Distributed Scientific

Overview

Who has access to the means of producing science? When we ask 'who are, and where are scientists?', we might first think that scientists are 'only those anointed with advanced degrees', and that we only find them 'in a lab' or 'at a university'. The laboratory (or large scale scientific computing resource), is where *real experimental science* happens (ignoring for the moment field sciences like ecology and the like). This has historically been true for a number of socio/cultural, as well as practical reasons. One reason is the fact that scientific equipment is expensive, and oftentimes the material for performing scientific experiments is costly and/or inaccessible. Even if this were not the case, and access to equipment and materials was both easy and low cost, we would still face the problem of not knowing exactly what to do with it to advance understanding. More troubling, *most of our newly trained professional scientists do not engage in basic research as a long term career*. Recently, an alternative view of scholarship has been proposed: *fractional scholarship* [1]. In this view, both the interested amateur and formally trained (though not currently practicing) scientist can be seen as resources to be employed fractionally towards a basic research goal. This is an intriguing idea, however not present in the formulation [1] is a view towards formally scaffolding and integrating distributed fractional efforts. Rather than fractional scholarship seen as an individual researcher applying for, and working on a grant solo, we instead see the most intriguing application that of *a large community of researchers contributing fractionally to a larger project*. If we are to enact such a vision, which tools and technologies will be necessary? Which communities will it be essential to inspire and engage with?

We have already crossed a technological threshold whereby it is now common place to engage in face to face real time communication with anyone, almost anywhere in the world. This has made possible collaboration and discussion among scientists with high speed internet accessible at the click of a button. Is it possible to treat scientific experimentation in a similar manner? We can, as the development of remote, robotically controlled 'Science/Experiment as a Service' facilities has emerged in recent years [2, 3]. Utilizing this technology, users from a web interface can control and specify numerous scientific experiments and protocols that are then run remotely in these highly automated scientific 'factories'. This represents a profound shift in the accessibly of the subset of experimental science which is feasible to automate, and this shift is poised in the coming years to revolutionize the practice of science. The question remains, who will benefit most from this transition? We are proposing to develop a curriculum that can be delivered at low cost, to anyone, anywhere in the world that is willing to put in the effort to learn. The basis of this course is the practice of *actual science*, as students perform and replicate key experiments in a selected field, *while simultaneously proposing and analyzing experiments of their own design*.

In a wider sense than just a class, a broader vision for this tooling is as an enabling technology for cohesive, community driven fractional scholarship. A class (and eventually full curriculum) acts as a means of leveling the playing field for interested potential scientists, where as an end goal institutionally scaffolded projects become the 'graduate' work of students, who then contribute when and where feasible, for an indefinite time period (or until a project is completed). In this view, if substantive scientific output is achieved, as in a traditional higher academic setting, the realities of the structure of IP and the market must be addressed. Here we offer a starkly different view than the norm, informed in particular from analysis [4] of the primary field we intend to impact the most: biotechnology. The problem is perhaps most clearly stated in the article 'Tragedy of the Anti-Commons' [5]. To quote the abstract:

"The anticommons thesis is simple: when too many people own pieces of one thing, nobody can use it. Usually, private ownership creates wealth. But too much ownership has the opposite effect – it leads to wasteful underuse."

In pharmaceutical development the problem of the anti-commons is particularly acute: parties must often assemble numerous patents, wherein any one of that number of patents might become blocking due to price or contract negotiations. For these reasons, an alternative academic institution would be remiss to not try to restructure the relationship between research and IP, and between IP and those who contributed toward its creation. To address this we propose two principle changes, and offer one forward looking point:

1. IP is granted Free license for all non commercial use.
2. IP is licensed commercially *non exclusively*, and for a royalty commensurate with firm market capitalization.
3. IP management is itself a research project (which we undertake), and we believe technologies like MPC [6] and blockchain can be brought to bear on the problem which can in the future significantly lower costs and barriers to innovation, as well as fairly assign shares based on contribution.

There are other key advantages to this remote, distributed scientific paradigm, over and above accessibility and the 'wealth of networks' [7] which we will briefly outline.

There is currently a reproducibility crisis in scientific research: numerous studies show that many experiments simply are *not reproducible*. It is difficult to overstate how large of an impediment this is to the progress of science. Given the volume of publications, and rates of replication, we are wading into an ever growing sea of noise and misdirection (take the current Alzheimer's debacle as perhaps the most visible example). The remote, automated science facility must be accessed via a computer interface, and experiments must be specified exactly and in code, which by default makes them far more reproducible than the current standard of ambiguous description in the methods section of published papers.

The second point is related to the first, and also related to the question of who should be qualified to do, or allowed to do scientific research. Research shows [8] that centralized scientific communities produce less replicable scientific research when compared to distributed communities. By opening up access to the means of scientific production, we can re-structure the networks that produce science, making them both more diverse and more resilient to error or manipulation. This is a synergistic effect of openness that has always been the hallmark of healthy and functioning scientific communities.

The next steps are to:

1. Refine, and get community feedback on this proposal
2. If sufficient interest is gauged, create a concrete outline/timeline and cost proposal for a proposed class/community scaffolding structure
3. Consult IP lawyers to determine feasibility of proposed IP management structures

References

[1] Fractional Scholarship PDF

[2] Emerald Cloud Labs

[3] Strateos Cloud Lab

[4] BioBazzar

[5] The Tragedy of the Anti-Commons

[6] Multip Party Secure Communication

[7] The Wealth of Networks

[8] []https://elifesciences.org/articles/43094

http://calteches.library.caltech.edu/51/2/CargoCult.htm :
"But this long history of learning how to not fool ourselves—of having utter scientific integrity—is, I'm sorry to say, something that we haven't specifically included in any particular course that I know of. We just hope you've caught on by osmosis."