



**Defining and Attempting to Automate the Data Collection Process
through Online Platform**

Thesis

Khaza Towfik Akbar

Master in Business Administration, May 2019

International Business Management

Acknowledgement

This work would not be possible without the constant support of Lionbridge Oy in Tampere, Finland. I am especially indebted to Dr. Petteri Vilén, who have been supportive to my idea and projects. I am extremely grateful to the team who worked actively to provide me with all types of support I needed in terms of collecting data to execute any plan.

I am grateful to all those people whom I have been working for the last one year in this project. I would like to thank all the members of my thesis who constantly supported me with grate comments on improving my thesis. I would especially like to thank Jari Peräkylä, the advisor of my thesis. As an instructor and a mentor, he has been supporting me with great comments and encouragements.

Last, but not the least, I would like to mention how much I have learned working with all my peers while doing all the courses. I would like to thank everyone for letting me grow with you and enrich myself from your knowledge and experiences that you have shared with me during the program.

Abstract

Tampereen ammattikorkeakoulu
Master's Degree Programme in International Business Management

AKBAR, KHAZA TOWFIK

Defining and Attempting to Automate the Data Collection Process through Online Platform

Master's thesis 49 pages, May 2019

The Lionbridge Oy Tampere local branch has been serving many Information Technology (IT) based firms in their projects. Translating and recording the documents into many different languages is one of the services that the firm has been providing. The current model of collecting data (recordings) has been a lengthy process and relatively costly both in monetary and opportunity costs.

This study has focused on the pros and cons of the current model that needed an improvement in the data collection strategy. The study aims at developing an online platform where all the stakeholders of the data collection process can work at the same time in the same platform. It has been designed to provide a safer, low cost, quick and efficient model of data collection that also focuses on the overall development of young generation and female workforce all over the world.

The research data was collected using an online survey sent to the clients and the participants who recorded the documents provided by the client. The firm also has been running a pilot project from where the study extracts the outcome of the project. The whole project has been developed with the help of a lean model and the project management tools. It employed the tools to develop, enrich and accelerate the model while taking care of the ethical concerns, safety measures and the cost and benefits of all the parties involved.

The main outcome of the study was that it provided a safer and user friendly module to collect, process and deliver the data to the clients. The proposed model has been found more efficient than the current model as it provides the data with a minimum costs and time used. The pilot project has been successful in proving the better quality products to the clients.

The results suggest that the pilot project should still be improved in many areas including research and development.

Key words: language, translation, data collection, online platform, recordings, youth, gender, project management, lean model, automation.

Table of Contents

i. Introduction	8
ii. Methodology	12
iii. Discussion	16
3.1. Current model	17
3.1.1. <i>Getting the orders from the buyers</i>	20
3.1.2. <i>Contacting the external party to do the job</i>	22
3.1.3. <i>Third party organize the task</i>	24
3.1.4. <i>Advertisement by recruiter (organization)</i>	25
3.1.5. <i>Sign Up participants in the third party website</i>	26
3.1.6. <i>Screening by recruiter</i>	28
3.1.7. <i>Time Booking of participants and checked by moderator</i>	30
3.1.9. <i>Support for participant while recording</i>	32
3.1.10. <i>Payment by the recruiter</i>	33
3.1.11. <i>Check and review the recordings</i>	33
3.1.12. <i>Deliver the product to the buyer</i>	34
3.2. Proposed Model.....	35
3.2.1. <i>Problems in current model</i>	35
3.2.2. <i>Solution</i>	36
3.2.3. <i>Key Matrices</i>	39
3.2.4. <i>Cost Structure</i>	42
3.3. Pilot project	42
3.3.1. <i>Project Integration Management</i>	42
3.3.2. <i>Project Scope Management</i>	44
3.3.3. <i>Project Time Management</i>	47
3.3.4. <i>Project Cost Management</i>	48
3.3.5. <i>Project Quality Management</i>	48
iv. Conclusion	50
References	52

List of Tables

Table 1: Age group and the rate of their participation	18
Table 2: Time allocation and the number of takes per participants.....	31
Table 3: Percentage of good quality recordings	33
Table 4: Features of the proposed model	37

List of Figures

Figure 1: Lean canvas..... 12

Figure 2: Lean Canvas Life Cycle 14

Figure 3: Block diagram for the test process 14

Figure 4: Automation Testing Process..... 15

Figure 5: Development and operation environment- DevOps..... 15

Figure 6: The respective steps..... 16

Figure 7: The key metrics for current model. 20

Figure 8: Enterprise Artificial Intelligence Market Revenue Worldwide (million U.S dollars) 21

Figure 9: Time required to deliver the task to third party 24

Figure 10: Recordings in each 1000 Days..... 25

Figure 11: Number of participants signed up for the tasks..... 27

Figure 12: Number of male and female participants signed up for the tasks 28

Figure 13: Percentage of Total Participants Recruited for a Project 29

Figure 14: The Key Matrices of the New Model 40

Figure 15: Project Integration Management Summary 43

Figure 16: Planning process for IT project 44

Figure 17: Sample Intranet Design 45

Figure 18: Mind Map for Ethical Project Management 46

Figure 19: Project Cost Management Summary 48

Figure 20: Project Quality Management system..... 49

i. Introduction

Online platforms has become a major contribution in business and research in modern world. Resent research on online platforms for businesses discusses about their multi-sidedness. Theoretical and empirical advances in technology brings the world in a small platform through online data collections and their uses in dynamic studies. An economically strong online platform for big data analysis and research is a critical need for any businesses. It model developed to comply with competition policy is needed to cope with recent business challenges (Just 2018).

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention. My projects identifies the issues in the current model of data collection and propose an effective model using Lean start-up model.

The business that I am focusing in the project is based on information technology. This is a multinational software and professional services company. It specializes on products focusing on language translation, localization, software testing, e-learning translation, internationalization, interpretation, software development & testing, and content development services. The project I am focusing on is related to the language translation. Hence, the first and very important step of this project is data collection. The current model involves manual ways to collect, clean, test and store the data. The team also use manual methods to sell the data to

prospective buyers. In this project, I propose an automated online platform that ease the procedure of language translation and selling the products.

Online platforms are now dealing with new market conditions that are different from the markets for other perishable or non-perishable goods. This entails a shift through a systematic inclusions of price and non-price factors in the analysis. Just (2018) discusses about the classical models of online platforms and how it should implement new strategies to make it an asset in recent digital economies. Just (2018) argues the importance of non-price competitions in the models, where the classical models only focusses on the price oriented competitions in the big-data platforms.

In the online platform, the mostly discussed and controversial fact is the externalities and welfare loss. This comes from information asymmetries arises from both the parties involved. In this model, we particularly focused on eliminating these facts and tried to make the platform as transparent as possible so that both the parties can gain keeping the business effectively strong. The project is motivated by a study by Yoshimoto and Nakabayashi (2019) who analyze the welfare structure of an online platform and look at buyer- seller matching. They have studied on the transaction cost of both the parties and analyzed the effect of the transaction costs. They have pointed on the fact that there are heterogeneous buyers and for the train tickets in the online platform that makes a significant deadweight loss in the market. In contrast to their study, the model designed for my study does not entail any deadweight loss, meaning that both the buyer and sellers in the platform has full information about the prices and products.

Contract programming can be a useful tool to experience and disseminate up-to-date knowledge (Pocock, 1986). Pocock (1986) discusses about freelancing opportunities in the time when the online freelancing were not even known to people. While studies discussed about the online platforms for manufactured products and services, online platforms for big data collection and analyzing has become a recent demand in business analysis. Studies (Wei, Lu, & Zhao, 2019) have discussed how the online platform for selling manufactured goods would be effective. Needham et al. (2019) discusses about the implementation of online platform for services like nursing. Hence, the major contribution of my project is to model an online platform to effectively collect, store and disseminate the data.

