Ripeta

Enhancing scientific integrity

Ripeta is an automated tool designed with a simple but vital objective: to enhance the credibility, reproducibility, and trustworthiness of scientific research papers. Its algorithms use several important metrics to assess the integrity of a paper, stemming the tide of untrustworthy science one paper at a time. Research Outreach caught up with Dr Leslie McIntosh, Ripeta's founder and CEO, to talk about the origins and objectives of the company, and to discuss the difficulties of open scientific publishing more generally.

In the absence of any codified criteria to assess the veracity of scientific research, open scientific publishing risks becoming inundated with questionable papers. This has never been clearer than during the COVID-19 pandemic. The rapidity with which research has been undertaken, reviewed, and published greatly increased the risks of misinformation. In addition, the borders of the scientific research community have become porous, as a general readership look to broaden their knowledge of public health.

In this context and beyond, it has never been more vital for papers to be assessed using rigid indicators of reliability. Ripeta does precisely this, offering a simple, clear, and rigorous tool for assessment. Dr Leslie McIntosh, founder and CEO of Ripeta, told *Research Outreach* more about the important task of the company.

Could you tell us about the professional experiences which led to you founding Ripeta?

The idea for Ripeta started when I was a faculty member at Washington University in St Louis, US. I worked with hundreds of researchers as the Director of the Center for Biomedical Informatics (CBMI) by providing them with data (e.g., electronic health record data, biospecimen data), which they would then use for their research. Unfortunately, the researchers did not cite the CBMI in publications. I

mentioned this omission of credit to physician-researchers in my class. One physician asked, 'Why don't you make it easy for us?'

After the initial frustration with that answer, I thought about the physician, who had just finished a long day of clinicals and was then taking an evening class to learn how to do research. Why shouldn't I make it easier for her to cite data rather than burdening her with more work?

Soon after, I approached the university's data librarian, Cynthia Hudson Vitale, to get her thoughts. One thing led to another as we explored what it takes to make a paper robust enough to be reproducible (i.e., what needed to be included in a paper for the research to be reproduced?). Combining my informatics skills with her informational knowledge, we eventually started Ripeta to identify, then automate, quality checks of research manuscripts.

How does Ripeta work, and how is it seeking to improve scientific publications?

Ripeta works through automation. We build algorithms to detect three key indicators of trust – research, professionalism, and reproducibility – using machine learning and natural language processing (NLP). In essence, we conduct hygiene checks of the reported research and offer suggestions for improvement when possible.

Ripeta offers a simple, clear, and rigorous tool for assessing scientific research papers.

Dr Leslie McIntosh is founder and CEO of Ripeta.

In terms of seeking to improve scientific publications, we were founded by researchers looking to solve the problems that many publishers, funders, institutions, and fellow researchers have come across: unreproducible and untrustworthy science. So, our goal is to make trustworthy science easy and accessible to different stakeholders. For example, the *ripetaReview* can quickly scan papers for the key quality indicators I mentioned before: trust in reproducibility, trust in professionalism, and trust in research. We are also in the process of creating a dashboard that publishers,

We conduct hygiene checks of the reported research and offer suggestions for improvement when possible.

funders, and institutions can use to get a sense of the quality of science across a variety of publications through an aggregated approach. In addition to our products, we work to publish papers, blog posts, and social media campaigns to inform and illuminate best practices in research publications.

Could you provide more detail on the factors which contribute to the integrity of a scientific research paper?

When determining the integrity of a paper, we first check to make sure it is indeed a research paper. This means that it includes a study objective as opposed to, say, an opinion piece. After that, we look for indicators of

trust in professionalism: does this study include an ethical statement? Is the author really a scientist? Then, trust in reproducibility: could this study potentially be reproduced with the current information available (e.g., data availability statements, detailed methods)? These are all questions we may ask when scanning a paper in our head. We have now automated this process to make it easier and faster for the user.

Has Ripeta worked with researchers and publishers across a range of scientific disciplines?

Definitely. Originally, we interviewed researchers from disciplines such as clinical medicine and anthropology

to understand different data sharing and publication practices. We now train our algorithms on a variety of disciplines and publications because research varies across disciplines and publications including: technology, social studies, humanities, environmental science, core

Ripeta works to ensure the veracity of scientific research.

sciences, clinical, biomedical, biology, and other health sciences.

Could you talk to us about some of the complications of open scientific authorship, particularly within a pandemic context?
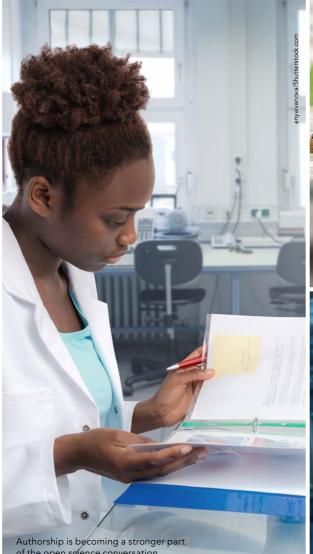Oh, I have a lot of thoughts here, so bear with me.

