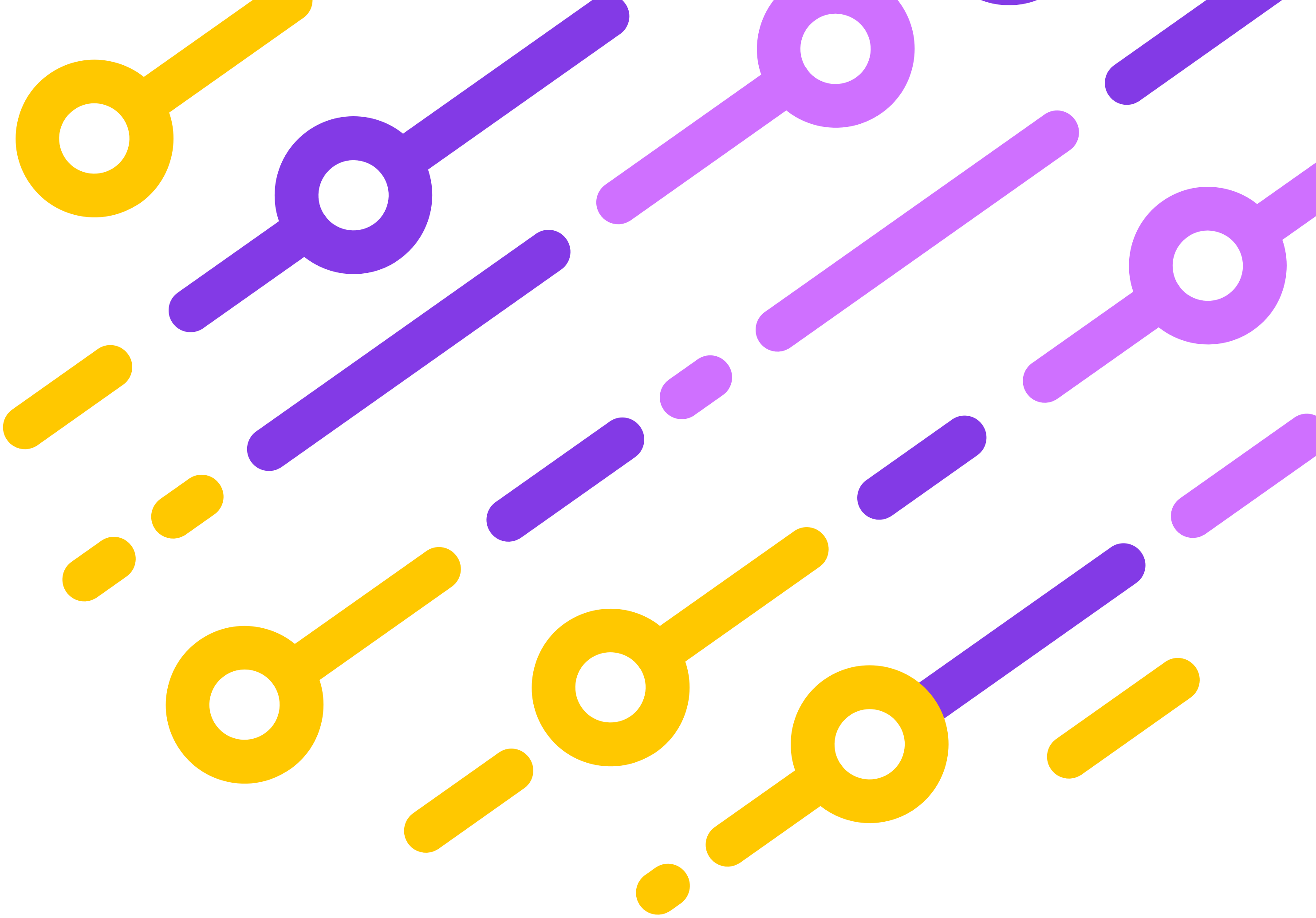


STM

Trusted Identity in Academic Publishing

OCT 2024

The central role of digital identity
in research integrity

Draft for community review

www.stm-assoc.org

Summary

For centuries, academic publishing has operated on a basis of trust, with an implicit assumption that individuals interacting with an academic journal do so in good faith and within established norms and practices. This high degree of trust means that researchers typically are not required to prove their identity or good intentions when they submit a paper for publication, act as peer-reviewer, or join an editorial board. In fact, most publishers require little more than a working email address to let users take part in the submissions and peer-review process.

Recent cases of mass retractions attributed to fraudulence, and a growing number of research integrity issues in academic publishing more generally, illustrate that this trust is increasingly vulnerable to exploitation. Paper mills and dishonest individuals have been able to subvert these processes for financial or reputational gain, risking pollution of the scholarly record and leading to a steep increase in retractions. As a consequence, there is now a gap between the level of trust that editorial systems need and the level that researchers can easily provide.

An instinctive solution is to increase security at the gates to editorial platforms, to insist on identity checks of the kind required when booking a plane ticket or hiring a car. But introducing measures of this kind is not a simple task. There are legitimate concerns about increasing friction for honest researchers, the risk of excluding researchers who do not have the means to pass such checks, and about user privacy.

While individual publishers who improve their researcher integrity checks may gain a strategic advantage by protecting their reputation, it's possible that some researchers would move to journals with the least stringent measures in place, simply because they present the lowest friction. Furthermore, inconsistencies between publishers' approaches adds complexity to the submission process. Any attempts to find a solution will need to block or deter fraud effectively while minimising burdens on researchers through a considered and proportionate approach.

This document aims to explain the background of this subject, and set the scene for forthcoming work that explores possible solutions. It also sets the direction for research into the most appropriate ways to make these solutions a reality, and to provide a way to measure their effectiveness.

1. Introduction

Academic publishing faces a growing research integrity challenge with an increasing number of retractions every year. These issues can take the form of entirely fabricated research, the misrepresentation of genuine findings to improve acceptance rates among reviewers and editors, or subversion of the peer-review process to increase the probability of a favourable outcome, for example by suggesting fake reviewers or even introducing corrupt guest editors. The repercussions extend beyond financial costs for editorial teams, posing a threat to the credibility of individuals, journals, and academic institutions, eroding trust in scholarly research on a broader scale.

Efforts to combat fraud include the detection of fake text and images, but detecting fraudulent practices in academic papers remains a complex challenge. Generative AI is becoming increasingly sophisticated, and the arms race between production and detection is unlikely to have a clear winner. Internal investigations by publishers have found that identity theft and manipulation often play a role in research integrity breaches. For this reason, publishers might consider factors beyond content, such as researcher identity, as a way to preemptively assess the integrity of submitted work.

