Ledizioni

Rossana Morriello

From Book Piracy to Predatory Publishing

A Journey Through the History of Printing and the Ethics of Scholarly Communication

Rossana Morriello

From Book Piracy to Predatory Publishing

A Journey Through the History of Printing and the Ethics of Scholarly Communication

LEDIZIONI

2024 Ledizioni LediPublishing
Via Boselli 10 – 20136 Milano – Italy
www.ledizioni.it
info@ledizioni.it

Rossana Morriello,
From Book Piracy to Predatory Publishing. A Journey Through the History of Printing and the Ethics of Scholarly Communication

First edition: December 2024

Print ISBN: 9791256002948
ePub ISBN: 9791256002955
PDF Open Access ISBN: 9791256002962

Catalogue and reprints information: www.ledizioni.it,
www.ledipublishing.com

Table of Contents

Science is but a perversion of itself unless it has as its ultimate goal the betterment of humanity.

Nikola Tesla

Preface
by Paola Castellucci

Scrolling through the rich and up-to-date Bibliography accompanying Rossana Morriello's work, one is immediately struck by the large number of question marks in the titles. Umberto Eco pointed this out almost half a century ago in *How to write a thesis* the use of direct interrogatives and exclamations should be avoided because it denotes a style that, at best - can be described as journalistic. And if it should be avoided in a dissertation, even more so in a scientific article; and all the more so if the topic is, as in the present work, the reliability of scientific literature taken by storm by *pirates* and *predators* (and infested with hoaxes, parodies, forgeries, plagiarism, counterfeits, whistleblowing, fakes, all phenomena that Morriello punctually analyses). It would seem like a contradiction in terms: an accurate, skillfully written, well-argued scientific essay on an important topic ends with references to sources using apparently unscientific language. Should those titles have been excluded from the bibliography? Should the presence of a question mark have been taken as an indication of a lack of seriousness? Certainly not. If anything, the large number of questions in the bibliography of *From Book Piracy to Predatory Publishing. A journey through the history of the press and the ethics of scholarly communication* immediately gives an account of the serious questions that arise for those who are dismayed when faced with an authority figure (the publisher) who fails in his noblest prerogatives, his role as evaluator and guarantor, and instead engages in unscrupulous commerce.

The transaction (certainly not a *philosophical transaction*!) is very clear: the publisher (who has lost values and therefore identity) manufactures, in exchange for money, products (not research products, as they say nowadays) for an author (who has lost values and therefore identity) who wants to strengthen his CV in order to gain access to an academic position, or to improve it. An author (one would like to put the name in inverted commas) does not undergo peer review, and sometimes even does not write, and in other cases does not even research; rather, he pays for a service that a publisher (one would like to put the name in inverted commas) cleverly packages. Perhaps the article is not even printed but exploits the credit of Open Access (the mask with which the predator operating online chooses to disguise itself). All

without the transparency of Open Access, a strongly ethical movement based on freedom, clear policies, and FAIR principles of publication and reuse.

The phenomenon examined by Rossana Morriello's work is predatory publishing. A topic treated here without sensationalist accents, without shouting and accusations, but with much study and many reference sources. If the immediate context is the age of the Net, Morriello relates the contemporary to the past (Ch.1, *Book piracy and the debate on intellectual freedom*). A similar path was taken by Robert Darnton (cited by Morriello) in *The Case for Books: Past, Present and Future*. If the phenomenon of publishing piracy was complex to define and curb, even now the identification of predator and prey is not immediate, nor are the motivations for choosing *the wild side* (according to the work of Bagues, Sylos Labini, Zinovyeva). Remaining within the metaphor, if the image of the pirate also exudes fascination and promises of alternative values, and if the address of the arXiv website, the first repository in history, flies the pirate flag, and if we are enthralled by Bruce Sterling's *Utopia pirate* story of Turin, we fail to feel empathy for the predators precisely because they do not position themselves as alternative heroes. Indeed, the contours of predatory publishing appear uncertain and sometimes overlap and blur with the healthy zone. As Ernest Abadal, for instance, points out in recent contributions, distortions with respect to the review system are also revealed in scientific journals, when monetary contributions are demanded for the practice of evaluation (APC- article processing charges). Morriello, therefore, highlights how difficult it is to draw a clear line between licit and illicit.

> *Predatory publishing is a* phenomenon that affects all types of publication but predominantly journals (*predatory journals*). These are journals that claim to be scientific when in fact they are not, even though they can be deceptive on the surface due to the presence of ISSNs, the use of titles that mimic, and thus resemble authoritative journals, and their presence in databases, including citation databases. It is not easy to define predatory publishing precisely, as it is by no means easy to establish all its characters and delimit its contours [...]

Predatory publishing exploits a 'middle ground': offering a ready-made publication, predatory journals tailor a product for a precisely targeted use. Not a product of research but rather a pure and simple product, i.e. the result of a market survey, so to speak. If the parameters of evaluation demand fast publication in international journals impacted and surveyed by databases, then predatory publishing offers a smuggled passport, without having to go through peer review evaluation. Titles to list on a CV can thus be 'fabricated' quickly. By paying, of course. These titles are expendable for a competition because, appar-

ently, they have all the prerequisites that the practice of quantitative evaluation requires, as well as non-bibliometric evaluation, which is in any case propped up by thresholds, medians, numerosity, ranked journals, internationalisation levels. Predatory publishing sets up a targeted business to meet these standards, at least in appearance. Like buying indulgences; or obtaining certificates of conformity for an ecological transition only in appearance (*greenwashing*). And if the article has not yet been written and if perhaps the research has not yet been conducted, the most unscrupulous predators offer, at an even higher price, a ready-made article, written by a *ghost writer* or assembled from other articles already published, thanks to an algorithm; and on this point, incidentally, it must be said that much remains to be investigated, as Maurizio Lana's research on the subject of *agency is* already revealing.

A true marauding story, then. A story that makes one tremble. A story that raises fears and questions posed initially in solitude, as if in front of a mirror that suddenly appears distorted. The researcher who chooses to tell this very story sees himself, talks about himself and his experiences, and thus about fears, risks, and obstacles in the adventure of knowledge. The approach is necessarily metadiscursive: in doing research, one wonders about the ways, values, aims, and dangers of research itself. Rossana Morriello therefore places the profound, existential, ideal, emotional, but also social and political reasons for research at the heart of the book (Chapter 2, *The ethics of scientific communication: a historical perspective*). All around (Chapters 1 and 3), as in a fable, the forest is full of dangers, pitfalls, in the past as in the present. The risk, as in any adventure, is to lose sight in every sense. For example, the loss of trust. If the publisher fails in his duty as a guarantor of quality, he breaks a pact of trust not only with authors and readers, but with the entire community that awaits the results of research and trusts and relies on the progress of science.

Here then are the *questions*, the *issues* that attract attention in the Bibliography (Chapter 3 is actually entitled *Predatory publishing and other issues*). These are not rhetorical questions, nor are they effect formulas. They are not lapses in style, as Umberto Eco stigmatised. If anything, faced with the fear of losing the references for a personal and collective identity, the Bibliography on the subject poses real and urgent questions. Faced with the ancient practice of publishing piracy, faced with the shocks suffered by the ethics of science, frightened by the more recent phenomenon of predatory publishing, questions remain.

Burning questions are put forward with all the disappointment and despondency of a researcher who sees values in which s/he believes offended. Let us try to list these questions. We repeat them in the simple alphabetical order in which they appear in the Bibliography and therefore not in a logical, consequential order with respect to the subject.

Yet even so, in the order in which they appear in the Bibliography (alphabetical by author), they offer a powerful synthesis of the problems addressed by Rossana Morriello:

*La frode scientifica va perseguita come un reato? (*Scientific fraud should be prosecuted as a crime*?)*

Why Should We Worry about Predatory Journals?

Image manipulation in scholarly publications: are there ways to an automated solution?

Already in the late 1990s, the famous Sokal hoax (named after the author who managed to have a nonsensical article accepted by an A-rank journal, counting on the fact that the reviewers did not actually read or were not experts in the field) had shaken up the apparently solid structure of double-blind peer review. On this point, too, Open Access has expressed itself ever since Paul Ginsparg's choice to build the arXiv repository by choosing the preprint as the documentary type since 1991, supported only since 2004 by the practice of endorsement (i.e. submitted by an author-guarantor of the quality of the contribution). In the bibliography selected by Morriello, Sokal's mockery is also present. But the mockery has a bitter tone:

The Sokal Hoax: At Whom Are We Laughing?

Is the solution, then, in reinforcing the very practice of peer review, rather than mocking it?

Digital magic or the dark arts of the 21st century - how can journals and peer reviewers detect manuscripts and publications from paper mills?

Or perhaps, in a now scalar dimension like that of the Internet (and also of scientific production), is a certain degree of criticality inevitable?

Is Biblioleaks Inevitable?
How many scientists fabricate and falsify research?

However, the fixed point remains: the ethics of science as a guarantee of truly useful results for human progress. As, moreover, appears in the exergue chosen by Morriello, and as the question reiterates in another title in the Bibliography:

Scientific publication - Is it for the benefit of the many or the few?

And again, how was it possible that even prestigious editorial boards were victims of misconduct and then had to ask for the article to be withdrawn?

The Lancet has made one of the biggest retractions in modern history. How could this happen?

Duplicate and fake publications in the scientific literature: How many SCIgen papers in computer science?

Quand les articles scientifiques ont-ils cessé d'être des communs?

Zombie papers: Why do papers by the most prolific fraudster in history keep getting cited?

And again:

Does the philosophy literature have a plagiarism problem?

Predatory Conferences: What Are the Signs?

Going even deeper into the phenomenon, isn't it precisely the practice of quantitative evaluation, the false mirage of merit and objective assessment, and isn't it the blackmail of *publish or perish* that drives illicit acts that are harmful to the entire community? On the other hand, the very word 'meritocracy' is a neologism taken from a post-war dystopian fable. So, you don't solve the problem by catching the marauders and hanging them from the highest pole. The root cause lies elsewhere and is not external:

What pushes scientists to lie?

Evaluation practices need to be rethought. We need to return to trust in the ability to evaluate; we need to trust in the responsibility that evaluation entails. Rather than the lightning speed of a pirate, rather than the sheer number of products amassed by a marauder, rather than the quality of seemingly fresh merchandise opportunistically designed for the international market, we need to rely on the slow and deliberate pace of reputation:

Time to assume that health research is fraudulent until proven otherwise?

And finally, an exclamation:

Congrats! Your paper was accepted. (Except if the acceptance letter was forged)

To all these questions, Rossana Morriello tries to give answers, tackling the problem with a confident and firm voice. After long service as a librarian, after a Ph.D. and years of research and many publications, Morriello offers us possible solutions, once again, inevitably, in the form of questions, issues, and dilemmas. On the last page, Rossana Morriello greets the reader and offers her work, with the noble humility that is proper to the ethics of service and research:

> The body of knowledge, the basis on which further knowledge is built in the scientific process, is undermined by numerous distortions, ethical violations, and fraudulent acts. University structures, departments, and laboratories are increasingly dependent on quantitative evaluations, the problems of which are well known and which drive deviant behaviour. We are faced with a varied set of problems that individually may not have particularly high numbers but collectively pose a threat to the culture of scientific research, the survival of the scientific method, and a major risk to the future of science. The solution is first of all to abandon, or at least scale back, quantitative methods for research assessment, as suggested by the important initiative on research assessment reform launched in 2022 jointly by the European Commission with EUA and Science Europe, which has produced the document *Agreement on Reforming Research Assessment* that universities and other non-profit organisations are called upon to sign. But it is equally imperative to try to strengthen scientific structures, starting with peer review, rethinking and refounding them, where necessary, to adapt them to a changing social and communicative context, and to curb the behaviour that weakens them. Finally, spreading awareness of these issues, both among editors and researchers, perhaps through specific training within universities, so that they can understand how to recognise and avoid them, is certainly essential. We hope to contribute to this goal with this volume.

Introductory note

It is widely believed that some of the problems that characterise today's scholarly communication are a phenomenon of modernity. Issues such as predatory publishing or the violation of research ethics and integrity are not infrequently considered as a consequence of the spread of publications in digital format, or the causes are traced back to the expansion of open access. The idea that making publications open access can facilitate plagiarism and scientific fraud is common among researchers and publishers and has been one of the limits in the spread of open science, especially in Italy. In fact, although open access may under certain conditions increase the risk of incurring illicit or ethically incorrect behaviour, it is also true that making the results of scientific research openly and freely accessible increases the visibility and circulation of publications, and thus facilitates the detection of distortion phenomena that would otherwise be more difficult to detect.

The peer review process, the foundation of a publication's scientificity, is carried out with the participation of a limited number of reviewers, whereas with open access the scientific content becomes public and extended to a much larger number of peer experts in the field. However, it is now well established that the *publish-or-perish* culture generated by today's scientific research dynamics, and especially by research evaluation systems, has produced the intensification of certain phenomena. We will dwell on these aspects, trying to provide an overview of the stresses that undermine the integrity of research today. However, to better understand the present and to look at current phenomena with a holistic approach and not flattened on contemporaneity, it is indispensable to position oneself in a historical perspective that allows one to define the present dimension as a stage in a journey and enables one to understand them in greater depth. Only from this perspective can one then attempt to identify some viable paths towards a sustainable future for research. Indeed, as Robert Darnton writes, "any attempt to see into the future while struggling with problems in the present should be informed [...] by studying the past"[1].

Looking back, in fact, it is clear that digital and open access cannot be interpreted in a cause-and-effect correlation with the issues we are

1 Robert Darnton, *The Case for Books: Past, Present, and Future,* New York, Public Affairs, 2009, p. XII.

going to deal with, because questions of research ethics and integrity and some less than legitimate practices in scientific publishing have their roots in a distant era, even before the birth of modern science and the first scientific journals. The history of science is traversed by innovations and transformations that have led to major changes at some stages, and these evolutions have gone hand in hand with the history of printing and the evolution of publishing, so much so that they have become inseparable aspects at some historical stages. While being fully aware of the substantial differences between the present, to which this work turns its final gaze, and the past on which it initially dwells, and keeping well away from the idea that certain phenomena are inevitable and irresolvable precisely because they have always existed, there is no doubt that retracing historical events can provide the diachronic context that is indispensable for looking to the future, all the more so if, as we can assume, we are in a phase of paradigm change. Indeed, as Thomas Kuhn argues[2] , the proliferation of conflicting theories, of different phenomena questioning the existing paradigm, is indicative of a phase of crisis that will probably lead to the emergence of a new paradigm. We are certainly going through such a phase, but it is too early to say what the new paradigm will look like. However, this is not the aim of this volume. Rather, the intent is to gather and analyse the clues of the change taking place. To achieve this goal, I have taken advantage of the advice and suggestions of several people to whom my sincere thanks go: Maria Teresa Biagetti, Andrea Capaccioni, Mauro Guerrini, Mario Infelise, Maurizio Lana, Giovanni Paoloni and Maurizio Vivarelli.

2 Thomas S. Kuhn*, The Structure of Scientific Revolutions,* Chicago, University of Chicago Press, 1962. .

Chapter 1
Book Piracy and the Debate on Intellectual Freedom

1. The origin of piracy

The origin of the concept of piracy in publishing dates back to 17th century England and has a history closely linked to that of movable type printing and the evolution of thought and regulation on intellectual property, which was gaining ground in the same century in a context of transformations initiated by the general renewed conception of knowledge during the Renaissance. In other European countries, as well as later in the American colonies, there were similar events, but what happened in England takes on a particular relevance because it was there that the first copyright law was born, and the process that led to its enactment is representative of the many intertwined dynamics at play when it comes to modern science and scientific publishing. Many of these dynamics can still be found today, along with the centrality of copyright in the relationship between authors, publishers, and libraries. Although we have obvious differentiations due to the intervening centuries, we find today many affinities with that historical period of great ferment that led to the first copyright law, that make us predict a similar ongoing process of transformation, to which we will return later. For this reason, and because dealing extensively with the events in the various countries would be beyond the scope of this work, we will take the English perspective as an example, not excluding, where useful, references to other nations as well. Moreover, the dominant position of English-speaking countries in the field of scientific publishing today is well known, and thus the history of those nations is particularly significant with respect to the current publishing scene.

A milestone in the history of science and scientific publishing is the work of the Royal Society and, particularly, Henry Oldenburg's initiative to found Philosophical Transactions, one of the first scientific journals. Without for the moment entering into the merits of the diatribe that sees Philosophical Transactions contending for the primacy of the first scientific journal with the 'Journal des Savants', there is no doubt that, in general, the first periodical publications of a scientific

nature had the function of cementing the printing revolution and the scientific revolution in such a way as to make this link seamless[1] and certainly indissoluble for centuries to come. However, this pivot point was preceded by complex events in publishing. In England, ever since the introduction of printing by William Caxton in 1471, the Company of Stationers operated in London, a corporation that brought together printers, booksellers, and bookbinders, whose activities were initially not very distinct from each other, and which maintained a register, the *Stationers' Register*, in which those who wanted to publish a book had to register the publication. The registration had the function of protecting the printer or bookseller against the possibility of someone else intending to print the same work. In general, innovations in knowledge in those days were welcomed and supported by the granting of privileges, which took on a formal aspect for the first time in 13th century Venice. Such privileges were granted by the sovereign for innovative activities capable of producing a benefit for the community, as printing was. As early as the 15th century, almost all European states granted them and were organised in this way, also for the activity of printing.

In the 17th century, turbulent social and political events in England were a stimulus for changes in the book trade. Printing had become a means of facilitating the circulation of political ideas, and for this reason the system of privileges granted by the king, as well as the Stationers' Register itself, began to be seen as a restriction on the free circulation of printed material and to be the subject of dispute and proposals for change. At the same time, the traditional distinction between the liberal arts and the applied arts was beginning to break down, not only in relation to printing, with the entire system of arts and sciences constantly being subject to pressures and pushes towards change. The figure of the printer was separating from that of the bookseller, and a distinction was thus being created between the latter's commercial activity and the skills and craftsmanship mastered by the printer. The bookseller sold copies of books without possessing the tools and techniques to produce them, and thus the activity became purely financial. Moreover, the complex English situation in the years of the Revolution, and later the Restoration, laid the foundations for the emergence of the concept of 'authorship' and then the idea of intellectual property. The emergence of these concepts went hand in hand with the evolution of piracy, which played a decisive role in the debate on authorship in an interconnected set of events that led to the enactment of the Copyright Act in 1710.

1 *Adrian Johns, Piracy. The Intellectual Property Wars from Gutenberg to Gates,* Chicago and London, The University of Chicago Press, 2009, p. 59.

The word piracy, in English *piracy*, has an Indo-European root[2] meaning 'an attempt, experiment, or experience' and thus originally had a positive connotation, associated with a certain creativity, exactly the opposite of how we consider it today.[3] The meaning began to change in antiquity, as attested in the works of numerous authors, such as Galen, Quintilian, Vitruvius and Cicero, who complained about the spread of falsely attributed and unauthorised works. The physician Galen in the second century wrote *De propriis libris liber*, "to present, carefully classified, his authentic works and to denounce those falsely attributed to him"[4] . Martial uses the term *plagiarius*, understood as tormentor, speculator, and tyrant to indicate one who appropriates someone else's writing[5] . However, these claims were mainly determined by the fear of seeing one's own ideas circulated in a manner distorted from the original. Non-original works were not considered condemnable, nor were the authors accused of any form of crime. As Adrian Johns recall, we find the same fear manifested in later centuries, for example, in some works by Shakespeare, Spenser, Marlowe, and Milton. In Elizabethan England, the presence of pirates extended to literary works as a reflection of what was happening in society, and, as in reality, the figures of the pirates were not entirely negative, for although they operated contrary to the laws, they represented in popular eyes the revenge of the poor and deprived (who often really in real life engaged in piracy out of necessity) against the rich and a way of subverting an unjust social system. Although it had been a phenomenon for a long time, it was only in the 17th century that the negative connotation in the meaning known to us was accentuated, and not only in relation to publishing. What contributed to it entering strongly into the collective imagination, in fact, was the propagation of news and tales of sea travellers telling of the raids of thieves (pirates), so that in the society of the end of the century, piracy was spoken of in every sphere and in every venue. In general, the attribute 'pirate' denoted an action that threatened public order, since it was outside the conventions of civilised society[6] . The linguistic element is obviously not the cause, but only the consequence of a process of cultural, social, and political transformation that will lead

2 Recall that the English and Italian languages, like most other European languages, are all descended from a single Indo-European strain and therefore the root is common.

3 A. Johns, *Piracy*, cit, p. 35.

4 Luigi Balsamo, *La bibliografia. Storia di una tradizione*, Milano, Unicopli, 2017, p. 17.

5 Walter J. Ong (1982), *Orality and Literacy.The technologizing of the Word*, London and New York, Routledge, 2005, p. 128.

6 A. Johns, *Piracy*, cit., p. 7-15.

to the intersection of cultural elements with commercial aspects in the modern conception of piracy.

In cultural and editorial circles, the first time the term '*wit-pyrats*' is attested is in a 1611 reference by John Donne, followed by a reference by Samuel Butler, who called a person accused of plagiarism a 'wit-caper,' with the word '*caper*' referring to a Dutch privateer. However, in both cases, the reference is to plagiarism on a personal level and not to commercial practices. Toward the end of the century, the concept of piracy emerged widely and is found in the words of major authors such as Defoe, Swift, Addison, Congreve, and Pope. In 1757, James Buchanan's dictionary *Linguae Britannicae vera pronunciatio* provides a definition of the word 'pirate' as "one who unjustly prints another person's copy",[7] thus shifting the responsibility from the person who writes the work to the person who prints it. The term flourished in the London book market environment and then spread to other geographical areas and fields. Between the end of the 17th century and the beginning of the 18th century, it was found in dictionaries in France, then in Italy and Germany. The transformation of the concept of piracy went hand in hand with the political events that characterised England in that century. As reported by Adrian Johns, in this process, a prominent role was played by Richard Atkyns, a fierce detractor of the Stationers' Company, which in the meantime had become an oligarchic and speculative corporation in which the group of booksellers was now completely separated from the group of printers. Atkyns, convinced that the printing system organised in this way was one of the causes of the turbulent fortunes of English society, induced Charles II to proclaim the art of printing the property of the Crown, restoring a privilege granted a century earlier by Elizabeth I. Thus, in 1662, the Press Act was enacted, which stipulated that published works, including the work of the Stationers, should be subject to the control of the Crown in an attempt to counteract the printing of antimonarchist material. But the restrictions imposed by the Press Act were often ignored. Unlicensed printers continued to print and used various ploys to publish without a licence, operating in a way that made it difficult to enforce fines or convictions, which, in fact, were not frequent. The system was also perfectly within the policy of the absolutist government of Charles' successor, James II Stuart, who ruled until 1688, the year of the great revolution.

From 1680, references began to appear in the Stationers' Company register to breaches of practice by people referred to as 'pirates.' The spread of the term was a consequence of the numerous stories and legends about piracy of the seas that reached England and spread to the press, as well as to other sectors (food, spices, and various artefacts), entering the collective imagination and contributing to the general adop-

7 *Ivi,* p. 23.

tion of the term in society. In the years preceding the Great Revolution, the climate of social and political ferment in English society had also been expressed and amplified through the press, and debate had become lively in every sector. The need for information and the circulation of ideas found an easy vector in the medium of the press. The practice of circulating news in oral and manuscript form, already frequent during the reigns of James I and Charles I, had gradually been joined by its transmission in printed form, with publications that were easy to print and circulate. Until 1620 these were mainly *broadside* (or *broadsheet*), i.e. single sheets, often illustrated, with heterogeneous material such as announcements and reported facts, poems, popular songs, or pamphlets. From 1620 onwards, the first journalistic publications began to appear, the *corantos* or *newsbooks*, forerunners of newspapers (the term 'newspaper' would be coined in 1670), which collected news from all over Europe and reported it at regular intervals, usually weekly. Their origin can be traced back to the *notices* printed in Venice until the mid-15th century. The first known English coranto was printed in Amsterdam, since a parliamentary law prohibited the publication of national news in England. It is a single sheet printed on both sides with news of the Thirty Years' War, entitled *The New Tydings Out of Italie Are Not Yet Come,* and dated 2 December 1620. The first English daily newspaper was the Daily Courant, which began publication on 11 March 1702. In 1709, one of the most influential periodicals was born, The Tatler, which later became The Spectator.

The first English printed newspaper was the Oxford Gazette, published in 1665, which became the London Gazette the following year and consisted of one sheet printed on both sides, published twice a week. The earliest forms of printed newspapers flanked, but did not supplant, the manuscript newsletters that had been widespread since the 14th century, known as 'relations', of which the earliest surviving copy is an account of the Battle of Flodden in 1513, entitled *Hereafter Ensue the Trewe Encountre or Batayle lately Don betwene Englande and Scotlande*. Henry Muddiman, the founder of the London Gazette, held the title of 'journalist to the king' and had previously published both handwritten newsletters and printed newsbooks for over a decade.[8] The content of handwritten newsletters and printed newsbooks or newspapers was often similar, sometimes complemented each other, and the readers were common because the two media together offered comprehensive coverage of events. The survival of handwritten newsletters was mainly due to greater freedom of expression of content that avoided press censorship. The first continental newspaper was instead from the German

8 Rachael Scarborough King, "The Manuscript Newsletter and the Rise of the Newspaper, 1665-1715", *Huntington Library Quarterly*, 79 (Autumn 2016) 3, p. 411-437: 418, https://www.jstor.org/stable/10.2307/huntlibrquar.79.3.411.

area, the Mercurius Gallo-Belgicus printed in Cologne in 1592 with the collection of news mainly from the Netherlands and France. The newsbooks were multipage pamphlets and contained not only news from abroad and perhaps already published elsewhere, but also local news, especially on political and social issues. They became widespread from the 1640s, and it is estimated that in the years between 1640 and 1650 about 300 titles came out with peaks of 1500 copies printed.[9]

The proliferation of new channels of information and communication was, of course, not limited to England. With different forms and names, often adapted to the nation or local community (such as mercurio, notice, newsbook, and others), from the 16th century onwards there was such an intense and internationally interconnected increase of such channels throughout Europe that any attempt to separate the channels through which news was spread and distinguish them with specific names, would be an artificial simplification of a context that was instead made up of a variety of formats and names and of networks through which they were distributed throughout Europe, which were, moreover, interconnected.[10] In Italy, in addition to *notices (avvisi),* news found a home in the *gazzette,* a term attested since 1570, that were collections of both manuscript and printed news often considered unreliable and not authoritative, and in the *fogli*. From the second half of the 17th century, the word '*giornale*' also appeared in the title of the first literary and scientific journals inspired by the Journal des Sçavans, such as the Giornale de' letterati in Rome (1668), the Giornale veneto de' letterati (1670), and the Giornale de' letterati d'Italia (1710)[11] . The different authoritativeness of the two types of channels, the former not very authoritative, the latter respected and credible, led to the clear distinction between gazettes and newspapers and the equally clear separation between the profession of gazetteer and that of journalist in the following century.

Alongside these different genres, the first scientific journals were born within the academies. In 1665, the aforementioned journal of the Royal Society, Philosophical Transactions, was founded, which still con-

9 Moira Goff, *Early History of the English Newspaper: 17th and 18th Century Burney Newspapers Collection*, Detroit, Gale, 2007, https://www.gale.com/binaries/content/assets/gale-us-en/primary-sources/intl-gps/intl-gps-essays/full-ghn-contextual-essays/ghn_essay_1718bnc_goff2_website.pdf.

10 Paul Arblaster, André Belo, Carmen Espejo, Stéphane Haffemayer, Mario Infelise, Noah Moxham, Joad Raymond, Nikolaus Schobesberger, *The Lexicons of Early Modern News*, in *News Networks in Early Modern Europe*, edited by Noah Moxham and Joad Raymond, Leiden, Brill, 2016, p. 64-101.

11 *Ivi*, p. 68. See also Jean-Pierre Vittu, "Du Journal des savants aux Mémoires pour l'histoire des sciences et des beaux-arts: l'esquisse d'un système européen des périodiques savants*", Dix-septième siècle,* 228 (2005) 3, p. 527-545, https://doi.org/10.3917/dss.053.0527.

tends today for the title of the first scientific journal with the Journal des Savants. The first was founded by Henry Oldenburg, secretary of the English Academy, of whom it was an emanation as a means of publicising the Society's activities, also to receive support from patrons. The second was published a few months earlier by the lawyer Denis de Sallo, using the pseudonym of Sieur d'Hédouville, with the intention, stated in the first issue, of reporting on scientific innovations, inventions in the field of engineering and meteorological observations, discoveries in the field of biology, but also to give an account of the activities of the Court and to cover any subject of interest to a man of culture of the time. The journal also included literary criticism and dealt with different topics as part of the cultural man's baggage. These hybrid characteristics of the content, combined with the fate of the journal whose publication was suspended soon after its inception due to religious pressure[12] , often lead one to consider the Philosophical Transactions as the first journal with exclusively scientific contents. This academic dispute, apart from a manifest nationalistic competition, is indicative of a separation that was being created at that time, at the dawn of modern science. Scientificity is recognised for a research method that is typical of the experimental and applied sciences, those prevailing at the Royal Society, while the humanities, which make use of other methods and of which the literary sciences promoted by the Journal des Savants are a part, remain separate. A separation between the areas of the hard sciences and the humanities that would persist over time and whose consequences we still see today, with the exaggerated division between the two areas in a society that would instead require their union in order to deal effectively with the growing complexity[13] .

In England, the proliferation of publications in the latter part of the 17th century was impressive, particularly in the years following the Glorious Revolution, with the deposition of James II and the arrival of his successors. Indeed, in 1695, with the accession to the throne of William III of Orange-Nassau and his wife Mary Stuart, the Press Act

12 The interruption lasted only a few months and in 1966 De Sallo was able to start publishing it again, but the issues were irregular for a long time, until 1724 when it began to have punctual monthly issues.

13 Instead of being resolved, this separation is now accentuated by the methods of evaluating scientific research. As Paola Castellucci writes, "the very fact of pre-defining evaluation methods by dividing them into bibliometric/non-bibliometric, has insinuated that the qualitative method can only be defined in negative terms with respect to the quantitative one" and adds "as if to say that for the scientific-technical area there are weights and measures, while in the humanistic area one goes ahead in an impressionistic way, judging "by eye"", see Paola Castellucci, *La visione del giudizio. Una prospettiva romantica*, in *Libri, biblioteche e società. Studi per Rosa Marisa Borraccini*, edited by Alberto Petrucciani, Valentina Sestini, Federico Valacchi, Macerata, EUM, 2020, p. 413.

was repealed by Parliament, and all government control over the press ceased to have effect. The years between 1695 and 1710 were consequently a period of proliferation and free circulation of printed works in all forms, authorised, and especially pirated. It was a significant period because, as the law fell away, the concept of piracy in the legal sense was replaced by the cultural concept of piracy, shifting the focus to other aspects[14] . The entry of piracy into the more purely cultural sphere meant that the debate on piracy, along with the spread of piracy and the multiplication of pirated works in every nation, became central throughout the next century, so much so that it can be said that without piracy, there would have been no Enlightenment[15] . The issue of piracy was too closely intertwined with the demands for intellectual freedom, freedom of the press and expression, and the relationship between intellectuals and power that characterised that century not to be fully integrated into it. Moreover, it was often pirate editions that facilitated the circulation of Enlightenment ideas within and across European nations. In England, for example, pirated copies of Philosophical Transactions helped to increase the prestige of the journal and the Royal Society[16] .

Looking closely at the situation at that time, one finds several similarities with the current coexistence of printed and online information, not only because even today the readership of printed and digital journals is often the same but also because of the greater freedom that is accorded to digital communication, at least when it is placed outside the constraints of various kinds imposed by the 'official' press'[17]. Even today, the constraints are mainly represented by the entrepreneurial ownership of newspapers, the configuration of the scientific publishing market, and in any case by commercial motives, and of course by copyright laws. Today, we must deal with the growing problem of fake news and the complete freedom of circulation of any kind of content online, without any ethical, appropriateness, or other considerations. Even for this aspect, whose declinations in the scientific sphere we will see later, a parallel can be drawn with newsbooks, gazettes, and the like, which boasted truth and accuracy, but were not infrequently studded with false news and errors and tended to be unreliable. Another interesting element of contiguity is the flexibility that characterised the

14 A. Johns, *Piracy,* cit, p. 43.

15 *Ivi* p. 50.

16 *Ivi, p. 46-50.*

17 Actually, open access aside, not only is scholarly publishing dominated by the platforms of large publishers but also the current configuration of the World Wide Web does not offer many spaces effectively free from the power of commercial platforms, as well explained in Josè van Dijck, Thomas Poell, Martijn de Waal, *Platform society. Public Values in a Connective World*, Oxford, Oxford University Press, 2018.

newsletter, which for a long time was also handwritten, both in terms of content in its various forms such as letters, oral reports, comments, and in terms of speed of updating, in comparison to the other print media that took a long time[18] . Today, digital expands both the concept of multimedia formats and the speed of content dissemination, with implications for the world of journalism and the world of scientific publishing, also in relation to issues of ethics and integrity, as we shall see later. Another analogy with contemporary times is the emergence, in the phase of lively concatenation of events discussed above, of new publication formats that flanked the already existing ones.

2. Piracy in the publishing system

The concept of piracy thus dates back to the mid-17th century, although the phenomenon precedes it, and will be influential in relation to the enactment of the first copyright law with the Statute of Anne, promulgated in 1709 and enacted in 1710. The Act, with the long title "An *Act for the Encouragement of Learning, by Vesting the Copies of printed books in the Authors or Purchasers of such copies, during the Times therein mentioned"*, is known as the first Copyright Act, although the term copyright is not present in the text of that Act but only appears around 1730, only to become established in the following decade and take on its current connotation in the second half of the 18th century. The law was urged by the *congers, the* alliances of booksellers that had been formed since 1670 and intensified after the repeal of the Press Act in 1695, which convinced Parliament to enact the law also in order to combat widespread piracy. The intention was to control the printing of works so that especially the most important and profitable works would remain in their hands, and thus prevent them from being printed by those outside the covenants. The booksellers of the *congers* considered themselves owners of the printed works and operated in such a way as to exclude others from the possibility of printing further copies.

Anna's Statute aimed to regulate the subject matter and promote learning and knowledge, as stated in the title, and set the duration of copyright at fourteen years (twenty-one for works published before its entry into force), renewable once if the author was still alive. Copyright was recognised as an author's right but held by the booksellers, still as a concession from the monarchs, in fact, as a form of privilege. The new law also limited the power of the booksellers in another way. The statute modernised the principle of legal deposit in libraries, a rule established in 1610 through an agreement between Sir Thomas Bodley and

18 R. S. King, "The Manuscript Newsletter and the Rise of the Newspaper", cit., p. 430.

the Stationers' which provided for the deposit of the best copy of every printed volume at the Bodleian Library in Oxford. In 1637 the rule was extended to Cambridge and the Royal Library. The legal deposit was repealed in 1662 and then incorporated into the Press Act, but in any case, it was always poorly respected by booksellers. The Statute of Anne strengthened it and added an additional six copies on legal deposit to be sent to university libraries in Edinburgh, Glasgow, St. Andrews, Aberdeen, Sion College, and the Faculty of Advocates in Edinburgh. For works sent to the libraries, the copyright was extended to 28 years. The intention was to make books available to everyone and to encourage the spread of knowledge. But booksellers interpreted the new law liberally and in their favour, intending that only books registered at Stationers' Hall were subject to the legal deposit requirement and began not to register important and, therefore, more saleable books, sending only works of lesser value and relevance to libraries, which were instead registered regularly. During the 18th century, book collecting was spreading, and legal deposit, which guaranteed the free availability of books in libraries, was also a problem for collectors as it could affect the market value of a work. In fact, in this century, the interconnections between cultural and economic aspects were strong, and the role of booksellers was important. The appearance of the term "bibliography" in the third decade of the century dates back to this period, with a redefinition of the concept that took on not only the meaning linked to the knowledge of books but also that of a technique for the description of books, especially rare books, and thus a form of mediation between book production and circulation[19] .

At the turn of the century, several publishers began to protest against copyright as well as against legal deposit, going so far as to declare, as Sir Egerton Brydges did in 1818, that the demand for free copies from university libraries was "the plea, not of the beggar..., but the robber!"[20] , adding the fear of reduced profits because potential buyers would be able to consult the library copy of the book instead of buying it. A fear that incidentally, still exists in our day, as many publishers declare. The issue is complex and multifaceted, but what we would like to emphasise is how since then the power of publishers has asserted itself clearly, without ever weakening to this day, and even strengthening in some historical phases, and how they have always found ways to ally and agree to work to their own advantage, even freely interpreting the laws, and not infrequently finding themselves very much out of sync with the mission of libraries. The aspiration of libraries, which after all is also the aspiration of scientists, as can be deduced from the impulse that led to the birth of Philosophical Transactions, to share and make knowl-

19 L. Balsamo, *La bibliografia,* cit., p. 13.

20 A. Johns, *Piracy,* cit. p. 235.

edge available to all, has clashed since the dawn of publishing with the commercial interests of publishers whose mission, on the contrary, was not usually aimed at the public good but mainly at economic interests.

Scotland, after the unification with England in 1707, was forced to suffer the Statute of Anne and the monopoly of the London booksellers, to which it rebelled through the proliferation of private and pirate reprints of works that were traded in the great fairs of continental European countries, such as France, Spain, Scandinavia, and the Netherlands, favouring the spread of culture internationally. In the 18th century, thanks to the Scottish booksellers, the question of the right to copy took on a different dimension, becoming fully part of the public debate. The work of London booksellers began to appear as a restriction on the progress and advancement of culture, and the free production of reprints was seen as a limitation of the public sphere and of intellectual and creative freedom. The heated debate on piracy, as well as the practice of unauthorised reproduction, also affected other European countries, starting with France, where the Parisian booksellers' guild had assumed a similar role to that of the Stationers[21] . In the Age of Enlightenment, what was being questioned was no longer just the act of piracy, but the entire publishing and authoring system, the forms of censorship, and even the broad and controversial question of the freedom of the intellectual. These issues intersect with each other and, on the whole, are hardly tolerated by many because they are perceived as the will to impose controlling actions by authority on book production. In particular, a broad movement of opinion was directed against the emerging concept of literary ownership of the work, as it linked the idea of literature to that of property, and hence to something marketable, and appeared to many as an unacceptable distortion. In France, the debate was fuelled by Marquis de Condorcet, who argued that the concept of literary property was useless because it conflicted with the public interest of works created to increase collective knowledge. An approach that once again finds an echo in the current action of the open access movement in favour of the free circulation of knowledge, freed from costs and limitations imposed by commercial publishers.

