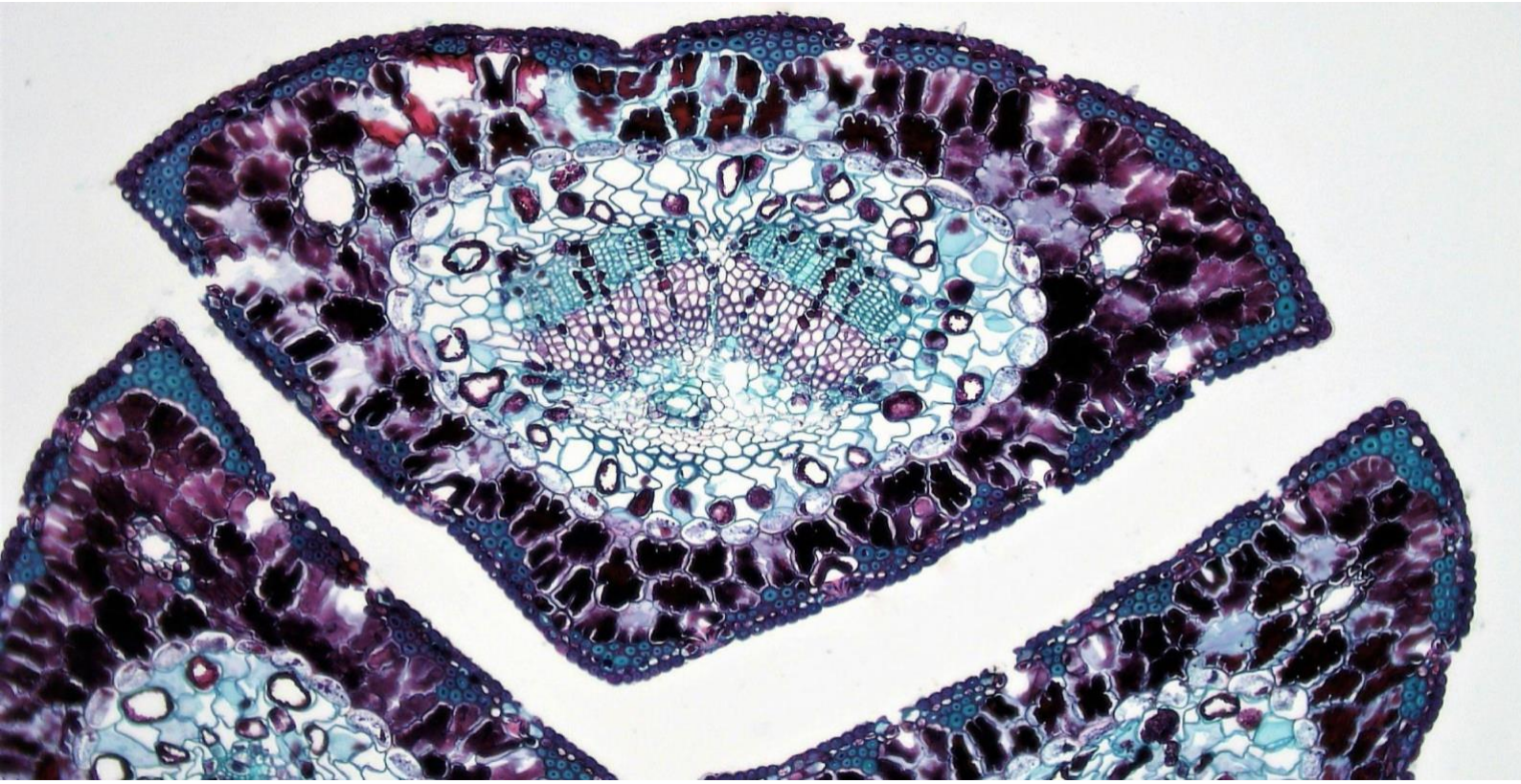
MOREBRAINS

Feasibility of technical solutions for the detection of falsified images in research

Phill Jones, PhD



November 13th, 2024

STM

This report was funded by STM

Photo by Fayette Reynolds M.S.: https://www.pexels.com/photo/close-up-of-shiny-rock-texture-11198509/

Table of Contents

Acknowledgements

Funding:
This report was funded by STM

The following people contributed ideas and acted as reviewers:
STM Image Alterations and Duplication Working Group
Joris Van Rossum PhD, STM Association
Hylke Koers PhD, STM Association
Alice Meadows, MoreBrains
Fiona Murphy DPhil, MoreBrains

Special thanks to the anonymous interviewees and technical reviewers

1 Summary and recommendations

In this report, we discuss the findings of an investigation into the feasibility of developing a research image integrity system. The focus was primarily on blots, and microscopy images, although the findings are relevant to other types of research images such as gel blots and spectroscopy, geophysical, or astronomical images, as well as to any digital data type. The investigation comprised a mix of desk research and semi-structured (anonymised) interviews with nine stakeholders representing publishers, researchers, technologists, and microscopy companies.

We found that it would be possible to implement a system to assert image integrity and that the technology to do so already exists. The key challenges are organisational and socio-technical; some form of shared governance and investment would be required to make it a reality.

Recommendation 1 - Create a cross-stakeholder working group to develop the concept of a research integrity system.

Developing the system would require collaboration and support from a variety of stakeholders:

- Researchers would be primary users of the system
- Instrument manufacturers would need to implement the system
- Editorial software providers would need to integrate the system into their software
- Publishers would need to incorporate image checks into their workflows
- Learned societies would need to support and educate their members in use of the systems, and ensure that the system supports and extends current best practice in their disciplines
- Journal editors would need to support and, eventually, require the system's use in research papers in their journals
- Institutions would need to support the system, encourage and educate researchers, and potentially host some of the infrastructure
- Funders would need to encourage and, eventually, mandate the system's use

Each of these stakeholders will have specific requirements of and concerns about the system that will need to be addressed. It is therefore vitally important that they are all involved in the governance, development, and implementation of the system., through the creation of a cross-stakeholder working group. Choosing a host organisation for the working group poses its own challenges, as it needs to be seen to be independent and not serving a particular stakeholder. We therefore recommend that the group is transparently governed and jointly hosted. Potential hosts might include many of the organisations listed in recommendation 2.

