Project Report

**Master of Technology in *Intelligent Systems***

Module: Intelligent Process Automation

**Intelligent Recruitment Helper**

*Submitted by:*

Soorej Mohanadas Ganga (A0195397J)

Rahul Jalan (A0195299H)

03 May 2020

Table of Contents

[Executive Summary 3](#_Toc39441705)

[Problem Description 5](#_Toc39441706)

[Project Objective 7](#_Toc39441707)

[Market Research 8](#_Toc39441708)

[Recruiting software trends and features 9](#_Toc39441709)

[Conclusion from Market Research 12](#_Toc39441710)

[Product Plan 13](#_Toc39441711)

[Launch Plan 13](#_Toc39441712)

[Development Plan 15](#_Toc39441713)

[Business Case / Financials 16](#_Toc39441714)

[Revenue Model 16](#_Toc39441715)

[Cost Model 17](#_Toc39441716)

[Profitability 18](#_Toc39441717)

[Solution – Intelligent Interview Helper 20](#_Toc39441718)

[System Architecture / Model 20](#_Toc39441719)

[Database Design 23](#_Toc39441720)

[IPA and RPA 25](#_Toc39441721)

[Limitations and Future Enhancements 26](#_Toc39441722)

[Conclusion 27](#_Toc39441723)

[Appendices 28](#_Toc39441724)

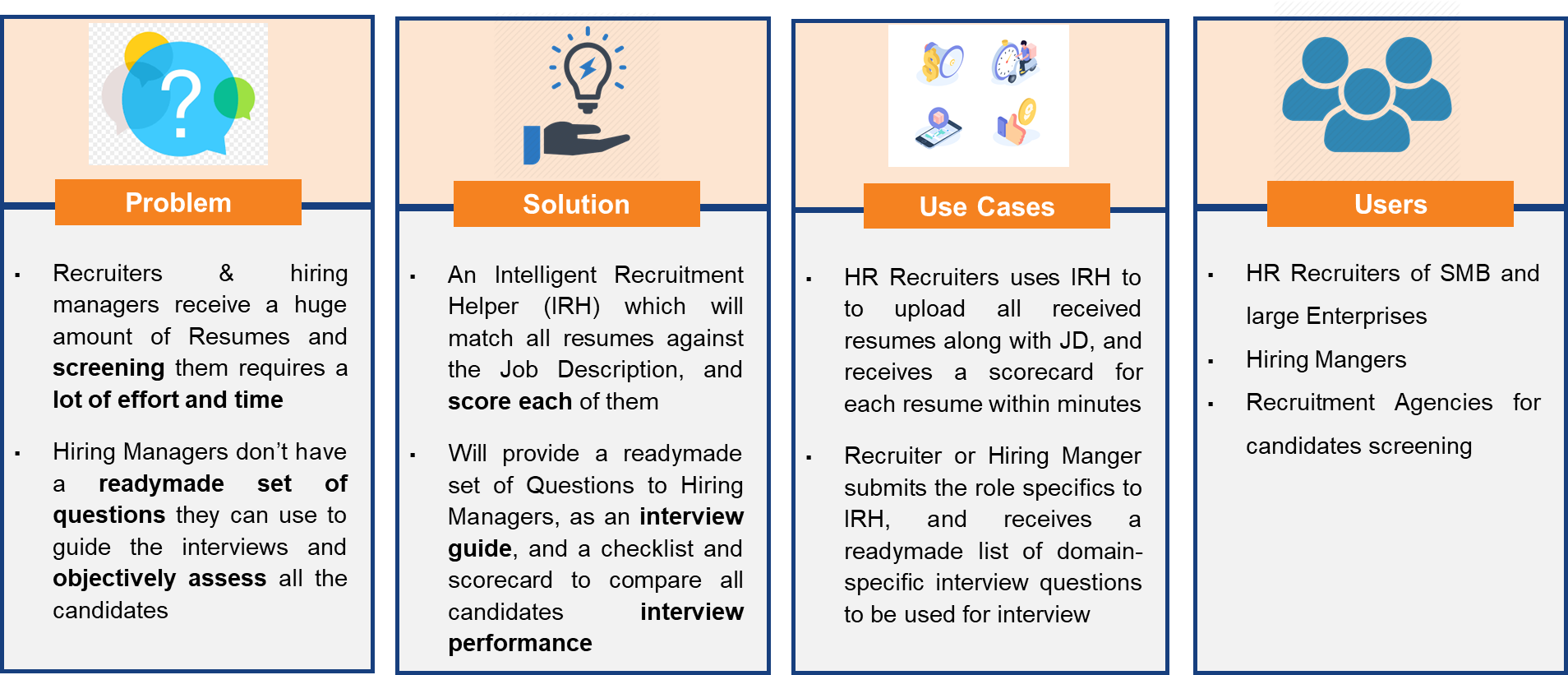
[APPENDIX A: Installation and User Guide 28](#_Toc39441725)

[Installation 28](#_Toc39441726)

[User Guide 30](#_Toc39441727)

[APPENDIX B: Individual Reports 38](#_Toc39441728)

# Executive Summary



**Fig 1** – Intelligent Recruitment Helper

Through this project we have tried to build a working AI system to address some of very common problems faced by Recruiters and Hiring Manager across all industries. Recruiters are challenged with the sheer number of resumes they receive every day and frequently face the problem of how to shortlist the right candidates. More often than not they send all the resumes to hiring managers which simply shifts the problem elsewhere. Second issue is that the hiring mangers or interviewing team does not have a structured list of questions to ask the candidates, nor do they have a proper methodology of accessing one candidate over another.

Our **Intelligent Recruitment Helper** solves these two problems by providing a scoring of each received resume against the Job description which gives a good indication of candidates to be shortlisted. Secondly it provides a structured list of contextualized questions to be asked during the interview which not help in effectively conducting the interview but also in objectively assessing the candidates against each other by using that questionnaire as a scoring checklist.

We have made extensive use of IPA and RPA techniques to build our system, as well as used process mapping and optimization techniques to find the right pockets for automation.

We have designed and developed our system up to a working MVP level with core functionality. We also have done extensive market research to evaluate the market size, competing products, customer segments and have found that although the market is quite saturated in terms of available products, there are unmet needs which are still fulfilled and which our system effectively addresses. The IRH efficiently supplements the existing products by plugging in the gaps of missing functionality, without directly competing with any of them. There is a business case accompanying the product with detailed financial analysis with all cost components and revenue streams, profitability analysis, investment needs, ROI and defined breakeven points. We also have comprehensive product development plan with features roadmap and launch plans.

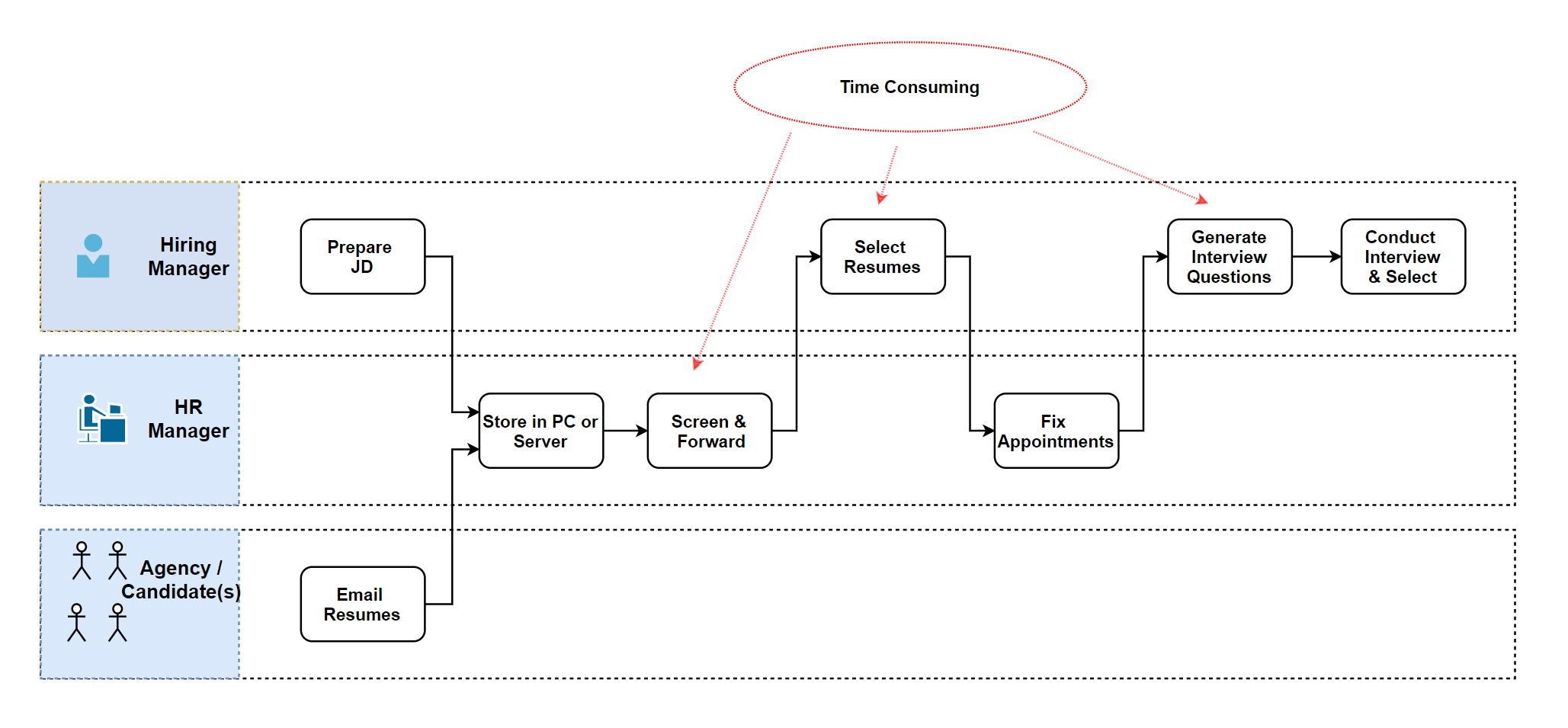
# Problem Description

A big challenge faced by Recruiters and Hiring Managers today is the number of resumes they receive, many of which may not even be remotely fit for the role. It takes a lot of time to carefully screen and evaluate every resume and this happens at various levels. First the HR recruiter evaluates the resumes, then he/she forwards a smaller set to the Hiring Manager who again evaluates the profiles and selects few of them whom he/she would like to interview and proceed further with. Hiring Manager may not be alone in interviewing the candidate and he may further distribute the set of resumes to other interview panel members, each of whom will spend a lot of effort and time in going through profiles and making selections.

As the HR recruiters usually support a number of Hiring Managers in the organization, the resume screening problem is especially acute for them since they receive hundreds of profiles in a day for multiple job postings, and they simply can’t do a meaningful job of preliminary selection. More often than not what happens is that they do a very basic sanity check of non-domain specific information and simply pass the whole set of resumes to the Hiring Managers, who end up receiving up to hundreds of profiles for each open role

A second issue is the prior preparation that the hiring panel needs to do for actually interviewing the potential candidates. For sure they can bank upon their expertise and experience to interview the candidates, but every interviewer wishes to have a readymade set of questions available to him, which could be used to drive the interview and also objectively evaluate all the candidates, especially when the number of candidates is high.

Through our Project we have tried to address above two problems – First to make the resume selection process automated, and second to build a set of readily available questions to help the panel interview the candidates