Literary property was indeed abolished in France following the Revolution of 1789, but the experiment did not have the expected results, as it intensified the accusations of piracy and at the same time led to a considerable increase in the number of printers and printed works. In 1830, a magazine with the explicit title Le Pirate was even published

21 Chiara De Vecchis - Paolo Traniello, *La proprietà del pensiero: il diritto d'autore dal Settecento a oggi*, Roma, Carocci, 2012.

for six issues, with a collection of articles already published elsewhere[22]. In Germany, Immanuel Kant was among those who spoke out against pirate publishing, but not because it was detrimental to the author's property, which for the philosopher was "inalienable – it was an inseparable extension of the creative self"[23] , but because of the damage that the circulation of counterfeit works caused to knowledge as a common good. Several authors had contrasting attitudes towards piracy, on the one hand condemning it but on the other hand using it when necessary to print unorthodox works, as Newton did for his religious writings, or Voltaire, well known to booksellers for his tendency to collaborate with abusive publishers behind the backs of his official publisher, reshaping his texts, adding passages, and creating countless pirate editions himself[24] .

Despite the concerns of many authors, in fact, being pirated became a sort of status symbol, as it was generally the works of the best known and most widely read authors that were subject to illicit reproduction. For foreign works, unauthorised reproduction on site was seen as an easier and cheaper route than having to bring in books from abroad, but authors whose fame had crossed national borders, whose works were famous, and, therefore, saleable, were reproduced. The pirated copy began to be seen as a sign of the author's authority and fame, who himself benefited from the free dissemination of his writings, and the printers, of course. There is also no denying how piracy contributed in various ways to the definition of the public sphere and the development of society in those years, to the spread of reading and knowledge. Adrian Johns lists them in his volume: first, it facilitated and increased the circulation of books and magazines especially outside the big cities; second, it had an impact on the type, quality, and price of books, as pirated copies of the most lucrative titles were printed in smaller formats and at much lower prices than the originals, consequently expanding the readership (albeit sometimes by mixing and recombining the contents and presenting them as 'improved' editions); third, it facilitated reading practices since the books were easily transportable and relatively easy to find; fourth, it raised the issues of accuracy and authenticity, since although pirate copies usually boasted a certain accuracy, this was not always the case, and sometimes, on the contrary, there were errors, while at other times, pirate printers tried to 'improve' the original, or rather make it more saleable, by abbreviating it, adding parts, creatively translating[25] . Many of the contributions recognised by Johns can be

22 All digitised issues of 'Le Pirate' are available on the BNF Gallica portal https://gallica.bnf.fr/ark:/12148/cb328400016/date1830.liste.

23 A. Johns, *Piracy, cit.,* p. 55.

24 R. Darnton, *The Case for Books,* cit.

25 A. Johns, *Piracy,* cit., p. 48-49.

observed in the current situation, particularly in relation to open access digital publishing. Open access publishing increases the dissemination of works for all readerships, the price for the reader is zero, ease of reading is guaranteed by digital, and accuracy and authenticity have become progressively more problematic issues.

In the 19th century, too, the question of the authenticity and accuracy of pirated works was a matter of concern for authors and contributed to the public debate. Improvements introduced by pirate printers were sometimes adaptations to the local context, as in the case of the Irish reprint of William Guthrie's *Modern Geography,* in which the part on Ireland was expanded from the original, but this contributed to making it one of the most popular books in the country. At other times, parts were removed or changed either in an attempt to improve them or for contingent reasons, such as the entire original copy being unavailable so that only the parts that were in their possession were reprinted, or different versions of a text were combined according to how the printer was able to obtain them. Ireland had a flourishing business of illicit reprints, as booksellers followed autonomous and rather free practices in the absence of domestic regulations (the Copyright Act was not extended until 1801), justifying them by the need for the dissemination of knowledge that the much cheaper copies allowed. The vibrant Irish free-reprint market then fuelled the equally flourishing development of piracy in the United States, to which an Irish printer-bookmaker, Mathew Carey, made a substantial contribution. Involved in some political issues in his homeland, where he had been charged with high treason for publishing in his newspaper, the Volunteers Journal, an indictment of a member of the government who had ordered the firing into a crowd during a demonstration, Carey fled in 1784 to self-exile in the United States where he exported the practice of illegal reprints. He became one of the most important publishers in the New World, so much so that in 1943 the Publishers Weekly instituted the *Carey-Thomas Award,* in honour of Mathew Carey and the American publisher Isaiah Thomas. In fact, in the colonies the practice of unauthorised reprints had already existed along the import of books from Europe, as there were no copyright laws and the knowledge and work of the various people who had emigrated from England, Scotland, and especially Ireland had been taken advantage of, but the birth of the new nation had also redefined the role of the press. The independence from the British Empire enshrined in the Declaration of 1776 also meant the consolidation of an American publishing business.

Italy was not less involved in the problems and debate on piracy, as evidenced by the many accusations of counterfeiting and "ugly prints"

recorded in Venice since the 16th century[26] . To bring further evidence of the similar situation in Italy, the invective of Melchiorre Gioia is significant. In 1827, he used the English definition of 'book piracy' to heavily condemn the practice of unauthorised reproduction of works by those he defined as printer-robbers. These behaviours, in Gioia's words, are adopted

> only by the printers most unfit in their profession, most despised in trade, usually bankrupt or close to bankruptcy, and only capable of earning a living by fraud, so that in their reprints, often made with the meanness and haste of thieves, they multiply spelling errors, spoil the senses, sometimes detach essential parts of the most esteemed works in order to be able to make a cheaper edition, and sometimes deface them with ridiculous additions in the belief of improving them[27] .

After giving a few examples, Gioia in his booklet accuses the 'printer-robbers' of not allowing authors to correct defects in the first printings of their works, adding that this 'impedes production in the sciences' and 'is a detriment to science', thus introducing a very topical concept regarding the deleterious effect of illicit behaviour on science. Gioia, still on the English front, is echoed by Charles Babbage, who in the preface to his 1830 volume "*Reflections on the* Decline of Science in England", heavily laments the decline of science in England. Babbage, moreover, states the necessity that those who are paid with public funds (*public purse*) must adhere to certain principles and the institutions that pay them (in those days these were companies, societies, and academies), must verify this adherence and have the right to discuss the work of those they finance and the ways in which they perform the tasks for which they are paid[28] . In the incipit of his collection of reflections, Babbage also identifies the clear link between science and education, defining it as a principle even too obvious to dwell on and valid in every nation. He states that it is evident that the state of science, in a translated sense of knowledge, substantially influences the country's education system. The quality of education received at university, in turn, influences what will be the likely future ruling classes in society. A series of cogent and topical considerations, although not always adequately considered.

In the United States, concern about the effects of piracy on science was voiced by, among others, Edward Youmans, an advocate of pub-

26 Carnelos Laura (ed.), *I pirati dei libri. Stampa e contraffazione a Venezia tra Sei e Settecento*, Venezia, Marsilio, 2012.

27 Melchiorre Gioia, *Cenni sulla pirateria libraria* in *Opere minori*, Lugano, Giuseppe Ruggia, 1837, p. 421-422.

28 Charles Babbage, *Reflections on the Decline of Science in England, And on Some of Its Causes,* London, Printed for B. Fellowes and J. Booth, 1830, p. xi.

lic science and the need to create a network of international collaboration between scientists and publishers, a goal for which he travelled to England in 1871 and founded the scientific journal Popular Science Monthly the following year, with the aim of popularising and counteracting the negative effects of piracy on the public. The progressive industrialisation of the printing process during the 19th century increased piracy, shortening the publication time of authorised copies and pirated copies, and shortening the interval between the release of the original and the pirated copy, with the result of making book piracy unprofitable and shifting printers even more towards scientific publishing and magazines. Contributing to this shift were successful new forms of periodical publication, such as *story papers*, which appeared in 1839, initially attached to newspapers but later autonomous, mostly sent to subscribers by mail, and devoted to a wide variety of topics. The *story papers* 'pillaged European periodicals, reprinted old works under new titles, and, at a pinch, stole from each other'[29] , without even bothering much about it, since one of the most famous was explicitly entitled The Corsair. In practice, the market for pirated books by the end of the nineteenth century had reached saturation point, and the business was no longer profitable, so the focus had shifted to periodicals that required less effort in terms of editorial preparation, fewer copyright complications, less time, and costs. However, during the course of the century, the extent of piracy had become clear in many countries, particularly starting with the ambiguous relationship, in terms of copyright, between England and its former North American colonies. From various fronts, and especially from America, the debate on the need for an international copyright law began to ebb, leading to the enactment in 1886 of the Berne Convention, which the United States would only join in 1988. In 1891, however, President George Washington signed the International Copyright Act that extended protection to foreign works in the United States.

Italy was one of the countries in the international network formed by Youmans. The first intellectual property law saw the light of day in 1865, shortly after the unification of the country. The debate that preceded the law shows the conception of two separate spheres, which still form the basis of copyright as conceived in Italian legislation. On the one hand, intellectual property, the inalienable moral right relating to the authorship of the work, and on the other the patrimonial right to reproduce the work in a specific form, which in English speaking countries is properly referred to as copyright, the right to copy and reproduce. Since the reproduction right is exercised not by the individual author but by a commercial entity, typically a publisher, the issue is from the outset declined in patrimonial terms, with emphasis on the

29 A. Johns, *Piracy*, cit., p. 304.

right of ownership. This is set in the context of the development of legal and economic thought that characterises the second half of the century. While moral rights remain in the hands of the author of the original form of thought, what can be transmitted, indeed is mostly made to be transmitted, is the expression of this thought that becomes a product through an industrial activity that allows the circulation of an artefact within which that content can pass from one mind to another and consolidate into a socially broader knowledge'[30] . The result is a separation between the artefact and the content of thought, which has been in place since the early days of copyright law. For magazines, this meant introducing a distinction between the collective work (the magazine) and its component parts (the individual articles).[31] It should be noted that the first intellectual property laws in European countries specifically covered only certain types of publication, i.e. books primarily, and journal articles enjoyed a separate regime until the early 20th century. Many journals, along the lines of newsbooks and newsletters, were initially collections of articles published elsewhere, national and international, and there was a kind of tacit reciprocity agreement between journals and intellectuals, in a common regime of openness from which all benefited[32] . This practice was endorsed by formal agreements between nations, starting with the bilateral treaty between France and England of 1852, which served as a model for similar treaties in other nations in the second half of the nineteenth century. The treaty provided for the journals of the two countries the free reproduction of articles that had appeared in other publishing houses, on condition that the original source was cited and unless otherwise stipulated by the author. This exception remained in force until the Berlin Convention of 1908, which revised the Berne Convention of 1886, but did not include journals in the provisions on the free licence to reproduce, which was provided for newspapers instead.

In Italy, in 1925, the Rocco Law made a clear distinction between patrimonial rights, and moral rights and this distinction was to enter, thanks to Italian action, into the revision of the Berne Convention in 1928. In 1941, the copyright law was enacted, which is still in force today, albeit with several subsequent amendments. The law came into being during the fascist era and with the intention of conforming the legislation to the needs of the regime. In fact, it provided for the establishment of a public register at the Ministry of Popular Culture (later

30 C. De Vecchis, P. Traniello, *La proprietà del pensiero,* cit., p. 105.

31 Roberto Caso, *La rivoluzione incompiuta: la scienza aperta tra diritto d'autore e proprietà intellettuale*, Milano, Ledizioni, 2020, p. 163.

32 Pierre-Carl Langlais, "Quand les articles scientifiques ont-ils cessé d'être des communs?"*, Sciences communes* 11 March 2015, https://scoms.hypotheses.org/409.

managed by the SIAE - Italian Authors' and Publishers' Association), just as it was in the 17th century when the political power wanted to keep publications under control although, unlike then, the ministerial register had purely documentary and evidentiary value, since the registration is authentic until proven otherwise, but does not influence the acquisition of the right[33] . In fact, the registry was the basis of the 1941 Copyright Act and was the discriminating factor for the identification of counterfeit works and therefore in some ways had a greater influence than stated. This law in articles 39-43 regulates collaboration in periodicals, granting the editor of the journal the power to introduce changes in the article deemed necessary in derogation of the 'moral right', and stipulates that, unless otherwise agreed, the author shall transfer to the publisher exclusive rights to the work, excluding any other party from the right of reproduction through the press[34] . The economic rights of collective works are also granted to the publisher. In addition, the law provides for administrative penalties for the offences of usurpation of the authorship of the work and of counterfeiting, which is more understood as reproduction on unregistered media, i.e. not marked SIAE, whereas the term 'plagiarism' does not appear in any passage since the appropriation of parts of another person's work to introduce them into another work under one's own name is defined as 'counterfeiting' or 'usurpation' when the plagiarism is total. Generally speaking, plagiarism is the subject of much debate among legal experts, but currently in most international legislation it is not recognised as a crime, at least not a criminal one, since it does not imply a violation of property rights and therefore cannot be equated with theft. Italian law provides for penalties for the violation of both moral rights and property rights, which differ depending on whether the violation of copyright is a civil, administrative, or criminal offence. In the first case, sanctioned by Articles 158-167, the punishment is withdrawal from the market and destruction of the plagiarised goods with compensation for damages; in the second case (Article 174), there is a sanction and possible suspension of commercial activity from six months to one year; in the case of a criminal offence, sanctioned by Articles 171-171nonies, the punishment may be a fine, suspension of commercial and professional activity, and in the most serious cases and proven fraud, imprisonment[35] .

The sanctions are mainly aimed at the commercial activity, and thus de facto at the publisher. In contemporary society, writes Paolo Traniello, producing and putting into circulation printed works is the primary responsibility of the publisher and the same protection of

33 C. De Vecchis, P. Traniello, *La proprietà del pensiero*, cit., p. 134

34 *Ivi*, p. 134-135.

35 Cristina Mantione, *Violazioni del diritto d'autore: le conseguenze*, "Altalex" 14/05/2021, https://www.altalex.com/guide/violazioni-del-diritto-d-autore.

the author, for example from counterfeiting, primarily concerns the publisher, who has invested capital for the printing of the work and is equipped, when it reaches the structural dimensions appropriate to a publishing enterprise, with the economic and judicial instruments (e.g., legal offices) suitable for protection[36] . He then adds the consideration that this is not a new phenomenon. As we have seen, since the dawn of printing, the power of booksellers and publishers has always been greater than that of any other actor in the publishing world and sometimes even of governing institutions. Today, we see this in an exaggerated manner in scientific publishing dominated by a few large publishing giants that control the market, impose limitations even greater than the legal restrictions, and have weakened the bargaining power of university libraries. This condition now extends beyond publications to include data, projects and various other products throughout the entire research life cycle, even evaluation, based on bibliometric databases produced by the publishers themselves[37] . Universities and libraries are therefore subjected to what Robert Darnton has called the 'heroin model', i.e. "a sales strategy that involves first bringing low-cost products to market, then relentlessly raising prices to the highest level once consumers have been hooked[38] .

The historical path outlined clearly shows how there are two elements at play when it comes to scientific publishing, which have always been closely intertwined and inherent in publishing: copyright and the evolution of printing techniques, which we can now extend to digital. A third factor is more recent and is represented by research evaluation systems, but affects just as significantly as the other two. Consequently, when discussing each of these three topics, one cannot avoid considering the other two. Research evaluation methods based on quantitative indicators play an important role in investigating distortions in scientific communication, but they are certainly not the only cause. Research evaluation is part of a system in which the loudest voice remains that of the publishers. It is no coincidence that at the beginning of the twentieth century in the United States and Great Britain there was an increase in cases of piracy, especially in the field of music. As Adrian Johns argues, there was a widespread sense of resentment towards traditional publishers, perceived as nepotistic, monopolistic, lacking creativity and imagination. Not only among pirate publishers, but even merchants complained about publishers, disapproving of their high pricing policy, their aura of secrecy about the

36 C. De Vecchis, P. Traniello, *La proprietà del pensiero*, cit.

37 See Rossana Morriello, "Lo sviluppo delle collezioni tra bibliometria e nuovi scenari dell'editoria scientifica", Biblioteche oggi Trends, 4 (2018) 2, p. 39-47, http://www.bibliotecheoggi.it/trends/article/view/854.

38 R. Darnton, *The Case for Books,* cit.

reasons for such prices, and the widespread perception was that "there was something seriously awry that the rise of the pirates was merely making manifest"[39] . Considerations that seem all the more appropriate to today's context in which publishing giants are forcing libraries to face a constant increase of the prices of journals and scientific monographs, without any clear reasons. In light of the above, it is clear that the entrepreneurship of publishing in the 1960s and then the introduction of the new digital technological medium have only exacerbated long-standing trends and lines of development, and that predatory and fraudulent publishing, which we will discuss at length later on, is nothing but a consequence of these trends and one of the signs of an ongoing paradigm shift. Just as in the 17th century, we are in a period of great ferment, in which divergent phenomena and many stresses characterise the world of publishing and scientific research.

3. At the origins of the new science

Alongside the numerous upheavals in the publishing world, and partly as a consequence of them, there were also great transformations in the scientific sphere in the 17th century. The impact of the discovery of the New World, the work of Galileo Galilei, and the publication of Isaac Newton's "Philosophiae Naturalis *Principia Mathematica"* in 1687 redefined natural philosophy and laid the foundations for the birth of modern science. In particular, with Galileo, the principles of the scientific method that had been elaborated in the Greek *polis* and the cities of Hellenism were re-appropriated; this reappropriation gave 17th century science the theoretical force that the culture of Humanism and the Renaissance had been unable to draw upon[40] . The new science would be built on the questioning of the credibility of existing knowledge, which had by then moved away from the heights of ancient science and decayed into a prescientific phase, of the predominance of transcendence and the intertwining of secular and religious power with scientific activity. Due to the general loss of trust in institutions, new science will be formed outside the institutions in charge. Traditional learning and therefore the universities, now largely subservient to ecclesiastical power, like many other institutions, had lost their role and authority, so that "self-appointed authorities were now springing up everywhere, generating a dangerous profusion of rival claims leveled at disparate constituencies"[41] . The scientific revolution took place outside

39 A. Johns, *Piracy,* cit., p. 331.

40 Maria Luisa Villa, *Scienza è democrazia. Come funziona il mondo della ricerca*, postfazione di Pietro Greco, Milano, Guerini e Associati, 2018. p. 52.

41 A. Johns, *Piracy, cit.*, p. 23.

the universities and was often condemned by canonical and traditional knowledge, as in the case of Galileo. To accomplish this, new and different institutions, the scientific academies, were needed, modelled on the literary academies of the Renaissance. As during the Hellenistic kingdoms, the new science was no longer based on explanations that traced phenomena back to divine intervention but on the rational study of the phenomena themselves, as observed and experienced by the scientist, described and then discussed with peers. This was the birth of the scientific method and science as a profession, which was supported in its initial development predominantly by the academies, at least throughout the 18th century and until the emergence of the new model of the universities in the 19th century. At the same time, towards the end of the 18th century the professionalisation of science led to its isolation and detachment from the public, in a process that began at that time and continued to the present day. The transformation of scientific activity into a profession was triggered by scientific progress, which led scientists to use increasingly sophisticated instruments and techniques and to conduct research in laboratories equipped with the latest technology and capable of housing large equipment, but in isolation from society[42] . The need to publicly present research and disseminate the results in a way that would make them understandable to all had been a priority for the academies, but it had fallen by the wayside as the state had replaced aristocratic patronage. Making public the research produced in the academies was meant to attract patrons, and it was necessary to communicate it in an open and easily comprehensible form precisely because financiers were more inclined to support research when they could understand it and see its results clearly. Between the end of the 18th and the beginning of the 19th century, universities were transformed on the basis of the Humboldtian model, embracing modern science and combining scientific research with teaching[43] . Those who remained outside the universities often went on to form the scientific societies whose role was essential[44].

At the same time, the debate on the need to adopt internationally valid copyright laws, born at the end of the nineteenth century, broadened to a more general perspective on the transmission and preservation of knowledge and, specifically, constituted the new strand of study

42 The 2021 Nobel Prize winner for Physics, Giorgio Parisi, recalls the large underground laboratories in the physics faculty at Rome's La Sapienza University back in the 1960s and 1970s, as a student and then at the beginning of his career, in his book *In un volo di storni. Le meraviglie dei sistemi complessi,* Milano, Rizzoli, 2021.

43 Wilhelm von Humboldt, *The Internal and External Organisation of the Higher Scientific Institutes in Berlin*, Italian translation by M.C. Pievatolo, 2017, https://archiviomarini.sp.unipi.it/735.

44 M.L. Villa, *Scienza è democrazia*, cit., p. 69-74.

of 'societal science,' i.e., the in-depth study of the ways in which science regulates society. This positivistic approach had among its main proponents Henry C. Carey, son of Mathew Carey, a staunch supporter of the nodal role of intellectual property in social science conceived "as the highest level of a coherent and universal system of knowledge"[45] . Carey was opposed to the idea of international copyright, and indeed to the very idea of a copyright at all, since, according to his thinking, authors, scientific and literary, build their works on the basis of previously accumulated knowledge, and there is therefore no reason to grant them a monopoly on works that represent only a step along the path of knowledge. This is a very modern idea, which we find in the substance behind the current thinking in support of open science. In addition to the continuing confrontation between the supporters and detractors of copyright, in this period, we see the first shoots of a perspective of analysis that has become central to the university today, such as the impact of scientific research on society, referred to as the Third Mission and Public Engagement.

The crisis of trust in institutions with the proliferation of self-styled authorities in the 18th century is, moreover, a description that is perfectly applicable to contemporary society and can be interpreted as a further indication of probable major changes taking place. Adrian Johns in his aforementioned volume dwells extensively on the ways in which in past centuries the press provided an opportunity to gain authority and credibility and how many authors, especially those with financial resources, shrewdly exploited this opportunity to gain prestige. Various cultural and political movements owed their success to the dissemination of printed material, which is not uncommon. A practice we see frequently today with digital and the Internet, where it is possible for many to have a voice, and to acquire and boast authority. The difference, apparently, is that today anyone can access the Internet, but on closer inspection, not unlike in the past, only those who are particularly skillful and also endowed with economic resources are able to turn the medium to their favour and gain a following and credibility, for better or worse. Trust in institutions is now reduced to a minimum, intermediate bodies have lost their role, digital publishing pirates, and fake news circulates freely. The similarities with that era of great transformation continue to appear significant.

One of the main topics of discussion regarding copyright was the concept of ownership of literary works, which in the 17th century did not exist as we know it today. A diatribe between London booksellers and Scottish booksellers led by Alexander Donaldson, a bookseller in Edinburgh, saw the former arguing in favour of a declaration of ownership of the literary work and the latter opposing the concept of owner-

45 A. Johns, *Piracy, cit.*, p. 312.

ship that bound works to booksellers. Among the arguments brought by Donaldson and the opponents was the idea of a similarity between literary invention and craft invention. The patent for technical inventions was regulated by the Monopolies Act of 1624 which granted a privilege by the king for a limited duration. It was not a property right, but a temporary concession at the end of which the invention reverted to the availability of other craftsmen. The underlying consideration was that each version of an object made by different craftsmen would be different due to the different skills, materials, and techniques used. However, this did not avoid disputes and cases of 'technological espionage' for misappropriation of patent ideas[46] . Based on this principle of similarity, according to Donaldson's view, Scottish booksellers could have continued to reprint. The opposing faction, on the other hand, argued that this principle could not be applied to books because the form that an author's intellectual work takes is that defined by the author himself, determined by his style, and copies of books, although different in form, are equal in style and intellectual content. Printing is only a means of transmitting that style and content, and unlike the handcrafted replication of objects that brings improvement and progress, reprinting has in itself no function of cultural advancement. The notion that the identity of literary work is not identified with its materiality, as asserted in those years, still remains the basis of the copyright debate by those who demand a revision of copyright in light of the spread of open science and the new dynamics of scientific research. In general, as Roberto Caso argues, there has been a shift in the scope of copyright from container to content. This has gone hand in hand with the commercialisation of science and the strengthening of the position of large commercial publishers, who are able to impose their conditions in acquisition contracts signed by libraries, and also vis-à-vis authors. The research evaluation systems that have produced the publish-or-perish culture force choices. Caso writes that the scientific author does not gain from the management of economic copyright but is instead interested in the acquisition of a good scientific reputation and a young researcher who finally sees the possibility of publication in a renowned scientific journal or with a prestigious publishing house materialise will not even attempt to negotiate a different copyright arrangement and will accept the clauses determining full and exclusive assignment to the publisher[47] .

46 As in the case of silk weaving, which passed from Bologna to Turin and from there to England where it was patented by John Lombe in 1718, see Marco Erriquez, Vittorio Marchis, "Lo spionaggio tecnologico nell'Italia del Seicento e del Settecento. Il caso delle macchine da seta", *Gnosis 2* (2018), p. 169-179.

47 R. Caso, *La rivoluzione incompiuta,* cit., p. 148.

The cession of copyright to publishers, and thus of the possibility of reproduction in any version and format, hinders the author's scientific freedom and affects educational freedom. Copyright laws thus understood stand as an obstacle to the free circulation of the author's ideas, giving publishers a right very similar to the right of ownership over printed works claimed by 17th and 18th century printers. In particular, copyright law in Italy, Caso writes, is focused on the expressive form of the intellectual work (the scientific text). Ideas, facts, and data - according to the traditional principle of the distinction between protected form and unprotected idea - remain in the public domain. The norms of science focus, by contrast, on the content of the theory[48] .

The focus of science on content and the attempt to break free from the constraints of publishing and copyright emerged when journals, one of the main means of disseminating the results of scientific research, began to be bound by copyright law, starting with the Berlin Convention of 1908. However, the need for a freer and easier way of circulating scientific information and research results than traditional journals allowed manifested itself not long after, once again pitting scientists against publishers. Since the 1920s at the American Chemical Society and since the 1940s at MIT (Massachusetts Institute of Technology), institutional repositories have been set up to hold preprints of articles submitted to journals, which were exchanged and commented on within the institution. In 1961, the National Institute of Health (NIH) in the US launched the Information Exchange Groups (IEG) programme to facilitate the circulation of biology preprints. The initiative involved 3,663 participants from 46 countries and generated 2,561 papers in the few years it remained active. The project was closed in 1967 because journals had begun to refuse to publish the articles that were already in circulation in preprint form, and publishers lobbied for the closure of the IEG as they feared a weakening of their role and consequently a reduction in economic revenues[49] . The success of the initiative also caused a stir in other disciplinary fields and other institutions, some of which had already been collecting their own preprints for some time, such as the library of the Stanford Linear Accelerator Centre (SLAC) and the library of CERN (Conseil Européen pour la Recherche Nucléaire) in the field of physics. Subsequently, the Physics Information Exchange (PIE) project was born, similar to IEG but with the difference that preprint copies were only sent to member libraries and not to individual applicants. The rise of IEG and the emergence of the other projects further alarmed the publishers. The first to

48 *Ibid.*

49 Matthew Cobb, "The prehistory of biology preprints: A forgotten experiment from the 1960s", *PLOS Biology*, 16 November 2017, https://doi.org/10.1371/journal.pbio.2003995.

come out against IEG was the American Association of Immunologists (AAI), which publishes *The Journal of Immunology, which accused the* National Institutes of Health *project of* jeopardising the publishing system, of constituting a threat to the integrity of science, and of being an inappropriate activity for a government institution. This was followed by the stance, through articles and editorials, of two prestigious journals such as Nature and Science. However, IEG's fate was sealed not by these large publishers but by a group of small publishers from the biochemical field, who met in Vienna in September 1966 and took the decision not to accept for publication articles that had previously been circulated as preprints. The closure of IEG in March 1967 also led to the decline of the PIE project, which in the meantime had reached a trial version, despite the fact that researchers had spoken out in favour of the use of such preprint archives, largely emphasising their value for research. The outcry generated by these positions led Franz J. Ingelfinger, editor-in-chief of The *New England Journal of Medicine* (NEJM), to declare in 1969 that he would not host articles that had already been published previously in other journals or whose results had been anticipated through other media, thus creating what would later be known as the 'Ingelfinger rule,' which was followed by most scientific journals. Other institutional archives, e.g. at Stanford University or in Europe at CERN, continued their activities and were the inspiration for the birth in 1991 of arXiv, which represented a turning point not only for physics, the discipline in which it was born, but for the general function and development of preprints. It was the physicist Paul Ginsparg at Los Alamos National Laboratory who created the first automated server for the dissemination of preprints[50] and effectively started the open-access movement. Again, as in previous centuries, the events described above show how the tension towards openness and sharing inherent in scientific research is contained and bounded by the economic interests of publishers.

It is not the purpose of this paper to delve further into the dynamics of scientific communication and the relationship between science and publishing, but the events mentioned are essential to establish the pillars on which the practices we will discuss in the following paragraphs are based. Dealing with complex and evolving phenomena, such as those we are observing, requires a holistic vision. Above all, one cannot ignore the historical context that shows how the idea of science as a common good, essential to the progress of society, was intrinsic to the new science, methods, and practices conducted by the Royal Society and other scientific academies. However, science needed to publish (in the etymological meaning of 'to make public') its results and thus re-

50 For the history of arXiv see Paola Castellucci, *Carte del nuovo mondo: banche dati e Open Access*, Bologna, Il Mulino, 2017.

late to the world of publishing, whose objectives were different. In particular, from the second half of the nineteenth century onwards, one can observe new dynamics brought about by industrialisation, which transformed the publishing process from an intellectual activity conducted by individuals to an industrial activity[51] . After the second half of the following century, a further turning point in scientific publishing occurred, with the shift from small science to big science theorised by Derek de Solla Price[52] . Research work becomes organised in large, even international networks (especially for the hard sciences), and there is an increase in the number of researchers and the number of scientific journals, with many publishers seeing a major development on an entrepreneurial basis in these years. A further crucial phase is the advent of digitalisation and the transformation of publishing in this sense. The tension that has emerged prominently since the 1980s along an axis leading towards open science thus lies in a line of continuity in the history of science and publishing and represents a path back to the original mission of scientific research. The new science and scientific publishing were born in the 17th century in a context that was in many ways similar to the present one, naturally with due proportions. As in that century, we are in a phase of great change, probably directed towards a redefinition of the canon, a revision of copyright, and an overall transformation of the publishing system. Every new means of communication and dissemination of knowledge causes metamorphosis. The printing press was fundamental to the Enlightenment debate on cultural and scientific freedom, which has never ceased to the present day and has been amplified by the emergence of digital technologies. Digital is only the latest in a series of impactful 20th century innovations such as radio, television, audio, and video recordings. The influence of some inventions on society in general, on piracy, and on the intellectual freedom debate in particular, has been extremely significant, as in the case of radio[53] .

Broadcasting was an innovation with at least as radical an effect as decades later the Internet and digital broadcasting, when compared to the time when it naturally occurred, the 1920s. In Italy, the first radio broadcast was transmitted on 6 October 1924 and also fuelled a lively patent debate. While Guglielmo Marconi's invention, patented at the end of the previous century in London and initially used for military purposes, was spreading to homes, with the timing of official chan-

51 The dynamics, which also differentiate the fate of publishing in various countries, as traced back in Nicola Tranfaglia, *Editori italiani ieri e oggi*, Bari, Laterza, 2001.

52 Derek de Solla Price, *Little Science Big Science*, New York, Columbia University Press, 1963.

53 A. Johns, *Piracy, cit.,* p. 357-399.

nels and the restrictions on licences imposed by the various ministries and bodies that took control of them in the various countries, a large number of amateur radios were being born in parallel. Radio piracy became a phenomenon of equal importance to book piracy during the Enlightenment and likewise generated an access debate on patents, copyright, intellectual freedom, commercial monopolies, and the status of knowledge as a common good. The spread of various new media and the attempt to impose commercial or political constraints on the results of scientific research brought the debate on science and copyright, on the relationship between science and politics, and on intellectual freedom back to centre stage. Just as it is happening again today with the spread of the digital medium, with the open access movement and with the large amount of thinking about changing copyright, about the need to counter the monopolies of commercial publishers with open access, and to set up sustainable research evaluation methods.

Each new medium provoked changes, redefined concepts and practices, and fueled claims for rights and freedoms by citizens, intellectuals, and scientists, and not only in the field of scientific research. Whether it was the printed book, the radio, the telephone (with phreaking, as telephone piracy flourished in the 1960s was called), or, in the following decade, computers with the debate on the intellectual property of computer code that would initiate the open source and free software movement, all the media produced a division between supporters of the concept of ownership and the need for copyright and supporters of free dissemination and free use, who when faced with laws and practices that restricted it sometimes found an outlet in piracy. Promoters of free access and commercial operators (publishers, computer companies, record labels, etc.) are often on opposite sides, but there are also different visions within the scientific community and public opinion. However, none of the commercial operators has been able to restrain the drive towards open source or file sharing (Napster has been closed down but resurrected in different and more sophisticated forms such as BitTorrent). This will, of course, also be the case for open access publishing, a movement that can hardly be stopped and toward which commercial publishers are in fact moving. The invention of the World Wide Web has the same profound implications as the printing press because it is a technological, but also a cultural revolution, redefining models and paradigms. But what above all unites the two inventions is that they were not patented. In Gutenberg's time, patents did not exist, and therefore his invention was necessarily in the public domain. Tim Berners-Lee, on the other hand, consciously decided not to patent the invention of the World Wide Web. Both inventions, left free and open, changed the world.

4. Scientific research as a common good: the case of shadow libraries

The aspirations for the free circulation of knowledge, rooted in the 17th and especially 18th centuries, and revolving around the intellectual property debate, will be a constant to this day. Book piracy devoted itself to reprinting cheap editions, cheaper than the originals, and the consequence was to make culture and science accessible to more people. The success of pirate publishers and the difficulty of other publishers to counteract their activities also induced the licenced publishers to eventually create cheap editions. The new copyright laws and international conventions signed at the turn of the 19th and 20th centuries did not reduce piracy, as we saw vibrant in the first decades of the 20th century. The gradual increase in the income of the lower and middle classes and general economic prosperity created a new and large audience for publishing products. It is not accidental that all the most famous paperback series that have survived to this day[54] and cheap paperbacks - pioneered by the Penguin Books in England (1935) and Robert de Graaf's Pocket Books in America (1939)[55] - were born in those years, with an increasing number of publishers, especially from 1945 onwards, publishing paperbacks exclusively or alongside bound editions.

Therefore, historically, it is precisely the difficulty of accessing knowledge that has moved people to action and generated great transformations and disruptive phenomena. In the digital world, this consideration remains valid because, actually, this is not a rupture, but only a change brought about by a new technological medium, as there have been so many throughout history, although the latter has profound implications for its 're-ontoligising and re-expistemologising' power[56] . The birth of the Internet and then of the World Wide Web, moreover, fit by their very nature, even before their effects, into the widely delineated line of sharing and universality of science. The Arpanet network was born in 1969 to enable the exchange of files between the computers of military experts and researchers in universities, and the World Wide Web was an implementation with the same purpose: to facilitate the exchange of documentation, and thus communication and collaboration between scientists. As is well known, in this spirit of openness and sharing, Tim Berners-Lee and Robert Cailliau decided not to patent the invention and to make the WWW free, releasing the source code into the public domain.

---

54 Siegfried H. Steinberg, *Cinque secoli di stampa*, Torino, Einaudi, 1982, p. 302 (orig. Five Hundred Years of Printing, 1951).

55 Ivi, p. 305.

56 Luciano Floridi, *Il verde e il blu. Idee ingenue per migliorare la politica*, Milano, Raffaello Cortina, 2020, p. 33.

However, the emergence of the Internet and of resources in digital format made it clear that, while the possibilities of disseminating information and knowledge were amplified and became easily and freely accessible, even greater barriers were created, as the power of a few private operators increased, and numerous inequalities were generated. In science, the consolidation of the digital format magnified the power of publishers. Pirates were not slow to circumvent these barriers, in the digital world even more so than before, since the typical digital intermediation concerns the creation of content as well as its violation. At the end of the 1990s, various file sharing systems sprang up for all kinds of resource, such as music, films, bibliographic material. A well-known case is that of Napster, the music file-sharing network set up in 1999 and closed in 2001 following a US Federal Court ruling for copyright infringement, after many complaints from record companies, music producers and musicians. But just as Napster closed, BitTorrent opened, and countless platforms based on the same peer-to-peer principle as Napster proliferated, but in decentralised mode, many of which are still active today. A similar but less-publicised fate, at least until recently, concerned the first file-sharing networks for books and journal articles that originated in Russia, such as Text.org, Monoskop, and Gigapedia. These are the progenitors of what are now called shadow libraries. In shadow libraries you can download articles and books for free that would normally be behind a paywall, i.e. for a fee on publishers' sites. Much of the material included in shadow libraries is deposited there in an unauthorised manner and in violation of copyright laws, but above all, of the rules imposed by publishers in their licencing agreements for subscribing to digital resources. These are sharing networks whose fate has been no different from that of other media: closed and reborn in even more sophisticated and consolidated forms. The ubiquitousness and fluidity of digital transformation transforms its creatures into phoenixes reborn from their own ashes. Just as had happened to Napster, when Gigapedia (also known as Library.nu) became too big a network, it began to worry publishers. In 2010, Gigapedia was shut down, following legal action by seventeen publishers led by Wiley[57] . The following year, Gigapedia's archive was incorporated into LibGen (Library Genesis), the shadow library established in 2008, also in Russia, thus strengthening its role. The definition '*shadow libraries*' was coined by Joe Karaganis of Columbia University in his 2011 book *Media Piracy in Emerging Economies* and is used together with another

57 Balázs Bodó, "The Genesis of Library Genesis: The Birth of a Global Scholarly Shadow Library", in *Shadow Libraries: Access to Knowledge in Global Higher Education*, edited by Joe Karaganis, Cambridge (Mass.), MIT Press, 2018, p. 27.

definition, 'biblioleaks', coined in 2014,[58] with clear reference to the well-known WikiLeaks affair, to define this kind of networks. Both definitions highlight the dark side of these resource repositories that operate in the shadows and on the edge of legality, but with the intention of making culture and science available to all.

The main shadow libraries emerged as a reaction to the difficulty of accessing a free culture not controlled by political power. In the former Soviet Union, the rate of schooling and the number of readers was always very high because for the communist regime, culture was a goal to be pursued and reading was almost as sacred as religion[59] . But culture was sacred because it was controllable by the regime and frequently subject to censorship. Literary works were hard to find, and authors such as Dostoevsky and Nabokov were banned. In the scientific field, it was difficult both to obtain texts in the original language (for those who could read them in different languages) and to obtain them in Russian translation. Books were the subject of a flourishing black market, which in the 1970s-1980s accounted for more than two thirds of book buyers[60] . In fact, a large underground publishing market had developed since the 1960s. Just as had happened in previous centuries with book piracy, an alternative but free 'market' had emerged in the Soviet Union that would expand with the spread of the Internet and digital media, moving from the simple exchange of already printed material to the production of digital works. Since the middle of the twentieth century, an undergrowth of illicit activities has formed with the aim of procuring literary and scientific works, especially foreign ones. Where access to culture is prevented for whatever reason, alternative routes are often taken to procure material and to circumvent the obstacles raised, in this case by ruling regimes. The birth of the Web and the spread of computers in homes facilitated the circulation of works and reproductions through unofficial channels, and the previously used photocopies were replaced by much easier digital formats.