The purpose of this work

STM Solutions assembled a group of experts from a number of academic publishers and other organisations in this space to form a `Task and Finish Group' to investigate this issue and develop potential solutions. By sharing data and developing these possible solutions, we hope that the challenges that have hindered progress in this area can be overcome, and that the potential solutions will aid in improving research integrity in academic publishing.

A guide to this document

1. Identity, trust and risk
A discussion of the core concepts in this area.

2. Identity methods
A look at the different ways that users identify themselves to websites.

3. Identity manipulation
Exploring the tactics that users can employ to subvert identity systems.

4. Mechanisms of trust
The ways that users can provide evidence of their trustworthiness.

5. Submissions Systems Survey
An investigation into the prevalence and severity of different manipulation tactics.

2. Identity, trust and risk

Conceptually, identity is the exchange of information about people in the real world with systems in the digital one. Referring to an individual user's identity in this context is to talk about the relationship between them and the system they are interacting with.

When someone visits a website, information is stored in their browser, and when they register an account they are setting up a way for this information to be remembered in future visits. By sharing something unique and something secret - typically an email address and a password - a user establishes a way for the system to know that it's the same person each time they sign in.

When users interact with digital systems of any kind, a level of trust is established. The basic assumption is that a user is who they claim to be, and that things they say about themselves can be relied upon, but this is not always the case. Some digital interactions involve individuals who seek to manipulate systems for their own benefit.

The domain of academic publishing suffers from the same issues as any other. Some users identify themselves fraudulently, or make false claims about themselves or their work, in an attempt to manipulate the publication process. This leads us to question whether the high degree of implicit trust that has been the basis of scholarly publishing for centuries is still tenable.

2.1 Two dimensions of trust

To develop a mental model of this subject, we can consider two dimensions:

- **Trust in the individual identity of a user:** how confident we are that the user's identity is being controlled by a real-world person, and not - for example - a bad actor impersonating someone else.
- **Trust in the things they claim about themselves:** how much evidence we have that a given user is a good actor, a genuine researcher acting honestly.

It's possible to be confident in a user's identity while having little or no evidence that they are a genuine researcher, and it is similarly possible that a user can point to an impressive research history without being able to strongly verify that they are the individual connected to that history. Verified identity without evidence of previous academic work (or equivalent) is preferable to evidence of credibility that can't be reliably linked to a given person, because identity at least provides a route to accountability.

Ideally it would be possible to reach a high level of trust in both dimensions, but this may not be possible or necessary in all situations. Instead, we need to find a way to get to a state where we have enough trust to allow a user to perform an action, given the nature and inherent risk of that action, the context in which it is to happen, and the ability of the user to demonstrate their integrity.

Although these two dimensions of trust are separate in theory, they are often combined in practice. For example, if a user is able to sign in via their institution, they can simultaneously provide information about their individual identity as well as evidence of a legitimate academic affiliation. Because there is a degree of inherent trust between publishers and known academic institutions, a researcher who can authenticate with their institution's identity infrastructure inherits that trust.

Where a user is unable to provide evidence of affiliation or their past academic work, it's possible that stronger proof of their individual identity could be accepted as an alternative.

2.2 Identity verification on the web

Identity verification serves as a means to hold users accountable for their actions. Measures to close the gap between complete anonymity and full identification increase the ability for users to be held responsible for what they do online, whether that's making a purchase, accessing sensitive information, or interacting with others. This acts as a deterrent to bad actors because knowing that their actions can be traced back to them personally can make individuals think twice before engaging in negative behaviour online.

In recent years, there has been a growing emphasis on identity verification across consumer web platforms. From financial services to social media networks, there's a trend towards implementing measures such as multi-factor authentication (MFA), biometrics, and external identity validation to enhance security and trust.

Many banking and investment platforms, as well as services that connect consumers to rental properties, modes of transportation and so on, now require users to undergo rigorous identity verification processes involving government-issued documents to prevent fraud and ensure compliance with regulations. Biometric verification is becoming increasingly common to authenticate users' identities and mitigate the spread of fake accounts. Services offering passport or driving licence validation have become a standard expectation for many digital consumers, adding a layer of accountability that is often absent in online interactions.

However, the effort and level of intrusiveness involved for users in strong verification of this kind is not trivial. It can take several complex steps to confirm your ownership of a passport, and the level of intrusiveness may make this approach inappropriate for some types of interaction. While in banking and other highly sensitive and regulated areas the benefits of robust verification measures outweigh the inconvenience or perceived intrusiveness, implementing disproportionately intrusive identity verification measures in academic publishing is likely to deter many researchers.

2.3 “Good” and “bad” actors

The range of actors in the world of scholarly publishing is fairly large, but the key individuals on each side of the process are the author and co-author, and the editor and reviewer. Any of these actors can be honest individuals who are seeking to work following the ethical and academic norms accepted in their community, or those who are seeking to manipulate the scholarly publishing process in dishonest ways. Separating the “good” from the “bad” actors in a fair and effective way is the persistent and challenging task at the centre of this work.

2.4 Primary and secondary users

Users directly interacting with an editorial submissions system can be considered primary users, whose identity and integrity as genuine academic actors needs to be trusted in real-time. This includes submitting authors and reviewers providing their reviews. Co-authors and suggested reviewers provided by the submitting author can be considered secondary users until they engage with the system directly as primary users. However, if secondary identities can be checked before this point, it may provide evidence of the trustworthiness of the primary user who suggested them.

2.5 Networks of trust

When a user identifies themselves to an editorial system through their institution, either by confirming their ownership of an email address or by signing in through the institution’s Identity Provider (IDP), the relationship between the editorial system and the institution is one of trust, equivalent to a direct phone call between people on each side. Service Providers (SPs) trust that IDPs will only allow genuinely authorised users - either academic staff, students or otherwise related people - to have credentials, and that any misuse of the system detected by the institution or reported to them by the SP will result in appropriate investigation.

