Task 2:

Statement by: Balinder WALIA

Steps which I should perform if I am given the opportunity to work on this project:

Step 2.1: Planning: Japanese people say, “An hour’s planning saves a day”. Starting with the right plan and right tools for this job is very important. As it could be a large database system, I believe we need a highly scalable & flexible database system. We have to continue to scrap data and fetch as much information about the influencer profiles as much possible from **social networks** such as Facebook, LinkedIn, Twitter, Instagram, Youtube, Snapchat and Pinterest etc. However we have to start building a fresh quality database of profiles which is owned by Social Circle, for it to be the Market leader. The influencer profiles with their good qualities in a database can be really valuable.

Let’s call it **master profiles database**. It can be linked with the scrapped raw data from other social channels and we can merge (or link) it with our own master database on regular basis, update the master database with available attributes, qualities and their social weight as per the social interactions of a given profile for a regular period of time or during the campaign. We can work out the best algorithm for rating an influencer based on information available from different social platforms, the raw scrapped data and their qualities.

It should be worth keeping a reference database table for various profile social platforms such as their tweets and retweets. Their video posts on youtube and how many people viewed that video. We should store & analyse audience’s demographics, behaviour and location information if available from scrapped data available through Social APIs etc. My plan would be to first learn about the end-to-end Social Circle idea, the final product and the targeted audience.

Step 2.2: Capacity Planning: As this is highly scalable project and big data is involved the main profiles database can be a SQL based app. However the related data volume can be very high which is associated with a profile, so that should be stored in NoSQL big data platform - such as MongoDB.

Step 2.3: Development SVN: Setup Dev environment on git private repositories or similar private open source version control system such as Tortoise SVN etc. for consistent development approach.

Step 2.4: Development: Data gathering, stats and database design. Carefully designed database structure is key for this project’s success. Normalised or flat file type data tables have to be discussed in details and then created as per the final requirements, so that the performance of the system is not compromised whilst high quality & efficient database of profiles can be maintained.

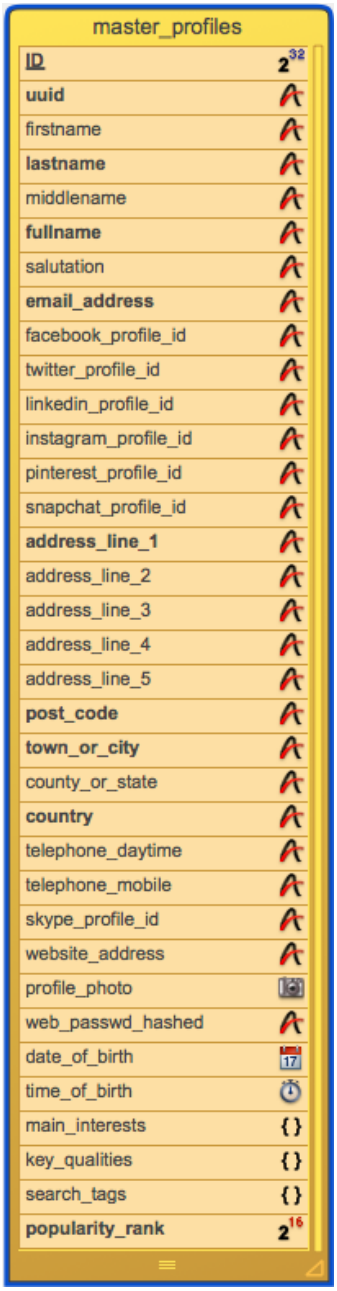
I think the biggest challenge in this project is going to be: Building a really clean database of high profile influencers in various marketing sectors in which SocialCircle and it’s customers are engaged in.

Building the master **profiles database**

Database name: **SocialCircle**
Database Table: **[master_profiles]**

Example fields for master profiles database table is shown in the left hand side illustration:

This database table is a just a quick representation of a profile in real world the database tables may be lot more normalised and social media profile ids may also be saved in different table example: **[profile_social_circles]**



Even addresses may be saved in separate database tables due to various types of addresses. Previous addresses. Home addresses. Work addresses. So different types of addresses can be all saved in different data table to make maintenance a lot easier. It depends on the project requirements and I can comment on database design in details once we have a detailed specifications in front of us.

