Research Integrity: Understanding our shared responsibility for a sustainable scholarly ecosystem

Martin Szomszor – BSc, PhD
Director, Institute for Scientific Information (ISI)

July 2021
The report encourages everyone involved in research to broaden their view of what it means to conduct research with integrity and to consider how certain research evaluation instruments and incentive mechanisms are leading to a rise in deviant publication behaviour.

The traditional focus on fabrication, falsification and plagiarism is no longer enough – new forms of manipulation are emerging as some stakeholders seek an unfair advantage.

Our report is intended as a guide:
- to expose the range of tactics used
- to describe our varied and collaborative responsibilities
- to highlight current and future technological enhancements that will help us all uphold the principles of research integrity.

https://clarivate.com/webofsciencegroup/campaigns/research-integrity-understanding-our-shared-responsibility-for-a-sustainable-scholarly-ecosystem/
Why is research integrity important?

- Scholarly publishing has a longstanding, crucial role in the communication of academic research, providing the essential substrate on which ideas can be exchanged, criticized, and improved.
- Much of the value of research is attributed to a shared ideology of integrity – the notion that honest, ethical behaviour coupled with sound methodology and rigorous peer review leads to results that can be trusted, replicated, and built on.
- Scholarly publication is used to communicate findings and form a shared record of human knowledge.
- Because the quality of research is often judged by examining this record, some actors seek short-cuts in order to gain an unfair advantage.

“If I have seen further it is by standing on the shoulders of Giants” – Newton, 1676
Brief History

• Misconduct became a topic of significant public interest in the 1970s due to a number of high-profile cases. Notably in 1974, William Summerlin used a permanent marker to darken a patch of skin transplanted between mice.

• The first hearing on the topic took place in 1981 by the Investigations and Oversight Subcommittee of the House of Science and Technology Committee (US).

• Since the 1980s, various institutions have been established to tackle the issues relating to research integrity, including The Office of Research Integrity in the US and the UK Research Integrity Office.

• Most agencies that fund research have established their own code of conduct, professional bodies and societies provide members with their guidance, and a range of intergovernmental organizations contribute recommendations (See appendix).
Growing topic of interest

- Publishers, Editors and Reviewers play a pivotal role in upholding the integrity of the publication record since they manage the review of submitted manuscripts and determine which are suitable for publication.
- In recent years, this has become a new ethical battleground as many have realized opportunities to exploit the process to their own advantage.
- The research community are increasingly engaged on this topic, publishing more papers in recent years on issues relating to Research Integrity.
Why is research integrity undermined?

"It is almost certain that misconduct has always been a feature of scientific research" – Lock, 1994

- The pressure to perform is a key influencing factor making assessment a critical component of the machinery that influences individual behaviour
- Various forms of measurement are taken, such as funding won, journals published in or citations received, and these figures are used to inform a variety of decision makers

| **Researchers** | want to improve their standing through the publication of many research articles in high-quality journals that receive a high number of citations. This improves their chances of receiving funding, enables them to take on better positions (institutional, editorial, advisory) and generally ensures longevity of their career. |
| **Journals** | want to attract and publish the very best research articles in their field or increase their publication volumes to ensure their profitability, long-term sustainability and growth in readership. |
| **Publishers** | want to build a portfolio of successful journals, possibly specialized by field, access model, threshold for acceptance or otherwise. |
| **Funders** | want to invest money in the teams and projects that will deliver high-impact outcomes. |
| **Institutions** | want to attract, develop, promote, and retain academics that produce world-leading research with wide socio-economic benefits. In turn, a better research profile improves their standing in rankings which bolsters student applications, increases alumni support, and enables recruitment of first-rate faculty. |
| **Governments** | want to build and invest in productive research systems with high quality governance that deliver political, economic, and cultural advantage. |
| **Database and Analytics providers** | seek to provide useful search and discovery features that help researchers work quickly and more efficiently, and to provide analytical tools (including metrics and indicators) that support research evaluation use-cases. |
We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity.

Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.

What are the different types of behaviour that undermine research integrity?
We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity.

Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.
What are the different types of behaviour that undermine research integrity?

We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity.

Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.
We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity. Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.
What are the different types of behaviour that undermine research integrity?

We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity.

Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.
What are the different types of behaviour that undermine research integrity?

We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity.

Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.
We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity. Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.

What are the different types of behaviour that undermine research integrity?

A. Research problem, literature review, hypothesis and plan
B. Research, experiment, and data collection
C. Data analysis, hypothesis testing, data preservation
D. Manuscript preparation
E. Choice of publication venue and submission
F. Editorial and peer review process, including revisions
G. Publication
H. Use of the publication record
What are the different types of behaviour that undermine research integrity?

We breakdown the research process into a number of stages and document the range of behaviours that can undermine research integrity.

Many of these are not captured under existing classifications (e.g. fabrication, falsification, and plagiarism – FFP, and questionable research practices - QRPs) including citation manipulation and misrepresentation.
Shared Responsibility

- There are many stakeholders responsible for upholding research integrity and there is no single group that can fix failures in research integrity - there are no citation police.
- It is a shared responsibility that requires each actor to seek out information on how to identify and tackle problems of misconduct, many of which will be varied depending on their role.
Researchers & Publishers
authors, co-authors, reviewers, editorial boards

<table>
<thead>
<tr>
<th>Perform literature reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Check for duplication of work and plagiarism</td>
</tr>
<tr>
<td>- Ensure appropriate credit is given to predecessors</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Confirm references are legitimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Check whether cited works have been corrected or retracted</td>
</tr>
<tr>
<td>- Ensure referenced material is relevant and avoid superfluous cited references</td>
</tr>
<tr>
<td>- Maintain a level of self-citation that is appropriate to the field or discipline</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ensure the provenance of any experimental data</th>
</tr>
</thead>
<tbody>
<tr>
<td>- While many cases of misconduct are deliberate, careless management of source data can lead to accidental use of dubious material</td>
</tr>
<tr>
<td>- Check for signs of image manipulation and falsification of data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uphold statistical validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Check that all variables are accounted for and that the expected statistical tests are in place</td>
</tr>
<tr>
<td>- When possible, verify published data against third-party sources such as clinical trial result databases</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verify author identities &amp; affiliations</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Make sure the names appearing on papers are for real individuals, the affiliations given are correct, and the named organizations are not fabricated</td>
</tr>
</tbody>
</table>
## Researchers & Publishers

### authors, co-authors, reviewers, editorial boards

<table>
<thead>
<tr>
<th><strong>Validate contributions</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Ensure that listed authors did contribute to the research</td>
<td></td>
</tr>
<tr>
<td>- Provide guidance on what constitutes an author, especially for highly collaborative works</td>
<td></td>
</tr>
<tr>
<td>- Consider group authorships</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Perform proper peer-review</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Ensure peer-review is not fake or self-directed</td>
<td></td>
</tr>
<tr>
<td>- Check suitability of suggested peer-reviewers</td>
<td></td>
</tr>
<tr>
<td>- Verify conflicts of interest</td>
<td></td>
</tr>
<tr>
<td>- Identify and suppress coercive behavior, e.g. when suggesting additional references</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Check Journal identities and validity</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Be aware that those soliciting submissions may not be from the journal advertised either because of blatant hijacking or subtle refactoring of journal names</td>
<td></td>
</tr>
<tr>
<td>- Avoid submitting to, reviewing for, or serving on editorial boards of journals that do not uphold basic scholarly standards</td>
<td></td>
</tr>
<tr>
<td>- Proactively identify and exclude low quality and fake sources when selecting content and indexing data</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Perform plagiarism detection</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Use software tools to enhance plagiarism detection capabilities</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Screen images</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Provide clear policies on what constitutes image manipulation</td>
<td></td>
</tr>
<tr>
<td>- Use experts to screen images prior to publication</td>
<td></td>
</tr>
</tbody>
</table>
Institutions & Funders