2.6 Risk levels

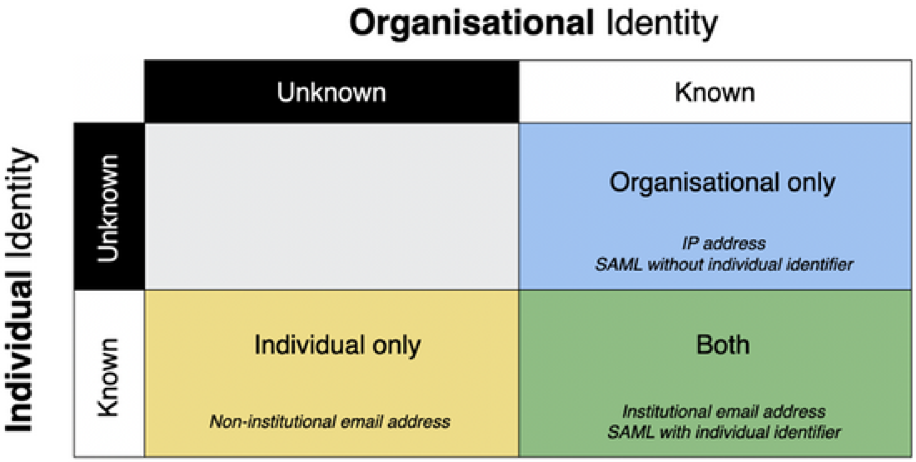
Trust levels required in editorial systems vary based on the threat of fraud and the seriousness of potential incidents. Certain journal subjects may attract more fraud, while not all actors or actions pose the same risk. The identity of corresponding authors holds greater importance than that of co-authors, given their primary responsibility for the manuscript. Reviewers wield significant power and thus require trust, though not to the extent of editors or guest editors who can make publishing decisions and operate at greater scales. For this reason, the default level of trust that users will be expected to meet will vary.

It’s important to acknowledge that there is unlikely to be a “one size fits all” solution, and that publishers will be in the best position to decide on the most appropriate balance for specific contexts.

3. Identity methods

In online interactions, users don't always need to be individually identified. Typically, users only sign in to allow information to be retained between sessions or where there is value in maintaining a long-term relationship. Organisational identity, on the other hand, is required whenever an interaction is based on a user's affiliation to an institution, for example when accessing resources paid for by their employer, or when submitting a manuscript to an Open Access journal and arranging for the payment of article processing charges (APCs).

These two sides of identity - individual and organisational - are conceptually independent of each other, and can occur separately or together. When they are presented in parallel, individual identity can be trusted more, as it is backed by trust in the organisation, although only to the extent that the organisation itself is trusted.



Organisational Identity

Individual Identity	Unknown	Known
Unknown		Organisational only *IP address* *SAML without individual identifier*
Known	Individual only *Non-institutional email address*	Both *Institutional email address* *SAML with individual identifier*

The methods described below evolved in the context of personalisation and resource access, but the relevance for trusted identity verification is important to examine.

3.1 Email

Email address is frequently used as a unique identifier, and confirming an email through a verification link increases confidence that the user providing it is its genuine owner. Confirmed institutional email addresses provide organisational identity in parallel, while publicly-available ones do not. Possessing an institutional email address doesn't guarantee that a user is a good actor, but it does provide a better way to trace suspicious activity and take action.

3.2 Federated Identity

Protocols like SAML allow delegation of the sign-in process to an IDP, so users don’t need to provide a password directly to the website they seek to use. Both individual and organisational identity can be shared in this way, and in situations where it is enough for an SP to know only the affiliation of a user, a user can remain anonymous. For privacy reasons, IDPs are often configured so that they don’t release individual identifiers, making this approach similar to the use of IP addresses as far as information disclosure is concerned. One crucial difference is that IDPs can keep an internal log of sign-in events, allowing institutions to investigate problems such as misuse if they arise.

The ability to sign in at an institution's IDP doesn't necessarily prove strict affiliation, as many institutions allow individuals other than active academics to have credentials. The eduPerson schema enables IDPs to share granular affiliation claims, providing roles such as student, staff, or alumni, but these claims are not always released.

The user experience of federated identity has been significantly enhanced by initiatives like SeamlessAccess, formerly known as RA21. SeamlessAccess makes identity journeys easier by allowing users to have their institution’s IDP remembered between visits to multiple service providers.

3.3 IP Ranges

Matching a user’s IP address against a list of known institutional IP ranges is a very common way for publishers to establish an institutional relationship. By design, there is no individual identity information shared. As there is only a passive interaction between publishers and the institutional network, it is not possible to maintain a log of usage in the same way that can be achieved via federated identity approaches, and security issues are difficult if not impossible to investigate.

3.4 No method is universal

It’s important to point out that not all authentication methods are available to be used by all researchers. The use of federated identity is widespread in areas including Europe and the US, but not all organisations in all parts of the world have the necessary technical infrastructure. Even institutional email addresses cannot be relied upon to be universally available, as many researchers use their personal email addresses to avoid the inconvenience of moving between multiple organisations, or don’t have access to an institutional inbox at all. The extent of these limitations in the real world is an area that needs to be investigated, and research on this is underway.

While the ability to access paid-for resources is by definition limited to those who can pay for them, or to those who can prove their relationship to an organisation that has, the ability to be a researcher and contribute to the academic literature is very deliberately not limited in this way. Open Access aims to ensure that scholarly content is available to everyone, and, in a similar way, it is a fundamental principle of publishing that everyone should be able to contribute their work.

A manuscript submitted by someone from an established institution should not carry more weight than unaffiliated research. For this reason, we must be careful not to assume that the methods of authentication that were established to facilitate paid-for content consumption are universally applicable to the submission process. Non-affiliated researchers, including independent scholars, citizen scientists, and researchers from smaller institutions or non-profit organisations, contribute significantly to scientific progress, and must not be excluded because they don’t have access to the identity infrastructure that established research institutions have.

4. Identity Manipulation

Bad actors seeking to manipulate editorial systems employ various tactics to gain fraudulent access. These tactics range from creating entirely fictional user accounts to impersonating legitimate users using false or misleading information, or even stealing the credentials of others to directly impersonate them. In the section below, we list and describe these different tactics.

4.1 Opaque addresses

The ease of creating user accounts with services like Google or Hotmail, without verification of real-world identity, allows bad actors to create entirely opaque, seemingly legitimate email addresses. Services offering temporary or anonymous email addresses provide valuable anonymity but also enable bad actors to conceal their identities and escape accountability for their actions.

4.2 Impersonation

Bad actors exploit the freedom to create any email address to deceive others, such as using a Gmail address resembling that of a reputable researcher. More sophisticated impersonation involves registering convincing false domain names, mimicking real researchers' email addresses, and creating plausible institutional domains, an approach often used in email scams and phishing attacks.