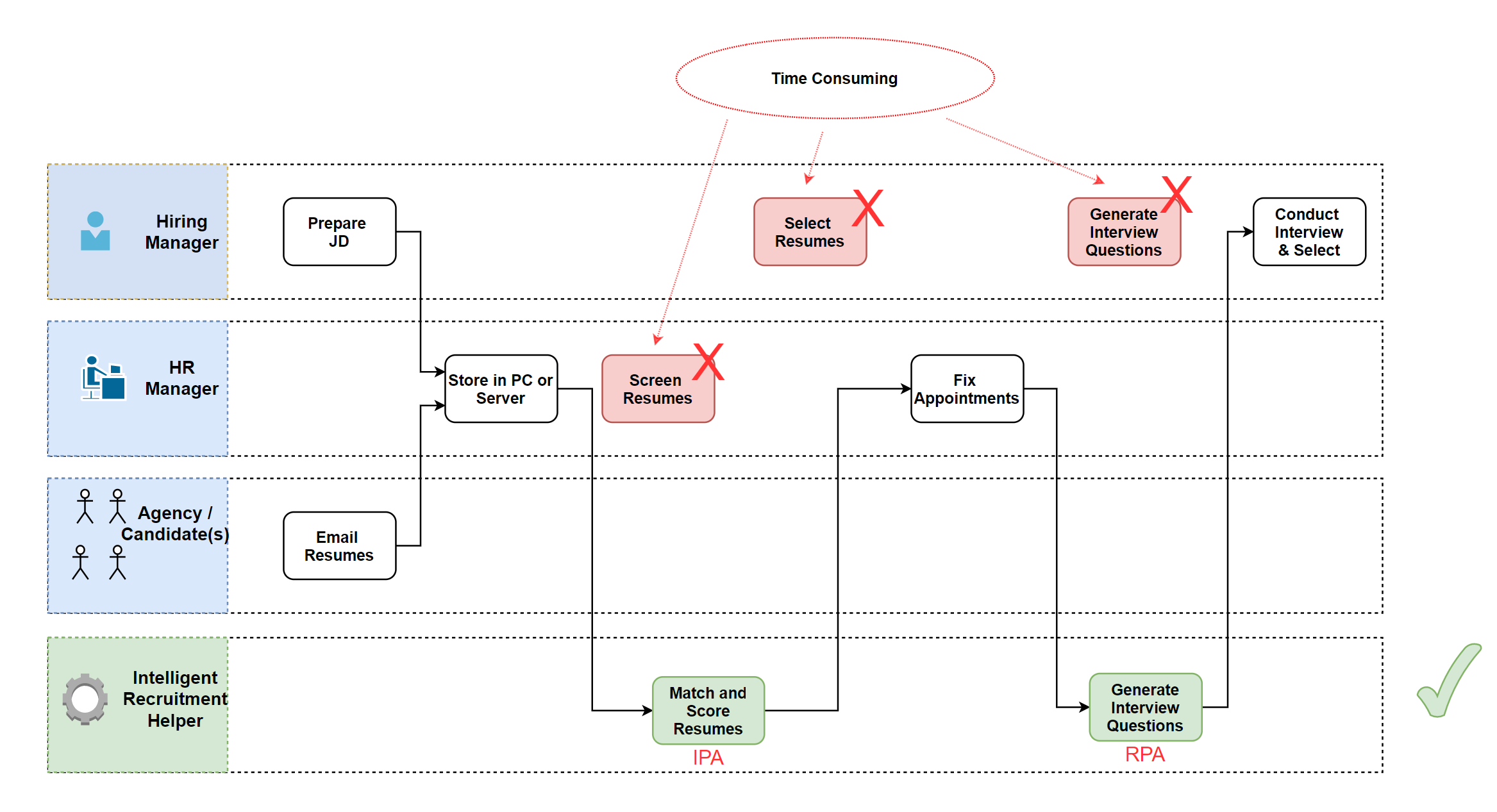


**Fig 2** – Problem Description

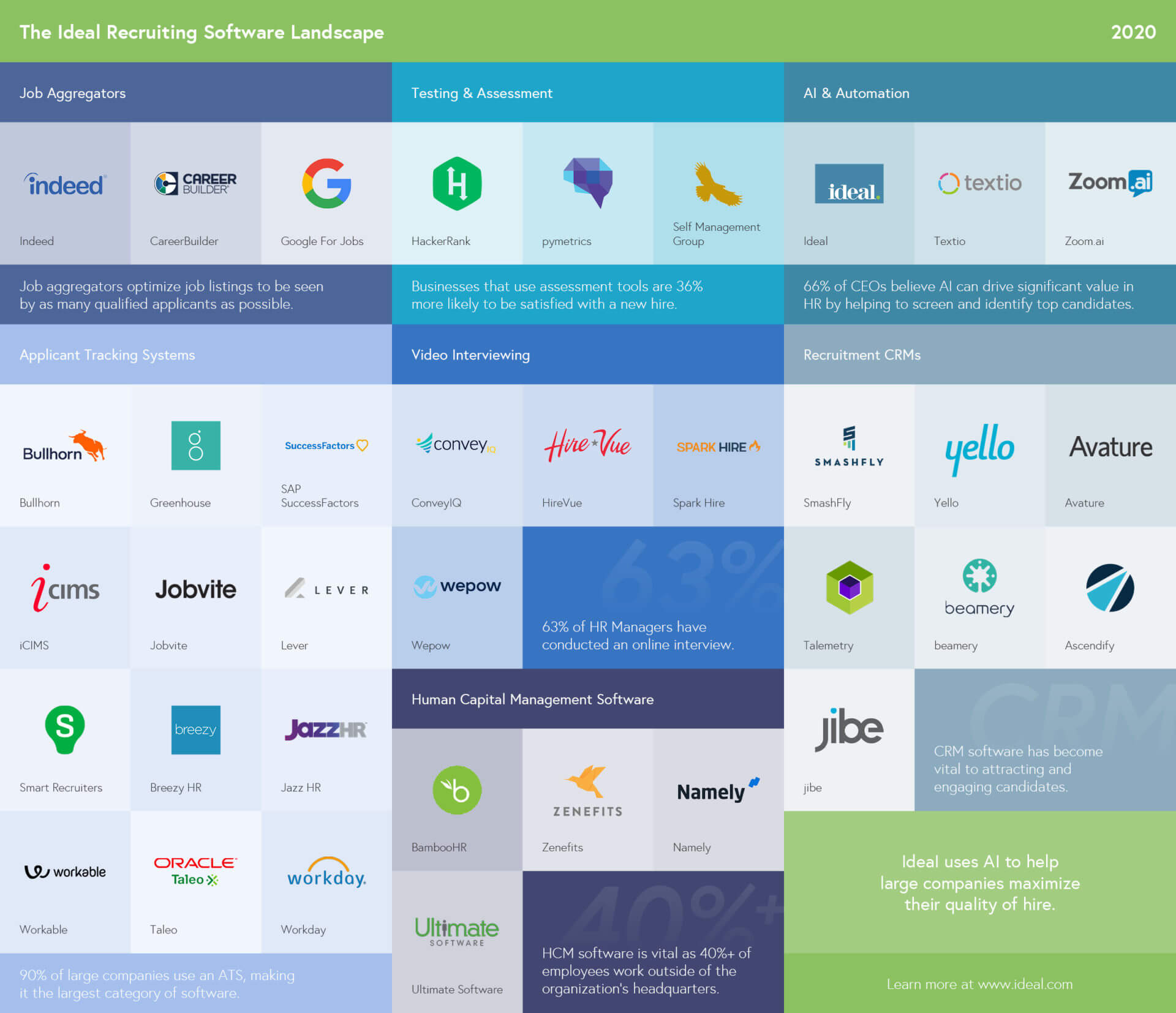
## Project Objective

Through this project we have tried to build a working system to achieve the following

1. An intelligent system which will take as input a Job Description (JD), and all the resumes and outputs a table containing a score for each of those resumes compared against the requirements specified in the JD. This score could be very useful for the recruiter to do a preliminary analysis for shortlisting, and could also be provided to the Hiring Manager to indicate priority and appropriateness of candidates
2. The second functionality of system is to provide a list or set of questions, specific to the domain and Job role, which can be readily used by the interview panel or Hiring Manager to conduct the interview. They can also use the questionnaire as an evaluation checklist and score all the candidates against those questions which will give them a very objective strategy for selection.

**Fig 3** – Solution using IPA and RPA

# Market Research



The recruiting process is time-consuming. Recruiters and hiring managers face resume overload. They sort, rank, and send applications to the correct person for review. Then managers collaborate with recruiters to screen, interview, and onboard qualified candidates. And along the way, everyone in the process must communicate with follow-up emails, internal messages, and offer letters.

It’s complex. It’s high volume. And losing track of candidates may mean losing a chance to hire the best employee for the opening or diminishing the company’s hiring brand in the candidate marketplace.

No wonder this is a lucrative market and hundreds of vendors now crowd this market with softwares trying to address specific areas and differentiating themselves on usability, customizations, analytics, and data management.

## Recruiting software trends and features

Functionality for most recruiting software is built around the following pillars:

* **Workforce Planning**: creating pools of internal and external candidates
* **Sourcing**: job posting management, advertising, social media promotion
* **Candidate Acquisition**: referrals, screening, assessments, and selection
* **Applicant Tracking**: interviewing, background checks, candidate communication
* **Onboarding**: meetings, orientation, training
* **Analytics**: reporting and metrics embedded throughout the system

The key trends in Recruiting Software are –

* Move towards Cloud and SaaS solutions
* AI and Automation capabilities
* Integration with existing HR systems and Suites
* Marketing capabilities such as posting jobs on multiple job boards and social media platforms
* Social recruiting functionality
* Mobile recruiting.

Below we have tried to categorize the Recruiting packages in terms of the customer segment they address

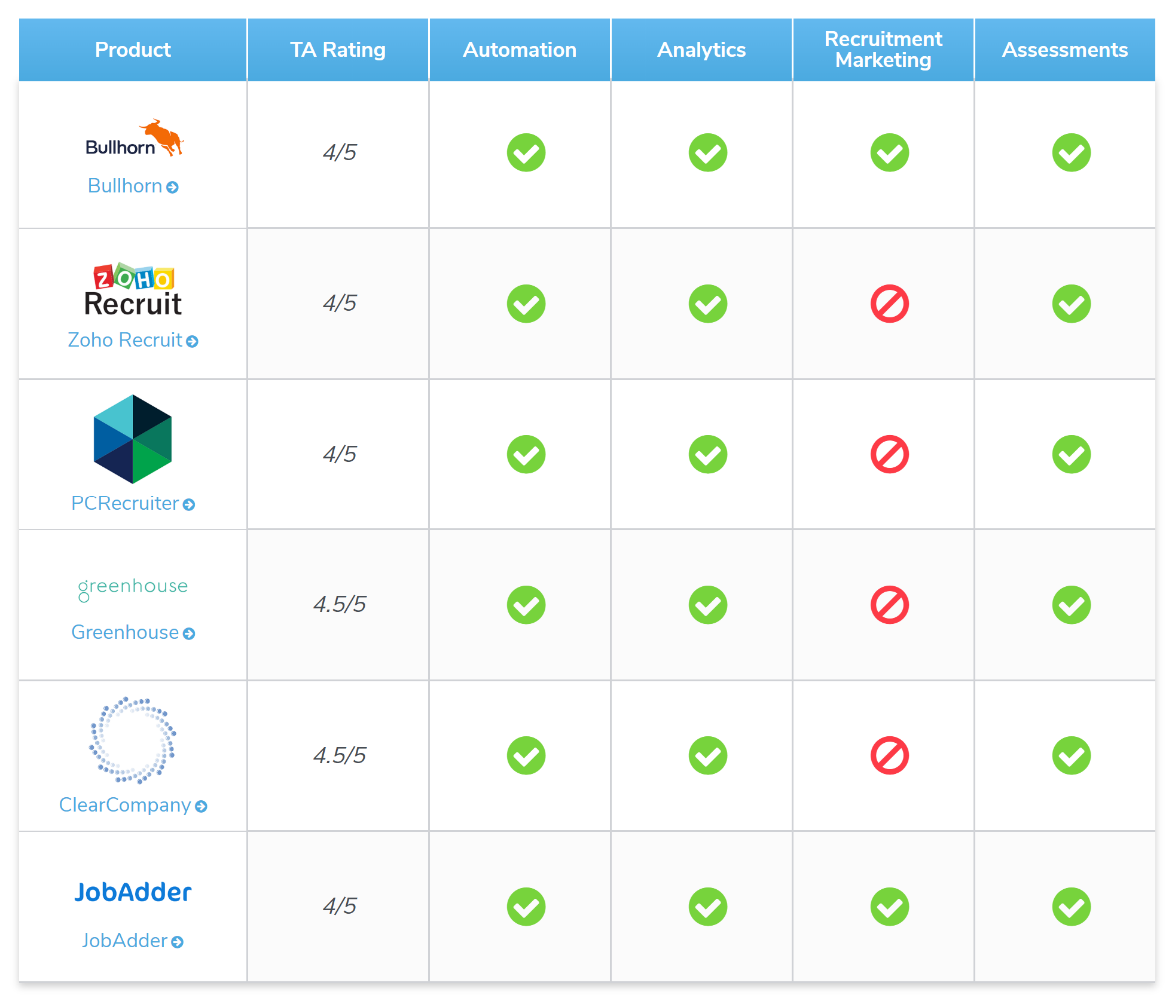
**Enterprise**

Enterprises and large organizations prefer packages which integrate with their HR or ERP systems, have strong collaboration features, and can support their internal Job portals. The top six are listed below



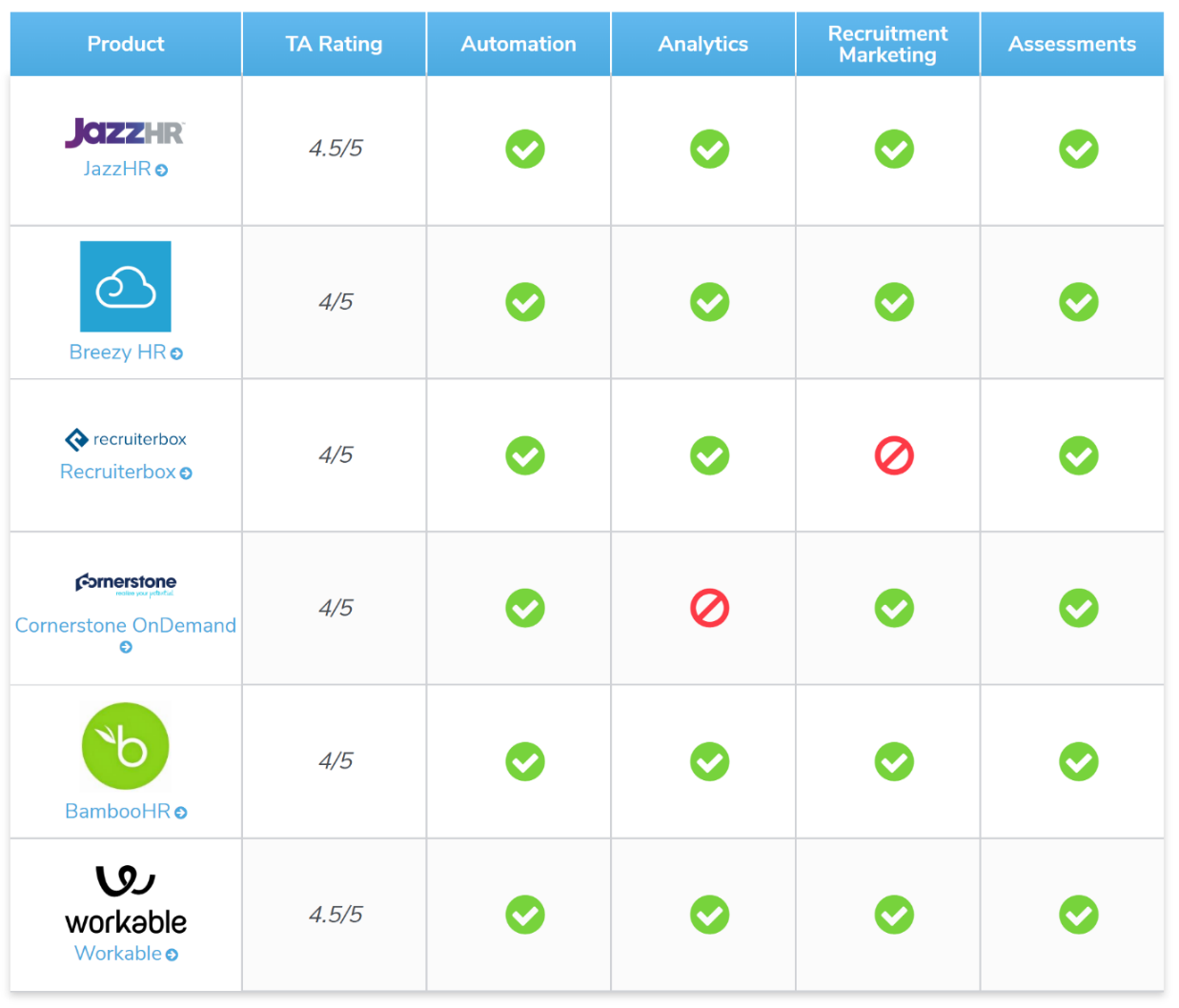
**Agency**

They have additional needs such as storing and searching large amount of data, advanced custom filtering, and email and CRM integration



**Small and Medium Businesses**

They often need simples and less expensive systems, which are scalable. SaaS and pay per use are a good model for them. Many large vendors also offer limited or free functionality for SMBs



