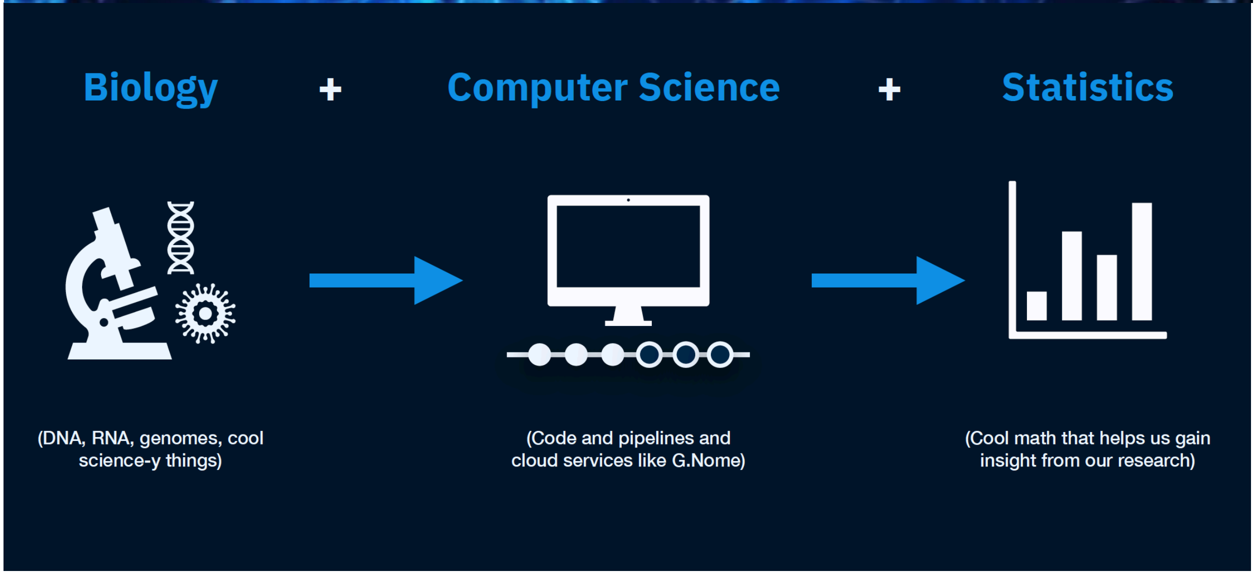


Biology + Computer Science + Statistics

(DNA, RNA, genomes, cool science-y things)

(Code and pipelines and cloud services like G.Nome)

(Cool math that helps us gain insight from our research)

G.Nome Workbench

Area: IBM Cloud and Cognitive Software
Time: 5 weeks (handed off to early career professionals for further development)
Completed this incubator project as a part of Patterns internship at IBM

Disclaimer – ***As this project is still in development, the details I can provide around the work I completed is very limited at the moment.***

Main Objective

The goal for this project was to structure the information architecture and interface of running bioinformatics pipelines on the cloud. The team's focus was more specific to locating errors in the workflow and having version control access to be able to quickly compare and contrast results or scripts.

Research

To understand our users more and the pain points they face we:

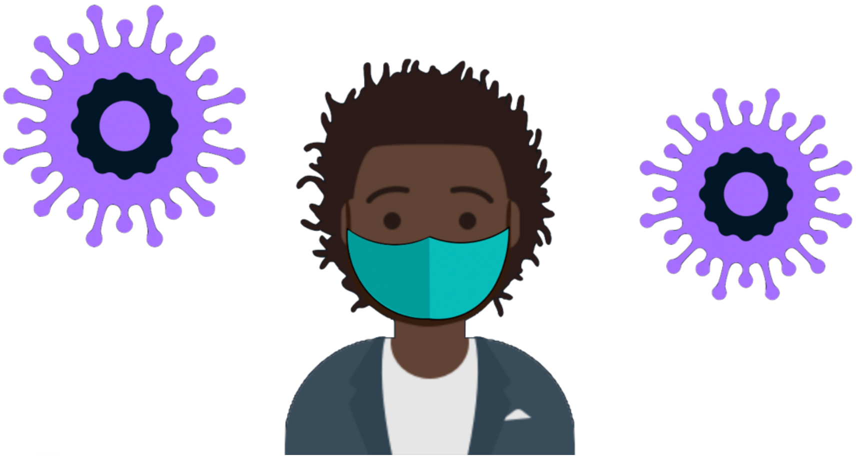
- Looked into competitors
- Analyzed ethnographic interviews
- Interviewed sponsor users

And from this we were able to learn more about alternatives the users turn to and how they feel when working with the current platform.

