, rather than be granted to journals, and encouraging the use by patent owners of licenses that enable uses of claimed inventions for purposes of research validation).

the interface between IP law and replication studies that has been addressed in the literature, most prominently in the work of Michal Shur-Ofry, is the low incentives that IP systems and other incentive mechanisms provide for engaging in replication studies and disseminating their results.[25] Another relevant aspect explored by scholars is the difficulty of replicating experiments that are described in patent documents.[26]

While acknowledging the importance of these aspects that have been discussed in the literature, our paper seeks to deepen the inquiry into the role that IP law plays in hindering access to replication materials. One contribution of this Article to existing scholarship is, first, in identifying and mapping all points of friction between IP law and replications that may have an impact on this matter. Our analysis reveals interesting findings: in contrast to the common view,[27] we demonstrate that copyright law does not have a significant effect on access to replication materials. Yet, the impact of patent law and trade secret law on sharing replication materials, which has not been sufficiently studied so far, is significant. To recognize this impact, one must consider not only the ex-post effect of IP rights on access, as is typically done, but also the ex-ante effect, i.e., how the desire to secure patent protection affects patent applicants' behavior.[28] As far as we know, we are the first to conduct a systematic examination of this important issue, while drawing on the available literature (both legal and scientific) and on a detailed analysis of the relevant statutes and case law. Another major contribution of the Article is our proposal of a novel policy tool to mitigate the replicability crisis—the Conditional Access Agreement (CAA)—designed to alleviate IP-related barriers to access to replication materials.

---

25. Shur-Ofry, *supra* note 21, at 370–79 (2016) (discussing the low IP incentives for the exposure and diffusion of negative knowledge, the bias against publishing negative findings in scientific journals, and the strong preference of funding schemes for research proposals that aim to produce positive results over research that aims to replicate or refute existing findings); *see also* W. Nicholson Price II, *The Cost of Novelty*, 120 COLUM. L. REV. 769, 778 (2020) (discussing the incentives provided by patent law "to pursue the divergent path" despite the social benefits entailed in "deepening innovation" by replicating existing studies).

26. *See, e.g.*, Janet Freilich, *The Replicability Crisis in Patent Law*, 95 IND. L. REV. 43, 456–58 (2020) (finding that experiments in patents have very poor methodological quality, which means that they are likely irreplicable at rates at least as high as experiments in scientific journals); Jacob S. Sherkow, *Patent Law's Reproducibility Paradox*, 66 DUKE L.J. 845 (2017) (maintaining that patents grounded in irreproducible research fail their constitutional bargain of property rights in exchange for working disclosures of inventions).

27. Stodden, *Enabling Reproducible Research*, *supra* note 19, at 1 ("[C]opyright law is a barrier to the sharing of scientific scholarship . . . ."); Stodden, *IP*, *supra* note 19 (arguing that copyright detrimentally affects the sharing of replication materials).

28. *But see* Stodden & Reich, *supra* note 23 (noting this potential effect on patent law); Carroll, *supra* note 21 (mentioning that data sharing prior to filing a patent application may destroy or impair patent protection).

The Article proceeds as follows. Part I provides further background with respect to replicability and the replicability crisis, focusing on the phenomenon of restricted access to replication materials. Part II examines the interface between copyright law and the replicability crisis. Part III discusses patents and trade secrets and demonstrates how these branches of IP law bolster incentives to restrict access to replication materials. Part IV presents our proposed solution to facilitate access to replication materials—the CAA—and discusses its merits and shortcomings.

I. THE REPLICABILITY CRISIS

This Part provides some necessary background for the ensuing discussion. First, we discuss replicability and its importance in science. Then, we explain the replicability crisis while focusing on a particular aspect of it—insufficient access to replication materials. We outline the most common factors that lead to the phenomenon of insufficient access and thus to non-replicability. This inquiry is essential to understanding the function that IP law may play in this domain. The last section of this Part reviews some of the initiatives the scientific community has taken to address the replication crisis.

A. Replicability & the Replicability Crisis

Replication is the process of verifying scientific findings. A scientific study is valid only if others can replicate its results. By replicating studies, scientists can substantiate—or, alternatively, refute—published studies, allowing the scientific community to advance certain theories and findings and eliminate others.[29] Science, after all, is built upon cumulative progress, i.e., it advances by relying on previous discoveries and insights. As Bernard of Chartres put it (and as later rephrased by Newton), we are "standing on the shoulders of giants."[30] Therefore, replication is justifiably considered a fundamental principle of the scientific method.[31]

To be precise, there are two main types of secondary studies that aim to confirm results and ensure the consistency of scientific findings:

29. *See* Shur-Ofry, *supra* note 21, at 364–70 (discussing, inter alia, the role of refutations in steering innovation efforts away from dead ends and guiding it toward viable solutions and the potential of exposed errors to spark paradigm shifts).

30. Originally in Latin: "nanos gigantum humeris insidentes." *E.g.*, Lodoen et al., *Nanos Gigantium Humeris Insidentes: Old Papers Informing New Research into* Toxoplasma Gondii, 51 INT'L J. FOR PARASITOLOGY 1193, 1193 (2021); *see also* Letter from Sir Isaac Newton to Robert Hooke (Feb. 5, 1675) (on file with the Historical Society of Pennsylvania and available at https://digitallibrary.hsp.org/index.php/Detail/objects/9792 [https://perma.cc/5JGC-T9WL]).

31. *See supra* note 3.

"replications" and "reproductions." These terms have no uniform definitions, however, and they are often used interchangeably. In connection with computational studies, a recent report by the National Academies of Sciences, Engineering, and Medicine defines "reproducibility" as the ability to obtain consistent computational results using the same input data, computational steps, methods, code, and conditions of analysis.[32] "Replicability," by comparison, is defined as the ability to confirm previous results by obtaining new data and applying the original methods to such data to see if the results are consistent across studies.[33] In this Article, we use the term "replications" to refer to both types of secondary studies without delving into fine distinctions between them.[34]

In theory, scientific studies should yield highly reliable results, i.e., the replicability rate should be very high. Failure (or inability) to verify previous studies undermines their credibility and calls into question all studies relying on them. In practice, however, scientists encounter great difficulties when attempting to replicate published studies.[35] The literature and the media commonly refer to this crucial problem as the *replication crisis*.[36]

Scientists have expressed concerns regarding replicability since the 17th century.[37] However, it is only in recent decades that such concerns have become a central debate within the scientific and public discourse. Since the early 2000's, various studies have pointed out problems with the replicability of scientific studies that are serious enough to warrant the term "crisis."[38] Researchers first revealed the magnitude of the crisis primarily in medicine and psychology, where efforts to verify results from both classic and cutting-

32. NASEM REPORT, *supra* note 1.

33. *Id.*

34. *See* Chen et al., *supra* note 10, at 114 (providing other terminologies for referring to secondary studies).

35. *See* NASEM REPORT, *supra* note 1; John P. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLOS MED. 124 (2005); Monya Baker, *1,500 Scientists Lift the Lid on Reproducibility*, 533 NATURE 452 (2016); Jonah Lehrer, *The Truth Wears Off*, THE NEW YORKER (Dec. 5, 2010), https://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off [https://perma.cc/M3UX-T2Q4].

36. *See, e.g.*, NASEM REPORT, *supra* note 1; Ioannidis, *supra* note 35; Baker, *supra* note 35; Lehrer, *supra* note 35.

37. *See* STEVEN SHAPIN & SIMON SCHAFFER, LEVIATHAN AND THE AIR-PUMP (1985).

38. *See, e.g.*, Harold Pashler & Eric–Jan Wagenmakers, *Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?*, 7 PERSPS. ON PSYCH. SCI. 528 (2012). Of course, the principal issue had been pointed at earlier. *See, e.g.*, Nathaniel C. Smith, Jr., *Replication Studies: A Neglected Aspect of Psychological Research*, 25 AM. PSYCH. 970 (1970).

edge research have failed.[39] Studies have found that the average replicability rate in psychology is a mere 36%.[40] When examined by journal, the rate of replicability ranges from 23% to 62%.[41]

The situation in medical publications is similar.[42] There are two major studies with particularly disturbing findings: First, Ioannidis examined 49 highly-cited medical studies published in top journals and found that only 44% of them were replicable.[43] The second is a study of the Reproducibility Project, which aimed to replicate experiments from top cancer publications and found that only a quarter of them are replicable.[44] Gradually, researchers have found evidence of a replication crisis in almost all scientific fields, including economics, chemistry, engineering, and physics.[45]

---

39. Ioannidis, *supra* note 35 (arguing that a large portion of medical publications cannot be replicated); Joseph P. Simmons et al., *False-positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant*, 22 PSYCH. SCI. 1359 (2011) (showing that the flexibility in methodologies increases false-positive rates, causing researchers to find evidence for effects that do not exist); Bradford J. Wiggins & Cody D. Christopherson, *The Replication Crisis in Psychology: An Overview for Theoretical and Philosophical Psychology*, 39 J. THEORETICAL & PHIL. PSYCH. 202 (2019) (providing an overview of the replication crisis, address theoretical and philosophical issues arising from it).

40. Open Science Collaboration, *Estimating the Reproducibility of Psychological Science*, 349 SCI. 943 (2015) (finding that 53%–61% of top 2008 publication experiments are non-replicable).

41. *Id.*; Colin F. Camerer et al., *Evaluating the Replicability of Social Science Experiments in Nature and Science Between 2010 and 2015*, 2 NATURE HUM. BEHAV. 637 (2018) (attempting to replicate twenty-one experimental studies in the social sciences published in Nature and Science in 2010–2015, and succeeding in 67% of the cases); Richard A. Klein et al., *Many Labs 2: Investigating Variation in Replicability Across Samples and Settings*, 1 ADVANCES METHODS & PRACS. PSYCH. SCI. 443 (2018) (conducting replication studies for 28 classic and contemporary published findings and finding that only 54% are replicable).

42. Florian Prinz et al., *Believe It or Not: How Much Can We Rely on Published Data on Potential Drug Targets?*, 10 NATURE REVS. DRUG DISCOVERY 712 (2011) (addressing the significant problem with replications in drug discovery).

43. John A. Ioannidis, *Contradicted and Initially Stronger Effects in Highly Cited Clinical Research*, 294 JAMA 218 (2005).

44. Tara Haelle, *A Massive 8-Year Effort Finds that Much Cancer Research Can't Be Replicated*, SCI. NEWS (Dec. 7, 2021, 8:00 AM), https://www.sciencenews.org/article/cancer-biology-studies-research-replication-reproducibility [https://perma.cc/5JZS-QQXL]; *Reproducibility Project Cancer Biology*, CTR. FOR OPEN SCI., https://www.cos.io/rpcb [https://perma.cc/84H9-76BR].

45. Baker, *supra* note 35; Lehrer, *supra* note 35; Joel Achenbach, *No, Science's Reproducibility Problem Is Not Limited to Psychology*, WASH. POST (Aug. 28, 2015, 2:31 PM), https://www.washingtonpost.com/news/speaking-of-science/wp/2015/08/28/no-sciences-reproducibility-problem-is-not-limited-to-psychology [https://perma.cc/TM7Q-JBHK]; Andrew Gelman, *The Experiments Are Fascinating. But Nobody Can Repeat Them.*, N.Y. TIMES (Nov. 19, 2018), https://www.nytimes.com/2018/11/19/science/science-research-fraud-reproducibility.html [https://perma.cc/Q7SB-CF9H].

With the recognition of the replication crisis in multiple disciplines, the discourse about the crisis has intensified, drawn much attention from the scientific community, and generated immense concerns.[46] Notably, the replication crisis is a severe, systematic malfunction, not a marginal phenomenon or sporadic failure.[47] It is troubling to think that a huge portion of scientific publications cannot be trusted and does not comply with the scientific method. Indeed, the replication crisis casts a dark shadow on the credibility of scientific publications and is one of the top problems in science today.[48]

One can divide the replication crisis into two separate sub-problems: (1) Inconsistent Results – cases in which replications of a study yield different results than the original study; and (2) Restricted Access to Replication Materials – cases in which conducting a proper replication is not possible due to no or limited access to *replication materials*.[49] Replication materials are the elements required for running a replication study. These include the original experiment's parameters and protocols, which commonly are not published. When replication materials are not available, one cannot easily investigate whether a publication is reliable. As a result, others may rely on findings that cannot be proven or disproven, some of which might be flawed.

This Article focuses on the second sub-problem—access to replication materials—as this is the context in which IP law may have significant impact.[50] The inconsistent results scenario is generally unrelated to IP law and is due instead to other problems such as low scientific skills or

---

46. Ramal Moonesinghe et al., *Most Published Research Findings Are False—But a Little Replication Goes a Long Way*, 4 PLOS MED. 0218 (2007); Simons, *supra* note 3; Francis Collins, Editorial, *Researching the Researchers*, 46 NATURE GENETICS 417 (2014).

47. NASEM REPORT, *supra* note 1; *supra* notes 39–46.

48. David Peterson & Aaron Panofsky, *Metascience as a Scientific Social Movement*, SOCARXIV (Aug. 4, 2020), https://osf.io/preprints/socarxiv/4dsqa/ [https://perma.cc/D94Z-RTQ8]; Stefan Schmidt, *Shall We Really Do It Again? The Powerful Concept of Replication Is Neglected in the Social Sciences*, 13 REV. GENERAL PSYCH. 90 (2016); Jonathan W. Schooler, *Metascience Could Rescue the 'Replication Crisis,'* 515 NATURE 9 (2014).

49. Carol Tenopir et al., *Data Sharing by Scientists: Practices and Perceptions*, 6 PLOS ONE e21101 (2011) (finding that 46% of scientists report they do not make their data electronically available to others); Benedikt Fecher et al., *A Reputation Economy: Results from an Empirical Survey on Academic Data Sharing* 9, 19 (DIW Berlin, Discussion Paper No. 1454, 2015), http://hdl.handle.net/10419/107687 [https://perma.cc/Z7BU-7X82]; Eric G. Campbell & Eran Bendavid, *Data-Sharing and Data-Withholding in Genetics and the Life Sciences: Results of a National Survey of Technology Transfer Officers*, 6 J. HEALTH CARE L. & POL'Y 241 (2003); Darren E. Zinner et al., *The Changing Nature of Scientific Sharing and Withholding in Academic Life Sciences Research: Trends from National Surveys in 2000 and 2013*, 91 ACAD. MED. 433, 439–43 (2016).

50. *See infra* Section I.B.

misconduct—serious issues that are beyond the scope of this Article.[51] In our discussion, we focus primarily on restricted access to replication materials in connection with *computational studies*. In computational studies, researchers design, implement, and utilize mathematical models to address scientific problems, often using an immense amount of data. Typically, the researchers set up a computational model to express the research question in numeric terms, regardless of the original discipline.[52] In theory, computational studies should excel in replicability, since digital simulations are unaffected by many factors that greatly influence lab experiments.[53] Yet, the replication crisis plagues computational studies as well.[54]

One may think that code and data alone are sufficient to allow for the replication of a computational study, but that view is mistaken.[55] *Replicators*, i.e., researchers who are interested in conducting a replication study, often need more than the code.[56] In computational studies, the pertinent replication materials may also include the algorithm, code documentation, datasets, machine details, workflows, run time environment information, and other know-how that underlies the study.[57]

---

51. *See, e.g.*, Michael Farthing, *Coping with Fraud*, 352 LANCET 11 (1998).

52. *Computational Science*, *supra* note 9.

53. Sharon M. Crook et al., *Learning from the Past: Approaches for Reproducibility in Computational Neuroscience*, *in* 20 YEARS OF COMPUTATIONAL NEUROSCIENCE 73, 73–74 (James M. Bower ed., 2013).

54. Stodden et al., *Verification of Computational Results*, *supra* note 8; Stodden et al., *Toward Reproducible Computational Research*, *supra* note 23; Stodden et al., *Journal Policy*, *supra* note 8; Matthew Hutson, *Missing Data Hinder Replication of Artificial Intelligence Studies*, SCI. (Feb. 15, 2018), https://www.science.org/content/article/missing-data-hinder-replication-artificial-intelligence-studies [https://perma.cc/CHA8-9XUU]; Roger D. Peng, *Reproducible Research in Computational Science*, 334 SCI. 1226 (2011); Nicolas P. Rougier et al., *Sustainable Computational Science: The ReScience Initiative*, 3 PEERJ COMPUT. SCI. e142 (2017); Marcin Miłkowski et al., *Replicability or Reproducibility? On the Replication Crisis in Computational Neuroscience and Sharing Only Relevant Detail*, 45 J. COMPUTATIONAL NEUROSCIENCE 163, 165–67 (2018); Benjamin Haibe-Kains et al., *Transparency and Reproducibility in Artificial Intelligence*, 586 NATURE e14 (2020); Krishna Tiwari et al., *Reproducibility in Systems Biology Modelling*, 17 MOLECULAR SYS. BIOLOGY e9982 (2021) (finding that over 50% of published computational models of physiological processes were irreproducible).