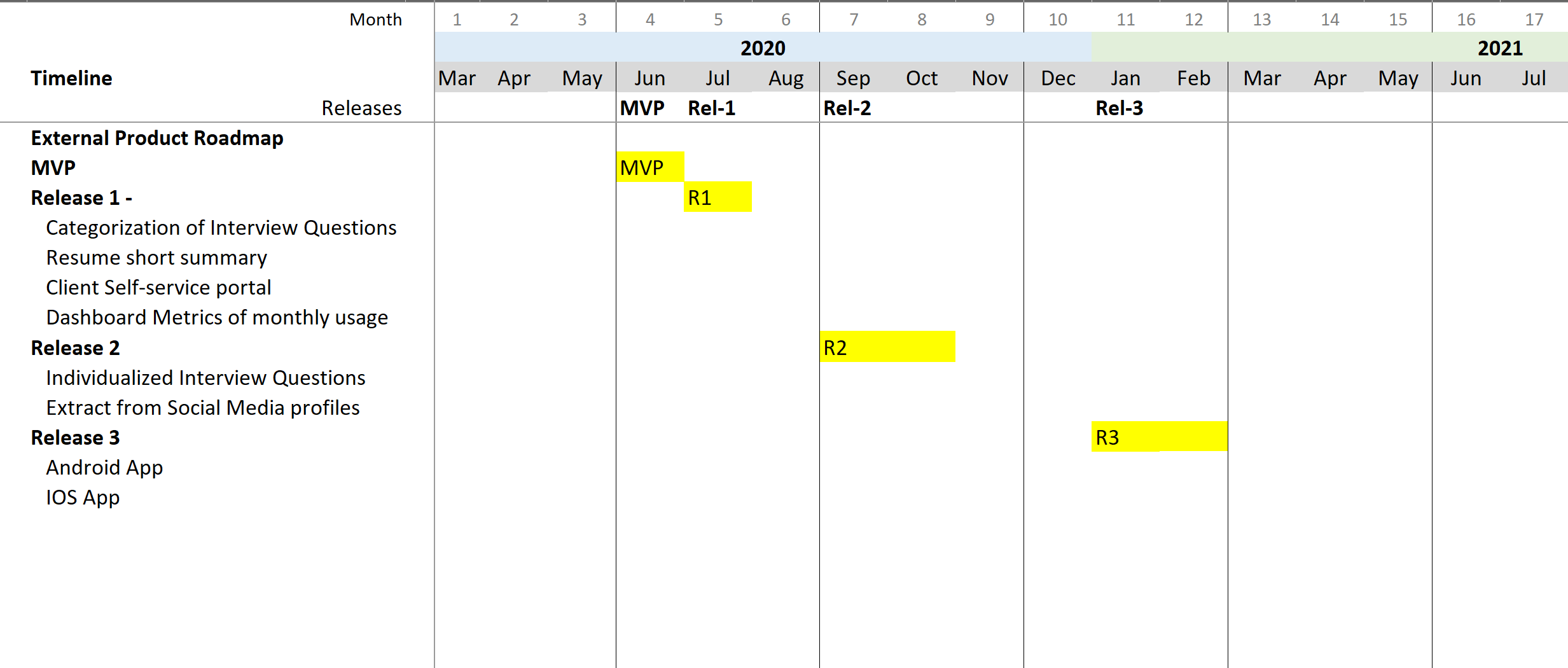
## Conclusion from Market Research

As we have seen the market is quite crowded and well addressed in terms of features, functionality, customer segments, and business models. There are high end systems which bring in advanced AI and Analytics capabilities as well. But we could not find any package which claims to have functionality of providing a **readymade list of questions** which hiring managers or panel can use to guide the interview and also rate the candidates objectively. Most packages leave that area to be handled by hiring managers as they seem fit. Also, the other key functionality of our Intelligent Recruitment Helper, **resume matching with Job description and scoring**, seem to be limited to high end expensive systems. **Thus, we conclude that we have a good business proposition across all customer segments, and our systems could supplement their existing systems to provide the added functionality**.

# Product Plan

## Launch Plan

A summary of Launch Milestones and features included in various releases is given below



**Fig 5** - Launch Milestones and Features

As of May’20 we are almost ready with a **Minimum Viable Product (MVP)**, which could be launched in the market for initial set of users. With some clean-up and packaging we believe that at start of June (month 4) could be the time for launch.

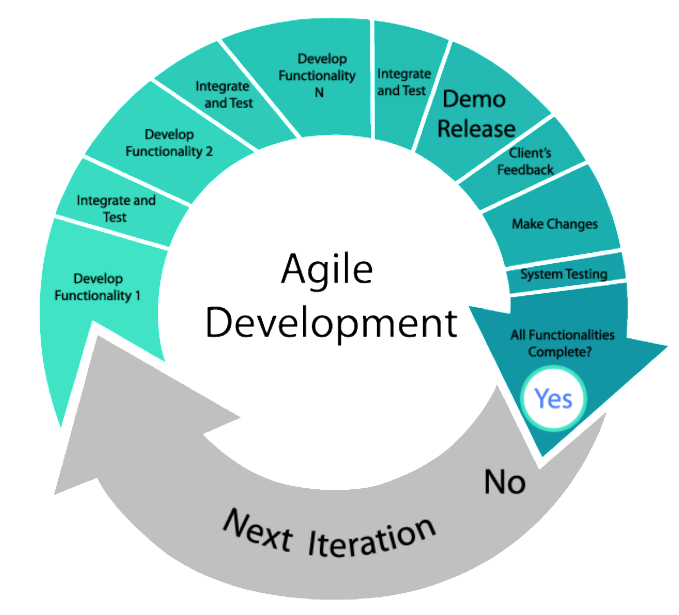
**MVP** will have the basic functionality of resume matching and score and providing a list of domain specific interview Questions to Interviewer based on Job Profile

**Release-1** would immediately follow in July (month 5) bringing in some improvements and new features such as Categorization of interview questions, short summary of resumes, Client self-service portal and Dashboard metrices

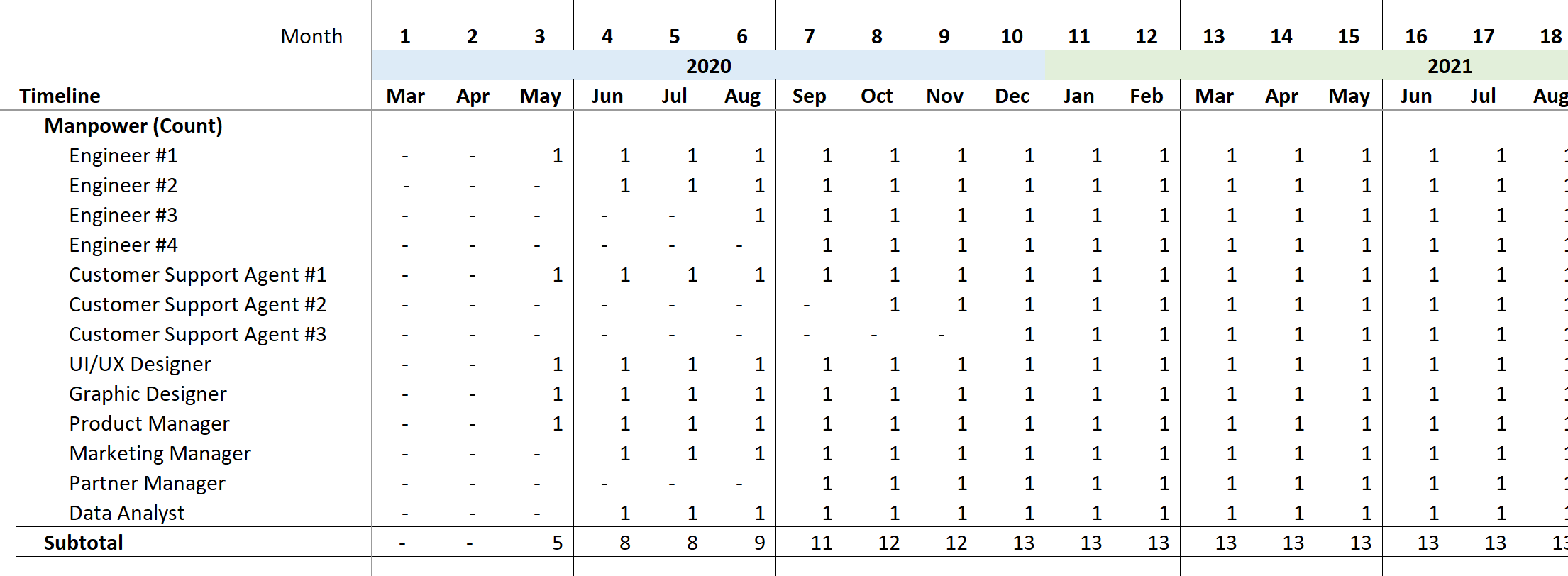
**Release-2** will follow after 2 months where we would develop the Interview Questions part further and individualize them based on interviewee profile. Also, we would add the functionality of providing a summary profile extract of candidate from social media sites such as LinkedIn which would include Recommendations and Skill endorsements from peers

**Release-3** will focus on building Android and IOS mobile Apps which could be installed and used by HR Recruiter and Hiring Managers to interact with our portal for various functionalities

## Development Plan

We will adopt an Agile approach for development of our product ‘Intelligent Recruitment Helper’. We’ll start with a Minimum Viable Product (MVP) with just enough features to satisfy the early customers and collect feedback for future developments. This would allow us to Go-to-market faster; also, following the same agile pattern we will be able to release features frequently and retain flexibility to change strategy - according to user feedback and signals from the market.

We will align our development resources as per the Launch plan and ramp up the team accordingly. A tentative resource plan is given below with counts and type of resources we will need



**Fig 6** - Resource Ramp up Plan

# Business Case / Financials

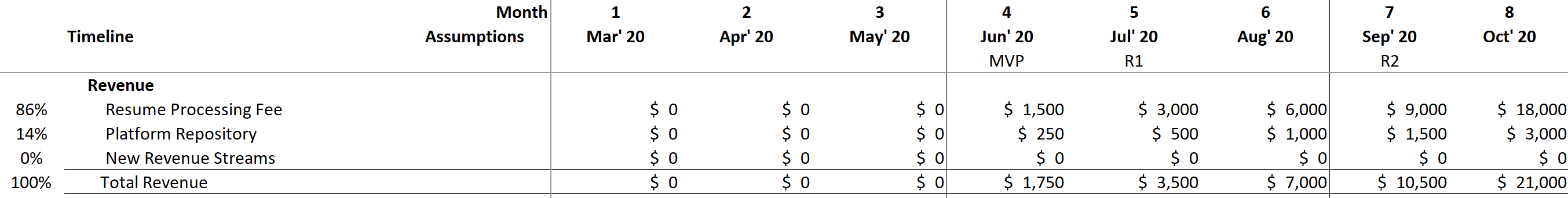
We have built a Business case including estimated Revenues and various costs components to give us an idea of Project profitability and initial investments needed to take us through the development cycle.

## Revenue Model

The Revenue model for our system would be **Pay as you Use** where we would bill the clients based on actual usage, the metrics that we plan to use is the total **number of resumes uploaded and processed** by our system during a month. We have assumed a price of **30 cents** per resume for the Business case calculation.

A second revenue stream which we have considered is Platform Repository functionality where HR Recruiters or Agencies can use our platform for **categorizing and storing** candidate profiles and Job postings. We have assumed a price of **5 cents** per resume for BC calculation

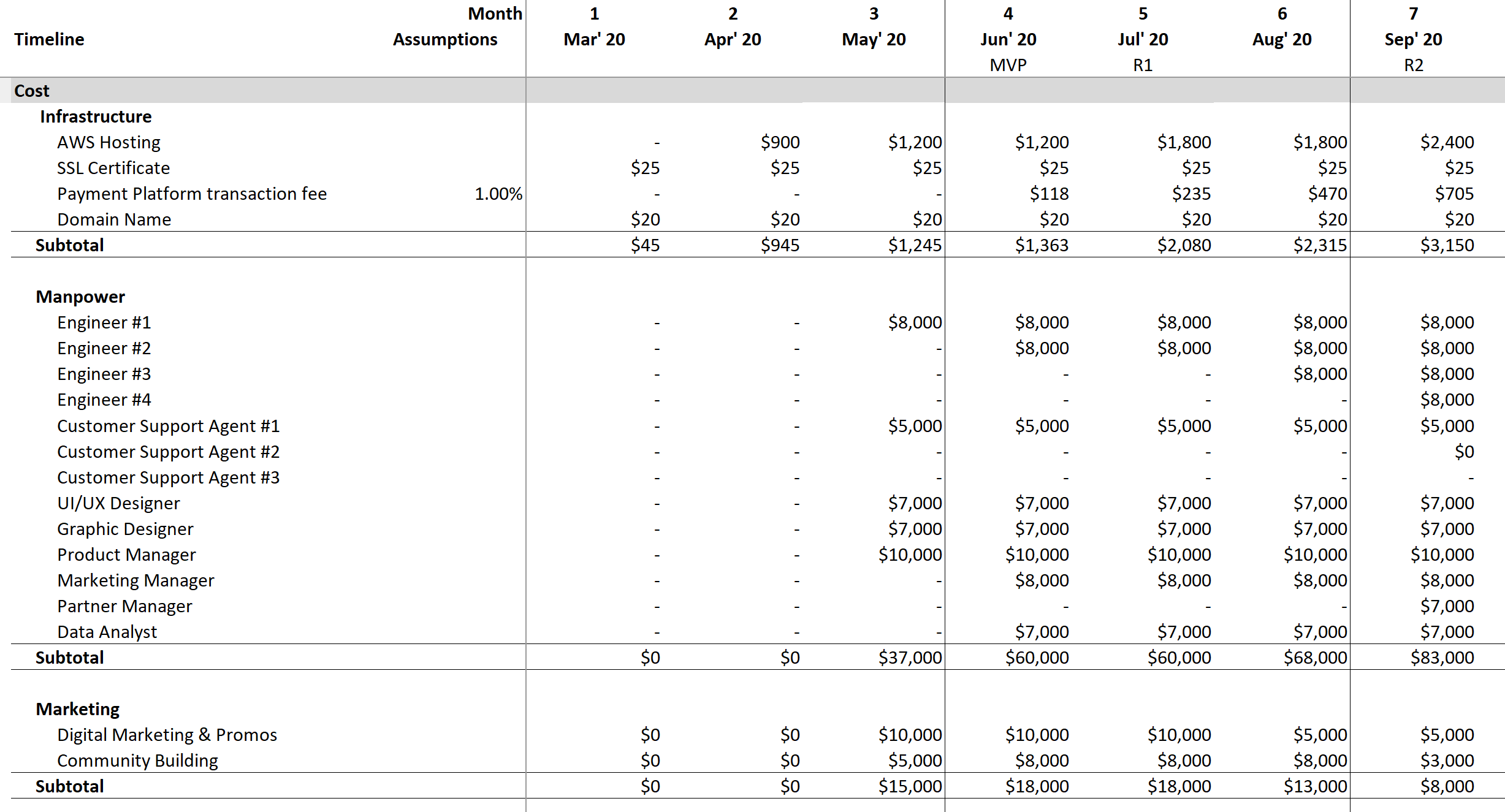
We understand that a successful business always creates multiple streams of revenue and we will look for further value add services that we can provide to our clients, and build functionality according to that based on the feedback.