The proposed model is not only focusing into the effectiveness of the business, but also concerns about the social issues. It encourages the employment and entrepreneurship not only in the country of the business, but also around the world by opening the opportunity or working online. Study has shown how employment opportunities have increased by the freelancing opportunities in IT sector around the world. The flexibility in employment has also increased because of this opportunity. They discussed the matter of fact that employees may not be fully committed to the employer when it comes to freelancing. With a survey analysis, Sub et. Al. (2010) identifies that there is a significant effectiveness in commitment among workers in the IT sector.

Sultana et al. (2019) analyzed about the entrepreneurial behavior of the workers in IT freelancing. They found that the skill level and the entrepreneurship increases due to be engaged in freelancing activities in IT sector. Damian and Manea

(2019) explored the fin-tech freelancers and the market strategies of the business. They have also found the entrepreneurial behavior among the freelancers in the business.

Shevchuk, Strebkov and Davis (2015) conducted a study on the satisfaction level of workers based on Russian-language internet freelancing. They found that the satisfaction level is different for male and female workers. Moreover the socio-economics outcomes for both the groups depends on the educational background. Shevchuk et al. (2015) argues that the freelancing opportunism encourages gendered career strategies and greater family demands for women in the economy.

Having analyzed all the debates about online platform for the business, the objective of the study is to propose an online platform that ease the business in every possible ways discussed. My model discusses the complete procedure to collect different types of data for different projects, for different clients. To build the model, Lean start-up model has been used. The model breaks up every step piece by piece and provides the idea to improve each step, which would finally improve the whole output of the data collection process.

There are few steps what the model follows for the whole process. The whole procedure is very long, and contributed by a lot of people, many third-party software, websites, and takes a lot of time to collect and process the data. Therefore, the main objective of the proposed model is to get rid of the third-party software and websites as much as possible, lessen the human involvement and time. The model expects to save money and time, which is the basic requirement to implement the new model to have a successful business.

The study has been broken up into three major parts. In the first part, it discusses the current issues and debates about the concern of the study and explains the importance of the proposed model. In the second section, it discusses about the methodology adopted for the study. In the third chapter, along with the discussion of current and proposed model, it analyzes the data collected to evaluate the proposed model.

ii. Methodology

The study adopts a Lean model to develop and test the hypothesis of the study. Lean canvas is one page business plan that deduce raw idea into key concepts. It is adapted from Alex Osterwalder's Business Model Canvas and optimized for Lean start-ups. It replaces elaborate business plans with a single page business model. Most of the business plans take too much time to write, but this lean canvas business model takes less time to write and contain all the value able information related to the business. The main reasons of adopting Lean model to construct the proposed model is that the lean canvas is fast, portable, concise and effective. It is a brilliant tool for a start-up (Miller). Lean model allows the model tested from the very beginning of developing the model. It saves the time and allows to modify the model during the process.

Figure 1: Lean canvas



Source: <https://leanstack.com>

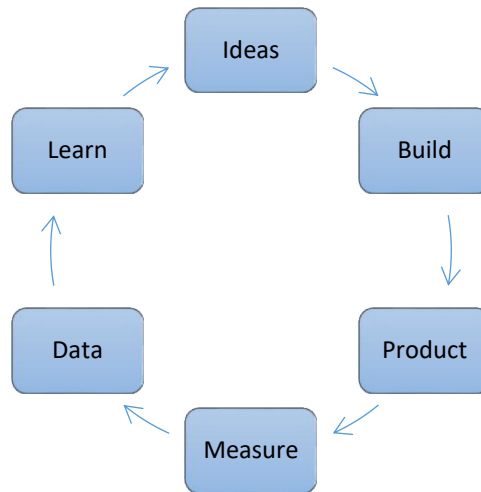
Lean Service Creation (LSC) is a combination of lean principles, design thinking and agile philosophy in a form of different tools like canvases to create successful services and products. LSC helps to reach business objectives in an iterative and human-centric way. Canvases aim to create concrete results; to facilitate co-creation, give and receive feedback, to experiment, try out and prototype.

Canvases are used for example by experienced service designers to remind them of business questions and technical issues, by business people to operationalize a customer-centric core into their work and start-ups to communicate their ideas and to see where to go next. What is needed to get started is basic knowledge of product or service development as well as creative and result-oriented mindset.

Full life cycle of lean canvas is demonstrated in figure-2. The canvas starts with the idea of the model. This is the idea of the model, which in my study is automation the data collection procedure. The second step is to build the model and then make it a product for the business. When the product is ready, it is measured with different

metrics that generates data. Using those data, the model is improved by modifying it. Hence the loop works until the product is finally ready to start-up the business.

Figure 2: Lean Canvas Life Cycle



Source: Nidagundi and Novickis, 2017

Test process involves the complete testing of the model developed in the study. To test all possible amalgamation including manual testing as well as the automation one, the team has to test thousands of combinations. The model incorporates the possibility of having a lot of tests and also accommodates the feasibility of running the tests to deliver a complete product in the market.

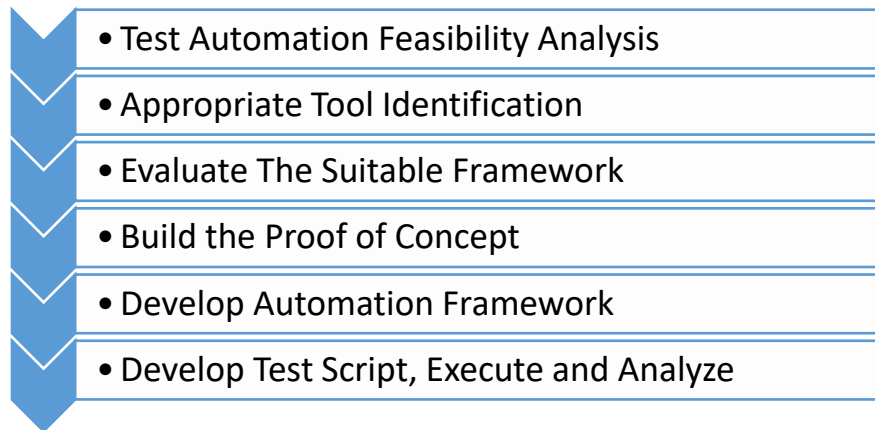
Figure 3: Block diagram for the test process



Source: Nidagundi and Novickis, 2017

The study adopts the automation testing process to speed up the process. This is the most important part of the model that is the main contribution of the project.

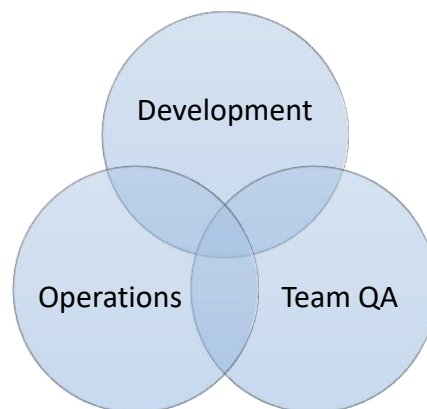
Figure 4: Automation Testing Process



Source: Nidagundi and Novickis, 2017

The model also develops the DevOps team. This term is introduced at the Agile conference held in 2008. This team focusses on the particular project and work together as a team to focus on the collaboration of the development, operation and QA team. The methodology of the model also adopt the DevOps to develop the effective tool.