The unofficial exchange of bibliographic material often started with those who had access to books and journals and made them available to all who needed them. Lecturers and students who had access to resources through university libraries were among those who contributed significantly to the networks. Entire collections of bibliographic resources, both scientific and fiction, were created, which with the arrival of digital could be exchanged via CD-ROMs and later via the Russian Internet network RuNet. These collections were the first nuclei from

58 In the article Adam G Dunn, Enrico Coiera, Kenneth D Mandl, "Is Biblioleaks Inevitable?", *Journal of Medical Internet Research* 16 (2014) 4, https://www.jmir.org/2014/4/e112/.

59 B. Bodó, "The Genesis of Library Genesis", cit. p. 29.

60 *Ivi,* p. 32.

which networks such as Gigapedia later emerged. Initially, the resources in Gigapedia were neither sorted nor categorised in any way, and searching in these huge databases was difficult until a librarian began to organise them by creating the *Library Genesis (or LibGen)* database. LibGen is a completely open and sharable site, where one can download the contents, the entire catalogue, and even the server code. The only condition for sharing and reproduction is that everything remains free and open to users. LibGen today includes millions of records with very high growth rates, as reported in the site's statistics. Data for 2018 indicate that more than 55,000 publishers are included, including the largest commercial publishers[61].

From the analysis of LibGen data, it can be seen that although large publishers account for the highest percentages in terms of supply (with Springer, CUP, Routledge, Wiley, OUP in the top five places), they do not account for the highest numbers in terms of downloads, which instead relate to small publishers and scholarly societies. Equally significant is the prevalence in the site archive of the social sciences category, both in terms of document volume (15%) and document demand, followed by technology and engineering (14.5%). The most represented Dewey classes are medicine and health, computer science, information science, general works, American literature in English, economics, mathematics, engineering and applied operations, social sciences, sociology and anthropology, management and public relations, social problems and services, English and Old English literature. If one looks at the geographical origin of LibGen users, one finds Russia in first place, as is obvious since much material is in Russian (along with German and English, while other languages are less represented), followed by countries such as Indonesia, India, Iran, Egypt, and China, i.e. countries with less access to published scientific literature for economic and political reasons. However, there are also countries such as the United States, Great Britain, and Italy, showing how current paywalls to digital resources represent an obstacle in all countries, albeit to different degrees. For less wealthy countries, shadow libraries provide a route to access knowledge that would otherwise not be possible and, therefore, are sometimes the only way forward.

The other shadow library that is now well known, especially in the scientific sphere, is Sci-Hub, which was created in 2011 on the initiative of Aleksandra Elbakyan. Then, struggling with her master's thesis in neuroscience and looking for bibliographic material to process it, she had the idea of transforming her personal sharing networks into a public on-line network. The first version of the site was set up in such a way that when a user typed in a request for an article, he was redirected to

61 The data are reported in Balázs Bodó, "Library Genesis in Numbers", in *Shadow Libraries:,* cit.

LibGen, and if the article could not be found on the other site, a search was automatically launched on the personal databases of colleagues (researchers, students) who had made them available on the sharing networks or directly on the publishers' sites using the colleagues' credentials to access them. Each time the article was retrieved, the system generated a copy of it to be stored in Sci-Hub, so that it would remain available for later retrieval, thus building a large database. In 2016, Sci-Hub had reached 50 million articles and in six months, between 2015 and 2016, more than 28 million downloads[62] . In 2015, Elsevier took Sci-Hub to court for copyright infringement, but neither Aleksandra nor her lawyers appeared, partly for fear of having to disclose the location of the server. Sci-Hub was convicted, and the closure of the scihub.org domain was ordered. However, this did not prevent the content from being moved to another domain or the use of P2P systems such as VPN or TOR to directly retrieve the material. Moreover, the affair paradoxically made Sci-Hub famous, increasing the number of users and turning the creator into a sort of champion of free access to knowledge. This has changed the game, altering the nature of the phenomenon. Previously, LibGen and Sci-Hub had tried to stay out of the spotlight, for obvious reasons, but after the ruling this was no longer possible. The interviews and awards given to the young woman became numerous, the site was strengthened, a donation campaign, and experimented with artificial intelligence techniques, periodically facing publishers' complaints.

Trying to stop such phenomena has the characteristics of a fight against windmills. There are now several such sites, and the ways to find 'pirated' bibliographic material are many. Social networks such as Facebook, Twitter (now X), and Reddit offer alternative access routes to scientific research products published through dedicated pages or channels. Reddit Scholar, for instance, is a subreddit that allows users to request and share articles available in various databases and has more than 80,000 members. It is a moderated channel and has very precise requirements for sharing files, which must have a title and author, a tag indicating the type (article, book, chapter, thesis, etc.), an identifier (DOI, PMID, ISBN), and an indication of whether the article is behind a paywall or not. Among other rules, one must not request material that is easily retrievable elsewhere, and one is invited to search first on sites such as Sci-Hub or LibGen or DOAJ, PLOS ONE, Google Scholar, or even in libraries. Furthermore, digital piracy is forbidden, which it is meant that material received via Reddit may only be used for personal and noncommercial purposes, in compliance with copyright. It is difficult to say how much this actually happens. In any case, on Reddit Scholar anyone can post material or a request for an article or book that is then

62 Joe Karaganis, "Introduction: Access from Above, Access from Below", in *Shadow Libraries,* cit. p. 2.

transmitted via P2P file sharing systems. Another way to obtain digital copies of articles or books via the social network Twitter is the use of hashtags such as #icanhazpdf, which derives from the meme "I Can Has Cheezburger?" to mean "I am looking for an article". This hashtag was created by Andrea Kuszewski, a cognitive scientist, in a tweet in January 2011. Simply use it by indicating which article or book you are looking for answers and certainly also the material you are looking for.

The phenomenon of shadow libraries, as well as the other avenues of alternative access to knowledge offered by the Internet, is the heir to book piracy. Tracing the diachronic axis, what we see reemerging in different shapes, sizes, and configurations is always the result of the never-ending tension between intellectual freedom and the desire for open sharing of science, and the economic interests that tend to limit it. The relationship of publishers with sites such as Sci-Hub is ambiguous and fluctuating, but the motivations are clear. Attempts to shut it down have been unsuccessful, and, not unlike with book piracy, the more attempts to do so, the more the multiple avenues of piracy open up. However, studies show that articles downloaded from Sci-Hub receive more citations[63] , as a consequence of being somewhat openly accessible, and thus Sci-Hub also contributes to the citation system that benefits publishers in other ways. Many of the users of Sci-Hub and LibGen are independent researchers or do not have an institution behind them that can subscribe to journals or purchase individual articles, the prices of which are often high. Single illicit access to an article is not too problematic, where publishers have an established market position and stable buyers in university library systems. However, more complex is the situation in potential but not yet conquered markets, such as India, where the three largest publishers, Elsevier, Wiley and the American Chemical Society, recently filed a copyright infringement lawsuit against Sci-Hub and Lib-Gen. The consequence was the closure of access to the sites by the Indian government, followed by protests from scientific societies and researchers[64] . India is a country where the number of scientific articles is growing, but where the penetration of publishers with their expensive digital journal package acquisition models is not as significant and dominant as in western countries. In contrast to countries where publishers already have an established role

63 Juan C. Correa, Henry Laverde-Rojas, Julian Tejada, Fernando Marmolejo-Ramos, "The Sci-hub Effect on Papers' Citations", *Scientometrics*, 127 (2022), p. 99-126, https://doi.org/10.1007/s11192-020-03806-w.

64 D. Chandrasekharam, "Sci-hub and Alexandra Elbakyan", *The Times of India*, June 4, 2021, https://timesofindia.indiatimes.com/blogs/dornadula-c/sci-hub-and-alexandra-elbakyan/; Richa Banka, "SCI hub, Libgen case: Delhi HC to hear students, researchers", *Hindustan Times*, Jan 7 2021, https://www.hindustantimes.com/india-news/sci-hub-libgen-case-delhi-hc-to-hear-students-researchers/story-PxBftjgAPFPfuXGok4qG7N.html.

and licencing contracts for digital journals and books, in countries where this is not the case, access to sites such as Sci-Hub and Lib-Gen could represent serious competition in a still partly free market.

Shadow libraries operate in violation of copyright laws and especially of the power of publishers over scientific publications, which, as we have seen, is an age-old issue. The author normally cedes all reproduction rights of the article to the publisher, depriving himself of them. An author who wishes to disseminate his article through a channel other than editorial publication, perhaps open access, cannot always do so. Going back to the debate of the past, it is no longer just a question of the editorial form of the article, but of the content, the idea that is given to the publisher and cannot be reproduced by the author. This violates the basic principles of science, the idea of communism described by Robert Merton[65] and avowedly adopted by Aleksandra Elbakyan. On the other hand, the scientific journal is one of the essential structures on which science rests and its function is fundamental. The central problem is the actions of large oligopolistic publishers that undermine this structure, imposing unsustainable costs and conditions on libraries and authors[66] .

65 Robert K. Merton, *Social Theory and Social Structure,* 3. ed., New York, Simon and Schuster, 1968.

66 On this aspects see Rossana Morriello, *Le raccolte bibliotecarie digitali nella società dei dati*, Milano, Editrice Bibliografica, 2020.

Chapter 2
The Ethics of Scholarly Communication: A Historical Perspective

1. Error and scientific misconduct

Scientific error is part of the very nature of science, which is based on attempts that sometimes turn out to be wrong and are overtaken by subsequent research, which is precisely why they allow science to advance. The greatest names in the history of science, from Galileo to Mendel, have made mistakes. The latter was even accused by Aylmer Fisher of falsifying data to conform to his theory, because "they were too good to be true"[1] . Karl Popper describes the scientific method itself, in all disciplines, as characterised by trial and error. The scientist who conducts an experiment or elaborates a theory presents the results and the method used in a codified form so that other scientists can verify them and detect errors, which are inevitable because the scientist, like everyone else, is conditioned by a system of prejudices and alone cannot achieve scientific objectivity. In fact, objectivity is closely linked to *the social aspect of the scientific method* and can only be achieved through peer-to-peer comparison, which takes place in the appropriate social institutions such as periodicals, congresses, and laboratories[2] . The Austrian philosopher writes:

> Scientific results are 'relative' (assuming that this term can be used) only insofar as they are the results of a certain stage of scientific development and are destined to be superseded in the course of scientific progress. But

1 More likely it was bias due to the knowledge Mendel had at his time, but the controversy is still rather debated, see Gregory Radick, "Beyond the "Mendel-Fisher Controversy", *Science* 350 (2015) 6257, p. 159-160, https://www.science.org/doi/full/10.1126/science.aab3846; Thomas F. Lüscher, "The codex of science: honesty, precision, and truth-and its violations", *European Heart Journal*, 34 (2013) 14, p. 1018-1023, https://doi.org/10.1093/eurheartj/eht063.

2 Karl Popper, *La società aperta e i suoi nemici*, volume unico, Roma, Armando Editore, 2018. All passages here and infra are translated from the book in Italian *(orig. The Open Society and Its Enemies, 1945).*

> this does not mean that the truth is 'relative'. If a statement is true, it is true forever. It only means that most scientific results have the character of hypotheses, i.e. statements for which demonstration is not conclusive and which are therefore subject to revision at any time[3] .

The scientific error described by Popper is part of the scientific process and is the discriminator that can prove or disprove the scientific validity of a methodologically logical path. As Francis Bacon stated even earlier, error can more easily bring out the truth, which is the ultimate goal of scientific research[4] . Popper again writes:

> There is only one way to ensure the validity of a chain of logical reasonings; and that is to put these reasonings in the form in which it is easiest to check them: we break it up into many small steps, each easy to inspect by anyone who has mastered the technique, mathematical or logical, of transforming utterances. If, after this, any person still raises doubts, all we can do is ask that person to point out an error in the steps of the proof, or to think again. In the case of empirical sciences, the situation is exactly the same[5] .

An error can be made because one starts from prejudices in which one is immersed without realising it, or because one does not have sufficient data to grasp the extent of a phenomenon that subsequent knowledge will be able to overcome and improve upon. In cases like these, the error is not intentional, it is made in good faith and with the intention of advancing knowledge. Therefore, the fundamental starting point is that not all scientific errors are the result of *scientific misconduct*.

However, one must recognise the difficulty, in some cases, of clearly defining what is meant by scientific misconduct. First, different disciplines have different practices, and if certain behaviour is not acceptable in one discipline, it may be acceptable in another. In addition, moral, ethical, and legal conventions change in different eras and what is not tolerable today may have been tolerable in the past and vice versa. Finally, laws and procedures vary between nations, and behaviour that is condemnable in one nation is not necessarily so in another.

Marcel LaFollette has proposed a convincing and effective distinction between types of misconduct, who distinguishes between *illegal acts*, prohibited by governmental, national and local laws, or national and local institutional regulations; *professional misconduct* that violates scientific standards, publication policies defined for authors, reviewers,

3 *Ibid,* p. 503. Translated from the book in Italian.

4 T. S. Kuhn, *The Structure of Scientific Revolutions*, cit..

5 Karl Popper, *Logica della scoperta scientifica. Il carattere autocorrettivo della scienza*, Torino, Einaudi, 2010, p. 93. All passages here and infra are translated from the book in Italian. Orig. *Logik der Forschung*, 1934).

and journal editors, or general rules governing professional conduct such as codes of associations or institutional regulations; and *immoral conduct* that violates the moral and ethical standards of society[6] . The discriminating factor is the intentionality or otherwise of the action. Actions that we define as misconduct are undertaken with the intent to deceive, are guided by unethically unacceptable motives, and must be distinguished from actions that produce deception unintentionally (bona fide errors) or intentionally but for ethically acceptable reasons, such as a scientific or literary hoax or satire. To give some examples of the second case, we recall Edgar Allan Poe's article known as "*The Balloon Hoax"*. In 1844, Poe published an article in the New York newspaper The Sun in which he recounted how Monck Mason, musician, writer, and balloon expert, together with some companions, had crossed the Atlantic Ocean on a balloon in just three days. In the article, Poe reported on the logbook account and the technical details of the crossing. The writer made the account so plausible that it triggered enthusiastic reactions and sold many copies of the newspaper. The article was followed by a note that revealed the ironic intent of the writing. Poe was probably inspired by a similar episode from nine years earlier, which in turn drew inspiration from a satirical piece of writing by Poe, "*The Unparalleled Adventure of One Hans Pfaall"*, which appeared in a magazine but went almost unnoticed.

The New York newspaper The Sun itself in 1835 was the venue for the publication of the "*Moon Hoax"* by Richard Adams Locke, a descendant of John Locke and a Cambridge graduate. With the intention of increasing the sales of his newspaper, the writer and journalist reported in a series of articles on the exploration of the lunar surface by the real British astronomer John Herschel, giving details of the moon's colourful topography and the presence of strange animals such as blue unicorns and even a tribe of flying men. However, fanciful, these descriptions seemed plausible in the nineteenth-century scientific context, in which countless innovations sparked the imagination, were translated and circulated in other countries, so much so that, for example, Locke's articles were collected in a small volume published in Naples the following year under the title *Delle scoperte fatte nella luna del dottor Giovanni Herschel (Figure 1)*.

6 Marcel C. LaFollette, *Stealing Into Print. Fraud, Plagiarism and Misconduct in Scientific Publishing,* Berkeley and Los Angeles, University of California Press, 1992.


DELLE SCOPERTE

FATTE NELLA LUNA

DAL DOTTOR

GIOVANNI HERSCHEL

Traduzione dal francese sulla 10ª edizione fatta in Parigi
NEL MARZO 1836, DAL SIG. F. P.
con una figura rappresentante gli abitatori di esse, e
disegnata al Capo di Buona Speranza dal sig. Kelk.

NAPOLI 1836 (APRILE)
PRESSO GAETANO NOBILE
Vico Concezione a Toledo n. 3.

Figure 1. *Frontispiece (Source Wikisource, CC BY-SA 3.0)*

The history of science and the history of literature are studded with similar episodes devised not always with malicious intent, but sometimes out of simple mockery and at other times to gain prestige and public visibility. The two fields, scientific and literary, are closely linked in a relationship of mutual influence that will remain indissoluble and lead to the creation of a true literary genre that brings them together, science fiction. In particular, the 19th century is characterised by major advances in science and industrialisation processes, which generate a powerful scientific collective imagination on which literature feeds. Mary Shelley, the author of the cornerstone of the genre *Frankenstein* (in which, moreover, she appeals to the scientific veracity of the story), was inspired by a hoax that occupied the pages of newspapers and magazines in the summer of 1826, known as "The Roger Dodsworth Hoax", for her short story *Roger Dodsworth: The Reanimated Englishman*, written that same year but published only 37 years later. The hoax was about the discovery of the body of a hibernated man, Mr Dodsworth, who was buried under an avalanche in 1660, found on the St Gotthard in the Alps after 166 years and brought back to life (with rather rudimentary techniques) and then became part of London social life[7] . One can easily identify as a scientific source, with its impact on the imagination,

7 Charles Robinson reports on the debate of the time, reproducing some pages of newspapers in which comments and inferences abounded as to Dodsworth's actual origin and birthplace, as well as on his affiliation to the Whigs or the Tories, a question

the first experiments to create ice in the laboratory dating back to the mid-19th century and the subsequent attempts to freeze spermatozoa at the end of the same century, which initiated the cryogenesis studies that developed from the mid-19th century onwards.

In 1783, the surgeon J. N. Foersch published in The London Magazine one of the most famous and impressive hoaxes in history. In the article, Foersch spoke of the 'upas' tree that grew on the island of Java and produced a poison that could kill every living thing within twelve miles. Foersch claimed to have seen it himself during his service on the island and was also able to collect other direct testimony. According to his account, poison was administered to convicted criminals who within five minutes began to tremble and then died. The way the hoax was constructed is accurate and denotes a profound knowledge of scientific language and the mechanisms of validation of discoveries by Academies such as the Royal Society, as well as the ability to anticipate every possible doubt in order to convince the public of the authenticity of his account. In reality, 'upas' is a Javanese word that generically indicates a poison of plant origin, and the false content of Foersch's article was based on a layered set of scientific and literary sources that over time had seen overlapping acts of plagiarism, errors in transcriptions due to cultural and linguistic misunderstandings, and ideological influences linked to the myth of the Orient that was being formed and the colonial expansion of European powers in Indonesia[8] . The reaction of the scientific community was one of scepticism from the outset. In 1789, the Batavian Society of Arts and Sciences in Java published an article refuting Foersch's theories and proving their fraudulent nature, which was followed by other objections from English, Dutch, and French scientists, and in 1810 by a harsh article by the botanist Jean-Baptiste Louis Théodore Leschenault de la Tour. Further direct confirmation of the scientific fraud perpetrated by Foersch came in 1811, when England occupied the territories previously controlled by Holland in Java. But of course, in the meantime, the article had circulated, entering the public imagination, and had been taken up by other authors, including Erasmus Darwin, Charles's grandfather, a botanist and poet, who included the reference to the upas tree in his 1791 poem, "*The Botanic Garden"*, calling it 'the tree of death.' Other literary figures were inspired by this poem, such as Lord Byron in "*Childe Harold's Pilgrimage"* (1812-1818), Samuel Taylor Coleridge in his 1813 tragedy *Remorse, and*

---

that involved various literary figures, cf. Charles E. Robinson, "Mary Shelley and the Roger Dodsworth Hoax", *Keats-Shelley Journal*, 24 (1975), p. 20-28.

8 They are explained in detail in Tim Hannigan, "Beyond control: Orientalist tensions and the history of the 'upas tree' myth", *The Journal of Commonwealth Literature*, 55 (2020) 2, p. 173-189, first published January 29, 2018, https://doi.org/10.1177/0021989418754345.

Alexander Pushkin in his 1828 poem "*Ancar*". The time that elapses between the publication of a fraudulent article and when one becomes aware of the fraud is a crucial aspect to which we shall return when dealing with the contemporary world, since even today it is a variable extension in which false information circulates freely, now moreover with the virality of the Net.

Figure 2. *People reach for an alcoholic drink falling from a pile of barrels of liquor likened to the upas-tree; skeletons litter the ground. Coloured etching by G. Cruikshank, c. 1842. Image in the public domain (source Wellcome Collection free museum and library in London https://wellcomecollection.org/works/k529wwv5).*

To give yet another well-known example, one may recall another of the biggest scientific frauds in history, that of the Piltdown Man, which came to light after a lapse of no less than forty years. In 1912, the palaeontologist Arthur Smith Woodward, curator of the geology section of the British Museum, and the antiquarian Charles Dawson announced in a publication the discovery of fossil remains of a skull and mandibular bone showing the link between ape and man in a hominid that was given the scientific name *Eoanthropus dawsoni*, after Dawson's surname. Although several scientists had, over time, raised suspicions about the authenticity of the material, it was not until the 1950s, thanks to the use of new analysis techniques, that it was definitely established that

it was, in fact, medieval animal remains, artificially modified with the intention of organising a hoax[9] .

The reasons for such actions are not always known or explicitly stated by the perpetrators, but it is clear that the main objective in those days was to achieve fame and prestige and to seek the scientific authority that acceptance by scientific academies could confer on the author of the discovery, starting with the influential Royal Society. Approaching the present day chronologically, however, we can see how the motivations are broadening. Symbolic capital, linked to prestige, as defined by Merton, continues to remain an important but not decisive lever given that even in the 1960s, the same author declared the absence of fraud in the annals of science, albeit admitting the lack of data, but maintained his belief in the public and controllable character of science[10] and its self-correcting dynamics. Certainly, it is in the following years that the problem begins to surface, with an accumulation of incisive factors, from the entrepreneurship of academic publishing and the consequent prevalence of economic aspects, to the procedures of research evaluation and career advancement that feed the culture of publish or perish, and even media popularity, which extend the range of causes but also, if not the number at least the visibility of fraud, as some of the most notorious episodes of bad science testify.

In the biological sciences, the story of Haruko Obokata, a Japanese stem cell researcher, and the article she published jointly with three coauthors, Yoshiki Sasai, Hitoshi Niwa, and Teruhiko Wakayama, in the prestigious journal Nature on 29 January 2014 is famous. The article confirmed a theory that had been tried for some time and thus became the subject of debate within the scientific community. The topic of the research was so important that it immediately led other scientists to verify and attempt to reproduce it and consequently to raise the first doubts as they failed to do so. A committee was even formed with the task of investigating the article. As a result of the investigation of the committee, Obokata was accused of using deliberately falsified data and an image that was already in her doctoral thesis, which was also falsified. Attempts to justify the action as an unintentional mistake were made to no avail. The scientific committee realised the malicious intent from certain clues, such as the fact that the photo caption had been changed[11] . The three coauthors who had supported Obokata's research were also found guilty of misconduct but limited to the fact that

9 Isabelle De Groote et al, "New genetic and morphological evidence suggests a single hoaxer created 'Piltdown man'", *Royal Society Open Science* 3 (2016) 8, https://doi.org/10.1098/rsos.160328.

10 R. K. Merton, *Social Theory and Social Structure,* cit.

11 David Cyranoski, "Stem-cell scientist found guilty of misconduct", *Nature* 01 April 2014, https://www.Nature.com/news/stem-cell-scientist-found-guilty-of-

they had not verified the data provided by the researcher. The committee could go no further in its accusation against the co-authors because the notes provided by Obokata lacked the data and information necessary to reconstruct the story in detail and ascertain the extent of the tampering. Attempts to reproduce the research with the data provided by the author all failed, further confirming the falsification, and the article was retracted in July of the same year by the journal Nature[12] . A month later, one of the coauthors, Yoshiki Sasai, director of a large research centre in Japan, took his own life. The affair attracted media attention for a long time, with bombastic headlines about scandal, lies, fraud, intrigue, and cheating in the world of scientific research[13] .

In the medical field, an equally resounding case concerned Don Poldermans, an eminent and prolific Dutch researcher, professor at the Erasmus Medical Centre in Rotterdam, and a reference in the field of cardiovascular medicine with more than 500 peer-reviewed publications to his credit, as well as a series of articles and speeches at congresses that were highly influential in the scientific community, particularly in relation to the use of beta-blockers in surgery. Poldermans was also a member of the European Society of Cardiology (ESC), where he was the coordinator of the European task force that developed guidelines in relation to cardiovascular surgery[14] . Doubts had been hovering over the professor's work for years due to his nonprotocol-compliant methods. In 2011, he was accused of falsifying data, inaccurate data collection, and using sensitive data from his patients without their consent. In situations of

misconduct-1.14974; Id., *"Stem-cell method faces fresh questions"*, *Nature* 18 March 2014, https://www.Nature.com/news/stem-cell-method-faces-fresh-questions-1.14895.

12 H. Obokata et al, *RETRACTED ARTICLE*: "Stimulus-triggered fate conversion of somatic cells into pluripotency", *Nature* 505 (2014) 641-647, https://doi.org/10.1038/Nature12968.

13 James Gallagher, "Stem cell scandal scientist Haruko Obokata resigns", *BBC News*, 19 December 2014, https://www.bbc.com/news/health-30534674 ; Margherita Fronte, "L'imbrogliona delle staminali", *Focus* 3 July 2014, https://www.focus.it/scienza/salute/l-imbrogliona-delle-staminali ; John Rasko, Carl Power, "What pushes scientists to lie? The disturbing but familiar story of Haruko Obokata", *The Guardian* 18 Feb 2015, https://www.theguardian.com/science/2015/feb/18/haruko-obokata-stap-cells-controversy-scientists-lie ; Dana Goodyear, "The stress test: Rivalries, intrigue, and fraud in the world of stem-cell research", *The New Yorker*, February 22, 2016, https://www.newyorker.com/magazine/2016/02/29/the-stem-cell-scandal.

14 Jan Petter Myklebust, "Netherlands: World heart research expert fired", *University World News* 23 November 2011, https://www.universityworldnews.com/post.php?story=20111123173136776; Larry Husten, "Prominent Dutch Cardiovascular Researcher Fired for Scientific Misconduct", *CardioBrief*, 17 November 2011, https://www.cardiobrief.org/2011/11/17/prominent-dutch-cardiovascular-researcher-fired-for-scientific-misconduct/; Vineet Chopra, Kim A Eagle, "Perioperative mischief: the price of academic misconduct", *American Journal of Medicine* 125 (2012) 10, p. 953-955, https://doi.org/10.1016/j.amjmed.2012.03.014.

this kind, which unfortunately are not uncommon in the medical field, the ethical aspects take on a double negative connotation: on the one hand, towards the scientific community that receives the false results presented in the articles, with the repercussions in terms of health risks in the development of further studies and cures, drugs, therapies that such unlawful behaviour can bring, and on the other hand towards the patients in relation to the use of health data, considered sensitive data. In Europe, GDPR (*General Data Protection Regulation*), enacted in 2016, definitively clarified its nature, regulating its use and penalties for violations. Furthermore, the article by Poldermans claimed that the procedure had been evaluated by two cardiovascular physicians, who denied knowledge of the study and participation in it and stated that they did not even know that their names appeared in the article. Requests to retract the article sent by the two scientists to the New England Journal of Medicine, one of the most prestigious journals in biomedicine, were ignored. Evidently, the credibility of the author, an established researcher and scientist, was greater than that of the person who requested the retraction[15] . The media campaign following the episode was similar to that of the Haruko Obokata case. The consequences of such behaviour in the medical field are enormous as it has a devastating impact on public health. In this case, the retracted article concerned a preoperative treatment that could put a person's life at risk during surgery, as it was later discovered that the substance suggested in the article actually increased the risk of death during surgery.

In addition, in the medical field, an Italian surgeon, Paolo Macchiarini, published a seemingly revolutionary article in The Lancet in 2008 on a trachea transplant in a cancer patient. The innovative technique described was used for the first time, and, according to him, it worked. The article brought him instant fame and prestigious appointments, as well as recognition from the Swedish institute where he was employed after the article came out, the Karolinska Institut. Thereafter, Macchiarini continued to publish articles in The Lancet about other operations. What he forgot to write in all the articles was that the operated patients had all died after the operation. Doctors at his own hospital, suspicious of the discrepancy between the facts they knew about the patients' deaths and what Macchiarini had written in the articles, raised the case, but it was 'covered up' by both the Karolinska Institut and The Lancet. Once again, the mass media brought the issue to the public's attention, a good eight years later, in 2016. First in an article that appeared in the popular magazine Vanity Fair, then with a documentary broadcast on Swedish TV, journalists brought the truth to the

15 "Critics of Poldermans' work baffled by NEJM stance on DECREASE papers", *Retraction Watch* September 29, 2014, https://retractionwatch.com/2014/09/29/critics-of-poldermans-work-baffled-by-nejm-stance-on-decrease-papers/.

surface and forced the rector of the Karolinska Institute, who had supported and defended the Italian doctor, to resign and fire Macchiarini. However, even after the surgeon's serious responsibilities were established, he was incredibly able to continue his career undisturbed in Russia[16] .

Pattium Chiranjeevi, professor of chemistry at Sri Venkateswara University in Tirupati, India, was accused of falsifying up to 70 articles published in 25 journals over a period of only three years. The journals were mostly Western, peer-reviewed, and also published by leading scientific publishers such as Springer and Elsevier. Five of these, specifically Talanta, Food Chemistry, Journal of Hazardous Materials, Analytica Chimica Acta and Chemosphere, were published by Elsevier, which then retracted the articles (as many as thirteen). Chiranjeevi's tactic was to send a large number of articles, from different email addresses, to many journals until someone accepted them. Forgeries and plagiarism were so obvious to chemistry experts that many scientists wondered how it was possible that the articles were accepted and passed peer review[17] . This is one of many unmistakable indications of the much-felt need to rethink peer review today.

In the field of economics, we find the case of two Harvard University scholars, Carmen Reinhart and Kenneth Rogoff, promoters of a theory with which they argued the futility of offering financial aid to governments with a public debt ratio above 90% of GDP because, according to them, aid does not contribute to improving the country's situation, and debt of this magnitude then tends to decrease naturally anyway. The statistical and economic data on which they based their theory were incomplete, as the authors intentionally omitted some of it from their studies and graphs used to visually support their theory (this is the phenomenon called p-hacking, which we will return to later). In fact, their calculations proved to be wrong in all attempts to reproduce them, exposing the fraud[18] .

The Sokal affair concerns the humanities. In 1996, Alan Sokal, a professor of physics at New York University, submitted a completely made-up article with no scientific basis, entitled "*Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity", in the journal Social Text': Toward a Transformative Hermeneutics of Quantum Gravity*", to the journal Social Text, with the intention of demonstrating the

16 Stuart Ritchie, *Science Fictions. Exposing Fraud, Bias, Negligence and Hype in Science*, London, The Bodley Head, 2020.

17 William G. Schulz, "A Massive Case Of Fraud", *Chemical & Engineering News* 86 (2008) 7, https://cen.acs.org/magazine/86/8607.html.

18 Adam Marcus, "Influential Reinhart-Rogoff economics paper suffers spreadsheet error", *Retraction Watch* April 18, 2013, https://retractionwatch.com/2013/04/18/influential-reinhart-rogoff-economics-paper-suffers-database-error/.

problematic nature of article selection mechanisms in cultural journals, particularly when they are ideologically aligned as in the specific case of the academic journal of the experiment, which is clearly left-wing feminist and poststructuralist in slant. Sokal considered such journals to easily be misled with titles that are consistent with the ideology followed, especially in the absence of a robust peer review process. The article was accepted, as the journal at the time did not practice formal peer review, and his theory confirmed[19]. Later, after Sokal, similar experiments were carried out with other journals in the cultural and humanistic fields[20] .

The examples could be numerous, but the objective pursued in these pages is, above all, to show how misconduct in science is not a recent phenomenon. However, the issue is growing, and the implications are much deeper today than in past centuries and even a few decades ago. The increased pressure to publish, and the quantitative rather than qualitative emphasis on publications in evaluation processes, lead to an increase in cases of misconduct, especially because the natural timeframe of the research process is shrinking, and it is not possible to take the time to go into detail on all aspects of a research and to go into depth in the analysis of data or scientific literature. The search for visibility and authority remains one of the main motivations, and the mechanisms facilitated by the virality of the Internet and the media's increasing attention to scientific issues and university life are triggered by it today. However, there is no doubt that the spread of open access to publications contributes to making the discovery of cases of scientific misconduct more likely and frequent, even for the media.

The growth in the number of cases of questionable ethics and integrity has resulted in increased attention to the problem, which has led to the creation of a site such as Retraction Watch, with the aim of keeping track of articles that have been retracted for various reasons of error or scientific misconduct[21] . To give an idea of the extent of the criticality, suffice it to say that a search in the Retraction Watch database of articles on Covid-19 that have been retracted since the beginning of the coronavirus pandemic, i.e. January 2020, returns more than 437 retracted articles[22] . This clearly indicates that the problem is serious, is growing and is heaviest at times of increased pressure, as a global health emergency

19 Mara Beller, "The Sokal Hoax: At Whom Are We Laughing?", *Physics Today*, 51 (1998) 9, https://doi.org/10.1063/1.882436.

20 Yascha Mounk, "What an Audacious Hoax Reveals About Academia", *The Atlantic* October 5, 2018, https://www.theatlantic.com/ideas/archive/2018/10/new-sokal-hoax/572212/.

21 Retraction Watch is a site that registers cases of retraction of scientific articles https://retractionwatch.com/. We will deal more with this site in a later section.

22 Given in August 2024.

can be, and makes clear the disastrous consequences that false science can have on people's lives. The case of the article on the alleged (and false) beneficial effects of hydroxychloroquine in the treatment of the Covid-19 virus, published once again in The Lancet and later retracted, was there for all to see, as was the damage it added to the already shaky (for some) scientific credibility during the pandemic[23] .

2. Violations of the ethics and integrity of science

As we have previously pointed out, ethically reprehensible behaviour that threatens the integrity of scientific research is not new. As early as 1830, Charles Babbage, in his reflections on the decline of science in England,[24] specifically addressed the types of illicit behaviour of those involved in scientific observations and experiments, classifying them into four categories: *hoaxing*, *forging*, *trimming*, and *cooking*. The English word *hoax, which* we could translate as hoax or prank, indicates an action deliberately planned to deceive one or more people, even for fun. To explain the way *hoaxing* operates in science, Babbage gives an example, deprecated as a deceptive action that cannot be justified[25] . In 1788, Giuseppe Gioeni (1743-1822), a Sicilian naturalist and volcanologist, duke of Anjou, knight of Malta, and professor at the University of Catania, published the news of the discovery of a testaceous mollusc named after him, *Gioenia Sicula, in* Naples, with great detail, describing its structure, movement, and accompanying the article with drawings. The description was reproduced in the *Encyclopédie méthodique*, a monumental work in more than 200 volumes based *on the* Diderot and Dalembert *Encyclopédie ou Dictionnaire raisonné des sciences, des arts, et des*

23 Elizabeth Redden, "Rush to Publish Risks Undermining COVID-19 Research", *Inside Higher Ed*, June 8, 2020, https://www.insidehighered.com/news/2020/06/08/fast-pace-scientific-publishing-covid-comes-problems ; James Heathers, "The Lancet has made one of the biggest retractions in modern history. How could this happen?", *The Guardian*, June 5, 2020, https://www.theguardian.com/commentisfree/2020/jun/05/lancet-had-to-do-one-of-the-biggest-retractions-in-modern-history-how-could-this-happen; Nathan M. Greenfield, "Why so many people mistrust science and how we can fix it", *University World News* 16 August 2022, https://www.universityworldnews.com/post.php?story=20220816072300893. The problem of mistrust in science is complex and has a lot to do with the lack of what is called 'scientific citizenship', but the episodes of maliciousness spread by the media certainly do not foster credibility.

24 C. Babbage, *Reflections on the Decline of Science,* cit.

25 Although it seems that Gioeni had meanwhile acknowledged the error, as recounted in Lisa Signorile's article, Lisa Signorile, "Un'apologia di Giuseppe Gioeni – naturalista", *Le Science Blog* 16 luglio 2021, http://lorologiaiomiope-lescienze.blogautore.espresso.repubblica.it/2021/07/16/unapologia-di-giuseppe-gioeni-naturalista.

métiers. But the fact is, writes Babbage, that such a mollusc does not exist, since what Gioeni saw was another type of mollusc, already known beforehand, to which he added the fruits of his imagination. *Forging* implies forgery or the creation of illegal copies of something with the intent of deceiving. Babbage explains how it differs from hoaxing because while in the former case the duration of the hoax may be limited in time as it is directed towards a specific goal, in the case of deliberate forgery, the intent is to acquire a scientific reputation and, therefore, can be protracted in time. Again, Babbage gives an example, the false observation of the second comet by Chevalier D'Angos in 1784, which was only seen by him and no other scientist or expert. Fortunately, Babbage adds, cases of *forging* are rare. *Trimming* consists of cutting data that differ too much from the average, making a 'fair adjustment', so as to remain prudently adherent to the average results, avoid guesswork, and thus give an idea of accuracy. In other words, one places oneself in a certain and safe vein of mainstream research. However, for Babbage, this is less serious than in the fourth case. *Cooking*, which today we might translate not quite literally as manipulating, means bending data in a way that is functional to one's research. The culinary metaphor fits perfectly with the different recipes used to "cook" data, ranging from choosing from different available data only those that demonstrate the results of one's research, to knowingly measuring them with some tools and methods rather than others for the same reason.

Today, we have a broader case study, but the main categories of fraudulent activities remain those indicated by Charles Babbage and, as already discussed, in all cases the fundamental distinction lies in the intentionality of the action on the part of the perpetrator. According to the LaFollette distinction proposed earlier, the violation may concern ethical and moral aspects, scientific standards, or laws. More recently, COPE (Committee on Publication Ethics), an association of publishers founded in 1997 with the aim of supporting publishers and the entire scientific community in preventing and dealing with unethical and fraudulent behaviour, has precisely listed the cases and defined a code of conduct[26] . To the categories formulated by Babbage, we can add a behaviour that can be partially likened to *cooking*, but nowadays declined in multiple further variants of possible bias. We will deal with fraudulent and bad conduct cases, such as p-hacking later on, but it is worth noting that even biases are not always based on bad intentions, as they are often unconscious.

Complete objectivity is utopian because each researcher is influenced by the context in which he or she operates and the state of knowledge of the era in which he or she lives, which only subsequent years, sometimes centuries, will be able to confirm or refute and over-

26 COPE publicationethics.org.

come with new knowledge. An example of this is the theories and the success achieved by the Italian psychiatrist Cesare Lombroso in his time, at least initially and despite the objections that came along with the enthusiastic exaltations. Lombroso, who became Professor of Forensic Medicine at the University of Turin in 1875[27] , perfected the preexisting theory of atavism, drawing on strands of study still considered central and important in his time, such as physiognomy, the origins of which can be traced back to ancient Greece, and phrenology developed by the German physician Franz Joseph Gall in the 19th century. To this must be added the further influence of Darwin's theories, popularised in Turin after 1871 by Filippo De Filippi, Michele Lessona and Giovanni Canestrini, but probably not yet received with sufficient depth in Italy, and other earlier writings. Lombrosian theories on the natural origin of the predisposition to crime are the fruit of genuine convictions based on the consultation of an extensive scientific literature but conditioned by the thinking of his time. The 19th century approach to scientific knowledge as definitive truth would only be challenged at the turn of the century with the ideas of Einstein, Freud, Russell, and others. Lombroso's theories were, moreover, amplified and extolled by the leading scientific journals in the field such as Archivio di psichiatria, founded by Lombroso himself (who was its director until 1909) and Raffaele Garofalo in 1880, and La scuola positiva, founded in 1891. However, from another journal, the Rivista penale, came harsh criticism from authors such as Enrico Pessina and Luigi Lucchini, along with various other strands of protest in Italy and abroad[28] .