55. *See supra* note 10; Robles, *supra* note 8; Piet Hut, *Virtual Laboratories*, 164 PROGRESS THEORETICAL PHYSICS SUPPLEMENT 38, 39 (2006); Konrad Hinsen, *ActivePapers: A Platform for Publishing and Archiving Computer-aided Research*, 3 F1000 RSCH. 1, 4–6 (2014); William A. Ingram & Edward A. Fox, *Preparing Code and Data for Reproducible Publication: A Hands-on Tutorial*, JOINT CONF. ON DIGIT. LIBRARIES 2020 (Aug. 3, 2020), https://fox.cs.vt.edu/talks/2020/JCDL%202020%20Reproducibility%20Tutorial.pdf [https://perma.cc/QW6P-CF3A].

56. *See supra* note 10.

57. Chen et al., *supra* note 10, at 114; Easterbrook, *supra* note 10; Donoho et al., *supra* note 10, at 4; Muhamad Fitra Kacamarga et al., *Lightweight Virtualization in Cloud Computing for*

We decided to focus on computational studies as they are becoming increasingly prevalent across almost all scientific disciplines due to the rise of machine learning and advancements in computing power.[58] Indeed, considering both software and hardware developments, it is reasonable to assume that computational studies will become even more popular in the coming years. The problem of restricted access to replication materials in the context of computational studies is acute.[59] Accordingly, addressing the issue is critical.

Nevertheless, this Article's contribution is not limited to computational studies. The analysis of the interplay between IP law and replicability is also relevant to other types of scientific analyses that involve IP-eligible subject matter. Furthermore, our policy recommendations could be applied, with minor tweaks and modifications, to other disciplines, methodologies, and types of replication materials as well.[60]

B. Restricted Access to Replication Materials

In contrast to what many may believe, providing access to the code, database, and other information used in the course of a scientific study is not generally a requirement for publication in scientific journals.[61] Many times, the journals themselves do not have access to all research materials, even materials that are crucial for verifying the study's results.[62] Moreover, even when there is a requirement to share research materials, there are multiple exceptions—many of which are very broad—that provide researchers with a convenient bypass.[63]

The following real-life scenarios illustrate how IP regimes can restrict access to replication materials. Joaquim, a self-reported researcher at INESC-ID, posted on ResearchGate—one of the largest academic social

---

Research, *in* 516 COMMUNICATIONS IN COMPUTER AND INFORMATION SCIENCE 439, 439–45 (Rolly Intan et al. eds., 2015); Andy Cockburn et al., *Threats of a Replication Crisis in Empirical Computer Science*, 63 COMMC'NS ACM 70, 78 (2020).

58. *See* Noble, *supra* note 9.

59. Stodden et al., *Verification of Computational Results*, *supra* note 8.

60. *See infra* Part IV.

61. Antti M. Rousi & Mikael Laakso, *Journal Research Data Sharing Policies: A Study of Highly-Cited Journals in Neuroscience, Physics, and Operations Research*, 124 SCIENTOMETRICS 131, 132 (2020).

62. *Id.*

63. For instance, see exceptions in Wiley's data sharing policies: *Wiley's Data Sharing Policies*, WILEY, https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html [https://perma.cc/U7CC-45NM].

networks[64]—the question "Will all Computer Graphics publications include source code in the future?"[65] Here is one reply: "Submitting the source code will waste original authors' efforts and . . . prevent original authors from possible commercialization or co-authorship [in] any future steps based on their work."[66] Another researcher supported the idea of sharing code, but acknowledged that "there is the issue . . . of protecting the non-trivial work inherent in creating functional code . . . It is predicated on the idea that the work of developing the code goes unrewarded if the code is released."[67]
Jason posted a similar inquiry on Quora: "Why don't academic papers include code?"[68] Will, a user who self-identified as a data scientist, contended that sharing a code (and *not* pseudocode[69]) "should certainly be a requirement."[70] Interestingly, however, Will noted that such a sharing requirement "breaks down in the cases where the computational system that has been invented and is being described is also proprietary."[71]

---

64. Richard Van Noorden, *Online Collaboration: Scientists and the Social Network*, 512 NATURE 126, 126 (2014).

65. Joaquim Armando Pires Jorge, *Will All Computer Graphics Publications Include Source Code in the Future?*, RESEARCHGATE (May 16, 2013), https://www.researchgate.net/post/Will_all_Computer_Graphics_Publications_include_source_code_in_the_future [https://perma.cc/DGQ2-PJPP].

66. Manal Ezzat Helal, Comment to *Will All Computer Graphics Publications Include Source Code in the Future?*, RESEARCHGATE (July 27, 2017), https://www.researchgate.net/post/Will_all_Computer_Graphics_Publications_include_source_code_in_the_future [https://perma.cc/X3WN-LTQV].

67. Sean Curtis, Comment to *Will All Computer Graphics Publications Include Source Code in the Future?*, RESEARCHGATE (May 22, 2013), https://www.researchgate.net/post/Will_all_Computer_Graphics_Publications_include_source_code_in_the_future [https://perma.cc/C4PG-G722].

68. It is worth noting that some of the discussion in this section relies on information from online forums. As such, some of the information, such as users' job title, is self-reported and may not be independently verified. Jack Rae, Comment to *Why Don't Academic Papers Include Code?*, QUORA, https://www.quora.com/Why-dont-academic-papers-include-code [https://perma.cc/FC7L-V4HY].

69. *Pseudocode* is an English-like description of an algorithm. It is intended for human reading rather than machine reading. Typically, pseudocode lacks details that are vital for machine understanding. In scientific publications, it is very common to document algorithms using pseudocode. *See* ELLIS HOROWITZ ET AL., COMPUTER ALGORITHMS 5–10 (1997); G. MICHAEL SCHNEIDER, INVITATION TO COMPUTER SCIENCE 44–50 (2018); STEVE MCCONNELL, CODE COMPLETE 218–20 (2004).

70. Will Lamond, Comment to *Why Don't Academic Papers Include Code?*, QUORA, https://www.quora.com/Why-dont-academic-papers-include-code [https://perma.cc/FC7L-V4HY]. Additionally, Will reinforces the argument made before that the code and data are not sufficient for running a replication study: "This [i.e., the necessary replication materials] includes machine details, source code, and any other run time environment details that affect the computation." *Id.*

71. *Id.*

Some participants in these blogs justified the non-publication of code by mentioning the costs of code writing or the relatively low income of researchers in academia. In a Quora post entitled, "Why are many academic papers in Computer Science accepted even though the authors do not provide the code?",[72] one of the participants explained: "Code immediately gets into the issue of copyright, ownership, etc. In most cases, implementing theories . . . involving lots of people, time, money, equipment, etc. and typically for a specific commercial purpose."[73] Another scientist wrote,

> [S]cientists get paid less than half what those same people could make in industry. In science, though, you are allowed to translate your findings, make them into a product and get royalties from commercializing it. Many universities have massive offices which will help you patent your idea, file for FDA clearance and even make your startup company. So, while people do publish the general idea behind the algorithm, the code itself is a trade secret.

The non-sharing practice is not incidental. Even when researchers actively pursue unpublished replication materials and solicit them directly from authors, sharing is far from guaranteed.[74] One researcher posted on Academia Stack Exchange the following question:

> It is often the case when I am reading a paper I start to wonder, *"Wow stunning results, however, I would like to prove that*." . . . As you can see, a lot of times the small things have a huge impact on the overall performance of the underlying methodology. Often they are not part of the paper or not revealed at all. My idea would be to contact the paper

---

72. *Why Are Many Academic Papers in Computer Science Accepted Even Though the Authors Do Not Provide the Code?*, QUORA, https://www.quora.com/Why-are-many-academic-papers-in-Computer-Science-accepted-even-though-the-authors-do-not-provide-the-code [https://perma.cc/7L3R-574J].

73. Brett Watters, Comment to *Why Are Many Academic Papers in Computer Science Accepted Even Though the Authors Do Not Provide the Code?*, QUORA, https://www.quora.com/Why-are-many-academic-papers-in-Computer-Science-accepted-even-though-the-authors-do-not-provide-the-code [https://perma.cc/7L3R-574J].

74. Andrew J. Nelson, *How to Share "A Really Good Secret": Managing Sharing/Secrecy Tensions Around Scientific Knowledge Disclosure*, 27 ORG. SCI. 265, 265–70 (2016). The phenomenon of authors refusing to share materials or simply not responding is not limited to computational studies. *See* Carolin Haeussler et al., *Specific and General Information Sharing Among Competing Academic Researchers*, 43 RSCH. POL'Y 465, 465–66 (2014); Bobby L. Houtkoop et al., *Data Sharing in Psychology: A Survey on Barriers and Preconditions*, 1 ADVANCES METHODS & PRACS. PSYCH. SCI. 70, 81 (2018); *Reproducibility Project: Cancer Biology*, CTR. FOR OPEN SCI., https://www.cos.io/rpcb [https://perma.cc/DK8L-ASZJ].

> writer, to ask for his research programs to recreate them and understand them. Can/should I do that? . . . What's your experience with that?[75]

One reply mentioned, "I've asked for parameters, codes, procedures etc. several times from the authors when the description in a paper has been vague. I've had a lot of different responses: Some have plain ignored me . . . others have given me everything I asked for . . ."[76]

Another participant stated, "Personally, I have asked the source code or datasets of other authors several times. If they don't want to share, they usually just don't reply to the e-mail, or they may say no."[77]

The picture is quite clear—access to replication materials is restricted in many cases, and authors are reluctant to share them.[78] The obvious question is why: what leads scientists to withhold information, specifically replication materials? The discussions above highlight commercial and proprietary issues, but there are other factors accounting for the wide phenomenon of restricted access. We categorize the different factors into two main groups:[79] academic/professional-related factors and commercial/proprietary-related factors.

(1) Academic/Professional-Related Factors: Researchers may limit access to the underlying materials simply because they are not interested in anyone else trying to replicate their studies. Even if there is nothing to hide in terms of bias or fraud in the original research, researchers could be concerned by

---

75. Carol.Kar, *Can I Request the Code Behind a Research Paper from the Author?*, ACADEMIA STACKEXCHANGE (July 20, 2014, 5:54 PM), https://academia.stackexchange.com/questions/26159/can-i-request-the-code-behind-a-research-paper-from-the-author/26162 [https://perma.cc/APW7-2MBL].

76. alarge, Comment to *Can I Request the Code Behind a Research Paper from the Author?*, ACADEMIA STACKEXCHANGE (July 20, 2014, 6:12 PM), https://academia.stackexchange.com/questions/26159/can-i-request-the-code-behind-a-research-paper-from-the-author/26162 [https://perma.cc/APW7-2MBL].

77. Phil, Comment to *Can I Request the Code Behind a Research Paper from the Author?*, ACADEMIA STACKEXCHANGE (July 30, 2015, 11:11 PM), https://academia.stackexchange.com/questions/26159/can-i-request-the-code-behind-a-research-paper-from-the-author/26162 [https://perma.cc/APW7-2MBL]. For a similar post, see StuckInPhDNoMore, *In Performing a Comparison, How Can I Get Authors to Actually Respond to a Request for a Source Code?*, ACADEMIA STACKEXCHANGE (Apr. 27, 2015, 3:17 PM), https://academia.stackexchange.com/questions/44333/in-performing-a-comparison-how-can-i-get-authors-to-actually-respond-to-a-reque [https://perma.cc/294S-CL83].

78. Merali, *supra* note 7, at 775 ("As a general rule, researchers . . . rarely release their codes, making it almost impossible to reproduce and verify published results generated by scientific software, say computer scientists."); *see* LeVeque et al., *supra* note 8.

79. Nelson, *supra* note 74, at 267 (offering a similar categorization).

the prospect of falsification or academic criticism.[80] Relatedly, in some cases, the materials may not be in good enough shape for the researcher to feel comfortable sharing them with the world.[81] Reluctance to share materials could also stem from a researcher's desire to stay ahead of his peers and maintain a relative advantage in using the underlying materials for subsequent research projects.[82]

(2) Commercial/Proprietary-Related Factors: Commercial and proprietary concerns may also play a significant role in a researcher's decision to restrict access to replication materials.[83] In some instances, the relevant information is owned by a third party that merely licensed it to the original researcher.[84] In other cases, researchers or the organizations they are affiliated with may avoid sharing information to preserve a commercial advantage over competitors.[85] This is where IP comes into play, as we discuss in Parts II-III.

---

80. Mark J. Costello, *Motivating Online Publication of Data*, 59 BIOSCIENCE 418, 421 (2009); Sophia K. Acord & Diane Harley, *Credit, Time, and Personality: The Human Challenges to Sharing Scholarly Work Using Web 2.0*, 15 NEW MEDIA & SOC'Y 379, 384–85 (2012); Neil Pearce & Allan H. Smith, *Data Sharing: Not as Simple as It Seems*, 10 ENV'T HEALTH 1 (2011).

81. *See* Tatiana Perrino et al., *Advancing Science Through Collaborative Data Sharing and Synthesis*, 8 PERSPS. ON PSYCH. SCI. 433, 439 (2013) ("[S]ome [researchers] expressed reservations about data sharing—they felt it would be time-consuming, that their data may not be correctly interpreted, and in particular that they wished to protect their intellectual investments."); Neela Enke et al., *The User's View on Biodiversity Data Sharing – Investigating Facts of Acceptance and Requirements to Realize a Sustainable Use of Research Data*, 11 ECOLOGICAL INFORMATICS 25, 30 (2012) ("[L]oss of control, possible misinterpretation of one's data by someone else, the time and effort required to prepare a data set for sharing, not being acknowledged for sharing data, missing data standards, missing infrastructure and unclear legal conditions."); Matthew Cooper, *Sharing Data and Results in Ethnographic Research: Why This Should Not Be an Ethical Imperative*, 2 J. EMPIRICAL RSCH. ON HUM. RSCH. ETHICS: AN INT'L J. 3 (2007).

82. Fecher et al., *supra* note 49, at 9; Zinner et al., *supra* note 49, at 434; Acord & Harley, *supra* note 80; Raymond Dalgleish et al., *Solving Bottlenecks in Data Sharing in the Life Sciences*, 33 HUM. MUTATION 1494, 1494–96 (2012); Hassan Masum et al., *Ten Simple Rules for Cultivating Open Science and Collaborative R&D*, 9 PLOS COMPUTATIONAL BIOLOGY e1003244 (2013); José M. Fernandez et al., *Ethical and Secure Data Sharing Across Borders*, *in* FIN. CRYPTOGRAPHY & DATA SEC. 136 (J. Blyth, S. Dietrich & L. J. Camp eds., 2012).

83. Daniel Gardner et al., *Towards Effective and Rewarding Data Sharing*, 1 NEUROINFORMATICS 289, 291–92 (2003); Wei Hong & John P. Walsh, *For Money or Glory? Commercialization, Competition, and Secrecy in the Entrepreneurial University*, 50 SOCIO. Q. 145, 153–54 (2009); Campbell & Bendavid, *supra* note 49; NATIONAL RESEARCH COUNCIL, SHARING PUBLICATION-RELATED DATA AND MATERIALS: RESPONSIBILITIES OF AUTHORSHIP IN THE LIFE SCIENCES 28–29 (2003); Eric G. Campbell et al., *Data Withholding in Academic Genetics: Evidence from a National Survey*, 287 JAMA 473, 475–79 (2002); NATIONAL RESEARCH COUNCIL, INTELLECTUAL PROPERTY RIGHTS AND RESEARCH TOOLS IN MOLECULAR BIOLOGY 28–37 (1997).

84. Carroll, *supra* note 21.

85. *See* Nelson, *supra* note 74, at 267.