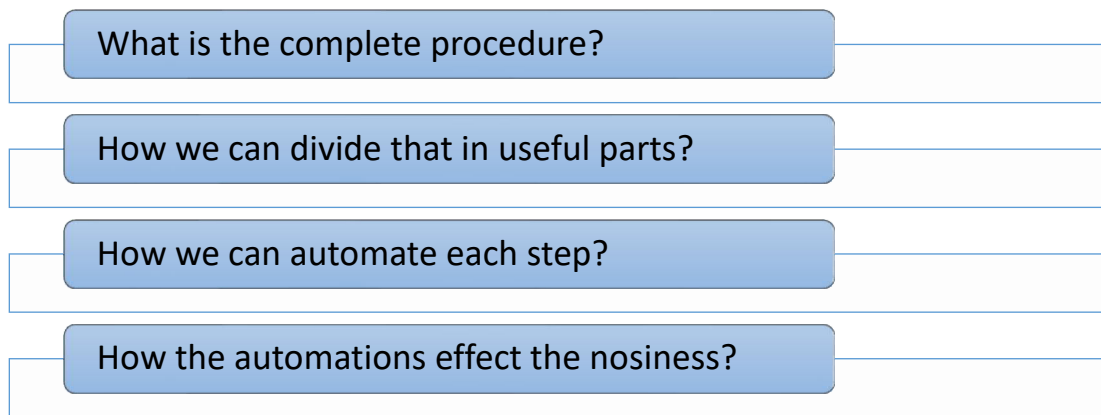
Figure 5: Development and operation environment- DevOps



Source: Nidagundi and Novickis, 2017

To automate the current model, first, we learn the complete procedure of data collection and break it piece by piece. And try to automate those pieces separately and connect them. The respective steps are as follows.

Figure 6: The respective steps



Using all the tools discussed above, the study is conducted as follows. In the first step, we analyze the current model and the pros and cons associated with the model. In the second part, we develop the tool and do the feasibility analysis. The third step involved the test analysis and the outcomes of the model.

iii. Discussion

The study discusses the analysis into three parts. In the first part, it discusses about the current model, its process, cost structure and problems. In the second part, it discusses about the proposed model and the structural change the study proposes for the model. It also discussed the opportunities that is created by the proposed model and the cost benefit analysis. In the final part, the study discusses about the project

management technique that the proposed model suggests for the model so that the model runs very efficiently.

3.1. Current model

The business that the study is concerned with involves with machine learning, languages and artificial intelligence (AI). The business is an American based Information Technology (IT) organization that deals with several projects coming from other information technology, software development firms. The project that I am dealing with in this study involves translating the documents into several different languages. The documents are sent by the clients to our firm and our firm manages the process of translating with the current model discussed in the following. Mainly the data collection process is the focus of the study and I analyze the pros and cons of the current model and have proposed an efficient and less costly model to update the format.

Currently, the data collection process involves several steps and people in the whole process. This involvement of several steps, which are connected to each other, makes the current model lengthy and time consuming. In the line of the business, the organization work with people from different countries. They are of different age groups who speak different languages. The translator are from different countries and their age is ranged from 20 to 62. It is interesting that the participation from the age category 20 to 40 is very high compared to older generation. However, it is also an interesting fact that the participation from the age category who are aged more than 51 is also a significant amount. It says that, people who are on the verge of their retirement are also

interested in this type of freelancing jobs. It helps in extra earning and improves skill level.

The age groups and their contribution in this translating freelancing is showed in Table 1. The percentage of the participation is almost similar for age group 20 to 30 and the age group 31 to 40, with a slightly higher percentage for the age category 21 to 30. The younger generation has become very skilled in freelancing jobs and they are also very eager to earn some extra money working on their free times. Moreover the freelancing opportunity helps in building their experience in the skill level that helps in the job market. Moreover, the study also looks at the participation from different countries and found a significant percentage of participation from the developing countries.

The main task that the organization do for the projects is to translate the documents and record it in many languages. The organization gets the order from other businesses who requires their documents to be translated to their required languages. The buyers of the programs ask the product to be recorded by the people of specific age groups.

Table 1: Age group and the rate of their participation

Age Group	Percentage of Interpreter
Less than 10	8.74
11 to 20	9.63
21 to 30	27.83
31 to 40	26.73

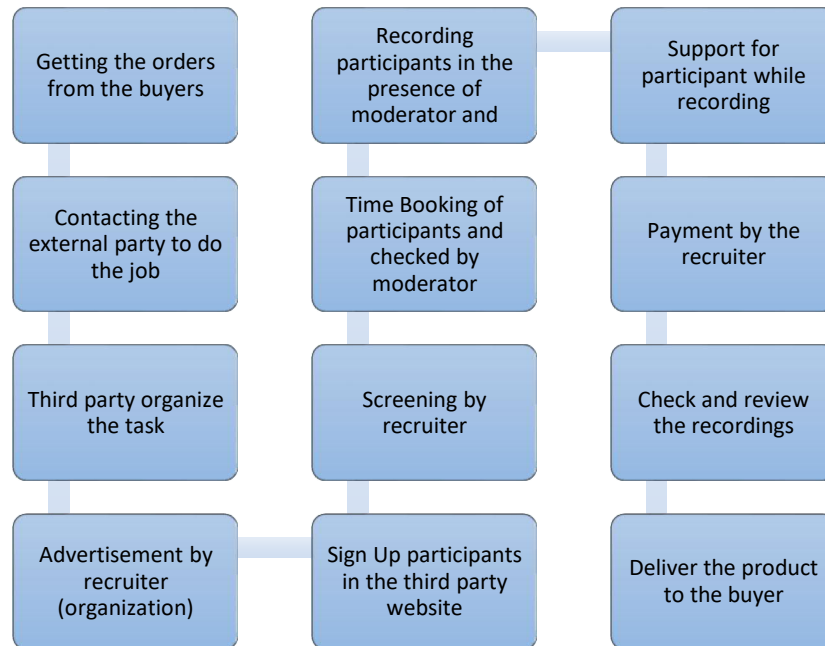
41 to 50	15.72
more than 51	11.09

When the organization gets the order, they find some external party to do the task. The third party then conduct the process and collect the data. The data collected by the third party is given to the team assigned by the organization to review and check the quality. Then the team puts the file to the folder for the buyers. The whole process is time consuming as there is third party involved in the whole process. The process of data collection has been described in the key matrices in figure 7.

From getting the orders from the buyers to deliver the product, there are 12 steps of data collection and processing steps involved. After getting the order from the buyer the organization contacts the external party to do the job. Once the third party gets the task from the organization, they organize the task and process it to circulate to the prospective interpreters. To do this job, the third party mostly circulate an advertisement about the task.

Once the participants see the advertisement and matches with the requirement, they can sign up for the job. When the sign up process is done, the recruiter starts screening and books the participant for the task. The moderator also checks the whole process. The participants then can start recording their voices in the presence of moderator. There is a 24 hour support team who moderates the whole recording process. Once the recording is done, the payment is processed by the recruiter. In the final step, the moderator and reviewers review the recordings and deliver the product to the buyer. The whole process is explained in the following section in details.

Figure 7: The key metrics for current model.