Therefore, the role of scientific journals has always been crucial in the dissemination of theories and scientific research results. And it is precisely the impact that one aspires to achieve through journals, and in recent times through the mass media as well, that often generates that long tail of unreliable studies, with erroneous data, conditioned by bias, whose results, even if later retracted, remain in the public perception for a long time. We have seen an example of this with the upas tree, and another, much more recent example is the case of the food psychology studies conducted by Brian Wansik, director of the Food and Brand Lab at Cornell University, author of successful books and hundreds of articles, in the early 2000s. He is credited with some wide-

27 The University of Turin preserves the archival documentation of his life and career at the Museum of Criminal Anthropology that bears his name, a precious testimony to his studies, as well as the numerous honours and awards he received during his career. The archive is accessible online at the URL https://www.museolombroso.unito.it/info/archivio-storico/.

28 Paolo Marchetti, "Il Contributo italiano alla storia del Pensiero – Diritto: Cesare Lombroso", in *Enciclopedia Treccani*, 2012, https://www.treccani.it/enciclopedia/cesare-lombroso_%28Il-Contributo-italiano-alla-storia-del-Pensiero:-Diritto%29/.

spread beliefs such as the idea that if you serve food on a large plate, you will eat more than you would have eaten and it is therefore preferable to use a smaller plate or that if you go shopping at the supermarket when hungry you will buy more high-calorie foods[29] . It was only in 2016 that Wansik revealed that experiments undertaken to prove such theories had come to nothing, but nevertheless he had published the data. Such ideas still circulate today and are periodically picked up on television and in popular magazines, so much so that they have now entered the collective imagination that many people, if questioned, would certainly prove to be aware of them.

The prestigious journal Nature is conducting a process of reflection and even self-criticism, on the role of journals in perpetuating scientifically unfounded ideas that later prove to be unreliable and harmful in the long run. An example of this is the articles on eugenics published in Nature starting with Francis Galton's speech that appeared in the pages of the journal in 1904[30] . Galton was the inventor of the term 'eugenics' and theorised its principles, which had a widespread diffusion in the first decades of the century and devastating consequences during Nazism, well known to all. A study of the archives of Nature has revealed anti-Semitic articles, despite the journal's openly anti-Nazi stance, sexist and racist articles, and discrimination of various kinds, up to recent times. The great issue that opens up is central in science and in the venues that host the results of scientific research, such as journals, which are often affected by the ideas and prejudices of their editors and scientific committees, often not very diverse in terms of gender, culture, nationality. The issue is extremely topical and important with respect to the permanent consequences that such aspects of ethics and integrity can generate and would certainly require a more in-depth and specific treatment than the objectives of this work allow. Sometimes discriminatory attitudes and ethical violations can be traced back to the personal convictions of an author, who may genuinely be convinced of his ideas, however radical and scientifically unfounded. Sometimes the bias in the approach to research and in the treatment of data is a voluntary action that may conceal economic motivations and is the consequence, for example, of an undeclared conflict of interest or of research financed by companies and thus tied to the achievement of a positive result.

Another emblematic case concerned an article published again in Nature in 2018 on research on quantum computers funded by Microsoft, which was retracted in 2021. The data contained in the research seemed to pave the way for the development of more powerful

29 S. Ritchie, *Science Fictions,* cit., p. 97-98.

30 "How Nature contributed to science's discriminatory legacy"*, Nature 609* (2022), p. *875*-876, https://doi.org/10.1038/d41586-022-03035-6.

computers based on quantum superconductors that can play an important role on global issues such as climate change. The release of the article was immediately picked up by Microsoft in the competition between companies working on this type of computer development, but it emerged that some data had been fraudulently altered to support the evidence of the discovery and an incorrect caption had been attributed to a graph, making it misleading. The article was retracted by Nature and, given the hype that Microsoft itself had given to the alleged discovery at the time of publication, the case was also picked up by the major newspapers and other media,[31] and fed to the public as an example of maliciousness. Journalists did not fail to point out that the journal Nature has recorded 79 cases of retracted articles since its foundation in 1869, including eight in 2020 alone, helping to undermine the credibility of one of the most authoritative scientific publications. There could be many more, but it is already clear from the few reported that the problem does not spare the most prestigious publishers and journals and, moreover, again brings up again the crucial issue of peer review. The forcing of scientific research towards positive results at any cost, and often only positive results to the exclusion of negatives, is an equally crucial issue.

The set of stresses that influence scientific research and have consequences on publishing production induce numerous distortions. Several studies conducted in Italy and abroad show certain trends and changes in researchers' behaviour, which are believed to be specifically attributable to the widely referenced rapid publication requirements. These include salami slicing, i.e. the choice of publishing split content in many different locations rather than a single overall one, such as a monograph. A survey published in GigaScience in May 2019[32] , and conducted on millions of publications outlined several relevant phenomena such as the increase in the number of publications, the increase in publications on so-called megajournals such as PLoS Online or Nature's Scientific Reports, the increase in the average length of articles and the increase in the average number of authors per article, which more and more reaches hundreds or even thousands of names.

According to the study reported in GigaScience, in general, the number of publications has increased (from 1 million in 1980 to 7 million in 2014) and the speed at which researchers publish. There have even been significant changes in the wording of article titles, with an increase in

31 Cody Godwin, James Clayton, "Microsoft-led team retracts quantum 'breakthrough'", *BBC News* 10 March 2021, https://www.bbc.com/news/technology-56328980.

32 Michael Fire, Carlos Guestrin, "Over-optimisation of academic publishing metrics: observing Goodhart's Law in action", *GigaScience*, 8 (2019) 6, p. 1-20, https://doi.org/10.1093/gigascience/giz053.

words and punctuation marks (such as exclamation marks, question marks, etc.) that are more likely to catch the eye. In science, the quest for popularity takes on the characteristics already widely present in clickbait-dominated journalism. The number of clicks is the measure against which newspapers now measure themselves, and it has become the criterion to sanction the effectiveness of an article on which the sale of advertising space can be calculated, as can easily be seen by consulting the pages of newspapers on the Internet, where the news is now flooded with advertising inserts, pop-ups, and spots in news videos. Even journalists are sometimes paid based on clicks received from their articles[33] . The mechanism being established in scientific research risks generating similar drifts. Scientists are looking for clicks because this increases the popularity and dissemination of an article (the more it is clicked on, the more it remains in circulation and visible in social networks, for example), and consequently also the citations that then contribute to the bibliometric indices on which they are evaluated. The increasing emphasis placed on the third mission of universities, and in particular on public engagement, represents another lever for scientific clickbait, extended to all disciplines, not only the bibliometric ones. The more the titles of articles attract the public, even the general public, the greater the interaction that can be measured in the various evaluation practices, such as the number of participants in a public initiative (and views if it is online), a popular conference, a television or radio broadcast. Even ANVUR, the National Agency for the Evaluation of the University System and Research in Italy, measures public engagement with such quantitative data. A particularly catchy title affects another type of measurement, such as altmetrics. Digital convergence has unified access routes to different types of information that were previously separate. To consult a printed scientific journal in the pre-Internet era, one had to go to the library and leaf through its pages, in a dedicated time and space. Today, the scientific article can be read from a smartphone, newspapers, social networks, or any website, and is part of an information overload to which we are all subjected, now without any distinction of time and space, as everything happens simultaneously and in a common space. Competitive pressure on researchers and universities, in which scientific publication is one of the main 'weapons', also for career purposes, for university reputation, for international rankings, to attract students, produces a tendency to exalt and amplify achievements.

During the coronavirus infodemic, for example, news continued to circulate that the Covid-19 virus had been created in a laboratory in China. This news was then picked up and republished by many pol-

33 See Ben Frampton, "Clickbait: The changing face of online journalism", *BBC News* 14 September 2015, https://www.bbc.com/news/uk-wales-34213693.

iticians and journalists, adding further confusion to the already confusing situation. But the news did not only originate from fake news created by network users and supported by certain statements by people with information influence, but also from scientific articles, such as the one that two Chinese researchers added to ResearchGate, one of the most popular sharing sites used by researchers to share their publications. The article was later withdrawn but was picked up and shared in time and contributed to the rise of the conspiracy theory, forcing a group of scientists to publish an open letter condemning the theory as having no scientific basis[34] . A similar case had occurred with Andrew Wakefield's article published in 1998 in the journal The Lancet, which claimed the existence of a link between vaccinations and autism, and which was withdrawn from the journal a few years later because it was based on falsified data, leading the British General Medical Council to expel the author from the Medical Council. But this was not enough to stop the spread of the theory that the vaccine (by now it has become 'any vaccine') causes autism. The reader who has no medical expertise, nor knowledge of the mechanisms of research, will not be able to verify false content, or what the presence of the word 'retracted' in front of an online article implies. Furthermore, as a study in Retraction Watch shows, many researchers also 'missed' the word 'retracted' and when they cited Wakefield's article, they did not report the fact that it had been retracted (the case, by the way, is exemplary of how misleading and damaging the use of citations in the evaluation of publications can be)[35] . We have seen many similar cases with regard to Covid-19, partly because most journalists and members of the public have no knowledge of the ethical issues involved in research, of phenomena such as the 'reproducibility crisis' and of the importance of reproducibility as a validation tool for any scientific experiment. Journalists draw on unsupervised sites such as ResearchGate or Academia, or on pre-print repositories, where articles not yet peer-reviewed are published. The reference context of an article in a repository can perhaps be easily reconstructed by a researcher but not by a general user who accesses that open repository without always having the means to assess its reliability.

The situation is obviously exacerbated in times of health emergencies such as the coronavirus. John Inglis, one of the managers of the

34 *Open letter. Call for a Full and Unrestricted International Forensic Investigation into the Origins of COVID-19,* https://int.nyt.com/data/documenttools/virus-inquiries-pandemic-origins/d7a097a4c758a65a/full.pdf.

35 Ivan Oransky, "Andrew Wakefield's fraudulent paper on vaccines and autism has been cited more than a thousand times. These researchers tried to figure out why", in *Retraction Watch*, 18 November 2019, https://retractionwatch.com/2019/11/18/andrew-wakefields-fraudulent-paper-on-vaccines-and-autism-has-been-cited-more-than-a-thousand-times-these-researchers-tried-to-figure-out-why/.

biology preprint archive bioRxiv and co-founder of the newer medicine archive medRxiv, pointed out in a recent article in The New York Times what has happened in the weeks since the pandemic spread[36] . Accesses to medRxiv have increased more than a hundred times, and the submissions of preprints by authors have also increased considerably. Authors who deposit on medRxiv must declare at the time of submission that the article is ethically acceptable, that it is deposited with the consent of all persons involved, and that it reports significant experiments; there is a validation process and verification that it is not plagiarism, that it has a scientific basis and that it cannot harm people. However, preprints are deposited at an embryonic stage, the contents of which are superseded by updated versions deposited some time later, or it happens that the deposited version differs widely from the final version published by a journal after peer review. Furthermore, according to the study, 30% of the preprints deposited are not published in any journal in the next two years. Some preprints are retracted and withdrawn from the archive by the authors themselves. However, before it is modified or withdrawn, a preprint circulates widely, is cited in other articles, picked up by journals, and even goes viral on the Web. If it contained incorrect information and was later amended or retracted, the first version will still circulate. In fact, medRxiv, BioRxiv, arXiv, as well as other repositories, have now added a banner with a warning stating that the preprints in the archive are preliminary studies not subject to peer review and therefore should not be considered definitive and reported as such in the media. In the general picture of the information that has emerged, effectively termed infodemics[37] , it is doubtful that this will be sufficient. In addition to these dynamics, not always determined by bad faith, of course, there are, on the dark side of open access, phenomena such as predatory journals and all deliberate practices of scientific fraud. Phenomena of maliciousness that do nothing but feed what is called 'science shaming', i.e., the continuous contestation of scientific authority, which we have witnessed abundantly since the beginning of the pandemic.

3. Considerations of authorship

The aspects related to authorship take on different declinations, but on the whole represent a significant part of the deviations from sci-

36 Wudan Yan, "Coronavirus Tests Science's Need for Speed Limits", *The New York Times* 14 April 2020, https://www.nytimes.com/2020/04/14/science/coronavirus-disinformation.html.

37 For the etymology of the term, see the Accademia della Crusca page https://accademiadellacrusca.it/it/parole-nuove/infodemia/19506.

entifically acceptable behaviour, such as to merit a specific in-depth study. These include extreme and sensational actions, but also many very subtle criticalities that are not easy to discern. One of the most striking episodes concerned the attempt, fortunately not so frequent in the scientific literature, to create an ad hoc journal to allow a researcher or his research group to publish with fictitious or piloted peer review. The case of Mohamed El Naschie, editor of the physics journal Chaos, Solitons & Fractals, is among the best known. The journal founded by him in 1991, and published by the Pergamon Group (bought by Elsevier in 1992), published 269 articles by its founder between 1991 and 2008, amounting to 85% of his total scientific output, of which 53 in 2008 alone, and, moreover, his articles were cited almost exclusively in the same journal[38] . When this phenomenon was discovered by other researchers during a network analysis study on citations, a debate ensued in the journal Nature. El Naschie sued Nature for defamation, but later withdrew the complaint in 2012. In the meantime, Elsevier forced him to resign from the editorship of Chaos, Solitons & Fractals.

In relation to the issues of ethics in authorship, the study published in GigaScience[39] recorded an increase in cases of *honorary authorship*, which in fact did not contribute significantly to the article, and *ghost authorship*, whose contribution, on the contrary, is not recognised in the article. Honorary authors are usually added because they hold important positions in the research group, in the department; perhaps they are responsible for the funding, or they are figureheads according to the dynamics traceable to the Matthew Effect[40] . The 'free' granting of authorship in an article may also take the form of *gift authorship, the* aim of which is the exchange of favours, in the hope, therefore, that the colleague will do the same later. *Ghost* authors, on the other hand, are often early-career researchers, especially Ph.D. students, who see their research results published by the research group without appearing as authors[41] .

Another widely manifested trend is the increase in the number of collaborations, known as *hyperauthorship* or *hypercollaboration.* The authorship of an article by numerous coauthors is especially characteristic of STM disciplines and particularly of physics and medicine, where

38 Yves Gingras, Mahdi Khelfaoui, "Scientific publication - Is it for the benefit of the many or the few?", *University World News*, 11 July 2020, https://www.universityworldnews.com/post.php?story=20200709125857707.

39 M. Fire, C. Guestrin, "*Over-optimisation of academic publishing metrics",* cit.

40 Robert K. Merton, "The Matthew Effect in Science" *Science 159*, no. 3810 (1968): 56–63, https://doi.org/10.1126/science.159.3810.56.

41 As recounted in Karishma Bisht, "Don't make early career researchers 'ghost authors.' Give us the credit we deserve", *Science 9 September 2021*, https://doi.org/10.1126/science.caredit.acx9061.

it is common practice to list all the contributors to an experiment, including those who participated only marginally in the article and the research that generated it. According to some studies, in 1981 the average number of authors for a physics article was 118, and this average has gradually increased to several hundred or even thousands. In 2015, the record number of 5,154 authors for an article published in the journal Physical Review Letters was reached, belonging to 344 different institutions and alphabetically listed on no less than 24 pages of the journal, as noted in a study conducted by Gianfranco Pacchioni, full professor of General and Inorganic Chemistry and then Pro-Rector for Research at the University of Milan Bicocca[42] . Delving deeper into the analysis, Pacchioni quantified how the first of the authors listed in the article in question, chosen as a case study only because he was the first on the list, had a production of 500 articles between 2012 and 2016, of which 123 in 2016 alone, amounting to a decidedly improbable average of one article every three days, including Saturdays, Sundays, and all holidays[43] . Evidently, he had been included in many articles as an author without having actually contributed to their writing. Blaise Cronin noted the dramatic nature of the phenomenon of hyperauthorship as early as 2005, particularly in the field of biomedicine and high-energy physics, where it appeared so widespread that, in his opinion, the reliability of scientific communication was called into question[44] . In his reflections, he then added the observation of the poor perception of the seriousness of the problem by scientists in those fields and even more so in other disciplinary fields.

As Cronin reminds us, the concept of author has undergone continuous evolution, culminating in the centrality it has assumed in our times. According to Alberto Manguel, quoted in Cronin's volume, the first to sign herself as author in history was Princess Enheduanna in ancient Mesopotamia, who affixed her name to the end of tablets on which were engraved, in cuneiform writing, hymns she composed in honour of Inanna, the goddess of love and war. The evolutionary path reaches the postmodern conception of writers such as Roland Barthes (who wrote *The Death of the Author*) and Michel Foucault (*What is an Author?*). Along this path, some moments marked significant transformations. The first is the birth of the scientific journal, which begins to link the author to his scientific discovery, but in a still very bland manner. The concept of the author will be strengthened later on, with Romanticism and the contemporary age, and also thanks to the copy-

42 Gianfranco Pacchioni, *The Overproduction of Truth, Passion, Competition, and Integrity in Modern Science*, Oxford, Oxford University Press, 2018, p. 49-50.

43 *Ibid.*

44 Blaise Cronin, *The Hand of Science. Academic Writing and Its Rewards,* Lanham, The Scarecrow Press, 2005, p. 13.

right laws that will consolidate the author's relationship with his work. The professionalisation of science introduces forms of collaboration in the conduct of experiments and then the coauthoring of articles. In the twentieth century, it is the postwar period, with the aforementioned transformation of science into a large collective enterprise, that encourages and strengthens collaboration between authors.

Cronin's definition of *hyperauthorship*, again quoting Mikhail Epstein, is among the most effective. Hyperauthorship is a form of "authorship dispersed among several virtual personalities which cannot be reduced to a single 'real' personality"[45] . The definition is interesting because it raises the crucial question of responsibility. In a publication where authorship is dispersed among hundreds or thousands of authors, who bears responsibility if doubts are raised about integrity, accusations of scientific fraud, or ethical misconduct? The *corresponding author* is the person named in multiauthored articles as the point of contact with the public, editors, and *stakeholders*, but does not assume any particular responsibility for the content of the article. This is a rather marked difference between the hard sciences and the humanities. In the HSS, publications tend to have single authors or at least a low number of coauthors, among whom it is easy to identify responsibility, sometimes even formally divided and explicitly stated in the article or monographic work. How, on the other hand, is it possible to identify the responsibility of over five thousand authors from 344 different institutions? Hyperauthorship, in addition, produces non-worthless side effects. On the one hand, as has been proven by numerous studies, a larger number of authors increases the number of citations and thus affects all bibliometric indicators, to the advantage of authors and journals. On the other hand, with hundreds or thousands of authors, it is easier to conceal practices such as guest, gift, and *surprise authorship*. The latter concerns articles to which authors are added without their knowledge, often authoritative names useful to get the article accepted by the journal or to pass it off as valid and scientifically sound, perhaps in predatory publishing practices, or even to increase citations.

The numerous distortions that have emerged over time have brought the topic of authorship and what is really meant by the definition of author to the centre of the debate. The International Committee of Medical Journals Editors (ICMJE)[46] defined the concept of authorship on the basis of four criteria relating to the functions performed.

> Make substantial contributions to the conception and design of work or

45 *Ivi, p. 48.*

46 ICMJE, "*Defining the Role of Authors and Contributors",* 2021, http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

to the acquisition, analysis, and interpretation of work data.
Creating the draught version of the article and critically reviewing it for important intellectual content.
Approve the final version for publication.
Take responsibility for all aspects of the work so that accuracy or integrity issues can be investigated and resolved appropriately.

Based on these criteria, in 2015, Mozilla Science Lab and BioMed Central, together with PLOS, ORCID, and others, and the CRediT project[47] , developed digital badges to be affixed next to the name of the authors of an article to identify roles[48] . CRediT is the taxonomy elaborated by NISO (National Information Standards Organisation), and widely used by editors, for the correct attribution of the role played by the different people in the realisation of a scientific contribution, and includes fourteen roles defined as follows: *Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualisation, Writing original draft, Writing - review & editing*. The badge was tried out by a couple of journals but soon abandoned due to the difficulty of implementation. Defining the contribution of different people to an article presents an inherent difficulty in being able to identify roles with precision[49] .

The problem of authorship continues to be underestimated, as Cronin already pointed out, and without the commitment of publishers and journal editors, it will not find an effective resolution. In fact, further investigation should be conducted with respect to the responsibilities not only of authors but also of journals and publishers. The practice of adding fictitious authors, sometimes even asking for payment in the case of predatory journals, has not insignificant consequences on the evaluation of research. Likewise, the buying and selling of citations has become flourishing. In addition to the exchange of citations between authors, there are publishers who pressure authors to cite other journals of the same publishing group in order to increase the count, which then goes to form indices such as the impact factor, which has become a surrogate for measuring the prestige of a journal[50] . There is great responsibility on the part of the publishers and editors of the journals, as well as the authors. Evaluation agencies, such as ANVUR in Italy, operate with the basic assumption that issues of research ethics

47 CREDIT, credit.niso.org.

48 Laura Paglione, "*Contributorship Open Badges on ORCID"*, October 20, 2015, https://info.orcid.org/contributorship-open-badges-on-orcid/.

49 B. Cronin, *The Hand of Science,* cit., p. 58.

50 Adam Marcus, "Publisher offers cash for citations", *Retraction Watch* August 31, 2021, https://retractionwatch.com/2021/08/31/publisher-offers-cash-for-citations/.

and integrity are addressed and resolved upstream by the editors and scientific committees of the journals, through the instrument of peer review and others at their disposal. When an article is published in a scientific journal, it should have gone through a certification and validation of its scientific authority, but this is not always the case.

4. Fabrication, Falsification, and Plagiarism (FFP)

Before addressing the fraudulent behaviour outlined in the title of this section, it should be clearly stated again that although many studies have shown that most cases of article retractions are due to fraudulent behaviour, which increased tenfold between 1975 and 2012[51] , it is not always the case that the violation of research integrity is due to fraudulent causes. Unintentional errors may occur at different stages of the life cycle of publication, from data analysis to peer review, and ethically unacceptable behaviour and questionable but not fraudulent practices may occur. Instead, there are a number of actions that presuppose the intention to commit outright scientific fraud, right up to the limits of the criminal act. We have seen descriptions of these by some authors of the past, according to the moral principles and practices contemporary to them, which, however, change over time and vary according to the era. The current standard for defining illicit practices in scientific research is summarised by the acronym FFP (Fabrication, Falsification, Plagiarism)[52] . One speaks of fabrication when data are created ad hoc, describing experiments that did not take place and results that were never obtained; in falsification, data exist but are falsified, i.e. the results of the research or experiment are altered so as to support the researcher's original theory or idea, for instance by omitting contrary data or selecting only positive data, falsifying images and graphical representations; plagiarism occurs when a researcher uses parts of text, data, experiments, taken from someone else's work without citing the source and without requesting permission from the actual authors, or

51 Ferric C. Fang., R. Grant Steen, Arturao Casadevall, *Misconduct accounts for the majority of retracted scientific publications*. *Proceedings of the National Academy of Sciences of USA,* 109 (2012) 42, p. 17028-17033, https://doi.org/10.1073/pnas.1212247109. In the analysis of NIS-funded publications 95.9% were retracted for fabrication or falsification of data, see Andrew M. Stern, Arturo Casadevall, R. Grant Steen, Ferric C. Fang, "Financial costs and personal consequences of research misconduct resulting in retracted publications", "Elife" 2014 3:e02956, published 2014 Aug 14., https://doi.org/10.7554/eLife.02956.

52 David B. Resnik, Talicia Neal, Austin Raymond, Grace E. Kissling, "Research Misconduct Definitions Adopted by U.S. Research Institutions", PubMed Central, 1 January 2015, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4269469/.

from his own previous work without declaring it, and in this case we speak of self-plagiarism.

The Research Council UK (RCUK) specifically lists misconduct[53] , in a classification shared by the scientific community and also adopted by the Medical Research Council[54] , which was initiated by a research integrity agreement signed by British universities. These behaviours, in addition to the cases that fall under the definitions of FFP, include:

False declaration:

Misrepresentation in relation to data, e.g. the omission of results and/or relevant data, whether deliberately or through obvious negligence, leading to misinterpretation of the data.

Undeclared duplication of publication, including submission of the manuscript to several journals without declaring it.

Failure to declare specific interests, including financial or funding interests in the research.

The false declaration of qualifications and/or experience and the declaration of qualifications and/or experience that one does not possess.

False declaration of authorship, either by attributing or attributing oneself the role of author and/or contributor when one has not contributed at all, or in cases of ghost authorship, i.e. omitting the name of an author who has contributed to the research.

Breach of ethical duties, whether deliberately or negligent:

Improperly disclosing the identities of individuals or groups involved in research without their consent or breach of confidentiality.

Endangering persons involved in research in any way, without their prior consent and without safeguarding their rights, including putting their reputation at risk.

Not taking all possible precautions to ensure that the risks and hazards, the general objectives and possible sponsors of the research are known to the participants or their legal representatives, and not ensuring that informed consent is obtained beforehand in an explicit and transparent manner.

Failure to comply with ethical principles and legal obligations with respect to animals, human organs used in research, and environmental protection.

Engaging in improper conduct in peer review, for proposals, projects, or research results (including manuscripts submitted for publication);

---

53 *RCUK Policy and Guidelines on Governance of Good Research Conduct,* 2013, https://www.ed.ac.uk/files/imports/fileManager/RCUKPolicyandGuidelinesonGovernanceofGoodResearchPracticeFebruary2013.pdf.

54 *MRC Policy and Procedure for Investigating Allegations of Research Misconduct*, 2014, https://www.ukri.org/wp-content/uploads/2021/08/MRC-12082021-Research-Misconduct-Policy.pdf.

this includes failure to declare conflicts of interest; failure to explicitly acknowledge limited expertise on the topic; misappropriation of the content of peer-reviewed material; breach of confidentiality, or misuse of material provided in confidence for the purposes of peer review.

Inadequate treatment of allegations of misconduct:

Do not address suspected cases of misconduct, including covering up such behaviour or threats against whistleblowers.

Failure to adequately deal with allegations of malicious behaviour that should be traced back to breaches of good conduct.

The UK Research Council, which became UK Research and Innovation (UKRI) in 2018, is a government body tasked with coordinating UK research and a funding agency for scientific research. In the document in which it classifies scientific malfeasance, it also declares sanctions, which are differentiated according to the seriousness of the breach but can go as far as demanding the return of funds received, either from the accused researcher or his or her institution, as well as exclusion from any future funding. The violation of research integrity not only has consequences in terms of the prestige of the researcher and his or her institution, general repercussions on the credibility of science, and potentially dangerous consequences for society, but also produces economic damage.

A research team from the University of Washington conducted an economic estimate based on funding provided by the US National Institute of Health (NIH) in relation to 291 publications retracted for scientific misconduct, as recorded by the Office of Research Integrity (ORI) in the years between 1992 and 2012. The result of the analysis shows a total cost in relation to these publications of USD 58 million disbursed, corresponding to approximately 1% of the total funds disbursed by the NIH. Each of the retracted articles cost an average of $392,582. Furthermore, the study found a positive correlation with the impact factor of the journals in which the articles appeared, so the cost was higher for articles published in journals with a higher impact factor[55] . It is interesting to note that the discovery of the fraud resulted in a significant reduction in the scientific production of some of the involved researchers, some of whom ceased publication completely in the years following the accusation, but, on the contrary, other researchers continued to publish without interruption, demonstrating that in many cases there is no consequence for their careers and the risk is therefore not a deterrent. The percentage compared to the NIH's total funding may appear low, but this is only one of the funding bodies and refers only to the publications that have been found to be fraudulent, i.e. only the part of the phenomenon that has come to light, which

55 A.M. Stern, A. Casadevall, R.G. Steen, F.C. Fang, *Financial costs,* cit.

the same article, echoing some previous studies, quantifies as 20%. In general, the cases detected by the ORI result in a drastic reduction of research funding by this body for those found guilty of misconduct and sometimes repercussions on their careers, and this is what happened at the NIH. However, the costs of fraudulent behaviour are not only directly related to the funding received and used in research. Resources of various kinds, from the institution's equipment to human resources, are used to publish the articles that are later retracted, both in the performance of the experiments and in the submission phase to the journal, whose validation process culminating in peer review has costs, if only in terms of time. Total indirect costs, if quantified, would greatly increase the above figures. Moreover, the system of science is based on the construction of knowledge from prior knowledge, and therefore other researchers could base their research on the results of falsified articles, wasting time and resources on false and misleading results, with the risk of generating further unreliable articles because they are based on incorrect data. The overall costs should also be quantified on their work and the long tail that fraudulent activities create. We must not forget the implicit and very serious cost of the effects of the spread of falsified science on society, for example, on people's health in the medical field, the general loss of credibility, and above all the crumbling of the foundations on which science is based.

Although these problems have always been there, as we have tried to show to explore their nature and certainly not to try to justify them, their relevance has changed today. On the one hand, because open access favours the circulation of good science as much as bad science and, on the other hand, because the number of publications globally is constantly growing. In STEM fields alone, whose counts are more frequent because they are based on databases such as Scopus and Web of Science, publications in 2018 amounted to 2.6 million (peer-reviewed articles and contributions in conference proceedings), with a global annual growth rate of 4% over the last ten years[56]. China's national impact over the same ten-year period grew at twice the rate of the global average, while the United States and the European Union grew at half the global average. In 2018, China became the country with the highest scientific output in these fields, followed by the United States and India. Italy ranks eighth in terms of the number of publications in the ranking compiled by the US National Science Foundation (NSF), with a growth rate of 2.41%, taking it from 56,157 publications in 2008 to 71,240 in 2018. The NSF study shows other significant data in relation to research trends measured by publications in the investigated fields, in which, in

56 National Science Foundation, *Publications Output: U.S. Trends and International Comparisons*, https://ncses.nsf.gov/pubs/nsb20206/executive-summary.

addition to China, other countries stand out with remarkable publication growth rates that are well above the world average, such as India (10.73%), Russia (9.88%), which is ahead of Italy in the ranking, and Iran (10.99%). Brazil (5.42%) and South Korea (4.17%) also have an above-average percentage increase in publications. But, as we shall see, some of these countries are also among the largest producers of predatory and fraudulent publishing.

During the pandemic, there was an exponential growth in the number of Covid articles, but several turned out to be unreliable[57]. The Retraction Watch website keeps track of retracted Covid articles and lists, as already mentioned, over 260, plus several dubious cases that have not been clarified. This is a low percentage compared to the total number of publications on the subject, but in the meantime, these are only the cases that have emerged in the three years since the pandemic began and a crucial issue of collective health that requires extreme caution. These include several articles that have been retracted by prestigious journals, such as The Lancet and The New England Journal of Medicine (NEJM), both of which were retracted in 2020 for articles based on unreliable data provided by Surgisphere, a data analytics company. The company had provided the data to researchers who processed them and published the results in a series of articles. The articles passed peer review and were published. However, the Surgisphere data, after independent expert analysis proved implausible, both due to the implausible number of patients tested, the quality of their data, and the lack of credibility of the claimed dosages[58]. It was the authors themselves (three of the four authors of the article published in The Lancet) who requested the article be retracted after realising the lack of integrity and reliability of the data provided by the data analysis company[59]. Moreover, when the researchers asked Surgisphere for the raw and complete datasets in order to actually verify their reliability, the company responded negatively, not allowing access and firmly based on reasons

57 Xiaojing Cai, Viola C. Fry, Caroline S. Wagner, "International collaboration during the COVID-19 crisis: autumn 2020 developments", *Scientometrics* 126, 3683-3692 (2021), https://doi.org/10.1007/s11192-021-03873-7; Paola Berchialla, Sara Urru, Veronica Sciannameo, "The effect of COVID-19 on scientific publishing in Italy", *Epidemiologia & Prevenzione* 2021, 45 (6) November-December, p. 449-451, https://doi.org/10.19191/EP21.6.136.

58 Kelly Servick, Martin Enserink, "A mysterious company's coronavirus papers in top medical journals may be unraveling", *Science* Jun. 2, 2020, https://www.sciencemag.org/news/2020/06/mysterious-company-s-coronavirus-papers-top-medical-journals-may-be-unraveling.

59 Charles Piller, Kelly Servick, "Two elite medical journals retract coronavirus papers over data integrity questions", *Science* Jun. 4, 2020, https://www.sciencemag.org/news/2020/06/two-elite-medical-journals-retract-coronavirus-papers-over-data-integrity-questions.

related to patient privacy. A third article that had used Surgisphere's data appeared only as a preprint and was subsequently removed from the repository where it had been deposited.

However, the procedure is not always so straightforward, and not always the authors agree to retract the article. Sometimes the journal does not agree and it is usually difficult to obtain a retraction, especially with the most powerful publishers and the most prestigious journals. Cases of fraud can affect all publishers, large and small, and all journals, and the reluctance to retract articles is not only due to the fear of looking bad or of damaging the prestige of the journal. By now, sensational cases of retraction of articles, for example by a journal such as The Lancet, have had no effect other than to highlight the journal's difficulty in admitting cases of fraud[60]. Nor does the historical fact act in its favour that The Lancet, a journal founded in 1823 by Thomas Wakley, "built its reputation on the abusive and serial piracy of medical lectures, and frequently had to defend itself in court for its actions"[61]. The reasons for journals' reluctance to retract are more likely to lie in citation calculations, as sometimes retracted articles, precisely because they present untrue and therefore surprising innovations and discoveries, are highly cited and contribute to the journal's impact factor. Retraction does not result in removal from the citation databases and does not block the possibility of citing the article. Although it makes sense to keep the retracted article in the journal in order not to create bibliographic ghosts, and in the citation databases for sociological analyses and bibliometric studies, their counting in the indices is one of the critical points. However, it should be noted that it is not always easy to establish fraudulent intent with absolute certainty. Consequently, sometimes the journal only communicates an 'expression of concern', perhaps waiting to obtain more information to support a substantiated choice before issuing an accusation of misconduct and proceeding with a retraction, which could have consequences for the researcher's career.

Sometimes such sensational cases emerge where the decision is easy and incontrovertible. An example of this is the 'Star Trek case', concerning the pediatrician and cardiologist Victor Grech, the author of 113 articles in Elsevier's journal Early Human Development (EHD), 19 of which dealt with the medical portrayals in the TV series Star Trek and dwelt on how doctors are portrayed in the series. An entirely legitimate approach would be the case if not for the fact that the particular slant of the articles was not stated, and the articles were classified by the journal as 'best-practice guidelines'. In December 2020, a student who noticed the anomaly reported it to the journal. The editorial board condemned

60 F.C. Fang, A.M. Stern, A. Casadevall, "*Misconduct accounts for the majority of retracted scientific publications"*, cit.

61 A. Johns, *Piracy*, cit. p. 257-258.

this explicitly misleading attitude, but by March 2021, only one of these articles had been retracted[62]. Elsevier withdrew as many as 26 of Grech's articles in Covid-19, in which the cardiologist dealt with crucial topics such as vaccines and mortality caused by the virus. Victor Grech is nevertheless part of the list of thirty authors with the highest number of retractions in the Retraction Watch database[63]. The list also includes an Italian name, a former editor in chief of the journal IEEE Transactions on Electromagnetic Compatibility from which 26 of his articles were retracted on a single day in 2018[64]. On that occasion, IEEE had retracted a total of 29 articles from various journals to which a further 49 articles were added in 2019, on the generic grounds of a violation of peer review procedures, but without adding further details and without responding to Retraction Watch's requests for further clarification.

Another of the many cases that have emerged since the start of the pandemic involved the resignation of several members of the scientific committee of the MDPI publisher's (a rather controversial publisher) journal Vaccines. It was an article that, just like Robert Wakefield's now famous fraudulent article[65], hypothesised a correlation between the Covid vaccine and the mortality of vaccinated people. In practice, it indiscriminately linked all deaths in the analysis to the vaccine, without pausing to verify the actual causes. The case is emblematic because it is an article on vaccines written in a journal specialising in the subject, but, as was noted in a comment in Science[66] , none of the three coauthors of the article has a background in vaccinology, virology or epidemiology. The authors are a psychologist and science historian, a physicist studying the dietary aspects of anticancer treatments, and an independent data analyst. Therefore, the first question to ask is therefore how it is possible that an article in the medical field submitted to the journal by authors without training or experience in this field was accepted and passed peer review. The three reviewers (two of whom were

62 Adam Marcus, "Beam us up! Elsevier pulls 26 Covid-19 papers by researcher with a penchant for Star Trek", *Retraction Watch*, March 31 2021, https://retractionwatch.com/2021/03/31/beam-us-up-elsevier-pulls-26-covid-19-papers-by-researcher-with-a-penchant-for-star-trek/.

63 *The Retraction Watch Leaderboard* https://retractionwatch.com/the-retraction-watch-leaderboard/.

64 Ivan Oransky, "Engineering prof in Italy earns 26 retractions in one fell swoop", *Retraction Watch* November 22, 2018, https://retractionwatch.com/2018/11/22/engineering-prof-in-italy-earns-26-retractions-in-one-fell-swoop/.

65 I. Oransky, *"Andrew Wakefield's fraudulent paper on vaccines and autism has been cited more than a thousand times",* cit.

66 Meredith Wadman, "Scientists quit journal board, protesting 'grossly irresponsible' study claiming COVID-19 vaccines kill", *Science* Jul. 1, 2021, https://www.sciencemag.org/news/2021/07/scientists-quit-journal-board-protesting-grossly-irresponsible-study-claiming-covid-19.

anonymous) published open reviews that even endorsed the method and scientificity of the article, allowing it to be published. The chain of resignations began the day after the article came out, which demonstrates its blatant fraudulent nature, which was immediately evident as soon as it reached the relevant scientific community through publication. Some suspicion as to how the peer review was conducted arises by analysing the dates on the article: the article was received by the journal on 2 June, reviewed on 19 June, accepted on 21 June and published on 24 June. Although there is no doubt that during the pandemic, peer review was shortened, it is indeed an anomalous speed sequence. The article was retracted relatively quickly, but in the meantime more than 350,000 people have had access to the article to read it, quote it, share it on social networks and use it to support no-vax theories. The problem is that the article should not have been published. Withdrawing it after publication is now ineffective in stopping its spread with the viral timing and modalities of the social web.

The motivations behind fraudulent actions have also evolved over time. In Babbage's time, when, moreover, fraud was more difficult because the circle of scientists in a given field was small, the goal was the recognition of scientific authority by peers and perhaps the public, admission to scientific academies, and, in general, personal affirmation. In the 20th century, the transformation of science and of the university system brought about new goals for the researcher, such as a university professorship, prizes, and rewards from associations, power in the scientific sphere but also in the political sphere since teachers and researchers are often called upon to hold political office or to collaborate in political bodies, with the consequent additional economic incentive. Finally, in all epochs, but perhaps more accentuated in the digital age, hovers the benefit exposed by physicist David Godstein in 2010, namely the immortality that scientific fame seems to guarantee[67]. This aspiration can be as much a cause of misconduct as the psychological motivations provided in 1959 by the American neurologist and psychiatrist Lawrence Schlesinger Kubie, who identified the basis of fraudulent behaviour in science as 'neurosis, unresolved childhood conflicts, and an unrealistic image of the profession', as well as 'poor preparation for the emotional pressures of the job'[68].