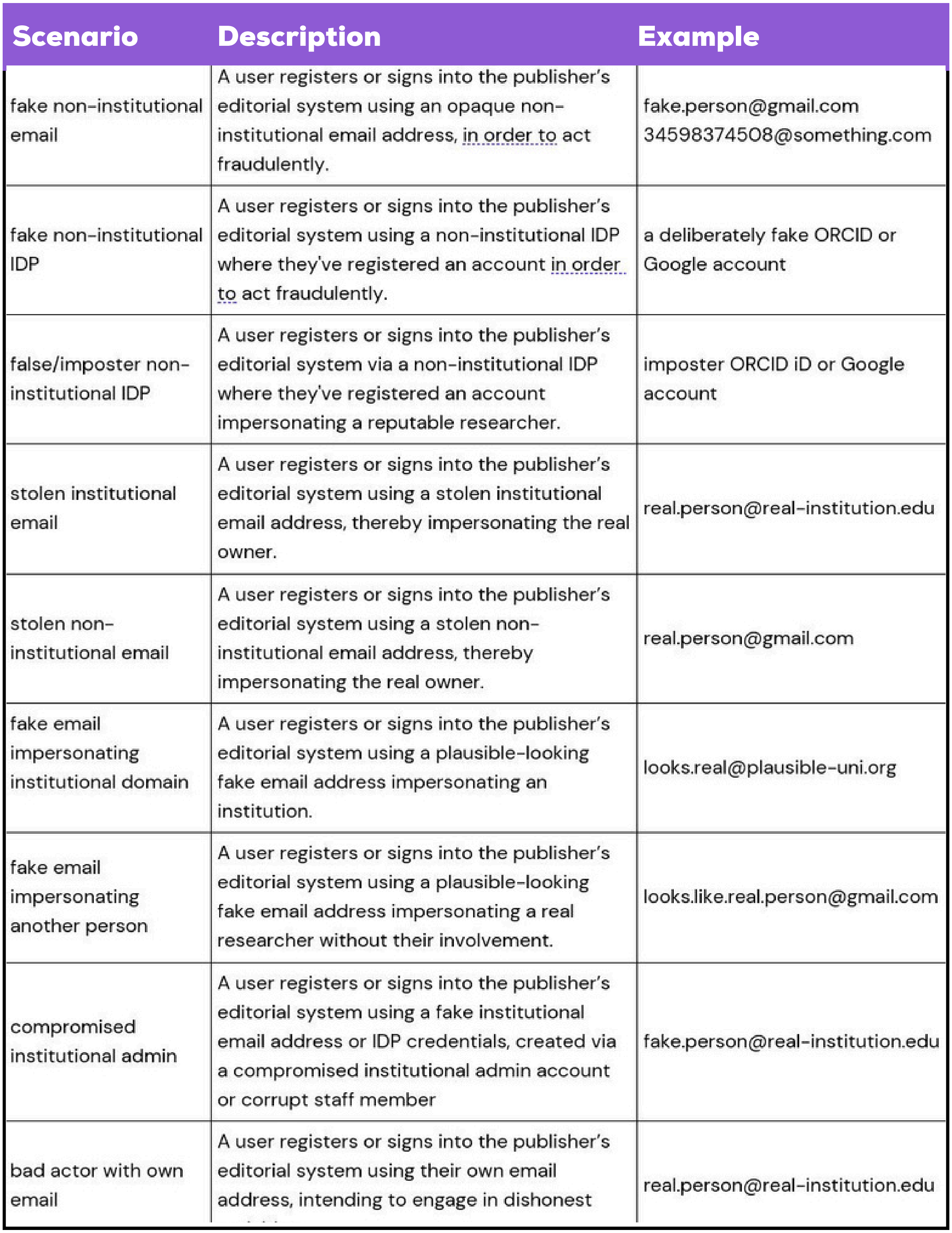
4.3 Credential Theft

Bad actors can steal a user's identity by obtaining their username and password. While password complexity rules aim to prevent easy guessing, unauthorised access remains possible with knowledge of the credentials. Multi-factor authentication (MFA) defends against this vulnerability by requiring additional verification beyond credentials, such as real-time mobile notifications, to ensure the rightful owner's identity. It’s important to note that if a user deliberately misuses MFA, by permitting the use of their account by another user, this cannot be straightforwardly defended against, and that some implementations rely on one-time passwords sent via SMS, and are vulnerable to being compromised by sophisticated attackers.

4.4 Corruption

The integrity of a system's security relies on every component, making it vulnerable if any part is compromised. If the administrator of an institution's email and identity infrastructure acts dishonestly, they can create user accounts for fraudulent activities, posing a significant challenge to defence measures.

4.5 Examples



Scenario	Description	Example
fake non-institutional email	A user registers or signs into the publisher's editorial system using an opaque non-institutional email address, in order to act fraudulently.	fake.person@gmail.com 34598374508@something.com
fake non-institutional IDP	A user registers or signs into the publisher's editorial system using a non-institutional IDP where they've registered an account in order to act fraudulently.	a deliberately fake ORCID or Google account
false/imposter non-institutional IDP	A user registers or signs into the publisher's editorial system via a non-institutional IDP where they've registered an account impersonating a reputable researcher.	imposter ORCID iD or Google account
stolen institutional email	A user registers or signs into the publisher's editorial system using a stolen institutional email address, thereby impersonating the real owner.	real.person@real-institution.edu
stolen non-institutional email	A user registers or signs into the publisher's editorial system using a stolen non-institutional email address, thereby impersonating the real owner.	real.person@gmail.com
fake email impersonating institutional domain	A user registers or signs into the publisher's editorial system using a plausible-looking fake email address impersonating an institution.	looks.real@plausible-uni.org
fake email impersonating another person	A user registers or signs into the publisher's editorial system using a plausible-looking fake email address impersonating a real researcher without their involvement.	looks.like.real.person@gmail.com
compromised institutional admin	A user registers or signs into the publisher's editorial system using a fake institutional email address or IDP credentials, created via a compromised institutional admin account or corrupt staff member	fake.person@real-institution.edu
bad actor with own email	A user registers or signs into the publisher's editorial system using their own email address, intending to engage in dishonest	real.person@real-institution.edu

The following are discrete methods of identity manipulation. Note that "fake" in the scenarios below means something that's not genuine, and that has been created to mislead.

5. Mechanisms of Trust

Trust in digital systems is not binary but encompasses various elements of information about a person's identity and attributes. Beyond the ability to sign in, users may share information such as their institutional affiliation, educational background, academic credentials, and past contributions to scholarly research, such as authored papers or conference presentations. These attributes contribute to establishing trust in the person's identity and expertise within the academic community.