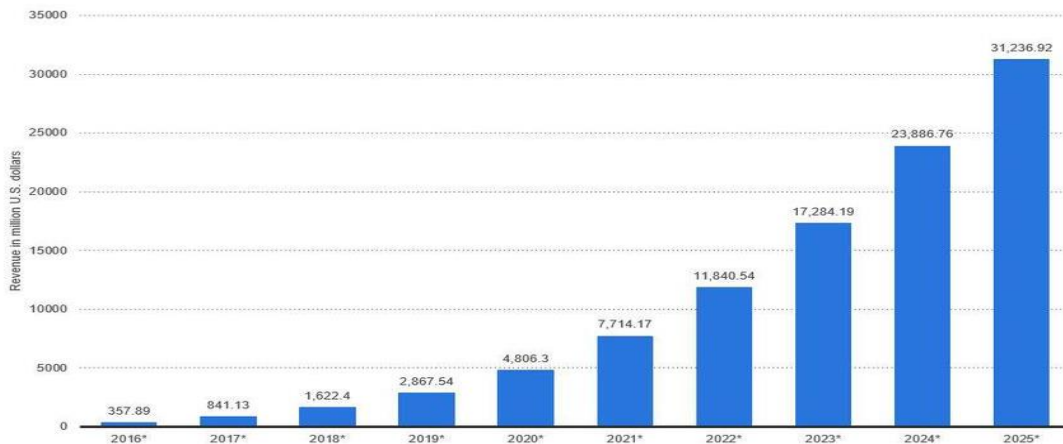
3.1.1. Getting the orders from the buyers

The information technology and software development firms require many support from other similar information technology firms. Our organization is such kind of a firm that provides supports to other IT based firms. From developing artificial intelligence to making a software, it requires a vast majority of the data. The one example is Siri from Apple products. Siri talks like a human and interacts with people in voices. It can talk in several tones both in male and female. Hence the programming of Siri required someone to record all the words and sentences that Siri can or will say while interacting with people. Someone had to give their voices that is currently the voice of Siri. Similarly there are a lot of such software's and applications that require voice recordings in different languages. If apple wants Siri to talk to in Spanish or French, Apple would require a vast amount of data to store in building the application.

Our organization deals with such supports that such other organization would require. The world is becoming very demanding in artificial intelligence. Robots now talk to people in different languages and we can hope that, world will develop such supports in long run to build its applications in artificial intelligence.

Current data shows that, in 2019 the market revenue from artificial intelligence is very high worldwide (figure 8), which is going to reach in its peak in next 5 years. The demand is projected to increase at incremental rate in next few years. Hence the demand for translating the speech into different language would be very demanding for those applications. Current and future expected demand for artificial intelligence is shown in the figure 8. This is the demand that the world would expected to face in next few years. The model is then very relevant in recent and future times. Making the best way to deliver the product (translating and recordings) has become a prior demand in recent relevance. Our organization is concerned with such demand and process the tasks that other firms demand.

Figure 8: Enterprise Artificial Intelligence Market Revenue Worldwide (million U.S dollars)



Source: Forbes.com

The organization gets the order from the buyer through different projects. Clients ask the organization to conduct the projects. Each project requires the provided documents to be translated and recorded in specific languages. Clients choose the languages and the documents to be translated. The organization has a dedicated team to keep updated with the client's requirements. The client may give several documents at a time or one or two documents in some time intervals. It cannot be estimated or projected that when the clients would give the documents. Hence, the project team has to be very efficient and devoted to the project requirements so that they can contact the third party at their earliest convenience.

3.1.2. Contacting the external party to do the job

Our firm does not directly conduct the job that the clients ordered. We employ a third party to manage and process the task for us. As there is another party involved in this whole process of data collection, cleaning and processing, it requires one team in our firm who handles all the jobs to be delivered to the third party and off course, there

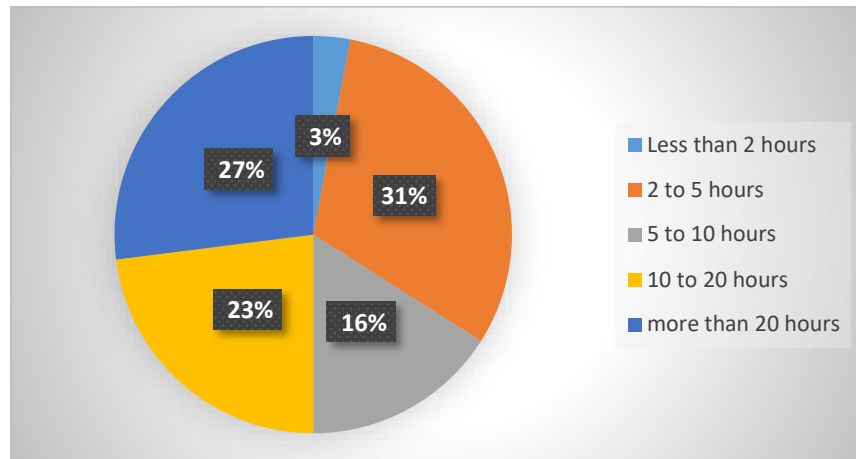
should be another management team in the third party's concern who would only work for our projects. In this way both the team can easily track the tasks and the quality is assured. We have a small team, consisted with 3-4 members in our firm who constantly monitors the whole process. In the third party's concern, there is no one fixed with only our firm. There are several employees in the third party's concern who deals with all translating jobs to be delivered to the participants. Hence, it is sometimes cumbersome and time consuming to explain things to the third party as there is none fixed with only our tasks. Our team has to explain same things to different people working at the third party. It consumes time and effort.

The reason that our firm employs a third party to do such job on our behalf is convenience. As the task involved continuous monitoring and processing, it would require a lot of employees who can constantly monitor and process the recordings if it was done in our firm. It would be more costly both in monetary units and opportunity costs. As our firm was initially very small both in market share and scope, it started this projects employing the third party in the year of 2009. However, at this time, our firm has become sufficiently bigger to afford such projects to be done by ourselves. Moreover, it has become renowned for its products and the demand for the products are increasing at increasing rate.

However, when the team gets the orders from the client, they contact the third party about the requirements. The average time recorded to deliver the work to the third party after receiving the task from the client is 14 hours. Figure 9 shows the time recorded to deliver the task to third party. It shows that, most of the tasks (31 percent) are delivered within 2 to 5 hours. The percentage of tasks delivered in less than 2 hours

are very small (3 percent). There are a large portion of the work that is delivered in more than 20 hours. The reasons for this delay include delay in detecting the project, processing time, casual or sick leave of any member in the team, weekends, third party availability and many more.

Figure 9: Time required to deliver the task to third party

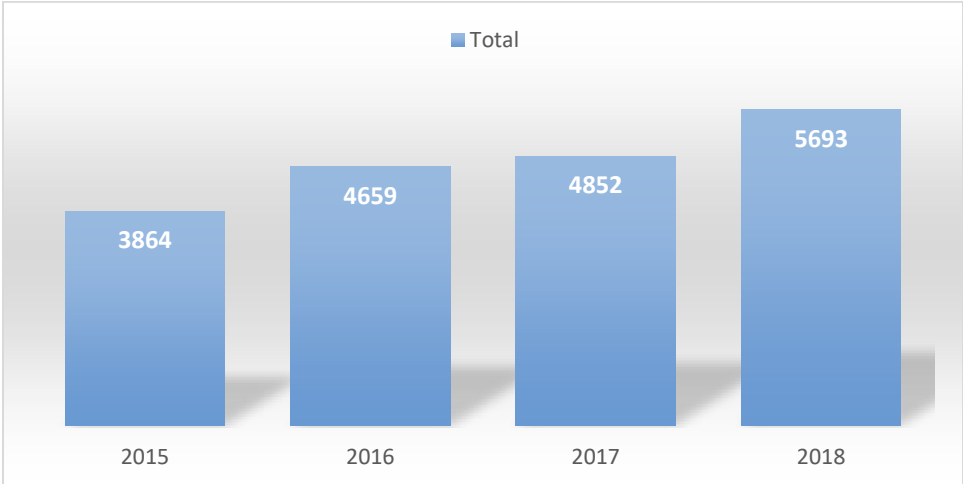


3.1.3. Third party organize the task

When the project is transferred to the third party, it processes the task. The project needs to record voice of around 1600 people from the age group 15-65 in multiple languages. The languages are Spanish, French, German, Danish, Swedish, Mandarin, Hindi, Portuguese, Arabic, Russian, Japanese, Lahnda, Korean, Marathi, Telegu, Bengali, and many more. They can be either male or female. The third party has to plan for the required task by recruiting the translator and record the voice. Every task is time sensitive. Hence, the tasks should be delivered to the clients within the time period. Hence, the management in the part of the third party is also important. Sometimes the third party fails to provide the tasks in the scheduled time. The reasons

for such delay involves difficulty in finding particular translators, delay in recording by the translators, delay in processing the records, delay in quality check etc. The average number of recordings by the third party is shown in figure 10. It has decreased over time as it gets efficient. In 2018, on an average around 5700 recordings were processed in every 1000 days.

