Thought Leader

Proofig AI

An AI tool to ensure image integrity in scientific publications

Proofig AI is an innovative and reflexive AI-powered proofing tool, aimed at researchers, editors, institutes, and publishers. The tool aims to identify the improper usage of images in scientific publishing and beyond. Research Outreach spoke with Dr Dror Kolodkin-Gal, the founder of Proofig AI. He shares insight into the challenges scientific publishing faces with the publication of fraudulent manuscripts, and how AI image proofing tools can help encourage proper practice in scientific publishing.

Being able to identify mistakes, and indeed demarcate between human error and intentionally fraudulent manuscripts, is a major issue in scientific publishing. What’s more, the detection of adverse image manipulation is increasingly difficult, requiring time and accuracy to find, monitor, and block. So, how can we continue to identify errors in scientific research, as well as identify whether they are intentional, or just honest human error?

Research Outreach was thrilled to speak with Dr Dror Kolodkin-Gal, the founder of Proofig AI. An AI-powered image proofing tool for scientific publishing, Kolodkin-Gal outlines how and why Proofig AI was founded and offers a unique insight into solving the present and future challenges scientific publishing is facing.

Please could you introduce yourself and explain what led you to found Proofig AI?

I’m Dr Dror Kolodkin-Gal, founder of Proofig AI. I started the company in 2017 after a long research journey in virology, obtaining a PhD, followed by cancer research, as well as my time as a post-doc researcher fellow at the Beth Israel Deaconess Medical Center and Harvard Medical School.

During my time as a researcher, colleagues and I submitted papers – knowing that publishing is highly competitive, and that leading scientific journals accept only a small portion of the research they receive. While the scientific community understands the competitive nature of publishing, few understand the issues that can occur post-publication if the paper contains mistakes. After seeing the toll paid by fellow researchers that experience reputational damage and even retractions after unintentionally including image duplications, I began to understand the importance of image quality assurance before publication.

Reviewing image duplications within a paper, such as rotations, overlaps, or flipped images can be very difficult by eye. These forms of duplication are often honest mistakes, but if they are reported post-publication by a reader, it can lead to a costly investigation. No matter the outcome, these investigations can damage the reputation of the author, their research institute, and the publication journal itself.

To help reduce the risk of publishing mistakes, I developed Proofig AI. This automated image integrity detection software uses artificial intelligence (AI) to rapidly detect most image duplication and some forms of manipulation in scientific manuscripts. The Proofig AI team works closely with reputable universities, publishers, and image integrity experts ensuring the tool is fit for purpose.

How common do you think image plagiarism or misuse is in research?

It is difficult to accurately determine how many papers contain forms of image misuse; first we must understand how many papers contain image duplications or manipulations, and then determine if these are innocent mistakes or intended misuse.

According to the Public Library of Science, 72% of all biology papers include image-based figures to convey results. Additionally, a single paper contains multiple sub-images of each microscopy sample and the more sub-images there are, the more comparison between them. For example, if a paper contains 100 sub-images, there will be 10,000 comparisons. Can you confidently say that you could accurately review 10,000 comparisons by eye?

While accusations of image misuse are regularly referenced, most investigations find that image duplication is an honest mistake.

These unintentional image duplications occur for a variety of reasons, but are often introduced when representing the various stages of experimentation. For example, when looking at a sample under the microscope, researchers may want an image of the entire specimen. Depending on the magnification, the researcher must move the microscope from left to right and up and down to document the slide – which may lead to overlaps.

Effective use of AI technology can help maintain the reputation of researchers, enhance the impact factor of scientific journals, and reduce the need for costly investigations and retractions.

With software like Proofig AI, users can complete checks quickly and accurately, streamlining the image review process.

Proofig AI aims to enhance scientific publishing by enhancing image integrity awareness.

Developing automated duplication detection is a major step towards maintaining the rigour of our review processes.
when capturing images. Moreover, if researchers change the magnification, they may accidentally capture the same part of the sample twice.

Data organisation can also cause unintentional duplications. During experiments, researchers may take thousands of images from different samples before choosing the representative images to include in the paper. Labelling these effectively is key to ensuring researchers use the correct images, and only use them once. This is particularly important when researchers from different institutions collaborate on a paper, if research data is disorganised, it may be very difficult to identify which figures have already been featured.