Recommendation 2 - Engage and coordinate with relevant initiatives and efforts already in place

Several initiatives related to research integrity already exist. It is important that the working group coordinate with these efforts to create legitimacy and buy-in, take advantage of potential synergies like integrations and common standards, and avoid recreating things that already exist. Below is an incomplete list of relevant organisations and initiatives, in the approximate order that they are mentioned in this report:

- Cytiva
- Research Data Alliance (RDA)
- Committee on Publishing Ethics (COPE)
- Open Microscopy Environment (OME)
- Various national microscopical societies
- The DONA Foundation
- The DOI Foundation and registration agencies such as DataCite and Crossref
- Various research funders that have been involved in research integrity issues
- Standards organisations including National Information Standards Organization (NISO)
- National Institute for Standards and Technology (NIST)
- Persistent Identifiers for eResearch (ePIC)
- Coalition for Content Provenance and Authenticity (C2PA)
- European Open Science Cloud (EOSC)
- Research Object Crate (RO-Crate)

Recommendation 3 - Develop a series of standards and tools to enable sector-wide adoption

Creating the conditions for universal adoption of the system will be paramount to its success. A proprietary or closed system will not earn the trust of the research or instrument manufacturer communities. The choice of technologies and approaches will ultimately be the responsibility of the working group. However, we recommend the following ideas as a starting point for discussion:

- Cryptographic hashing, coupled with a public key infrastructure approach
- A network of distributed trust stores, using a DOI registration agency or a new implementation of the Handle system to verify certificates
- Integration with persistent identifiers for instruments that are currently being developed by an RDA working group[1]
- Partnerships with equipment manufacturers and instrument software plugins
- Development of open source verification software, and integrations into existing publishing workflow software like manuscript submission systems

[1] More information about the working group can be found here: https://www.rd-alliance.org/groups/persistent-identification-instruments-wg/

2 Authenticity and integrity of images

In an ideal world, it would be possible for some mix of technical and workflow processes to guarantee the veracity of images and other data published in or alongside research articles. Unfortunately, veracity itself is too high of a bar for any quality control mechanism as it requires integrity, authenticity, honesty, and, not least, infallibility. However, a solution to prevent, or at least significantly impede, bad actors – like paper mills and fraudulent researchers – from publishing faked, copied, or inappropriately manipulated images must, at a minimum, be able to assert two things:

- Authenticity - the image originates from a particular instrument
- Integrity - the image has not been altered in a way that hasn't been declared

The technology to certify the authenticity and integrity of an image (which we will refer to as an 'image integrity system') exists and could be implemented with effort and coordination across multiple stakeholders. While the potential system we describe here would not guarantee image integrity in and of itself, it would allow for a certified version of an original image to be stored or shared for comparison with a final, published version. It would then be up to whoever is investigating or assessing the image to decide whether the published version accurately reflects the original, and whether the author has adequately described any manipulations or enhancements made to the published image.

As is often the case in multi-stakeholder environments where a technical solution is warranted, the biggest challenges to implementation are sociotechnical. For it to be adopted, and to work reliably, the system must be broadly accepted and implemented in a standardised way by all stakeholders. For such a system to become a requirement for the publication of research, it must be available to everyone who wants to publish. It must, therefore, be implemented in a way that enables all equipment manufacturers to easily adopt it, and it must also be flexible enough to cater to various national and disciplinary requirements.

3 The importance of image integrity

Research integrity is an increasingly important concern across many disciplines (1). In 2022, an investigation by STM and the Committee for Publishing Ethics (COPE) published a report into the prevalence of paper mill-generated articles in the scholarly literature (2) that found alarming levels of fake papers, including fake images, in many journals.

The rise of generative AI poses an increased level of risk. Faked AI content, including images, is significantly harder to detect than manually replicated or altered content. There have been high-profile examples of illustrations in some articles being obviously AI-generated, with nonsense labels and ridiculous cartoonish illustrations. These examples frequently come to light, as nonsensical illustrations of anatomy or cellular pathways are easily spotted by the community. However, data images such as microscopy or blots would not be as readily spotted, and many believe that these may already be widespread in the literature (3). As Rongshan Yu from Xiamen University (China) was quoted in a 2022 article in *Chemistry World* (4):

> *I think we have reached the point where we can no longer tell if the paper is real or fake*

It is important to note that the manipulation of images is not always a sign of poor research integrity, and opinions on how serious it is vary in the research community. During our interviews, every current or former researcher we spoke to confirmed that they had personally witnessed varying levels of manipulation of images. This included manipulations by colleagues, and manipulations they made themselves (as junior researchers) because it was an accepted part of research culture at their institution or lab. As one interviewee explained:

> *It was common. People just put what they wish to see or what they think they should be seeing. But sometimes, in some cases, there's something else going on that explains the absence, or that explains the difference, [it] could be very interesting science [but colleagues and reviewers] wouldn't believe it.*

Some researchers make the argument that manipulation of so-called 'representative' images is not, in and of itself, very important, because the images are never really representative. We spoke with a senior researcher with a longstanding interest in research integrity and reproducibility, who said:

> *Why do we put these images in papers? To what extent does an epistemological claim rest on a visual piece of information? [A researcher might say] well, we showed in this Western [blot] that the knockout doesn't have any of the protein*[2]*. You know, sure, but I know, and every reader who's done a Western knows that that's not the typical Western blot. It's the Western blot that they spent three months getting this very atypical example … it's not the median [result]. So, part of my response is to say, anyone who judges the validity of claims made in a research publication by looking at the immuno or looking at the Westerns is daft, because it can never provide that.*

[2] A 'knockout' is a genetically modified animal in which a specific gene has been disabled so that the animal is incapable of manufacturing a specific protein within its own cells. A common use case is to create a 'disease model' so that experimental medical interventions can be tested.

While that quote may seem surprising to many, it is widely acknowledged among researchers that images are seldom representative examples; they are invariably at least highly selected – and often enhanced – to emphasise a point. This is not to say that faked images are not a concern. The same researcher quoted above also noted that images that are faked or manipulated in undisclosed ways can serve as a 'canary in the coal mine' – potentially alerting readers to generally poor research practice by a particular researcher, group, or lab. Consequently, in the context of paper mills in particular, the ability to assert with confidence whether an image is legitimate or not would be a valuable tool in the fight.

4 Current approaches to published image integrity

The integrity of images, along with other aspects of research integrity, has been a specific concern for some time. Surprisingly, however, there is very little consistency in how the publishing industry deals with this type of quality control. During the editorial process, editorial and peer-review work are the first two lines of defence; when a reviewer flags an integrity concern with an image, or any part of a manuscript, editors do their best to ascertain the validity of those concerns. If that cannot be done based on the evidence at hand, editors may ask authors to supply more information, and may, in more serious cases, notify an author's institution. Responses by institutions to concerns are varied, with each institution having its own set of policies and its own culture around research integrity.