Persona

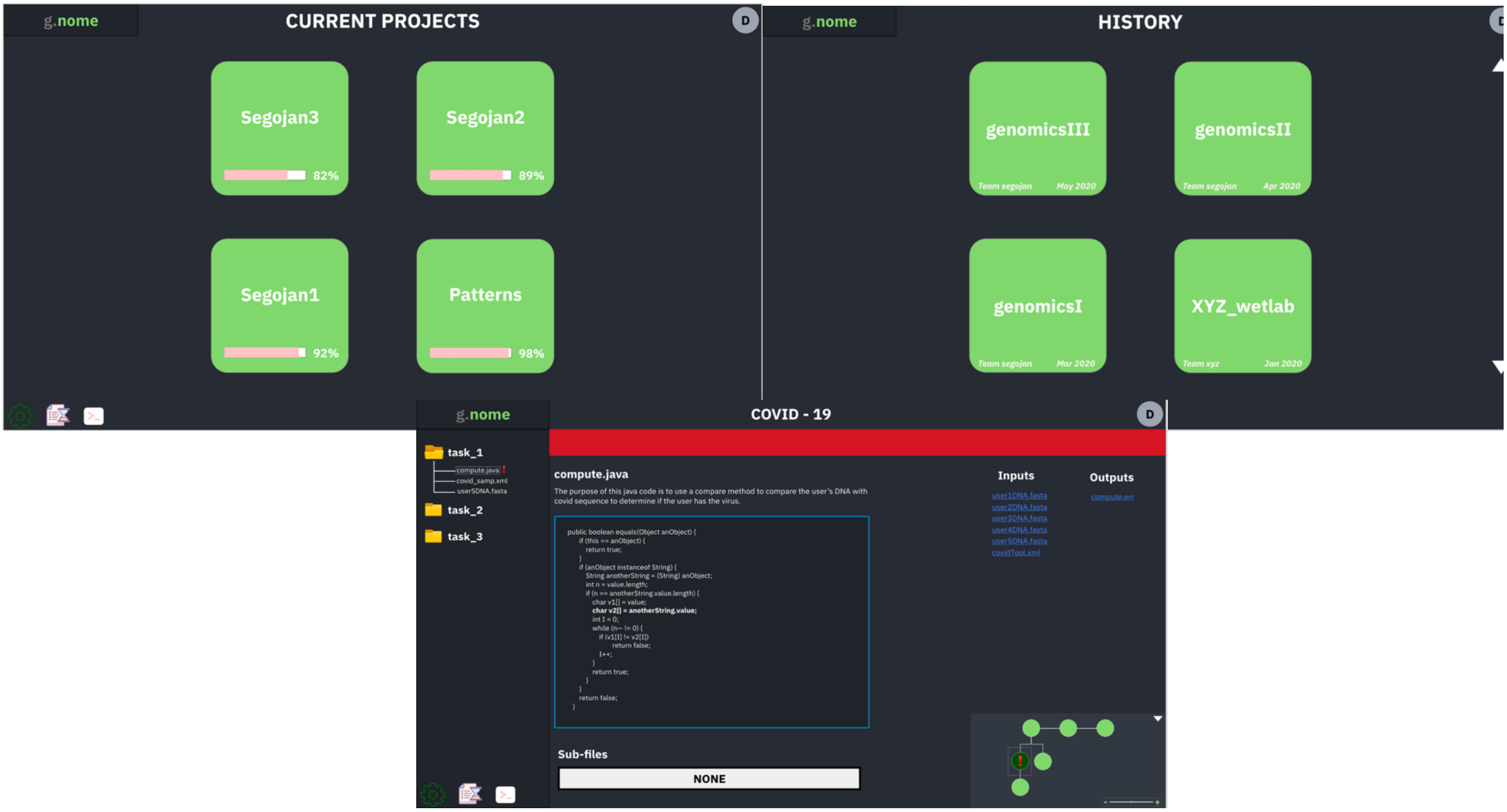
So we created our user Dan, who's a senior bioinformatician that works at XYZ research institute for almost 23 years now! His institute has been currently working on trying to combat COVID-19 – specifically his current project is working on using genomic data and bioinformatics processes to predict how the virus may involve.

Every time Dan runs into a nasty error, he faces the misfortune of having to sift through dozens of log files which is extremely annoying and overwhelming for him.



What are the main features?

1. Tree directory structure – this was a must. All of our users are familiar with this structure so why not include it? It provides users with an easy access to navigate between different folders – tools and scripts
2. Alert system – in the current platform users have no clear indication of where errors are located which results in them sifting through a lot of unnecessary log files. With these alerts, it provides the users a quick and clear indication of where to navigate to – cutting down the error locating time by a lot
3. Mini map – this is also another way to view an overview of the pipelines and tools you are working with. It is another way to view error locations in a graphical view



compute.java

The purpose of this java code is to use a compare method to compare the user's DNA with covid sequence to determine if the user has the virus.

```
public boolean equals(Object anObject) {
    if (this == anObject) {
      return true;
    }
    if (anObject instanceof String) {
      String anotherString = (String) anObject;
      int n = value.length;
      if (n == anotherString.value.length) {
        char v1[] = value;
        char v2[] = anotherString.value;
        int I = 0;
        while (n— != 0) {
          if (v1[I] != v2[I])
              return false;
          I++;
        }
        return true;
      }
    }
    return false;
  }
```