# Continuous Research

Setting Collaboration as the Default

### 

### Introduction

Science today is built almost entirely on a culture of competition. Formal integration of research from different groups working on the same topic happens very rarely. When it does happen, for example when results/algorithms are explicitly compared, these results are often highly biased(ref?) or incomplete.

Continuous Research is the concept that all scientific outputs, every new dataset, technique, and hypothesis, can be can be integrated such a way that they are re-evaluated against all related research. In developer-speak we could say that we’re “bringing Continuous Delivery to research”.

### What is Continuous Delivery?

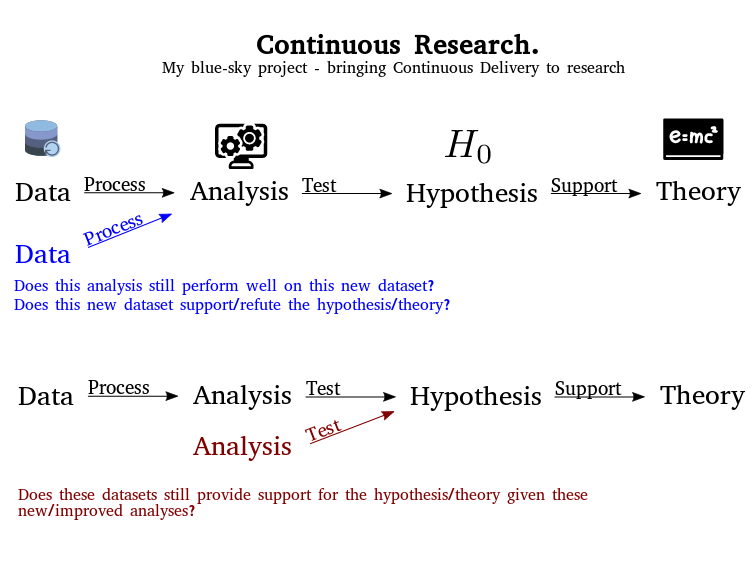
The requirement to maintain large, distributed, and complex projects, along with the high availability of low-cost computing has driven the widespread adoption of “Continuous Delivery” in software development. The mechanics can be complex, but the principle is simple: automated tests are frequently run against the software so that even when a brand new feature is added, you can be sure that the software is still performing correctly and (at least close to) functional enough to be released immediately.

### How is research done now?

You collect a dataset, you do some analysis to test a hypothesis, and then try to integrate that hypothesis with current theories. Great, there’s your paper, submit it to Journal of $your\_field and move onto the next project. But what happens when someone else collects a dataset that could also be used to test that hypothesis? Or a computational specialist writes a better version of your analysis. Best case scenario they might use your code/data in their papers, otherwise they might just mention or cite you. But the work will never be truly integrated, and many papers might be linked to tens/hundreds of other pieces of research, meaning that using your code/data with their work will not always be possible.

### How does that change under Continuous Research?

Under Continuous Research we re-evaluate *all* hypotheses when any new data, analysis, or pipeline becomes available. In such a way, we obtain a live view of how much support there is for a particular theory, and under what conditions such a theory really holds.



### Use Cases

Adding a “full” study - *“The Virtual Paper”*

Adding a new dataset - *“At the Coalface”*

Adding a new analytical pipeline / node - *“Have you tried...”*

Adding a new visualisation - *“No Sunglasses Required”*

Adding a new hypothesis / Editing a theory - “*The Comments Section”*

**Prototype/Visual Demonstration**

### Ethos

While technology continues to improve, the methods underlying much of research haven’t quite entered the 21st Century. The pace of change is now increasing which could be a great thing, but I am concerned that we are still baking in many out of date values into our new practices. “Open Science” is not really open unless it is truly participatory. Small steps are great, and we encourage everyone to get started in whatever way is most accessible to them. Nonetheless simply putting the code you used for a piece of research isn’t fulfilling the potential of what Open could be. But if your code is running live somewhere and anyone can come along and play with the parameters to see what effect that would have had on the outcome of your experiments - that’s getting a bit closer. If there is a low-barrier to someone recreating your visualisations in a way that’s accessible to them even better. In a technical sense open for me additionally nesseciates *accessibility*, *extensibility* and *interoperability*. And this project is about showing how powerful research can be when it meets these criteria.