Some journals are working towards increased requirements for certain types of images. One editor in chief of a journal we spoke to has instituted a requirement for authors to submit original, uncropped, unmodified images of blots, so that they can be checked against the final figure for consistency. On the other hand, the same editor, who is also a researcher, warned against creating 'huge amounts of bureaucracy', highlighting the need to make any system a net benefit for researchers. Similarly, a leader at a learned society that publishes a number of journals which rely heavily on imaging, spoke about balancing the need to ensure research integrity with not overloading researchers, or accidentally creating processes that aren't relevant to their specific community.

> *We are basically going slowly on this because it's very complex … there's many different types of data … and communities with different … data standards they require. So we are a little bit careful not to be too prescriptive. …our starting point right now is to encourage as many people to post their source data [as possible].*

In many cases, concerns are raised after publication, sometimes months or even years later. Image sleuths like Elizabeth Bik[3] and Jana Christopher[4] have been identifying manipulated and copied images for many years, and a number of publishers have adopted some of the same techniques to find manipulated images in submitted manuscripts. As one editor told us:

> *… there'd be a PubPeer thread typically saying this image looks dodgy … the problem is that with most of these cases, it's really difficult to track the original images … so we … have what we're given, [if] the original files are still in the archive then we can try and access those and then [in] Photoshop…, changing the contrast or exposure … to be able to highlight differences in background… non-homogenous backgrounds.*

Recently, some startups including Proofig AI, ImageTwin, figcheck, and ImaChek[5] have developed ways to detect copied images in submitted research papers; the more advanced of these can detect replication even if images are cropped, rotated, or modified in other ways. There is at least one tool available that is intended to detect AI generated images, created by Proofig AI, although the accuracy of such tools has not yet been fully validated. Many publishers are experimenting with

[3] https://scienceintegritydigest.com/
[4] https://image-integrity.com/
[5] More details about these companies can be found on their respective websites: https://www.proofig.com/, https://imagetwin.ai/, https://www.figcheck.com/, and https://www.imachek.com/

partnerships with new companies like these. A university press employee described their work in this area as follows:

> *…working alongside the product team, I also work with our trust and integrity program, and [have] some communication with our journals ethics teams …. I look at these technological advances that are coming up, as we look at these products for piloting*

Recently, Cytiva life sciences[6] have developed a system for digitally signing microscopy images. Their Image integrity checker, which is compatible with some of their Western blot imaging systems, generates a report that can be attached to a journal article submission and verified with free software that can be downloaded from their website. The image integrity checker is proprietary to Cytiva in its current form but is effectively a proof of concept for the use of digital signatures for image verification.

Considering all the approaches that are currently being tried to ensure image integrity, or to identify faked or manipulated images, it is clear that policing image integrity through editorial and review processes is not only time consuming for publishers and peer-reviewers, but also increasingly difficult – to the point that many experts consider it an impossible task. For this reason, we propose that editorial efforts be supplemented with a mechanism for assessing the integrity of images, by going upstream in the research process and certifying images at the point of original creation.

There are a variety of ways in which image certification could be done. Options include proprietary software applications supplied by each equipment manufacturer; a distributed trust store model similar to that used by secure websites and the https protocol; integration with a persistent identifier (PID) system; blockchain; and even trusted research environments, where every digital manipulation of an image could be recorded for later auditing. The requirements for the solution are that it is robust, scalable, trustworthy, and as automatic as possible to minimise the additional burden for all stakeholders. In the following sections of this report, we discuss the various technology choices available for creating a research integrity system and explain our rationale for each component of the proposed system.

[6] https://www.cytivalifesciences.com/

5 Asserting image integrity and authenticity

In this section, we describe a potential image integrity system that would enable an editor, peer-reviewer, investigator, or interested researcher to verify that an original image was created using a specific instrument, and that neither the image itself, nor its associated metadata, have been altered. The technologies we describe already exist, and the workflow has been validated by two of our interviewees who are technical experts in scholarly research infrastructure. While technically feasible, much of the system would require various stakeholders, like instrument manufacturers, manuscript submission systems vendors, and publishers, to implement software and integrate with shared infrastructure systems. It would also require the assent and buy-in of the research community. There would also be a need for cross-stakeholder collaboration and consideration of governance requirements.