These honest mistakes may be published unintentionally because editors will find it very difficult to see every form of duplication by eye, meaning there is no guarantee that they will detect issues before submission and publication.

How can researchers, editors, institutes, and publishers benefit from using Proofig AI?

With software like Proofig AI, users can complete checks quickly and accurately, streamlining the image review process. This will also help to improve quality assurance and maintain the reputation of researchers, enhance the impact factor of scientific journals, and dramatically reduce the necessity for costly investigations and retractions. The tool is an automated software that uses AI to detect all the sub-images in a scientific paper. It compares images against one another, producing a report outlining different forms of duplication within minutes. Once complete, the user can review the flagged issues. By identifying these issues before publication, the user can make informed decisions by asking questions such as: is it an error made in good faith? Does it impact the research results? Does it warrant correction, or should the editor reject the manuscript instead?

How can researchers and publishers strive towards high standards of publication integrity?

As we build more awareness about the importance of image integrity, researchers, editors, and publishers can take proactive steps to make improvements. Publishing houses for example can look at technologies and processes that will help them detect any image issues at every stage of article submission. Integrating software, such as Proofig AI, into editorial systems enables editors to automate the image review process, effectively scanning every image and comparing them. Each image is checked against itself and against the rest of the images to detect any anomalies and produce them in the report. Editors can use the report to check any flagged images in more detail, deciding if they are allowable edits, or if they should return the paper to the researcher to make amendments, or if they should reject the paper entirely.

Researchers can also take proactive steps to reduce the risk of submitting articles with unintentional image integrity issues. During the experimentation process, labelling files clearly helps researchers keep track of images. They can also use software to scan their own images beforehand so they can resolve any honest mistakes before submission – just like they’d check their grammar before submitting written content.

How can we better identify image integrity issues in research and publishing?

While we have found that most duplications are unintentional and can therefore be resolved by checking papers, we must also improve image integrity to prevent the publication of fraudulent papers and paper mills, which are organisations that produce and sell research papers with many manipulations and integrity issues. The recent advancements in AI writing assistants have sparked some interesting conversations in the scientific publishing community, with people discussing how we’ll be able to determine between human and artificial writers. However, we also need to think about images; AI tools already exist and are consistently improving, so soon it may also be possible to create fraudulent figures for scientific papers.

As these AI tools and paper mills become more advanced in hopes to get through the publishing process, the industry must act. Fraudulent papers are difficult to detect in isolation, and it requires investigators to compare multiple papers from different sources, looking for unusual similarities in text and figures. To effectively reduce the publication of fabricated research, we must take collective action to increase awareness in the scientific community, develop a process to detect fraudulent manuscripts, and develop technologies to address the problem.

What’s on the future agenda at Proofig AI?

While the scientific research community is aware that there can be issues with paper manipulation and duplication, few understand how frequently these issues occur. Our current and future aim is to increase awareness around image integrity in scientific publishing so that we can collaborate with publishers and researchers to improve quality assurance.

As we understand more about image integrity, the scientific publishing community can take more effective steps to reduce the risk of retractions caused by honest mistakes, while also understanding how to prevent fraudulent paper and paper mills from getting through the review process. Paper mills are constantly looking for creative ways to slip through publication integrity standards. As we see further advances in AI image creators, we need increased awareness and to develop sophisticated tools to identify inappropriate use of these assistants. We currently work with integrity teams to help develop our software and stay on top of any new trends of image issues.

In our opinion, image integrity checks will still require human intuition. From our experience working with editors and publishers on Proofig AI, image detection is working well; AI is there to detect images, make comparisons, and flag potential issues. The user must then, however, investigate the flagged issues themselves – looking in more detail to understand if the image is legitimate duplication, for example an enlargement of the same specimen that is also properly referenced in the text, or a real issue.

We must take collective action to increase awareness in the scientific community, develop a process to detect fraudulent manuscripts, and develop technologies to address the problem.

Dr Dror Kolodkin-Gal, Founder of Proofig AI.

www.proofig.com

www.researchoutreach.org