Figure 10: Recordings in each 1000 Days



3.1.4. Advertisement by recruiter (organization)

The third party would require the participants to translate the documents into their languages and record. This process require a lot of participants as they are screened by the recruiter for different projects based on the requirements of the clients. This job is done by our organization as there are certain categories that needs to be fulfilled by the participants who gives their voices. The client may specifically ask for translating and recording of some document into Spanish language by a female person aged 24 to 27. In this particular case, the recruiter has to find some participant who fulfills the basic

requirements. The recruiter has to reach out vast majority of the population who speaks the desired language. Hence, reaching out through daily newspapers and social media are very good options.

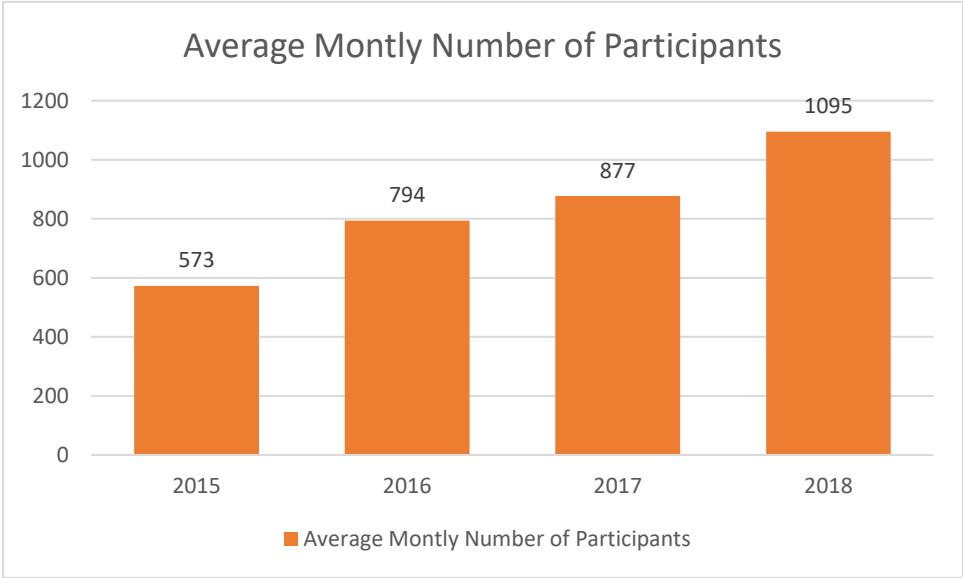
Recruiter from the Tampere office posts on social network sites and local newspapers about the task in each of the language-based countries. The advertisement includes the description of the job, the remuneration for doing the job and the last day for the signup. That also includes the website (<https://www.formsite.com/>) to sign up for the interested peoples. This website is mainly the website for the third party. Once the participants sign up, the third party can access all the informations from the site. The form includes information about the gender, age, language spoken by the participants, address, nationality, education and employment information. On the basis of such information, the moderator may check the categories mentioned by the clients for each task.

3.1.5. Sign Up participants in the third party website

Interested participants sign up in <https://www.formsite.com/>. We use this third party tool for our projects. We create the project profile, and people sign up there. After the deadline we export the users' information from the site. Every client has some requirements on the type of recordings. Some ask for a record of the documents in some language that has been recorded by a men aged 30-40 years. Some required a voice recording of a female of 25-30 years. Hence it is very important to have the information about the demographic characteristics of the participants. It becomes easier to recruit the participants when we know about their demographics. The average

number of participants that enroll or sign up for the projects are listed in the figure 11. It has increased in the monthly rate of signing up for the project. The main reason for this increase is increase in the demand as well as the increased percentage of freelancing opportunity. Working in online projects have become very popular for younger generation because of time flexibility and income. The average month registration of participants is 1095 in the year of 2018. It is expected to increase in 2019 and following.

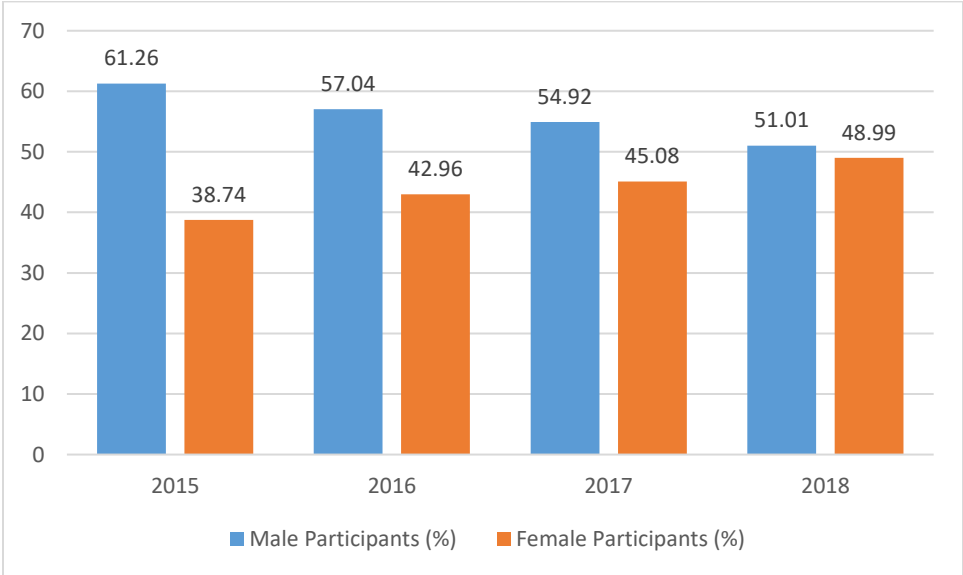
Figure 11: Number of participants signed up for the tasks.



As the clients demands the records in some specific form, the participants who records the tasks vary to a large extent. The male and female ratio of the participants is also close to one, having more male participants than females. The participation of female has increased in recent years. The percentage of participants that are female in 2018 is around 49 percent. The percentage of female participants in 2017 was 45 percent, while in 2016 it was around 43 percent (figure 12). Hence, the increase in the percentage of female participants also tells us about the demand for the projects as well

as the opportunities that has become available to the female population around the world. The projects this also contribute in building a skilled female labor force around the world that helps in women empowerment. Though the world is changing very fast with its technological advancement, women in many developing countries are far from the fruit of the development. Bringing females in the labor force by encouraging their participation in such activities is one of the main contribution of the project.