Note: In the table shown above each profile is assigned universally unique **UUID - Universal unique identifier**.

Also we will require associated tables for storing raw data for each profile from various Social platforms such as Facebook, LinkedIn, Twitter, Instagram, Youtube, Snapchat and Pinterest.
The profiles database must allow a way of inviting famous influences to come and register and update their profile. We need to make sure the profiles are accurate and we have consent from the relevant parties to keep their data and use it.

There is a regular automated script I can imagine to download UK census data and import into Social Circle's [uk_census] data tables, to keep it in the local database cache for easy database queries and presenting on a dashboard along with influencer's profiles and stats. So we can present it to customers in neat, easy to use and visual ways and customers can then make informed decisions efficiently. So storing qualities of the influencers profiles is the key component of this project's database system.

Let's take example of my friend Dipna Anand. She is renowned Indian chef and co-owner of Indian restaurant in Southall. She posts plenty of Facebook and Twitter posts about her cooking skills, new recipes, corporate chef pop-up events on regular basis. She is also an Indian food cookery teacher for schools, colleges, and has degree in Hospitality management from University of West London.

Example qualities or tags we can save in a profile can be:

-Indian food

- Cook
- Graduate in Food technology
- Hospitality Management
- Indian cookery lessons
- Pop in Chef events for corporate businesses
- Chef
- Two Books published in Indian cookery
- Best Indian Cooking recipes

The popularity rank or trust factor algorithm can go through all of the above qualities like education, work experiences and other volunteering positions. Then it can also add extra points to overall ranking or trust factor for her social engagements, for example let's say she has

- 1 million views for the videos for her indian recipes on youtube
- 25K followers on twitter and 28K likes on Facebook, LinkedIn, Twitter, Instagram, Youtube, Snapchat and Pinterest etc.
- 20 tweets retweeted

Also worth having their online presence showing up on their profiles like their web pages or blogs. Save links to their website(s) and profiles for constant updates and checks for new tweets / like etc: **http://www.beyondbrilliantbook.com** & **http://dipna.com**

Dipna does a lot of pop-up chef events and she has a lot of followers. So based on her Facebook, LinkedIn, Twitter, Instagram, Youtube, Snapchat and Pinterest likes, videos and website information we can start to build master profile in Social Circle database and set a suitable rank based on her posts, tweets, followers, demographics and above qualities.

Lots of rice importers or indian spice manufacturers, resellers and wholesalers would like Dipna to be there brand ambassador. Here qualities can be used to promote their brands in her recipe youtube videos, Facebook likes and tweets.

So for above marketers the challenge is to find a good influencers and then audience who fit into their social marketing campaigns. So building a social demographic profile for Dipna would be: Scrap tweets, Facebook posts etc. where conversations include: indian food, cookery, chef, indian recipes, learn how to cook indian food etc.? Find the influencers profiles and then start to build their master profile.

Also their target should be no. of people of Indian and asian ethnicity, living in UK and or various targeted regions (Analysis through UK Census data). I would also interrogate if it is possible to get Google searches data, Youtube searches data and related conversations form other major search engines and social networking sites.

Automation

Setup automated backgrounder scheduled tasks to run a bot on regular basis to scrap influencers information where possible. In my opinion however a manual admin might also be required (to start with at least) to make sure the rank of a profile set automatically is a good quality rank. In certain cases system should be able to put the profile verification into manual reviews for certain profiles where system can't determine the veracity of their data. I need more thinking as I do not have the real data in front of me but once I have the data in front of me and I can analyse and see what we can do about veracity of the data.

Influencer Marketing Idea (My understanding so far)

Building a influencers database which is accurate can also be achieved through aggressive marketing, one idea is perhaps some kind of incentive for influencers to register directly with Social Circle Website.

So Social Circle will need it's own database of talented profiles (influencers) that is clear and Social Circle admin can verify information using best ways possible.

Imagine the Recruitment companies: Their biggest selling point is a really good database of talented candidates. I think once that's achieved other things should fall into place itself. Social Circle is a great idea for Marketing and if done right, this system can be a game changing technology for Marketing.