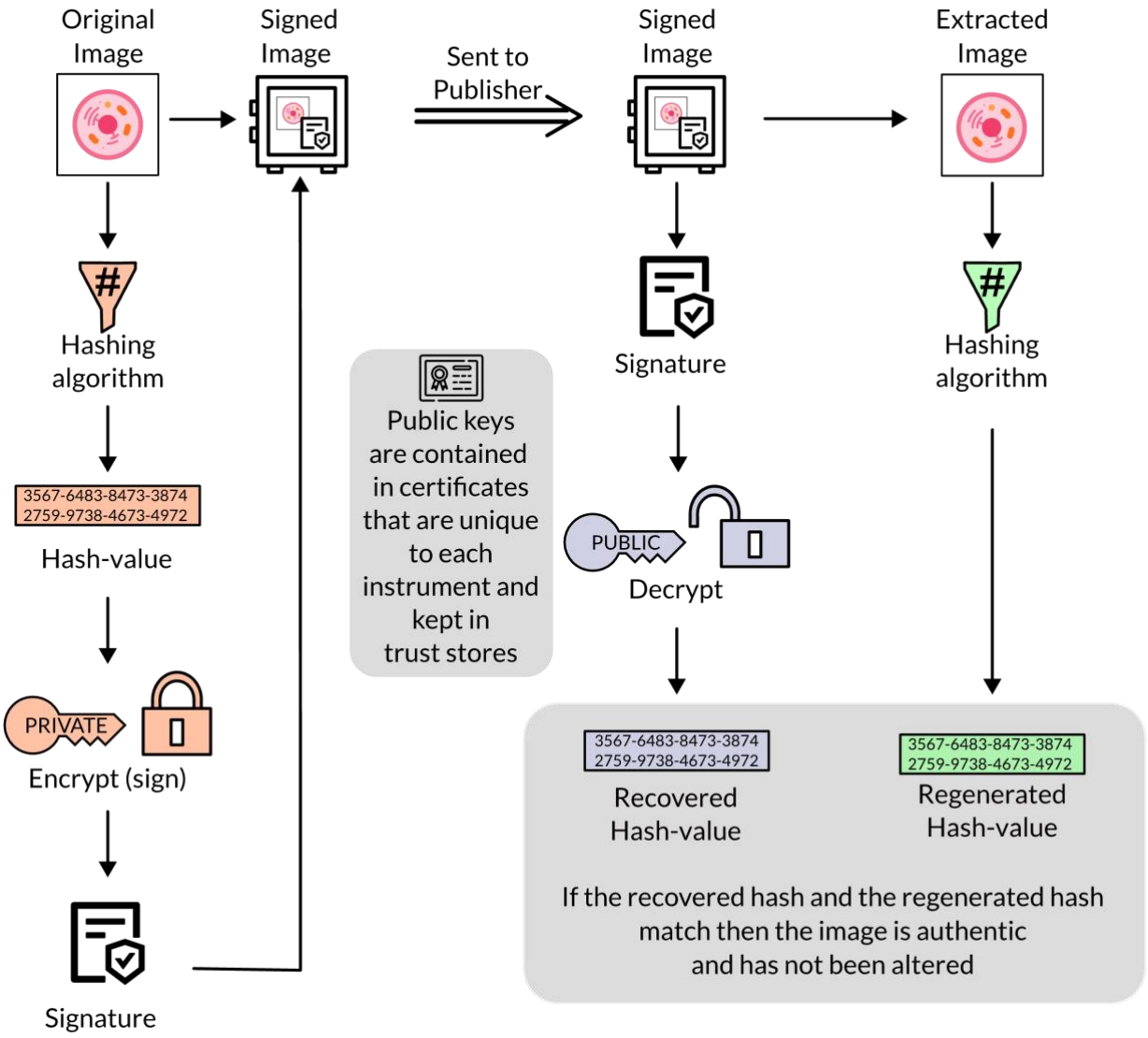


Figure 1: A graphical representation of how public cryptography can be used to assert the authenticity and integrity of a research image for publication

5.1 Digital signatures to assert the authenticity of an image

There are two parts to creating a digital signature, which can be sent alongside an image and then verified by the recipient.

1. A one-way process, in which the image (including its metadata) is reduced to a hash-value – a relatively short numeric or alpha-numeric string that is unique to the image
2. The hash-value is encrypted (or signed) using a private key. The resulting digital signature could also contain metadata about the image, for example, which instrument was used, a time stamp, or other information that might be useful in establishing authenticity of the image. The image and the signature can then be sent together as a package
3. Upon receiving the image and the signature, the recipient can regenerate the hash-value using the same algorithm from step 1 and decrypt the signature using a public key to recover the original hash-value. The image is verified if the regenerated and recovered hash-values match.

One challenge for the use of digital signatures is how to verify the signature so that the recipient can trust that the public key they use for decryption is associated with the instrument that the author claims. We will discuss options for establishing trust in sections 5.2 - 5.7.

5.1.1 Cryptographic hashes

Cryptographic hashing takes a digital object and transforms it, using an algorithm called a hash function, into a short alpha-numeric sequence, referred to as a digest or hash-value (5).

The power of this technique lies in the fact that the hash-value is not an encrypted version of the original data, but the result of a calculation based on the data. It is therefore impossible to recreate the original data based on the hash-value. On the other hand, the probability of a given hash-value is so vanishingly small that two datasets will practically never result in the same hash-value. The practical upshot is that, if two people run the same hash function on the same object, and generate the same hash-value, they can say with confidence that the objects are exactly the same.

On their own, hashes are not sufficient evidence that an image is authentic. For hashing to be useful, the function used to generate the hash-value must be known and well specified. Unfortunately, it would then be trivial for a bad actor to simply implement the hash function and apply it to a faked image. So, for hashes to address our problem, a method of securing and authenticating the hash-value is needed.

5.1.2 Digital signatures using public-key cryptography

In order to assure authenticity, it is necessary to show that the image originates from a trusted source. In this case, the trusted source would ideally be the instrument used for creating the image itself.

One way to do this would be for a piece of software installed on the instrument to combine the image with its metadata, for example, using the the Exchange Information File format (EXIF) for digital still cameras (6) or the Open Microscopy Environment data model (OME-XML) format (7) for microscopes, into a single object. While it is theoretically possible to then encrypt (sign) the data

object itself, the computational burden of doing so would be too high, especially for large files. Instead, the combined object would be hashed and only the hash-value would need to be signed.

The signing process uses another cryptographic algorithm that combines the hash-value with a secret alphanumeric string (private key) kept only on the instrument and not visible to anybody, including the users. The resulting signed hash-value is called a digital signature.

5.1.3 Verifying the signature

Associated with the private key is a corresponding public key, which can be used to decrypt the signature. The public key should be kept in a database or trust store (see section 5.3) in the form of a digital certificate. Importantly, the certificate would need to be publicly available so that anybody can extract the public key and use it to decrypt the signature, thereby recovering the hash.

In practice, the original data file could be sent to any interested party, say a journal editor, research integrity officer, or investigator, along with the signature. The interested party would use their own software to hash the image file, just as the instrument did originally, and would also decrypt the signature using the public key. Publishers might use stand-alone image verification software, or it could be built into existing editorial tools like manuscript submission systems as a feature of those applications. If the two hash-values match, then the image is authentic and has not been altered.

Certificates could also contain metadata about the instrument, such as the manufacturer name, model number, or even configuration and physical location, though any information that would be subject to change would have to be updated when needed. An additional level of integrity checks could, therefore, be made by comparing the instrument metadata contained in the certificate, with the metadata embedded in the signature and any information the author provides. For example, if the author claims that they generated the image using a particular instrument based at a particular university, but the certificate belongs to a different instrument, or the metadata in the signature has a timestamp that doesn't match the author's claimed research timeline, then those discrepancies could be investigated.

5.2 Establishing a trust domain

For the system to work, there is a socio-technical challenge to overcome – how does the interested party know which hash function and public key to use?

It is worth drawing a comparison at this point with how internet browsers ensure that a web page is authentic. On the web, trusted certificates are issued by certificate authorities (CA). They contain public keys and other metadata about both the CA and the certificate itself, such as the issuer's identity, the algorithm used for encryption, and the period of validity. Web browsers come preloaded with a list of trusted root certificates, which acts as the base layer of a domain of trust. In the system that Cytiva has developed, the image checker software comes preloaded with Cytiva's certificates, making the checking process simple and transparent to the user. That is to say, the end user doesn't need to know anything about the certificates, as long as they trust the software.

In order for digital signatures for images to become widely accepted and used, the system must be instrument manufacturer agnostic. Specifying exactly how to do that is beyond the scope of this

investigation, however, two possible approaches emerged from interviews: a centrally managed database of instruments with certificates; and distributed certificate registration.