As the quotes above indicate, scientists do take commercial interests into account when considering whether to share research materials. Although anecdotal, these statements illustrate the problem we address—commercial matters, and specifically IP-related interests—may detrimentally affect access to replication materials and consequently the publications' replicability. These stories are not unique; they join an empirical thread in the literature that indicates that proprietary and commercial interests (including IP-related ones) influence scientists' and institutions' decisions regarding the sharing of research materials.[86]

This is not to say that commercial interests have no place in academia. Many important inventions, such as magnetic resonance imaging,[87] frequency modulation synthesis,[88] and glatiramer acetate,[89] were developed in research institutes and transferred to the market thanks to knowledge commercialization. Moreover, insulating research from industry would ignore the changing technological environment[90] and the need for great minds in both academia and the industry. Accordingly, we do not argue that commercial interests should be entirely taboo within academia. Rather, we observe certain negative consequences that can arise from the influence of commercial interests on science and research.[91] In particular, we point out

86. *See* Campbell & Bendavid, *supra* note 49, at 243, 250; Hong & Walsh, *supra* note 83, at 153–54, 161–64; *see also infra* notes 173–189 and accompanying text; Stephen Hilgartner, *Access to Data and Intellectual Property: Scientific Exchange in Genome Research*, *in* INTELL. PROP. RIGHTS AND RSCH. TOOLS IN MOLECULAR BIOLOGY 28, 36–38 (National Research Council ed., 1997) (arguing that patenting increases the practice of secrecy among scientists); David Blumenthal et al., *Data Withholding in Genetics and the Other Life Sciences: Prevalences and Predictors*, 81 ACAD. MED. 137, 142–45 (2006) (demonstrating how commercial interests adversely affect scientists' sharing practice). *Cf.* Shur-Ofry, *supra* note 21, at 385 ("[T]he disclosure of negative information that was accumulated as part of R&D activities may allow competitors to benefit from the effort and investment expended by the party possessing the negative knowledge."), 387 ("[I]n commercial settings, the disclosure of negative information revealed during R&D activities may allow free riding on part of competitors, result in loss of lead time or jeopardize first-to-market advantage.").

87. *See generally* DONALD W. MCROBBIE ET AL., MRI FROM PICTURE TO PROTON (2007) (providing background on magnetic resonance imaging).

88. *See generally* John M. Chowning, *The Synthesis of Complex Audio Spectra by Means of Frequency Modulation*, 21 J. AUDIO ENG'G SOC'Y 526, 526–34 (1973) (providing background on frequency modulation synthesis).

89. *See generally* G. Comi et al., *Effect of Glatiramer Acetate on Conversion to Clinically Definite Multiple Sclerosis in Patients with Clinically Isolated Syndrome (PreCISe Study): A Randomised, Double-Blind, Placebo-Controlled Trial*, 374 LANCET 1503 (2009) (providing background on glatiramer acetate's potential application to multiple sclerosis patients).

90. *See infra* notes 173–189 and accompanying text.

91. *See, e.g.*, Don Chalmers & Dianne Nicol, *Commercialisation of Biotechnology: Public Trust and Research*, 6 INT'L J. BIOTECHNOLOGY 116 (2004) (analyzing the effect of commercial interests on public trust in biotechnological research).

one problematic effect: that IP-related interests create a disincentive to share research materials and thus exacerbate the replication crisis.

To conclude, a variety of professional and commercial factors may deter scientists from sharing research materials. Empirical evidence suggests that to mitigate these factors, scientists seek to maintain control over their research materials.[92] This ultimately leads to restricted access to replication materials.

C. Existing Policy Responses to the Replication Crisis

The scientific community is extremely concerned—and for good reason—with the replication crisis. Scholarly associations, funding agencies, scientific journals, and activist scientists have responded with various initiatives aimed at tackling the problem.[93] In the following paragraphs, we describe some of the major initiatives, then discuss their limitations and explain why it is difficult to rely on them as a comprehensive solution to the replication crisis. Various initiatives purport to enhance replicability, some with the specific aim of encouraging the disclosure of replication materials.[94] For instance, the Open Science Framework (OSF), a project founded by the Center for Open Science, provides a platform for scientists who wish to share their research materials in a comfortable and collaborative manner.[95] Other initiatives in this vein are the findable, accessible, interoperable, and reusable ("FAIR") Data Principles, the Registry of Research Data Repositories ("re3data"), and the Transparency and Openness Promotion guidelines.[96]

---

92. *See generally* Kristin Eschenfelder & Andrew Johnson, *The Limits of Sharing: Controlled Data Collections*, 48 PROCS. AM. SOC'Y FOR INFO. SCI. & TECH. 1 (2011) (examining why and how repositories control access to and use of research data and finding that commercial interests play a major role in this domain).

93. *Reproducibility PI Manifesto*, LORENA A. BARBA GRP. (Feb. 6, 2013), https://lorenabarba.com/gallery/reproducibility-pi-manifesto [https://perma.cc/G8U2-TTML] (outlining a pledge by Professor Barba to lead her group with a consistent reproducibility policy); Sven, *Dutch Research Funder Grants 3 Million Euros for Replication Studies*, EDAWAX (August 2, 2016), https://www.edawax.de/2016/08/dutch-research-funder-grants-3-million-euros-for-replication-studies [https://perma.cc/Q4WE-WNMY] [hereinafter EDaWaX]; Rougier et al., *supra* note 54 (explaining the ReScience Initiative, which encourages the explicit replication of already published research).

94. OPEN SCIENCE FRAMEWORK, https://osf.io [https://perma.cc/7E3E-F4UG]; *Reproducibility PI Manifesto*, *supra* note 93; *Code Share*, 514 NATURE 536, 536 (2014).

95. OPEN SCIENCE FRAMEWORK, *supra* note 94.

96. *FAIR Principles*, GOFAIR, https://www.go-fair.org/fair-principles [https://perma.cc/NQ6Z-K4NH] (describing how the FAIR Principles are intended to provide guidelines to improve the findability, accessibility, interoperability, and reuse of data); RE3DATA, https://www.re3data.org [https://perma.cc/5PWX-V54N]. For more initiatives in this direction, see Brian A. Nosek et al., *Promoting an Open Research Culture*, 348 SCI. 1422 (2015).

Funding agencies have significant power to influence replicability.[97] They can, of course, provide funding to encourage replication studies as the Netherlands Organisation for Scientific Research has done, allocating three million euros for the sole purpose of performing replication studies.[98] In addition, funders may support replicability indirectly, for example by requiring researchers to commit to sharing replication materials in connection with funded projects.[99] Indeed, some funding agencies have already taken such measures, e.g., the National Institute of Health ("NIH"),[100] the National Science Foundation ("NSF"),[101] and other funders.[102]

Scientific journals also strive to mitigate the problem of restricted access to replication materials.[103] Leading journals such as Nature, Science, and PLoS have published recommended standards and editorial policies regarding the disclosure of research materials.[104] These reflect a wide range of standards and policies, including some for code, some for data, some for miscellaneous replication materials, and some for different combinations of these elements.[105]

Some activist scientists have suggested their own reforms to promote better access to replication materials.[106] Lorena Barba, for example, a

---

97. Benedikt Fecher et al., *What Drives Academic Data Sharing?*, 10 PLOS ONE e0118053 (2015).

98. EDAWAX, *supra* note 93.

99. NASEM REPORT, *supra* note 1, at 5; Youngseek Kim & C. Sean Burns, *Norms of Data Sharing in Biological Sciences: The Roles of Metadata, Data Repository, and Journal and Funding Requirements*, 42 J. INFO. SCI. 230, 232–34, 241–42 (2016).

100. *Data Management & Sharing Policy Overview*, NIH, http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm [https://perma.cc/MCB4-CYB8].

101. *Scientists Seeking NSF Funding Will Soon be Required to Submit Data Management Plans*, NAT'L SCI. FOUND. (May 10, 2010), http://www.nsf.gov/news/news_summ.jsp?cntn_id=116928 [https://perma.cc/98MN-7WC9].

102. Toronto International Data Release Workshop Authors, *Prepublication Data Sharing*, 461 NATURE 168, 168–70 (2009).

103. *Code Share*, *supra* note 94; *Reporting Standards and Availability of Data, Materials, Code and Protocols*, NATURE, https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards#replication-studies [https://perma.cc/5LKW-G95Z]; *Science Journals: Editorial Policies*, SCI., https://www.science.org/content/page/science-journals-editorial-policies#research-standards [https://perma.cc/C9YB-C8FL]; *Data Availability*, PLOS ONE, https://journals.plos.org/plosone/s/data-availability [https://perma.cc/D2CH-Q5T7].

104. Rousi & Laakso, *supra* note 61.

105. *Id.* at 134–40 tbl.1.

106. *See, e.g.*, *Reproducibility PI Manifesto*, *supra* note 93; Stodden, *Enabling Reproducible Research*, *supra* note 19; Romain-Daniel Gosselin, *Statistical Analysis Must Improve to Address the Reproducibility Crisis: The Access to Transparent Statistics (ACTS) Call to Action*, 42 BIOESSAYS p.e1900189 (2020); John P. A. Ioannidis, *Why Most Clinical Research Is Not Helpful*, 13 PLOS MED. p.e1002049 (2016); Rafael Jimenez et al., *Four Simple Recommendations to Encourage Best Practices in Research*, 6 F1000 RSCH. 876 (2017).

professor at the School of Engineering and Applied Science at George Washington University, has proposed the *Reproducibility PI Manifesto*.[107] She calls upon principal investigators ("PIs"), who play a key role both in decision-making with respect to materials sharing and in nurturing sharing practices among young researchers, to commit to norms that would increase replicability.[108] In the 1990's, David Donoho, a statistics professor at Stanford University, and his group initiated a research project and shared their research-related software with other scientists.[109] After 15 years, they summarized their experience of sharing computational studies materials and offered advice to inspire colleagues to follow their lead.[110]

All such initiatives—whether led by organizations, funding agencies, journals, or scientists—are beneficial as they encourage a culture of replication materials sharing. Yet, these initiatives alone do not provide a comprehensive solution to the replication crisis. To begin with, the majority of these initiatives, which are designed to encourage information sharing, are voluntary. As a result, sharing replication materials ultimately boils down to the authors' goodwill. This is the case, for example, with respect to all scientist-initiated initiatives.[111]

With respect to funding agency initiatives, although there is a growing trend toward requiring the disclosure of replication materials, there is no uniformity in the scope of the requirements and some funding agencies do not mandate it at all.[112] Moreover, even when such requirements apply, enforcement is inconsistent.[113] For instance, the NSF has had a sharing requirement for funded studies since 2011.[114] However, the NSF does not consistently enforce it.[115] Moreover, because not all studies are funded through funding agencies, agency initiatives offer, at best, a partial solution.

---

107. *Reproducibility PI Manifesto*, *supra* note 93.

108. *Id.*

109. Donoho et al., *supra* note 10.

110. *Id.*

111. Rousi & Laakso, *supra* note 61, at 146.

112. Xiaolei Huang et al., *Willing or Unwilling to Share Primary Biodiversity Data: Results and Implications of an International Survey*, 5 CONSERV. LETTERS. 399, 404 (2012) ("[O]nly one-third of respondents reported that sharing data was encouraged by their employers or funding agencies."); Paul Schofield et al., *Post-publication Sharing of Data and Tools*, 461 NATURE 171, 171 (2009).

113. Perrino et al., *supra* note 81, at 438 (arguing that funding policies show varying degrees of enforcement when it comes to data sharing).

114. *Digital Research Data Sharing and Management*, NAT'L SCI. FOUND. (Strategy & Budget Task Force on Data Policies, Arlington, Va.), Dec. 14, 2011, at 1, 4.

115. Christine Borgman, *The Conundrum of Sharing Research Data*, 63 J. AM. SOC. INF. SCI. TECHNOL. 1059, 1063 (2012) ("[NSF] has not enforced the [sharing] requirement consistently.").

Regarding journal policies, most journals do not have any materials sharing guidelines,[116] and those that do generally express the relevant policies or standards as expectations rather than mandates.[117] Moreover, journal policies often contain various exceptions that allow authors to avoid sharing.[118] Common among these are privacy concerns, ethical issues, third-party restrictions, and commercial considerations.[119]

Another significant limitation of journals' sharing policies is the scope of materials to which they apply.[120] In Section A, we discussed the need for a comprehensive set of materials—and not only code or data—when replicating computational studies.[121] However, most policies address code and data only, leaving aside crucial elements such as software documentation and workflows. Indeed, empirical data confirm that there is insufficient access to replication materials in computational studies.[122]

Finally, note that these pro-sharing initiatives are oriented almost exclusively toward the first group of impediments to sharing, that is, the academic/professional-related factors.[123] Regardless of how valuable these measures could be, they must be complemented with policies that address the second thread of factors that hinder sharing of replication materials, namely, the commercial/proprietary-related factors. IP-related considerations often play an important role in connection with such commercial motivations to inhibit access to replication materials.[124] Yet, IP law's impact in this context is nuanced. As our analysis demonstrates, different IP rights and regimes have different effects on access to replication materials.

In the next two parts, we investigate the effects of three IP regimes—copyright, patents, and trade secrets—on access to replication materials. We find that copyright law's detrimental effect on access is probably not very significant. In contrast, patent law and trade secret law likely encourage the withholding of replication materials by individual scientists and organizations, and accordingly exacerbate the replication crisis.

---

116. Rousi & Laakso, *supra* note 61, at 136.

117. WILEY, *supra* note 63.

118. Rousi & Laakso, *supra* note 61, at 135.

119. *Id.*

120. John Maunsell, *Announcement Regarding Supplementary Material*, 30 J. NEUROSCIENCE 10599 (2010) (pointing out problems with data-sharing in journals' policies).

121. *See supra* Section I.A.

122. *See* Stodden et al., *Verification of Computational Results*, *supra* note 8, at 2588.

123. In fact, some of them, as explained above, grant an exemption from sharing in case of counter-interests of commercial or proprietary aspects. *See* WILEY, *supra* note 63.

124. Note that IP rights may be correlated with non-commercial interests or connected only indirectly to commercial considerations. Yet, in most of the cases commercial interests play a certain role in the game. *See* Or Cohen-Sasson, *The Patent Medium: Toward a Network Paradigm of the Patent System*, 32 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 857, 857–59 (2022).

II. COPYRIGHT & REPLICABILITY

One branch of IP law that could have an impact on replicability is copyright law. Many components of a computational study may be protected by copyright. The scientific paper itself, the underlying software code, and the documentation may all qualify as literary works eligible for copyright protection.[125] In addition, under U.S. copyright law, while data are not copyrightable per se, the selection and arrangement of the dataset can be copyrighted if it displays sufficient creative choices.[126] In the EU, in addition to copyright for original selection and arrangement, a database can also be protected under certain terms by a sui generis right.[127] Hence, copying or adapting any of these elements for purposes of replication could constitute copyright infringement.[128]

Nevertheless, copyright law's impact on replications is probably not as significant as it may seem at first blush. In the next paragraphs, we present a two-fold argument to support this claim. In short, we posit that where the materials are not publicly accessible, they are not susceptible to copying, making copyright law largely irrelevant; while to the extent the materials are accessible, existing policy levers can be invoked to alleviate concerns about liability for copyright infringement.

A. Non-Accessible Replication Materials

When the materials are not publicly available, the real barrier to replicability is not copyright law but rather lack of access. In such cases, copyright law has no impact on the ability to replicate, since the replication materials cannot be accessed regardless of copyright.

In these cases, because copyright is not the problem, copyright law cannot offer an adequate solution. Copyright law regulates lawful access to materials that are otherwise available to the user. Among the exclusive rights of the copyright owner, the law provides for certain limitations and exceptions—including, for example, the fair use doctrine in U.S. copyright law, discussed in the next sub-part. Such limitations and exceptions enable certain uses of copyrighted materials that would otherwise be considered infringing. In recent years, a growing thread of legal scholarship suggests viewing

125. 17 U.S.C. § 101.
126. Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 357–60 (1991).
127. Council Directive, *On the Legal Protection of Databases*, 1996 O.J. (L 9) 1.
128. 17 U.S.C. § 106.

copyright's limitations and exceptions as "user rights."[129] The concept of user rights has been used by courts in certain jurisdictions, including Canada and Israel.[130] The conceptual switch from exceptions to user rights can have significant doctrinal implications in copyright law. Nevertheless, even the broadest proposed constructions of user rights do not suggest imposing on copyright owners affirmative duties to supply copyrighted materials, and there is certainly no doctrinal basis in current copyright law for such a claim. Thus, even where copying replication materials is permitted under copyright law—for instance, under the fair use doctrine—there is no legal basis for a potential user's claim that the copyright owner must grant them access to unavailable materials.

There is one context in which copyright law could hinder access to non-published materials—that is, when the materials are protected by a digital lock (i.e., a technological protection measure, or "TPM"), such as passwords or encryption, that restricts access to the work.[131] In such cases, hacking the TPM to retrieve the information may be possible (although the chances that one would do so for the purpose of conducting a replication study are probably small). Copyright law, though, prohibits the disabling of a TPM that a copyright owner has placed on a work, and thus restricts access to the materials.[132] Nonetheless, this obstacle is probably more theoretical than real in the context of replication materials. TPMs are typically used by copyright holders or content providers (e.g., website owners) when the copyrighted works are circulated, such as when they are embedded in a product or included in an online database; they are unlikely to restrict access to replication materials.