I would like to also emphasise on the importance of a connection between matching profiles and given customer for their marketing engagement.

So a good reviewing system based on the profile rating and previous engagements can also be a good to get quality results. Like LinkedIn profile of a candidate shows his/her recommendations. If Social Circle system could also show a profile's previous recommendations for their great work it can also provide a good incentive for profiles to register with Social Circle.

I think, In order to make sure the database is accurate and up-to-date we will have to find a way to invite influences themselves to update the information. It has to be really easy to use database for influencers as well. Perhaps each influences can earn a point or a reward to keep their database and profile qualities up-to-date.

Profile weight or [popularity_rank]

A ranking value is an important factor to store for each profile as when the customer searches for the matching profiles; Social Circle have to make sure it delivers quality high ranking profiles on top in the search results. This database field should be updated as soon profile related new information is available. Influencers rank or weight is probably one of the most important and interesting in the decision making process for customers who are looking to engage influencer in their next Marketing assignment.

Name and or Face recognition software techniques is a great idea. We will also need a way to authenticate users by their email address for example, or OTPs (One time password or codes sent to their mobiles to verify their mobile no. and identity) and on smart phones we can also use the finger prints match and verify their identities. Also validating their identity via their **bank account and address verification** is a possibility - it means we can make sure we have the right identity for the right candidate as financial information is very likely to be associated with that person.

As it is an online app, caution should be taken as there are a lot of bogus profiles on Twitter, Facebook, LinkedIn and there is no way to verify the identity of the bogus profiles. For example in Bollywood their are a lot of actors and actresses where it is hard to identify whose profile is the real one. Perhaps with financial and address verification we can make sure the identity of the profile is verified in as many ways as possible. On regular basis we should let users change their passwords for their accounts again to make sure system is secure in as many ways as possible.

Address verification can be done in different ways. We send a automated letter to the influencer to his/her address with a special encoded code in it which they enter online to verify their address. For each of the verification step we can add points to their profile, let's say field name is [profile_completion_percentage] and if they perform all the min. required steps to verify their profile then we make their profile 100% complete.

Like Recruitment companies must check many things for a given candidate, the work eligibility or DBS checks for certain type of jobs [optional but mandatory step for certain jobs]. I think we need similar system which includes various predefined steps to verify profiles, automated as many as we can.

Customers database table

Customers table will allow customers to register possibly with different levels of access to the Social Circle. Social Circle can also have contacts table related to one customers table in one-to-many relationship. Different contacts / users from same organisation can register and login to the system and regenerate quotes and run reports and run stats on their campaigns and marketing activities. We also have to store customer profiles and qualities and their interests so we can present them with matching profiles.

I have lot of experience working in Recruitment industry, while working with Jobshout project I created CV search to match candidates to Job ADs using database search tags. That experience will come handy.

I have also implemented postcode or location based radius search for many job board websites to match candidates with the right jobs ADs, An example: **http://www.cvscreen.co.uk**/
It may be useful in certain cases to find location based local talent if client wants to meet them and prefer local talented profile only. For example we may provide radius search to find nearest resources and visa versa.

I would think about Implementing data archiving and data mining techniques to make sure the system delivers high quality results. Also to do house keeping and preventive database maintenance we should implement Data Archiving and House keeping in this system.

Customer strategy, plans, quotes and estimates:

I would like to create a really easy to use and flexible UI to allow customers to search the talent they are after, and enter their marketing budget for a given range in time and then system outputs the estimates in real time. For this to work we will want to save a rate cards for each profile. Their rate. For example in Recruitment Software systems we save Rate cards. If it is a time based assignment system can calculate total time a given talent is being engaged for and calculate costs based on their rate card.

Another interesting question is do we need to produce estimates for fixed-cost based customer assignments? In which case Social Circle system should save preferred ways of pricing the assignments to produce profitable estimates - for all the parties involved talents, customers and Social Circle itself. We could also think about creating like a products and packages database with some pricing information.

Pricing is another question I am thinking about. How to create a good price matrix database to make sure Social Circle and profitable and successful venture.