5.3 A centrally managed database of instruments with certificates

One or more central databases of instruments could be created to contain certificates. They might be maintained by trusted organisations or by multi-organisational initiatives of research institutions, microscopy companies, and/or learned societies or publishers.

There are already some databases of microscopes and research instruments, often at the national or regional level. For example, the Royal Microscopical Society maintains a list of microscopy facilities in the UK, including their instruments[7]. These databases currently don't have exposed structured metadata about equipment, there are no standards associated with how the information is presented, and they don't offer computational integrations like APIs. Nevertheless, it should be possible to expand their use to act as trusted certificate authorities, with appropriate data structures, standards, and integrations. If one or more agreed databases of certificates were to be created and openly licensed, it– or they – could be integrated by software manufacturers (including image viewing and editing software like ImageJ[8], OMERO[9], GIMP[10], or Adobe Photoshop[11]), as well as by manuscript submission systems or other editorial software.

The advantage of this approach would be that, with appropriate governance, it would be manufacturer agnostic and leverage the reputation of the hosting organisation(s). On the other hand, centralised databases of this kind are challenging to scale, and present a significant curation burden for the hosting organisation. They present a single point of failure and can be vulnerable to attack by bad actors.

[7] https://www.rms.org.uk/community/facilities-database.html
[8] https://imagej.net/
[9] https://www.openmicroscopy.org/
[10] https://www.gimp.org/
[11] https://www.adobe.com/uk/products/photoshop.html

5.4 Distributed certificate registration

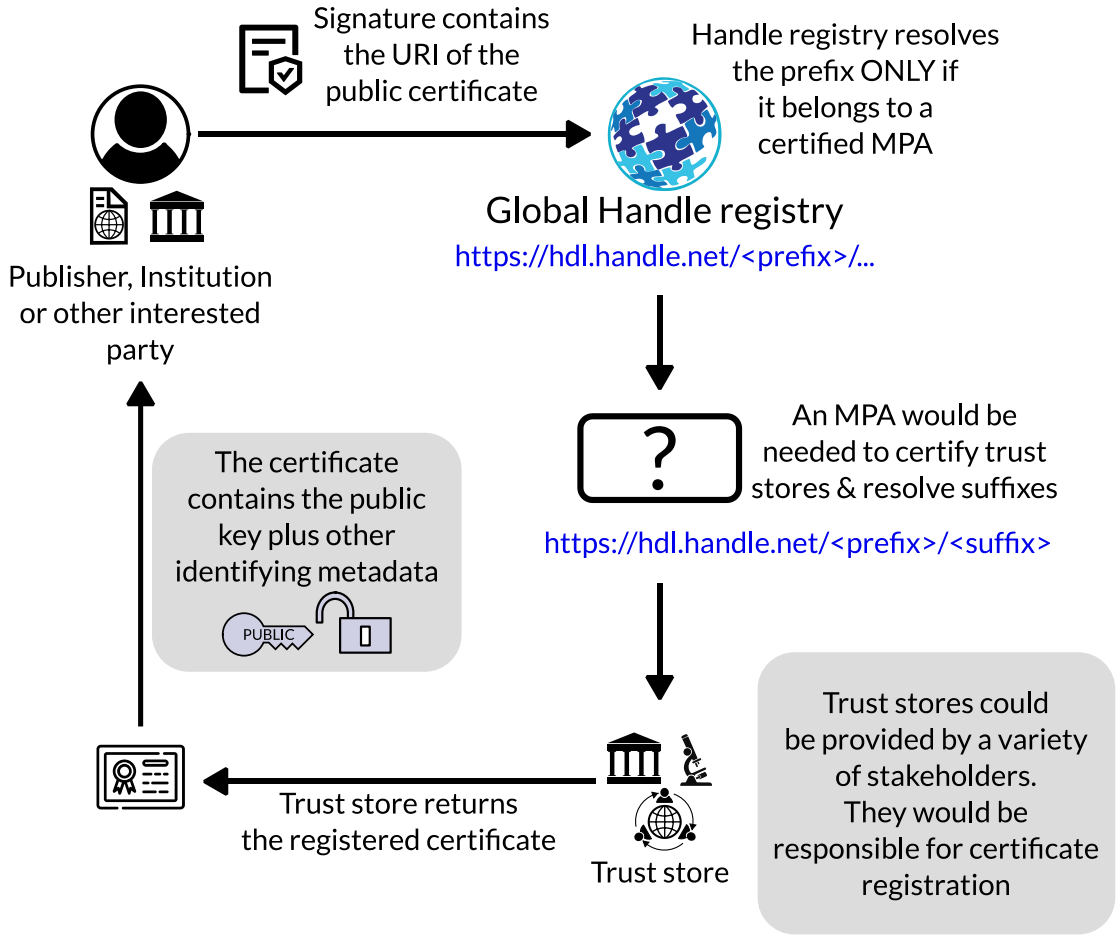


Figure 2: A graphical representation of one way that cryptographic certificates could be obtained from a distributed system of trust stores supported by the existing Handle infrastructure

A less centralised method for housing and curating certificates would be to use the Handle system or DOIs. Handles are a globally distributed network that persistably resolve identifiers to information resources (8) . Key to the system's scalability is that Handles are resolved in two stages, which correspond to two levels of distributed trust.

Figure 3: The two components of Handle identifiers are resolved in a two-stage process where the Handle system first looks up the prefix in the global Handle registry. The suffix is resolved by the specific Handle service

The prefix is handled by the DONA foundation[12], which resolves the Handle to the correct Multi-Primary Administrator (MPA). MPAs are accredited by DONA, with each MPA having one or more assigned two-digit credentials; for example, the DOI Foundation[13] has been assigned the credential 10. MPAs are responsible for resolution of the suffix, and in turn, they can accredit Registration Agencies (RAs), which register objects and may also provide other services, such as metadata registries.

One option would be for an existing DOI RA to provide a certificate registry as one of their products. Alternatively, a separate system could be developed, potentially in collaboration with DONA. This latter approach would have the advantage of enabling a range of organisations (national microscopy associations, research institutes, government agencies, universities, publishers, and others) to host trust stores of certificates, with unique identifiers resolvable through a consistent URI format.

5.5 Integration with a persistent identifier system

A way to drive adoption through synergy would be to integrate with a PID system. As will be described in section 6.6, considerable work has already been done towards developing a PID for instruments (9).