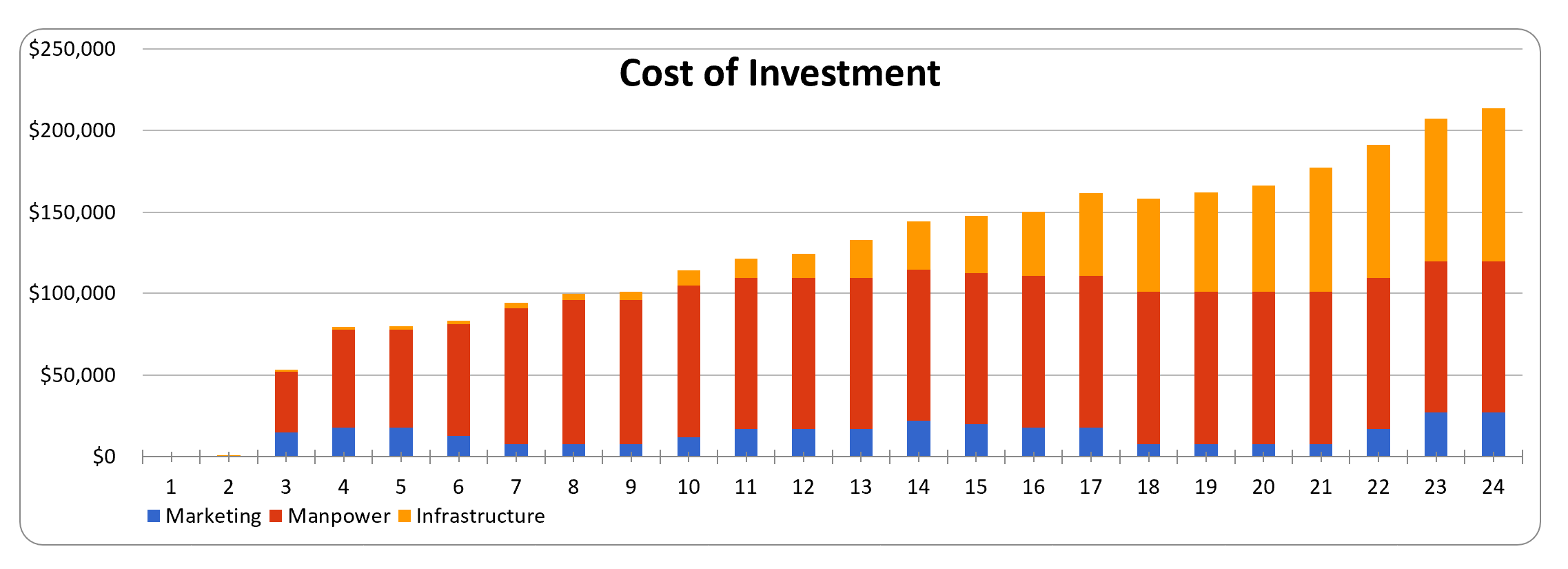
**Fig 7** - Revenue Streams

## Cost Model

We have primarily considered three main cost components for business case calculation – Infrastructure, Manpower, and Marketing.



**Fig 8** - Cost Components



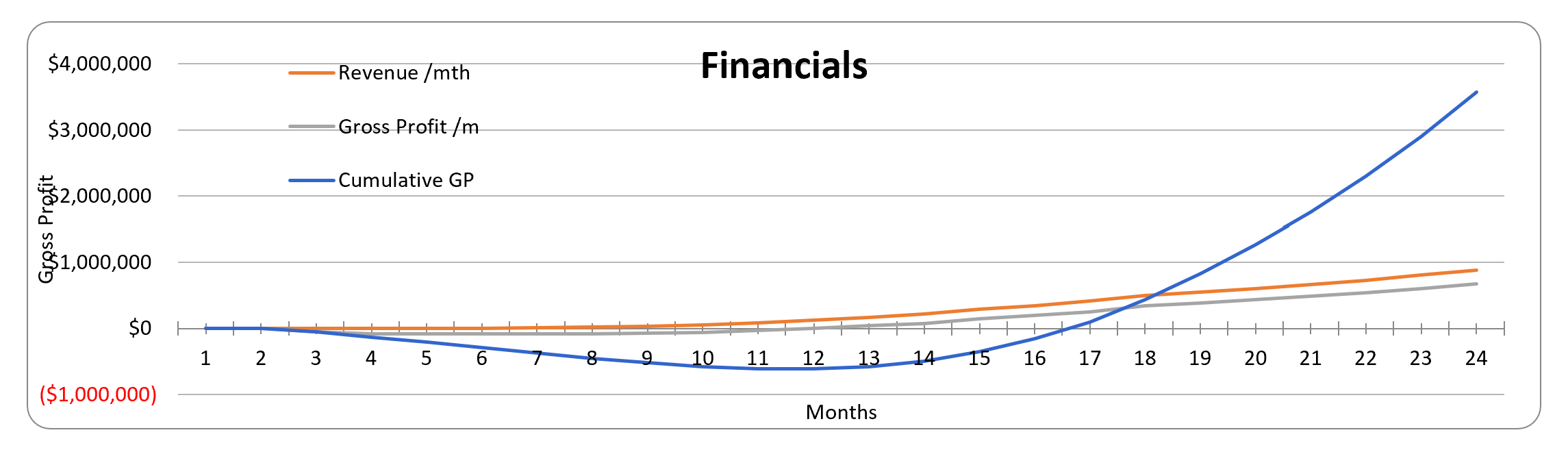
**Fig 9** - Cost breakup over time

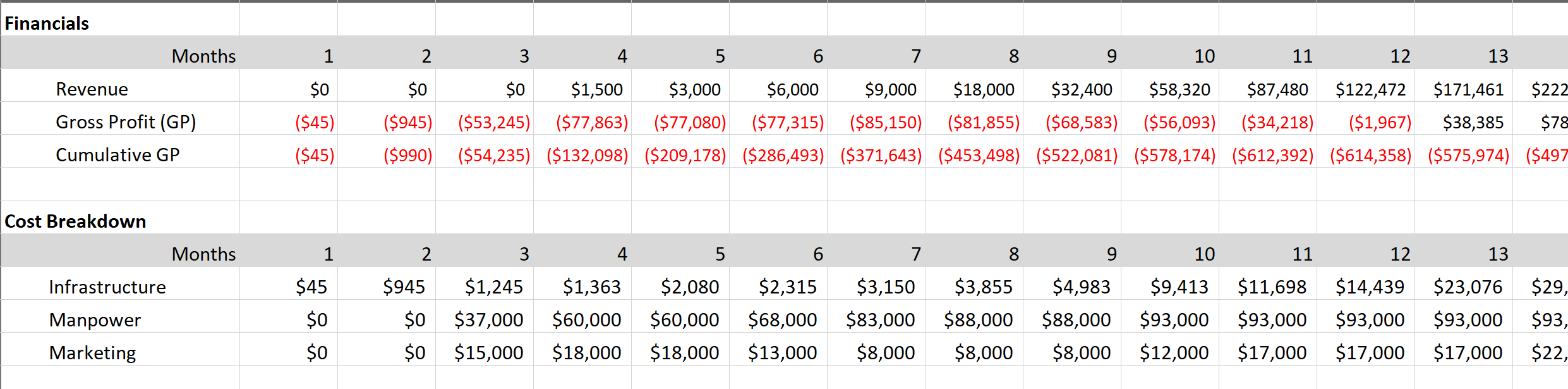
## Profitability

The overall profitability of the project over a period of 2 years is show below. We can see that overall, we achieve a **Gross Profit of 55%,** and a **ROI of 120%** which are very healthy



Below charts and tables provide a detailed view of our projected financials and project profitability





**Fig 10** - Business Case financials

We can see from our profitability projections that we only become profitable in the 13th month, and on a cumulative basis we breakeven in month 17.

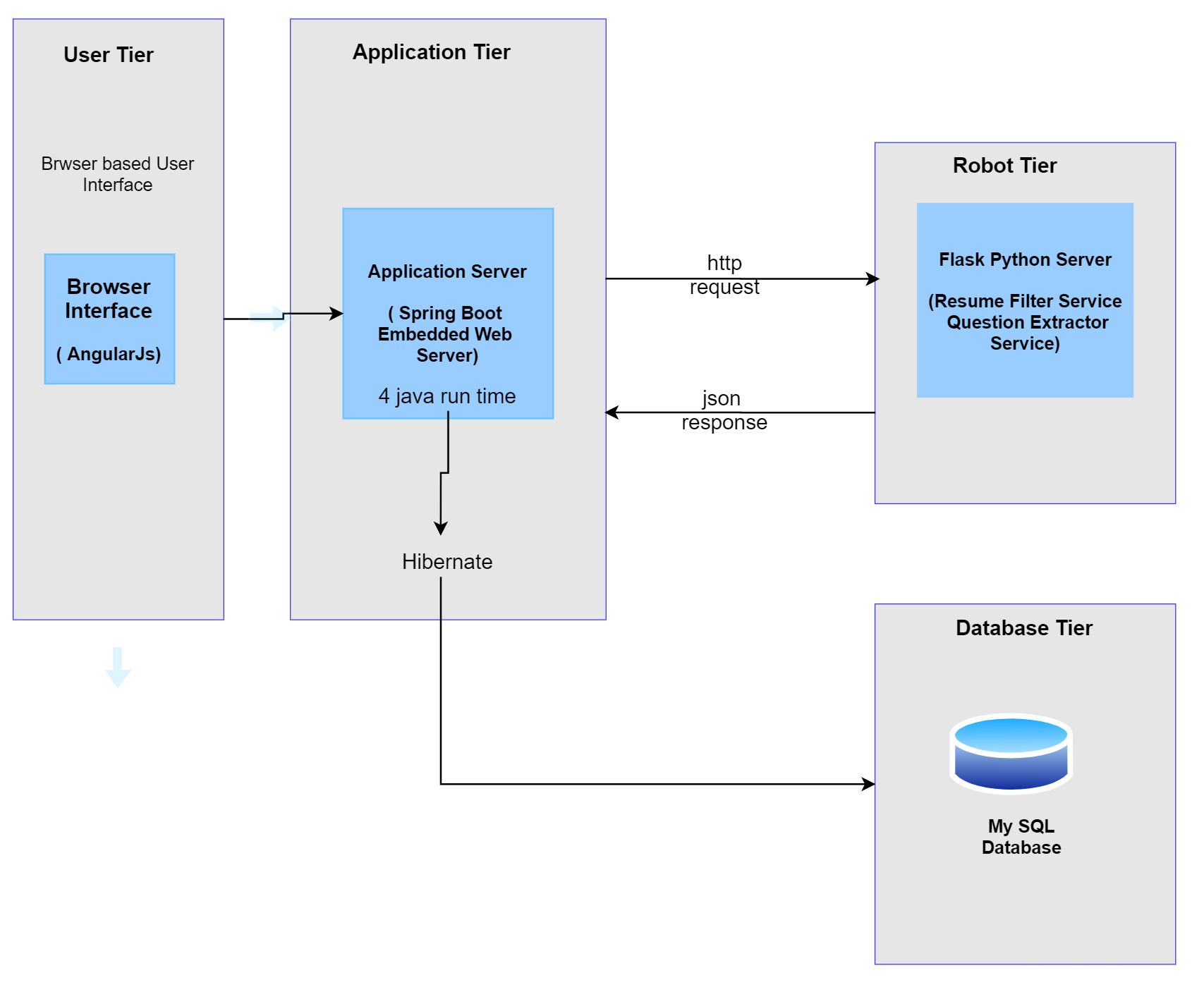
Also, we see that lowest figure for Cumulative GP is $ -614,358 in month 12, which gives us an indication that we need at least this much investment or seed capital to make the project viable.

The business case is attached herewith



# Solution – Intelligent Interview Helper

## System Architecture / Model



**Fig 11** - High Level System Architecture

Below are the various components we have used

* **Front-end User Interface**: Angular JS
* **Middle ware**: Spring boot - 6 microservice separate
* **Database**: My SQL
* **AI and Automation Layer**: Flask, Python, TagUI.

We have used a Microservices based architecture.

There are 4 separate Microservices -

1. **Job/Recruitment Microservice** - for Post job, List job, all job-related features
2. **Employ module Microservice** - for create user, edit user (like HR Recruiter, Hiring Manager who uses the application). Calling the resume filter, question filtering all come in the module
3. **Authentication Microservice** - only for authentication using Java token authentication JWT. All request for all the microservices module first reaches the Authenticate module, then it is checked whether Authentication token is already available in auth (session) and only if valid, gets access.
4. **DB service microservice -** All Database operations are handled through this microservice.

All these microservices can run independently. So, code maintenance is easy. E.g.- if any change needs to be made to a particular Microservice, only that part would be changed. All these microservice connect between each other during execution

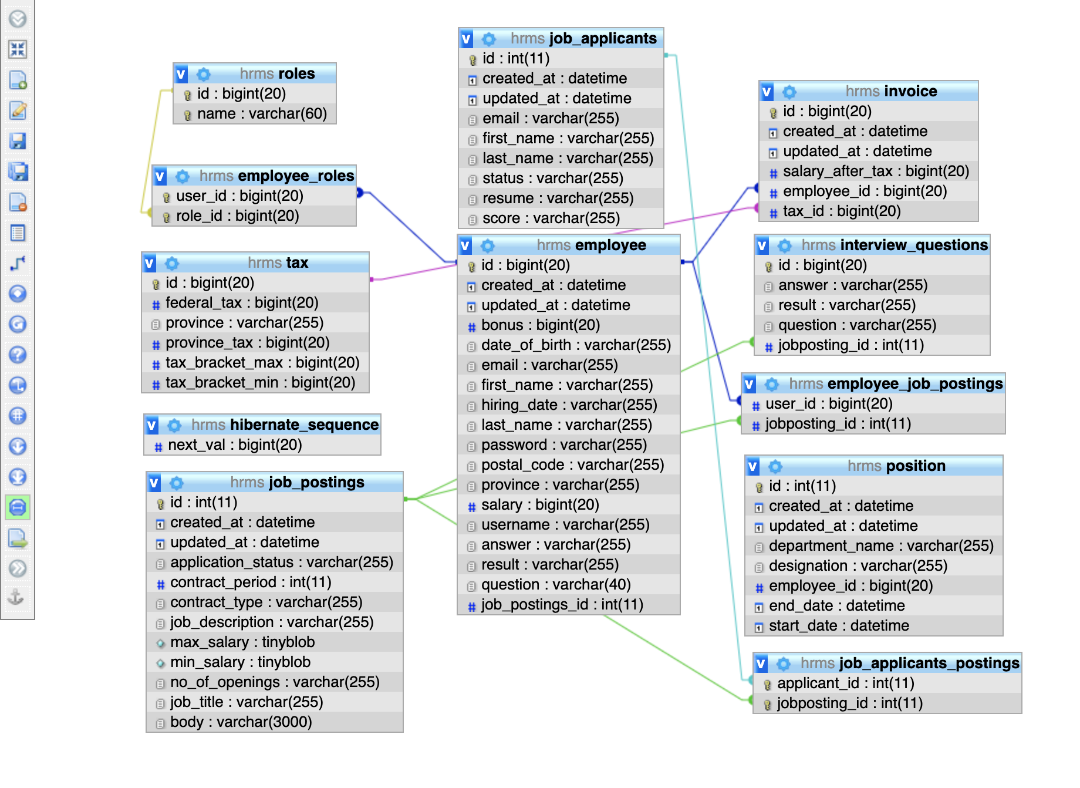
* SUUL API gateway interconnects all the microservices.
* EUROKA is a monitor that monitors all these microservices. Status being Up/Down etc.