We may have to store customers activity in database, so we can analyse their steps, as it helps to build a clear picture for Social Circle what customers want and what kind of talent they are looking for? A good match finding search would be really useful which matches the talent from the pool of Social Circle influencers database using the tagged searches and ranking algorithms. System can be written to produce good quality matches, based on their budget and product list of influencers showing their existing achievements, profile and sorted by their popularity rank. We can store profiles **popularity_rank [0-10]** for each influencer to sort the data presented for the customers. Also a way of seeing active profiles is necessary for customers to save time.

Eventually for customers benefits we need a way to quickly determine whether the influencer's audience is right for their brand. They'll see interesting data such as age, income, location, and profession of followers.

Step 2.5: Development: Design patterns - create open source libraries and set of tools which are going to used for this job:

As a Full stack developer I would start with:

Redhat (Centos, Ubuntu Stack)
PHP
JavaScript
jQuery
Angular
MySQL
Node.JS Express

NoSQL MongoDB (for big data)
NGINX HTTP server
Charts libraries such **plot.ly**
Security Layer for web application - referring to **https://www.owasp.org/index.php/Main_Page**
possibly **4D.com** for certain complex technical analysis

Frameworks such as Laravel or Symphony could be also good for certain parts of the application. Subject to final evaluation, once I have full detailed spec I can do detailed technical analysis.

Setup a **Development** environment, **Staging** environment and Auto Scaling Deployment tools for Continues Development and Deployment. Specially when team grows we will need a solid base to get high quality software out to the Market quickly. We must look at using Configuration and Deployment automation tools to save time. Also using CD (Continuous Development) and CI (Continuous Integration) tools can make the process a lot more efficient, consistent and we can get quality software out of the door quickly. Some of the well known CD, CI & Deployment and config tools are **Jenkins**, **Ansible, Docker etc**.

Make use of databases such as SQLite for embedded preference and apps into desktop environment for saving preferences relevant to each developer and testing staging machine and preferences which can be pushed from dev to live if need arises. Implement Unit testing using Assert commands in the core of the project. Use Agile methodologies and Scrum to implement bug tracking software and perhaps make use of Paired programming in certain complex tasks. All this are important while preparing the grounds and in-future for easy-to-manage source code, development processes, and system maintenance. Must avoid developer bad habit which leads to **Software Entropy**.

Step 2.6: Hosting : We may have to think about this at some later stage but it is important to know how to setup a auto scaling environment and way of auto-spinning new servers on demand. The profiles registering and searching done all the time the Stats database can grow quickly. A really robust database replication is required for profiles and associated database. NoSQL such MongoDB offers great replication and auto scaling functionality

Step 2.7: Backups : We may have to think about this at some later stage but it is important to know how to setup a auto scaling environment and way of auto-spinning new servers on demand. And a really good back up system. Daily backups of the main database sets. Weekly backups and Monthly back ups are a must.

Step 2.8: Rest APIs : Well thoughtful REST APIs can save a lot of time. We have to use REST services where ever possible due to the performance benefits REST can offer and it's flexibly, global adoption and ease of implementation. Socket programming is also helpful to make some of the complex tasks more efficient and manageable.

Step 2.9: Time scale : Timescales to create, test, and deploy a profile verification system can take from 3-6 months and above mentioned steps including if not done already in the existing system then we should be looking at least 9 - 12 months of work. ETA also depends on how many people are working on this project and many other factors but this is a rough estimate based on my understanding so far. A lot of it depends on the APIs we use for various social platforms and how they work. Technically creating our own influencer profiles database system is easy and should not take long. It is the third party attributes and data which can be time consuming and tricky. Again I would like **not** to speculate too much and once I have clear picture in front of me how system should be implemented with full list of requirements, then I can workout more realistic estimates.

Step 2.10: R&D: I believe in Research and Development. In order to find out the best way to validate the accuracy of the data; I would normally divide the various tasks in different steps and then try out in a test environment. Depending on the outcome of those test, I can find out the best

available methods to work with for that given task. We also can also create micro-services based architecture of the entire system to manage complex tasks.