A combined PID system for instrument identification, with an appropriate metadata registry associated with a trust store for certificates, would be extremely powerful and serve many use cases. For example, it would: help funders track the impact of capital investments in instrument infrastructure; enable researchers to find collaborators with particular instruments; allow meta-scientists and researchers writing systematic reviews to connect grants and research outputs that use the same or similar instruments; and be an invaluable source of instrument metadata for publishers and learned societies as they build the next generation of research discovery databases and tools.

5.6 The possible use of blockchain technology

As an alternative to using public key infrastructure as described above, another approach to decentralised trust would be to use blockchain to create immutable, time-stamped records of images. Each record or 'block' in the chain might include both a hash-value for the image and a set of metadata documenting the history and provenance of the image.

The advantages of using blockchain are that it would be entirely decentralised and immutable, with no hierarchy of trust. Blockchain is less vulnerable to hacking as there are no organisational IT systems to target. It would also, if designed correctly, be impossible to alter metadata or provenance information without the change being documented and traceable.

The disadvantages are that blockchains can become slow if overwhelmed with high numbers of certificate transactions, causing scalability issues, and they are inherently computationally inefficient, resulting in carbon-footprint concerns. Meanwhile, traditional public key approaches are

[12] https://www.dona.net/
[13] https://www.doi.org/

well established, with more predictable costs, mature regulatory frameworks and standards, and they work well with existing infrastructure. For these reasons we have not recommended blockchain in this report.

5.7 Open source authentication software

Once a mechanism has been put in place for certificates, the final remaining challenge is the availability of software to authenticate images.

Cytiva have developed an 'Image Integrity Checker'[14] — free software for anyone to download, with the certificates for their digitally signed images pre-loaded. The advantage of this approach is that it is simple for an individual to authenticate that an image comes from a specific instrument with the claimed configuration at the correct date and time. From a sector-wide perspective, this solution represents a proof of concept for image authentication but doesn't on its own solve the problem because the image checker software, although free, is proprietary.

A solution that works for all equipment will need software that is manufacturer agnostic. The most effective way to deliver this is through a set of open source tools, including a stand-alone image verification tool, plugins for manuscript submission systems and repositories, and APIs that enable third-party integrations.

[14] https://www.cytivalifesciences.com/en/us/solutions/protein-research/products-and-technologies/western-blotting/image-integrity-checker

6 Socio-technical challenges

The technology to enable certification and verification of images already exists; technology and organisational choices will need to be made but, the solution proposed here includes a combination of embedded metadata, cryptographic hashes, public cryptography, and a trust store.

As previously noted, the socio-technical challenges around adoption of technology are significant. Addressing image integrity, like all aspects of research integrity, requires collective action from instrument manufacturers, funders, research institutions, publishers, learned societies, and researchers themselves.

6.1 Establishing the need among instrument manufacturers

Instrument manufacturers, like many businesses, are led by the demands of their customers. As one representative of a microscopy company stated during an interview:

> *For us as a company, we can assist, of course, maybe we also have some experts sitting in our software departments who can say something in this context, but in the end, whatever they decide, [it's about] whatever the community decides …*

So, from the perspective of many instrument manufacturers, for a new feature to be put into place, there needs to be a strong signal from the community that it is needed. This is sometimes a challenge in the microscopy space, as the same interviewee stated:

> *We are [involved in] a lot of initiatives from the community, and a lot of them absolutely make sense, but following them for 10 or 15 years, still there are no standards … before we, as a company, and I could imagine, the same is also true for [our competitors] to start thinking about investing in R & D, we really want to have the guidelines*

The work that Cytiva has done in demonstrating that it is possible to certify a microscopy image and verify it in a separate piece of software shows that there is interest among instrument manufacturers to advance research integrity. That said, for instrument manufacturers to adopt an image integrity solution, they must agree that it is both the right solution and that it is required by the community.

From our interviews, we learned that there is no single, consistent way that instrument manufacturers learn about requirements. Rather, there is a mix of user groups, community initiative participation, engagement with learned societies, outreach efforts, and collaborations. The most effective way to engage with instrument manufacturers is, therefore, likely to be through a committee or working group made up of a variety of stakeholders. Demonstrating broad, cross-sector support will be vital to getting buy-in and adoption.

6.2 Standards and standardisation

For a system to be implementable in a standard way, there will need to be a set of standards for implementation, including a choice of cryptographic algorithms, metadata, and file formats. In addition, there needs to be agreement among all stakeholders about how certificates are

distributed, whether via a centralised database or a decentralised series of trust stores. As a product manager at a large microscopy company said:

> *I absolutely could imagine that this will be a standard requirement in the future…so if the community and journals … agree and that something like that makes sense … there has to be a standard, a clear definition …What are the specifications? What does the metadata have to look like?*

6.3 Earning the trust of researchers

Researchers are time poor, largely due to unnecessary bureaucracy and poorly implemented information systems (10–13). Many researchers and research managers have understandably grown suspicious of initiatives which they suspect may result in more bureaucratic overhead for little benefit. It is therefore crucial to the success of any system that it does not add burden for researchers and does not, unless absolutely necessary, involve extra administrative work. If the system is designed correctly, then it should not create extra work for researchers, as the signing process should be seamless and automated for the end user.

For researchers to buy into a system, it must generate value for them that is immediately obvious from the beginning. A good illustration of this need comes from something a senior researcher said:

> *…I think it'd be super useful for me to be able to track our images back to a raw image and track them all the way through, if it was useful to your research, people would buy into it, if it was useful to your research and didn't cause huge amounts of admin bureaucracy even better. It would have to, at a minimum, comply with the regulations that we already have. And you would not believe how many regulations we have about various types of things that we can do electronically with data*

6.4 Impact on publisher workflows

Any research integrity initiative carries a risk of negatively impacting editorial workflows if extra burden if extra checks or processes need to be put in place, particularly if there is a lot of manual administrative work associated with them. On the other hand, it is possible to reduce that impact with well-thought-out workflow design (14) and appropriate automation.

Careful attention should be paid to ensuring that as much as possible of the process of verifying images is automated. This could be done by embedding verification into existing workflow software as a feature or by creating simple, user-friendly, stand-alone tools that enable quick and batch screening of images.

6.5 Establishing and governing trust stores

For the system to be universally adopted and accepted as a source of trust, it must be developed and implemented in such a way as to inspire that trust. Requirements for trust stores may vary by geography, discipline, and image type. Balancing simplicity of implementation with flexibility requires careful decision-making about technology, architecture, and organisational structures. It's

also important that the system is seen as fair, and not serving the interests of any particular stakeholder or stakeholder group.

