



STANDARDISED DATA ON INITIATIVES (STARDIT)

Additional Resources for the WikiCite 2020 Virtual conference session by Jack Nunn and Thomas Shafee

About this document

This document contains links to additional resources relevant to the pre-recorded session by Jack Nunn and Thomas Shafee for the WikiCite 2020 Virtual conference.

More information about STARDIT can be found here: ScienceForAll.World/STARDIT

Contents

About this document	1
Contents	1
Additional resources	2
Download the video of the presentation.....	2
Presentation.....	2
Get involved	2
Example STARDIT reports	2
Original PDF STARDIT reports as supplementary files	2
Free text STARDIT reports hosted on Wikispore	2
Structured data hosted on Wikidata.....	2
Video Transcript.....	3
Acknowledgements.....	5

Contact Jack.Nunn@ScienceForAll.World or Twitter: [@JackNunn](https://twitter.com/JackNunn). Learn more at ScienceForAll.World

This document was created by Jack Nunn ([Orcid.org/0000-0003-0316-3259](https://orcid.org/0000-0003-0316-3259)) for [Science for All](https://ScienceForAll) on 26th October 2020. This document is licensed under [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/).

Additional resources

Download the video of the presentation

You can download the video of the presentation [here](#).

Presentation

[Download](#) this short presentation 'Making STARDIT open, machine readable, interoperable and multilingual' by Thomas Shafee.

Get involved

[Join our online community here](#).

Example STARDIT reports

Original PDF STARDIT reports as supplementary files

- doi.org/10.21203/rs.3.rs-54058/v1 (supp file 5)
- doi.org/10.21203/rs.3.rs-62242/v1 (supp file 2)

Free text STARDIT reports hosted on Wikispore

- [STARDIT/Q98539361](https://www.wikispore.com/wiki/STARDIT/Q98539361)
- [STARDIT/Q100403236](https://www.wikispore.com/wiki/STARDIT/Q100403236)

Structured data hosted on Wikidata

- [Q98539361](https://www.wikidata.org/wiki/Q98539361)
- [Q100403236](https://www.wikidata.org/wiki/Q100403236)

Contact Jack.Nunn@ScienceForAll.World or Twitter: [@JackNunn](https://twitter.com/JackNunn). Learn more at ScienceForAll.World

This document was created by Jack Nunn ([Orcid.org/0000-0003-0316-3254](https://orcid.org/0000-0003-0316-3254)) for [Science For All](https://ScienceForAll) on 26th October 2020. This document is licensed under [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/).

Video Transcript

So Hello, my name's Jack Nunn. I'm director of the charity Science For All. I'm also the Strategy Liaison for the Wiki Journals. And I'm here today to talk about a project I've been working on with Thomas and many other people called [Standardised Data on Initiatives \(STARDIT\)](#), which we're helping develop with the wiki journals. Being British, I drink tea, and I sometimes sit on my sofa and eat a biscuit. And it struck me that I can't sometimes know for sure that my tea has been picked by people who are paid fairly, that the rain forest hasn't been cut down to grow the cocoa or the tea leaves, or in fact, the timber for the sofa that I'm sat on, or the clothes I'm wearing haven't been stitched by slaves. I don't know this because the data isn't available for me to make an informed choice. And I think the long term aim for STARDIT is that we wouldn't just include research that we'd be including all initiatives, whether it's government policy, political manifestos, data on manufacturing processes, it's public domain data that describes who was involved, who did what, and who funded it. And what were the outcomes, what and what were the methods used to involve people. So standardized data on initiatives. It's something I've been thinking about for many years, and something that I'll talk about in more detail about where it came from. For a bit of background, I've been working for over a decade, to involve the public and patients in improving health services, health care and research. And certainly in what's called public and patient involvement in some areas. We're asking the question, what does it mean when we involve people in research? Well, who was involved? And which tasks are they doing? Data sharing on this is quite variable and quite poor, according to a scoping review that I did recently about public involvement in genomics research out of this work, and, you know, I discovered there were a few frameworks already where people could report who was involved in research and how all these frameworks are very specific. So all of them are specific to health research, or specific to systematic reviews. And they're also English language only. So the idea of STARDIT was to try and create a way that people could report data about initiatives such as research, using machine readable link data. And as director of Science For All, I'm also involved in some environmental projects where we're doing environmental research, and we're involving citizen scientists, or involving the public in going out and gathering environmental DNA samples or helping us write protocols. And so I was sat here going, is there a way that I can report how people were involved in a standardized way across disciplines, so whether it's health research, Environmental Research, or public policy, government policy, etc, these were, these were the foundation thoughts for STARDIT. I then got very involved with the wiki journals. And my good friend Thomas here told me about wiki data, which was something I wasn't particularly aware of even existed. And the more I learned about it, the more I realized a lot of things I was hoping to do with STARDIT were already possible with wiki data. So you can share information about things in a machine readable way. It's linked data. And it's also and most importantly, it can be edited by anybody. And, and perhaps just as importantly, it's automatically translated into other languages. So it isn't just English language specific. So what we've got already here is a data backbone of data about initiatives, people, research, what have you that already exists. And so we decided to work in that, in STARDIT standardized data on initiatives STARDIT for short.