Figure 12: Number of male and female participants signed up for the tasks

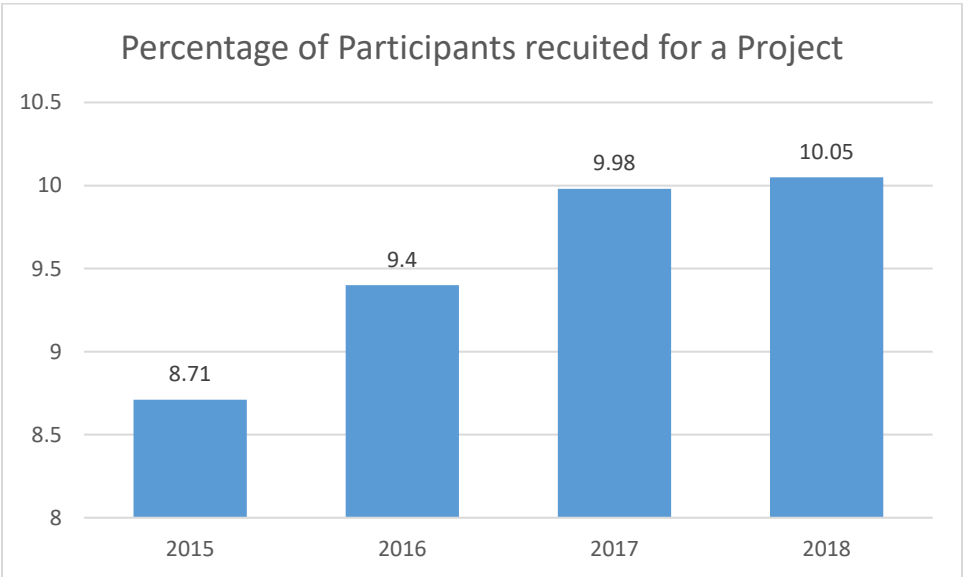


3.1.6. Screening by recruiter

Once the participants sign up for the projects, the recruiter starts the screening process. It is an important step for this project as different task require different set of participants to record their documents. This screening is done by the recruiters in our organization. After we export the users' profiles, the recruiters match them with the project criteria. Like in this example project, we need people from the age group 15-30

and can be any of male or female. It is not necessarily true that most of the task require voices from younger population. In many projects, the client requires voice recordings of kids who are 4 to 7 years old. Most of those projects are used for making kids toys and kids' applications like games and others. It is also very common that some project also require voice from older people aged more than 50 years. Those kinds of projects are mostly used for constructing elder friendly applications for mobile phones and computers. In the applications, where the user can select the voice, it is also important to have multiple voice recordings from different age group of people. It is also true for male and female participants. The average rate of participants that has been hired for a project is not more than 10 percent of the total participant pool. It has also increased over time as both the number of participants and the demand of such projects has increased (figure 13). After selecting the people, the recruiter contacts with the selected people via their personal mail, which they used to sign up for the project in formsite.

Figure 13: Percentage of Total Participants Recruited for a Project



3.1.7. Time Booking of participants and checked by moderator

This step is done by the participants. After screening by the recruiter, they select the participants for the tasks according to the task requirements. Participants get notified through the portal and they can select the time schedule for the task. There are several time scheduled uploaded in the scheduling options from where the participant can choose one. This scheduling port is managed by the third party. When the participants sign up and select their preferred time slots for recording, the recruiter again checks the time and the eligibility of the participants. This screening process is done by the recruiter from our firm. After the screening, the recruiter checks the time booking of the selected participants and informs them to be in our local office for do the recording in that time. The participants has to go to the local office, verify their identity and then start for the recording.

3.1.8. Recording participants in the presence of moderator

Recordings by the participants takes place in the local offices. After matching everything, the participants arrive in the local Lionbridge office to record in the booked time. Lionbridge has local offices in mostly every countries. The relevant opportunity cost of the participants involves the traveling cost and the time preferences. Not all the participants live in the nearby areas where the local offices are located. In this case, they have to travel to the local office to record the task. There is a traveling cost associated with this process. Moreover, not all the prospective participant register for the task because of this opportunity cost. It is sometimes inconvenient for the

participant. Furthermore, the time schedules are mostly in the office hours. Hence, the participants, who works for other full time jobs, cannot participate in this program. The whole process of recording takes on an average 26 minutes per participant. It off course varies with the length of the task and the recording quality. Some participants take several takes to record the whole task, some takes one. The whole recording process is controlled by the moderator. The time allocation of each participants and the number of takes for each participants are listed in the table below.

Table 2: Time allocation and the number of takes per participants

Year	Average Recording time per participant(minutes)	Average number of takes per participants
2015	32.75	4.7
2016	28.95	4.2
2017	28.7	3.8
2018	26.64	2.9

The time allocation per participants has decreased over time because of the increased efficiency of the monitors and the participants. Moreover the number of takes also decreases over time as the participants become used to the process. The probability of the participants being selected in future tasks is higher as they become experienced over time and the quality of the work has been observed.

3.1.9. Support for participant while recording

In the whole process of recording the tasks, there is a need of constant support. Not only the participant's need the support in the recording process, the moderator and the recruiter also need support in the whole process. Hence, the firm organizes a support team to give 24 hours support to any participant, recruiter or the moderators. The support team provides technical, resource and managerial support to the participants, moderator, recruiter and manager. The support team consists of around 20 people. In this whole process, whenever the participants face any issue in any point, our support team provides the support. Support for participants include scheduling time, changing the schedule, cancelling the schedule, payments, and other disability supports. Supports for the moderator is mostly technical. Sometimes the recording through the internet does not work or face some difficulties. Supplying the technical tools to local office is firstly organized by the support team. The recruiter also needs the support of the support team in managing the time schedule and the participants' eligibility and scheduling time. Major task for the support team is maintaining the technical supports such as IT supports, hardware and software malfunction etc. the manager needs to update the schedule for all the local offices and the participants. It also needs a constant support by the technical support team. The support team is online 24 hours every day, where they allocate 5 support manager all time. They work in shifts so that they can support any problem any time.

3.1.10. Payment by the recruiter

After the recording goes correctly, we need to pay the contracted amount to the participant. And it goes from the Lionbridge Tampere office. For this, the moderator needs to contact with the manager every time and asks for his\her permission and after the approval he\she proceeds with the payment. For now, we are paying using PayPal, but if the participants' country doesn't support PayPal, money sends to the moderator to the local office. Then the moderator pays the participants in cash. The payment varies with the type and length of the task. The average payment per participants is USD13 per hour in the year 2018.

3.1.11. Check and review the recordings

When the recordings are submitted, the support team checks and review the quality of the recordings. Some of the recordings are very low quality and the quality assurance team has to drop those records. The quality assurance team works under the firm. The team consists of 5-6 employees where they constantly checks all the recordings. If the quality of the recording does not satisfy a minimum quality criteria, it is sent to the third party again to re-record the task. The rate that the recordings does not satisfy the minimum quality has decreased over time. Still the rate is very high. The percentage of the good quality works are listed in the table below.

Table 3: Percentage of good quality recordings

Year	Good Quality Recordings (%)
2015	73.73

2016	76.94
2017	80.04
2018	79.95

In 2018, on an average 80 recordings passed the quality check among 100 recordings. Hence 20 percent of the recordings goes back to the third party to process this again. This not only involves with a loss of time and effort, but also associates with a huge cost. The recruiter has to hire more participants for the additional recordings and the third party works for the whole process for the second and sometimes for the third times. The firm thus has to pay the additional participants and pay for the additional times worked by the moderator, recruiter, third party and the support teams. The average cost involves with this process accounts for 12- to 15 percent of the total cost of the recordings. This percentage was even higher in the previous periods. While all the recordings are collected by the firm and assessed by the quality assurance team, they are ready to deliver to the clients.

3.1.12. Deliver the product to the buyer

After finishing the recordings and the quality check, the team sends the recordings to the client. This is not the last step of the processing. The client may get back to the organization involving many issues regarding the products. The issues are mostly related to the quality of the recordings and the time preferences. Most of the recordings are involved with some time bounds. For an example, a recording cannot be more than 15 seconds long. On some cases they are associated with the bounds such as 12.30 seconds to 12. 59 second. It is obvious that the recordings are used in some

software and applications. They are frequency sensitive as well as time sensitive. Hence, if the demand for client is not met, he comes back to the organization and ask us to run this recordings again. The team has to go over with each issue separately and try to solve them. This also brings the cost up. It is also a loss of efficiency for the organization.

3.2. Proposed Model

This section contains the description of the Lean Canvas Model based on our product. It discusses the problem of the current model and possible solutions in the first part. In key metrics, it argues about how the participants use the portal. Furthermore, we also discussed how the unique value proposition and the advantages that both the participants and the clients receive. Moreover, it discusses about the cost structure and the revenue streams for the business.