---

129. *See* Niva Elkin-Koren, *Copyright in a Digital Ecosystem: A User-Rights Approach*, *in* COPYRIGHT LAW IN AN AGE OF LIMITATIONS AND EXCEPTIONS 132 (Ruth Okediji ed., 2017); Guy Pessach, *Toward a New Jurisprudence of Copyright Exemptions*, 55 IDEA 287, 287–89 (2015); Niva Elkin-Koren & Orit Fischman-Afori, *Taking Users' Rights to the Next Level: A Pragmatist Approach to Fair Use*, 33 CARDOZO ARTS & ENT. L.J. 1, 41 (2015); Pascale Chapdelaine, *Copyright User Rights and Remedies: An Access to Justice Perspective*, 7 LAWS 1, 21–26 (2018).

130. *See* David Vaver, User Rights in Canadian Copyright Law, Keynote Speech for Ontario Library Association Copyright Symposium (Dec. 8, 2021) (transcript available at https://digitalcommons.osgoode.yorku.ca/conference_papers/10 [https://perma.cc/LBX5-SA9J]); Niva Elkin-Koren, *Users' Rights*, *in* AUTHORING RIGHTS: READING THE NEW ISRAELI COPYRIGHT ACT (Michael Birnhack & Guy Pessach eds., 2009) (in Hebrew).

131. Aside from such "access control measures," there is also a different type of TPMs, "copy control measures," which does not restrict access but rather what can be done with the work. *See, e.g.*, *Technological Protection Measures (TPM) – Fact Sheet*, SIMON FRASER UNIV. (Apr. 4, 2022), https://www.lib.sfu.ca/help/academic-integrity/copyright/technological-protection-measures [https://perma.cc/E394-BNVY] (defining TPMs as "digital management tools used to restrict what users can do with digital materials").

132. 18 U.S.C. § 1201. The U.S. Copyright Office is authorized to adopt exemptions to this prohibition, but such exemptions are temporary and narrow in scope. *Id.*

B. Accessible Copyrighted Replication Materials

In regard to publicly available materials, copyright law supposedly serves a more direct role in impeding replication studies. As the materials are accessible to the public, the only thing supposedly stopping a scientist from reproducing and adapting them as part of a replication study is a legal barrier—namely, the concern for copyright infringement liability.

Fortunately, existing copyright law doctrines seem to alleviate this concern to a large extent. In particular, the fair use doctrine, which exists in the United States and elsewhere, permits the unlicensed use of copyright-protected works under certain circumstances. [133] The fair use doctrine enables courts to balance the interests of copyright owners against the interests of subsequent creators and the public.[134] As we argue below, the use of copyrighted works in the context of replication studies could (and should) qualify as fair use in many cases.

Section 107 of the United States Copyright Act provides the statutory framework for determining whether certain activities constitute fair use. The statute identifies certain types of uses—such as criticism, comment, news reporting, teaching, scholarship, and research—as examples of activities that may qualify as fair use. Replication studies seem to fall under both criticism and research. In any event, this list of uses that potentially qualify as fair use is not meant to be exhaustive. Section 107 calls for consideration of four different factors in evaluating whether the use of a copyrighted work made in a particular case is a fair use.[135] As the following paragraphs demonstrate, these factors lean in favor of construing replication studies as a fair use.

The first factor is the *purpose and character of the use, including whether the use is of a commercial nature or is for nonprofit educational purposes*.[136] Under this factor, courts look at the goal of the party claiming fair use. Using copyrighted materials for the purpose of replicating scientific studies serves an important public interest, as explained above.[137]

When the replication study is conducted within a research university or a nonprofit organization, this may further increase the likelihood that a court would find the use to be fair. This does not mean, however, that a replication

---

133. 17 U.S.C. § 107.

134. *See* Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 WASH. L. REV. 579, 620 (2018).

135. In addition to the four factors above, other factors may also be considered by a court in weighing a fair use question, depending upon the circumstances. *See U.S. Copyright Office Fair Use Index*, U.S. COPYRIGHT OFF., https://www.copyright.gov/fair-use/ [https://perma.cc/DD58-YR8C] [hereinafter *Copyright Office*].

136. 17 U.S.C. § 107.

137. *See supra* Section II.A.

study conducted by a commercial enterprise could not be considered fair use. In recent years, courts in the United States have often considered socially valuable uses of copyrighted materials by a commercial entity to be fair.[138] Indeed, distinguishing among users based on the character of their motivation (commercial or non-commercial) is particularly difficult in the context of research. Most research organizations today are driven, at least partially, by a desire to earn a profit. Even universities are no longer simply research institutions but rather behave like "firms" in many respects, including in their drive to commercialize their research results.[139] At all events, even where a commercial purpose weighs against fair use in a certain case, courts ultimately balance the purpose and character of the use against the remaining factors.[140]

Under the first factor in the fair use analysis, courts also tend to look more favorably at "transformative" uses. Transformative uses are those that add something new, with a further purpose or different character, and do not substitute for the original use of the work.[141] A replication study, by its nature, seeks to reuse the original materials in the same way they were used in the original study. But as the Supreme Court recently emphasized, literal transformation of the copyrighted work is not necessary so long as the purpose and character of the use is transformative.[142] Replication serves a different purpose and has a different character than the original study, and thus could plausibly qualify as transformative. In any event, courts should not hesitate to find that the first factor supports the classification of a non-transformative use as fair if it has a clear social value, which is certainly the case here.[143]

Another relevant consideration under the first factor is the necessity of using the particular set of replication materials to achieve the goal of replicating a given study. Indeed, the replications scenario seems to fit

---

138. *See, e.g.*, Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183, 1204 (2021) ("There is no doubt that a finding that copying was not commercial in nature tips the scales in favor of fair use. But the inverse is not necessarily true, as many common fair uses are indisputably commercial.").

139. *See, e.g.*, Karl-Heinz Leitner et al., *The Role of Heads of Departments in the Commercialization of University Research*, 91 J. BUS. ECON. 353, 353 (2021) ("Today, the acquisition of third-party funds, the filing of patents and the foundation of spin-offs are seen as similarly relevant to other academic activities such as conducting research and teaching."); *see also* Hong & Walsh, *supra* note 83 (describing how increasing commercialization and scientific competition cause greater secrecy in academic science).

140. *See Copyright Office*, *supra* note 135.

141. *Id.*

142. *See Google LLC*, 141 S. Ct. at 1202–04.

143. *See generally* Rebecca Tushnet, *Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It*, 114 YALE L.J. 535, 547–82 (2004) (discussing the free speech value of non-transformative copying).

squarely within the larger category of "necessity or right-of-reply" fair use cases.[144] In these cases, because the user arguably cannot promote her socially beneficial goal without using a specific copyrighted work, the use receives favorable treatment under the first factor. This consideration has led many courts to favor parody over satire.[145] In Justice Souter's words in the famous *Campbell* case, parody "needs to mimic an original to make its point," while satire need not use the original work to make its point about society.[146] Replications—like parodies—require use of the original materials to achieve their essential purpose: validating (or refuting) the original study.

The second factor is the *nature of the copyrighted work*.[147] Under this factor, courts consider the degree to which the work relates to copyright's purpose of encouraging creative expression.[148] The closer the work is to the core of what copyright is intended to protect, the more difficult it is to establish that its reproduction amounts to fair use.[149] Thus, the use of a creative or imaginative work (e.g., a fictional short story) is less likely to be deemed fair than the use of a factual work (e.g., a biography).[150] The inherently objective nature of replication materials, particularly in computational studies, makes them highly susceptible to fair use.[151]

The third factor is the *amount and substantiality of the portion used in relation to the copyrighted work as a whole*.[152] Under this factor, courts consider whether the portion used is "reasonable in relation to the purpose of the copying."[153] The use of an entire work can still qualify as fair use when it is reasonably necessary for the purpose.[154] In the context of a replication

144. Bruce P. Keller & Rebecca Tushnet, *Even More Than Parody the Real Thing: Parody Lawsuits Revisited*, 94 TRADEMARK REP. 979, 980–81 (2004) (discussing this category of fair use cases).

145. *See id.* at 979 (critically examining this trend); *see also* Richard A. Posner, *When Is Parody Fair Use?*, 21 J. LEGAL STUD. 67, 73 (1992) (arguing for a narrow copyright exemption for parody).

146. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 580–81 (1994).

147. 17 U.S.C. § 107.

148. *Copyright Office*, *supra* note 135.

149. *See Campbell*, 510 U.S. at 586 ("This factor calls for recognition that some works are closer to the core of intended copyright protection than others, with the consequence that fair use is more difficult to establish when the former works are copied.").

150. *Id.*; *see also* Stewart v. Abend, 495 U.S. 207, 237 (1990) (holding that "fair use is more likely to be found in factual works than in fictional works").

151. The analysis of this factor could change in regard to replications of TDM studies, where the underlying materials often include creative and imaginative works. Yet even in such cases, those copyrighted works are not used for their expressive value. *See infra* notes 160–161 and accompanying text.

152. 17 U.S.C. § 107.

153. Authors Guild v. Google, Inc., 804 F.3d 202, 221 (2d Cir. 2015).

154. Levendowski, *supra* note 134, at 627.

study, the use of the code, database, or other materials in their entirety is arguably warranted, since the study could not be conducted properly otherwise.

The fourth factor is the *effect of the use upon the potential market for or value of the copyrighted work*. In assessing this factor, courts review whether, and to what extent, the unlicensed use harms the current or future market for the copyright owner's original work.[155] Using copyrighted works to replicate a study may decrease the value of the study—e.g., if the replicator reveals that the findings of the original study are unsubstantiated. In extreme cases, the original publication might even be retracted.[156] But a decrease in the paper's value that may result from such circumstances is not the type of harm that copyright is intended to prevent. Again, comparison to the parody scenario illustrates the point: As the Supreme Court held in *Campbell*, "when a lethal parody, like a scathing theater review, kills demand for the original, it does not produce a harm cognizable under the Copyright Act."[157] The same is true in the replication context.[158]

In short, each of the foregoing factors seems to suggest that the use of replication materials to conduct replication studies would present a strong case for fair use. Nevertheless, the outcome of any given case will depend on a fact-specific inquiry. The high level of uncertainty involved in fair use analysis is a major shortcoming of this measure, as it results in a chilling effect on expressions and discussions that use copyrighted materials, including replication studies.[159] Still, we think that courts can (and should) apply the fair use doctrine in a manner that supports replication studies. With sufficient precedent applying fair use in this manner, the level of uncertainty in this context would certainly decrease.

---

155. *Copyright Office*, *supra* note 135.

156. *See generally* S.P.J.M Horbach & W. Halffman, *The Ability of Different Peer Review Procedures to Flag Problematic Publications*, 118 SCIENTOMETRICS 339 (2019) (providing and analyzing data concerning retractions of scholarly publications and the cause of such retractions).

157. Campbell v. Acuff-Rose Music, 510 U.S. 569, 592 (1994). As the Supreme Court's observations reflect, the same logic applies to critical reviews of creative works. *See also* Posner, *supra* note 145 (maintaining that the harm to an author that comes from exposing the weaknesses of his effort is not the kind of harm that copyright law protects against). In these cases, while the use arguably reduces the value of the work, it is not because it supplies the demand for it.

158. Posner, *supra* note 145. In fact, the decrease in value of the scientific paper may be entirely irrelevant to the inquiry under the fourth factor, when the copyrighted work at stake is the code or other replication materials (apart from the paper itself).

159. Leah Chan Grinvald & Ofer Tur-Sinai, *Intellectual Property Law and the Right to Repair*, 88 FORDHAM L. REV. 63, 108–11 (2019). For relevant empirical evidence, see Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978-2005*, 156 U. PA. L. REV. 549, 575 (2008) (reporting that 30.4% of preliminary injunctions found in favor of fair use and 24.1% of bench trial opinions did).

Thus far, our discussion has assumed that the databases used for replication studies include raw data, which are not protected by copyright. However, we must also address a special type of studies that use databases comprising copyrighted works, such as texts and images. Such works often have the potential to generate information that goes beyond what their individual authors expressed.[160] Various methods of computational and statistical analysis—often referred to as text data mining ("TDM")—can yield that type of additional information.[161]

A researcher conducting a TDM study may have purchased licenses from the owners of the copyright in the underlying materials or relied upon the fair use doctrine or designated TDM exceptions that exist in certain jurisdictions.[162] To the extent the original TDM study is considered fair use, we maintain that further use of the same materials for replications should generally qualify as fair use as well. This may not be straightforward, as the use of the materials made by the author is novel and creative, while the replicator merely seeks to repeat the same study.[163] This, however, should not impact the analysis. When evaluated in relation to the original materials, the replication study should count as transformative to the same degree as the initial TDM study. Most importantly, from a policy perspective, as replications serve a significant social interest and form an essential component of the scientific enterprise, they should not receive less favorable treatment than the original study simply because, by their nature, they do not strive to present novel findings.

To recap, it seems that at least in the United States, with proper construction and consistent application of existing policy levers, copyright law should not pose a major obstacle to replicability.

III. PATENTS, TRADE SECRETS & REPLICABILITY

This Part deals with two other branches of intellectual property law—patent law and trade secret law—that may have a significant effect on access to replication materials, both in general and specifically in the context of computational studies. The first section will discuss patent law. We show that researchers and institutions may avoid public disclosure of replication

160. Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291, 359 (2019); *see also* Michael W. Carroll, *Copyright and the Progress of Science: Why Text and Data Mining is Lawful*, 53 U.C. DAVIS L. REV. 893, 909 (2019).

161. Sag, *supra* note 160.

162. *See id.* (arguing that fair use is likely to apply in these cases).

163. For the importance of transformativeness in the context of the first factor, see *supra* notes 143–145 and accompanying discussion.

materials due to patent-related considerations. Then, the second section will deal with trade secrets, where to maintain protection, one must keep information confidential, away from the public domain.

A. Patent Law

Patent law is seldom discussed in the literature concerning the replicability crisis, despite its potential impact.[164]

It is not uncommon for a scientific publication and a patentable invention to have overlapping subject matter.[165] For example, a computational study reported in a scientific journal may involve the use of a software invention that is eligible for patent protection. While the patentability of software inventions raises various complexities,[166] software patents are generally available in many jurisdictions, including the United States.[167] The Supreme Court of the United States held in *Alice Corp. v. CLS Bank International* (2014) that implementing an abstract idea on a computer is unpatentable in itself.[168] Following this decision, commentators have raised the possibility that "software patents are now a thing of the past."[169] However, empirical evidence from the years following the decision show that *Alice* has not had a substantial impact on software patenting.[170]

In addition to a potential patent for any software-related components, a computational study may also yield an invention in a different technological

164. For a notable exception, see Stodden & Reich, *supra* note 23. *See also* Freilich, *supra* note 26; Sherkow, *supra* note 26 (both discussing the difficulty of replicating experiments that are described in patent documents).

165. *See, e.g.*, Chiara Franzoni & Giuseppe Scellato, *The Grace Period in International Patent Law and Its Effect on the Timing of Disclosure*, 39 RSCH. POL'Y 200, 204 (2010) (conducting an empirical study of "duals," i.e., paired patents and scientific articles describing the patented innovations).

166. *See generally* Michael Guntersdorfer, *Software Patent Law: United States and Europe Compared*, 2 DUKE L. & TECH. REV. 1 (2003); Tanner Mort, *Abstract Ideas: The Time Has Come for Congress to Address the Patentability of Software and Business Method Inventions*, 56 IDAHO L. REV. 383 (2020).

167. Guntersdorfer, *supra* note 166, at 7–8.

168. Alice Corp. v. CLS Bank Int'l, 573 U.S. 208, 225–26 (2014).

169. Nathan Hakimi, *The Status of Software Patents in the Post-Alice Era*, CHI.-KENT J. INTELL. PROP. (Nov. 20, 2016), https://studentorgs.kentlaw.iit.edu/ckjip/status-software-patents-post-alice-era/ [https://perma.cc/7REP-PKM8]; *see also* Neil Gandal et al., *Out of Sight: Patents That Have Never Been Cited*, 126 SCIENTOMETRICS 2903, 2920 (2021) (noting that the *Alice* decision raised the threshold of patentability in the software field).