To begin, we do not have a framework, governance, or process for how open science should be trusted – even within the scientific community. For reference, open source code has norms within programming communities that people outside of that group may not understand. In open science, we have yet to establish these norms.

The pandemic has futhered this conversation; we now see the challenges and opportunities of open science in a more foundational way. There has been a surge in preliminary data release on preprint servers (which is where we find many of the issues in open science), rapid reviewing, and more. Providing research and data early and making them openly available can greatly improve the speed and spread of scientific information – as well as misinformation.

In terms of professionalism, this nebulous ecosystem has allowed fake authorship of preprints to move into the system without a coordinated means to question the scientific work. For example, we highlight one author in our paper, Imposters and Impersonators, named Kira Smith (whom our team has become infinitely closer to due to all our research on them). Smith uses different open science platforms to share the same work, which is generally poor practice.
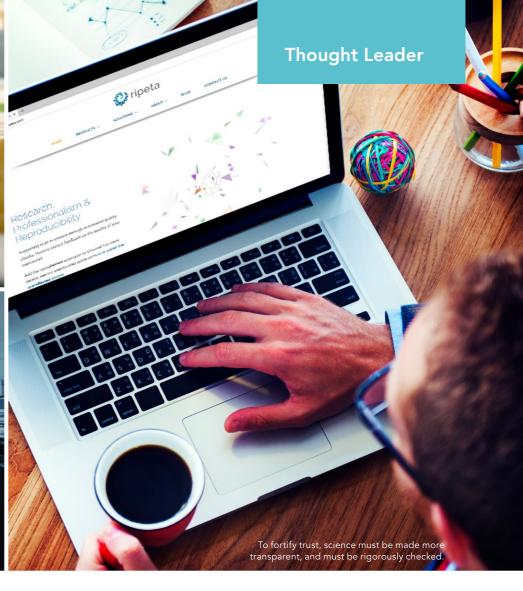
Thus, complications of open science result in asking 'How do scientists trust scientific authorship?' and 'How do non-scientists trust the scientists?' Authorship, a pillar of scientific trust, is becoming a stronger part of the science and open science conversation because it is being questioned now more than ever.

To fortify open science, we need to bring to light these issues that

Authorship is becoming a stronger part of the open science conversation.

How do scientists trust scientific authorship?

The *ripetaReview* will quickly automate quality checks.

To fortify trust, science must be made more transparent, and must be rigorously checked.

> **Providing research and data early and making them openly available can greatly improve the speed and spread of scientific information – as well as misinformation.**

weaken science and work towards solutions to minimise their impacts.

At Ripeta, we are also concerned with the quality and trustworthiness of research, the ways in which they are understood, and how to more quickly check these trust indicators. There have been more conversations on trust, legitimacy, and authorship in science because COVID-19 has highlighted that it is not just scientists engaging in the research – it's *anyone* interested in scientific research.

Do you have any general suggestions for how to fortify trust and maintain ethics within open scientific publishing?
Through a combination between awareness and action. We need to not only *understand* but also be able to *explain* how this messy thing of science works, so that we can know when it doesn't work. Going a little deeper, we need to talk more about the governance of open scientific publishing because, frankly, it will not govern itself.

So, how do you fortify trust? Make science more transparent, make it easier, and have it checked. Science is hard! Yet places exist to not only check the work but automate the checks. There are ways to check within and across scientific publications to find (potentially harmful) trends and how to address those less-than best practices. In building Ripeta, we continue to focus on issues of science and trust using our automated checks for research, reproducibility, and professionalism.

What are your hopes for Ripeta in the coming months and years?
The aim is to make better science easier. We want to expand the adoption of the *ripetaReview* and *ripetaReport* to quickly automate quality checks, and we also want to educate people on the internal checks they can do to assess trust.

We plan to integrate products like *ripetaReview*, software built to check individual papers on their quality indicators, into the researcher workflow to make their job of maintaining

integrity-driven research easier. We created *ripetaReport* to help inform stakeholders on how they can build upon their current and projected guidelines for open science. In the coming months, we will also expand the *ripetaReport*, to offer more tailored views into the quality of thousands of research papers and make recommendations for key stakeholders in science (funders, institutions, and publishers).

The *ripetaReport* dashboard will also be expanded, giving stakeholders a bird's eye view on the current status of their portfolio. It will become an interactive space to see how journals, publishers, and institutions are doing in their journey towards higher quality science.

Beyond our products, we will continue to offer information and resources for researchers, institutions, publishers, and funders on best practices in reporting research, practicing open science, and fortifying trust in science. Between the research our team conducts, as well as our general-audience-friendly blog

posts, we hope to be a reliable and consistent source of information.

Ultimately, I created this company to make better science easier for researchers, and with that goal, we have brought the mission of building trust in science to the forefront. We are excited to see what the future brings, and we want to be there to help bring trustworthy information to others who are as excited, skeptical, critical, and hopeful as we are.

W: https://ripeta.com/
@ripetaReview
@ripeta