3.2.1. Problems in current model

The study identifies some areas of the current model that can be improved. In the current model, the lengthy process needs a lot of time and human efforts. It is involved with a large human labor just in the support and quality assurance teams. Every team has to wait for next available task and for this reason everyone is dependent on other teams. Moreover the process is costly as it involved third party to conduct some of the work for the organizer. As the quality assurance teams gets the recordings after they are recorded and processed by the local recording centers, it involved time laps. Delivering the jobs done by one team to the next team requires

some time. This is the first inefficiency of the current model. The second inefficiency involves with the effort or the opportunity cost. If the recordings cannot pass the quality check, it is returned to the thirst party again to re-record the task. The efforts and time involved from selecting the participants to the delivery of the recordings are lost. Not only the firm supper for this loss, but also the employees, have third party and participants also suffer from this loss. The participants have to travel again to the recording center to record the exact same task. The moderator has to organize the recording and the quality assurance team also put their effort.

From the IT perspective, this whole process is very less secured as we are using two 3rd party tools and some of the vital information of the project is online. As almost all our projects seek the confidentiality of our clients' upcoming products. The current model could not assure this security because of the involvement of several parties including the third party. Hence we propose a new model where everything is done in the same portal and all of the parties can access the portal at the same time. Whenever there is a problem associated with any recording or any process, the concerned team is notified instantly and can solve the problem. In this way all the parties can work on the efficient level and the opportunity cost of every party is minimized to a large extent.

3.2.2. Solution

To solve all the problems identified in the current model, the study proposes a new model where the process is being automated. The new model do not involve the third party to do the work. There is only one online portal where the clients submit their

tasks and participants sign up to do the work. The main features that the new model is adopting are listed in the table 4.

Table 4: Features of the proposed model

Automated screening process
Usage of attributed email templates which enable communication date logging
Participant progress statuses semi-automated
Detailed participant progress column
Editable participant details
Enhanced participant selection for future tasks
Support for document signing such as NDAs or Parental Release forms
Export to excel
Project summaries with calculated completion date and agreed deadline.

We adopt an automated screening process where the participants’ information goes directly to the screening algorithm and the process identifies the participants who are eligible to certain task. In the portal, the participants can create their own account where they put all their information along with the national identification number or the passport. In this way the portal has all the information about the participants. We maintain ethical requirements in storing and maintaining the privacy of the participants.

As soon as the participant sign in, the portal automatically process the screening for each participants and send it to the recruiter section. Participant's informations are categorized in the portal and it helps the recruiting algorithm to find the eligible participant for each task. We use the attributed email templates which enable communication date logging. This system automatically sends the notification to each eligible participants informing the type of the task and the deadline to sign up for the task. Participant progress statuses are semi-automated in this new model. The program is set in the way that the monitor and the quality assurance team can check the progress of the work done by the participants. The team can see if the participant agreed to record the task, is in progress or completed. Detailed participant progress column stores this information for all the parties involved in the process. We do not need the recruiter or moderator in this new model.

However we need extended support and quality control team for this model. Editable participant details allows the participant to edit the information or update if needed. We also offer an enhanced participant selection for future tasks. The program stores the participants' data along with the tasks they have done so far. It also stores if their work passed through the quality check. The more the participant is active and the more the rate of the good quality task. The participant is rated high. The rating system is also automated by the system. This also allows the support team and the algorithm to choose the efficient participants for the next available job. Moreover, the algorithm is set in such a way that every time it chooses the new participants randomly from the participant pool. This accounts almost 20 percent of fresh participants to participate in

the new task. The algorithm also allows at least 40 percent of the female participants to choose from the pool whenever the task is not gender specific.

We also follow the ethical background using the support for document signing such as NDAs or Parental Release forms. This allows the participants below 18 years old to be verified by their parents. We export all the data to an excel spreadsheet for convenience. We also do the project summaries with calculated completion date and agreed deadline. The quality assurance team checks the recordings instantly as soon as they are updated in the portal. We also use automated system to check the frequency level and the time allocation so that it meets the criteria. When it is checked by the quality assurance team is is updated and available to the clients for the final verification.

3.2.3. Key Matrices

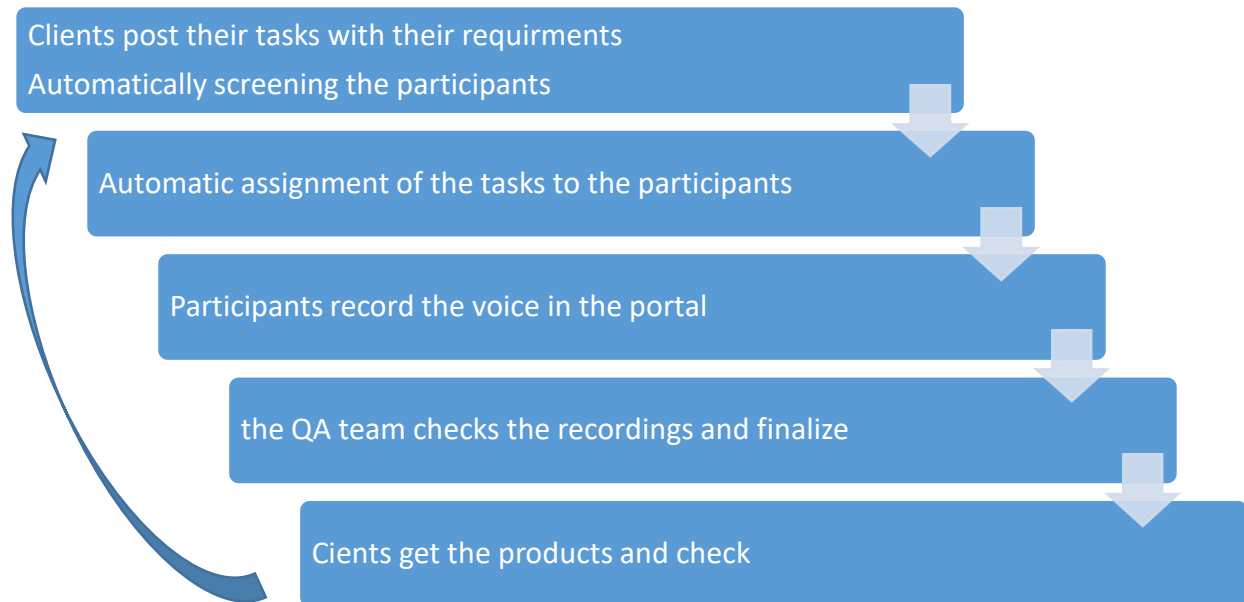
The first task is to open the online platform. Then there are specific areas for clients, participants, quality assurance team and the support team in the same portal. The organization releases the advertisement in social media and other popular websites so that the participants from around the world sign up in the portal and get paid by their freelancing jobs. Encouraging the sign up is easier through social media. The marketing team of the online platform work together to collect the participants from around the world because the platform requires people speaking different languages.

When the participants sign up in the portal, there is an automatic screening process through which it will make a list of participants for different projects. Some projects may ask 25-35 year old female to record the documents uploaded in the portal.

The portal then automatically assigns the work to those who are eligible to work with the task. Here language is the main criteria to select the participants. All the documents provided by the clients are in English. Hence the participants have to translate the documents in their languages and record it with their own voice. In this case they record the voice in same platform to maintain the quality of all records being similar. It is also very convenient to the quality assurance team because the initial check-up would then be automated through the website. The platform would then just send the records those are not in the given limit in Harz.

When the quality assurance team checks the quality of the recordings, they become available to the clients and they can also check the recordings and see if they meet their requirements. If the client is not satisfied with any product they can send it back to the quality assurance team. The process then be repeated until the client is satisfied with the product.