170. Hakimi, *supra* note 169.

field. For example, a study in the field of computational immunology may advance the development of a new vaccine.[171]

One may wonder to what extent the motivation to publish research results in an academic journal co-exists with the desire to obtain patent protection for innovative findings resulting from the same study. Indeed, patent protection may seem to be in conflict with traditional norms of open communication and free flow of information within the scientific community.[172] One must recognize, though, that today's reality is more complex, and the lines separating academic and commercially oriented research have become blurred over the last decades.[173] In fact, as the following paragraphs demonstrate, there is a dual drift: commercial firms publish in scientific journals, while universities often seek patent protection for research findings.[174]

Commercial firms publish a considerable amount of their R&D in scientific journals.[175] Publication fulfils researchers' personal and professional aspirations and enables researchers to maintain links with—and gain reputation and prestige in—the wider academic community.[176] Thus, to attract and retain top scientists, particularly in research-intensive industries, firms recognize that they must enable researchers to publish study results.[177] Publications can also help a firm to establish credibility for its research results and serve as a meaningful signal to investors and collaborators regarding the scientific and technological competence and capabilities of the firm.[178]

---

171. Kamal Rawal et al., *Identification of Vaccine Targets in Pathogens and Design of a Vaccine Using Computational Approaches*, 11 SCI. REP. 1, 1–2 (2021); Lassi Liljeroos et al., *Structural and Computational Biology in the Design of Immunogenic Vaccine Antigens*, 2015 J. IMMUNOLOGY RSCH. 1, 1–2 (2015).

172. Rebecca S. Eisenberg, *Property Rights and the Norms of Science in Biotechnology Research*, 97 YALE L.J. 177, 178 (1987).

173. *See supra* note 139 and accompanying text; Wesley M. Cohen et al., *Not in the Job Description: The Commercial Activities of Academic Scientists and Engineers*, 66 MGMT. SCI. 4108, 4108–09 (2020) (examining how scientists' motives to engage in commercial activities differ across fields).

174. For related discussions, see Hong & Walsh, *supra* note 83 and accompanying text; *supra* note 139 and accompanying text.

175. Roberto Camerani et al., *Do Firms Publish? A Multi-Sectoral Analysis* (Oct. 23, 2018) (unpublished working paper).

176. *Id.* at 8.

177. *Id.* at 7–10.

178. *Id.* at 9; Rebecca S. Eisenberg, *The Promise and Perils of Strategic Publication to Create Prior Art: A Response to Professor Parchomovsky*, 98 MICH. L. REV. 2358, 2359 (2000); *see also* Francisco Polidoro & Matt Theeke*, Getting Competition Down to a Science: The Effects of Technological Competition on Firms' Scientific Publications*, 23 ORG. SCI. 1135, 1135 (2012) ("[F]irms need to demonstrate the merits of their innovations to outside parties, such as regulatory agencies and professional communities, whose assessments influence the commercialization of new products.").

Publishing is also conceived as a "ticket of admission" to the academic network, which enables the firm to access knowledge and resources available only to the members of that network.[179]

At the same time, efforts by universities and state-sponsored research institutions to secure patent protection for their research results is a common practice nowadays.[180] Universities, as institutions, have an interest in patenting simply as a source of revenue.[181] Technology transfer offices are tasked with ensuring that such patents are properly managed and commercialized.[182] Moreover, it is very common nowadays for university-based scientists to establish business enterprises (most often startup companies).[183] The university or research institution generally has a commercial stake in such ventures, and they often fund and support them.[184] Universities and research institutions are thus clearly oriented and organized towards commercialization and not only towards advancing their research goals.[185]

Individual researchers involved in academic research may also benefit from patenting through royalty revenues from university licenses and spin-offs.[186] In addition, researchers may have non-pecuniary incentives to patent their inventions, including the reputational value of receiving credit for a

---

179. Camerani et al., *supra* note 175, at 6.

180. Franzoni & Scellato, *supra* note 165; Katherine J. Strandburg, *Curiosity-Driven Research and University Technology Transfer*, *in* 16 ADVANCES IN THE STUDY OF ENTREPRENEURSHIP, INNOVATION, AND ECONOMIC GROWTH: UNIVERSITY ENTREPRENEURSHIP AND TECHNOLOGY 93, 94 (Gary D. Libecap ed., 2005) (pointing to the drastic increase in university patenting).

181. Strandburg, *supra* note 180.

182. *See, e.g.*, Gail A. Van Norman & Roï Eisenkot, *Technology Transfer: From the Research Bench to Commercialization: Part 1: Intellectual Property Rights—Basics of Patents and Copyrights*, 2 JACC: BASIC TO TRANSLATIONAL SCI. 85, 85 (2017).

183. *See generally* MANUEL STAGARS, UNIVERSITY STARTUPS AND SPIN-OFFS: GUIDE FOR ENTREPRENEURS IN ACADEMIA (2015); Christopher S. Hayter, *Harnessing University Entrepreneurship for Economic Growth: Factors of Success Among University Spin-offs*, 27 ECON. DEV. Q. 18 (2013); Sonali K. Shah & Emily Cox Pahnke, *Parting the Ivory Curtain: Understanding How Universities Support a Diverse Set of Startups*, 39 J. TECH. TRANSFER 780, 780–82 (2014); Achim Walter et al., *Championship Behaviors and Innovations Success: An Empirical Investigation of University Spin-Offs*, 28 J. PROD. INNOVATION MGMT. 586, 588 (2011).

184. *See supra* note 183 and accompanying text; Strandburg, *supra* note 180.

185. Shah & Pahnke, *supra* note 183.

186. *Id.* at 787; *see also* Lisa Larrimore Ouellette & Andrew Tutt, *How Do Patent Incentives Affect University Researchers?*, 61 INT'L REV. L. & ECON. 105883, 105883 (2020) (noting that under the framework established by the Bayh Dole Act, academic grant recipients have a direct financial stake in the success of their inventions as a result of the requirements that universities share the resulting patent royalties with inventors); Stodden & Reich, *supra* note 23 (demonstrating that the number of software patents granted to academic researchers has more than doubled from 2000 to 2010 among top patenting universities).

particular novel idea.[187] Universities, in fact, increasingly consider faculty members' patenting history in decisions regarding tenure and career advancement.[188]

Collaborations between academic institutions and commercial firms should thus be expected, as well as a mingling of commercial and academic incentives.[189] All of this leads to frequent situations in which an individual or entity seeks to register a patent for innovations that have been developed as part of a study that is (or will be) published in a scientific journal. The potential for patenting a study's research findings could impact decisions regarding the publication and sharing of research-related materials. To understand this, a brief description of patentability standards is in order.

One of the conditions for patentability is that the invention be novel when the patent application is filed.[190] In simple terms, the invention must be new—i.e., not publicly disclosed previously by the applicant or third parties.[191] For purposes of determining novelty, patent law defines what types of public disclosures may qualify as "prior art."[192] In the United States, prior art consists of, *inter alia*, all printed publications, including patents, patent applications, and non-patent literature, such as magazine articles, newspaper articles, electronic publications, on-line databases, websites, or Internet publications.[193] If any single reference within the prior art contains the claimed invention, the patent is "anticipated" and cannot be granted.[194] Note that to be considered prior art, the relevant information must be available to the public.[195]

---

187. Ouellette & Tutt, *supra* note 186.

188. *See, e.g.*, Paul R. Sanberg et al., *Changing the Academic Culture: Valuing Patents and Commercialization Toward Tenure and Career Advancement*, 111 PROC. NAT'L ACAD. SCIS. 6542, 6542 (2014) (commending this practice).

189. Albert Banal-Estañol & Inés Macho-Stadler, *Scientific and Commercial Incentives in R&D: Research Versus Development?*, 19 J. ECON. & MGMT. STRATEGY 185, 185–86 (2010); Henry Sauermann et al., *Doing Well or Doing Good? The Motives, Incentives and Commercial Activities of Academic Scientists and Engineers*, (Jan. 2010), https://www.researchgate.net/publication/228555271_Doing_Well_or_Doing_Good_The_Motives_Incentives_and_Commercial_Activities_of_Academic_Scientists_and_Engineers [https://perma.cc/6DJF-6FHA]; Aled Edwards, *Team Up with Industry: Combining Commercial and Academic Incentives and Resources Can Improve Science*, 531 NATURE 299 (2016).

190. 35 U.S.C. § 102(a).

191. *Id.*

192. Price II, *supra* note 25, at 782.

193. *Id.*; U.S. PATENT AND TRADEMARK OFFICE, MANUAL OF PATENT EXAMINING PROCEDURE §§ 2127–2128 (9th ed. 2020).

194. Price II, *supra* note 25, at 782.

195. For an elaborate discussion of the case law regarding what is considered public in this context, see Camilla A. Hrdy & Sharon K. Sandeen, *The Trade Secrecy Standard for Patent Prior Art*, 70 AM. U. L. REV. 1269, 1282 (2021).

In addition to being novel, the invention must also be non-obvious (or, in some legal systems, must represent an "inventive step") as of the filing date.[196] In the United States, this requirement is codified in Section 103 of the Patent Act, which provides that no patent shall be registered "if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious . . . to a person having ordinary skill in the art to which the claimed invention pertains."[197]
To summarize, the general rule is that a patent claim will be rejected if there is prior art that anticipates it or renders it obvious (on its own or in combination with other prior art references). Nevertheless, this rule is subject to exceptions for "non-prejudicial disclosure." Many patent systems worldwide recognize two such exceptions: first, a disclosure made during an international exhibition recognized by the Convention on International Exhibitions;[198] second, a disclosure made as the consequence of an abuse, like disclosing an applicant's work without her consent.[199] In these cases, the disclosure does not undermine patentability, provided that the application is filed within a specified period.[200] Notably, no legal system has a designated exception for disclosures made for purposes of replications.

In addition, a few countries, including the United States, Japan, and Canada, contemplate a "grace period,"[201] which operates as a broader exception to the general rule.[202] The grace period is a specified period prior to the filing of a patent application during which certain public disclosures—typically, those made by the applicant or derived from the applicant's work—are not considered prior art for purposes of assessing the claimed invention's novelty or non-obviousness.[203] Different countries' regimes differ in terms of

196. For the use of the term "inventive step," see, for example, The Patents Act 1977, § 3 (UK) ("An invention shall be taken to involve an inventive step if it is not obvious to a person skilled in the art.").

197. 35 U.S.C. § 103.

198. *See, e.g.*, Franzoni & Scellato, *supra* note 165, at 201.

199. *See* Franzoni & Scellato, *supra* note 165, at 201. For example, one of the exceptions listed in Article 55 of the European Patent Convention is for a disclosure occurring no earlier than six months prior to the filing date that is due to an evident abuse of the applicant.

200. *See* Franzoni & Scellato, *supra* note 165, at 201 (noting, as an example, the case of an invention that is unlawfully disclosed after it has been stolen from the inventor who held it in confidence).

201. *See* Franzoni & Scellato, *supra* note 165, at 201–02. *See also* Robert P. Merges, *Priority and Novelty Under the AIA*, 27 BERKELEY TECH. L.J. 1023, 1046 (2012) ("[T]he AIA diverges from the international norm which approximates an 'absolute novelty' standard.").

202. *See, e.g.*, Merges, *supra* note 201, at 1030.

203. *See, e.g.*, William G. Giltinan, *The Disclosure Function, Academic/Private Partnerships, and the Case for Affirmatively Used, Multinational Grace Periods*, 22 TEX. INTELL. PROP. L.J. 109, 114–15 (2014); Ouellette & Tutt, *supra* note 186. Only in the United States, the

the length of the grace period, the types of disclosures that are not deemed to be prior art, and other factors.[204]

In the United States, a disclosure made by an inventor or another who obtained the subject matter therefrom, one year or less before the filing date, is not considered prior art.[205] As a result, in a significant number of cases, an inventor who files a patent application after a disclosure that would otherwise qualify as prior art has appeared in the field may still be entitled to a patent.[206] As noted above, the grace period concept is not universal. In European jurisdictions, there is no grace period regime but only narrow non-prejudicial exceptions along the lines described above, which some of the EU member states allow.[207] As a result of this lack of harmonization, inventors who file for patent protection in the United States after disclosing the invention during the grace period would probably not be eligible for patent protection in European countries.[208]

The foregoing issues are highly relevant to understanding how patent law bears on the sharing and disclosure practices of individual researchers and organizations. Generally speaking, due to the requirements of novelty and non-obviousness, when a scientific publication and a patentable invention have overlapping subject matter, the preferable strategy from the vantage point of securing patent protection is to minimize disclosure before filing a

---

grace period may apply in certain circumstances to independent disclosures made by third parties. *See* 35 U.S.C. § 102(b)(1)(B) (removing from the prior art any disclosure, made one year or less before the effective filing date of the claimed invention, of subject matter that had been previously publicly disclosed by the inventor, a joint inventor, or another who obtained the subject matter directly or indirectly from the inventor or joint inventor).

204. For a comparative account of different grace period regimes, see Tomasz Ozyhar et al., *When Speed Matters: A Discussion on the Benefits of a Grace Period in Patent Law to Accelerate Pharmaceutical Innovation in Times of Pandemic*, 9 J. L. & BIOSCIENCES 1 (2022).

205. 35 U.S.C. § 102(b)(1)(A). In addition, certain disclosures made by third parties during the grace period are removed from the prior art (provided that the same subject matter had been previously disclosed by the inventor).

206. Merges, *supra* note 201.

207. *See, e.g.*, Merges, *supra* note 201, at 1043 (noting that Europe currently operates under a strict novelty requirement, without a grace period); Ozyhar et al., *supra* note 204, at 6 ("European law, apart from very narrow exceptions provided in the EPC, does not allow a grace period . . . ."); Margo A. Bagley, *The Need for Speed (and Grace): Issues in a First-Inventor-To-File World*, 23 BERKELEY TECH. L.J. 1035, 1055 (2008) ("The European Patent Convention (EPC) operates on an absolute-novelty basis, with limited (and virtually meaningless) exceptions for certain types of disclosures occur ring within six months of the application filing date.").

208. Franzoni & Scellato, *supra* note 165. Under Article 55 of the European Patent Convention, a disclosure of the invention shall not be taken into consideration if it occurred no earlier than six months prior to the filing date and it was due to: (a) an evident abuse of the applicant, or (b) a display of the invention at a recognized international exhibition.

patent application (or, in the United States, at the very least—a provisional patent application).[209]

This general strategy extends to the publication of the scientific paper itself.[210] As Margo Bagley states, "the unforgiving nature of patent novelty rules encourages a culture in which the dissemination of even very early-stage research, sometimes no more than a proof of concept, is delayed while a provisional patent application is prepared by the university TTO."[211] Commercial firms and entities engaged in public-private research collaborations probably adhere even more strongly to a practice of delaying publication.[212]

Nevertheless, empirical evidence shows that in many cases, publication of a scientific paper precedes the filing of a correlating patent application.[213] A combination of factors may account for this phenomenon. First and foremost, urgency is a relevant consideration not only in connection with patenting, but also in the realm of academic publications.[214] Scientific careers depend largely on the reputation built upon scientific publications and conferences, and being the first to claim a scientific discovery can have crucial importance in this regard.[215] "Publish or perish" is an aphorism describing the need to publish frequently and continuously to obtain tenure and achieve academic

---

209. Since 1995, the USPTO has offered the option of filing provisional patent applications (PPAs) that are easier and substantially cheaper to file than a full-fledged patent application. *See* Miriam Bentwich, *Changing the Rules of the Game: Addressing the Conflict Between Free Access to Scientific Discovery and Intellectual Property Rights*, 28 NATURE BIOTECHNOLOGY 137, 137 (2010).

210. *See, e.g.*, Margo A. Bagley, *Academic Discourse and Proprietary Rights: Putting Patents in Their Proper Place*, 47 B.C. L. REV. 217, 218 (2006) ("[T]oday, academic researchers are being encouraged by technology transfer offices ("TTOs") and industry sponsors to delay publishing and presenting their work until after filing a patent application and sometimes even longer than that.").

211. *Id.* at 221.

212. *See* Franzoni & Scellato, *supra* note 165, at 209 (finding that "[t]he presence of a firm among the assignees appears to be related to relatively longer time lags between the patent priority date and the date of publication of the matched paper," and reporting a survey revealing that a high proportion of firm-university agreements on joint research included explicit delayed disclosure clauses).

213. *See* Franzoni & Scellato, *supra* note 165, at 207–08 (reporting patent-publication pairs that showed a negative time lag—i.e., cases where the publication preceded the filing of the patent application).

214. *See, e.g.*, *see* Franzoni & Scellato, *supra* note 165, at 204 (listing a large corpus of works that attest to the fact that academic inventors may more urgently feel the need to disclose their inventions (including at conferences or in scientific papers) in a timeframe that does not always correspond to that required by patent procedures).

215. *See* Franzoni & Scellato, *supra* note 165, at 203.

success.[216] The pressure to publish early also stems from expectations entrenched in the academic tradition of advancing the science rather than sitting on important discoveries, as well as the need to show results for purposes of securing funding and finding collaborators for future research projects.[217] One can see that there is a tension between the advantages of early patenting and the desire to avoid delayed publications. Ultimately, universities and other relevant actors must balance these considerations in forming their policies and implementing them on a case-by-case basis.