It is important to separate the idea of trust in an individual's identity from their reputation as a legitimate researcher. The former is analogous to establishing that someone is the genuine owner of a bank account, for example by challenging them to enter their PIN number when they present their bank card at a cash machine, while the latter is more like their credit history, containing a record of previous transactions. Just as money can only be accessed by someone who can authenticate with their bank, an established researcher can only be recognised as such by a digital system if they can prove that they are the genuine owner of the identity linked to claims about their history.

The following are ways in which users may provide information to do this.

5.1 Institutional affiliation

The term "affiliation" can have several meanings, because of the different kinds of relationship that a user may have with an institution. An affiliation might be long-term or temporary, or one of several existing in parallel. Despite this potential ambiguity, proof of a relationship of any kind is valuable as it links an individual to an organisation, providing a route for accountability where there would otherwise be none.

5.1.1 Federated Identity

Federated identity can simultaneously provide confidence in both a person's identity and their associated attributes. Trust provided in this way directly reflects an SP's trust in the IDP, which is responsible for ensuring that only genuine individuals can sign in, as well as optionally releasing additional information about them. Broad institutional affiliation is always released as a claim, while personally identifiable information is only released where the IDP is configured to do so.

Not all IDPs can be necessarily trusted to the same extent, however. Some institutions may have more robust security measures in place than others, making them more trustworthy indicators of a researcher's affiliation and legitimacy.

5.1.2 Institutional email

Confirmation of institutional email acts in the same way as federated identity, in that it provides evidence of a genuine institutional relationship. While it's possible for an IDP to release granular affiliation information, email confirmation is not a precise indication of the nature of a user's relationship, revealing only that the institution has given access to an inbox. This is nevertheless meaningful, as it shows that the institution trusts the user enough to let them use their infrastructure, and has a record of their activity.

5.1.3 Validated ORCID affiliation

An validated affiliation on an ORCID profile can be relied upon by editorial systems, as one or more of the same methods will have been used to determine the validity of the relationship as would have been used directly.

5.2 Previous academic activity

Publishers often use manual processes to investigate a user's history and assess their trustworthiness, looking for both positive and negative signals in a researcher's background. Elsevier, for example, explains in an article called "Fighting the problem of fraud in publishing" how they encourage editors to "keep an eye open for any potential things that might be wrong". They ask "...[i]s the author on the editorial board for that journal? Have they authored papers at that journal before? Have they been a reviewer? All things that might establish more confidence that they are the person they say they are, and that they do have expertise in the field."

Checking these things manually and on a case-by-case basis is costly. ORCID offers a more scalable way to make these checks, using "Trust Markers", which are claims made by a trusted source about a researcher. For example, when a user says that they have published work, acted as a reviewer on a journal, or received funding, that can be independently verified or directly provided by the institution or funding body rather than being purely self-asserted by the user. It is, in theory, possible for an unscrupulous organisation to falsely verify a researcher's claims, but if this happens then the organisation in question can be identified, and appropriate action taken. ORCID may remove false data and terminate the membership of any organisations that make knowingly false assertions which they refuse to correct.

5.3 Connection to another trusted user

Where a user cannot provide direct evidence of their trustworthiness, association with other trusted individuals could provide a proxy for that trust. This could be through one individual directly vouching for another, or through patterns of association indicative of trust, for example frequent co-authorship on manuscripts not associated with fraudulent behaviour. Reputation-based identity verification processes are not without their challenges; for example, there is the question of consequences for someone who has vouched for someone who goes on to act fraudulently.

5.4 Identity verification services

There are many companies offering identity verification services, examples include https://www.veriff.com/, https://id.me, and https://www.yoti.com/. They are widely used by a range of corporate, government and healthcare organisations, and support users in hundreds of countries. They generally charge on a per-verification basis, but are extremely valuable as they allow users to securely prove their ownership of official documents such as passports and driving licences, which provides strong confidence in their individual identity. While this doesn't say anything about their academic legitimacy, it does offer a powerful route to accountability.

5.5 Direct contact

Where no automated method is available, manual verification will become the last resort. Publishing teams can contact submitting authors directly, as they do where they need to clarify aspects of their submissions, request additional information, or address concerns related to the peer review process. However, the feasibility of this approach as a pre-emptive step will depend on the volume of submissions and the resources available to journals.

5.6 Watch-lists

Paper mills have been known to generate sophisticated fraudulent identities, designed to appear trustworthy even after careful scrutiny. Publishers may choose to keep a record of known bad actors and check for their reuse, which would make this approach less effective, as the effort required to fabricate identities is quite significant. However, the use of AI to generate fakes is likely to make this easier, and would require equivalent approaches to defend against.

5.7 Trust over time

Just because a user can demonstrate their trustworthiness at a given moment in time does not mean that they can be trusted indefinitely, without any further verification. But a certain level of trust in a user can be established once and then stored so that the entire process doesn't have to be endlessly repeated. The length of time that a claim can be relied upon will vary on the nature of the claim and the context in which it's being used, and so in some cases it's necessary to revalidate frequently.

At the same time, the age of a claim can add to its trustworthiness. An identity that was created shortly before its use should be treated with caution, while an account that has existed for many years - provided it has not been stolen - may be less likely to be fraudulent.

6. Submissions System Survey

In order to get a picture of the ways that identity manipulation is experienced in the real world, we designed a set of questions and invited a number of publishers to provide input. We wanted to explore:

- how frequently different identity fraud tactics are seen in practice
- the frequency and severity of different actions taken by bad actors during the editorial process

12 publishers anonymously responded to the survey (see appendix), and the results are presented below.

6.1 Mechanisms used by Bad Actors

We presented respondents with a set of options - the scenarios listed in section 3 above - and asked them to indicate how often they come across the different scenarios when incidents of fraudulent activity are later investigated.