67 Enrico Bucci, *Cattivi scienziati: la pandemia della malascienza*, Torino, Add, 2020, p. 63.

68 *Ibid,* p. 56

5. Three exemplary problems

In order to delve into the issues of fraud, forgery, and plagiarism, we will briefly focus on some of the most frequent causes of retraction, as indicated in the Retraction Watch database. In particular, we will analyse the three main causes of publication retraction relating to Italy (of which we will report and comment on the national data specifically in a later section), i.e. data manipulation and p-hacking, image manipulation, plagiarism. On a scale of the severity of maliciousness, as defined by Lüscher,[69] they are positioned at different levels (Fig. 3), but are nevertheless problematic, not least because they are sometimes difficult to detect and have very blurred boundaries. Whereas fabrication and plagiarism, once discovered, tend to be easier to prove objectively, manipulation of data and images is much less so. So, the cases under the surface of the iceberg are probably very numerous.

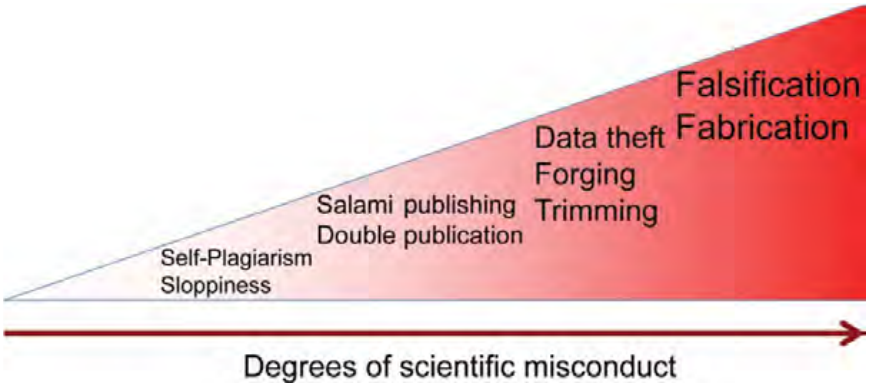


Figure 3 (Lüscher, 2013)

Data Manipulation and P-Hacking

The importance of having accurate and transparently processed data is one of the elements underlying the reproducibility of science. The question of data, like that of reproducibility, although alive for many centuries, has taken on a different dimension since the 1990s with the

69 T. F. Lüscher, *The codex of science: honesty, precision, and truth-and its violations*, cit.

emergence and spread of digital technologies. On the one hand, experiments, data collection, and analysis carried out with technologies became more complex and quantitatively relevant. On the other hand, the same complexity in data processing makes the proof of reproducibility both more difficult and crucial. Data are subject to the same types of research integrity violations, typically fabrication, falsification, and plagiarism, and to the same ethical problems as publications. In fact, it is often the data themselves that are subject to violations in scientific articles.

We have seen some examples of data falsification in the history of science, so let us turn to another recurring case of data misuse, p-hacking, also called data dredging or inference. This is a specific type of manipulation aimed at shaping the data in such a way as to construct statistics that can support the author's initial thesis and lead to a P-value. The p-value indicates in statistics the value below which the data obtained from a statistical test lead to the rejection of the null hypothesis. According to the Wikipedia definition, "it is the probability, for a hypothesis assumed to be true (null hypothesis), of obtaining results that are equally or less compatible than those observed during the test with the said hypothesis. In other words, the p-value helps to understand whether the difference between the observed and the hypothesised result is due to the randomness introduced by the sampling or whether this difference is statistically significant, i.e. difficult to explain by the randomness due to the sampling".

The p-value was theorised by statistician Ronald Fisher in the 1920s as part of his studies of crop variations in the Hertfordshire area of England. The theory was proposed to allow scientists to arrive at research results that were not too affected by noise, but in fact today it is sometimes used, in contrast, to precisely publish the noise part, the insignificant data, of the research[70] . The p-value is determined in advance and is usually set between 0 and 1, and if it is very small (usually less than 0.05), the null hypothesis is rejected. In p-hacking, there is no falsification of the data, but there is manipulation, as bias occurs in the choice of the data, with different connotations and consequences depending on the stages of the research in which it occurs. It may occur at the stage of data collection or at the stage of data processing or presentation of results and take different forms. Practices of p-hacking can consist, for instance, of collecting or selecting only positive data in an experiment or demonstration of a theory, avoiding collecting contrary data or not considering them in the processing, or collecting only data with $p<0.05$; or choosing to present in a publication only positive results among the several available after analysis, i.e., only the data with

70 Andrew Gelman, Eric Loken, "The Statistical Crisis in Science", *American Scientist* 102 (2014) 6, https://doi.org/10.1511/2014.111.460.

$p<0.05$; in stopping data collection when the value is $p<0.05$; and again disregarding obvious biases to obtain $p<0.05$ or in deliberately excluding participants in an experiment or statistical survey to obtain $p<0.05$. Gathering or analysing only the data that are useful to argue one's own theory or position, or to arrive at certain desired results in an experiment, i.e. selecting only the acceptable fruits in terms of data, is called *cherry-picking*. Finally, when one performs tests many times, somewhat randomly, until one obtains a certain unexpected result but instead declares that this result was the initial goal of the research, one speaks of a form of misconduct called HARKing (Hypothesising After the Results are Known), while the random search for results is defined as a '*fishing expedition*'[71].

In addition to these cases, there is another type of manipulation called the 'file drawer effect', or publication bias, which consists in the choice of many researchers not to publish results that are insignificant or negative, or that would be less popular, thus selecting the 'drawer' in which to file (or bin) them according to their potential to have an impact. The phenomenon is complex and multifaceted. Today's science generally pushes toward the publication of only positive results because they are better received and get more citations and because they attract popularity, including media attention, especially when dealing with critical topics such as pandemic issues. However, the researcher's choice is undoubtedly conditioned by the fact that journals are less likely to accept publications that report negative or insignificant results, for the same reasons, i.e. less attractiveness in terms of citations and impact. But science also needs negative and insignificant results to progress. If we return to the theories of Popper, for whom scientific error contributes to the progress of science, it is clear that it is indispensable that the error, the experiment that leads to a negative result, be shared with peers through publication. The prevalence of positive results biases scientists who rely on those studies and creates a collective prejudice (bias), as well as providing an unrepresentative and untruthful picture of the state of knowledge in a field. Clinical trials, to take one example, need negative data as much as positive ones[72]. The omission

71 Chittaranjan Andrade, "HARKing, Cherry-Picking, P-Hacking, Fishing Expeditions, and Data Dredging and Mining as Questionable Research Practices", *Journal of Clinical Psychiatry,* 82 (2021) 1, 20f13804. https://doi.org/10.4088/JCP.20f13804. PMID: 33999541; S. Ritchie, *Science Fictions*, cit, p. 81-121; N.L. Kerr, "HARKing: hypothesizing after the results are known", *Personality and Social Psychology Review* 2 (1998) 3, p. 196-217, https://doi.org/10.1207/s15327957pspr0203_4. PMID: 15647155.

72 Jérôme Adda, Christian Decker, Marco Ottaviani, "P-hacking in clinical trials and how incentives shape the distribution of results across phases", *PNAS*, 117 (2020) 24, 13386-13392; first published June 2, 2020; https://doi.org/10.1073/pnas.1919906117.

of negative results, or contrary data, in a trial, theory, or experiment leads subsequent studies to start from an unfounded and misleading basis[73]. Of course, this is the case when the p-value is used improperly and with misleading intent (p-hacking), but it must be emphasised that the exclusion of data in statistical analyses can, on the other hand, have scientific foundation in some cases[74]. Sometimes, only the analytical dataset necessary to make the research in question reproducible in specific terms is provided in the publication, and not the entirety of the data or data irrelevant to the research objective are excluded. Finally, the error that may lead to p-hacking may also be unintentional but due to the inadequate mastery of statistical methods and data management techniques. A self-assessment of one's ability to work with data analysis and statistical techniques (possibly to be perfected by appropriate training), as well as a clear a priori definition of the research objectives and the methods to be used to achieve them, are useful measures to avoid deviations and errors in data processing.

Image manipulation

The problem of image manipulation is on the rise due to the possibilities offered by digital technologies, but even for this kind of fraudulent behaviour, we are not dealing with anything new. The cases described by Babbage included examples of image manipulation. Until the invention of photography in the 19th century, images accompanying scientific publications consisted of drawings that could already be falsified, forged, and plagiarised, like any other scientific data. Even in the predigital era, fakes sometimes had a great impact on society and the collective imagination, as in Foersch's tale of the tree of death and its various literary and pictorial representations, or in the detailed drawing of the alleged mollusc discovered by Giuseppe Gioeni. The image helps to make a piece of writing credible, and therefore a forgery plausible, so that it could (and can) itself be the object of forgery.

The birth of analogue photography facilitated forgery long before digital technologies. The case of the official photograph portraying Nikolai Yezhov, head of the Soviet secret police, next to Joseph Stalin and another officer is well known. After the years of the Great Terror, which saw Yezhov's arrest and execution, it was altered by removing the figure of the officer[75]. In the scientific sphere, in 1961 the journal

73 Megan L. Head, Luke Holman, Rob Lanfear, Andrew T. Kahn, Michael D. Jennions, "The Extent and Consequences of P-Hacking in Science", *PLOS Biology* Published: March 13, 2015, https://doi.org/10.1371/journal.pbio.1002106.

74 Brad Verhulst, "In Defense of P Values", *AANA Journal* 84 (2016) 5, p. 305-308, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5375179/.

75 The website *Altered Images: 150 years of posed and manipulated documentary photography* shows this and numerous other cases of altered analogue images

Science publicly apologised for publishing an article by an Indian researcher in which he claimed to have discovered a parasite in chicken eggs and identified it as a possible cause of toxoplasmosis. The proof of this discovery was a microscope photograph that turned out to be false, as it was the detail of another image simply enlarged and rotated horizontally[76].

However, with analogue photography, an expert photographer was still needed to make a sophisticated forgery, and the operation was quite complex. With digital photography, anyone can access software, such as Photoshop or many similar ones, even open source, that can alter the original photograph. Photo-editing applications are now available on all mobile phones. The falsification of a scientific photograph certainly requires more care, but the tools to do so are now within everyone's reach. In fact, cases of image manipulation, even in the scientific field, are on the rise, and the examples are numerous. We have described the case of Haruko Obokata. Another equally sensational episode was that of Woo-Suk Hwang in 2004. Once again, it was an article published in an authoritative journal such as Science in which Hwang claimed to have cloned human embryos[77]. The publication had great resonance and in his home country, South Korea, the scientist became a kind of national hero hailed by the media, receiving large amounts of funding for his research. Until it was discovered that the images that were supposed to prove the existence of the cloned embryos were fake, and in a chain of events, other ethical violations and fraudulent behaviour emerged in the use of the data of the women involved in the experiments. Additionally, it emerged that the funds obtained for the investigation had been diverted to a personal account and used to support political figures and private purchases, such as a new car for his wife. As soon as it had hailed him, the press stigmatised him. Moreover, Hwang was dismissed from the university where he worked and was criminally convicted even though he managed to avoid prison on a technicality. Yet even today, he still continues to research cloning, albeit at a less prestigious university.

The cases exposed, including those of Obokata, concern prominent journals. The weakness of the article review process, which is often unable to detect dubious cases, is further aggravated by the objective difficulty of determining the manipulation of a digital image, especially with the naked eye. The same forensic science that deals with the different ways in which an image can be falsified, for obvious reasons related to the crucial role an image can play as judicial evidence, highlights

throughout history: http://www.alteredimagesbdc.org/stalin.

76 S. Ritchie, *Science Fictions,* cit., p. 55.

77 *Ibid,* p. 55-58.

how complex it is to establish whether an image has been manipulated. The most recent research identifies five categories of alteration[78]:

Object transfer: consists of transferring an object or an area of an image to another photo or moving it within the same image. It is the most common type of forgery and can be carried out by mounting or copy-and-paste operations.

Inserting or manipulating objects: Inserting false parts into an image or altering objects in an image to change their properties.

Lighting: altering the light and lighting aspects in an image.

Delete: Remove an object or an area of the image to hide it.

Enhancements and retouching: this is the most generic category and includes all kinds of retouching compared to the original image, with changes to objects, colours, perspective, and more.

Such forms of manipulation are not always illegal. They are not, for example, when they are used for artistic, creative, or recreational purposes. However, they take on negative connotations when the manipulation is not declared and is carried out with the aim of making an image that is not real appear real, to support a theory, a discovery, or the content of a scientific article. The association of scientific, technical and medical publishers, STM (International Association of Scientific, Technical and Medical Publishers), is currently working on a project to develop guidelines on the manipulation of images in scientific articles, in which it identifies three degrees of seriousness classified on the basis of the extent and relevance of the aberration, the intentionality of the alteration, and thus the presence of evidence to refute voluntariness or, conversely, the unintentionality and genuine error in the deviation detected, and finally the importance of the data affected by the manipulation with respect to the final result presented in the article[79]. Each of the criteria is then also calibrated by verifying the possible repetition of integrity violations by the author and their severity.

Image forgery has become one of the most widespread forms of scientific fraud. The Office of Research Integrity (ORI) of the United States stated in 2016 that it represented more than 70% of the cases handled, although it is not possible to accuse the perpetrators of forgery with certainty, as it is difficult to prove the voluntariness of the action[80]. If

78 Victor Schetinger et al., "Image forgery detection confronts image composition", *Computers & Graphics* 68 (2017), https://doi.org/10.1016/j.cag.2017.08.014.

79 STM Working Group on Image Alteration and Duplication Detection, *Recommendations for handling image integrity issues,* Final draft, July 2021, https://osf.io/kgyc6/.

80 Miriam Shuchman, "False images top form of scientific misconduct", *CMAJ: Canadian Medical Association Journal* 188 (2016) 9, p. 645, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4902687/.

spotting a counterfeit image is not easy for forensic experts and it is not possible in many cases to determine whether the manipulation was deliberate, it is even less so for those involved in the peer review of an article submitted to a scientific journal for publication. For this reason, specialised figures in analysing images to discover manipulations and duplications are emerging, to whom journal editors, publishers, research institutes, and individual researchers turn when in doubt[81]. Several artificial intelligence algorithms are also tried out by journals to find cases of image manipulation[82]. Artificial intelligence offers an important aid in finding suspicious cases, but it is also a sophisticated technique for manipulating images, the product of which is often difficult to detect even with other algorithms.

Plagiarism

Plagiarism can affect any aspect of scientific research, the written text of articles, data sets, images, ideas, and, of course, not only the scientific field. Plagiarism is also not a recent phenomenon. A famous case of plagiarism of ideas is that of Giuseppe Meucci, one of the Italian inventors whose inventions were not finally realised in Italy. He is credited with the idea behind the telephone, in its first version called the telettrophone. In 1871 Meucci founded a company to study its development, but was forced to close it down within a year because he found no entrepreneurs willing to support it financially and develop the invention. On 28 December 1871, Meucci filed Caveat No. 3335 entitled *Sound Telegraph,* in which he described his invention at the United States Patent and Trademark Office in Washington. The caveat was a general prior art request, in force at the time, that an inventor could file to show his intention to patent an invention. To obtain the $250 needed to file a regular patent, Meucci tried to find sponsors, but failed. The caveat expired after one year, and so the invention became unencumbered and was taken up and patented by Alexander Graham Bell, a British engineer naturalised in the USA, who with a string of industrialists behind him succeeded in perfecting the invention and patent-

81 Helen Shen, "Meet this super-spotter of duplicated images in science papers", *Nature* 13 May 2020, https://www.Nature.com/articles/d41586-020-01363-z ; Richard Van Noorden, "Publishers launch joint effort to tackle altered images in research papers", *Nature* 13 May 2020, https://www.Nature.com/articles/d41586-020-01410-9.

82 Jinjin Gu, Xinlei Wang, Chenang Li, Junhua Zhao, Weijin Fu, Gaoqi Liang, Jing Qiu, "AI-enabled image fraud in scientific publications", *Patterns* 3 (7) 2022, https://doi.org/10.1016/j.patter.2022.100511; Richard Van Noorden, "Journals adopt AI to spot duplicated images in manuscripts", *Nature News* 21 December 2021, https://www.Nature.com/articles/d41586-021-03807-6; Thorsten Stephan Beck, "Image manipulation in scholarly publications: are there ways to an automated solution?" *Journal of Documentation* 78 (5) 2022.

ing it under his own name. Meucci filed a lawsuit for fraud against Bell, which he probably would have won had he not died in 1889, before it was completed.

Many other examples could be cited, but we have amply highlighted how widespread the phenomenon of piracy was since the 17th century and how, in fact, it often consisted of plagiarism. The development of certain reproduction technologies favoured plagiarism over time. For instance, in the 19th century, the invention of photolithography, which made it possible to create copies identical to the originals, some of which 'could only be distinguished from genuine articles by an expert' since 'the quality of the paper might differ, but, in general, in terms of readability, a pirated reproduction was impeccable'[83] . In the publishing world, this was happening well before digital but again the difference is that today it is within everyone's reach.

The sticking point for defining plagiarism remains the intentionality or unintentionality of the action, as defined by LaFollette, which makes it possible to speak of two types of plagiarism, intentional and unintentional, although it is difficult to distinguish them. It should be emphasised that plagiarism does not necessarily imply copyright infringement, since the two concepts and the related actions, are different. Plagiarism consists of using the words, data, and images of other authors without declaring it by quotation or bibliographic reference. For example, if you quote a sentence from a work in the public domain without citing the source, you commit plagiarism, but not copyright infringement. Plagiarism is an ethical violation, the disregard of tacit, or even written rules within codes of conduct and research integrity regulations formally issued by universities or other institutions, which form the basis of scientific communication (it does not only concern the scientific sphere of course, but this is our spectrum of interest). Copyright is regulated by national laws and therefore infringement is an action that has legal consequences. Copyright infringement may be related to the unauthorised reproduction of a work, its transposition from one format to another, transmedia transposition, public representation, etc. In scholarly communication, copyright is usually transferred from the author to the publisher when publishing an article or book, and infringement concerns the publisher in the first place. Large publishers are indeed diligent in suing for copyright infringement, as we have seen, for instance, in the case of shadow libraries. However, they are not as diligent in recognising and declaring cases of plagiarism.

Publishers are very slow to acknowledge allegations of plagiarism and act. This is undoubtedly partly due to the difficulty of ascertaining cases of intentional plagiarism with the utmost certainty, but in general the reluctance is also due to the consequences on the authorita-

83 A. Johns, *Piracy,* cit., 330.

tiveness and prestige of the journal. The consequences can be severe, especially in extreme cases such as the one reported in connection with the "*Annals of Internal Medicine* of the American College of Physicians", which was the subject of data plagiarism by one of the reviewers who conducted the peer review[84] . Some publishers use alternative formulas instead of explicitly stating that an article has been retracted for plagiarism. In the humanities, publishers have been found to be even more reluctant to retract an article or declare plagiarism than STM publishers, as the harm resulting from plagiarism is perceived to be less than in other fields such as medicine[85] .

Plagiarism is among the most frequent cases of retraction surveyed by Retraction Watch and now takes many extreme forms and variants; it may concern parts of the text or the entire article. There are cases where articles are retracted because they were accused of plagiarism of articles that themselves had already been accused of plagiarism, as happened in the field of astrophysics[86] . Duplication of articles in journals is also on the rise. Duplication occurs when the same article is submitted to several journals with minor variations or whole paragraphs that are similar without stating it. This phenomenon, which is called self-plagiarism, since it is in fact an author 'plagiarising' himself from one article to another, is starting to worry journal editors themselves, who wonder how to proceed in these cases that are increasingly being discovered. The discovery may occur during the author's submission of the article with the help of anti-plagiarism software[87] or after publication, especially if the articles are open access. It is therefore necessary to define common procedures to act, on which publishers are wondering. Should both articles be withdrawn? Should one notify the other journal even if one finds out and decides not to publish the article? The publishers' association COPE (Committee on Publication Ethics), in collaboration with BioMed Central, has drawn up guidelines, specifically for dealing with recycled texts, addressed to journal editors, in which

84 Adam Marcus, "Researcher who stole manuscript during peer review earns second retraction", *Retraction Watch* August 1 2017, https://retractionwatch.com/2017/08/01/researcher-stole-manuscript-peer-review-earns-second-retraction/.

85 Alison McCook, "Does the philosophy literature have a plagiarism problem?", *Retraction Watch* May 19 2017, https://retractionwatch.com/2017/05/19/philosophy-literature-plagiarism-problem/.

86 Ivan Oransky, "Astrophysics retraction trail includes paper that plagiarised another already retracted for...plagiarism", *Retraction Watch* May 18, 2012, https://retractionwatch.com/2012/05/18/astrophysics-retraction-trail-includes-paper-that-plagiarized-another-already-retracted-for-plagiarism/.

87 *Like CrossCheck* https://crosscheck.ieee.org/crosscheck/ ; *iThenticate* http://www.ithenticate.com/; *Turnitin* https://www.turnitin.com/solutions/plagiarism-prevention; also open source like *Plagiarism Detector* https://plagiarismdetector.net/; *Grammarly* https://www.grammarly.com/.

they propose to assess how much has actually been recycled and how much is new in the article, to check for duplication of data, and then to decide how to act[88] . Let it be clear that no one prevents a journal from republishing an article that has already been published, except for considerations of expediency and convenience to publish original content that may increase the reader's interest, so it is a common requirement that publishers include in the publishing agreement. Moreover, since the 17th century, there have been journals built as collections of articles already published elsewhere, which today take on the connotation of *overlay journals* and draw on preprint repositories as well as open and closed access trade journals, where granted. The critical issue lies in the concealment of this information by the author from the journal editor. From the authors' point of view, this may also be dictated by the need to publish extensively in order to meet the thresholds set by research assessment or career advancement procedures.

6. Repeatability, Replicability, and Reproducibility

In the biographical film *The Man Who Knew Infinity*, made in 2015 by American director Matt Brown and dedicated to the figure of Indian mathematician Srinivasa Aiyangar Ramanujan, one of the fathers of number theory, there is a significant sequence of scenes. Ramanujan, of humble origins but great talent, arrives from India at the prestigious Oxford University and immediately wants to publish his theories. Professor Hardy (played by Jeremy Irons) stops him and convinces Ramanujan to first find the demonstrations of his theorems. One of his proposed theorems is tested and fails. But this allows him to revise it and pick up new nuances and implications, with the result that he can publish an excellent paper and later become a member of the Royal Society. Ramanujan's theory subjected to the reproducibility test fails and must be further refined to be publicly presented to the scientific community in the form of a published article.

The possibility of reproducing the process and results of research is the very foundation of the scientific method. Robert Merton speaks of 'organised scepticism' as the suspension of judgement with respect to a scientific discovery or thesis until one is able to validate it through the verification of data and method. Karl Popper identifies the possibility of reproducing an experiment by other scientists as the very essence of scientific objectivity underlying the scientific method. The principle of falsifiability defines whether a hypothesis or a theory is scientific. Falsifiability must be distinguished from falsification. Falsifiability

88 COPE, *Text recycling guidelines for editors,* https://publicationethics.org/text-recycling-guidelines.

is 'the criterion for establishing the empirical character of a system of assertions', whereas a theory can only be considered falsified when a number of conditions are met:

> We say that a theory is only falsified if we have accepted basic assertions that contradict it. This condition is necessary but not sufficient; for we have seen that single, non-reproducible occurrences have no meaning for science. It is therefore difficult for a few unrelated basic assertions that contradict a theory to lead us to reject it as falsified. We will only consider it falsified if we discover a *reproducible effect* that disproves the theory[89] .

This objectivity, which is the basis of what he calls the 'game of science' and which determines scientific progress, is not an activity attributed to the individual researcher because science and scientific objectivity do not (and cannot) result from the efforts of a single scientist to be 'objective', but from the cooperation of many scientists. Objectivity can be called the intersubjectivity of the scientific method. The scientific method consists of expressing theories in a codified form that allows other scientists and "anyone who has mastered the technique of understanding and demonstrating scientific theories [to] repeat the experiment and judge for themselves"[90] .

The intrinsic value of reproducibility is proven by its affirmation as a fundamental criterion at the birth of modern science and by the review procedures conducted by the Royal Society, which would later find a public outlet in the journal Philosophical Transactions. In fact, the first famous case of controversy over reproducibility dates back to Newton's time. In 1672, Isaac Newton submitted a design for a new telescope to the Royal Society for peer review and the institution's endorsement. A short time later, he also sent a letter explaining his theory on the refraction of light and colours. Robert Hooke, tasked with verifying and reproducing Newton's experiments, failed to do so and disputed them, giving rise to a dispute that later extended to other topics.[91] A few years after Newton's affair, in 1680, the Irish chemist and naturalist philosopher Robert Boyle published *The aerial noctiluca,* considered to be the first work in which all the data from his experiments were reported in detail to facilitate their reproducibility.

Therefore, the issue of reproducibility is not a recent topic, but its profound implications have not been addressed for a long time. As with many other issues we have dealt with, the non-reproducibility of scientific research can be perceived as an act intended to undermine the trust

89 K. Popper, *Logica della scoperta scientifica*, cit., p. 76-77.

90 K. Popper, *La società aperta e i suoi nemici,* cit., p. 499-500.

91 See details in A. Johns, *Piracy,* cit., p. 64-71.

that underlies and governs relations between scientists. And, as with other cases of maliciousness, it must be pointed out that the non-reproducibility of research is not always linked to fraudulent intent, as it may be due to carelessness and inaccurate data collection or processing, and it is not always easy to attribute malicious intent with certainty. In any case, since the early 2000s, there has been an increase in the number of cases of articles with non-reproducible results, as a probable consequence, firstly, of the increased availability of publications through digital and open access formats and thus an increased possibility of peer review. A study a few years ago on reproducibility problems identified, in the first place, the frequently incomplete account of the experiment, of which perhaps only the main lines are drawn in the article, making it difficult to reproduce it; in the second place, the pressure to publish, followed by a lack of statistical competence and analytical skills, and then various other causes[92] .

Stuart Richtie, professor of psychology at King's College London, reported on an emblematic case that concerned him personally[93] . In 2011, the psychologist Daryl J. Bem, then professor at Cornell University, published an article entitled "*Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect" in* a leading peer-reviewed journal, the Journal of Personality and Social Psychology. The article reported on image perception experiments conducted on students that demonstrated the possibility, under certain circumstances, of predicting the images they would later choose to see. The hypothesis aroused many doubts, starting with Ritchie himself, who with some colleagues tried to reproduce the experiment without success. The details are recounted in the cited volume and can be found in other sources, as the case is fairly well known, but what we want to emphasise here is what happened next. Ritchie and the co-authors of the counter experiment wrote an article in which they demonstrated that the experiment had not produced the results described by Bem at all, and indeed had produced no significant results, and sent it to the same journal in which Bem's article had appeared. The article was rejected with a letter of reply from the editor explaining their policy of never publishing articles that repeated, either positively or negatively, experiments that had already been published. The article later appeared in another journal, but the editor's attitude is symptomatic of a general problem.

92 Monya Baker, "1,500 scientists lift the lid on reproducibility", *Nature*, 533 (26 May 2016) p. 452-454, https://doi.org/10.1038/533452a.

93 S. Ritchie, *Science Fictions,* cit.

The field of psychology is among the most studied in relation to reproducibility, probably due to the specific interest in such behaviour[94] . Ritchie reports several case studies in the field, with reproducibility rates ranging from 38% to 77%[95] . The phenomenon also affects other subject areas, as the same author reports and as can be seen from the growing studies. Taken together, these behaviours are the cause of the so-called reproducibility crisis that affects science today and undermines its integrity[96] . Although we usually speak generically of reproducibility, we can, in fact, find three different situations that fall under the definitions of repeatability, replicability, and reproducibility. *Repeatability* is the reproduction of an experiment conducted by the same original researchers using the same data and the same analysis procedure to verify that the same results are obtained. In theory, it is a verification that should always be followed, with several attempts, before the results of a research are considered valid and publishable. *Reproducibility* involves different groups of researchers using different experimental procedures repeated several times with independently collected data and arriving at the same results as the original. *Reproducibility, in the strictest sense of the word,* is when groups of researchers other than those who conducted the original experiment repeat the same analysis with the same data as the original and arrive at the same results[97] .

According to the US National Science Foundation, reproducibility refers to the condition by which it is possible to achieve the same results as another researcher using existing data from the previous study and is the minimum condition for research to be credible and informative, as opposed to replicability, which instead implies collecting new data to check whether the same results as in a previous study can be achieved[98] . The National Science Foundation's guidelines therefore recommend transparency in methods, the adoption of policies for open data with their deposit in repositories where they can be publicly available, and the allocation of a portion of a research project's resources for data

94 See also the Center for Open Science's significant study in the field of psychology, conducted from 2011 to 2014: Open Science Collaboration, "Estimating the reproducibility of psychological science", *Science* 349 (2015) 6251, https://doi.org/10.1126/science.aac4716.

95 S. Ritchie, *Science Fictions,* cit., p. 31-34.

96 Monya Baker, "*1,500 scientists lift the lid on reproducibility*", cit.

97 Evanthia Kaimaklioti Samota, Robert P. Davey, "Knowledge and Attitudes Among Life Scientists Toward Reproducibility Within Journal Articles: A Research Survey", *Frontiers in Research Metrics and Analytics*, 29 June 2021, https://doi.org/10.3389/frma.2021.678554.

98 The National Science Foundation - The Institute of Education Sciences, U.S. Department of Education, *Companion Guidelines on Replication & Reproducibility in Education Research*, Nov 28, 2018, https://www.nsf.gov/pubs/2019/nsf19022/nsf19022.pdf.

documentation, management, and sharing activities. These principles are also the basis of the *Transparency and Openness Promotion Guidelines (TOP)*, the guidelines of the Centre for Open Science, a non-profit association founded in 2013, to which several thousand journals and scientific societies have already adhered[99] .

Indices have also been proposed to assess the reproducibility of research. These include the R-Factor in the biomedical field. The R factor would measure the number of times an attempt was made to reproduce the research whose data and results were presented in an article. The index is set equal to zero for the article just published and then related to the number of times the research has been attempted to be reproduced, so if ten studies try and all ten succeed, the R-Factor is 1, while if of these ten only two succeed, the R-Factor is 0.2 (2/10)[100] . Such a measure, in the intentions of the proponents, would allow a reproducibility check to discourage attempts at scientific fraud and the publication of research without any prior verification. The limitation of such an index is that it must be calculated from the citation indices, and thus it is mainly the producers of those indices themselves, Elsevier and WoS, who can work out a possible R-Factor.

Some studies on reproducibility provide significant data on the extent of the problem. A study from a few years ago, reported on the blog of the journal Nature, refers to data provided by the well-known pharmaceutical company Bayer according to which at least 50% of research published by universities proves not to be reproducible[101] . This implies that the company must either re-verify all data, or rely on research conducted in-house. The consequences are extremely heavy for universities, especially for certain scientific fields, which have always had collaborative relationships with companies and nurtured industrial and applied research. The credibility of academic scientific research is severely undermined. The type of evaluation carried out in a company such as Bayer, which focusses on the ability of research to provide a solid basis for the creation of a medicine, indirectly shows the weakness of quantitative evaluation systems, which are based on the container rather than the content, and the risks of the trends currently visible in academic science. Publishing falsified or incorrect data also implies a

99 *Transparency and Openness Promotion Guidelines (TOP),* https://www.cos.io/our-services/top-guidelines.

100 Joshua Nicholson, Yuri Lazebnik, "The R-Factor: A Measure of Scientific Veracity", *The Winnower*, August 17, 2014, https://doi.org/10.15200/winn.140832.20404 ; attempts to apply it also in psychology are collected on the Replication Index site https://replicationindex.com/.

101 Brian Owens, "Reliability of "new drug target" claims called into question", *Nature News blog* 05 Sept 2011, http://blogs.Nature.com/news/2011/09/reliability_of_new_drug_target.html.

cost in terms of public funds. Not only are funds spent on research that proves to be unreproducible and, therefore, in some way unreliable and unworkable, but other researchers will spend resources to continue that research, to extend the field of interest from erroneous data, or even to attempt to replicate experiments.

Reproducibility is also a sensitive issue and should be approached with due caution, as some cases in history have shown. For example, in the 19th century, that of the chemist John Dalton, who formulated a theory that no one was able to reproduce for a long time, until recent times when modern techniques made it possible to understand its scope and reliability, so much so that the vision defect called daltonism is named after his studies. Also well known is the aforementioned controversy between Mendel and Fisher, which largely relied on the fact that the latter had not correctly understood the data collected by the former[102] .

102 Walter W. Piegorsch, "Fisher's Contributions to Genetics and Heredity, with Special Emphasis on the Gregor Mendel Controversy", *Biometrics* 46 (1990) 4, p. 915-924, https://doi.org/10.2307/2532437.

Chapter 3
Predatory publishing and other issues

1. Predatory journals and ghost conferences

Predatory publishing is a phenomenon that affects all types of publication but predominantly journals (*predatory journals*). These are journals that claim to be scientific when in fact they are not, even though they can be deceptive on the surface due to the presence of ISSNs, the use of titles that mimic, and thus resemble authoritative journals, and their presence in databases, including citation databases. It is not easy to define predatory publishing precisely, as it is by no means easy to establish all its characters and delimit its contours, but a group of scholars from ten different countries tried and came up with the following definition of a predatory journal:

> Predatory journals and publishers are entities that prioritise self-interest at the expense of scholarship and are characterised by false or misleading information, deviation from best editorial and publication practices, a lack of transparency, and/or the use of aggressive and indiscriminate solicitation practices[1] .

The definition highlights the preeminent characteristics of predatory publishing such as deviating from publication practices while not declaring it and indeed providing untruthful information, and in particular with reference to the peer review that underpins the scientific nature of a journal, the lack of transparency, and an invasive way of proposing oneself to researchers, typically by sending emails soliciting the publication of an article, not infrequently even of one already published elsewhere. Within this general definition, case studies are numerous and often intertwine and overlap. For instance, predatory journals may declare fake editors, fake scientific committees, peer review processes that are not carried out, non-existent publication locations,

1 Agnes Grundniewicz et al., "Predatory journals: no definition, no defence", *Nature Comment* 11 December 2019, https://www.nature.com/articles/d41586-019-03759-y.

invented editors, and fictitious contributors. Sometimes they even use names of deceased persons, or use stolen identities, including names of persons on editorial boards without the knowledge of those concerned, and not infrequently publish pseudo-knowledge with false and fabricated data. A published study found that 4000 Australian researchers (or 7 percent of the country's total number of researchers) were included on the editorial boards of journals considered predatory, and found countless errors and confusing, even improbable data, such as the exchange of affiliation institutions (when indicated) between Australia and Austria, and vice versa[2] . In an attempt to avert these risks, Open Editors was created, a database containing data on the editorial boards of more than 6,000 journals from 17 publishers[3] . It is also not uncommon for these journals to charge authors to publish articles, with the mechanisms used for open access such as *article processing charges* (APC). The phenomenon of predatory publishing is frequently associated with open access, but is, in fact, not a direct consequence of it, even though it is one of the elements that facilitates its dissemination (but at the same time also its discovery). The practice of falsifying dates, names, and publishers in order to circulate unauthorised publications is at least as old as the invention of printing and, therefore, cannot be attributed to the spread of open access. But what changes not so much with open access as with the emergence of the Web is that all forms of intermediation between publisher and author are eliminated and it therefore becomes easier to operate directly.

Many predatory publications spread false science, but even in this case, one cannot generalise since it can happen that good articles are found even in predatory journals maybe because they are written by unaware authors. In fact, authors, especially researchers at the beginning of their careers and therefore inexperienced, may happen to run into such publications without realising it and thus hand over valid articles. The World Association of Medical Editors (WAME) distinguishes between predatory and pseudo-journals precisely on the basis of the author's awareness and voluntary choice[4] . According to WAME, if the choice to publish there is made by the author consciously, i.e. knowing that the journal did not meet the accepted standards of the scientific community, one cannot speak of 'predatory' and thus of predatory journals but one must speak of pseudo-journals. Knowledge is the discriminator between an error and a malicious action.

2 Mike Downes, "Thousands of Australian academics on the editorial boards of journals run by predatory publishers", *Learned Publishing* 33 (2020), published online 19 March 2020, https://doi.org/10.1002/leap.1297.

3 *Open Editors* https://openeditors.ooir.org/.

4 World Association of Medical Editors, *Identifying Predatory or Pseudo-Journals*, February 18, 2017, https://www.wame.org/identifying-predatory-or-pseudo-journals.

One of the common characteristics of such journals is to inundate potential authors with continuous email messages in which they propose to publish an article in a certain journal. A similar practice concerns predatory lectures, i.e. invitations to give a talk at a conference, perhaps for a fee, which may then turn out to be non-existent. Sometimes the offer is blatantly suspicious, especially when the message proposes to publish in a journal an article that has already appeared elsewhere or contains an invitation to a journal or conference that has nothing to do with the specialisation and research topics of the person to whom it is addressed. At other times, it is more difficult to distinguish them. Predatory conferences are on the increase, with similar characteristics to those of journals such as the absence of scientific committees and the charging of fees for participation, and with organisations that are specialising in this activity, now blatantly denounced in the scientific literature, such as OMICS and WASET[5] . What certainly does not help in identifying and surveying predatory journals and conferences is their unstable and precarious nature, as they are born and cease all the time, but with a birth rate greater than the cessation rate. Their number fluctuates continuously, reaching, according to the most recent analyses, peaks of two thousand titles[6] .

The functioning of a predatory journal was exemplified by Rick Anderson in an article in which he recounts having experienced such a situation himself[7] . Anderson received an email message inviting him to publish an article in a medical journal, the Journal of Cardiothoracic Surgery and Therapeutics. Anderson is a librarian at the University of Utah and has never even remotely dealt with topics related to health and medicine, which made him think of a predatory journal that had probably been swapped, perhaps between homonyms. However, to conduct an experiment, the librarian thought about responding to the proposal and accepted it. He then created an article by randomly taking a few paragraphs from an authoritative cardiothoracic medicine jour-

5 Tracey Elliott et al., "Perceptions on the prevalence and impact of predatory academic journals and conferences: A global survey of researchers", *Learned Publishing* 35 (2022) 4, p. 516-528, https://doi.org/10.1002/leap.1458; Emanuel Kulczycki, Marek Hołowiecki, Zehra Taşkin, Z., Güleda Doğan, G., "Questionable conferences and presenters from top-ranked universities", *Journal of Information Science*, preprint 10.31219/osf.io/guaw3 in OSF Preprints, April 04, 2022, Last edited June 29, 2022, Peer-reviewed Publication, https://doi.org/10.1177/01655515221087674; Diane Pecorari, "Predatory Conferences: What Are the Signs?" *Journal of Academic Ethics* 19 (2021), p. 343-361, https://doi.org/10.1007/s10805-021-09406-4.