Another factor that may explain why academic publications sometimes precede patenting is the nature of academic research, which may be worthy of publication even in its early stages, when it is not yet ripe for patent protection because, for example, the inventor cannot yet show the invention's utility[218] or supply an enabling disclosure, i.e., a specification of the invention demonstrating how it can be utilized.[219] Similarly, in some cases, the findings of a particular study are only one component of what may give rise to a patentable invention after additional aspects are developed.

In any event, even if a patent application related to a study can be drafted, the invention may have only speculative commercial value at an early stage.[220] Before an organization—particularly a TTO with limited resources—[221] decides whether a given invention merits patenting, it may want to allow for some time to further develop the invention by working out specific embodiments, constructing prototypes, collecting data to show utility, and investigating the potential market value of the invention.[222] In addition to these factors, the partial shelter offered by the grace period regime certainly plays a role in encouraging earlier publications in the United States,

---

216. *See* Bagley, *supra* note 210, at 239 (maintaining that "issued patents and start-up companies, lucrative or not, may be considered poor substitutes for scholarly publications by tenure committees").

217. Ouellette & Tutt, *supra* note 186. *See also* Jordan S. Joachim, *Is the AIA the End of Grace? Examining the Effect of the America Invents Act on the Patent Grace Period*, 90 N.Y.U. L. REV. 1293, 1315 (2015) (maintaining that delayed publications "will force researchers to sit on important discoveries rather than disclosing them").

218. According to 35 U.S.C. § 101, one of the conditions for patenting is that the invention is useful.

219. 35 U.S.C. § 112.

220. *See, e.g.*, Bagley, *supra* note 207, at 1046 (noting that inventions generated in universities often have only speculative commercial value).

221. *See* Joachim, *supra* note 217, at 1314 (maintaining that "university technology transfer offices only have limited funds to allocate towards patent applications").

222. *See* Ouellette & Tutt, *supra* note 186 (noting the need for such activities); Franzoni & Scellato, *supra* note 165, at 208 (pointing at the need to collect information about the industrial relevance and potential profitability of inventions prior to the decision whether to patent or not).

where disclosures may be made in the hope that within twelve months, a patent application will be filed.[223]

All of this directly affects researchers' approach to sharing replication materials. To maximize the chances of obtaining patent protection, individual researchers or organizations would be inclined to keep whatever materials they can out of the prior art. When a paper is published before a corresponding patent application is filed (and the grace period cannot be relied upon—for instance, if patenting within a year is not likely,[224] or submissions in other jurisdictions are expected)—any additional information disclosed along with the paper may be deemed prior art for the forthcoming patent application.[225]

The motivation to conceal materials is generally weaker when a patent application is submitted prior to the publication of a paper, since anything published after the filing date does not count as prior art. Still, organizations may have reasons to minimize the disclosure of materials (and may instruct individual scientists not to disclose them) in this scenario as well. One reason could be to preserve the option of filing additional patents that are derived from the same findings. Minimizing disclosure may also be desirable to maximize the inventor's ability to file amendments or continuations to the patent application.[226]

The secrecy norms cultivated by patent law can disincentivize information sharing even where sharing would not necessarily affect patenting. To illustrate, a certain study may not even yield an invention that qualifies for patent protection.[227] In the same vein, the publication of a scientific paper may bar novelty or non-obviousness, regardless of whether any replication materials are disclosed. In other cases, the invention may ultimately satisfy

---

223. *See* Franzoni & Scellato, *supra* note 165, at 203 (concluding that grace period favors early disclosure in open science).

224. *See* Bagley, *supra* note 207, at 1054 ("Given the realities of academic research and TTO practices, even a one-year grace period is often not long enough to accommodate the needs of many researchers. It is not uncommon for more than a single year to pass before academic research progresses to the point where a TTO can effectively assess the research's commercial potential.").

225. Note that the paper itself does not necessarily negate novelty or non-obviousness (for instance, if it only describes the software invention used in the research in general non-enabling terms), while disclosing the replication materials may change that. For the enablement standard that determines the sufficiency of a disclosure in prior art publications, see Eisenberg, *supra* note 178, at 2362 ("A publication that merely suggests a technological advance, without disclosing enough to allow a person of ordinary skill in the relevant field to make it without undue experimentation, cannot defeat a subsequent patent claim.").

226. *See* Franzoni & Scellato, *supra* note 165 (explaining that the refinement of patent applications could consist of filing additional patents on the process or on complementary or substitute technologies).

227. 35 U.S.C. § 101.

the conditions for patentability, while the disclosure of replication materials does not add information that anticipates or renders the invention obvious. Nevertheless, in all of these situations, given the inherent uncertainty involved in patent prosecution, risk-averse decision makers operating in an environment that regards patent protection as important may prefer to keep as much information as possible out of the public domain. Notably, decisions regarding information sharing are typically made by the authors of the scientific paper, who typically lack comprehensive understanding of patent law. Not being able to know *ex ante* whether putting replication materials in the public domain will jeopardize patent protection creates a chilling effect on sharing replication materials. This demonstrates a significant cost of the patent system in impeding knowledge dissemination in variety of situations, including when patenting is not possible at all or when researchers ultimately choose not to apply for a patent.[228]

In sum, the prospect of patent protection for the findings of scientific studies may cause researchers and organizations to be reluctant to publicly disclose any information with respect to their studies that they are not obligated to share. In the context of computational studies, such information may encompass replication materials, including code and documentation.

B. Trade Secret Law

Another branch of IP law that may impact access to replication materials is trade secret law.[229] Generally speaking, any piece of information that is not widely known and gives its holder a competitive edge qualifies for trade secret protection, provided that the holder takes reasonable measures to keep it secret.[230] As the following paragraphs demonstrate, a wide range of replication materials could be protected as trade secrets. As protection depends on the holder's ongoing efforts to keep the information

228. Jeremy M. Grushcow, *Measuring Secrecy: A Cost of the Patent System Revealed*, 33 J. LEGAL STUD. 59, 60 (2004) (finding that patent law led scientists to conceal data even when eventually not seeking patent protection). *See also* Stodden & Reich, *supra* note 23 (describing the rise in software patent applications amid scientists, which can possibly have detrimental effects on sharing research materials); Stodden, *IP*, *supra* note 19 (discussing the disincentive of information sharing due to software patents).

229. *See* Mark A. Lemley, *The Surprising Virtues of Treating Trade Secrets as IP Rights*, 61 STAN. L. REV. 311 (2008) (arguing that trade secrets are best understood as a form of intellectual property rights).

230. *See, e.g.*, Deepa Varadarajan, *Trade Secret Fair Use*, 83 FORDHAM L. REV. 1401, 1403 (2014) (noting that "virtually any useful information can be a trade secret, so long as the information is relatively secret, economically valuable, and subjected to reasonable secrecy precautions by the owner").

confidential—here, too, IP law may incentivize concealment of replication materials.

Until recently, trade secret law in the United States was governed by state law. Over the years, almost all states have enacted a form of the Uniform Trade Secrets Act (UTSA), formulated in 1979 to promote harmonization of this legal field.[231] In 2016, Congress enacted the Defend Trade Secrets Act (DTSA), the first federal civil trade secret act, to supplement state trade secret laws.[232] The DTSA is largely modeled after the UTSA,[233] and it does not preempt state law claims.[234]

The UTSA defines a trade secret as any information, including a formula, pattern, compilation, program, device, method, and process that: (i) derives independent economic value, actual or potential, from "not being generally known to, and not being readily ascertainable by proper means" by, other parties who can benefit from it, and (ii) is the subject of reasonable efforts by its holder to maintain its secrecy.[235] A trade secret is protectable as long as these conditions are kept;[236] thus, trade secret protection is potentially infinite.[237]

Once information is protected as a trade secret, the law prohibits its *misappropriation*.[238] Misappropriation can occur in a variety of ways, including by the acquisition, use, or disclosure of another's trade secret by "improper means" (e.g., trespass or bribery) or by a person who breaches a

---

231. *See* Sharon K. Sandeen, *Lost in the Cloud: Information Flows and the Implications of Cloud Computing for Trade Secret Protection*, 19 VA. J.L. & TECH. 1, 38 (2014) ("Today, the predominant source of trade secret law is the Uniform Trade Secrets Act (UTSA), which was first adopted in 1979 and has now been enacted in substantial part by forty-seven states and the District of Columbia.").

232. Pub. L. No. 114-153, 130 Stat. 376 (2016) (codified at 18 U.S.C. § 1836 (2012)).

233. Jeanne C. Fromer, *Machines as the New Oompa-Loompas: Trade Secrecy, The Cloud, Machine Learning, and Automation*, 94 N.Y.U. L. REV. 706, 709 (2019).

234. Camilla A. Hrdy, *The General Knowledge, Skill, and Experience Paradox*, 60 B.C. L. REV. 2409, 2433 (2019).

235. UNIF. TRADE SECRETS ACT § 1(4). Note that clause (i) includes two separate requirements: First, the information must be secret, i.e., not generally known or readily ascertainable, and second, it must derive independent economic value from not being generally known or readily ascertainable. *See* Sandeen, *supra* note 231; Rex N. Alley, *Business Information and Nondisclosure Agreements: A Public Policy Framework*, 116 NW. U.L. REV. 817, 825 (2021).

236. Camilla A Hrdy & Mark A. Lemley, *Abandoning Trade Secrets*, 73 STAN. L. REV. 1, 12–13 (2021).

237. Fromer, *supra* note 233, at 710. For the famous examples of the Chartreuse liquor and Google's search engine algorithm, see respectively: Thornton Robison, *Confidence Game: An Approach to the Law about Trade Secrets*, 25 ARIZ. L. REV. 347, 362 (1983); Andrew A. Schwartz, *The Corporate Preference for Trade Secret*, 74 OHIO ST. L.J. 623, 651–52 (2013).

238. *See* UNIF. TRADE SECRETS ACT §§ 2–3.

confidentiality obligation.[239] Under the UTSA, neither reverse engineering nor independent discovery is considered "improper means."[240]

With this brief summary of the basic principles of trade secret law behind us, we can return our focus to the replications scenario. In connection with a computational study, an individual author or affiliated institution may choose to claim trade secret protection for elements of the study that are not disclosed as part of the publication. Essentially all replication materials in connection with a computational study can qualify for trade secret protection. This includes the algorithm, code, database, workflows, details about the computational environment, methodology, and any other know-how regarding the study.

For this reason, the impact of trade secret law on the availability of replication materials could be much broader than that of patent law. Patent law protects only a defined set of inventions, and only certain types of public disclosure will undermine patent protection for an otherwise patentable invention. Because trade secret protection is available for a much broader set of information items, the incentive to keep everything confidential is generally stronger.

The motivation for keeping replication materials confidential could stem from various factors, including preserving an option to exclusively use the materials for future research projects, as well as maintaining the potential to commercialize the information (e.g., via licensing agreements).[241]

Note that with respect to some types of information, it may be hard (or even impossible) to rely on the protections offered by other IP regimes, making trade secret protection particularly appealing. This is the case, for example, with respect to most datasets, which normally do not qualify for copyright protection and are not considered patentable subject matter.[242]

To be sure, even when patent or copyright protection is available, trade secret protection may be attractive as a complementary or alternative measure. Trade secrets have certain advantages over patent protection—for example, their potentially infinite duration.[243] Another factor that may strengthen the motivation of businesses to rely on trade secret protection is

---

239. *See* Fromer, *supra* note 233, at 710–11 (summarizing the law on this point and explaining that "improper means" encompass both criminal or tortious behavior and some lawful conduct (e.g., aerial photography of a manufacturing plant under construction)).

240. *See* UNIF. TRADE SECRETS ACT § 1 cmt. (amended 1985), 14 U.L.A. 539 (1985).

241. *See supra* notes 78–92 and accompanying text (providing a general survey of the factors that could motivate concealment of replication materials); *see* David S. Levine & Ted Sichelman, *Why Do Startups Use Trade Secrets?*, 94 NOTRE DAME L. REV. 751, 768–70 (2019) (explaining the practice of licensing trade secrets).

242. *See supra* note 126 and accompanying text (explaining the copyrightability of datasets).

243. *See supra* note 237 and accompanying text.

the relatively low cost of maintaining secrecy, especially in comparison with patent protection.[244] Also, modern trade secrets are also often more resistant to reverse engineering than in the past, making trade secret protection more effective.[245] For these and other reasons, recent years have seen a growing reliance on trade secret protection, particularly in the computing industry.[246] As noted above, trade secret protection is not used exclusively as an alternative to patent protection but may serve as a complementary form of protection against misappropriation.[247] For example, trade secrecy can be used to protect information before a patent application is filed (and even later, up until the publication date of the application).[248] In addition, while the law requires a patent applicant to disclose the manner and process of making and using the invention, applicants often manage to keep certain aspects of the patented invention secret.[249]

If an author wishes to use trade secret protection for replication materials, she must take reasonable measures to keep them confidential.[250] Clearly, if the author shares this information with the public, it can no longer be protected as a trade secret. Indeed, even if the shared information does not become generally known, once it is shared with others, it risks losing protection.[251]

Nevertheless, pursuant to the doctrine of relative secrecy, trade secret holders can share their secrets with third parties without losing protection if their disclosure is limited, and the persons to whom the disclosure is made have an express or implied duty of confidentiality.[252] Of course, if a party bound by a confidentiality agreement discloses the information or uses it in a

---

244. Levine & Sichelman, *supra* note 241, at 782.

245. Fromer, *supra* note 233; Hrdy & Lemley, *supra* note 236.

246. *See, e.g.*, Fromer, *supra* note 233, at 718 (discussing trade secrecy's appeal for computing innovation)*; see also* Levine & Sichelman, *supra* note 241.

247. Brenda M. Simon & Ted Sichelman, *Data-generating Patents*, 111 NW. U. L. REV. 377, 383 (2016); Cohen-Sasson, *A Hidden Technological Assumption in Patent Law: The Case of Gene Patents and the Disclosure Requirement*, 22 J. WORLD INTEL. PROP. 272, 282 (2019) (discussing simultaneous use of patents and trade secrets in genetic inventions).

248. 35 U.S.C. § 122; 37 C.F.R. § 1.14.

249. *See, e.g.*, Michal Shur-Ofry & Ofer Tur-Sinai, *Constructive Ambiguity: IP Licenses as a Case Study*, 48 U. MICH. J.L. REFORM 391, 420–21 (2015); Rebecca S. Eisenberg, *Patents and the Progress of Science: Exclusive Rights and Experimental Use*, 56 U. CHI. L. REV. 1017, 1029 n.52 (1989) (referring to patent applicants' practice of withholding information from patent specifications and protecting their know-how through trade secrecy); Pamela Samuelson & Suzanne Scotchmer, *The Law and Economics of Reverse Engineering*, 111 YALE L.J. 1575, 1620 (2002) (discussing the practice of platform developers to patent some components of their systems while maintaining Application Programming Interfaces (APIs) as trade secrets).

250. *See supra* note 235 and accompanying text.

251. Sandeen, *supra* note 231.

252. *Id.*

manner that violates the agreement, and the information becomes generally known, trade secret protection is lost. In such a case, however, the trade secret holder would, at a minimum, have recourse against the wrongdoer for breach of contract and misappropriation of the trade secret.

Taken together, our analysis of patent law and trade secret law reveals an undesired phenomenon: These legal regimes encourage the concealment of replication materials. Given the high societal interest in replications, scholars and policy makers may consider limiting the scope of IP rights.[253] But narrowing IP protections is not imperative to promote replicability.[254] In lieu of this binary view—either IP rights or replicability—we propose an inclusive approach, maintaining that IP protection and replicability can coexist.

IV. FACILITATING ACCESS TO REPLICATION MATERIALS: CONDITIONAL-ACCESS-AGREEMENT

In this Part, we suggest a novel way to facilitate replicability of scientific findings. Our proposal centers around a new policy tool: the *Conditional-Access-Agreement* (CAA). We have formulated the proposed mechanism in a way that seeks to enable replication studies without depriving the original scientists of existing or prospective IP rights. The gist of the CAA mechanism is the establishment of a private, controlled channel for the transfer of replication materials between original scientists and replicators. By enhancing the ability to gain access to replication materials, the CAA would reduce the number of publications that cannot be verified (or disproved).

Recall that our analysis in Part IV shows that *public* access to replication materials is what may destroy both trade secret protection and the ability to secure patent protection. For replication purposes, however, access does not have to be public. This is precisely the core tenet of our solution. The CAA mechanism strives to establish a *private* and *controlled* access to replication materials. This design both protects IP rights and facilitates replication studies. To establish the controlled channel of access to replication materials, we harness the power of a non-disclosure agreement (NDA).