AI Service:

1. **Resume Filter** – The HR Recruiter uploads the .zip file containing all the resumes and job description. The middleware service forwards the .zip file to the AI Tier, where zip file is extracted, each resume is read individually, and score is calculated for each resume against the job description
2. **Questionnaire Extractor** – When HR Recruiter clicks the ‘Generate questions’, the job description, role, and domain is provided to the middleware service, which in turn forwards it to the AI Tier. The AI tier then uses TagUI to perform Google Searches, selects relevant websites to visit, and filters out the questions from those websites. Finally, the questionnaire list is returned back to the User

### Database Design

Below is a description of all the Databases we have used in our system



**Fig 12** - Database Design

* **hibernate\_sequence:** Table for hibernate framework for db handling
* **employee\_roles:** Mapping of employee and role, to give privileges to users
* **job\_applications\_posting:** Used for mapping of job applications
* **employee\_table:** To enter details of users who use the application
* **job\_applicants:** Details of all applicants, score, resume names
* **interview\_questions:** This stores the result of question extract service
* **roles:** This stores all the user roles in the system
* **job\_posting:** This is to store all the new job postings created by HR Recruiter
* **employee\_job posting:** to keep records of all the job postings created by a particular user



### IPA and RPA

Our Intelligent Recruitment Helper System uses IPA and RPA technologies in two major sections

1. **Resume Scoring and Filtering** (Intelligent process automation) –

* Used Cosine Similarity TFID
* Bulk upload of .docx file as zip file along with Job Description
* Extract all .docx file from zip file and convert to text files.
* Used nltk python library to clean up data using English stop words
* TFIDF calculation using all the resumes and job description
* Cosine similarity calculation between each resume and job description to generate score
* Return the score to front end

1. **Question extraction** (Robotic Process Automation)

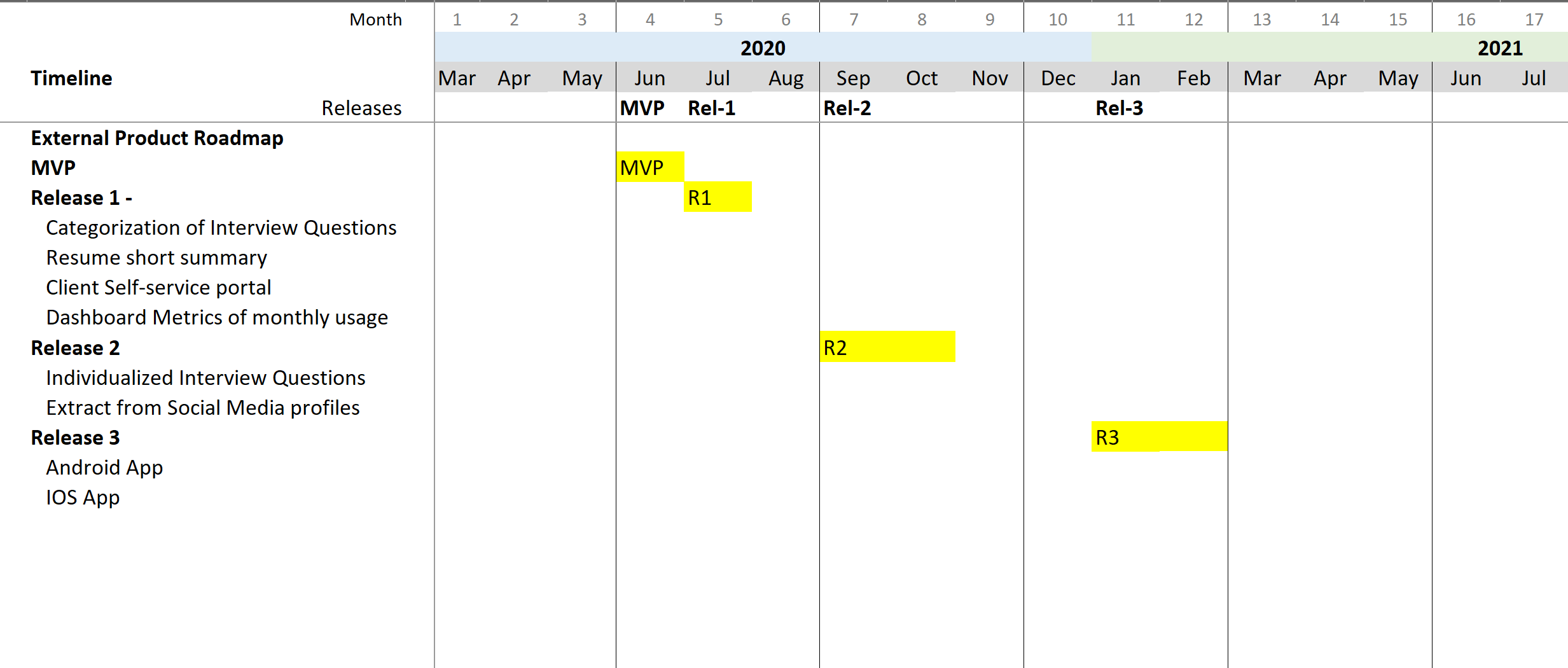
* Take Job description, domain, and the role as inputs
* Using TagUI RPA framework explore google to get all sites URL that provide FAQ’s for the domain role
* Using TagUI, visit each URL and pull the whole text from the sites
* Use python regular expression to extract Questions (sentence that have question marks ‘?’) from the full text.
* Return the list of questions to front end

## Limitations and Future Enhancements

The Intelligent Recruitment Helper System that we have developed is at MVP stage where with a bit of User interface improvement could be launched as a commercial product. MVP will include the core functionalities of Resume matching and scoring, and generating a list of questions to be used by Interviewers.

One of the limitations currently is that for the Interview Questions searching and extraction we are only using the Job role details, and the questions list is not individual specific. It would be very powerful to contextualize the questions to specific candidate and also take into account extract from professional profile portals such as LinkedIn for the search. We have pushed that functionality to Rel-2 of the product

Below is a roadmap of future enhancement to the product (details in [Launch Plan](#_Launch_Plan) section)



**Fig 13** - Future Enhancements

# Conclusion

Considering the small size of our team (2 persons), we believe we have done a meaningful job of putting all the learnings from various modules covered under RPA / IPA course. We have used TagUI covered in RPA module, Process mapping and Analysis that we learnt in RPA strategy module, and IPA techniques that we learnt in ISA / IPA module. We have also done a comprehensive business case analysis and kept the MVP concept in mind while designing the system.

We sincerely hope that we can take this product to market as we genuinely believe that our system addresses an unmet need in the market. We are at a stage where we can approach the VCs, present our product, and seek funding for further development and marketing efforts.

# Appendices

## APPENDIX A: Installation and User Guide

### Installation

**Remote Installation**

We have deployed our application for easy access at URL <http://cfuat.xyz/hrms>

**Local Installation**

AI service

* Download the PyCharm project from Git-Hub
* Install PyCharm in system
* Open the project using PyCharm
* Install all dependencies as mentioned in the requirements file
* Run the app.py from the PyCharm project

Microservices Installation

* Step for setting up the system
* Install Java 1.8
* Install MySQL on local host system with port number – 3306
* Create database name: cfuat\_hrms and username/password: cfuat\_spring/Codeface@123
* import hrms mysql.sql dump into MySQL (import the dataset)
* Run all micro service module in following order

1. Eureka Service (eureka.jar)

comment: java -jar eureka.jar

2. API Gateway Service (gateway.jar)

comment: java -jar gateway.jar

3. Db opera%on micro service (db-service.jar)

comment: java -jar db-service.jar

4. Authen%ca%on Service (auth.jar)

comment: java -jar auth.jar

5. Employee Service (employee.jar)

comment: java -jar employee.jar

6. Recruitment Service (job.jar)

comment: java -jar job.jar

7. Payroll Service (payroll.jar)

comment: java -jar payroll.jar

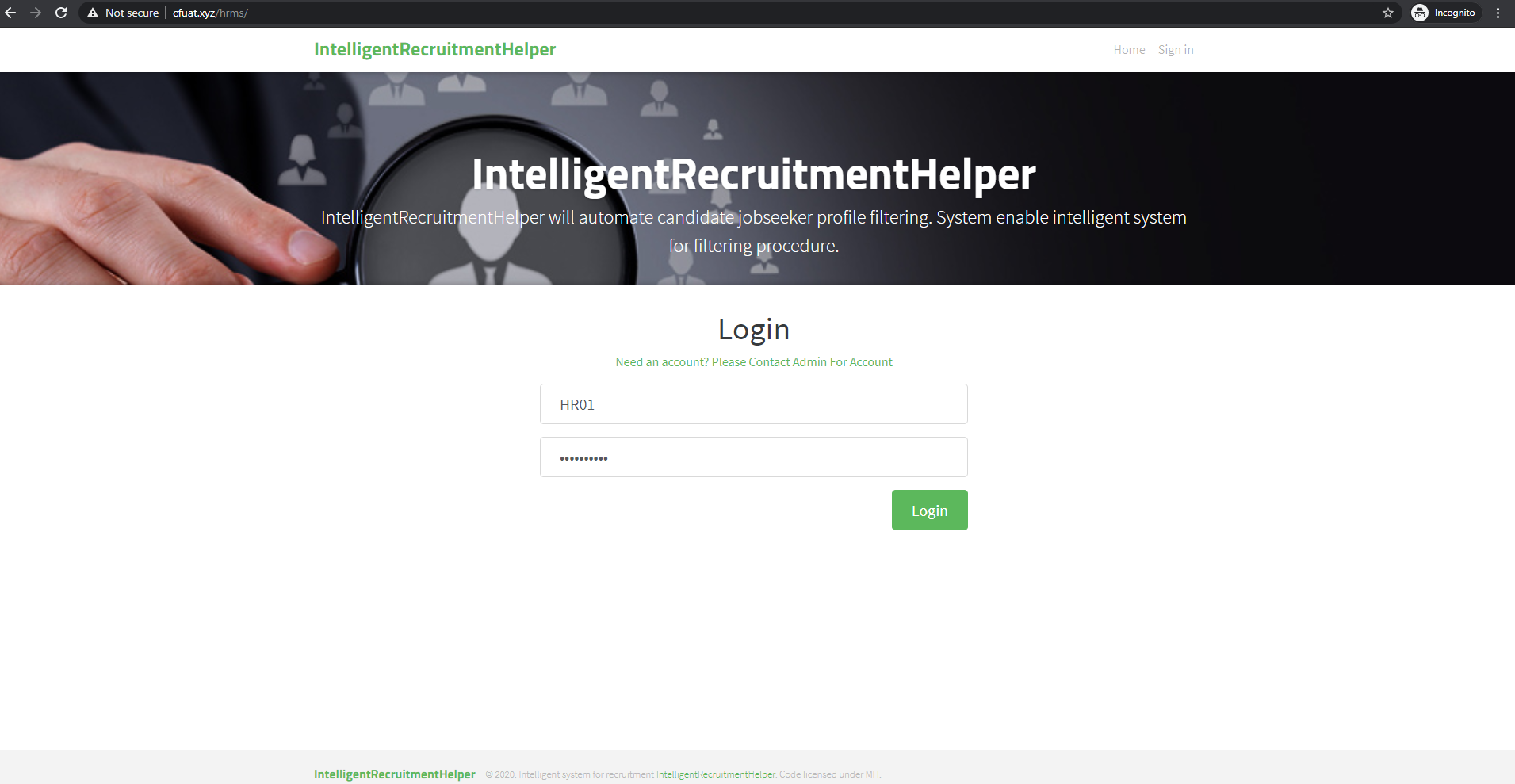
8. Posi%on Service (posi/on.jar)

comment: java -jar posi/on.jar

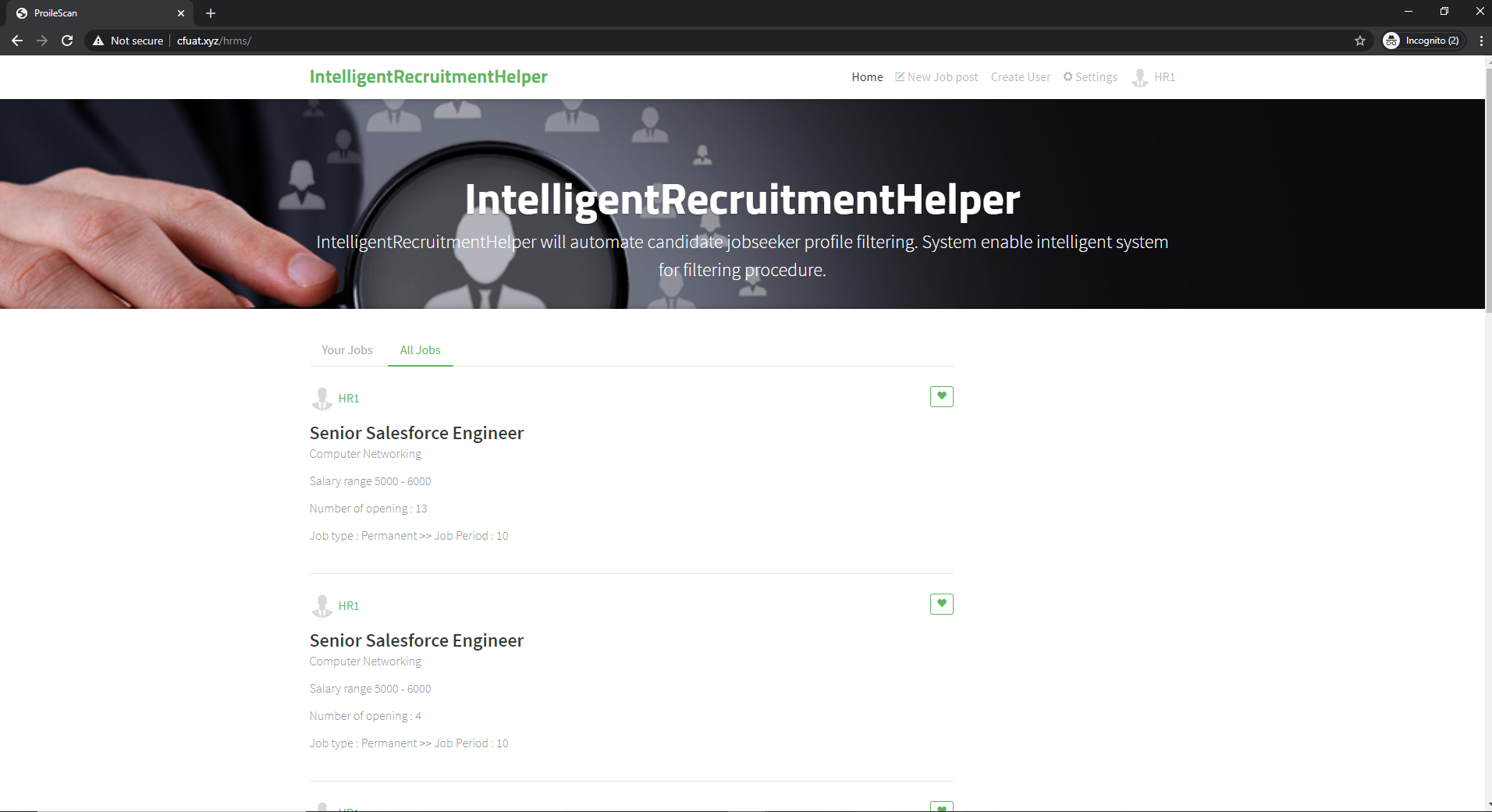
6. Deploy Client side applica/on on any applica/on server