Because the governance structure, the legal structure is already managed by the Wikimedia Foundation, which is obviously a hugely trusted organization, and the processes for creating data are already very

Contact Jack.Nunn@ScienceForAll.World or Twitter: [@JackNunn](https://twitter.com/JackNunn). Learn more at ScienceForAll.World

This document was created by Jack Nunn ([Orcid.org/0000-0003-0316-3254](https://orcid.org/0000-0003-0316-3254)) for [Science For All](https://ScienceForAll.org) on 26th October 2020. This document is licensed under [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/).

well established. So the idea of STARDIT is to create a kind of a ring around some data fields. So who was involved? What initiatives? How was it funded? And a STARDIT report can be the minimum contribution is just what's the minimum information we have, you know, conflicting, competing interest who is involved who wrote the report, and then the largest STARDIT report has more data, and we're currently collaboratively designing what kind of data that should be. And I can say more about that. So let's take the example of a Wikipedia page that gets submitted for peer review, you know, that might have had hundreds of editors over the past 10 years. And what STARDIT allows is to report who those people were, maybe what their tasks were to a resolution that everyone agrees is useful and helpful. And I think the other thing to add to what Thomas said that's very important about the wiki journals is that it's free. And it's open access. And I think the importance of that really, I can't stress enough, everyone has probably had the experience of coming up to a paywall or finding peer reviewed information that they can't access because they can't afford it. This is fundamentally against the principle of how science should work. Science just means knowledge. And if people cannot access knowledge, then you're limiting their ability to make all kinds of decisions and informed judgments. And this is a huge issue in the developing world in particular, the point of open access is that it's free for everyone. And that's great. However, anyone who's tried to publish in an open access journal might have been faced with something called an article processing fee, which can be as much as 2000 US dollars, which is a lot of money. And I can tell you it is more than most people would afford from their own pockets to publish something. Now, as a caveat, publishers do an important job, quality control is very important and people need to be paid for their time that needs to be valued. Of course, wiki journals are different. At the moment, we rely almost entirely on volunteers to do the whole process. And that means that it's free for people to publish. And I think that's really important. There are other free to publish open access journals out there. And the importance of that as it gives us independence as well. It's rare, but it's not impossible that there might be an open access journal that happens to be owned by, you know, a big publisher. And this is a multi billion dollar annual industry, by the way, publishing, academic publishing more than film and music, a huge industry that makes a lot of people a lot of money. And there are people who have certain interests, financial or other that may wish to have certain issues not explored in journals that they own. And there are examples of this happening online. I think the importance of something like Wikipedia is that it is demonstratively independent, it's funded by donations, the entire process for deciding what's in or out is completely transparent. And anyone can be involved. One of the issues at the moment, though, is that trust in who's involved, you know, and there are many examples of people in government editing their own entries about themselves or other political leaders, or what have you, not saying that's good or bad. But wouldn't it be nice to have a bit more transparency about who those people are, and bringing things like, you know, the ORCID numbers, which is an identification number that a lot of academics used to associate with themselves? I think the first thing to say about STARDIT is it's probably everything that appears in the reviewed paper is in terms of you know who the authors were, etc. The advantage as well as it can be updated over time. Unlike a journal article, which you know, it's published, and then it's out there. So, one thing that you can do with the STARDIT report is you could say, we're going to do this research. So STARDIT provides a protocol. And you can say, we're going to share data about this research in this way. So let's say you're working on the Hadron Collider, and you say, we're going to share all our data, according to the fair principles, findable, double, accessible, interoperable, usable, I think. So we're going to share our data contract