---

253. IP-narrowing solutions could take the form of specific exceptions from liability for replications (though such exceptions would not be of much help where the materials are not available at all) or go in the direction of excluding or limiting IP protection for replication materials. *Cf.* Stodden, *Enabling Reproducible Research*, *supra* note 19; Stodden, *IP*, *supra* note 19 (suggesting licensing copyright-protected data under the Creative Commons license, which allows, inter alia, commercial use).

254. *Cf.* Shur-Ofry, *supra* note 21, at 388–89 (encouraging regulators to devise schemes that mandate disclosure of negative findings, on the one hand, while recognizing the potential harm for the disclosing party and providing it some sort of compensation, on the other hand).

An NDA is a contract between two or more parties that establishes a confidential relationship with respect to certain information that the parties wish to share with one another. The party receiving the confidential information typically commits to: (a) using the information only for a specified purpose, (b) keeping it confidential, and (c) avoiding sharing it with anyone else. An NDA defines what constitutes confidential information, the use for which the information is disclosed, the obligations of the receiving party, the period of confidentiality, as well as certain exceptions to the confidentiality obligations.[255]

By creating a confidential relationship between parties, an NDA can protect any type of sensitive information.[256] NDAs thus play a major role in protecting proprietary interests.[257] Consequently, NDAs are very common in the business world, for instance, in the startup-venture capital relationship;[258] between employees and employers;[259] and in dispute settlements.[260]

In the CAA context, the NDA would prohibit the disclosure or use of the replication materials for any purpose other than replication. Since patent law and trade secret law bar protection only in the case of public disclosure, information shared through a private channel, pursuant to an NDA, would not constitute prior art for prospective patent applications, nor would it negate trade secrecy. Accordingly, the NDA would alleviate major concerns that deter authors and organizations from sharing research materials.[261] More generally, this mechanism would allow authors to maintain control over who has access to their research materials.[262] All in all, the CAA would facilitate on-demand access to replication materials on the one hand, while protecting the researcher's proprietary interest in those materials on the other.

Our proposed CAA mechanism would operate as follows: When submitting a paper for publication, an author would execute a CAA vis-à-vis

---

255. Sometimes parties mutually disclose confidential information, in which case each party is both a disclosing and a receiving party.

256. Illegal contracts are an exception. *See* Vasundhara Prasad, *If Anyone Is Listening, #MeToo: Breaking the Culture of Silence Around Sexual Abuse through Regulating Non-Disclosure Agreements and Secret Settlements*, 59 B.C. L. REV. 2507, 2513–15 (2018).

257. Note that an NDA can be used to protect non-commercial interests, for instance, privacy-related interests.

258. Charles F. Wieland & Scott W. Cummings, *How Successful Startups Capitalize on IP*, 1 IEEE NANOTECHNOLOGY MAG., Dec. 2007, at 11, 12–15.

259. David R. Hannah, *Should I Keep a Secret? The Effects of Trade Secret Protection Procedures on Employees' Obligations To Protect Trade Secrets*, 16 ORG. SCI. 71 (2005).

260. *See generally*, Scott A. Moss, *Illuminating Secrecy: A New Economic Analysis of Confidential Settlements*, 105 MICH. L. REV. 867 (2006) (describing the use of NDAs in dispute resolution scenarios).

261. Hilgartner, *supra* note 86; Blumenthal et al., *supra* note 86.

262. Eschenfelder & Johnson, *supra* note 92.

the journal, pledging to provide full access to replication materials upon a replicator's demand. The CAA would specify that any replicator seeking access to the materials would be required to sign an NDA prohibiting disclosure or use of the information for any purpose other than replication.

This framework is feasible thanks to the existence of a powerful intermediary—scientific journals.[263] Journals are key players within the academic publishing domain, and they have continuous involvement in a publication's lifecycle. Publishing a study in a journal involves various procedural steps, including uploading the paper, declaring a conflict of interests, and signing a copyright-related agreement.[264] These steps require the same infrastructure necessary to operate the CAA mechanism—e.g., submission platform, editors, and staff—so there is no need to establish a new infrastructure. The procedural process through an official gatekeeper[265]—the journal—is a convenient path to implement the CAA policy.

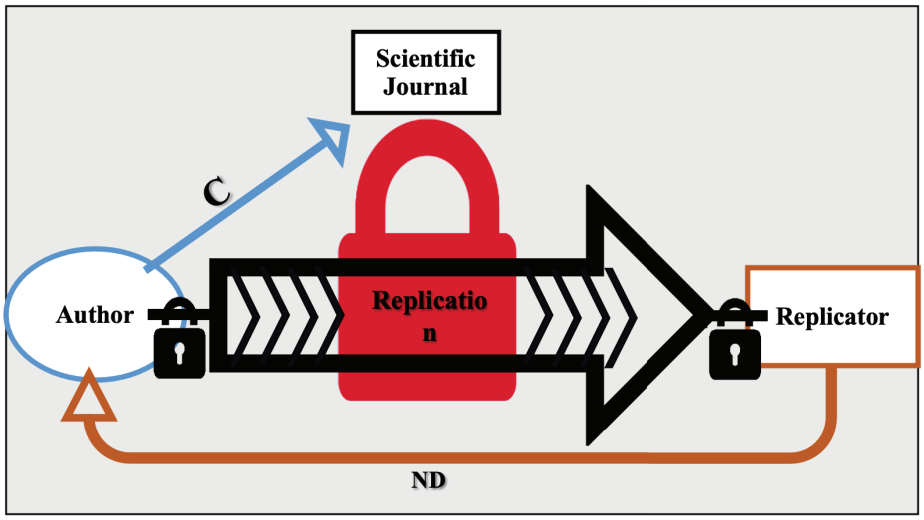

Figure 1: The CAA Mechanism: An Overview

263. *See generally* DAN LAUGHEY, KEY THEMES IN MEDIA THEORY 23–26 (2007) (describing the two-step flow theory of communication); ELIHU KATZ & PAUL F. LAZARSFELD, PERSONAL INFLUENCE 32–33 (1955) (explaining the power of intermediaries and the two-step communication model).

264. *See generally Paper Publishing Process*, AIJR, https://aijr.org/paper-publishing-process/ [https://perma.cc/RH5B-RFKF] (describing the general steps in publishing a paper). For more on conflicts of interest, see, e.g., *Conflicts of Interest*, SAE INT'L (Aug. 2016), https://www.sae.org/binaries/content/assets/cm/content/publications/journals/resources/conflicts.pdf [https://perma.cc/8LEH-DHHT].

265. *See generally* Mohammadreza Hojat et al., *Impartial Judgment by the "Gatekeepers" of Science: Fallibility and Accountability in the Peer Review Process*, 8 ADVANCES HEALTH. SCIS. EDUC. 75 (2003) (highlighting the substantial power journal reviewers have as "gatekeepers" of scientific information).

On top of the procedural-technical advantages of implementing the CAA mechanism through scientific journals, there are substantive benefits as well. Scientific journals are agents of change in society, particularly in the scientific community.[266] Some journals actively embrace this role and occasionally pursue important public missions.[267] Accordingly, it is reasonable to believe that they would welcome the opportunity to participate in measures that mitigate the replication crisis. Indeed, many have already done so voluntarily with various standards, guidelines, and policies designed to improve the replicability rate.[268] The CAA is another instrument that can facilitate replication through an intermediary that is eager to advance this goal—and journals perfectly fit this role.

Notably, scientific journals already play a key role in assuring the quality of scientific publications.[269] Their involvement spans the publication process, from establishing technical requirements and performing editorial evaluations[270] to facilitating a scholarly peer review by experts,[271] and also includes the handling of post-publication matters such as letters to the

266. *See* Ann C. Schaffner, *The Future of Scientific Journals: Lessons from the Past*, 13 INFO. TECH. & LIBRS. 239, 241 (1994); George E. Axtelle, *Technology and Social Change*, 25 EDUC. F. 133, 138–39 (1961). For an inverse effect, see André C. R. Martins, *Modeling Scientific Agents for a Better Science*, 13 ADVANCES COMPLEX SYS. 519 (2010).

267. For example, journals provided fast-track review for and free access to COVID-19-related articles during the pandemic to aid society in finding treatments faster. *See, e.g.*, *COVID-19 Articles Accepted for Fast-Track Publication in Psychological Science*, SAGE J., https://journals.sagepub.com/page/pss/covid-19 [https://perma.cc/727D-VGE4]; *Coronavirus (COVID-19) Research Highlights*, SPRINGER NATURE, https://www.springernature.com/gp/researchers/campaigns/coronavirus [https://perma.cc/CR6Y-MXGY].

268. *See supra* Section I.C.

269. *See* John C. Bailar & Kay Patterson, *Journal Peer Review: The Need for a Research Agenda*, 312 NEW ENG. J. MED. 654, 654 (1985); *see also* Phil B. Fontanarosa, et al., *Thanking Authors, Peer Reviewers, and Readers – Constancy in a Time of Change*, 283 J. AM. MED. ASS'N 2016, 2017 (2000); American Psychological Association, *Summary Report of Journal Operations 2000*, 56 AM. PSYCH. 693, 693–94 (2001).

270. IRENE HAMES, PEER REVIEW AND MANUSCRIPT MANAGEMENT IN SCIENTIFIC JOURNALS: GUIDELINES FOR GOOD PRACTICE 1–4 (2007).

271. *Id.*; Jacalyn Kelly et al., *Peer Review in Scientific Publications: Benefits, Critiques, and a Survival Guide*, 25 EJIFCC 227, 229–31 (2014).

editor[272] and retraction procedures.[273] As the primary channel for publishing research findings, journals play the most prominent role in assuring the quality and credibility of scientific studies.[274] And because replication is a cornerstone of credibility, journals are crucial stakeholders who should take active steps to enhance replicability.

Moreover, the business of scientific journals—nearly all of which are lucrative, for-profit organizations[275] owned by mega-publishers such Elsevier, Black & Wiley, Taylor & Francis, Springer Nature, and SAGE[276]—is the scientific enterprise.[277] In return for the value they derive from their work, scientific journals owe a duty to the scientific community. This duty includes the elementary obligation of ensuring the quality of the research they publish, including by facilitating replication studies.

Indeed, quite apart from the monetary rewards of operating within the scientific domain, journals have an innate duty to maintain the essential tenets of the scientific method. Replicability is one of these tenets, and journals

---

272. "Letters to the editor" or "Correspondence" are public, critical reviews by readers regarding previous publications, which are published in following issues of a journal. *See generally* Houcemeddine Turki et al., *The Value of Letters to the Editor*, 117 SCIENTOMETRICS 1285 (2018); Khalil G. Falavarjani et al., *Letter to Editor, a Scientific Forum for Discussion*, 28 J. CURRENT OPHTHALMOLOGY 1 (2016).

273. *See generally* Vedran Katavić, *Retractions of Scientific Publications: Responsibility and Accountability*, 24 BIOCHEMIA MEDICA 217 (2014); Grant Steen et al., *Why Has the Number of Scientific Retractions Increased?*, *in* METHODOLOGICAL ISSUES AND STRATEGIES IN CLINICAL RESEARCH 557 (Alan E. Kazdin ed., 2016); Elizabeth Wager & Peter Williams, *Why and How Do Journals Retract Articles? An Analysis of Medline Retractions 1988–2008*, 37 J. MED. ETHICS 567 (2011); Daniele Fanelli, *Why Growing Retractions Are (Mostly) a Good Sign*, 10 PLOS MED. e1001563 (2013).

274. Hojat et al., *supra* note 265, at 75–77.

275. Thomas J. Walker, *The Electronic Future of Scientific Journals*, U. FL. IFAS, figs.3 & 4 (1997), https://entnemdept.ufl.edu/walker/aedraft.htm [https://perma.cc/4XT2-HT28].

276. *See* Vincent Larivière et al., *The Oligopoly of Academic Publishers in the Digital Era*, PLOS ONE, June 2015.

277. The business model of journals is toll access; authors transfer their copyright, fully or partially, to a journal that holds the major commercial rights inherent in the manuscript. Journals generate income mainly from print subscriptions, site-licenses, and one-time purchases. In recent years, some journals have moved the toll from the reader to authors and institutions through Open Access models. Yet, regardless of the payer's identity, i.e., the reader or the author, the access to publications generates income to journals. It is a quite lucrative model. Journals generally do not compensate authors or peer reviewers, and production costs have plunged dramatically with the rising popularity of e-publications. Moreover, the fees for Open Access not only lower the journals' risk (e.g., the risk of non-sufficient subscriptions to cover costs), but also removes the apprehension of shadow libraries, that is piracy of academic materials, such as Sci-Hub, Z-Library, and Library Genesis. Martin Hagve, *The Money Behind Academic Publishing*, TIDSSKRIFTET (Aug. 17, 2020), https://tidsskriftet.no/en/2020/08/kronikk/money-behind-academic-publishing [https://perma.cc/L4Y6-44T7].

should do all they can to promote replicability. The CAA framework represents a promising step in this direction.

We now turn to the technicalities of implementing the CAA solution. We propose that the CAA be signed and submitted via the manuscript submission platform. Almost all journals run an electronic system through which authors submit their manuscripts. In addition to uploading the manuscript, the author can attach further documents such as tables, figures, and a cover letter; provide required information, e.g., the manuscript's subject and keywords; and digitally sign legal documents concerning any conflict of interests, originality, ethics-related approvals, and a copyright license, for example. The submission platform offers the possibility for seamless implementation of the CAA, as execution of the agreement would be merely an additional step in an existing process.

The CAA would include an appendix identifying the elements that comprise the replication materials. The list should be exhaustive, as authors must have certainty with respect to the specific materials they may be required to share with replicators. Additionally, because an exhaustive and detailed list—as opposed to an open or representative one—is not subject to interpretation, having such an appendix would minimize cases of insufficient sharing.[278] Journals, as the expert gatekeepers, should build the list based upon previous materials lists other professionals have suggested.[279] Note that these lists may differ depending on the nature of the studies published by each journal. In the event a replicator is unable to conduct a replication study because the materials required are not specified on the list, the parties can communicate regarding the necessary materials. If the author agrees to provide further materials, the NDA would cover them as well.[280] Of course, if a certain element is routinely missing, journals should update the CAA appendix accordingly.

The CAA could be signed either at the outset of the process, when a manuscript is submitted, or after it has been accepted for publication. We

278. That said, the CAA establishes a channel of communication between the author and the replicator. Through this channel, replicators can ask for and authors can provide further information, protected by the NDA.

279. *See, e.g.*, *Reproducibility PI Manifesto*, *supra* note 93; Donoho et al., *supra* note 10; Yale Law School Roundtable on Data and Code Sharing, *Reproducible Research*, COMPUTING SCI. & ENG'G 8, 9–11 (2010); Geir Kjetil Sandve et al., *Ten Simple Rules for Reproducible Computational Research*, PLOS COMPUTATIONAL BIOLOGY, Oct. 2013; Herbert M. Sauro, The Practice of Ensuring Repeatable and Reproducible Computational Models (July 17, 2021) (unpublished manuscript) (on file at https://arxiv.org/pdf/2107.05386.pdf [https://perma.cc/PG5J-ZK5R]) (providing a list of materials need to be shared to enable replication studies, and suggesting a scoring system to determine how well authors are doing).

280. The NDA would be formulated in a way that applies to any replication materials, not only those specified in the CAA appendix.

recommend that the submission platform inform authors of the CAA requirement but emphasize that, like other requirements journals routinely impose, it applies only if the manuscript is accepted.[281] Then, if the paper successfully passes the review process, the author will receive a request to sign the CAA. In our view, this procedure would generate less reluctance on the authors' side.[282] Once they have successfully cleared the highly selective threshold of peer review, authors may be more motivated to sign the CAA than at the beginning of the process.[283]

The question whether the CAA mechanism should be voluntary or compulsory is a difficult one. On the one hand, it may seem too burdensome to require authors to commit to sharing materials. On the other hand, leaving the choice to authors seems unlikely to produce widespread compliance given the disincentives to sharing we identify throughout this article.[284] On balance, we argue that a compulsory model is preferable; yet journals disinclined to impose a mandatory requirement should, at a minimum, offer a voluntary CAA. Even under a voluntary arrangement, there are efficient ways—such as nudging through reward systems, defaults, and social-proof heuristics[285]—to encourage authors to sign the CAA.[286] Moreover, in contrast to other

---

281. For instance, one such requirement is authorship verification. *See, e.g.*, *Information for Authors*, SCI., https://www.science.org/content/page/science-information-authors [https://perma.cc/CS6U-6U5Z] (describing Science's policy regarding authorship confirmation).