### User Guide

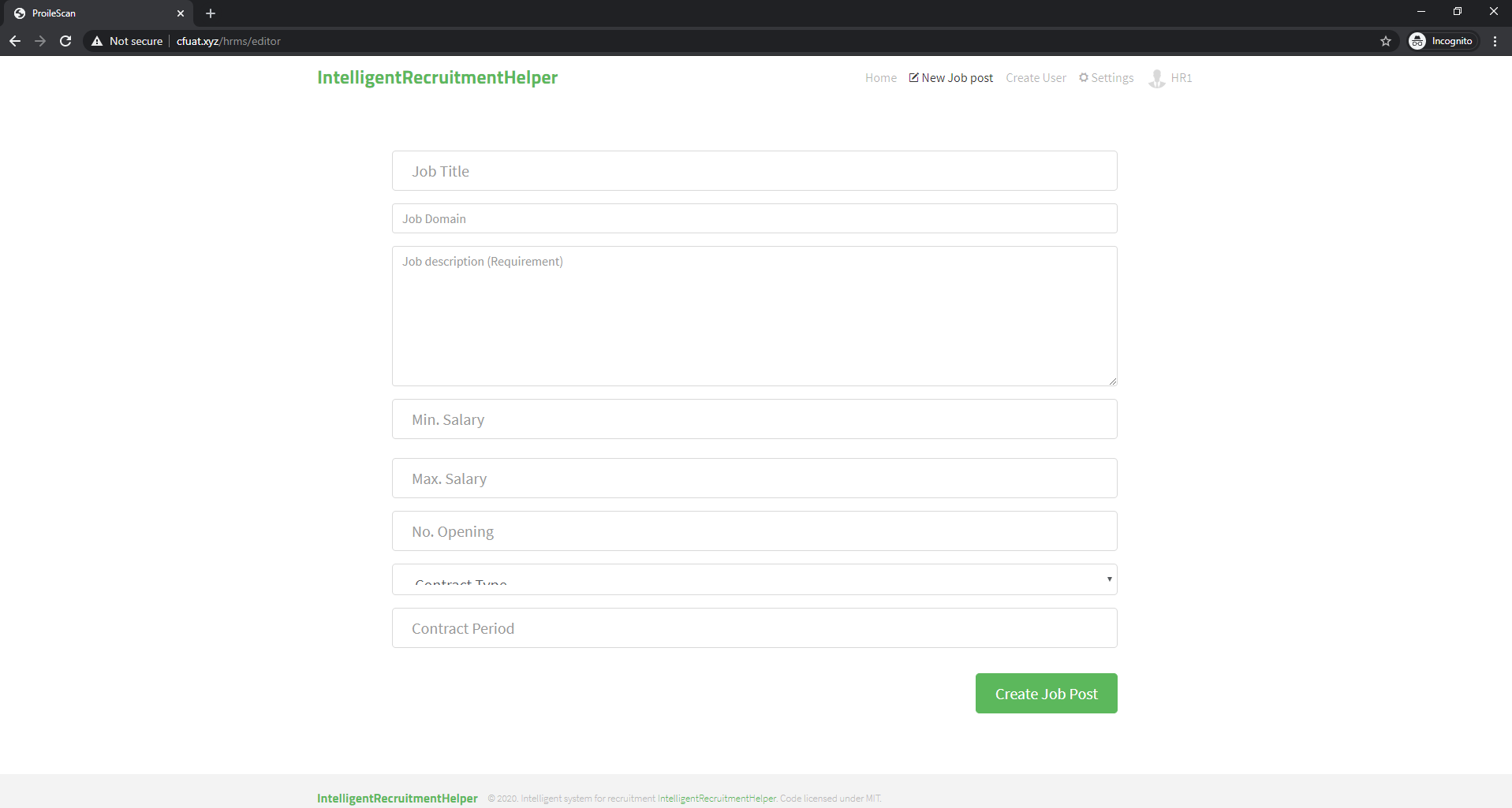
1. Access the URL at <http://cfuat.xyz/hrms>