Contact Jack.Nunn@ScienceForAll.World or Twitter: [@JackNunn](https://twitter.com/JackNunn). Learn more at ScienceForAll.World

This document was created by Jack Nunn ([Orcid.org/0000-0003-0316-3259](https://orcid.org/0000-0003-0316-3259)) for [Science for All](https://ScienceForAll) on 26th October 2020. This document is licensed under [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/).

principles, and we're going to be funded by the government. And then you can have another update report. And here's all the data. It's all in the public domain. And what STARDIT helps do, is link to public domain data and link to that it's not necessarily hosting all data from Hubble or the Hadron Collider, it's just saying, well, this is where it is, and this is how it's structured. This is who owns it, and this is how it's governed. And then more importantly, certainly for organizations like Cochrane or the Campbell Collaboration, which do systematic reviews into health or social issues, knowledge translation, how has this knowledge helped or informed or improved lives or ecosystems? Then the STARDIT report can go Okay, this paper is 10 years old, or this research was done 10 years ago. What was the outcome? Or what are the impacts? The potential applications of STARDIT are limited only by our imagination. And it can be anything from a Community Arts Initiative, to an educational intervention to public policy or foreign policy. You know, there's a lot of things humans choose to do because they expect certain outcomes. And we follow those outcomes and sometimes call that research. But it's very important to bring the scientific method where possible to questions like educational interventions, you know, we are trying to achieve. For example, if humanity wants to lower birth rates, slow down climate change, and improve air pollution, etc. We know one of the best things we can do is educate girls, specifically girls, because that reduces infant mortality, amongst many other benefits of course, and thus also helps populations have less pressure. To have more children, etc. Um, so suddenly education is this hugely, it's always been this centrally important thing in all cultures. But wouldn't it be nice to have a bit more evidence informed methodology for it, you know, for a government to come in and say, well, let's retrain ballerinas in cyber security was a recent suggestion from the UK Government. It didn't go down very well, because, frankly, it was ideologically based. It wasn't evidence informed. And there was very little evidence that it was a sensible suggestion, and it was rightly ridiculed. I'm speaking from my personal opinion. Now, by the way, just to be clear, I'm not speaking on behalf of any organizations. Rightly ridiculed. But wouldn't it be nice if there was educational policy, where they could say, we've got this idea? We're going to let students lead their own learning, we're going to have teachers learn with students were going to do this or that. And here's the evidence that we've used to inform this policy decision. And then 10 years ago, wow, guess what your evidence informed educational policy achieved the outcomes you said it was? How often does that happen in government or in manifestos? Not enough and we're talking about the most important thing humanity can do, which is education, which is why Science For All is an education charity.

Acknowledgements

This document was prepared by Jack Nunn for Science for All. Thank you to Thomas Shafee for his time in supporting this project, Adam Weiss for editing the video and Nyssa Braid for editing the transcript.

Thank you to everyone who has supported STARDIT so far!

Contact Jack.Nunn@ScienceForAll.World or Twitter: [@JackNunn](https://twitter.com/JackNunn). Learn more at ScienceForAll.World

This document was created by Jack Nunn ([Orcid.org/0000-0003-0316-3254](https://orcid.org/0000-0003-0316-3254)) for [Science for All](https://ScienceForAll.org) on 26th October 2020. This document is licensed under [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/).