6.6 Aligning with other initiatives to create greater value

As described in section 4, there are a few approaches currently being employed to try to assure both image integrity and research integrity in general. With so much going on in the publishing and scholarly communications sector, it's important to collaborate across organisations and initiatives when there are synergies and alignments to be gained. A thorough landscape analysis is beyond the scope of this investigation, but we highlight three particularly relevant projects below.

The Research Data Alliance (RDA)[15] is a community organisation, started in 2013 by the European Commission, the American National Science Foundation (NSF), the National Institute of Standards and Technology (NIST), and the Australian Department of Innovation. Through a series of interest groups and working groups, RDA establishes communities of practice and creates recommendations in a standards-like process. The RDA's Persistent Identification of Instruments Working Group (PIDINST) has done excellent work towards developing a PID infrastructure and common metadata schema (9), working with infrastructure providers DataCite[16], ePIC[17], and institutional instrument centres Helmholtz-Zentrum Berlin für Materialien und Energie[18] (HZB) and the British Oceanographic Data Centre[19] (BODC).

The Coalition for Content Provenance and Authenticity[20] (C2PA) seeks to address issues of online misinformation through technical standards for certifying media content, such as news photographs and video. While this initiative is not specific to scientific imaging, there is significant overlap in use cases and technology. C2PA is hosted in the Microsoft Azure cloud and makes use of the Microsoft key vault[21] as a trust store.

The European Open Research Cloud (EOSC)is a European Commission initiative to provide a shared environment for research data, services, and tools. Open research and research integrity are important to EOSC, so there may be potential for them to support or collaborate on this work. As will be described in section 7, in the future there is a potential for signed research workflows that could enable entire workflows to be verifiably replicated.

[15] https://www.rd-alliance.org/
[16] https://datacite.org/
[17] https://www.pidconsortium.net/
[18] https://www.helmholtz-berlin.de/
[19] https://www.bodc.ac.uk/
[20] https://c2pa.org/
[21] More information on the Microsoft key vault can be found here: https://azure.microsoft.com/en-us/products/key-vault

7 Going beyond microscopy and blots

The technical approach described above would, in principle, work with any data type where the primary raw data is created under controlled conditions. For example, any instrument, from a cell sorter to a seismograph, could have software embedded to certify the raw data files.

On the other hand, in real research environments, not all data originates with instruments that are purchased from a manufacturer. In some cases, measurements are taken with instruments built by the researchers themselves. Data may also be read from analogue instruments, entered by hand into spreadsheets, or written into physical notebooks. Some disciplines use well-established digital standards and standard software. Some of these are developed and supplied by manufacturers, like the SPCM software for fluorescence lifetime imaging developed by Becker and Hickl (15), or the computer-aided qualitative analysis software atlas.ti (16); some are sponsored by research institutes, like ImageJ from the National Institutes for Health (17); and some are community-driven efforts, such as the Ancient Demographic Modelling Using Radiocarbon (ADMUR) software (18).

Adding image integrity features to proprietary software will require outreach efforts to software developers. Community-driven and institute-sponsored software tends to be open source and might be modified by the community to include image integrity plugins – provided that the community understands and supports the feature. However, as we noted in section 6, integration with research software is only part of the requirements. For the system to become multi-disciplinary, it must be flexible enough to serve the needs of each community and may require customisation to disciplinary and regulatory needs. Distributed certificate registration would be helpful here, for its ability to scale, and flexibility to adapt to the needs of different communities.

8 Limitations and further considerations

The image integrity system described above would be an effective mechanism to certify that a particular image originates from a specific instrument, and to make information about that instrument easily accessible and verifiable. This characteristic of the system makes it an extremely effective weapon against paper mills and those that might use generative AI to create images.

One significant challenge in adopting this system would be the availability of instruments that support the generation of certificates. It may be possible that some instruments can be updated to support the system, but it is quite likely that significant numbers of non-certifying instruments would remain in use for years into the future. There is also a risk that patchy coverage might create inequality; if publishers begin to require certified images for publication, that may exclude researchers in less well-funded institutions from publishing. Careful consideration will need to be given to how to encourage adoption without excluding those who have access to older instruments. One mitigation could be to create image repositories that could contain certified images on upload. These could be associated with institutions, or other authorities and, while not as secure as certification through the instrument itself, such an approach could be considered as a fall-back approach when needed.

As with all digital architecture, environmental issues should be considered. As the system is designed and developed, careful attention should be paid to minimising its carbon footprint. If the system were to be based on Handles, as suggested in section 5.4, then the Handle implementation would likely be a significant contributor. CNRS, which administers Handles, is involved in efforts to decrease the environmental impact of their infrastructure by optimising data centres, using lower-energy hardware, and increasing the use of renewable energy sources, in line with France's GES 1point5 framework (19).

As mentioned in section 2, the system we describe in this report is not intended to guarantee that images or data are factually accurate, because that would be impossible. For instance, proving that an image came from a particular instrument does not preclude the possibility that many other images have been generated and that the published one is not representative. Even worse, it's impossible to prevent a rogue experimenter from making inaccurate claims about what was imaged, for example, which antibodies or stains were used before the instrument was used to image the sample.

Another limitation is that there is no tracking of what happens to the image or data after it leaves the instrument. It would therefore be necessary for an editor or investigator to compare the original, certified image with the final published one, making a judgement as to whether the author's description of their data pipeline adequately explains any difference between the two images.

On the other hand, the system we describe above, or something similar, could eventually be extended to record, certify, and verify an entire digital research environment. There are significant challenges to achieving this, however, some of the necessary components for creating certified reproducible research environments have already been developed.

In 2007, the myExperiment project (20), funded by BioVeL[22], EPSRC[23], Jisc[24], and Microsoft, released an online research environment to support the social sharing of bioinformatics workflows. The technology built on work done to deploy data and tools as web services by the European Bioinformatics Institute (22) and the DNA database of Japan (23), among others. It enabled the documentation of a research workflow in a way that could be reproducibly recreated using a standard tool set by any researcher. More recently, the European Open Science Cloud (EOSC) has been seeking to create a cross-disciplinary research environment for data and research tools and services (24). Meanwhile the RO-Crate project (25) is taking a structured approach to packaging all items associated with a research outcome, along with their descriptive metadata, which is currently being used in bioinformatics, digital humanities, and regulatory sciences.