6 Mike Downes, "Why we should have listened to Jeffrey Beall from the start", *Learned Publishing* 21 July 2020, https://doi.org/10.1002/leap.1316.

7 Rick Anderson, "Why Should We Worry about Predatory Journals? Here's One Reason", *The Source*, March 3, 2020 https://blog.cabells.com/2020/03/03/guest-post-why-should-we-worry-about-predatory-journals-heres-one-reason/.

nal, combining them, and giving the article a credible title, attributing it to completely invented authors. The article was sent to the journal, who replied to him announcing that it was proceeding with double-blind peer review and after a while informed him that the article had been accepted. However, the acceptance was accompanied by a request for APC's payment of USD 1100. Anderson then replied that they had never mentioned APC before. The journal retorted that it had been a staff error and explained the advantages of the business model that would allow everyone to access the article in open access. Anderson again replied that he could not support that amount and asked that the authors consider the article proposal withdrawn. Not caring about his request, the article was published in the journal anyway, accompanied by the DOI and all bibliographic data. Anderson linked it in his article of denunciation published in The Source (Figs. 1 and 2), and one can verify not only the obvious nonsense of what is contained therein (evident even to those who are not medical experts) but the blatant obviousness of the forgery, starting with the authors who answer to the names of Jackson S Pollock, Rosa Q Luxemburg, Hercule CM Poirot, and Friedrich X Engels.

Figure 1 *(Anderson, 2020)*

Vol 4 | Issue 1 | Pages 43-45

Journal of Cardiothoracic Surgery and Therapeutics

ISSN: 2643-5780

Case Report

DOI: 10.36959/582/420

Non-Invasive Laser Modalities in Treatment of Posterior Ventral Infarction

Jackson S Pollock[1], Rosa Q Luxemburg[1], Hercule CM Poirot[2] and Friedrich X Engels[3]*

[1]*Brockton State University, USA*

[2]*U.S. National Clinics of Health, USA*

[3]*Massapequa University, USA*

Background

Postoperative myocardial infarction is not a well-studied topic after lung surgery. From the few reports in the literature, the frequency to range between 0.7% to 2%, but it is difficult to determine an exact standardized incidence, in part due to differences in postoperative presentation, differences in EKG interpretation and a variation in the choice of serum cardiac markers.

circular shape, the CSA LVOT was calculated as: 3.14 × (D/2)², where D is the inner diameter of the left ventricular outflow tract. The V LVOT was obtained with pulsed-wave Doppler in the left ventricular outflow tract proximal to the aortic prosthesis from the apical five-chamber view; the V Trans prosthetic was obtained with continuous-wave Doppler from the apical five-chamber view [3]. The EOAi was calculated by the formula EOA/BSA (body surface area). LVM was derived from Devereux's formula and LVMI from LVM/BSA. The DVI was

Figure 2 *(Anderson, 2020)*

The volatility of the phenomenon has not prevented attempts to identify and list predatory journals. The first list, which had the merit of highlighting the problem, was compiled starting in 2008 by the American librarian Jeffrey Beall[8] , but was closed in 2017 due to pressure received by the librarian from various fronts and objections to certain weaknesses in his list. Beall had listed 1155 predatory publishers and 1294 predatory journals based on criteria partly derived from documents of associations such as COPE, WAME, DOAJ, OASPA (Open Access Scholarly Publishers Association)[9] . It should be noted that Beall's list always mentions predatory'suspects,' both to avoid repercussions but mainly because, as already mentioned, it is often indeed difficult to accuse journals of predatory practices with certainty. Beall's criteria can be found on the website of WAME, the association of medical editors and publishers, and they give an idea of the quantity and variety of suspicious behaviour, but also of the limitations of his analysis:

Criteria for the editorial responsibility of the journal

8 *Beall's list of Potential Predatory Journals and Publishers* https://beallslist.net/.

9 Updated and merged in COPE, DOAJ, OASPA, WAME, *Principles of Transparency and Best Practice in Scholarly Publishing*, January 15, 2018, http://wame.org/principles-of-transparency-and-best-practice-in-scholarly-publishing.

Some of the criteria mentioned in this section are important clues, such as the lack of transparency in the publishing operations, the absence of any indication of the editor or the editorial and scientific board, or that there are names without adequate experience and, expertise, and of course, that there is a fake editorial board, with people whose permission has not been sought (although this is difficult for an author to understand, a thorough analysis is needed). However, the other criteria of Beall are excessive. Although it is regrettable that the scientific board of a journal has no women or no geographical representation, it can certainly not be taken as an indicator of its predatory nature.

Economic and Publisher-Related Aspects

Here again, the observation that the journal presents insufficient information or hides information on the fees authors have to pay is a possible sign, but the lack of guidance on digital preservation practices or the protection of PDFs with systems that make anti-plagiarism control difficult is certainly not.

Integrity

Beall's warnings about titles are partially helpful, as predatory journals usually have titles that are incongruent with the stated mission or are very generic or resemble the titles of prestigious journals. Equally important are indications of the declaration of an impact factor that does not actually exist, or a bogus or predatory impact factor or indexing in databases and abstracting and indexing services (all easily verifiable data, by the way). It is equally significant that the publisher sends peer review requests to unqualified scholars or does not send them at all, and that it does not devote sufficient resources to preventing and eliminating misconduct, but in the first case it is not known to the author sending in his article, and in the second case it cannot be considered as a feature denoting a predatory journal.

More about the publisher

The clues listed by Beall with respect to the publisher are significant, but even then it sometimes takes in-depth investigation to uncover them: it republishes articles already published elsewhere without providing proper credit*;* it uses pompous language calling itself a 'market-leading publisher' even though it may be a fledgling or inexperienced organisation*; it* operates in a western country mainly for the purpose of functioning as a vanity press for scholars in a developing country (e.g., using a PO box or post office station in the US while actually operating from a developing country; it provides minimal or no editorial review and proofreading for submitted articles; it publishes

articles that are not scientific at all, written by non-expert people, polemical editorials or evident pseudoscience.

Other indications of poor quality and non-compliance with standards

Lack of information on the publisher is relevant, such as in cases where it does not provide sufficient contact information, including a lack of clear information on the administrative address or misleading information (e.g., through the use of addresses that are, in fact, general delivery). Whereas it is risky to attribute a predatory character to a journal because it retains the copyright on the content or demands its assignment to the author (not uncommon in all journals).

These are a number of clues that can be a useful point of reference but must be applied with caution. Certainly, one cannot assume that the presence of a single criterion is sufficient to define a journal or publisher as predatory, since many of the criteria listed by Beall are found in journals that are not predatory at all. In fact, it was precisely this one of the limitations of Beall's list and the subject of criticism. Furthermore, Beall's work has been accused of referring to commercial databases such as Web of Science as proof of journal reliability, whereas some predatory journals are present in both Web of Science and Scopus, as some articles have shown[10] . His perspective related to the commercial database system, combined with a lack of knowledge of the open access world, led him to make big mistakes, such as accusing the SciELO platform, a collaborative open science initiative born in Brazil and extended to many other countries in Latin America, of predatory attitude. Beall had called the SciELO platform 'an editorial favela" and noted its lack of care and visibility, confirmed, in his opinion, by the absence of data on the Web of Science. In contrast, the SciELO platform is sober because it

10 A specific article on predatory journals in Scopus was later retracted: Vít Macháček - Martin Srholec, "Predatory publishing in Scopus: evidence on cross-country differences", *Scientometrics* 126 (2021), p. 1897-1921, https://doi.org/10.1007/s11192-020-03852-4. The article was retracted for 'errors in analysis, method and unreliable results' as stated in the Retraction Watch database. But there are other studies on the subject such as: Tove Faber Frandsen, "Authors publishing repeatedly in predatory journals: An analysis of Scopus articles", *Learned Publishing* early view version 04 August 2022, https://doi.org/10.1002/leap.1489; Nguyen Minh Duc et al., "Predatory Open Access Journals are Indexed in Reputable Databases: a Revisiting Issue or an Unsolved Problem", *Medical Archives* 74 (2020) 4, p. 318-322, PMCID: PMC7520066, https://doi.org/10.5455/medarh.2020.74.318-322. The hope is that articles of this kind will not be taken as unsubstantiated allegations and buried, but will instead prompt even large publishers to take up the issue. Moreover, Elsevier claims to do so https://www.elsevier.com/connect/predatory-vs-trustworthy-journals-what-do-they-mean-for-the-integrity-of-science, also for Scopus https://www.elsevier.com/connect/the-guardians-of-scopus, as do now most of the large publishers and producers of digital scientific resources such as Springer https://www.springer.com/gp/editorial-policies/predatory-journals-and-references, Clarivate https://clarivate.com/blog/bealls-list-gone-but-not-lost/.

was created with few resources and precisely as a reaction to commercial databases. A blog post on the platform defines Beall's attitude as full of "prejudice, classism, imperialism, and boorish commercialism"[11] .

Based on Beall's pioneering work, however, other lists have subsequently sprung up, also in specific subject areas. *Scholarly Outlaws* offers a further list of publishers, suspect conferences, and other predatory practices[12] . *Cabell's Scholarly Analytics* is the most comprehensive and cited list[13] . Established in 2017 with 4,000 suspicious journal titles, it now gathers almost 15,000 (data as of September 2021). Cabell's criteria are derived from Beall's list, but have been expanded, more detailed, and updated (for example, the criterion of gender bias in journal editorial boards has been removed). In both cases, the sites state that they are 'suspected' predators, for the reasons outlined above. In fact, there are cases where it is easy to unmask predatory publications by the language used in the articles, perhaps written with one of the many automatic generators of scientific articles[14] or other artificial intelligence. These articles sometimes have nonsensical titles, such as those listed in Mike Downes' analysis of 1,165 publishers on Beall's list[15] . In the analysis we find titles such as "*The Physics of the Third Millennium*", "*The Conquest of Mars*", or "*Global consciousness (cognitivism to connectivism) & better worldliness - data study of spiritual consciousness measurement*", "*Life Universe, Universe Body and Something from Nothing*", "*General Law of the Universe and Unity of All Universal Forces*", "*Disclosing of Thousand Years' Mystery - Origin of the Book of Changes*".

Various other tools have been created to help researchers defend themselves against predatory journals. *Think Check Submit*[16] helps them identify the appropriate and reliable publication venue through a checklist. The Centre for Science and Technology Studies (CWTS) at the University of Leiden, together with Quality Open Access Market (QOAM), developed a check matrix that, instead of starting with the list of predatory journals, certifies definitely non-predatory journals[17] . The *Bona Fide Journals* system is based on the authentication of a jour-

11 H. Momen, "Jeffrey Beall and Blacklists", *SciELO in Perspective*, 2015, https://blog.scielo.org/en/2015/08/04/jeffrey-beall-and-blacklists/.

12 Scholarly Outlaws https://scholarlyoutlaws.com/.

13 *Cabell*'s https://www2.cabells.com/about-predatory.

14 Like SCIGen https://pdos.csail.mit.edu/archive/scigen/; Essay Generator https://www.essaysoft.net/essay-generator.html; Mathgen https://thatsmathematics.com/mathgen/

15 M. Downes, "*Why we should have listened to Jeffrey Beall from the start*", cit.

16 Think Check Submit thinkchecksubmit.org.

17 *Bona Fide Journals - Creating a predatory-free academic publishing environment* https://leidenmadtrics.nl/articles/bona-fide-journals-creating-a-predatory-free-academic-publishing-environment

nal's validity by librarians, and on other verified parameters such as presence in the Directory of Open Access Journals (DOAJ). On a more general level, WAME proposed an algorithm, based on Beall's criteria, to verify the reliability of the journal and identify predatory journals (Figure 3)[18] .

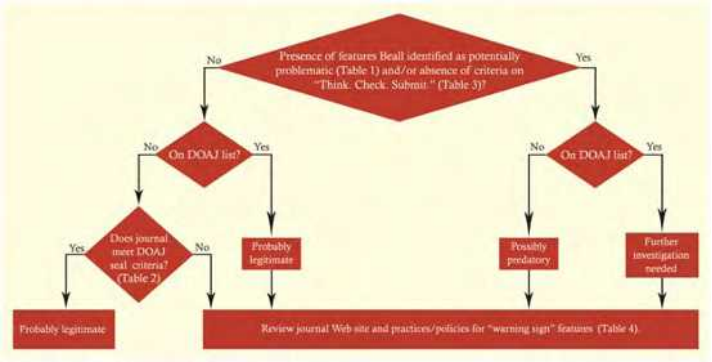


Figure 3 - Predatory Journal Algorithm (source WAME)

Jeffrey Beall had proposed that this type of fraud should be checked in research evaluation procedures and that publication in a predatory journal should not be considered in career advancement, competitions, and other evaluation procedures or should result in a penalty, up to and including the exclusion of candidates who present this type of publication in their CV. Indeed, publication in predatory journals causes a distortion not only in the publication count of the individual researcher, but sometimes in the calculation of bibliometric indices, as it is not uncommon for such journals to be cited and included in citation databases. Beall did not achieve any results in this respect, but he certainly raised awareness and laid the foundations for later initiatives. On closer inspection, his proposal with respect to the evaluation of research was unworkable because it would unfairly penalise the many researchers, perhaps in the early stages of their careers, who publish in predatory journals without realising it. But the observation remains that when predatory journals are present in databases, bibliometric

18 Christine Laine, Margaret A. Winker, *Identifying Predatory or Pseudo-Journals*, WAME February 18, 2017, https://www.wame.org/identifying-predatory-or-pseudo-journals#Table%204.

calculations are made on journals that are anything but scientific and lack peer review.

Italy is not exempt from such critical issues. A joint study by the University of Pisa and Aalto University in Finland looked at our country and the 46,244,000 researchers who participated in the first round of Abilitazione Scientifica Nazionale (ASN) 2012[19] . Of these researchers, the publications submitted for the ASN were compared with the titles on Beall's list. The result shows that of the 1,800,000 articles processed in the analysis, 5798 were published in Beall-listed journals, and that 5% of the authors included in the sample (i.e., 2225 researchers/doctors) published in a journal considered predatory (although 70% only once). According to survey data, among these researchers, the highest percentage is located in southern Italian universities, although the sample consisted of two thirds of researchers from northern Italy, one third from the south, and one-third from the centre. The analysis reveals a prevalence among articles in predatory journals of the subject areas of economics and business sciences, in which 4.1% of articles in English were published in a predatory journal. The other scientific fields follow with engineering (1.6%), humanities and social sciences (0.7%), and medicine (0.4%). Another finding is that in many cases these are early career researchers, who have numerous articles to their credit, but few with a high impact factor and who work in departments characterised by lower research quality levels (as measured by the results of the VQR 2004-2010). Also interesting is the average figure on the quantity and type of publications per researcher, amounting to 45 products over the ten-year period, of which 24 (53%) are journal articles, 8 (18%) are contributions in conference proceedings, 8 (18%) are monographs and book chapters, and 5 (11%) are other types such as abstracts in conference proceedings, databases, translations, commentaries on court decisions. The typology varies widely between the different disciplines, with journal articles prevailing in medical-scientific fields, contributions in conference proceedings prevailing in the engineering sciences, and books and book chapters prevailing in the social sciences, humanities, and economics. These data were then cross-referenced with a citation analysis based on Google Scholar to check the impact of predatory journals and articles in terms of citations, and with face-to-face interviews administered to a sample of 1,000 researchers. Following these verifications, the authors of the study found that only

19 Manuel Bagues, Mauro Sylos Labini, Natalia Zinovyeva, "A walk on the wild side: 'Predatory' journals and information asymmetries in scientific evaluations", *Research Policy* 48 (2019) 2, p. 462-477, https://doi.org/10.1016/j.respol.2018.04.013. ASN is a national qualification that must be achieved for the academic career https://abilitazione.mur.gov.it/public/index.php?lang=eng.

38% of the sampled journals met the criteria for inclusion in Scholar, i.e., having an h-index valued at least in relation to the last five years (which is the criterion on which Scholar's h-index is based) and having certain formatting rules, and that most of the articles had a very low impact, which is significant since Google Scholar tends to have very high indices and generally higher than other databases.

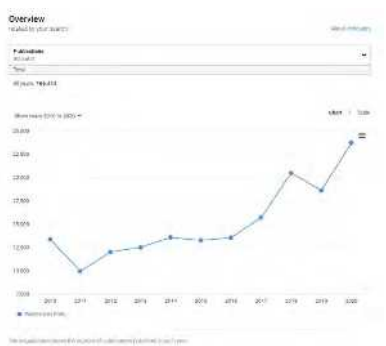
Since these were only 'suspected' and 'probable' predatory journals, as specified by Beall, the authors of the survey decided to interview a sample of 1088 Italian researchers among the 2012 ASN, belonging to 779 different departments, to directly verify their experience with the practices of these journals and to dispel any doubts. Several clues emerged from the interviews that should have made the authors suspicious. One-third of the interviewees answered that the journals in question had not returned any peer review reports to the authors during the submission process or that they had behaved suspiciously in terms of integrity and fairness. The researchers were then asked why they had chosen to publish in those journals, especially when they had suspicions of irregularities, and the answers were, predictably, mostly related to the need to be evaluated for career advancement, since, moreover, several journals were included in both Scopus and ANVUR's lists of scientific journals.[20] The authors of the study therefore wished to further ascertain these statements and indeed found that 131 of the titles on Beall's list were included in Scopus and 273 were part of the ANVUR scientific journal lists (213 of the sample investigated for ASN 2012), while only two were part of the Class A journal lists. They then analysed, as far as possible with respect to data availability, the results of the ASN round, and found that the success of researchers who published in predatory journals was lower than the others (by 3.5 points or 9%), but these researchers still exceeded the quantitative thresholds. A correlation also emerged between evaluator qualification and candidate evaluation, i.e. candidates who published in predatory journals were better evaluated by evaluators who themselves had a low scientific profile and few articles with a high impact factor. In many cases, the interviewees stated that they were misled by the statements provided by the journals. Indeed, the way in which predatory journals sometimes present themselves, with ISSNs, declared scientific committees, even though they may turn out to be false, their inclusion in citation databases and the insistent policy implemented, with continuous invitations to publish sent by email, may induce them to accept the offer. In other cases,

20 The national agency for evaluation of research in Italy, ANVUR, produces lists of journals considered "scientific" and "scientific class A" for non-bibliometric sectors, mainly humanities and social sciences, in which researchers must published for career advancement and other evaluation.

however, respondents stated that they were aware of dubious integrity, but published there anyway for reasons of career advancement.

The phenomenon of publication in predatory journals affects western and upper middle-income countries equally and less wealthy and developing countries, although there is a higher incidence in the latter. Negative effects do not only affect researchers and their evaluation but have a much broader and more impactful scope. Meanwhile, it is publicly funded research that ends up in the hands of less than honest publishers. But above all, they are publications that often have no control because there is no peer review or any kind of filter, so predatory articles could contain any kind of false and misleading content that becomes harmful and dangerous in certain contexts, such as those related to health. These and numerous other problems, such as the publication of unreliable data and the description of experiments, even on animals, without any statement regarding compliance with ethical principles, were found in the analysis of two thousand biomedical articles from two hundred journals considered predatory, conducted by an international group of scientists, and published in 2017 in Nature'[21] .

The growth in the number of predatory publishing cases and in the attention paid to the phenomenon are demonstrated by the data in the *Dimensions* citation database. A search carried out with the words "predatory journals" searched in the title, abstract, and keywords of the publications indexed by the database for the years 2010-2020, shows a significant and constant increase in publications on this topic (Figure 4).



21 David Moher et al., *"Stop this waste of people, animals and money"*, Nature 549 (2017) 7670, https://www.Nature.com/news/stop-this-waste-of-people-animals-and-money-1.22554.

Figure 4 - *Articles with the topic 'predatory journals' for the years 2010-2020 (source: Dimensions).*

A not dissimilar result is provided by Ngram Viewer with reference to the contents of Google Books, where the growth curve of the string 'predatory journal' clearly shows growth over the last ten years.

Figure 5 - *Search result for the string 'predatory journals' in N-Gramme Viewer.*

The predatory attitude not only concern the production of articles for journals, conferences, and other publications, but extends to various stages of research, such as evaluation. Various predatory impact factors such as the Journal Impact Factor (JIF - jifactor.org/) or the Global Impact Factor (GIF - globalimpactfactor.com) have emerged, which give an idea of how wide-ranging the issue is. The presence of a journal in such citation indices is considered an indication of predatory attitude by lists such as Cabell's. Some journals claim to be indexed by databases that produce the true impact factor or by another citation index, such as the Directory of Open Access Journals (DOAJ), when, in fact, they are not. Therefore, the confirmation of predatory attitude requires, as a first step, to check whether what is claimed is true.

A particular case of a predatory journal is the 'hijacked journal', defined as a journal that uses a name that is the same or very similar, in such a way as to be misleading, to that of a prestigious and well-known journal, and perhaps even indexed in the main citation databases such

as Scopus and Web of Science and thus endowed with an impact factor[22] . The intention of such journals is to attract authors, in particular by inducing them to pay APC fees or other services. Like other predatory journals, they set up websites and send email to potential authors. Their number is not known, but Beall's list listed more than a hundred, mainly in the STM fields. In the humanities, the proliferation of hijacked journals seems difficult since, unlike in the STM, the journal is not the main means of disseminating research results, and therefore journals are fewer in number within fields that tend to be more fragmented into specialised fields of research with not too large a number of researchers, and therefore a clone of a journal would not go unnoticed. As with other predatory publishing, the hijacked journals use fake editorial boards, even including names of real people without their knowledge, and offer little information on websites about the scientific committees and the journal in general. In not uncommon cases, services sold include co-authoring[23] , i.e. the possibility of appearing as a co-author in an article, or the promise of publication, with even fake acceptance letters being sent for the sole purpose of getting paid[24] .

One can run into these journals by mistake or voluntarily, thus falling into the case of pseudo-journals. But special attention must be paid to this risk since publication in predatory or hijacked journals is never without consequences, even for the author. As Salim Moussa[25] explains, and as confirmed by the numerous cases surveyed and commented on in Retraction Watch, an article published in a predatory journal is very difficult to have withdrawn because once any APC fees or payments for other services have been pocketed, the journals often do not even respond to attempts to contact them or disappear altogether. The published article remains in the author's curriculum vitae and is also considered already published, and therefore cannot be republished in other serious publishing venues. The predatory journal can close down from one day to the next, and the article is no longer published and therefore not presentable for the purposes of evaluation pro-

22 Salim Moussa, "Journal hijacking: Challenges and potential solutions", *Learned Publishing*, first online 26 July 2021, https://doi.org/10.1002/leap.1412.

23 Anna Abalkina, "Unethical Practices in Research and Publishing: Evidence from Russia", *The Scholarly Kitchen*, Feb 4 2021, https://scholarlykitchen.sspnet.org/2021/02/04/guest-post-unethical-practices-in-research-and-publishing-evidence-from-russia/; Alison McCook, "7 signs a scientific paper's authorship was bought", *Retraction Watch* Oct 24 2016, https://retractionwatch.com/2016/10/24/seven-signs-a-paper-was-for-sale/; see also, as an example https://science-publisher.org/detailed-description-of-the-co-authorship-service/.

24 Alison McCook, "Congrats! Your paper was accepted. (Except if the acceptance letter was forged)", *Retraction Watch* June 5 2018, https://retractionwatch.com/2018/06/05/congrats-your-paper-was-accepted-or-was-it/.

25 S. Moussa, "*Journal hijacking",* cit.

cedures or career advancement, and if the journal has falsely claimed to be indexed in Scopus or WOS and the article is not indexed there, it will not result as a bibliometrically assessable product. The author's home institution finds itself involved in a veritable form of malfeasance, and its prestige is undermined, as well as being forced to take note of having directed research results to fraudulent operators. The original journal that is seized also suffers damage. First of all, because APC fees may be paid to the predatory journal and not to the original, to which is added damage to authoritativeness and image; moreover, proximity in journal names may induce authors to cite the wrong journal and, in this way, distort the citation count for the impact factor. The issue is not as simple as it might appear, as this also implies that reviewers of a journal must check, at the peer review stage, that citations and bibliographic references are for genuine, non-predatory journals. While this might not be too complicated for smaller scientific fields, in very large fields a spot check would be complex.

In addition to the implications for individual authors and individual journals, at a general level, the effect that such practices are having on the world of scientific research and on one of the backbones of science is becoming increasingly evident. In fact, two types of journals are being created, which, moreover, pander to the needs of the current research system: on the one hand, journals that only serve to publish at any cost, and on the other, journals that actually serve to disseminate the results of academic research and advance science. The gradual corporatisation of research and the commercialisation of the sources through which to disseminate its results is leading to the emergence of a veritable bad science industry, of which the paper mills are another striking case in point.

2. Paper mills: the flourishing market for buying and selling articles

A further widespread phenomenon, this time born recently, is paper mills. These are private profit-orientated, and usually illegal, organisations that create articles, entire datasets or images in the guise of scientific research results and then sell them to authors who request them[26] . Sometimes the sale also takes place after the article has passed peer review by a journal, whereby the organisation offers for sale to the researcher the possibility of adding their name to an article that has already been accepted for publication. The price may vary depending on the position one wants among the authors, since in some disciplines

26 However, the phenomenon is also widespread among students where 'essay mills' are spreading, offering essays and term papers, cover letters and other useful material for students for sale. The British government has openly declared essay mills illegal in the recent reform of the education system, the *Skills and Post-16 Education Act 2022, see* https://www.gov.uk/government/news/essay-mills-to-be-banned-under-plans-to-reform-post-16-education.

the position in the list of authors of an article indicates different functions and responsibilities, or the journal's impact factor. Articles are published not only in predatory journals, but also in serious and authoritative journals with an impact factor.

The request may come from the researcher contacting the paper mill, or in 'push mode' from the organisation sending an email informing that an article on a certain topic has been accepted for publication in a journal and is for sale. This industry is rather sophisticated and large-scale, as articles are produced in large quantities using standard templates to which variations are applied as appropriate but making sure not to overdo the falsification of data, so that manuscripts can appear real and be accepted by journals[27] . Often, paper mills have real laboratories and researchers behind them who study how to construct the falsified articles, data, or images, in a plausible manner, or take already published articles and modify them, or even translate them into English from unfamiliar languages, such as Chinese. They usually present themselves on the surface as companies offering support services to researchers for the publication of research results, such as language services related to the need to publish in English, or editorial services and assistance in finding the right journal in which to publish and in the manuscript submission stages, but in many cases the offer is much less ambiguous. Sometimes, paper mills take advantage of the collaboration of researchers and members of the editorial boards of journals with few scruples.

The problem is assuming considerable and worrying dimensions in some countries such as China, where it has become a de facto business sector with a lively market of supply and demand. A few years ago, a survey reported in an article in Science showed how widespread it was and how easy it was, to find one of the many paper mill companies with a simple search on Baidu, China's main search engine[28] . The pressure on researchers, who are required to publish in journals with high bibliometric indicators, and on doctoral students, who have to publish scientific articles indexed in Web of Science, makes the purchase of articles a practice in China anyway, especially in the medical sector. Additionally, non-university hospital doctors are also asked to publish articles for career advancement, so their lack of research and publishing habits leads them to have to resort to such practices. In fact, one of the journals that most frequently fall into the meshes of this flourishing

27 Jennifer A. Byrne, Jana Christopher, "Digital magic, or the dark arts of the 21st century-how can journals and peer reviewers detect manuscripts and publications from paper mills?", *FEBS Letters* 594 (2020) 4, p. 583-589, https://doi.org/10.1002/1873-3468.13747.

28 Mara Hvistendahl, "China's Publication Bazaar", *Science* 342 (2013) 6162, p. 1035-1039, https://doi.org/10.1126/science.342.6162.1035.

market is the Chinese Medical Journal, the open access journal published by the Chinese Medical Association.

The phenomenon is spreading, and many publishers are beginning to openly denounce it. In 2020, the Royal Society of Chemistry (RSC) carried out a thorough analysis of its publications and found 68 articles that had been retracted for this reason, which were then listed on its website and forced the scientific society to apologise to the readers of its journals[29] . An analysis conducted on a specific Russian paper mill for the years 2019-2021 identified 434 articles of that origin published in predatory journals and authoritative journals by authors from 39 different countries[30] . According to a study in the journal Nature, in the year between January 2020 and January 2021 alone, journals withdrew 370 articles explicitly accused of being products of paper mills[31] . The Institute of Physics (IOP) withdrew 494 articles from its journals and *conference series,* as announced in September 2022[32] . The Retraction Watch database lists just under 1600 officially retracted articles with the reason 'paper mill'[33] . Data from other searches conducted to investigate such cases of fraud were collected in a blog, which in March 2021 counted 1,300 articles suspected of coming from paper mills and around sixty subject to an expression of concern', i.e. suspected to be products of a paper mill but not yet ascertained or not yet retracted[34] . Of these, 26% were retracted, others were still being analysed by the journals. Comparison of the two data, recorded at a distance of time, shows the growth of the phenomenon and confirms how, even for these cases of fraud, it is often difficult to detect with certainty, making the time lapse from publication of the article to the time when it is retracted problematic. After publication, there is a more or less long period before it is discovered (or at least suspected) to be probable fraud, followed by the time needed for actual verification and to convince the journal of the need to retract the article. The time for retraction can

29 *RSC Advances retractions*, https://www.rsc.org/news-events/articles/2021/jan/paper-mill-response.

30 Anna Abalkina, "*Publication and collaboration anomalies in academic papers originating from a paper mill: evidence from a Russia-based paper mill"*, deposited in arXiv 26 December 2021, version 2 revised 26 March 2022, arXiv:2112.13322v2.

31 Holly Else, Richard Van Noorden, "The fight against fake-paper factories that churn out sham science", *Nature News Feature* 23 March 2021, https://www.Nature.com/articles/d41586-021-00733-5.

32 Ivan Oransky, "Physics publisher retracting nearly 500 likely paper mill papers", *Retraction Watch* September 9 2022 https://retractionwatch.com/2022/09/09/physics-publisher-retracting-nearly-500-likely-paper-mill-papers.

33 The figure is updated to the beginning of October 2022.

34 Reported in Else and Van Noorden, "*The fight against fake-paper factories that churn out sham science"*, cit.

therefore be very long, even years, and during this period the article circulates freely.

The extent of the phenomenon may be significantly greater than has already emerged since it is relatively new, and not all publishers bother to carry out systematic analyses of this kind, except for cases reported by *whistleblowers*. The high probability of not being discovered is one of the reasons behind the flourishing 'business model' of paper mills[35] . The aforementioned COPE publishers' alliance, which is moving extensively to safeguard the publishing world precisely in order to avoid cases of scientific fraud such as these, points out in the forum dedicated to paper mills that it is sometimes very complex to detect these articles if one does not have the possibility of comparing articles with each other and that, in order to do so, it is often necessary to compare different publishers, which is not always easy[36] . Only in this way is it possible to detect cases of duplication and manipulation, although the sophistication in modifying and adapting paper mills' templates for different needs has now reached a high degree, also thanks to the use of artificial intelligence techniques[37] , and especially when researchers are behind these organisations. In the analysis of the individual article, paper mills may not be obvious, even if one has the raw datasets used in the article, as experts on very specific topics are often required to notice them. Therefore, it may happen that the article also passes peer review in a journal. However, even when one has suspicions and attempts to contact the author to unravel them, one cannot be sure that it is actually the author who answers and not the paper mills directly. When this is the case and the article cannot therefore be retracted on objective grounds, editors label suspicious articles with an 'expression of concern' (EOC), so that it is clear that doubts remain about the integrity and scientificity of the article. The consequences are tangible if one thinks of the medical field which, as COPE denounces, is among

35 Jennifer A. Byrne, Jana Christopher, "*Digital magic, or the dark arts of the 21st century",* cit.

36 COPE, *Systematic manipulation of the publishing process via 'paper mills'*, https://publicationethics.org/systematic-manipulation-paper-mills.

37 Software for automatic article generation has existed for a long time and is now perfected with the use of artificial intelligence techniques. Sci-Gen is one of the most popular software for creating articles https://pdos.csail.mit.edu/archive/scigen/. See also Cyril Labbé, Dominique Labbé, D., "Duplicate and fake publications in the scientific literature: how many SCIgen papers in computer science?", *Scientometrics* 94 (2013), p. 379-396, https://doi.org/10.1007/s11192-012-0781-y; Richard Van Noorden, "Publishers withdraw more than 120 gibberish papers", *Nature News* 24 February (2014), https://doi.org/10.1038/Nature.2014.14763; Ian Sample, "How computer-generated fake papers are flooding academia", *The Guardian* 24 Feb 2014, https://www.theguardian.com/technology/shortcuts/2014/feb/26/how-computer-generated-fake-papers-flooding-academia.

the most affected sectors, not only in countries like China, but globally. However, the phenomenon is widespread in all disciplines, as can be seen from the COPE forum, where numerous cases discovered in fields such as computer science, engineering, social sciences, and the humanities are reported.

In June 2022, a joint report by COPE and the association of scientific and technical publishers STM[38] was published, the aim of which is first to make publishers aware of the existence of these fraudulent companies and, at the same time, to raise awareness among research evaluation agencies and academic institutions so that they change their evaluation systems so that authors are not forced to turn to these operators in order to publish quickly. The report reproduces some examples of advertisements published by paper mills to attract customers and highlights an important aspect in relation to the punishability of such activities, which are not considered as illegal activities in all countries and, therefore, not always punishable and condemnable. In compliance with privacy and GDPR, publishers are striving to create a network to effectively address the problem, by exchanging information and best practices to deal with paper mills and other violations of ethics and scientific integrity in publications, and by strengthening procedures to detect fraud during the peer review process and thus before the publication of an article, through the establishment of ethics and integrity committees to which reviewers of an article can refer if they have suspicions. There are also initiatives by individuals such as the one created by Guillaume Cabanac and Cyril Labbé, two French computer scientists who in 2015 set up the site *The Problematic Paper Screener* to automatically detect suspicious articles using algorithms[39] .

Generally speaking, many publishers are increasing their checks on authors and tightening their standards for accepting articles, providing for checks on the context in which the research was conducted, in some cases requiring one of the authors to take responsibility for the integrity of the article by acting as a guarantor throughout the publication process. Some governments, such as China, are taking action in this regard. In China, the Ministry of Universities recently announced that it would change the research evaluation system in an attempt to curb fraud and other illegal behaviour. Following the DORA Declaration[40] ,

38 COPE & STM, *Paper Mills - Research report from COPE & STM* 2022, Committee on Publication Ethics, Version 1: June 2022, https://doi.org/10.24318/jtbG8IHL.

39 Diana Kwon, "Nature's 10. Ten people who helped shape science in 2021. Guillaume Cabanac: Deception sleuth", *Nature* 15 December 2021, https://www.Nature.com/immersive/d41586-021-03621-0/index.html#section-gM9iO4XBRl.

40 *San Francisco Declaration On Research Assessment (DORA)* https://sfdora.org/read.

evaluation will be carried out with systems different from the current one based on Web of Science indicators, not only in universities but also in hospitals, and the peer review system will be preferred. Some scholars have raised concerns about peer review-based evaluation, due to its subjective nature and the presumed risk that loss of the objective of achieving bibliometric indicators would lead to a decline in China's scientific competitiveness[41] . This fear is indicative of the change in values in scientific research, where the objective no longer is to achieve results in scientific terms, but positioning of research in international rankings is perceived as a priority objective, thus shifting the focus to the level of the institution and not of the individual researcher or research group and moving away from the scientific objective itself. The international rankings represent an evaluation system that emphasises this shift toward the institution. Almost all international rankings, such as the Shanghai Academic Ranking (ARWU), Quacquarelli Symonds Ranking (QS), Times Higher Education (THE), U-Multirank, GreenMetric, include, among their evaluation parameters, to varying degrees, the bibliometric indices on publications produced by the Scopus and WoS citation databases, and they generate the same distorting effects on research but also on teaching, especially if they are linked to funding, and thus the same attempts to cheat the metrics, as the Columbia University affair on the misleading data provided for the rankings, denounced by a *whistleblower* of the same university, has shown.[42]

3. Zombie papers: Sometimes they come back

In the bibliometric literature it is a well-known case of *sleeping beauties*, articles that remain 'dormant', i.e. without citations, or with few citations, for a long time and then suddenly reawaken when, for various reasons, the topic of the article returns to interest. To give just one example, the article by Ettore Majorana on the relativistic theory of particles with arbitrary intrinsic momentum, the subject of one of his publications in 1932 in Il Nuovo Cimento, remained dormant until the

41 Smriti Mallapaty, "China Bans Cash Rewards for Publishing", *Nature News* 579 (2020) 5 March, https://www.Nature.com/articles/d41586-020-00574-8.

42 Chris McGreal, "Columbia whistleblower on exposing college rankings: 'They are worthless'", *The Guardian* 16 September 2022, https://www.theguardian.com/us-news/2022/sep/16/columbia-whistleblower-us-news-rankings-michael-thaddeus; Orsola Riva, "Università e ranking, lo scandalo della Columbia: ha fornito dati «ingannevoli». Ora è scesa dal secondo al 18esimo posto", *Corriere della Sera* 13 settembre 2022, https://www.corriere.it/scuola/rientro-a-scuola/notizie/universita-ranking-scandalo-columbia-ha-fornito-dati-ingannevoli-ora-scesa-secondo-18esimo-posto-45c9e7da-3343-11ed-80fb-2302675b77bf.shtml.

mid-1960s when particles of this kind were discovered in accelerators that were unknown in the 1930s[43] . Sometimes, however, the articles are indeed deceased, and not just dormant but at some point, resurrected as zombies. Zombie papers are articles retracted by journals after publication that continue to be cited positively by other articles. The phenomenon does not only concern articles, but also different stages of the scientific process, so we also speak of zombie trials, for example. An analysis of 153 trials in articles submitted to the journal Anaesthesia for the years 2017-2020 found that 44% of the trials were false and 26% were zombie trials[44] . The consequences of such behaviours are enormous, particularly still in the medical field, as they distort statistics, lead down the wrong path, and lead to false results whose effects have repercussions on public health and at the same time slow the progression of science. They are like whirlpools that always rewind and stop the natural course of science.