282. Once authors receive a positive response from a journal, it is tempting to publish the paper with this journal. Also, the submission process, including the time, efforts, and anxiety it incurs, are sunk costs, and as such they may influence the author's decision. *See generally* Hal R. Arkes & Catherine Blumer, *The Psychology of Sunk Cost*, 35 ORG. BEHAV. & HUM. DECISION PROCESSES 124 (1985) (discussing sunk-cost bias and how it influences decision making); Roch Parayre, *The Strategic Implications of Sunk Costs: A Behavioral Perspective*, 28 J. ECON. BEHAV. & ORG. 417 (1995).

283. *See generally* Arkes & Blumer, *supra* note 282; Parayre, *supra* note 282. Of course, once the CAA becomes a standard requirement in most journals, the timing of executing the CAA becomes less relevant as all platforms would anyway request the author to sign it at some point.

284. *See* Ellen M. Key, *How Are We Doing? Data Access and Replication in Political Science*, 49 PS: POL. SCI. & POL. 268, 270 (2016) (finding that with a mandatory provision, it is twenty-four times more likely that any materials will be shared and seventeen times more likely that a full replication package will be published).

285. *See generally* Yashar Saghai, *Salvaging the Concept of Nudge*, 39 J. MED. ETHICS 487 (2013) (describing nudge theory).

286. *See* Kathleen H. Jamieson et al., *Signaling the Trustworthiness of Science*, 116 PNAS 19231 (2019); Hendrik P. Van Dalen & Kène Henkens, *Signals in Science: On the Importance of Signaling in Gaining Attention in Science*, 64 SCIENTOMETRICS 209 (2005); Marcus R Munafò et al., *A Manifesto for Reproducible Science*, 1 NATURE HUM. BEHAV. 1 (2017) (demonstrating empirically that badges increase sharing by more than tenfold). *See also* text accompanying notes 303–304.

voluntary initiatives targeting the replication crisis,[287] the CAA mechanism does not require the authors' active involvement, and compliance is easy.[288] Turning to how the replicator would receive the NDA, we propose that all publications include a link to a generic NDA designed to provide disclosure for replication purposes only.[289] Upon clicking the link, the replicator would see the terms of the NDA and can then create a user account to receive the official, individualized NDA by e-mail through the journal's electronic platform.[290] Then, the replicator can sign the NDA and submit it to the journal's platform. Importantly, the replicator will be the first party to sign the NDA to ward off frivolous requests for disclosure and to prevent trolling. Thereafter, the author will provide the replication materials using the contact details in the NDA.[291]

In the remainder of this Part, we address critiques that our proposal may face and offer our responses. We also discuss the main advantages of the CAA mechanism.

To be sure, the CAA is not a risk-free solution. The CAA preserves the IP rights of authors only to the extent the replicators respect their contractual obligations under the NDA and do not disclose the replication materials. Of course, in the event of a breach, the authors can seek damages through a lawsuit, but that is not an ideal solution. Lawsuits dramatically increase authors' costs, and even if they win damages, they may lose the ability to claim trade secret protection for their work vis-à-vis third parties once the

---

287. *See supra* Section I.C.

288. The submission procedure itself brings all pertinent information about the CAA to the authors, provides the necessary forms (e.g., CAA, NDA), and all technical steps are done through the same platform through which the paper is submitted.

289. When available, the link can be embedded within the already existing Crossmark button. *See Crossmark*, ELSEVIER, https://www.elsevier.com/about/policies/crossmark [https://perma.cc/Q8F3-TUDR] ("Crossmark, a multi-publisher initiative from CrossRef, provides a standard way for readers to locate the authoritative version of a document. . . . Clicking on the Crossmark icon will inform the reader of the current status of a document and may also provide additional publication record information about the document."); *Crossmark*, CROSSREF, https://www.crossref.org/services/crossmark/crossmark [https://perma.cc/JM7F-4ZTM] ("The Crossmark button gives readers quick and easy access to the current status of an item of content, including any corrections, retractions, or updates to that record.").

290. Like in the case of authorship verification. *See supra* note 281.

291. For now, we do not think it will be efficient to require sharing the materials through one specific platform; instead, it is better to allow several common methods of research materials sharing to encourage more collaboration. *See, e.g.*, CODE OCEAN, https://codeocean.com/ [https://perma.cc/H8TD-CHZS]; BINDER, https://mybinder.org/ [https://perma.cc/9DSX-9EVB]; COLABORATORY, https://colab.research.google.com/?utm_source=scs-index [https://perma.cc/RN2T-8QCE]; GIGANTUM, https://github.com/gigantum [https://perma.cc/APL7-BMDZ]; NEXTJOURNAL, https://nextjournal.com/ [https://perma.cc/8TYA-BMV4]; ACTIVEPAPERS, https://activepapers.github.io/ [https://perma.cc/F38L-VTFE].

pertinent information is disclosed.[292] Moreover, the timeline for filing a patent application, if relevant, would be accelerated.[293]

We have two responses to these observations. First, legal enforcement is not the only tool for influencing behavior.[294] Social norms—even when not anchored in law—are powerful behavioral drivers and play a significant role in social control.[295] This is particularly true within defined groups and communities.[296] Some of the most important norms in the scientific community are intellectual honesty and respect for the law, which are part of the ethos of modern science.[297] For example, journals and other academic entities have policies against plagiarism,[298] and the scientific community has developed various technological tools to detect and fight plagiarism.[299] Indeed, the scientific community punishes plagiarism harshly, including by

---

292. *See supra* notes 230, 250–252 and accompanying text.

293. *See supra* notes 201–208 and accompanying text.

294. *See* LAWRENCE LESSIG, CODE: VERSION 2.0 120–37 (2006).

295. *See generally* Jeffrey W. Legro, *Which Norms Matter? Revisiting the "Failure" of Internationalism*, 51 INT'L ORG. 31 (1997) (applying the power of social norms to the international relations context); Bryan H. Druzin, *Social Norms as a Substitute for Law*, 79 ALB. L. REV. 67 (2015) (arguing that policymakers can use social norms to support or replace regulation); Peter H. Huang & Ho-Mou Wu, *More Order Without More Law: A Theory of Social Norms and Organizational Cultures*, 10 J.L., ECON. & ORG. 390 (1994) (analyzing the role of emotions in maintaining social order without the need for formal law).

296. *See generally* Christian S. Crandall, *Social Contagion of Binge Eating*, 55 J. PERSONALITY & SOC. PSYCH. 588, 592–95 (1988) (analyzing how defined groups impact binge eating behavior); Laura Doering & Amandine Marie Ody-Brasier, *Time and Punishment: How Individuals Respond to Being Sanctioned in Voluntary Associations*, 127 AM. J. SOCIO. 441 (2021) (analyzing social norms and punishment in defined groups); ROBERT C. ELLICKSON, ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES (1994) (arguing that societal norms create order in communities and regularize interactions between members). In the context of the scientific community, see ROBERT K. MERTON, THE SOCIOLOGY OF SCIENCE 223–80 (1973); *For Building Trust in Science: Good Scientific Practice (GSP)*, FRITZ LIPMANN INST., https://www.leibniz-fli.de/research/good-scientific-practice [https://perma.cc/9DCB-QKH6].

297. *See* Izet Masic, *Plagiarism in Scientific Publishing*, 20 ACTA INFORMATICA MEDICA 208, 208 (2012). For an example where such principles were not obeyed, see Dennis Normile, *Chinese Scientist Who Produced Genetically Altered Babies Sentenced to 3 Years in Jail*, SCI. (Dec. 30, 2019), https://www.science.org/content/article/chinese-scientist-who-produced-genetically-altered-babies-sentenced-3-years-jail [https://perma.cc/SQ5T-X7SV].

298. *See* Stuart P. Green, *Plagiarism*, *Norms, and the Limits of Theft Law: Some Observations on the Use of Criminal Sanctions in Enforcing Intellectual Property Rights*, 54 HASTINGS L.J. 167 (2002); Neil S. Morton, *Publication Ethics*, 19 PEDIATRIC ANESTHESIA 1011, 1012 (2009); Deepak Juyal et al., *Plagiarism: An Egregious Form of Misconduct*, 7 N. AM. J. MED. SCIS. 77 (2015); Ben Rosamond, *Plagiarism, Academic Norms and the Governance of the Profession*, 22 POL. 167 (2002); MARCEL C. LAFOLLETTE, STEALING INTO PRINT: FRAUD, PLAGIARISM, AND MISCONDUCT IN SCIENTIFIC PUBLISHING (1992).

299. *See* Ksenija Baždarić, *Plagiarism Detection–Quality Management Tool for All Scientific Journals*, 53 CROATIAN MED. J. 1 (2012); Tara C. Long et al., *Responding to Possible Plagiarism*, 323 SCI. 1293 (2009).

retracting publications, expelling scholars from academic positions, and damaging their reputations.[300] The norm of intellectual honesty—and the sanctions imposed for nonconformity—would discourage violation of the NDA.

Second, replicators' affiliated institutions can also be required to sign the NDA and to be jointly accountable for any violation.[301] This would create additional pressure to abide by the NDA, as a breach would harm replicators' home institutions economically and reputationally.[302] And while damages suits remain a sub-optimal solution, the potential for joint liability may change the calculus from the researcher's perspective as it may be both psychologically easier to sue an institution than to sue another scientist (as a lawsuit against another scientist may be perceived as non-collegial) and practically easier to obtain recovery if successful. On balance, the legal and professional risks associated with violating the NDA would create a strong incentive for replicators to comply.

Another potential criticism of our proposal concerns enforcement of authors' sharing obligations. What should journals do if an author declines a replicator's request for replication materials, despite signing the CAA? While the journals may have a valid claim for breach of the CAA, we do not expect them to engage in long, tedious, and costly battles. Indeed, while the CAA is a legal tool, its value is primarily to safeguard authors' rights in their IP. Instead of lawsuits, journals can consider ways to use the CAA as both a carrot and a stick, again harnessing the power of social norms. To induce authors' compliance with the CAA, journals could stamp a "Replication-friendly Authors" badge over the first page of a publication, for example. In contrast, the publications of authors who breach their duty to deliver replication materials pursuant to the CAA could be stamped, "Authors Refuse to Share Replication Materials." Both badges would deliver a powerful signal about the publication and its authors to the scientific community, which affects the authors' decision.[303] In fact, the badges concept—as a signal—can

300. Michael Koziol, *Plagiarism, Plagiarism, Plagiarism: Five Recent Cases*, RETRACTION WATCH (June 23, 2016), http://retractionwatch.com/2016/06/23/plagiarism-plagiarism-plagiarism-five-recent-cases [https://perma.cc/ZH6P-7DVL]; Green, *supra* note 298, at 197; Normile, *supra* note 297.

301. Importantly, the institution shall not be the sole party that is accountable for a replicators' infringement, as such a structure may minimize the replicators' risk and thus may encourage infringement.

302. *See, e.g.*, Yudhijit Bhattacharjee, *Saudi Universities Offer Cash in Exchange for Academic Prestige*, 334 SCI. 1344 (2011).

303. For background on the importance of signals in scientific publications, see *supra* note 286.

positively influence authors' willingness to use the CAA mechanism in the first place, and not only to comply with it.[304]

One of the main advantages of the CAA solution is that it imposes minimal transaction costs on replicators, authors, and journals. Replicators can address authors directly to ask for materials, using the automated NDA procedure, and journals would be contacted only as a last resort if authors refuse to share materials despite signing the CAA. Even then, journals are not required or expected to take legal measures against authors, but rather to apply simple negative reinforcements.[305] For authors, the CAA would entail minimum transaction costs, as the paper submission process remains almost the same, and the journal provides the applicable documents. Meanwhile, authors' affirmative obligations—sharing replication materials—are triggered only upon a replicator's request.

Note that a journal needs to formulate a CAA only once and embed it as part of its electronic submission platform.[306] All CAAs will be very similar, with minor modifications that the author can be make electronically through the submission platform itself, as is done for other submission forms, e.g., a conflict-of-interest form or a statement of ethics approval. The same argument applies to the formulation of the NDA—a general, strict NDA will be linked to all publications, and the electronic submission platform can automatically send it—upon a replicator's click-request—with zero to minimal human intervention. Accordingly, and considering the limited recourse journals are expected to take in the event of an author's breach,[307] the cost of the CAA mechanism is negligible. And while the CAA mechanism does place some burden on journals, it is reasonable to expect them to carry the burden for reasons discussed earlier—journals occupy a unique position in the scientific community and are best suited to implement the CAA, as they have an interest in mitigating the replicability crisis, and bear responsibility in making replication possible due to their very function as distributors of science.[308]

This Article focuses on the CAA's value as a potential solution to the problem of access to replication materials due to commercial/proprietary-related factors. However, the CAA mechanism is likely to alleviate other concerns a researcher may have with respect to sharing replication materials,

---

304. Jamieson et al., *supra* note 286; Van Dalen & Henkens, *supra* note 286; Munafò et al., *supra* note 286–303.

305. *Supra* note 286 and text accompanying notes 303–304.

306. Considering that there are about 2.5 million new publications a year, the marginal cost of the CAA aspires to zero. *See* Mark Ware & Michael Mabe, *The STM Report: An Overview of Scientific and Scholarly Journal Publishing*, 25 INFO. STANDARDS Q. 27 (2013).

307. *Supra* notes 298–299 and accompanying text.

308. *Supra* notes 263–277 and accompanying text.

such as the fear of being "scooped" or plagiarized[309] as the NDA would prohibit any use of the shared materials other than for replications.

Above all, the CAA mechanism advances a win-win situation.[310] Often, society is forced to balance competing interests, such as between the preservation of IP rights and replicability. Rather than pursue one at the cost of the other,[311] however, the CAA mechanism harmonizes the two. It provides both authors and the scientific community a way to promote their interests, i.e., IP-related considerations and replicability. Satisfying all parties becomes possible thanks to the unique channel of information sharing, which delivers complete replication materials through a confidential process. Of course, there are costs to such a solution, e.g., some complexity and risk. Nevertheless, it allows all parties to accomplish their goals.

A final caveat is in order: The CAA is not a panacea for the replication crisis, which is rooted in various factors, not only IP-related ones.[312] Whereas the CAA is designed to solve the problem of access to replication materials caused by IP, the scientific community needs to address other issues, including the lack of incentives generally to replicate studies.[313] We think that initiatives that target academic and professional incentives (rather than the commercial/proprietary interests the CAA is designed to protect) have a synergistic effect with the CAA solution, creating a virtuous circle of replications. Implementing the CAA policy may thus remove a significant barrier to replicability. As a result, the scientific community can expect alleviation of the replicability crisis in computational studies, and likely in other types of studies as well.

---

309. Heidi Laine, *Afraid of Scooping: Case Study on Researcher Strategies Against Fear of Scooping in the Context of Open Science*, 16 DATA SCI. J. 29, 29–30 (2017).

310. Bagley, *supra* note 210, at 224 ("In the interest of the public good, researchers should not have to choose between engaging in early-stage academic discourse and obtaining proprietary rights.").

311. Stodden, *Enabling Reproducible Research*, *supra* note 19; Stodden, *IP*, *supra* note 19; Shavell, *supra* note 19.

312. Although addressing mainly IP-related barriers, the CAA mechanism can facilitate other obstacles. *See* Laine, *supra* note 309 and accompanying text (regarding scientists afraid of scooping and plagiarism).

313. Aside from the CAA mechanism, there is a need for incentivizing further replicators, which can be best done through academic treats, such as establishing replication journals with a fixed impact factor or nurturing rewarding citation norms. For instance, one possible rewarding practice is to encourage researchers to also cite replication studies when referring to an original publication. By generating more citations, both the impact factor of replication journals and the academic indexes of replicators (e.g., h-index) will increase.

CONCLUSION

IP law and policy play an important part in shaping the incentives of scientists and organizations where scientific inquiry is conducted. This is truer today than ever, as the commercialization of academic research is on the rise. This Article examines the impact of IP law on the willingness of scientists to provide access to research materials required for replication. As our analysis shows, while copyright law does not play a significant role in this domain, the conditions for legal protection under patent law and trade secret law may strengthen the incentives to keep replication materials confidential, exacerbating the replication crisis.

To address this issue, we propose a contractual mechanism—the CAA—that, if implemented by scientific journals, would enable scientists to share replication materials while minimizing the detrimental effect that disclosure of those materials can have on their ability to secure IP protection. We firmly believe that adopting this policy would reduce the number of scientific publications that cannot be replicated due to restricted access. This would promote open science for the benefit of the scientific community and the public at large.