Results

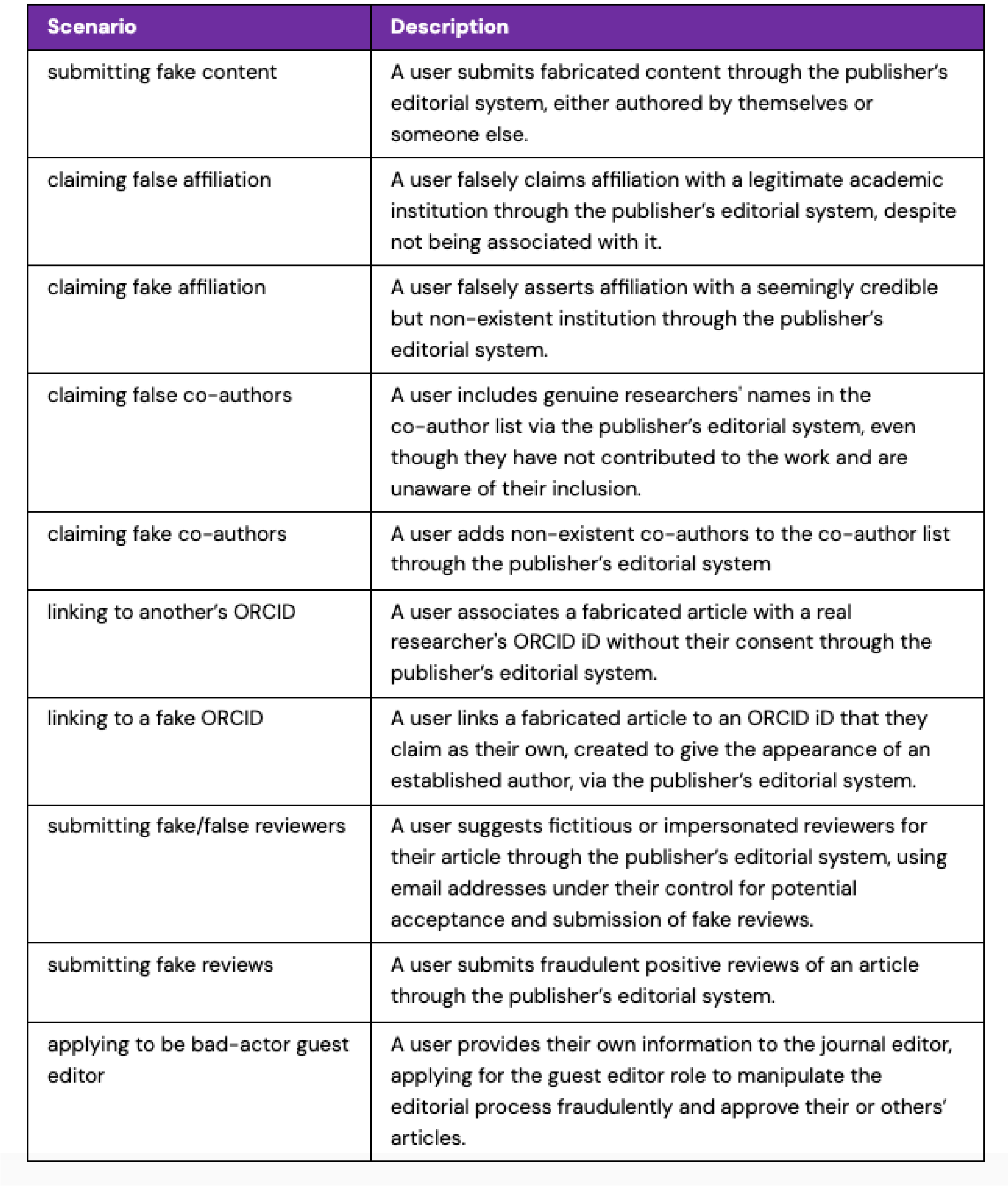
Responses to the survey revealed the following order of prevalence.

1. fake non-institutional email
2. bad actor with own email
3. fake non-institutional IDP
4. fake email impersonating institutional domain
5. fake email impersonating another person
6. false non-institutional IDP
7. stolen non-institutional email
8. stolen institutional email
9. bad actor with shared email
10. compromised institutional admin

The fact that non-institutional emails and IDPs are most frequently seen in instances of fraud is significant, because it suggests that - while other methods will be necessary - a way to defend against bad actors would be to insist on confirmed institutional identity for all users.

6.2 Actions taken by Bad Actors

As there are many different parts of a publishing workflow, and therefore a number of places where fraud may happen, we need to understand which ones need to be protected by verification.