The example of the well-known article on vaccines by Andrew Wakefield, to which we have already referred, clearly demonstrates this and is an exemplary case study for zombie papers. In the article in question, published in 1998 in The Lancet, the author claimed a link between a type of vaccination and autism. The issue generated a huge debate on vaccines, which is still present today, and at the time of its release naturally triggered further studies on a topic of great interest in the medical field. However, no scientist was able to reproduce the results of the article, which, in fact, turned out to be fraudulent, based on falsified data and written with the intention of benefiting the marketing of a new vaccine competing with the one the article was about. Wakefield was involved in the production of the competing vaccine and, in addition to falsifying the data, had not declared a conflict of interest. Wakefield's article was withdrawn from the journal several years after it came out, when, thanks to the work of a journalist who initiated the case, it was established that it was based on false data, which led the British General Medical Council to disbar the author from the Medical Council. This was not enough to stop the spread of the content and the spread of the theory that the vaccine (in the meantime extended to any vaccine) causes autism, nor was it enough to stop the citations of the article. A study in Retraction Watch found that Wakefield's article had received more than a thousand citations and that many researchers missed the word 'retracted' when they cited it, and therefore

43 Yves Gingras, *Bibliometrics and Research Evaluation, Uses and Abuses*, Cambridge (Ma), MIT Press, 2016, p. 30-31.

44 J. P. A. Ioannidis, "Hundreds of thousands of randomised trials circulate among us"*, Anaesthesia* 76 (2021), p. 444-447, https://doi.org/10.1111/anae.15297.

did not report in any way that the article had been retracted[45] . This case makes it clear that the quality and impact of an article, especially when measured in terms of citations, are by no means synonymous. Wakefield's article had a considerable impact which, had it been measured by citation count, would have given it rather high indicators but it is clear that the quality of the article was not high at all; indeed, the article was a fraud. Furthermore, the citations of that article, as of all retracted publications, continue to affect the calculation of indicators in the databases in which it is included, such as the impact factor. This is not the only case, as the citations received by articles after retraction are usually many, and some studies show that on average only a quarter of the citations report retraction. Retraction Watch lists the ten articles that received the most citations after retraction, and one can see that these are significant percentages of the total citations counted in Web of Science[46] . As much as it may make sense to keep articles in citation databases, as it is important, we reiterate, both for bibliometric and network analysis citation studies and for studies in the sociology of science, it is in any case misleading and wrong for this indicator to be associated with an assessment of the quality of a publication considered fraudulent.

More recently, similar behaviour was observed during the Covid-19 pandemic, during which several articles were published that were later retracted. Retraction Watch has created a specific section listing the retracted Covid articles, the number of which is constantly being updated and is over 260, as we have already seen. This is a significant number if one considers, as previously pointed out, that it represents only the emerged part of a phenomenon that is undoubtedly larger. On the sidelines, it is also necessary to consider how fast the transmission times through digital channels are today and how much the world of journalism has changed, in which the speed and sensationalism of the news now prevail over the slowness required to accurately verify sources. Journalists draw extensively on the scientific literature, especially when available in open access and preprint archives, and from the moment a study is published to when it is retracted, there is plenty of time for the news to circulate widely and even go viral. Often, journalism adopts vicious behaviours and mechanisms such as clickbait, whereby the metric for measuring the success of an online article is the number of clicks,

45 I. Oransky, "*Andrew Wakefield's fraudulent paper on vaccines and autism has been cited more than a thousand times",* cit.

46 "Top 10 most highly cited retracted papers" *https://retractionwatch.com/2015/07/14/half-of-anesthesiology-fraudsters-papers-continue-to-be-cited-years-after-retractions/;* "Half of anaesthesiology fraudster's papers continue to be cited years after retractions", *https://retractionwatch.com/the-retraction-watch-leaderboard/top-10-most-highly-cited-retracted-papers/.*

also because it is perhaps based on the number of clicks that sponsorships are obtained[47] . Add to this the political or other instrumentalisation that certain news items are often subject to. Covid was an example of this, as it remained a 'hot' topic for a long time. The health emergency has increased the pressure on researchers to discover something new and useful about a previously unknown virus, generating a great deal of research and, thus, publications. However, the pressure also led to increased scrutiny by the scientific community, especially since many publications on Covid were made open access available by publishers[48] , and fast turnaround times. However, this was not enough to prevent the proliferation of zombie papers. The journal Science conducted an analysis of two highly influential and then retracted articles and found that in the space of just over six months, starting in June 2020, when the articles were retracted less than a month after their release, as many as 200 other publications cited those two articles, not to refute them but to support further studies, thus in a positive way, and of these, the majority, 52.5%, did not report in any way that the cited articles had been retracted[49] . Of course, this may depend on the timing of publication of the articles, so there is no way to verify that the cited article has been retracted in the meantime, but with such a rapid pace of retraction as in the case studied by Science, it is nevertheless something to think about. The two incriminated articles were published in major journals such as The New England Journal of Medicine (NEJM) and The Lancet. The articles supported the efficacy of hydroxychloroquine in the treatment of Covid, claiming to have conducted the study on a large database of hospital patients from around the world produced by Surgisphere, a company owned by the vascular surgeon Sapan Desai, coauthor of both articles. None of the articles that subsequently appeared, with bibliographic references and citations of the two hydroxychloroquine studies, took notice that they had been retracted, and none of the reviewers who conducted the peer review noticed.

As evidence of how the phenomenon of retractions also affect authoritative journals, a search in the Retraction Watch database retrieves for The Lancet, a journal published by Elsevier, 37 retracted articles with more than a hundred total reasons for retraction (thus an average of three per article), and for NEJM 38 articles with almost 90 reasons. Ivan Oransky, one of the cofounders of Retraction Watch, blames the

47 B. Frampton, "*Clickbait",* cit.

48 Open access has favoured the greater circulation of research results and also a more capillary control over these results by the scientific community. On these aspects I refer to R. Morriello*, Le raccolte bibliotecarie digitali nella società dei dati,* cit., in particular p. 7-15.

49 Charles Piller, "Disgraced COVID-19 studies are still routinely cited", *Science* 371 (2021) 6527, p. 331-332, https://doi.org/10.1126/science.371.6527.331.

authors who often quote articles without having read them (and this is no stranger to the need to publish, at a fast pace, publish or perish), but there is no doubt that there are also other motivations. In several of the cases found, the authors self-cite their own retracted articles, articles whose fate they were therefore well aware of, and the action cannot be justified by superficiality in reading[50] . Part of the responsibility then lies with the journals themselves. Peer review is one of the pillars of science for the purpose of verifying the scientific validity of a piece of research. In fact, the more rigorous the peer review and the higher the rejection rate of a journal, the more prestigious it is considered. Scientific research evaluation procedures are based on these foundations, which rarely include forms of verification with respect to the integrity of the research and cases of retraction, at least in Italy[51] . The bibliometric classification and the division of journals into 'scientific' and 'class A' journals by ANVUR is based on the recognition of an authority attributed to journals as guarantees of a quality standard. It is evident that in many cases this no longer corresponds to the truth. The reasons are many and intersecting. Meanwhile, the ever-increasing number of published articles makes it difficult to keep up with peer review, as it has become difficult to find reviewers, considering that this is a costly activity that is not evaluated in career advancement procedures. Furthermore, it should be mentioned once again that it is not always easy to detect a forgery, especially when it is sophisticated and carefully elaborated. There is often a lack of tools to do so, as is the case with trials, for instance, where the complete dataset used in the experiment must be available. There is also sometimes an ambiguous attitude on the part of journals with regard to retractions[52] . It is difficult to obtain a retraction of an article as this could affect the journal's prestige, real or perceived, and cast doubts on the procedures adopted, such as peer review. Furthermore, a factor that may contribute to slowing down retraction is the payment by an author or his institution of APC fees for the article to be retracted. But it is also true that one needs to be absolutely certain that it is manipulation and fraud before retracting, as one could jeopardise the career and authority of a researcher as well as the journal. Retracted

50 Enrico M. Bucci, "On zombie papers", *Cell Death & Disease* 10 (2019) 189, https://doi.org/10.1038/s41419-019-1450-3.

51 The first research evaluation exercise conducted by ANVUR, the VQR 2004-2010, provided that in proven cases of plagiarism or fraud, the publication would be negatively weighted with a -2 weight. However, this system has disappeared in subsequent evaluation exercises, probably also due to the difficulty of ascertaining cases of plagiarism and fraud.

52 Adam Marcus - Yoshitaka Fujii, "Zombie papers: Why do papers by the most prolific fraudster in history keep getting cited?", *Retraction Watch* April 1 2020, https://retractionwatch.com/2020/04/01/zombie-papers-why-do-papers-by-the-most-prolific-fraudster-in-history-keep-getting-cited/.

articles remain in the citation databases and contribute to the impact factor, with positive and negative citations. As is well known, one of the limitations of citations and the bibliometric indicators built on them is, in fact, that they do not distinguish positive from negative citations, nor do they distinguish between positive and negative citations on retracted articles[53] . Evaluation agencies rely on the judgment of journals, which are often no longer able to perform the necessary certification and validation. Evaluation agencies and universities should begin to financially support instruments such as Retraction Watch and perhaps even consult and use them regularly, as well as begin to think deeply about the long-term consequences of the emerging scenario, which is largely a consequence of the publish-or-perish culture.

4. The Italian situation

To contextualise the historical analysis and the definition of research ethics and integrity issues with respect to the Italian system, I analysed the data extracted from Retraction Watch in order to obtain an overview of the extent of these issues in our country. Although I was unable to hypothesise a cause-and-effect correlation with the research evaluation exercises, the analysis offered interesting and relevant information for thought. The Retraction Watch website was established in 2010 and in 2018 launched a searchable database that now counts more than thirty thousand retraction news reports[54] . It should be noted that retractions are not always related to cases of fraud or ethical violations since sometimes it is the authors themselves who notice errors and withdraw articles. However, the site is an essential tool, rich in information and constantly updated. Collaboration with Zotero allows those using Retraction Watch to be notified when an article is retracted[55] .

The Retraction Watch database contains 639 retractions for Italy, retrieved through the country search, by which is meant the nationality of at least one of the authors of the article. The database also censuses cases in which the publisher has expressed doubts, indicated as 'expressions of concern', and cases of 'corrections,' in which articles have been amended to make changes to remedy problems of various kinds. For Italy, there are 124 *expressions of concern*, 44 *corrections,* and 5 *reinstatements*, i.e. articles prudently retracted and then reinstated. Therefore,

53 Although some citation databases are trying to go in this direction by using artificial intelligence techniques to signal the tone of the citation, whether negative or positive, see for instance Scite.ai with 'smart citations'.

54 The Retraction Watch Database [Internet], New York, The Center for Scientific Integrity, 2018, http://retractiondatabase.org/.

55 *Retracted item notifications with Retraction Watch integration https://www.zotero.org/blog/retracted-item-notifications/.*

the total number of problematic cases is around 800[56] . In the following analysis, I have excluded the cases of expression of concern and corrections, as the former definition includes suspicious publications that might not result in an allegation leading to retraction, and in the latter case, it is assumed that the flaws found in the articles have been remedied and the errors corrected. Of the 639 retracted articles, 611 are open access, the others behind paywalls, or those data were unknown. A first check carried out for Italy concern the disciplinary areas of the withdrawn products. The graph in Figure 1 shows the proportion in the macro-areas, for which it must be taken into account that each of them is subdivided into several sub-areas, which are not homogeneous in number. For example, ENV-Environment has six (climate change, climatology, ecology, environmental sciences, food science, ground/surface water), while other areas have many more, according to a classification prepared by the site.

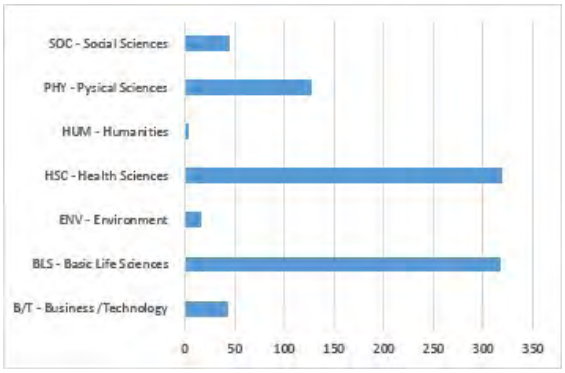


Figure 1 - Retractions Italy by disciplinary grouping in Retraction Watch

The graph shows how the problem of retraction in Italy, which is in line with studies carried out in other countries, mainly concerns the STM sectors, and in particular the pure sciences and, to a much lesser extent, the humanities. Among the former, the physical sciences account for a less significant portion, a fact that seems to give reason to Blaise Cronin, who argues that in the field of physics there is less tendency to fraud because there is a structured international collaboration network, and the practice of sharing preprints with the partial results

56 Last data extraction 16 October 2022.

of ongoing research exposes the research and data to peer scrutiny even before being published in a journal[57] . In any case, the numbers are not low for physics either, so it is an indication that something is changing in that area as well. The difference between the disciplines depends on many factors, one of which is probably the greater number of articles that are published in STM areas compared to HSS where the monograph continues to play a significant role, but as the above-mentioned studies show, the pressure that research evaluation systems based on exclusively quantitative and bibliometric criteria place on researchers is not irrelevant, unlike HSS publications that usually go through peer review, partly because they are not sufficiently represented in bibliometric databases.

Another data extraction affected the types of publication subject to retraction (Figure 2), which confirms the prevalence of research articles over other types, followed by review articles and clinical studies.

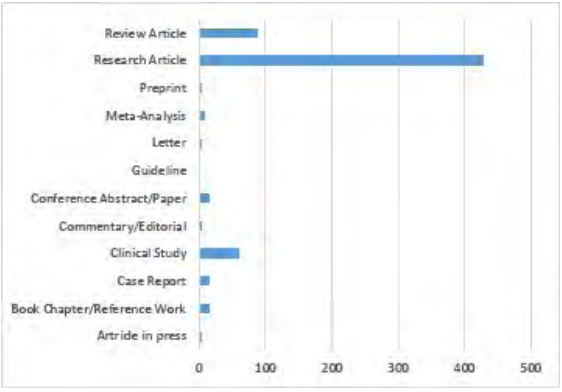


Figure 2 - *Italian Retractions by Publication Type in Retraction Watch*

The publishers that had to retract articles, again limited to retractions of articles with authors of Italian affiliations, are many, and the list includes the major international scientific publishers: Elsevier (100 retractions), Springer (72 retractions), Wiley (53 retractions) and PLoS (38 retractions) are the publishers that lead the ranking. On the other hand, these are the most important scientific publishers that cover most of the market and possess the highest number of prestigious journals in which researchers seek to publish, so it is obvious that they have

57 B. Cronin, *The Hand of Science,* cit., p. 61.

a high number of retractions as they have a high number of articles. The almost total absence of Italian publishers is striking, but retractions are mainly attributable to STM areas characterised by a prevailing tendency to publish in English in foreign journals, as opposed to the HSS areas more tied to a local publishing market and national languages. Nevertheless, the list shows that no publisher, not even the most authoritative, is exempt from the problem.

The reasons for retraction listed on Retraction Watch are many and often overlapping, so an article is frequently retracted for more than one allegation of ethical or integrity violations. Furthermore, definitions may differ for the same phenomenon, as the data is based on what the journals state. Journals define cases of retraction differently, so that in the database we find very general expressions such as 'problems with the data', 'problems with the images' or 'journal investigation','misconduct' and even 'retracted' without further specification. For example, 81 retractions report 'investigation by a journal' as the reason, 69 'investigation by the institution / company', and 36 'euphemism for plagiarism'. In any case, the range of motivations for Italian retractions covers almost the entire spectrum of possibilities envisaged by Retraction Watch, and there are more than 80 in total. As reported by a study published in the journal JLIS.it that analysed the causes of retractions by Italian authors, plagiarism is the main cause[58] . Overall, plagiarism, divided into four reasons, in descending order of incidence, of plagiarism of articles, text, data and images, accounts for around 10% of the total number of retractions. The order also reflects the difficulty in discovering cases of plagiarism. Not infrequently, discovery occurs only after the article has been published. For plagiarised whole articles or parts of the text, anti-plagiarism software can be used, and adoption during peer review, in the event of suspicion, could lead to the plagiarism being discovered and, thanks to the use of information technology, to reasonable certainty. But the operation becomes more complex for data and images. Manipulated and falsified images are difficult to detect, especially by the human eye of the peer reviewer, and reviewers are certainly even less likely to suspect image manipulation than is already the case for text plagiarism. *Reverse image search* software, now commonly used, helps, but only in the case where the entire image is plagiarised and has already been published, while it is less supportive when laboratory images, graphics, or portions of an image are plagiarised, in which case scientific knowledge is indispensable to detect tampering. In addition to the retractions explicitly labelled as plagiarism in

58 Gonzalo Marco-Cuenca, José Antonio Salvador-Oliván, Rosario Arquero-Avilés, Chiara Faggiolani, Brenda Siso-Calvo, "Scientific publications of affiliated authors in Italy retracted due to fraud. Review and analysis", *JLIS.it* 12 (2021) 2, http://dx.doi.org/10.4403/jlis.it-12711.

Retraction Watch, there are also cases labelled 'duplication' of articles, texts, data, and images. In general, retractions for image-related issues attributable to plagiarism, duplication, errors, manipulation, falsification/manufacturing, and other issues not specified further account for about 13% of retractions.

The third element emerging from the analysis of what is in Retraction Watch concerns the data, divided into the cases of error, manipulation, falsification/manufacturing, plagiarism, and various other more specific causes, to which a generic section called problems with the data'. Recall that the definitions of the reasons are applied to retractions by the publishers, and the database reproduces what the publisher states. To this must be added the justifications described as non-reproducibility of the research since, in all likelihood, the inability to reproduce the research is due to problems in the data. The 45 cases of 'fake peer review' also stand out among the motivations, which would suggest predatory journals that declare peer review but do not carry it out. Only two accusations of a 'paper mill' are found for Italy among the total number of retractions recorded in the database.

A look at the Italian reality cannot ignore certain considerations. Retractions take place after a period of time that can be very long, even years, during which the publication can participate in all kinds of evaluation. The reviewers in the VQR and the commissions for the ASN rely on the trust placed in the fundamental structures of science represented by the journals, which should act as a filter in this respect, and on the bibliometric databases. But both sources are not always able to contain and control abnormal and illicit behaviour.

5. Investigation of bad science

The widespread belief that fraudulent practices and predatory journals can be attributed exclusively to the emergence of open access is unfounded, as we have tried to demonstrate by tracing a diachronic path that we hope has made it clear that scientific fraud and predatory publishing have existed as long as modern science has existed. It is plausible that with open access, and in particular with the dynamics triggered by certain acquisition models based on APC payments, there has been an increase in the number of cases, but it should not be forgotten that open access encourages the discovery of predatory journals and turns the spotlight on the phenomena of distortion, violation of ethics and integrity. If, therefore, the number of cases is growing, it has also become easier to discover them. In fact, it has never been easy to precisely define the extent of fraud and ethical violations and their evolution over time. Before the emergence of Retraction Watch, there were no uniform sources keeping track of them. Several studies

and statistical surveys were reported in the scientific literature, but they were partial and often difficult to compare, as they were not homogeneous. Daniele Fanelli, currently a researcher at the London School of Economics, tried to systematise them in a 2009 article[59] . From the analysis conducted and described in the article, Fanelli found that on average 1.97% of researchers admitted to having fabricated, falsified or modified data or research results at least once. A recent study conducted in the Netherlands raises this figure to 8%[60] . As previously emphasised, the widespread certainty is that only a small proportion of cases emerge, and the phenomenon is much more extensive than researchers self-report and can be discovered. In any case, the scientific community's attention to these issues has certainly increased. In the 1960s and 1970s, few cases could be found in the literature, and behaviour that undermined the integrity of research was seen as exceptions, condemned as such, but as occasional deviations from the self-correcting nature of science. From the 1980s onwards, the incidents began to intensify, clearly showing how they also affected prestigious journals and researchers from important universities[61] . Moreover, this was the decade in which the spread of digital journals began. The first scientific periodical in digital format with peer review is considered to be New Horizons in Adult Education (later New Horizons in Adult Education and Human Resource Development), now owned by the publisher Wiley, and came into being in the autumn of 1987. The beginning of the 1990s then saw the proliferation of the digital format and the introduction of new models for the acquisition of journals by university libraries in which the publishing offer was structured in packages aggregating hundreds or even thousands of journals on a single platform subscribed to by the universities (this is the model known as the big deal)[62] . This has given researchers access to a very high number of journals compared to the traditional acquisition of the printed format. Therefore, it cannot be ruled out that the increased ease of access for researchers to a wide range of titles is one of the reasons why cases of fraud and ethical violations have become more visible. A gradual increase that has led to very high percentages of fraudulent articles in some areas such as medicine

59 Daniele Fanelli, "How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data", *PLOS One* 4(5) (2009), e5738, https://doi.org/10.1371/journal.pone.0005738.

60 Dalmeet Singh Chawla, "8% of researchers in Dutch survey have falsified or fabricated data", *Nature* 22 July 2021, https://doi.org/10.1038/d41586-021-02035-2.

61 There is also a Wikipedia page listing the most notorious cases of scientific misconduct by discipline https://en.wikipedia.org/wiki/List_of_scientific_misconduct_incidents.

62 See R. Morriello, *Lo sviluppo delle collezioni tra bibliometria e nuovi scenari dell'editoria scientifica,* cit.

and in some countries such as China, India, Egypt and Iran. So much so that the former editor of the British Medical Journal, Richard Smith, asked himself, somewhat provocatively, whether the time has not come when instead of starting from the idea that research is ethical and intact and that one should look for the rotten apple, one should not start from the idea that research is fraudulent until proven otherwise, since it is no longer just a question of one rotten apple but of entire orchards[63] .

The issue is complex and cannot only be traced back to a linear correlation with the increased accessibility of journals, nor addressed without considering other factors. The publish or perish culture drives researchers to publish in any publishing venue that allows them to do so quickly in order to increase their bibliometric and quantitative indicators. Moreover, in response to this need, predatory journals have sprung up, which do not ask too many questions about the quality of an article before publishing it and guarantee short timeframes, also because they skip the peer review step. Furthermore, in some countries, there are government agencies for the evaluation and funding of research that require the publication of publicly funded research results to be open access. Some studies attribute the increase in the number of articles published in journals considered predatory to this requirement, which has risen from an estimated 53,000 in 2010 to 420,000 in 2014, with 75% of the corresponding authors coming from India and other Asian countries and Africa[64] . In the many cases where predatory journals charge APCs, the increase in the number of articles also means more revenue for the journal.

However, the growth in the number of cases of fraudulent science has undoubtedly been one of the reasons for greater general attention to the phenomenon and the creation of bodies to monitor and bring them to light, as well as sanction them, at least in some countries. In the United States, 1989 saw the creation of the Office of Scientific Integrity (OSI) under the NIH and the Office of Scientific Integrity Review (OSIR), which in 1992 merged into the Office of Research Integrity (ORI). Also in 1989, the Whistleblower *Protection Act* was enacted (*whistleblowers* are those who report illegal and fraudulent activities), a federal law to protect the ethics and integrity of any activity carried out by public and private entities, including, therefore, academic research activity, with sanction rules for those accused of violating these principles. Since the

63 Richard Smith, "Time to assume that health research is fraudulent until proven otherwise?", *The BJM Opinion* 5 July 2021, https://blogs.bmj.com/bmj/2021/07/05/time-to-assume-that-health-research-is-fraudulent-until-proved-otherwise/.

64 Paul Newton, Katepalli Sreenivasan, "Commentary: The publication pandemic", *Physics Today*, 26 May 2021, https://physicstoday.scitation.org/do/10.1063/PT.6.3.20210526a/full/.

1980s, the issue of research integrity has therefore been at the centre of attention of the federal government, which has continued to regulate the matter, adopting a *Federal Policy on Research Misconduct in* 2000, and to keep it at the centre of its political work. In 2009, President Barack Obama's White House inaugural address contained a Memorandum on Research Integrity[65] . In 2021, Joe Biden, the 46th President of the United States of America, released a similar document entitled *Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking*[66] a few days after taking office.

What happened in the United States in the 1980s and prompted federal regulation was the discovery of some glaring cases of scientific fraud. Among these, an emblematic episode is described by Marcel C. LaFollette, who was directly involved in the affair he reports on in his book, as he was the editor of the American magazine Science, Technology, & Human Values (STHV). The case is that of John Darsee, a promising young researcher at Harvard Medical School. In 1981, Darsee was accused of manipulating and falsifying data in more than a dozen co-authored articles and more than fifty abstracts based on his research in cardiology at Harvard[67] . The news caused quite a stir and was picked up by the media mainly because it involved one of the most important universities in the United States. People began to wonder how the people involved could have been unaware of the fraud, not only the journals but especially the other authors of the articles, and it was discovered that many were honorary authors who had not participated at all in the research but had nevertheless agreed to be included among the authors. The affair opened the debate on the concept of authorship and honorary authorship and was followed in detail by the media. In particular, what was discovered outraged two biologists, Walter W. Stewart and Ned Feder of the National Institute of Arhritis, Diabetes, and Digestive and Kidney Disease (NIAD-DKD), who decided to investigate the matter further and denounce it publicly by writing an article in which they condemned the practice of honorary authorship and argued the need for those who accept such practices to take responsibility for them. The article made explicit reference to a specific group of Darsee coauthors. Considering the scientific relevance of the topic, the

65 Barack Obama, *Memorandum for the Heads of Executive Departments and Agencies*, 3-9-09, https://obamawhitehouse.archives.gov/the-press-office/memorandum-heads-executive-departments-and-agencies-3-9-09.

66 Joseph R. Biden Jr., *Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking*, Jan 27 2021, https://www.whitehouse.gov/briefing-room/presidentialactions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/.

67 *M.C. LaFollette, Stealing Into Print, cit.*

article was submitted to Nature and Cell, but Stewart and Feder were not able to get it published, partly because the manuscript had in the meantime come into the hands of Eugene Braunwald, one of the coauthors and director of the laboratory where Darsee worked, who filed a libel suit, which was followed by similar complaints from other coauthors. One of the lawyers defending the accused coauthors spread the rumour that if a journal published Stewart and Feder's article, they would face legal action. In light of this situation, LaFollette, then editor of STHV, took the trouble to inform himself in detail when, in 1985, he received the proposal to publish the article from the two NIAD-DKD biologists. The journal he edited was published in collaboration between Harvard University and the Massachusetts Institute of Technology (MIT) and by John Wiley & Sons. LaFollette explains well the delicate position in which he found himself as editor. The freedom he enjoyed as editor of the journal clashed with the risk of leading the two universities and Wiley into a lawsuit. The results of LaFollette's confrontation with the three publishers are very interesting and representative of what is at stake, even today, in situations of this kind. On the publisher's side, there was little opposition, and the publication of the article was seen as an opportunity to stir up even more debate and, after all, as a way to gain publicity and increase the journal's readership. However, on the side of the two universities, the issue was more complex. Firstly, because the two universities in the editorial agreement had guaranteed the quality and integrity of the proposed contributions. Darsee and several of the coauthors of the offending article worked at Harvard, one of the two publishing universities. Harvard's legal department would therefore have had to defend both the authors of the article accused of fraud and forgery and the journal in which the article denouncing its fraudulent nature was published. The conflict of interest was obvious and, in fact the article was not published. However, the case had meanwhile gone ahead, and Stewart and Feder were heard by a US court of law and the Committee on Science and Technology, a federal institution established in 1957 with jurisdiction over research and development and public organisations operating in this field. The Committee dropped the defamation charge and published the article in the official register of congressional proceedings, recognising that certain scientific publication practices were the consequence of an increasingly competitive climate among researchers due to organisational change in research and development activities. When the risk of legal repercussions ceased, Nature published the article, in 1987, with a defensive comment by Eugene Braunwald.

In 1988, another case was discovered that caused a stir, that of Stephen J. Breuning, a psychology researcher. Breuning had been accused of misconduct in 1983 and years later was again accused of fal-

sifying reports by the National Institute of Mental Health (NIMH), for which he was sentenced to leave his post, to pay a fine, and to community service. What is interesting to note is that the psychologist had published extensively on the subject between 1980 and 1983, and a citation analysis of his publications between 1981 and 1985, conducted, moreover, by Eugene Garfield himself, showed a considerable impact in terms of citations[68] . These were two of the many cases that surfaced in the 1980s in the United States, which were widely reported in the press and triggered action by the federal government to safeguard scientific integrity, especially as a counterbalance to the reactions of the scientific community and publishers, as both had taken defensive positions intended to minimise the incident. For the scientific community, including the funding body NIH itself, these were only sporadic events, exceptions to most good science, and in any case practices that have always been part of science. This was the opinion of the editor-in-chief of the journal Science, Daniel E. Koshland, who claimed in an editorial in 1987 that 99.99 per cent of science is good and that cases of fraud are inevitable and attributable to interdisciplinarity and the overworked scientists who were forced (even then) to work at a fast pace in order to be competitive in projects and in applying for funding[69] . Of course, his words are true, bad science remains a minority part of science, albeit a growing one, but it is certainly not to be underestimated and minimised. In general, it is difficult for those involved in scientific research to admit that there can be fraudulent and dishonest behaviour in an activity that by its very nature has as its objective the search for truth. The scientific method established over the centuries, with the endorsement of validation procedures such as peer review, tends to lead researchers to trust in the 'self-correcting power' of science and thus to hardly accept that someone might publish falsified or otherwise fraudulent articles[70] . A belief that would seem to be particularly deep rooted in universities, where there is less awareness and action than in other types of institutions. To remain in the United States, in addition to the governmental institutions mentioned above, the National Academies of Sciences, Engineering and Medicine, a grouping of private institutions, is among the organisations that pay great attention to the issue,

68 Eugene Garfield, Alfred Welljams-Dorof, "The Impact of Fraudulent Research on the Scientific Literature. The Stephen E. Breuning Case", *JAMA* 263 (1990) 10, p. 1424-1426.

69 Daniel E. Koshland, "Fraud in science", *Science* 235 (1987) 4785, p. 141, DOI: 10.1126/science.3798097;

70 Theorised among others by Michael Polanyi, see Id., "The Republic of Science: Its Political and Economic Theory", *Minerva* I (1): 1-32; Id. *La società libera. Pensieri liberali*, a cura di Massimo Baldini e Antonello Malavasi, Roma, Armando Editore.

through the constant publication of reports and guidelines on research ethics and integrity, which it makes available for open access at[71] . The association of scientific and humanities academies ALLEA has also taken action in Europe with the publication in 2017 of the *European Code of Conduct for Research Integrity*, translated into several languages, including Italian. [72]

Alongside academies and public institutions, publishers themselves must play an essential role in curbing fraud and preserving the ethical dimension of research. In contrast, as we have seen, many publishers are reluctant to admit and denounce dubious incidents relating to articles published in their journals. Although guidelines on publication ethics appear on the websites of most major scientific publishers, in practice, it is difficult to induce them to take serious action. Other publishers, however, are beginning to be concerned about a phenomenon that they recognise is on the rise and to initiate forms of collaboration to share information on cases detected so as to protect themselves, and to equip themselves with guidelines and shared practices useful for dealing with such problems. In Great Britain, the non-profit association COPE (Committee on Publication Ethics), composed of publishers, journal editors, and others affiliated with the publishing world, has been active since 2007 and has become an international point of reference on ethics and integrity issues from the point of view of what publishers and journal editors can do[73] . An evaluation checklist called REAPPRAISED was drawn up by a group of researchers, some of whom were also journal editors, after realising the difficulties of obtaining a retraction of an article and the time, sometimes even more than ten years, to do so[74] . REAPPRAISED aims to provide anyone interested with a useful and agile tool to discern the elements of a publication that may lead to suspicions about the reliability of the article and therefore possibly need to be investigated. The initiative has an important implication in that the emphasis is shifted from the author's wrongdoing, which is sometimes ambiguous and difficult to attribute, to the scientific objectivity of the article and the reliability of the research results it proposes. In Europe, several other similar associations and collaborations have emerged, some of which have merged into the ENRIO network

71 The National Academies of Sciences, Engineering, Medicine https://www.nap.edu/.

72 ALLEA, *The European Code of Conduct for Research Integrity,* 2017, https://allea.org/code-of-conduct/#toggle-id-12.

73 *COPE* https://publicationethics.org/.

74 Andrew Grey, Mark J. Bolland, Alison Avenell, Andrew A. Klein, C. K. Gunsalus, "Check for publication integrity before misconduct", *Nature* 577 (2020), p. 167-169, https://doi.org/10.1038/d41586-019-03959-6.

(The European Network of Research Integrity Offices)[75] . The list of participants in the ENRIO network shows the presence of a single Italian institution, namely the National Research Council (CNR), which since 2009 has had an ad hoc commission on research ethics and integrity issues.

In Italy, the problem has not yet become sufficiently clear. Several universities have issued ethics codes and recommendations for the integrity of scientific conduct, but these often remain on paper. However, the absence of a government institution that, as is the case in the United States and other countries, clearly defines what conduct is condemnable and what action should be taken because of violating the principles of ethics and integrity is a weakness that makes it difficult for universities to implement any action. The Italian system is based on the autonomy of universities and professors, but the integrity of research is a collective social problem, particularly when fraudulent research is conducted with government funding. The difficulties of acting in this direction are undeniable, starting with the reality that there are different perceptions within different disciplines, and practices that appear ethically unacceptable in one area may be normal in another. However, we need to start addressing this and thinking more about the future of scientific research that is being built. Fraud affects various aspects of research, not only the publication of results but also the experiments and projects for which funding is sought. In general, evaluation systems, not only in Italy, have dealt very little with these aspects, and in any case have done so unevenly[76] . In research evaluation procedures, criteria are developed that refer to what is expected and not to what is not expected and not wanted. Universities are rewarded for output on the basis of quantitative indicators that incentivise fraudulent behaviour, but without any checks on this and without any reflection on the consequences. The scientific system should self-correct through the work of its structures, but some, such as journals, have problems maintaining quality standards and are not always interested in exposing fraud and retracting articles. Moreover, even universities and other research institutions have no incentive to raise problematic cases, which can be defined as "any published or unpublished study that raises questions about the reliability of data or results, regardless of whether the study has been formally retracted or not"[77] . In general, problematic cases not only undermine the prestige and authority of the institution but may

75 ENRIO http://www.enrio.eu/.

76 Maria Teresa Biagetti, Aldis Gedutis, Lai Ma, "Ethical Theories in Research Evaluation: An Exploratory Approach", *Scholarly Assessment Reports*, 2 (2020) 1, https://doi.org/10.29024/sar.19.

77 Stephanie L. Boughton, Jack Wilkinson, Lisa Bero, "When beauty is but skin deep: dealing with problematic studies in systematic reviews", *Cochrane*

also affect the funding received or the chances of receiving it in the future. In fact, even in the face of proven cases of scientific misconduct that lead to the dismissal of the lecturer, the institution sometimes endeavours to keep the matter under wraps, as happened with the recent expulsion of physicist Eric Noji from the U.S. National Academy of Medicine (NAM), which was covered by the journal Science[78] .

The phenomenon is taking on major proportions and the consequences could be severe for the future of science and society. Scientific credibility is threatened by these incidents, and the long-term consequences are likely to be enormous. The cases of fraud now end up on the pages of newspapers and circulate quickly on the Internet. This contributes not only to the loss of trust in science per se, but to the collapse of trust in scientific and research institutions. The Covid pandemic has clearly shown us a generalised problem of trust in institutions as well as in science. We have been witnessing for some time now the crisis of intermediate bodies due to the weakening of their function,[79] and the risk, which is already very real, is that the private operators, even improvised ones, who proliferate on the Internet will step into this void. If institutions do not appear solid and ethically unassailable, there is no reason in the eyes of many people why the information coming from institutions should be any different from that coming from any source on the Internet. University systems, governments and ministries, and research evaluation agencies should stop and think carefully about what is happening.

The meticulous detection of scientific misconduct, together with the creation of institutions for verification, equipped with the regulatory tools to act, appears to be a potentially effective and impactful set of solutions vis-à-vis the public. In some countries there are forms of sanctions for misconduct, ranging from the return of funds received as research funding to the suspension of the possibility of applying for further funds, from dismissal to imprisonment, in extreme cases, and in countries where scientific fraud is considered a criminal offence. The criminalisation of scientific fraud is a current topic of heated debate[80] ,

---

Database of Systematic Reviews, 2021, Issue 6. Art. No.: ED000152, https://doi.org/10.1002/14651858.ED000152.

78 Meredith Wadman, "Top secret: U.S. National Academy of Medicine keeps expulsions quiet", *Science* 29 September 2021, https://doi.org/10.1126/science.acx9256.

79 See in this regard Giovanni Solimine, Giorgio Zanchini, *La cultura orizzontale*, Bari-Roma, Laterza, 2020.

80 Lee Harvey, "Research fraud: a long-term problem exacerbated by the clamour for research grants", *Quality in Higher Education* 26 (2020) 3, https://doi.org/10.1080/13538322.2020.1820126; William Bülow, Gert Helgesson, "Criminalization of scientific misconduct", *Medicine Health Care and Philosophy* 22 (2019), p. 245-252, https://doi.org/10.1007/s11019-018-9865-7; Francesco Aiello, "La

which in order to move in a genuinely useful and sustainable direction should involve universities, publishers, journal editors, and research evaluation agencies in a discussion. The question of ethics and integrity of scientific research should be addressed and managed upstream. Additionally, there is a need for specific training for young researchers that reinforces the foundations of the scientific method and clearly defines the areas within which ethics and integrity should be respected. On closer inspection, these principles should be the subject of training activities aimed not only at doctoral students embarking on a research pathway, but also at all university students, of any degree course, becoming part of the study curricula from the first year.

In general, the underlying problem is that the structures on which science is based are weakening. Publications are changing, the fundamental peer review process of journals is in crisis, editorial boards in a similar situation due to the difficulty of sustaining effective peer review, contributions to conferences decreasing and subject to fraudulent activity not unlike other forms by which research results are made public. The body of knowledge, the foundation on which further knowledge is built in the scientific process, is undermined by numerous distortions, ethical violations, and fraudulent acts. University structures, departments, and laboratories are increasingly dependent on quantitative evaluations, the problems of which are well known and drive deviant behaviour. We are faced with a varied set of problems that individually may not have particularly high numbers but collectively represent a strong strain on the culture of scientific research, the subsistence of the scientific method, and a major risk for the future of science. The solution is first and foremost to abandon, or at least scale back, quantitative methods for research assessment, as suggested by the important initiative on research assessment reform launched in 2022 jointly by the European Commission with EUA and Science Europe, which produced the document *Agreement on Reforming Research Assessment* that universities and other non-profit organisations are called upon to sign[81] . But it is equally imperative to try to strengthen scientific structures, starting with peer review, rethinking, and refounding them, where necessary, in order to adapt them to a changing social and communicative context, and to stem the behaviour that weakens them. Finally, spreading

frode scientifica va perseguita come un reato?*", Scienza in rete* 23/07/2014, https://www.scienzainrete.it/articolo/frode-scientifica-va-perseguita-come-reato/francesco-aiello/2014-07-23; Geoff Maslen, "Scientists sent to prison for fraudulent conduct", *University World News* 25 April 2013, https://www.universityworldnews.com/post.php?story=20130425143432184.

81 EUA (European University Association), Science Europe, Karen Stroobants, and (EC) European Commission, *Agreement on Reforming Research Assessment*, 2022, https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/reforming-research-assessment-agreement-now-final-2022-07-20_en.

awareness of these issues, both among editors and researchers perhaps through specific training within universities, so that they can understand how to recognise and avoid them is certainly essential. We hope to contribute to this goal with this volume.

References

ADDA, J., DECKER, C., OTTAVIANI, M., "P-hacking in clinical trials and how incentives shape the distribution of results across phases", *PNAS,* 117 (2020) 24, 13386-13392, https://doi.org/10.1073/pnas.1919906117.

AIELLO, F., "La frode scientifica va perseguita come un reato?", *Scienza in rete* 23/07/2014, https://www.scienzainrete.it/articolo/frode-scientifica-va-perseguita-come-reato/francesco-aiello/2014-07-23.