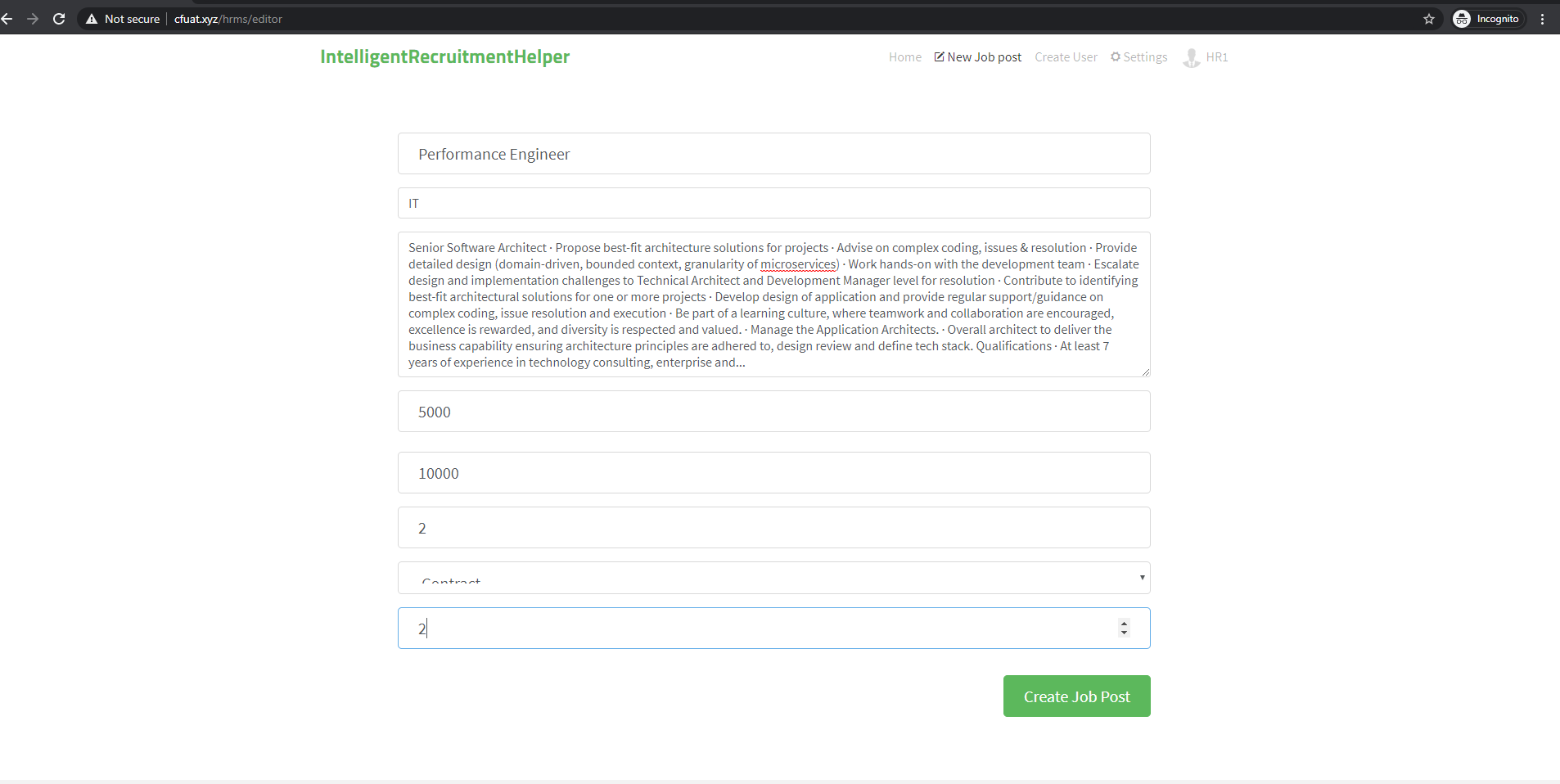


2. Login using HR1 and password 1234567890

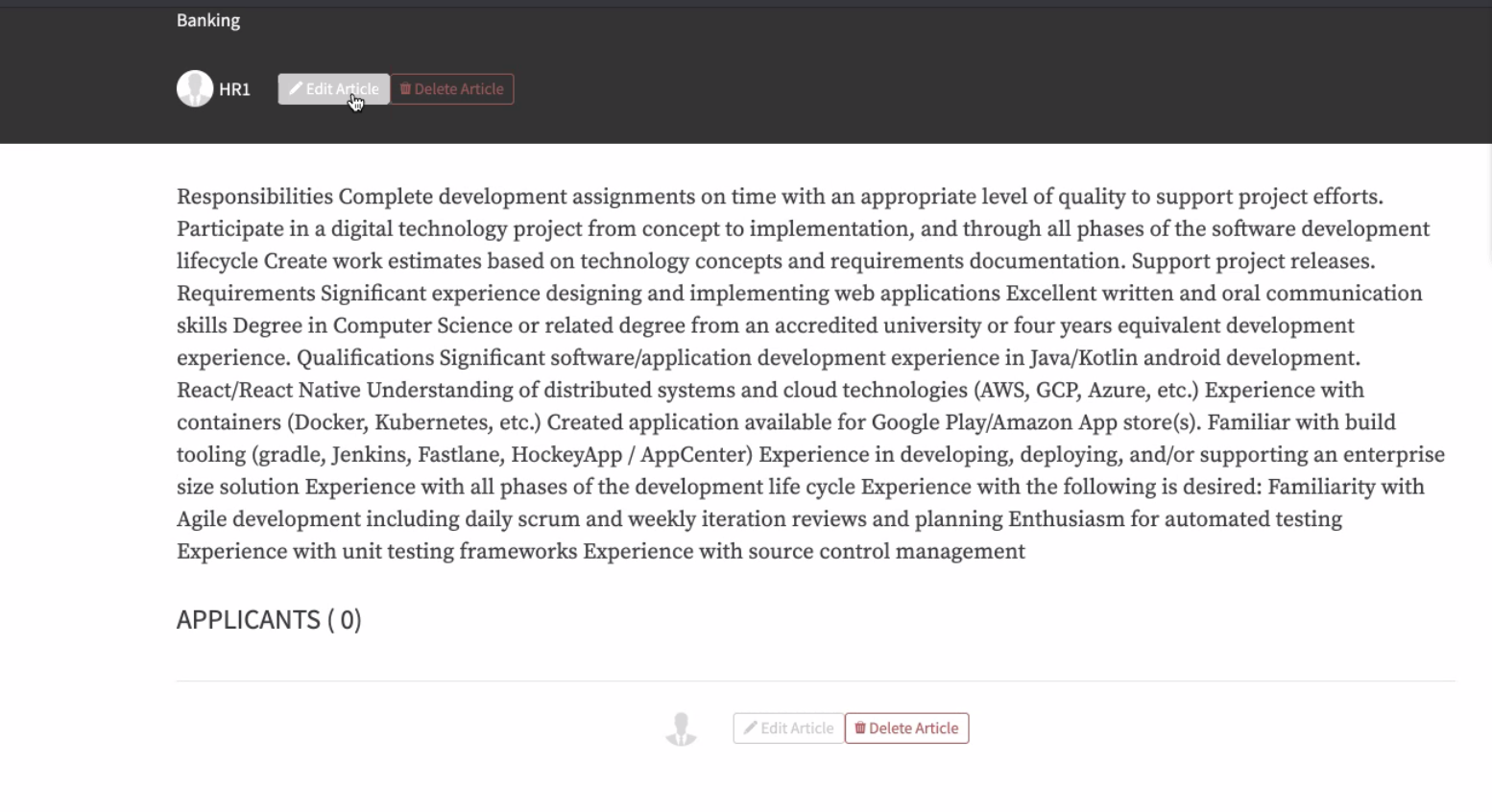


3. Create a new job opening by clicking create job and enter all mandatory details

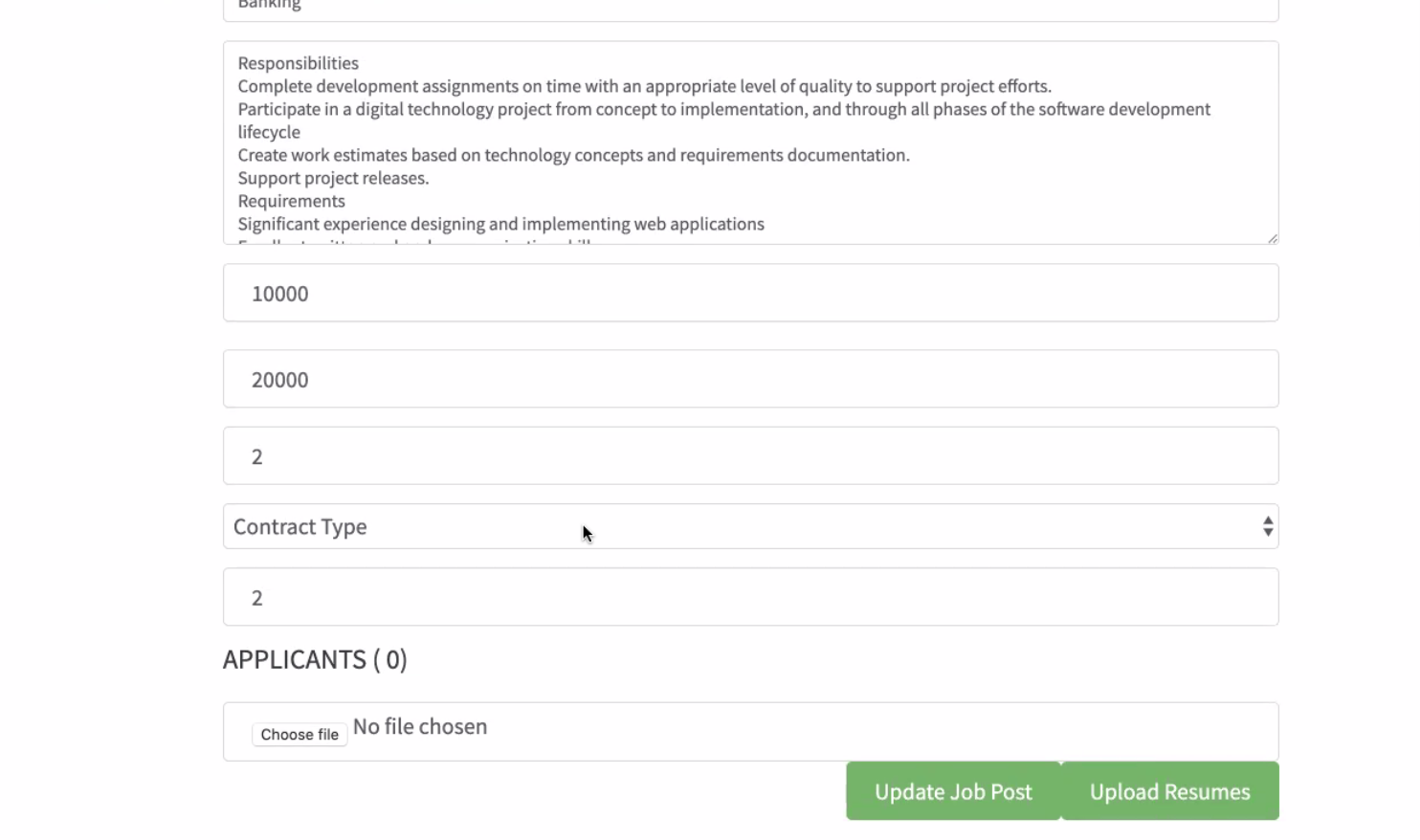




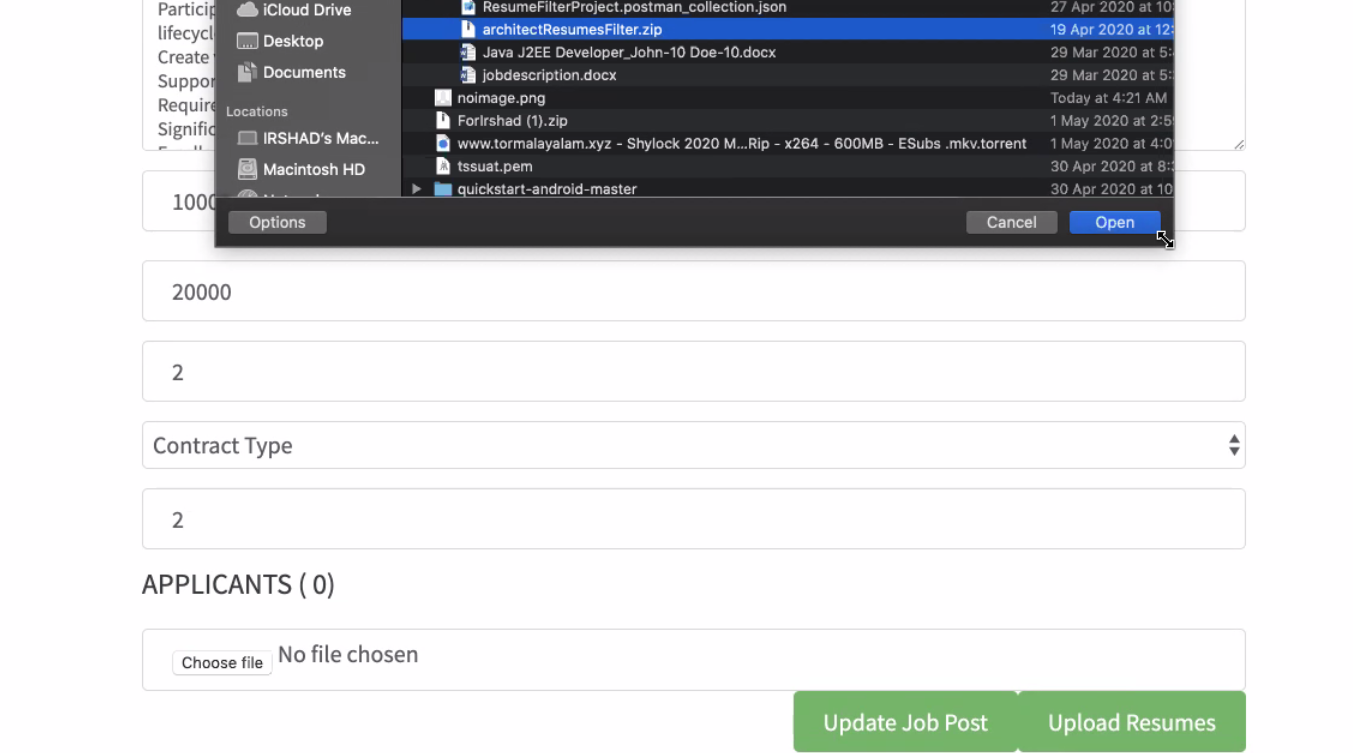
4. Click Edit Article

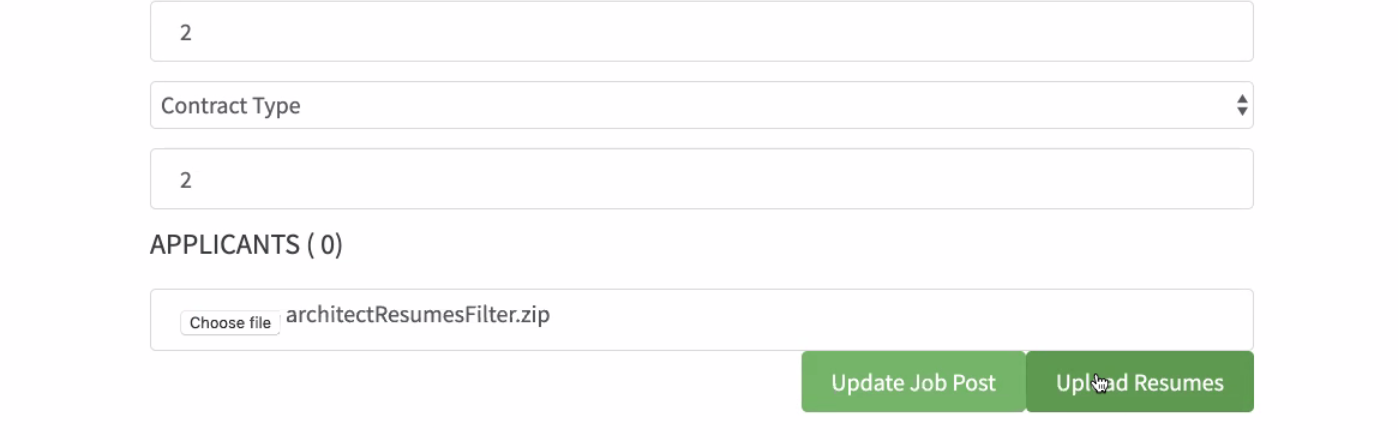


5. Click Upload Resume



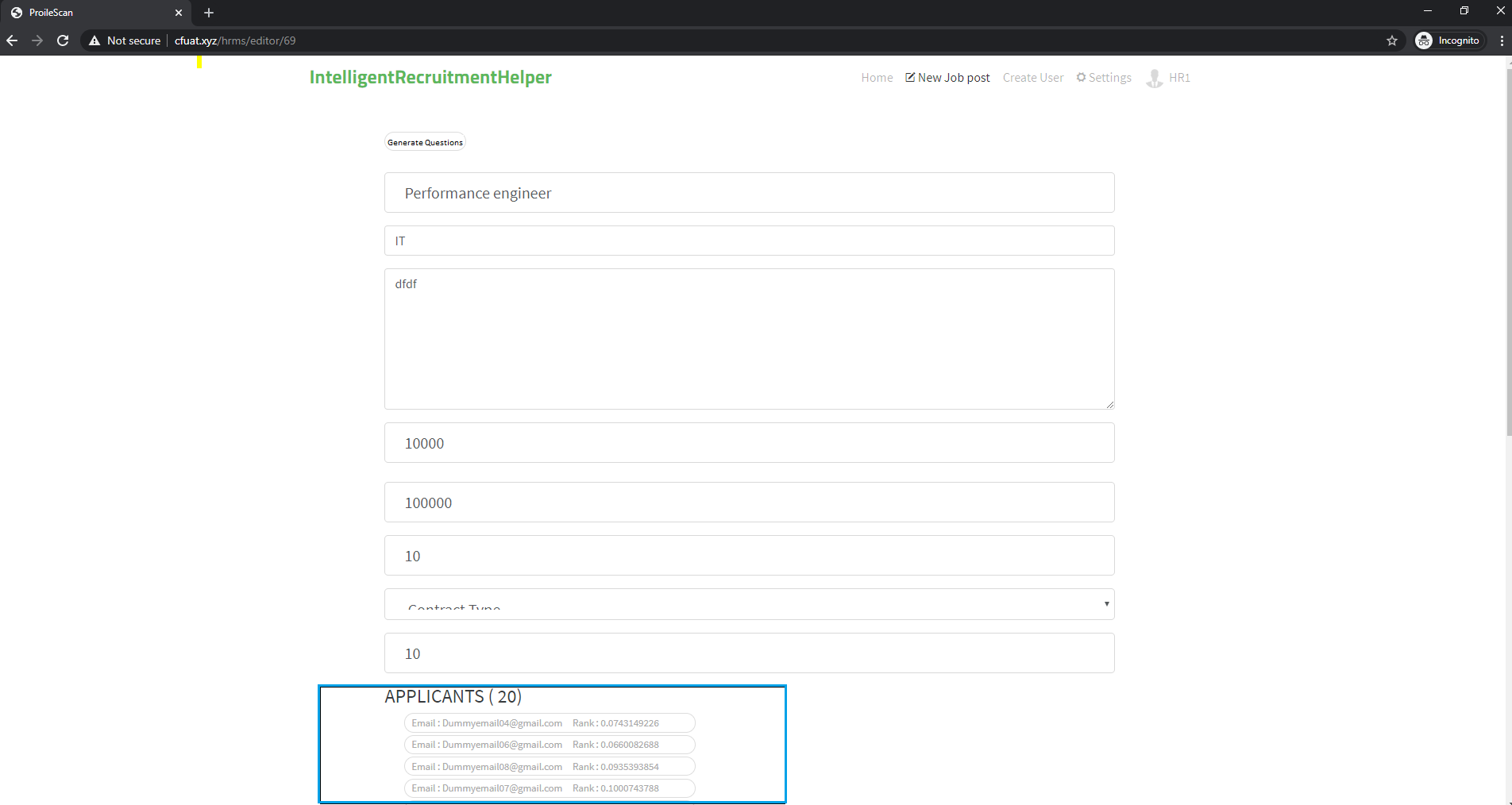
6. Go to created job and then upload the zip file containing resumes, and job description in docx format (keep the Job Description file name as jobdescription.docx, also the docx extension need to be in small letters).



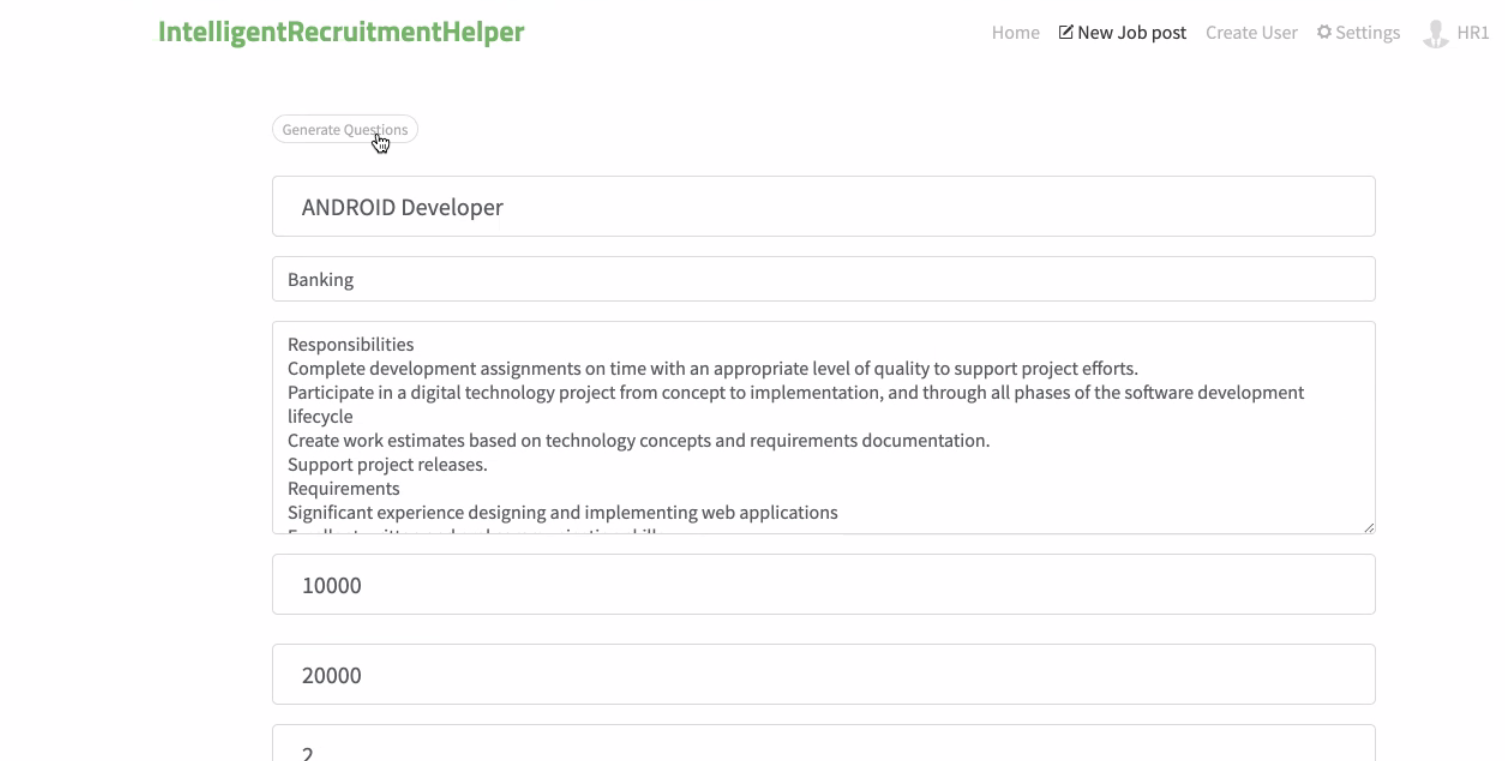


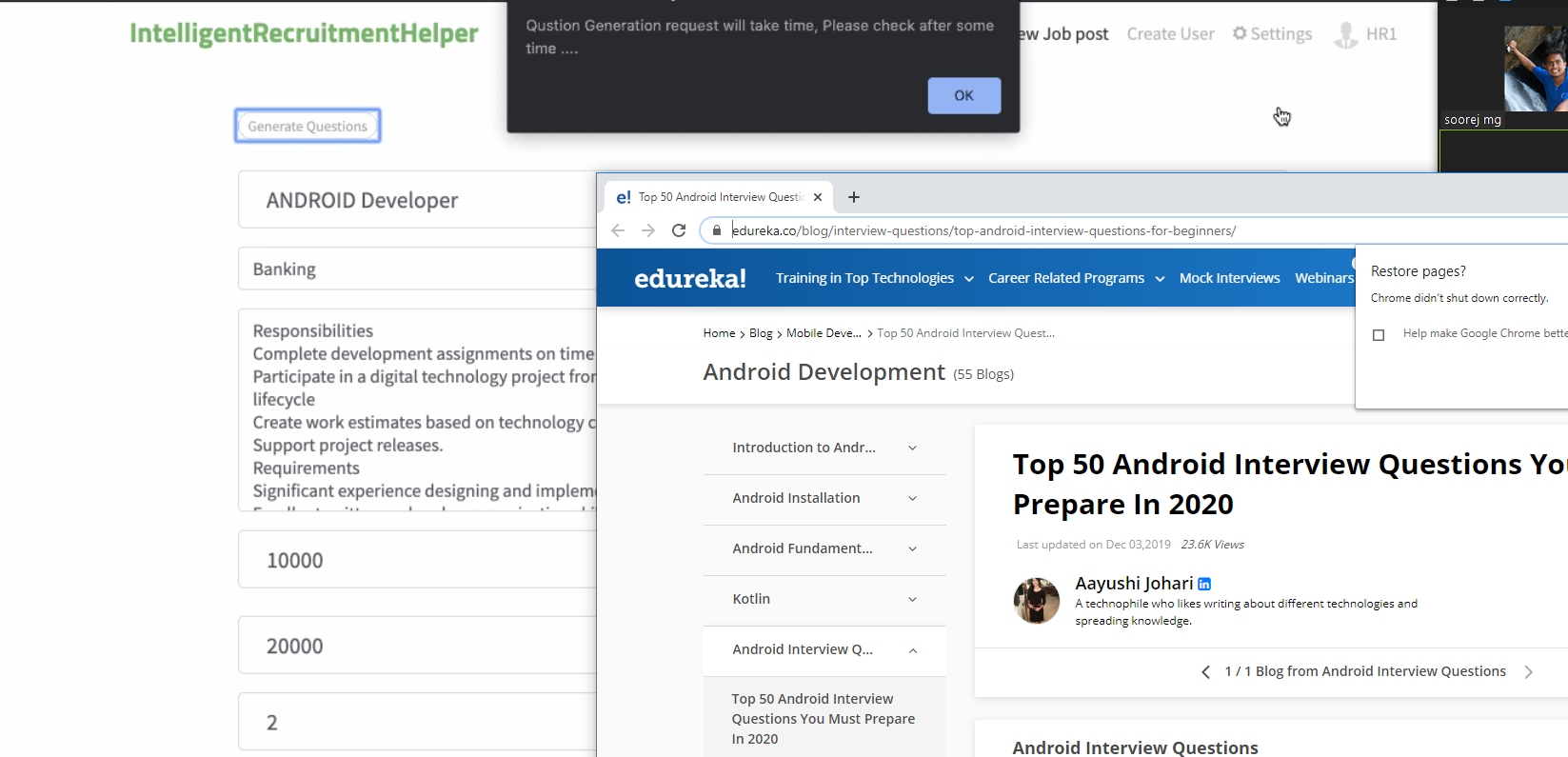
The GitHub repository has a ‘Miscellaneous’ folder and inside that there is a samplefolder.zip file containing resumes and job description inside. Two sample job description files have also been uploaded which need to be renamed as ‘jobdescription.docx’.