These efforts to create common research environments that can be used by any researcher offer the promise of both more efficient working through resource-sharing and greater reproducibility. While the various technologies and systems required are still being developed, and much work is needed in terms of both adoption and scalability, a workflow package could eventually be used to reproducibly document every step in a research process – from data acquisition through to final creation of a publication figure. In the event of investigation, or if another researcher needed to replicate the work, the entire toolchain could be redeployed and each step inspected.

Separately, there has also been rapid progress in the field of electronic lab notebooks (ELNs) in recent years. A number of companies have developed commercial offerings including RSpace[25], Labguru[26], Dotmatics[27], and several others; while uptake in academia has been inconsistent, the pharmaceutical and related industries have made good use of ELNs to improve both discoverability and reliability of their research, particularly at the pre-clinical stage (26). There is potential to integrate ELNs with certified workflows to establish proof of the integrity of a reported research workflow.

[22] A forerunner to the Apache Taverna workflow management system(21)
[23] https://www.ukri.org/councils/epsrc/
[24] https://www.jisc.ac.uk/
[25] https://www.researchspace.com/
[26] https://www.labguru.com/
[27] https://go.dotmatics.com/

9 Bibliography

1. Nosek BA, Hardwicke TE, Moshontz H, Allard A, Corker KS, Dreber A, et al. Replicability, Robustness, and Reproducibility in Psychological Science. Annu Rev Psychol. 2022 Jan 4;73:719–48.
2. Paper mills research [Internet]. Committee on Publication Ethics and STM; 2022 Jun [cited 2024 Apr 16]. Available from: https://publicationethics.org/node/55256
3. Christopher J. Systematic fabrication of scientific images revealed. FEBS Lett. 2018 Sep;592(18):3027–9.
4. Krämer K. AI-generated images could make it almost impossible to detect fake papers. Chemistry World [Internet]. 2022 May 24 [cited 2024 Sep 19]; Available from: https://www.chemistryworld.com/news/ai-generated-images-could-make-it-almost-impossible-to-detect-fake-papers/4015708.article
5. Menezes AJ, Oorschot PC van, Vanstone SA. Handbook of Applied Cryptography. CRC Press; 2018. 811 p.
6. Exchangeable image file format for digital still cameras: [Internet]. 2023 [cited 2024 Sep 6]. Available from: https://www.cipa.jp/std/documents/download_e.html?DC-008-Translation-2023-E
7. Goldberg IG, Allan C, Burel JM, Creager D, Falconi A, Hochheiser H, et al. The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging. Genome Biol. 2005 May 3;6(5):R47.
8. Lannom L, Boesch LCBP, Sun S. Handle System Overview [Internet]. Internet Engineering Task Force; 2003 Nov [cited 2024 Sep 11]. Report No.: RFC 3650. Available from: https://datatracker.ietf.org/doc/rfc3650
9. Stocker M, Darroch L, Krahl R, Habermann T, Devaraju A, Schwardmann U, et al. Persistent Identification of Instruments. Data Sci J. 2020 May 5;19:18.
10. Tickell A. Independent Review of Research Bureaucracy: final report [Internet]. 2022 Jul p. 63. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1094648/independent-review-research-bureaucracy-final-report.pdf
11. Havergal. REF reforms to seek 'accountability without complex bureaucracy'. Times Higher Education (THE) [Internet]. 2020 Oct 20 [cited 2022 Nov 11]; Available from: https://www.timeshighereducation.com/news/ref-reforms-seek-accountability-without-complex-bureaucracy
12. Husain M. The three deceits of bureaucracy. Brain. 2022 Jun 1;145(6):1869.
13. Brown, Josh, Jones, Phill, Meadows, Alice, Murphy, Fiona. Incentives to invest in identifiers: A cost-benefit analysis of persistent identifiers in Australian research systems [Internet]. Zenodo; 2022 Sep [cited 2022 Nov 7]. Available from: https://zenodo.org/record/7100578
14. Holt J, Walker A, Jones P. Introducing a data availability policy for journals at IOP Publishing: Measuring the impact on authors and editorial teams. Learn Publ. 2021 May 6;leap.1386.
15. Becker W. Fluorescence lifetime imaging--techniques and applications. J Microsc. 2012 Aug;247(2):119–36.
16. Soratto J, Pires DEP de, Friese S. Thematic content analysis using ATLAS.ti software: Potentialities for researchs in health. Rev Bras Enferm. 2020;73(3):e20190250.
17. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012 Jul;9(7):671–5.
18. Timpson A. ADMUR: Ancient Demographic Modelling Using Radiocarbon [Internet]. 2020 [cited 2024 Sep 22]. p. 1.0.3. Available from: https://CRAN.R-project.org/package=ADMUR

19. Mariette J, Blanchard O, Berné O, Aumont O, Carrey J, Ligozat A, et al. An open-source tool to assess the carbon footprint of research. Environ Res Infrastruct Sustain. 2022 Sep 1;2(3):035008.
20. Goble CA, Bhagat J, Aleksejevs S, Cruickshank D, Michaelides D, Newman D, et al. myExperiment: a repository and social network for the sharing of bioinformatics workflows. Nucleic Acids Res. 2010 Jul 1;38(suppl_2):W677–82.
21. Hardisty AR, Bacall F, Beard N, Balcázar-Vargas MP, Balech B, Barcza Z, et al. BioVeL: a virtual laboratory for data analysis and modelling in biodiversity science and ecology. BMC Ecol. 2016 Dec;16(1):49.
22. Mcwilliam H, Valentin F, Goujon M, Li W, Narayanasamy M, Martin J, et al. Web services at the European Bioinformatics Institute-2009. Nucleic Acids Res. 2009 Jul 1;37(Web Server issue):W6–10.
23. Kwon Y, Shigemoto Y, Kuwana Y, Sugawara H. Web API for biology with a workflow navigation system. Nucleic Acids Res. 2009 Jul;37(Web Server issue):W11-16.
24. European Commission. Directorate General for Research and Innovation. Realising the European open science cloud: first report and recommendations of the Commission high level expert group on the European open science cloud. [Internet]. LU: Publications Office; 2016 [cited 2020 Nov 26]. Available from: https://data.europa.eu/doi/10.2777/940154
25. Soiland-Reyes S, Sefton P, Crosas M, Castro LJ, Coppens F, Fernández JM, et al. Packaging research artefacts with RO-Crate. Peroni S, editor. Data Sci. 2022 Jul 20;5(2):97–138.
26. Edfeldt K, Edwards AM, Engkvist O, Günther J, Hartley M, Hulcoop DG, et al. A data science roadmap for open science organizations engaged in early-stage drug discovery. Nat Commun. 2024 Jul 5;15(1):5640.