**Researcher training & enforcement**
- Train researchers on fundamental skills including literature review, manuscript preparation, and peer-review
- Create clear policies regarding expected behavior, monitor researcher activity, and take punitive action where appropriate
- Be accountable to funders and governments

**Community Wide**
Including database and metrics providers

**Check Journal identities and validity**
- Be aware that those soliciting submissions may not be from the journal advertised either because of blatant hijacking or subtle refactoring of journal names
- Avoid submitting to, reviewing for, or serving on editorial boards of journals that do not uphold basic scholarly standards
- Proactively identify and exclude low quality and fake sources when selecting content and indexing data

**Produce and use bibliometric indicators responsibly**
- Produce responsible metrics and withhold scores when anomalous behavior is identified
- Use metrics to support decision making, not supplant it
- Seek out multidimensional quantitative and qualitative indicators
- Consider incentives created by evaluative frameworks carefully and the effects they will have on researchers
How technology, data & analytics can help

Researcher Self-citation analysis

• The rate at which any individual or group (i.e., journal institution, region) references itself is a topic that has received continuous interest since the 1960s (Kaplan 1965)

• We face this problem head on when compiling our annual list of Highly Cited Researchers
  
  — Factors such as retractions, misconduct, and extreme self-citation—all of which would detract from true community-wide research influence—may lead to an author being excluded or suppressed from the list.

• Our paper in Scientometrics (Szomszor et al 2020) describes how we approach this problem using a combination of statistical analysis and visualization

---

Another issue that must be tackled by a database and metrics providers is that of journal-level citation manipulation.

DORA states:

- **For organizations that supply metrics** (13) Be clear that inappropriate manipulation of metrics will not be tolerated; be explicit about what constitutes inappropriate manipulation and what measures will be taken to combat this.

Since 2004, the Journal Citation Reports (JCR) have suppressed publication of a Journal Impact Factor (JIF) for journals with anomalous citation patterns.

More complex behaviour has emerged where combinations of journals coordinate efforts, often evident through shared editorial board membership.

How technology, data & analytics can help
Spotting coordinated journal citation manipulation
How technology, data & analytics can help

Other areas of interest

• **Plagiarism detection**
  – Technology originally developed for educational settings has been successfully repurposed for analysis of academic text
  – Difficult to identify cross-language plagiarism
  – Citation-based approaches have been proposed that may provide additional utility

• **Image Manipulation**
  – Image screening software now used in the editorial process, mostly in the biomedical domain
  – Much of the work is still manual - expert detectives search for tell-tale signs such as cropping, selective enhancement, colour adjustment, cloning, fabrication, etc

• **Anomalous reviewer activity**
  – Software can be used to track reviewer activity (e.g. through IP address tracking and browser fingerprinting) to help identify spoofing and fake reviews.
  – Open Peer Review platforms (e.g. Publons) and advanced submission software (e.g. Scholar One) will both provide valuable information to editorial teams to help them ensure high quality review standards

• **Authoring tools**
  – By incorporating up-to-date information on retractions as well as journal-level metrics (both quantitative and qualitative), authors will have better tools to ensure their manuscripts are high quality
  – Could be used to help authors identify unusual rates of self-citation
What’s in store for the future?

• Collective, proactive effort is required to address multifarious opportunities for misconduct
• Increasing Open Access publishing will change how budgets for publication are allocated (from library subscriptions to institutional & funder mandates)
  – Sanctioned journal lists may play a crucial role in deciding where researchers are allowed to publish
  – Transparent evaluation criteria will be essential
• Since many of the subversive behaviours discussed are based around misrepresenting identity, blockchain may provide a useful framework for identity management
• Research evaluation must evolve to avoid goal displacement and reduce the incentive for gaming
Thank you for your attention

Martin Szomszor
Martin.Szomszor@Clarivate.com

clarivate.com/webofsciecegroup/solutions/isi-institute-for-scientific-information/