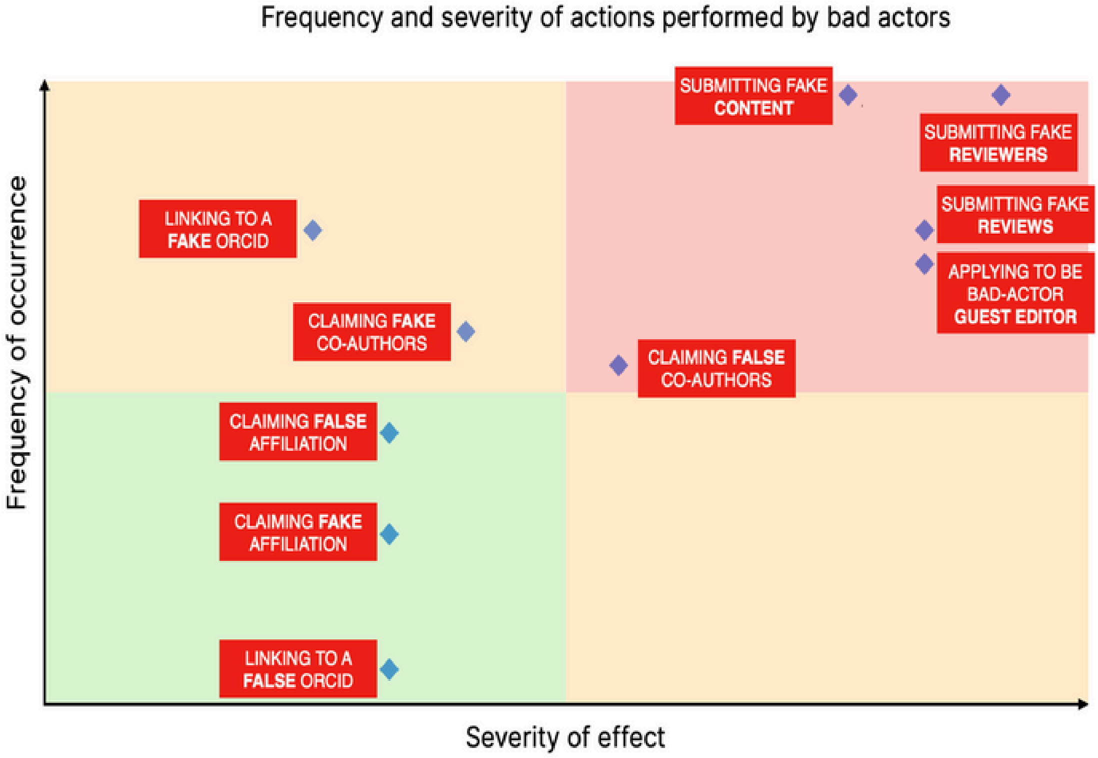


Scenario	Description
submitting fake content	A user submits fabricated content through the publisher's editorial system, either authored by themselves or someone else.
claiming false affiliation	A user falsely claims affiliation with a legitimate academic institution through the publisher's editorial system, despite not being associated with it.
claiming fake affiliation	A user falsely asserts affiliation with a seemingly credible but non-existent institution through the publisher's editorial system.
claiming false co-authors	A user includes genuine researchers' names in the co-author list via the publisher's editorial system, even though they have not contributed to the work and are unaware of their inclusion.
claiming fake co-authors	A user adds non-existent co-authors to the co-author list through the publisher's editorial system
linking to another's ORCID	A user associates a fabricated article with a real researcher's ORCID iD without their consent through the publisher's editorial system.
linking to a fake ORCID	A user links a fabricated article to an ORCID iD that they claim as their own, created to give the appearance of an established author, via the publisher's editorial system.
submitting fake/false reviewers	A user suggests fictitious or impersonated reviewers for their article through the publisher's editorial system, using email addresses under their control for potential acceptance and submission of fake reviews.
submitting fake reviews	A user submits fraudulent positive reviews of an article through the publisher's editorial system.
applying to be bad-actor guest editor	A user provides their own information to the journal editor, applying for the guest editor role to manipulate the editorial process fraudulently and approve their or others' articles.

The survey asked respondents to indicate how frequently these scenarios are found to have been part of an incident of publishing fraud, and how severe they judge the effect was.

Results

The actions identified as high occurrence and severity were broadly as expected.



Overall, the most common and serious scenarios were:

- authors suggesting fake candidates for peer review
- the submission of fake content
- reviewers submitting fake reviews
- users applying to become guest editors
- claiming false co-authors

In reality, it's likely that the frequency and judged severity of these scenarios will vary between journals. As mentioned, not all actions have the same level of inherent risk, as some have more serious consequences than others. Some journals will be particular targets for fraud, and may be more or less prepared to tolerate this. It will be important to develop a way for each editorial system to be able to reach its own position on risk, based on a consistent model.

7. Conclusions

Editorial systems generally allow freely-available non-institutional email addresses to be used when users sign in to submit a manuscript or act as a reviewer. This opens the door to the use of opaque and false emails that are cheap and easy to create, reducing accountability and lowering the barrier to submit or review fraudulent research through the peer-review system.

At the opposite end of the publishing process, when it comes to researchers gaining access to paid-for scholarly resources, there is an established approach to identity. Users need to be able to provide proof of their affiliation to a paying institution, either by being inside a recognised IP range, by being able to sign in at a known IDP, or by confirming their ownership of an institutional email address. Similarly, many Open Access publishers need institutional identity to be confirmed when handling article processing charges.

There are important differences between paid-for access and content submission, but the fact that most researchers are able to prove institutional affiliation for the purposes of the former may provide an important part of the solution to the problem of identity manipulation.

In subscription scenarios, publishers trust that a request coming from someone authenticating in one of these established ways really does have a relationship with a paying customer organisation and so has a right to access content. There has been considerable effort over many years to build the infrastructure to make this possible, and - with projects like SeamlessAccess - to make the user journeys involved as simple and consistent as possible. In a similar way, it will be important and valuable to work collaboratively to arrive at a coordinated approach to addressing the challenge of identity fraud in scholarly publishing.

While access to subscription content doesn't require individual identity and can be based on anonymous organisational identity, the submission process does. By combining individual and organisational identity, coupled with prior academic publishing activity, the established trust that publishers have in customer institutions could be used to provide additional assurance that a user is genuine and trustworthy.

7.1 Next steps and open questions

Addressing identity fraud in academic publishing is a complex issue that requires a coordinated approach. While there is no single solution, we have identified some areas for further work that may help to improve trust and integrity in the publishing process.

- Strengthening identity verification through the use of federated identity and requiring institutional email verification should provide a higher level of assurance about a user's legitimacy. These methods rely on existing trust infrastructure and are likely to cover a significant proportion of researchers.

- The use of ORCID trust markers can help verify a researcher's previous work and affiliations, offering an alternative for those who do not have access to an IDP or institutional email address.

- For researchers unable to prove their affiliation through these means, exploring third-party identity verification services might provide additional options. However, it will be essential to balance the need for security with the potential invasiveness and effort required from users. In cases where automated methods are not feasible, manual verification through direct contact with submitting authors will remain a last resort.

Several questions remain for further exploration

- What alternative verification methods can be offered to researchers who cannot prove their affiliation through traditional means or point to verified previous work?

- Is it appropriate to use services that validate government documents, considering the potential effort, invasiveness, and cost?

- How can we offer a simple and consistent set of recommendations that balances the need for trust and accountability with the need for inclusiveness, low friction, and respect for privacy across various risk levels?

- How can we design pilots to test a range of approaches and measure their effectiveness?

By working together to address these questions, we hope to enhance the integrity of scholarly research while ensuring that the process remains accessible and trustable for all researchers.

8. Appendix

8.1 Group members

- Aaron Wood, American Psychological Association
- Adam Hough, Elsevier
- Adam Sewell, IOP Publishing
- Andy Heard, IEEE
- David Flanagan, Wiley
- Helen King, Sage
- Hylke Koers, STM Solutions
- Jacob Kendall-Taylor, JAMA
- Jennifer Wright, Cambridge University Press
- Joris van Rossum, STM Solutions
- Kevin Lawson, Aries
- Liv Davies, IOP Publishing
- Ralph Youngen, American Chemical Society
- Richard Northover, STM Solutions, coordinator
- Sam Parker, Wiley
- Tim Lloyd, LibLynx

8.2 References

- More than 10,000 research papers were retracted in 2023 — a new record (Nature 624, 479-481)
- Trust Markers: Interpreting the trustworthiness of an ORCID record (ORCID, 11 August 2021)
- There's far more scientific fraud than anyone wants to admit (The Guardian, 9 Aug 2023)
- Revealed: The inner workings of a paper mill (Retraction Watch, 20 Dec 2021)
- Imposters and Impersonators in Preprints: How do we trust authors in Open Science? (Scholarly Kitchen, 21 March 2021)
- Taxonomy of Disinformation (arXiv, 19 Nov 2023)