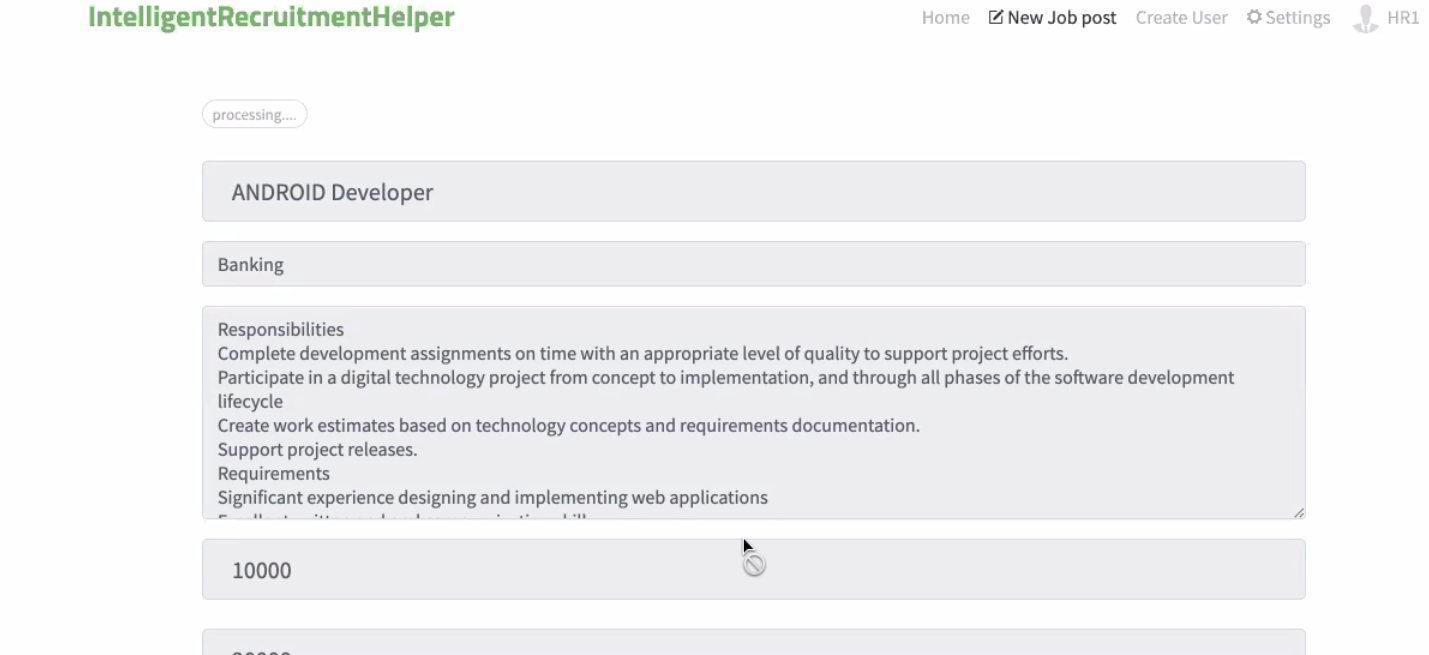
7. After uploading the resume, view the generated score for each resume. Screen below shows only a few as others are hidden due to scroll



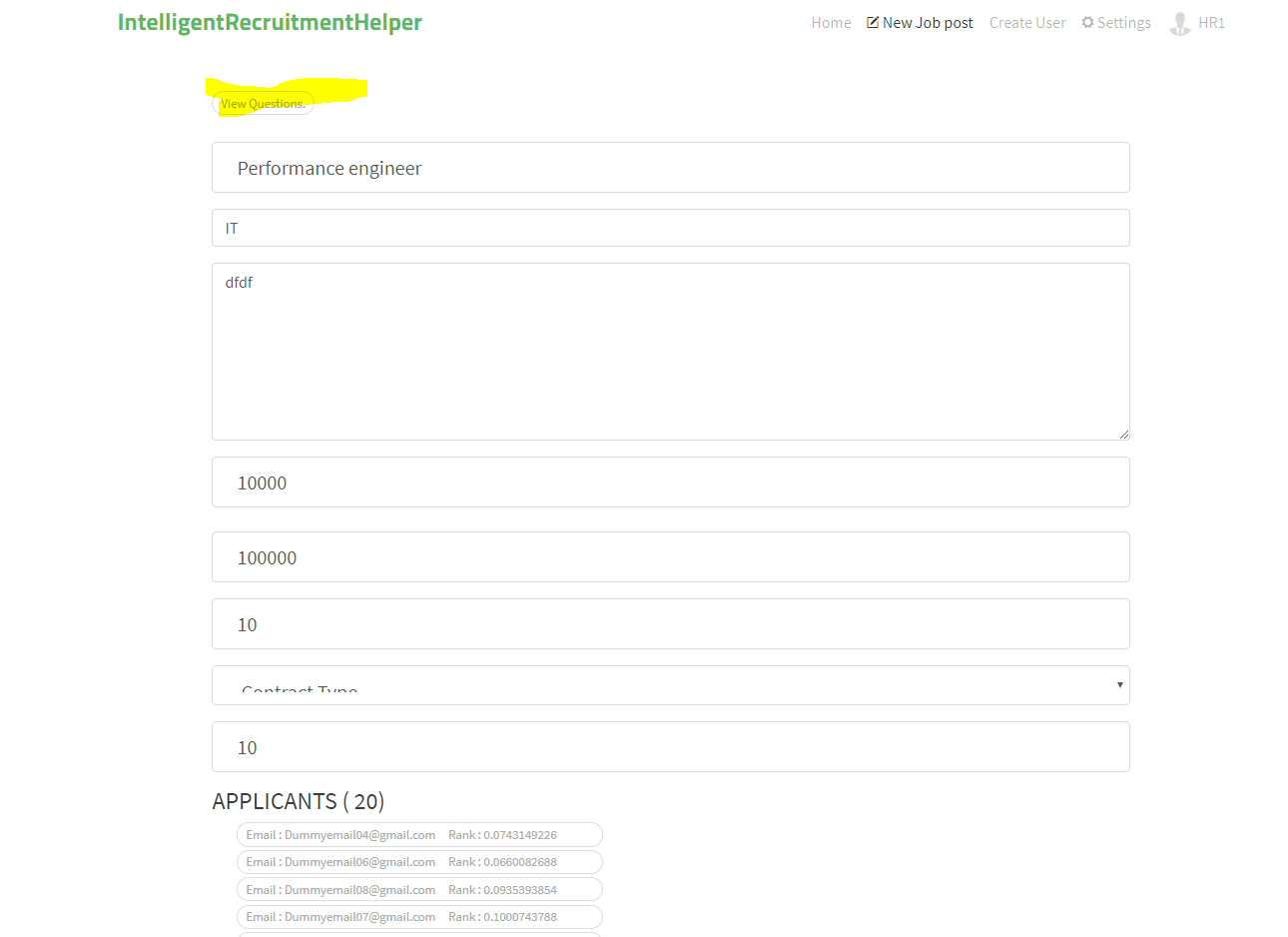
6. Click ‘*Generate Questions’* to generate questions. Wait for 3 minutes for questions to be fetched and prepared. The questions are pulled from multiple web sites which takes around 3 minutes

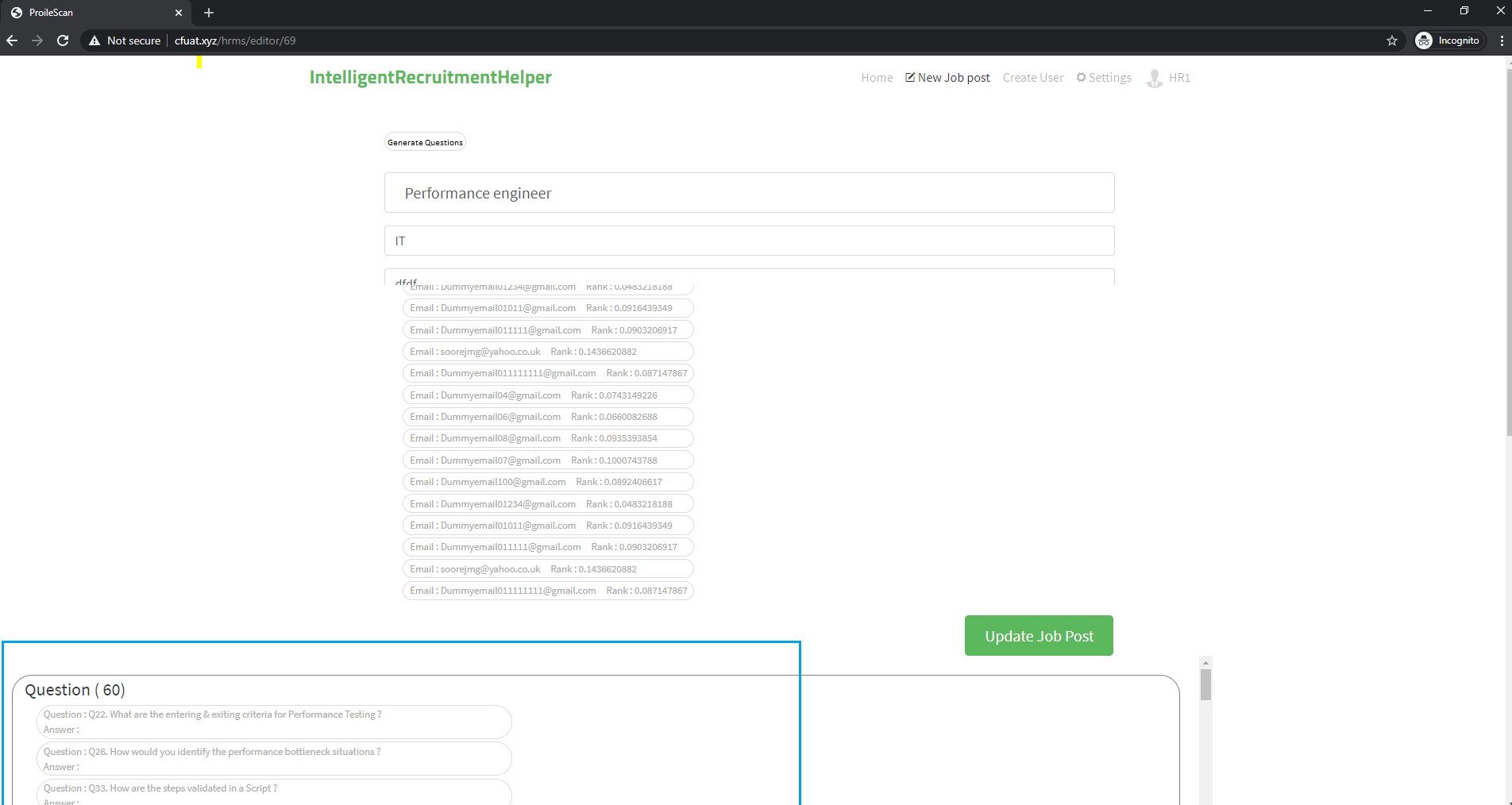






7. Click ‘*View Questions’*





## APPENDIX B: Individual Reports

**Rahul Jalan**

--------------

**Personal contribution to group Project**

Being only a 2 person group, each of us has been deeply involved in all parts of the project. Since I have wide experience in program management and other business roles, I have focused on initial ideation of identifying and articulating the market need, market research, business case and financial estimate creation, process mapping, product and launch plans, and where to apply RPA / IPA techniques etc. Apart from this I have also structured and formatted the whole project report to give it a professional look and feel.

**What have I learnt?**

Working on this project has been a great learning experience. Personally, I feel this module is one of the most practical and applicable from industry point of view, as demand for automation will keep on increasing in coming years. Especially for resource crunch countries like Singapore there is a lot of push and support from government to reduce reliance on manpower and bring in automation wherever possible. Also, the process mapping techniques have wide application in both personal and professional life. The business case creation template provided by one of our professors in IPA module is a very useful tool to quickly put a business case together and was used to create the business case for this project.

**How I can apply the knowledge and skills in other situations**

There is wide applicability of techniques and processes that we have learnt in this module and applied during working on this project. In this project we used TagUI for performing web searches and pulling in interview questions. The other tool UIPath is equally powerful and we have used in in our Capstone project for data scraping from websites. The process mapping techniques used in this project are also very generic and can be applied to a variety of situations to understand and improve upon the processes.

**Soorej Mohanadas Ganga**

----------------------------------

**Personal contribution to group Project**

Being only a 2 person group, each of us has been deeply involved in all parts of the project. Mainly worked on the resume scoring part and the question filtering modules in PY charm from scratch. Both of us worked on the Spring boot middleware and angular js modification to fit as a front end for our AI service requirement. Other than this, helped in getting raw inputs to report preparation.

**What have I learnt?**

Working on this project has been a great learning experience. Really understood the pain developers undergo while developing an application, which I never understood as a tester. It was totally a new learning experience and understood that all parts of the development life cycle is critical and has that value. Understood that more care should be taken towards the deployment part , error handling, exception handling etc while developing the app.

**How I can apply the knowledge and skills in other situations**

Tagui can be used in many automation tasks, especially in report generation integrating with python will be a combination to use.