

The miniReproducibility project

The idea

A valuable but hidden set of data: research trainees spend months to years confirming previously published data. These experiments are not considered publishable. Other trainees in the same field “reinvent the wheel” due to lack of publicly available information.

A simple idea: Imagine the savings in time and resources, reduction in retractions and the impact of each experiment if we knew the following: 10 labs repeated it, 7 found a phenotype and 3 did not. What if a single click could highlight differences in methodology between each experiment?

Our vision is to create a gamified portal for researchers (initially focused on students and postdocs) to publish replicated data, compare methodology and provide reproducibility statistics for individual experiments. A user ladder with increasing prestige (pioneer-->scholar--->catalyst), reviewing and commenting rights, and goodies donated by biotech (antibodies, kits etc.), coupled with the real benefit of having publications and peer-reviewing experience will attract trainees and encourage return usage.

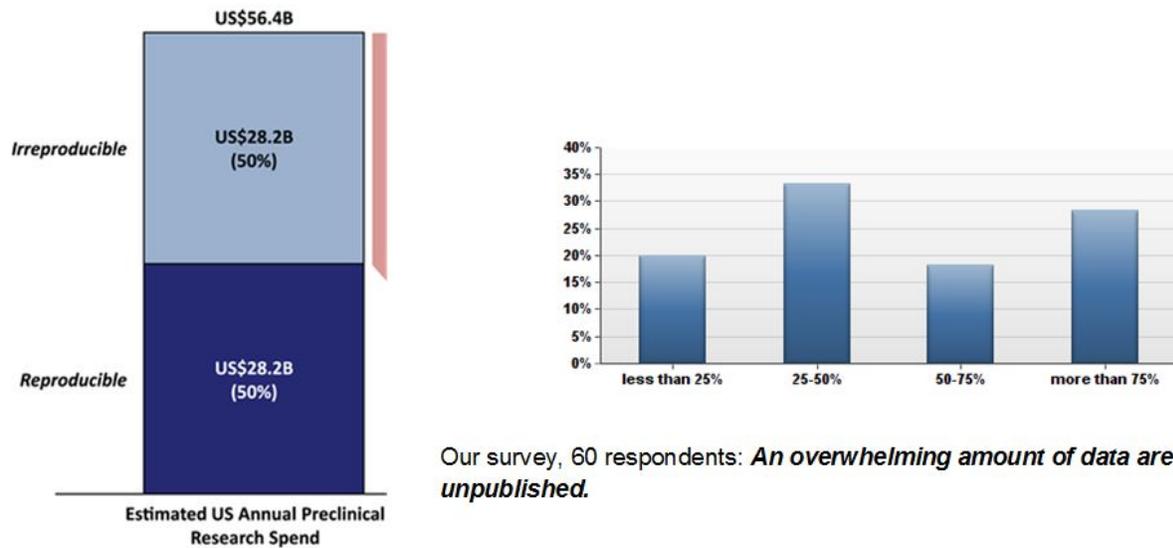
The act of comparing several submissions and associated methodologies itself has received overwhelming support from a large network of fellow trainees and mentors (~100). We request a technical partnership to allow single experiment submission, review by several “players” and assessment of reproducibility.

Background

Open science is believed to be the answer to the reproducibility crises. Current initiatives focus on repositories for data and code and the now famous "Reproducibility project". However, data without peer review in a repository doesn't give the researcher any recognition and nor does it lead to impact. Similarly, how many iterations of the reproducibility project will it take to confirm findings in different fields years after they are published?

Consider the dichotomy highlighted in Figure 1. On one hand, we rue the billions lost to irreproducible research; on the other “reproduced data” is not considered publishable. We conducted a survey and found that a large proportion of data generated in laboratories (including reproduced data) is never published. The miniReproducibility Project would both

make the years of reproduced work impactful and belated verifications of published studies unnecessary.



Modified from Freedman et al (2015) The Economics of Reproducibility in Preclinical Research. PLoS Biol 13(6): e1002165. doi:10.1371/journal.pbio.1002165

Figure 1: Reproduced data is not publishable and published data is not reproducible

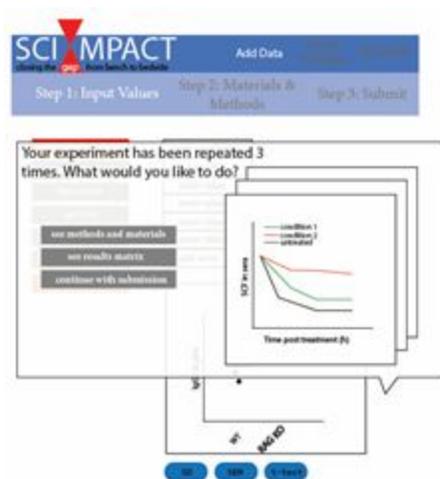
Execution

How could one execute the idea described above: after talking to many experts, it seemed overwhelming to rely on technology to identify replicates, align their results and compare methodology. So we envisioned a community of contributors who could curate machine or human flagged data. We focused on trainees for these reasons:

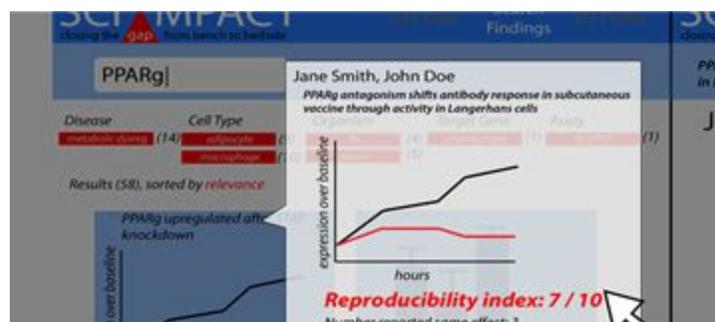
1. One of the main values of our startup is recognition for trainees and their effort
2. Trainees are likely to be conversant with social media and we want to use that to drive the gamification aspect of the portal.
3. Trainees need outlets for outputs currently considered unpublishable
4. The return on investment –a publication or recognition for peer-reviewing in return for few minutes spent “gaming” on the website (akin to tweeting/facebook) is more valuable to trainees than established researchers.

We consider our model that of a social business: we aim to recognize trainees for their efforts and improve the reproducibility and impact of experimentation. We have considered having an independent submission platform or partnering with repositories such as Figshare to build a linked reproducibility site.

Mockups:



Computer aided user submission index



Search by gene name showing reproducibility

User ladder and gamification privileges to incentivize engagement

User level-No' "reproducibility reviews" conducted	Title	Free Submissions of data	Commenting privileges	Reviewing privileges	Enhanced privileges
1	Beginner	0	2 comments	1 more	None
2	Entrepreneur	1	5 comments	3	
3-5	Innovator	2	10 comments	5	We will publicize their contribution to science on social media
5-10	Inventor	3	Unlimited	Unlimited reviews	We will provide linked-in endorsements
10-20	Catalyst	5		Can submit mini-reviews on data they have reviewed	
>20	Leader	Unlimited			

Market research

1. Conversations with >100 colleagues to discuss the vision and impact of miniReproducibility project.

2. Poster presentation at MIT Sloan Bioinnovations 2015
3. Mentoring from Boston angel investors
4. Survey: online survey to which 70 scientists largely between the ages of 22-45 responded. An overwhelming 68% said that they published less than 1 paper a year, a number often cited as the benchmark for getting prestigious academic and industry jobs. The survey highlighted an awareness that change in scientific publishing is required and a willingness to try alternative methods.

Business Model/ Financials

We consider ourselves a social business. We would like to charge a modest fee to make the project self sustaining and to provide funding for expanding its scope (see “Long-term plans”). The US has ~100,000 researchers. Our aim is to capture the value in the data that are not being submitted to traditional journals. Our conservative estimate is each researcher performs at an average of 10 experiments/year to confirm previous findings for which our platform can provide an outlet and citations. At a nominal fee of 10 USD/submission, our market size is 10,000,000 USD/year. We expect to achieve 20% US market penetrance in 2 years. Our business is easily scalable to all other countries.

We are intensively talking to mentors, collaborators and potential investors. However the most immediate plan for fundraising is through crowdfunding. We believe the idea is fun, has a social conscience and would appeal to trainees as well as the informed layperson worrying about the state of science.

Timeline and costs

Phase I (6 months, \$30,000)

- *Outsourced development of alpha version: computer guided user submission and citation platform*
- *Data crowdsourcing: “We don’t want your money. We want your data!” campaign. Goals: populate the database and spread the message*
- *Development of beta version: building and testing the platform based on crowdsourced REAL data*

Phase II (6 months, \$100,000)

- *Launch of beta version: pay per submission or yearly subscription*
- *Small complement of full time staff-editor, systems managers*

Phase III: Full launch

Long-term plans

We have big ideas but are starting small. Our vision is to seed a fun online researcher community with the miniReproducibility project. Once engagement with researchers reaches critical mass, we hope to expand into novel (previously undiscovered) findings, continuing in a single-figure format. Currently, researchers perform experiments in a linear fashion. Sometimes unbeknownst to us, the researcher next door is doing something very similar. Our semi-automated gamified model will allow linking of singly submitted experiments to identify novel biology. It is a **novel vision of science-parallel and crowdsourced**. It allows harnessing of different talents: some researchers are very good at conducting benchwork, others at generating hypotheses and others at finding links between apparently unconnected findings.

The team

Cofounders



Girija Goyal

Postdoctoral researcher at the Wyss Institute of Biologically Inspired Engineering

Years on the bench: 12

Business experience: Intern at Mass Medical Angels

“Question everything”

goyalgirija@gmail.com (*preferred*) , 617-953-2360



James Akin

Postdoctoral researcher at Novartis

Years on the bench: 12

Business experience: 2 years managing \$2 million in SBIR funding supporting critical technology for DoD; pioneer participant of Harvard Business School's intensive 9 week introduction to business-HBX CORE

"A good plan executed now is better than a perfect plan next week." -George S. Patton

jffour@gmail.com



Market analyst and Consultant-of-many-talents

Chen-Yuan “Charlie” Yang

PhD candidate at BU School of Medicine, Department of Anatomy and Neurobiology

Years on the bench: 9

Business/Startup experience: Developing a mobile app for diet tracking

“Failure is success in progress” – Albert Einstein



Consultant, crowdfunding lead

Karrie Wong

Research Investigator at Novartis

Years on the bench: 8

"I have always imagined that Paradise will be a kind of library"