8.3 Terminology

Term	Definition
bad actor	A person intending to fraudulently manipulate the editorial process
fake identity	A user account created with details that don't represent the person truthfully, intended to be used for fraudulent purposes
fake	Something that's not genuine, and has been created to mislead
fraudulent activity	Action that's taken with a dishonest motivation
IDP	An Identity Provider, controlled by a university or other organisation, which can be used to identify a person
impersonate	To misleadingly pretend to be somebody else
institutional email address	An email address that could only have been issued by a given institution
Multi-Factor Authentication (MFA)	The use of more than one method of authentication from independent categories of credentials to verify the user's identity.
non-institutional email address	An email address that could be owned by anyone, not tied to a specific institution, such as a Gmail or Yahoo email account.
SP	A Service Provider, typically a web application or service that relies on an external identity provider (IDP) for user authentication and authorization
Open Access (OA)	A publishing model that allows free, immediate access to research outputs such as journal articles, enabling anyone to read and download them without financial or legal barriers.
plausible-looking	Something that appears genuine to an observer
Paper Mills	Entities that produce and sell fraudulent research papers, often using fabricated data and authorship.
reputable/real researcher	A genuine person that a bad actor might impersonate
shared email address	An address that more than one person may legitimately use, for example research.group@insitution.org
stolen credentials	The email/password for an account that a bad actor has obtained

8.4 Researcher Identity Survey

Below is the text from the survey sent to editorial system owners for input.

This is a survey to learn about identity fraud in academic publishing.

You'll be presented with some descriptions of things that can happen in incidents of fraudulent use of publishing systems.

We want to know how common you think these scenarios are, and how serious their impact is. This will help us to prioritise areas for further analysis.

You don't need to base your answers on concrete data; we are looking for your best educated guesses.

This survey was created by the Researcher Identity Working Group in STM Solutions' Access & Identity Cluster. It should take you less than 30 minutes.

PART 1:

Ways that people identify themselves

This first section is about approaches "bad actors" use to register or sign in to a publisher's editorial system.

In our examples, "John" represents a bad actor who is looking to subvert the editorial process.

- What do I need to do?
- You should think about previous instances of fraudulent activity, and consider how that activity was carried out.
- For example, did the person behind the activity use stolen credentials, or create a fake identity using an email address that they created specifically?
- If you've never seen the scenario, or think it's irrelevant, mark it as "very low".
- If it's something you recognise as a common pattern in cases where identity fraud is at play, mark it as "very high".

...and feel free to choose answers that sit between these two.

John registers or signs into the publisher's editorial system using a fake non-institutional email address.
e.g. fake.person@gmail.com
34598374508@something.com

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a non-institutional Identity Provider (IDP) where he's registered a fake account.
e.g. fake O**RCID iD or Google account**

- **How common is this scenario?**

John registers or signs into the publisher's editorial system via a non-institutional IDP where he's registered an account that impersonates a reputable researcher.
e.g. imposter ORCID iD or Google account

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a stolen institutional email address, thereby impersonating the real owner of that address.
e.g. real.person@real-institution.edu

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a stolen non-institutional email address, thereby impersonating the real owner of that address.
e.g. real.person@gmail.com

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a plausible-looking fake email address that impersonates an institution.
e.g. looks.like.real.person@plausible-uni.org

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a plausible-looking fake email address, impersonating a real researcher but without their involvement.
e.g. looks.like.real.person@gmail.com

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a fake institutional email address or set of IDP credentials, created via a compromised institutional admin account, corrupt member of staff, etc.
e.g. fake.person@real-institution.edu

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using his own email address, intending to do dishonest things.
e.g. real.person@real-institution.edu

- **How common is this scenario?**

John registers or signs into the publisher's editorial system using a shared email address, intending to do dishonest things
e.g. research.group@real-institution.edu
research.group@gmail.com

- **How common is this scenario?**

- *Are there any identity fraud methods we've missed?*
- *How do they happen? How common are they?*
- *Would you like to say anything else about this section?*

PART 2:

The things that bad actors do

This section is about the actions that are performed once someone has identified themselves.

What do I need to do?
You should think about previous instances of fraudulent activity, and consider what happened. This might be as a result of investigations that happened some time after the event, or from actions that were picked up and prevented.

For example, did the person submit fake content, or fake information about that content? Were they using real names, impersonating others, or what?

As well as how often you see the scenario, think about how serious the impact is on the whole editorial process.

Rate the scenario from "very low" to "very high" on these two dimensions. The rating is meant to be relative, to get a sense of the priority order between the various scenarios.

Remember that it is fine to answer based on your instinct and experience, not necessarily concrete data that you would need to gather and process.

When we ask you to rate how common a given scenario is, we are not looking for prevalence compared to all submitted content (including all bona fide submissions), but relative to other "bad actor" situations.

If you've never seen the scenario, or think it's irrelevant, mark it as "very low". If it's something you recognise as a common pattern in cases where identity fraud is at play, mark it as "very high".

For impact, we're interested in how serious the scenario can be. If the action has minimal consequences, or can easily be mitigated, mark it as "very low". If it's something that is very likely to result in sizeable corruption of the scholarly record, and is impossible to mitigate, mark it as "very high".

John uses the publisher's editorial system to submit fake
which he or someone else has generated.
The content could be a research paper, or elements of it.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to claim that h
affiliated to a genuine academic institution that he is not
part of.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to claim affiliat
plausible-sounding but fake institution.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to add the nan
legitimate researchers to the co-author list, where in real
those researchers have not contributed to the work at all
not even know their names have been added.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to add the names of fake co-authors to the co-author list, where in reality those researchers don't exist at all.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to link his fake article to an ORCID iD that he claims belongs to him, which has been created to make him look like an established author.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to link his fake article to an ORCID iD that belongs to a real researcher but without their involvement.

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to suggest fake or impersonated reviewers for his article, using email addresses that he controls. (He can later accept them and submit fake reviews.)

- **How common is this scenario?**
- **How serious is the impact?**

John uses the publisher's editorial system to submit fake favourable reviews of an article that was submitted fraudulently.

- **How common is this scenario?**
- **How serious is the impact?**

John submits his own details to the journal editor, applying for the role of guest editor in order to fraudulently control the editorial process and approve his own or others' articles.

- **How common is this scenario?**
- **How serious is the impact?**

What is the name of your organisation?
This question is optional, but will help us to ensure good coverage of survey responses.

How many journals does your organisation publish?

- **Less than 20**
- **Between 20-50**
- **Between 50-100**
- **Between 100-200**
- **More than 200**
- **Rather not say**

Does your organisation focus on a particular subject area(s)?
If so, what are they?

Are there any other relevant identity-related parts of the editorial process that you see involved in incidents of fraud?

How common are they?

How serious is the impact when they're seen?

What solutions – or suggestions for solutions – are you aware of to any of these issues?

Please provide as much information as you can.

Do you have any final comments or suggestions?