ALABASTER, P., BELO, A., ESPEJO, C., HAFFEMAYER, S., INFELISE, M., MOXHAM, N., RAYMOND, J., SCHOBESBERGER, N., "The Lexicons of Early Modern News"*, in News Networks in Early Modern Europe*, edited by Noah Moxham and Joad Raymond, Leiden, Brill, 2016.

ALBAKINA, A., "Publication and collaboration anomalies in academic papers originating from a paper mill: evidence from a Russia-based paper mill", deposited in arXiv 26 December 2021, version 2 revised 26 March 2022, arXiv:2112.13322v2.

ALBAKINA, A., "Unethical Practices in Research and Publishing: Evidence from Russia", *The Scholarly Kitchen*, feb 4 2021, https://scholarlykitchen.sspnet.org/2021/02/04/guest-post-unethical-practices-in-research-and-publishing-evidence-from-russia/.

ALLEA, *The European Code of Conduct for Research Integrity*, 2017, https://allea.org/code-of-conduct/#toggle-id-12.

ANDERSON, R., "Why Should We Worry about Predatory Journals? Here's One Reason", *The Source*, March 3, 2020 https://blog.cabells.com/2020/03/03/guest-post-why-should-we-worry-about-predatory-journals-heres-one-reason/.

ANDRADE, C., "HARKing, Cherry-Picking, P-Hacking, Fishing Expeditions, and Data Dredging and Mining as Questionable Research Practices"*, Journal of Clinical Psychiatry*, 82 (2021) 1, 20f13804. https://doi.org/10.4088/JCP.20f13804. PMID: 33999541.

BABBAGE, C., *Reflections on the Decline of Science in England, And on Some of Its Causes*, London, Printed for B. Fellowes and J. Booth, 1830.

BAGUES, M., SYLOS LABINI, M., ZINOVYEVA, N., "A walk on the wild side: 'Predatory' journals and information asymmetries in scientif-

ic evaluations", *Research Policy*, 48 (2019) 2, p. 462-477, https://doi.org/10.1016/j.respol.2018.04.013.

BAKER, M., "1,500 scientists lift the lid on reproducibility", *Nature* 533 (26 May 2016) p. 452–454, https://doi.org/10.1038/533452a.

BALSAMO, L., *La bibliografia. Storia di una tradizione*, Milano, Unicopli, 2017.

BANKA, R., SCI hub, "Libgen case: Delhi HC to hear students, researchers", *Hindustan Times*, Jan 7 2021, https://www.hindustantimes.com/india-news/sci-hub-libgen-case-delhi-hc-to-hear-students-researchers/story-PxBftjgAPFPfuXGok4qG7N.html.

BECK, T. S., "Image manipulation in scholarly publications: are there ways to an automated solution?", *Journal of Documentation* 78 (5) 2022.

BERCHIALLA, P., URRU, S., SCIANNAMEO, V., "The effect of COVID-19 on scientific publishing in Italy", *Epidemiologia & Prevenzione* 2021, 45 (6) novembre-dicembre, p. 449-451, https://doi.org/10.19191/EP21.6.136.

BELLER, M., "The Sokal Hoax: At Whom Are We Laughing?", *Physics Today*, 51 (1998) 9, https://doi.org/10.1063/1.882436.

BIAGETTI, M. T., GEDUTIS, A., MA, L., "Ethical Theories in Research Evaluation: An Exploratory Approach", *Scholarly Assessment Reports*, 2 (2020) 1, https://doi.org/10.29024/sar.19.

BIDEN, J. R. Jr., *Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking*, Jan 27 2021, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/.

BISHT, K., "Don't make early career researchers 'ghost authors.' Give us the credit we deserve", *Science* 9 September 2021, https://doi.org/10.1126/science.caredit.acx9061.

BODO, B., "The Genesis of Library Genesis: The Birth of a Global Scholarly Shadow Library", in *Shadow Libraries: Access to Knowledge in Global Higher Education*, edited by Joe Karaganis, Cambridge (Mass.), MIT Press, 2018, p. 25-52.

BODO, B., "Library Genesis in Numbers: Mapping the Underground Flow of Knowledge", in *Shadow Libraries: Access to Knowledge in Global Higher Education*, edited by Joe Karaganis, Cambridge (Mass.), MIT Press, 2018, p. 53-78.

Bona Fide Journals - Creating a predatory-free academic publishing environment, https://leidenmadtrics.nl/articles/bona-fide-journals-creating-a-predatory-free-academic-publishing-environment.

BOUGHTON, S. L., WILKINSON, J., BERO, L., "When beauty is but skin deep: dealing with problematic studies in systematic reviews", *Cochrane Database of Systematic Reviews*, 2021, Issue 6. Art. No.: ED000152, https://doi.org/10.1002/14651858.ED000152.

BUCCI, E., *Cattivi scienziati: la pandemia della malascienza*, Torino, Add, 2020.

BUCCI, E., "On zombie papers", *Cell Death & Disease* 10 (2019) 189, https://doi.org/10.1038/s41419-019-1450-3.

BULOW, W., HELGESSON, G., "Criminalization of scientific misconduct", *Medicine Health Care and Philosophy* 22 (2019), p. 245–252, https://doi.org/10.1007/s11019-018-9865-7.

BYRNE, A. J., CHRISTOPHER, J., "Digital magic, or the dark arts of the 21st century—how can journals and peer reviewers detect manuscripts and publications from paper mills?", *FEBS Letters* 594 (2020) 4, p. 583-589, https://doi.org/10.1002/1873-3468.13747.

CAI, X., FRY, V. C., WAGNER C. S., "International collaboration during the COVID-19 crisis: autumn 2020 developments", *Scientometrics* 126, 3683–3692 (2021), https://doi.org/10.1007/s11192-021-03873-7.

CARNELOS, L. (ed.), *I pirati dei libri. Stampa e contraffazione a Venezia tra Sei e Settecento*, Venezia, Marsilio, 2012.

CASO, R., *La rivoluzione incompiuta: la scienza aperta tra diritto d'autore e proprietà intellettuale*, Milano, Ledizioni, 2020.

CASTELLUCCI, P., "La visione del giudizio. Una prospettiva romantica", in *Libri, biblioteche e società. Studi per Rosa Marisa Borraccini*, edited by Alberto Petrucciani, Valentina Sestini, Federico Valacchi, Macerata, EUM, 2020.

CASTELLUCCI, P., *Carte del nuovo mondo: banche dati e Open Access*, Bologna, Il Mulino, 2017.

CHANDRASEKHARAM, D., "Sci-hub and Alexandra Elbakyan", *The Times of India*, June 4, 2021, https://timesofindia.indiatimes.com/blogs/dornadula-c/sci-hub-and-alexandra-elbakyan/.

CHAWLA, D. S., "8% of researchers in Dutch survey have falsified or fabricated data", *Nature* 22 July 2021, https://doi.org/10.1038/d41586-021-02035-2.

CHOPRA, V., EAGLE, K. A., "Perioperative mischief: the price of academic misconduct", *American Journal of Medicine* 125 (2012) 10, p. 953-955, https://doi.org/10.1016/j.amjmed.2012.03.014.

COBB, M., "The prehistory of biology preprints: A forgotten experiment from the 1960s", *PLOS Biology*, 16 November 2017, https://doi.org/10.1371/journal.pbio.2003995.

COPE, STM, *Paper Mills — Research report from COPE & STM 2022*, Committee on Publication Ethics, Version 1: June 2022, https://doi.org/10.24318/jtbG8IHL.

COPE, *Text recycling guidelines for editors* 2019-2021, https://publicationethics.org/text-recycling-guidelines.

COPE, *Systematic manipulation of the publishing process via 'paper mills'*, 2020, https://publicationethics.org/systematic-manipulation-paper-mills.

COPE, *Core practices*, 2017.

COPE, DOAJ, OASPA, WAME, *Principles of Transparency and Best Practice in Scholarly Publishing*, January 15, 2018, http://wame.org/principles-of-transparency-and-best-practice-in-scholarly-publishing.

CORREA, J. C., LAVERDE-ROJAS, H., TEJADA, J., MARMOLEJO-RAMOS, F., "The Sci-hub Effect on Papers' Citations", *Scientometrics* 127 (2022), p. 99-126, https://doi.org/10.1007/s11192-020-03806-w.

CRONIN, B., *The Hand of Science. Academic Writing and Its Rewards*, Lanham, The Scarecrow Press, 2005.

CYRANOSKI, D., "Stem-cell scientist found guilty of misconduct", *Nature* 01 April 2014, https://www.*Nature*.com/news/stem-cell-scientist-found-guilty-of-misconduct-1.14974.

CYRANOSKI, D., "Stem-cell method faces fresh questions", *Nature* 18 March 2014, https://www.*Nature*.com/news/stem-cell-method-faces-fresh-questions-1.14895.

DARNTON, R. (2009), *Il futuro del libro*, Milano, Adelphi, 2011.

DE GROOTE, I., et al.., "New genetic and morphological evidence suggests a single hoaxer created 'Piltdown man'", *Royal Society Open Science* 3 (2016) 8, https://doi.org/10.1098/rsos.160328.

DE VECCHIS, C., TRANIELLO, P., *La proprietà del pensiero: il diritto d'autore dal Settecento a oggi*, Roma, Carocci, 2012.

DIJCK, J. van, POELL, T., de Waal, M. (2018), *Platform society. Valori pubblici e società connessa*, Italian edition by Giovanni Boccia Artieri and Alberto Marinelli, Milano, Guerini, 2019.

DOWNES, M., "Why we should have listened to Jeffrey Beall from the start", *Learned Publishing*, 21 July 2020, https://doi.org/10.1002/leap.1316.

DOWNES, M., "Thousands of Australian academics on the editorial boards of journals run by predatory publishers", *Learned Publishing* 33 (2020), published online 19 March 2020, https://doi.org/10.1002/leap.1297.

DUC, N. M., et al.., "Predatory Open Access Journals are Indexed in Reputable Databases: a Revisiting Issue or an Unsolved Problem",

Medical Archives 74 (2020) 4, p. 318-322, PMCID: PMC7520066, https://doi.org/10.5455/medarh.2020.74.318-322.

DUNN, A. G., COIERA, E., MANDL, K. D., "Is Biblioleaks Inevitable?", *Journal of Medical Internet Research* 16 (2014) 4, https://www.jmir.org/2014/4/e112/.

ELLIOTT, T., et al.., "Perceptions on the prevalence and impact of predatory academic journals and conferences: A global survey of researchers", *Learned Publishing* 35 (2022) 4, p. 516-528, https://doi.org/10.1002/leap.1458.

ELSE, H., NOORDEN, R. van, "The fight against fake-paper factories that churn out sham science", *Nature* News Feature 23 March 2021, https://www.*Nature*.com/articles/d41586-021-00733-5.

ERRIQUEZ, M., MARCHIS, V., "Lo spionaggio tecnologico nell'Italia del Seicento e del Settecento. Il caso delle macchine della seta", Gnosis 2 (2018), p. 169-179.

EUA, Science Europe, Karen Stroobants, e European Commission, *Agreement on Reforming Research Assessment*, 2022, https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/reforming-research-assessment-agreement-now-final-2022-07-20_en.

FANELLI, D., "How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data", *PLOS One* 4(5) (2009), e5738, https://doi.org/10.1371/journal.pone.0005738.

FANG, F. C., STEEN, R. G., CASADEVALL, A., "Misconduct accounts for the majority of retracted scientific publications", *Proceedings of the National Academy of Sciences of USA*, 109 (2012) 42, p. 17028–17033, https://doi.org/10.1073/pnas.1212247109.

FIRE, M., GUESTRIN, C., "Over-optimization of academic publishing metrics: observing Goodhart's Law in action", *GigaScience*, 8 (2019) 6, p. 1-20, https://doi.org/10.1093/gigascience/giz053.

FLORIDI, L., *Il verde e il blu. Idee ingenue per migliorare la politica*, Milano, Raffaello Cortina, 2020.

FRAMPTON, B., "Clickbait: The changing face of online journalism", *BBC News* 14 September 2015, https://www.bbc.com/news/uk-wales-34213693.

FRANDSEN, T. F., "Authors publishing repeatedly in predatory journals: An analysis of Scopus articles", *Learned Publishing* early view 04 August 2022, https://doi.org/10.1002/leap.1489.

FRONTE, M., "L'imbrogliona delle staminali", *Focus* 3 luglio 2014, https://www.focus.it/scienza/salute/l-imbrogliona-delle-staminali.

GALLAGHER, J., "Stem cell scandal scientist Haruko Obokata resigns", *BBC News*, 19 December 2014, https://www.bbc.com/news/health-30534674.

GARFIELD, E., ALFRED WELLJAMS-DOROF, "The Impact of Fraudulent Research on the Scientific Literature. The Stephen E. Breuning Case", JAMA 263 (1990) 10, p. 1424-1426.

GELMAN, A., LOKEN, E., "The Statistical Crisis in Science", *American Scientist* 102 (2014) 6, https://doi.org/10.1511/2014.111.460.

GINGRAS, Y., *Bibliometrics and Research Evaluation, Uses and Abuses*, Cambridge (Ma), MIT Press, 2016.

GINGRAS, Y., KHELFAOUI, M., "Scientific publication – Is it for the benefit of the many or the few?", *University World News*, 11 July 2020, https://www.universityworldnews.com/post.php?story=20200709125857707.

GIOIA, M., "Cenni sulla pirateria libraria" in *Opere minori*, Lugano, Giuseppe Ruggia, 1837.

GODWIN, C., CLAYTON J., "Microsoft-led team retracts quantum 'breakthrough'", *BBC News* 10 March 2021, https://www.bbc.com/news/technology-56328980.

GOFF, M., "*Early History of the English Newspaper: 17th and 18th Century Burney Newspapers Collection*", Detroit, Gale, 2007, https://www.gale.com/binaries/content/assets/gale-us-en/primary-sources/intl-gps/intl-gps-essays/full-ghn-contextual-essays/ghn_essay_1718bnc_goff2_website.pdf.

GOODYEAR, D., "The stress test: Rivalries, intrigue, and fraud in the world of stem-cell research", *The New Yorker*, February 22, 2016, https://www.newyorker.com/magazine/2016/02/29/the-stem-cell-scandal.

GREENFIELD, N. M., "Why so many people mistrust science and how we can fix it", *University World News* 16 August 2022, https://www.universityworldnews.com/post.php?story=20220816072300893.

GREY, A., BOLLAND, M. J., AVENELL, A., KLEIN, A.. A., GUNSALUS, C. K., "Check for publication integrity before misconduct", *Nature* 577 (2020), p. 167-169, https://doi.org/10.1038/d41586-019-03959-6.

GRUNDNIEWICZ, A., et al., "Predatory journals: no definition, no defence", *Nature* Comment 11 December 2019, https://www.*Nature*.com/articles/d41586-019-03759-y.

GU, J., WANG, X., LI, C., ZHAO, J., FU, W., LIANG, G., QIU, J., "AI-enabled image fraud in scientific publications", *Patterns* 3 (7) 2022, https://doi.org/10.1016/j.patter.2022.100511.

HANNIGAN, T., "Beyond control: Orientalist tensions and the history of the "upas tree" myth", *The Journal of Commonwealth Literature* 55 (2020) 2, p. 173–189, first published January 29, 2018, https://doi.org/10.1177/0021989418754345.

HARVEY, L., "Research fraud: a long-term problem exacerbated by the clamour for research grants", *Quality in Higher Education* 26 (2020) 3, https://doi.org/10.1080/13538322.2020.1820126.

HEAD, M. I., HOLMAN, L., LANFEAR, R., KAHN, A. T., JENNIONS, M. D., "The Extent and Consequences of P-Hacking in Science", *PLOS Biology* Published March 13, 2015, https://doi.org/10.1371/journal.pbio.1002106.

HEATHERS, J., "The Lancet has made one of the biggest retractions in modern history. How could this happen?", *The Guardian* June 5, 2020, https://www.theguardian.com/commentisfree/2020/jun/05/lancet-had-to-do-one-of-the-biggest-retractions-in-modern-history-how-could-this-happen.

HUMBOLDT, W. von (1809-10), *L'organizzazione interna ed esterna degli istituti scientifici superiori a Berlino*, Italian translation by M.C. Pievatolo, 2017, https://archiviomarini.sp.unipi.it/735.

HUSTEN, L, "Prominent Dutch Cardiovascular Researcher Fired for Scientific Misconduct", *CardioBrief*, 17 November 2011, https://www.cardiobrief.org/2011/11/17/prominent-dutch-cardiovascular-researcher-fired-for-scientific-misconduct/.

HVISTENDAHL, M., "China's Publication Bazaar", *Science* 342 (2013) 6162, p. 1035-1039, https://doi.org/10.1126/science.342.6162.1035.

ICMJE, *Defining the Role of Authors and Contributors*, 2021, http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

IOANNIDIS, J. P. A., "Hundreds of thousands of zombie randomised trials circulate among us", *Anaesthesia* 76 (2021), p. 444–447, https://doi.org/10.1111/anae.15297.

JOHNS, A. (2009), *Piracy. The Intellectual Property Wars from Gutenberg to Gates, Chicago, The University of Chicago Press, 2009*.

KARAGANIS, J., "Introduction: Access from Above, Access from Below", in *Shadow Libraries: Access to Knowledge in Global Higher Education*, edited by Joe Karaganis, Cambridge (Mass.), MIT Press, 2018, p. 1-24.

KANT, I. (1785), *L'illegittimità della ristampa dei libri*, Italian translation by Maria Chiara Pievatolo, Bollettino telematico di filosofia politica, https://btfp.sp.unipi.it/dida/kant_7/ar01s06.xhtml.

KERR, N.L., "HARKing: hypothesizing after the results are known", *Personality and Social Psychology Review* 2 (1998) 3, p. 196-217, https://doi.org/10.1207/s15327957pspr0203_4. PMID: 15647155.

KING, R. S., "The Manuscript Newsletter and the Rise of the Newspaper, 1665–1715", *Huntington Library Quarterly*, 79 (Autumn 2016) 3, p. 411-437, https://www.jstor.org/stable/10.2307/huntlibrquar.79.3.411.

KOSHLAND, D. E., "Fraud in science", *Science* 235 (1987) 4785, p. 141, DOI: 10.1126/science.3798097.

KUHN, T. S. (1962), *La struttura delle rivoluzioni scientifiche*, Torino, Einaudi, 2009.

KULCZYCKI, E., HOLOWIECKI, M., TASKIN, Z., DOGAN, G., "Questionable conferences and presenters from top-ranked universities", *Journal of Information Science*, preprint 10.31219/osf.io/guaw3 in *OSF Preprints*, April 04, 2022, Last edited June 29, 2022, Peer-reviewed Publication, https://doi.org/10.1177/01655515221087674.

KWON, D., "Nature's 10. Ten people who helped shape science in 2021. Guillaume Cabanac: Deception sleuth", *Nature* 15 December 2021, https://www.Nature.com/immersive/d41586-021-03621-0/index.html#section-gM9iO4XBRl.

LABBE, C., LABBE, D., "Duplicate and fake publications in the scientific literature: how many SCIgen papers in computer science?", *Scientometrics* 94 (2013), p. 379–396, https://doi.org/10.1007/s11192-012-0781-y.

LAFOLLETTE, M. C., *Stealing Into Print. Fraud, Plagiarism and Misconduct in Scientific Publishing*, Berkeley and Los Angeles, University of California Press, 1992.

LAINE, C., WINKER, M. A., *Identifying Predatory or Pseudo-Journals*, WAME February 18, 2017, https://www.wame.org/identifying-predatory-or-pseudo-journals#Table%204.

LANGLAIS, P., "Quand les articles scientifiques ont-ils cessé d'être des communs?", *Sciences communes* 11 mars 2015, https://scoms.hypotheses.org/409.

LÜSCHER, T., "The codex of science: honesty, precision, and truth—and its violations", *European Heart Journal* 34 (2013) 14, p. 1018–1023, https://doi.org/10.1093/eurheartj/eht063.

MALLAPATY, S., "China Bans Cash Rewards for Publishing", *Nature* News 579 (2020) 5 March, https://www.*Nature*.com/articles/d41586-020-00574-8.

MASLEN, G., "Scientists sent to prison for fraudulent conduct", University World News 25 April 2013, https://www.universityworldnews.com/post.php?story=20130425143432184.

MANTIONE, C., "Violazioni del diritto d'autore: le conseguenze", Altalex 14/05/2021, https://www.altalex.com/guide/violazioni-del-diritto-d-autore.

MARCHETTI, P., "Il Contributo italiano alla storia del Pensiero – Diritto: Cesare Lombroso", in *Enciclopedia Treccani,* 2012, https://www.treccani.it/enciclopedia/cesare-lombroso_%28Il-Contributo-italiano-alla-storia-del-Pensiero:-Diritto%29/.

MARCO-CUENCA, G., SALVADOR-OLIVAN, J., ARQUERO-AVILES, R., FAGGIOLANI, C., SISO-CALVO, B., "Scientific publications of affiliated authors in Italy retracted due to fraud. Review and analysis", *JLIS.it* 12 (2021) 2, http://dx.doi.org/10.4403/jlis.it-12711.

MARCUS, A., "Publisher offers cash for citations", *Retraction Watch* August 31, 2021, https://retractionwatch.com/2021/08/31/publisher-offers-cash-for-citations/.

MARCUS, A., "Beam us up! Elsevier pulls 26 Covid-19 papers by researcher with a penchant for Star Trek", *Retraction Watch* March 31 2021, https://retractionwatch.com/2021/03/31/beam-us-up-elsevier-pulls-26-covid-19-papers-by-researcher-with-a-penchant-for-star-trek/.

MARCUS, A., FUJII, Y., "Zombie papers: Why do papers by the most prolific fraudster in history keep getting cited?", *Retraction Watch* April 1 2020, https://retractionwatch.com/2020/04/01/zombie-papers-why-do-papers-by-the-most-prolific-fraudster-in-history-keep-getting-cited/.

MARCUS, A., "Researcher who stole manuscript during peer review earns second retraction", *Retraction Watch* August 1 2017, https://retractionwatch.com/2017/08/01/researcher-stole-manuscript-peer-review-earns-second-retraction/.

MARCUS, A. "Critics of Poldermans' work baffled by NEJM stance on DECREASE papers", *Retraction Watch* September 29, 2014, https://retractionwatch.com/2014/09/29/critics-of-poldermans-work-baffled-by-nejm-stance-on-decrease-papers/.

MARCUS, A., "Influential Reinhart-Rogoff economics paper suffers spreadsheet error", *Retraction Watch* April 18, 2013, https://retractionwatch.com/2013/04/18/influential-reinhart-rogoff-economics-paper-suffers-database-error/.

McCOOK, A., "Congrats! Your paper was accepted. (Except if the acceptance letter was forged)", *Retraction Watch* June 5 2018, https://retractionwatch.com/2018/06/05/congrats-your-paper-was-accepted-or-was-it/.

McCOOK, A., "Does the philosophy literature have a plagiarism problem?", *Retraction Watch* May 19 2017, https://retractionwatch.com/2017/05/19/philosophy-literature-plagiarism-problem/.

McCOOK, A., "7 signs a scientific paper's authorship was bought", *Retraction Watch* Oct 24 2016, https://retractionwatch.com/2016/10/24/seven-signs-a-paper-was-for-sale/.

McCOOK, A., "*Top 10 most highly cited retracted papers*", July 14 2015, https://retractionwatch.com/2015/07/14/half-of-anesthesiology-fraudsters-papers-continue-to-be-cited-years-after-retractions/.

McGREAL, C., "Columbia whistleblower on exposing college rankings: 'They are worthless'", *The Guardian* 16 September 2022, https://www.theguardian.com/us-news/2022/sep/16/columbia-whistleblower-us-news-rankings-michael-thaddeus.

MERTON, R. K. (1971), *Teoria e struttura sociale III. Sociologia della conoscenza e sociologia della scienza*, Bologna, Il Mulino, 2000.

MOHER, D., et al., "Stop this waste of people, animals and money", *Nature* 549 (2017) 7670, https://www.*Nature*.com/news/stop-this-waste-of-people-animals-and-money-1.22554.

MOMEN, H., "Jeffrey Beall and Blacklists", *SciELO in Perspective* 2015, https://blog.scielo.org/en/2015/08/04/jeffrey-beall-and-blacklists/.

MORRIELLO, R. *Le raccolte bibliotecarie digitali nella società dei dati*, Milano, Editrice Bibliografica, 2020.

MORRIELLO, R., "Lo sviluppo delle collezioni tra bibliometria e nuovi scenari dell'editoria scientifica", Biblioteche oggi Trends 4 (2018) 2, p. 39-47, http://www.bibliotecheoggi.it/trends/article/view/854.

MOUNK, Y., "What an Audacious Hoax Reveals About Academia", *The Atlantic* October 5, 2018, https://www.theatlantic.com/ideas/archive/2018/10/new-sokal-hoax/572212/.

MOUSSA, S., "Journal hijacking: Challenges and potential solutions", *Learned Publishing*, first online 26 July 2021, https://doi.org/10.1002/leap.1412.

MRC, Policy and Procedure for Investigating Allegations of Research Misconduct, 2014, https://www.ukri.org/wp-content/uploads/2021/08/MRC-12082021-Research-Misconduct-Policy.pdf.

MYKLEBUST, J. P., "Netherlands: World heart research expert fired", *University World News* 23 November 2011, https://www.universityworldnews.com/post.php?story=20111123173136776.

THE NATIONAL SCIENCE FOUNDATION, *Publications Output: U.S. Trends and International Comparisons*, 2019, https://ncses.nsf.gov/pubs/nsb20206/executive-summary.

THE NATIONAL SCIENCE FOUNDATION, THE INSTITUTE OF EDUCATION SCIENCES, U.S. DEPARTMENT OF EDUCATION, *Companion Guidelines on Replication & Reproducibility in Education Research*, Nov 28, 2018, https://www.nsf.gov/pubs/2019/nsf19022/nsf19022.pdf.

NATURE, "How Nature contributed to science's discriminatory legacy", *Nature* 609 (2022), p. 875-876, https://doi.org/10.1038/d41586-022-03035-6.

NEWTONS, P., SREENIVASAN, K., "Commentary: The publication pandemic", *Physics Today* 26 May 2021, https://physicstoday.scitation.org/do/10.1063/PT.6.3.20210526a/full/.

NICHOLSON, J., LAZEBNIK, Y., "The R-Factor: A Measure of Scientific Veracity", *The Winnower* August 17, 2014, https://doi.org/10.15200/winn.140832.20404.

NOORDEN, R. van, "Journals adopt AI to spot duplicated images in manuscripts", *Nature News* 21 December 2021, https://www.*Nature*.com/articles/d41586-021-03807-6.

NOORDEN, R. van, "Publishers launch joint effort to tackle altered images in research papers", *Nature News* 13 May 2020, https://www.*Nature*.com/articles/d41586-020-01410-9.

NOORDEN, R. van, "Publishers withdraw more than 120 gibberish papers", *Nature* 24 February (2014), https://doi.org/10.1038/*Nature*.2014.14763.

OBAMA, B., *Memorandum for the Heads of Executive Departments and Agencies*, 3-9-09, https://obamawhitehouse.archives.gov/the-press-office/memorandum-heads-executive-departments-and-agencies-3-9-09.

ONG, W. J. (1982), *Oralità e scrittura. Le tecnologie della parola*, Bologna, Il Mulino, 1986.

"Open letter. Call for a Full and Unrestricted International Forensic Investigation into the Origins of COVID-19", https://int.nyt.com/data/documenttools/virus-inquiries-pandemic-origins/d7a097a4c758a65a/full.pdf.

OPEN SCIENCE COLLABORATION, "Estimating the reproducibility of psychological science", *Science* 349 (2015) 6251, https://doi.org/10.1126/science.aac4716.

ORANSKY, I., "Physics publisher retracting nearly 500 likely paper mill papers", *Retraction Watch* September 9 2022 https://retractionwatch.com/2022/09/09/physics-publisher-retracting-nearly-500-likely-paper-mill-papers.

ORANSKY, I., "Andrew Wakefield's fraudulent paper on vaccines and autism has been cited more than a thousand times. These researchers tried to figure out why", *Retraction Watch* 18 November 2019,

https://retractionwatch.com/2019/11/18/andrew-wakefields-fraudulent-paper-on-vaccines-and-autism-has-been-cited-more-than-a-thousand-times-these-researchers-tried-to-figure-out-why/.

ORANSKY, I., "Engineering prof in Italy earns 26 retractions in one fell swoop", *Retraction Watch* November 22, 2018, https://retractionwatch.com/2018/11/22/engineering-prof-in-italy-earns-26-retractions-in-one-fell-swoop/.

ORANSKY, I., "Astrophysics retraction trail includes paper that plagiarized another already retracted for…plagiarism", *Retraction Watch* May 18, 2012, https://retractionwatch.com/2012/05/18/astrophysics-retraction-trail-includes-paper-that-plagiarized-another-already-retracted-for-plagiarism/.

OWENS, B., "Reliability of 'new drug target' claims called into question", *Nature News blog* 05 Sept 2011, http://blogs.*Nature*.com/news/2011/09/reliability_of_new_drug_target.html.

PACCHIONI, G., *The Overproduction of Truth, Passion, Competition, and Integrity in Modern Science*, Oxford, Oxford University Press, 2018.

PAGLIONE, L., *Contributorship Open Badges on ORCID*, October 20, 2015, https://info.orcid.org/contributorship-open-badges-on-orcid/.

PARISI, G. *In un volo di storni. Le meraviglie dei sistemi complessi*, Milano, Rizzoli, 2021.

PECORARI, D., "Predatory Conferences: What Are the Signs?", *Journal of Academic Ethics* 19 (2021), p. 343–361, https://doi.org/10.1007/s10805-021-09406-4.

PIEGORSCH, W. W., "Fisher's Contributions to Genetics and Heredity, with Special Emphasis on the Gregor Mendel Controversy", *Biometrics* 46 (1990) 4, p. 915-924, https://doi.org/10.2307/2532437.

PILLER, C., "Disgraced COVID-19 studies are still routinely cited", *Science* 371 (2021) 6527, p. 331-332, https://doi.org/10.1126/science.371.6527.331.

PILLER, C., SERVICK, K., "Two elite medical journals retract coronavirus papers over data integrity questions", *Science* Jun. 4, 2020, https://www.sciencemag.org/news/2020/06/two-elite-medical-journals-retract-coronavirus-papers-over-data-integrity-questions.

POLANYI, M., *La società libera. Pensieri liberali*, edited by Massimo Baldini e Antonello Malavasi, Roma, Armando Editore.

POLANYI, M.., "The Republic of Science: Its Political and Economic Theory", *Minerva* I (1): 1-32.

POPPER, K. (1945), *La società aperta e i suoi nemici*, volume unico, Roma, Armando Editore, 2018.

POPPER, K. (1934), *Logica della scoperta scientifica. Il carattere autocorrettivo della scienza*, Torino, Einaudi, 2010.

PRICE, D. (1963), *Sociologia della creatività scientifica*, Milano, Bompiani, 1967.

RADICK, G., "Beyond the 'Mendel-Fisher Controversy'", *Science* 350 (2015) 6257, p. 159-160, https://www.science.org/doi/full/10.1126/science.aab3846.

RASKO, J., POWER, C., "What pushes scientists to lie? The disturbing but familiar story of Haruko Obokata", *The Guardian* 18 Feb 2015, https://www.theguardian.com/science/2015/feb/18/haruko-obokata-stap-cells-controversy-scientists-lie.

RCUK, Policy and Guidelines on Governance of Good Research Conduct, 2013, https://www.ed.ac.uk/files/imports/fileManager/RCUKPolicyandGuidelinesonGovernanceofGoodResearchPracticeFebruary2013.pdf.

REDDEN, E., "Rush to Publish Risks Undermining COVID-19 Research", *Inside Higher Ed* June 8, 2020, https://www.insidehighered.com/news/2020/06/08/fast-pace-scientific-publishing-covid-comes-problems.

RESNIK, D. B., NEAL, T., RAYMOND, A., KISSLING, G. E., "Research Misconduct Definitions Adopted by U.S. Research Institutions", *PubMed Central* 1 January 2015, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4269469/.

RETRACTIONWATCH, "Half of anesthesiology fraudster's papers continue to be cited years after retractions", *Retraction Watch* December 2020, https://retractionwatch.com/the-retraction-watch-leaderboard/top-10-most-highly-cited-retracted-papers/.

RITCHIE, S., *Science Fictions. Exposing Fraud, Bias, Negligence and Hype in Science*, London, The Bodley Head, 2020.

RIVA, O., "Università e ranking, lo scandalo della Columbia: ha fornito dati «ingannevoli». Ora è scesa dal secondo al 18esimo posto", *Corriere della Sera* 13 settembre 2022, https://www.corriere.it/scuola/rientro-a-scuola/notizie/universita-ranking-scandalo-columbia-ha-fornito-dati-ingannevoli-ora-scesa-secondo-18esimo-posto-45c9e7da-3343-11ed-80fb-2302675b77bf.shtml.

ROBINSON, C. E., "Mary Shelley and the Roger Dodsworth Hoax", *Keats-Shelley Journal*, 24 (1975), p. 20-28.

RSC, "*Advances retractions"*, https://www.rsc.org/news-events/articles/2021/jan/paper-mill-response.

SAMOTA, E. K., DAVEY, R. P., "Knowledge and Attitudes Among Life Scientists Toward Reproducibility Within Journal Articles: A Re-

search Survey", *Frontiers in Research Metrics and Analytics* 29 June 2021, https://doi.org/10.3389/frma.2021.678554.

SAMPLE, I., "How computer-generated fake papers are flooding academia", *The Guardian* 24 Feb 2014, https://www.theguardian.com/technology/shortcuts/2014/feb/26/how-computer-generated-fake-papers-flooding-academia.

San Francisco Declaration On Research Assessment (DORA), https://sfdora.org/read.

SCHENTINGER, V., et al., "Image forgery detection confronts image composition", *Computers & Graphics* 68 (2017), https://doi.org/10.1016/j.cag.2017.08.014.

SCHULZ, W. G., "A Massive Case Of Fraud", *Chemical & Engineering News* 86 (2008) 7, https://cen.acs.org/magazine/86/8607.html.

SERVICK, K., ENSERINK, M., "A mysterious company's coronavirus papers in top medical journals may be unraveling", *Science* Jun. 2, 2020, https://www.sciencemag.org/news/2020/06/mysterious-company-s-coronavirus-papers-top-medical-journals-may-be-unraveling.

SHEN, H., "Meet this super-spotter of duplicated images in science papers", *Nature* 13 May 2020, https://www.*Nature*.com/articles/d41586-020-01363-z.

SHUCHMAN, M., "False images top form of scientific misconduct", *CMAJ: Canadian Medical Association Journal* 188 (2016) 9, p. 645, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4902687/.

SIGNORILE, L., "Un'apologia di Giuseppe Gioeni – naturalista", Le Science Blog 16 luglio 2021, http://lorologiaiomiope-lescienze.blogautore.espresso.repubblica.it/2021/07/16/unapologia-di-giuseppe-gioeni-naturalista.

SMITH, R., "Time to assume that health research is fraudulent until proven otherwise?", *The BJM Opinion* 5 July 2021, https://blogs.bmj.com/bmj/2021/07/05/time-to-assume-that-health-research-is-fraudulent-until-proved-otherwise/.

SOLIMINE, G., ZANCHINI, G., *La cultura orizzontale*, Bari-Roma, Laterza, 2020.

STEINBERG, S. (1951), *Cinque secoli di stampa*, Torino, Einaudi, 1982.

STERN, A. M., CASADEVALL, A., STEEN, R. G., FANG, F. C., "Financial costs and personal consequences of research misconduct resulting in retracted publications", *Elife* 2014 3:e02956, published 2014 Aug 14., https://doi.org/10.7554/eLife.02956.

STM WORKING GROUP ON IMAGE ALTERATION AND DUPLICATION DETECTION, *Recommendations for handling image integrity issues*, Final draft, July 2021, https://osf.io/kgyc6/.

TRANFAGLIA, N., *Editori italiani ieri e oggi*, Bari, Laterza, 2001.

VERHULST, B., "In Defense of P Values", *AANA Journal* 84 (2016) 5, p. 305-308, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5375179/.

VILLA, M. L., *Scienza è democrazia. Come funziona il mondo della ricerca*, postfazione di Pietro Greco, Milano, Guerini e Associati, 2018.

VITTU, J., "Du Journal des savants aux Mémoires pour l'histoire des sciences et des beaux-arts: l'esquisse d'un système européen des périodiques savants", *Dix-septième siècle*, 228 (2005) 3, p. 527-545, https://doi.org/10.3917/dss.053.0527.

WADMAN, M., "Top secret: U.S. National Academy of Medicine keeps expulsions quiet", *Science* 29 September 2021, https://doi.org/10.1126/science.acx9256.

WADMAN, M., "Scientists quit journal board, protesting 'grossly irresponsible' study claiming COVID-19 vaccines kill", *Science* Jul. 1, 2021, https://www.sciencemag.org/news/2021/07/scientists-quit-journal-board-protesting-grossly-irresponsible-study-claiming-covid-19.

WORLD ASSOCIATION OF MEDICAL EDITORS, *Identifying Predatory or Pseudo-Journals*, February 18, 2017, https://www.wame.org/identifying-predatory-or-pseudo-journals.

YAN, W., "Coronavirus Tests Science's Need for Speed Limits", *New York Times* 14 April 2020, https://www.nytimes.com/2020/04/14/science/coronavirus-disinformation.html.

Some of the problems characterized by contemporary academic communication are widely believed to be rooted in recent times. Issues such as predatory publishing and violations of research ethics and integrity are often considered to be the result of the expansion of digital publishing, especially open access. Although open access increases the risk of illegal or ethically wrong behaviour under certain conditions, access to research results increases the visibility and dissemination of publications. Not only is this a means of detecting bias cases that are otherwise more difficult to detect. A number of other critical issues related to the ethics and integrity of scientific research are rising, but also become apparent thanks to the potential of technology, such as predator magazines and conferences, paper factories that make scientific papers, zombies that emerge from bad management. The book describes harmful practices in scientific communication, destroying its credibility today, and draws backwards, showing the line of continuity with the past centuries. The integration of media history, the emphasis on book piracy, and the debate on intellectual freedom, which led to the establishment of the first copyright law and ignited Enlightenment, were crucial in modern times and were considered a possible transition to a new paradigm.

Rossana Morriello is Researcher in Library Science at the University of Florence, Department of History, Archaeology, Geography, Art and Performing Arts (SAGAS), where she teaches Library Science, Digital Library Science, Cataloguing and Metadata. PhD in Documentary, Linguistic and Literary Science, Book and Documentary Science curriculum, at the Sapienza University of Rome; magistral degree cum laude in Archivistics and Library Science at the Ca' Foscari University of Venice and in Foreign Languages and Literatures at the University of Turin. She is author or co-author of more than 150 publications including books, articles in national and international journals, contributions in books and conference proceedings.

www.ledizioni.it

9 791256 002948

€ 24,00