Figure 14: The Key Matrices of the New Model



In the new model, participants and clients have more flexibility in work. The client can rate each product and thus automatically the participants become rated through their work. Hence the quality of the product can be improved. On the other hand, the organization now requires less human labor to manage the whole process. It saves the costs in business.

Moreover, the new model contributes to the world economy by engaging a huge lot of people in quality jobs. People from different countries can sign up online and do their work and get paid. Participants can do this kind of job in their leisure time and earn some extra money. On the other hand, it is a great opportunity for students who search for work with flexible hours. It is also a very good opportunity for females to participate in the labor market while staying at home.

3.2.4. Cost Structure

The main areas in the current model where the new project works on are human labor, expensive and lengthy process and safety. By looking into all the issues, the best solution is to create own HTTPS Lionbridge site, which needs to be accessible from internet all over the world. The site contains the platform where people sign up, book the preferred time and get paid. The managers also get notification for the approvals from the site before the payment. It needs some more research for making the site perfect.

3.3. Pilot project

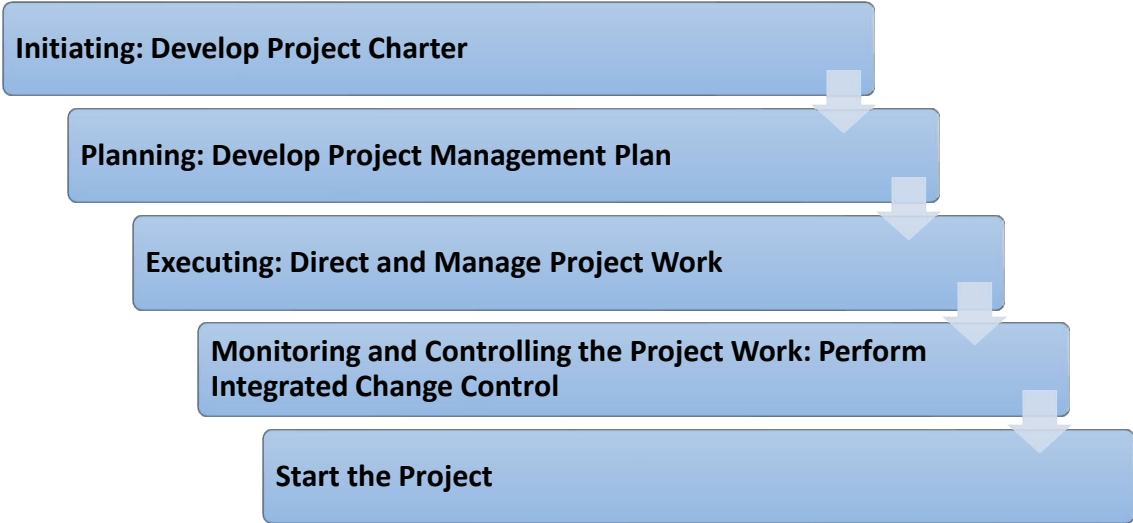
The firm has already started the pilot project following the proposed model. It shows that the total cost per recording has decreased by 28 percent. It also makes the whole process faster and efficient. The pilot project is described in the following section. While describing the project the study explores all the managerial areas that need to be updated observing the pilot project. We follow Schwalbe (2015) management areas in the analysis.

3.3.1. Project Integration Management

Developing the project charter is the first step to start the project. It identifies the project information and the key stakeholders. The project management plan then offers a consistent and comprehensible document that facilitate the action of the project. The

managing of the project is involved with a high budget. It manages all the teams and provides the support for every group. Monitoring and controlling the project offers the assessment on the project performance. Integrated control offers identifying and managing the project life cycle.

Figure 15: Project Integration Management Summary

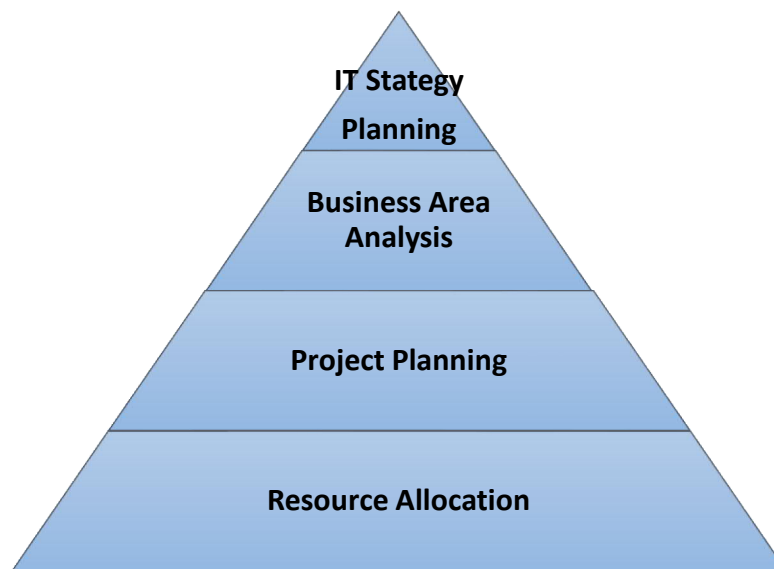


Source: Schwalbe (2015)

We first identify the resource available for the project and allocate it to start the project. The Information Technology based projects such as our model is involved with lower physical recourses and higher human capital. Hence it has been difficult to allocate the recourses for this project. The pilot project did not require much resources. We have used our existing capacities to make the pilot project run. Further research in this area is required as we extend the project. While making the project plan, we focus on the scope, benefits and the constraints of the project. We also analyze the business area analysis that includes documenting key business process that could be benefited by the project. We found that the clients would appreciate the change as it becomes

higher quality in delivering the services and it takes less time to make and process the tasks. We also identifies the mission and vision of the project to be integrated with the information technology strategy planning.

Figure 16: Planning process for IT project



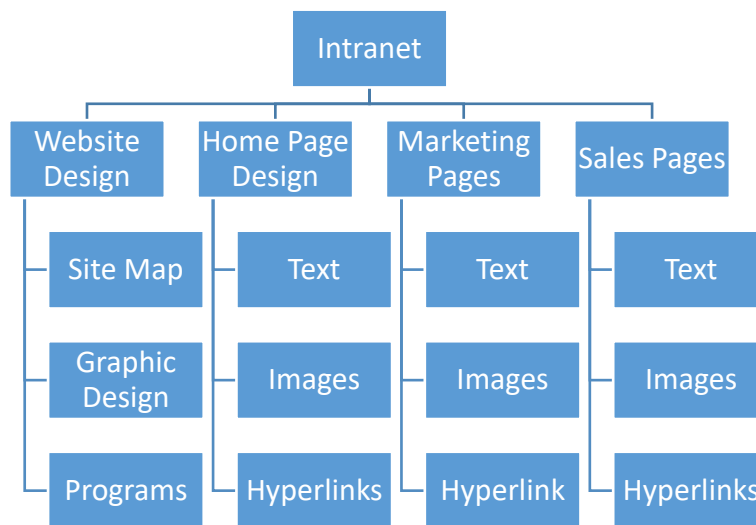
Source: Schwalbe (2015)

We perform the financial analysis where we conduct the net present value analysis, investment analysis, payback analysis, and the analysis on the return associated with the risk. As soon as we gets the significant results from the pilot project, the firm will extend this analysis to the larger context.

3.3.2. Project Scope Management

Scope management is important to have a detailed definition required for the project. It includes the definitions of work needed to process and improve the accuracy of time, cost and resource estimation. The scope management is done in several steps. Our project scope management follows defining the key schedule management, budget information, tasks for project manager, defining the objectives and the main criteria for the success of the project. We strictly follow the rules for ethical project management and security. The design for intranet follows the process on what we can provide maximum security of the data and provide an easy interface. The sample intranet design is provided in figure below.

Figure 17: Sample Intranet Design

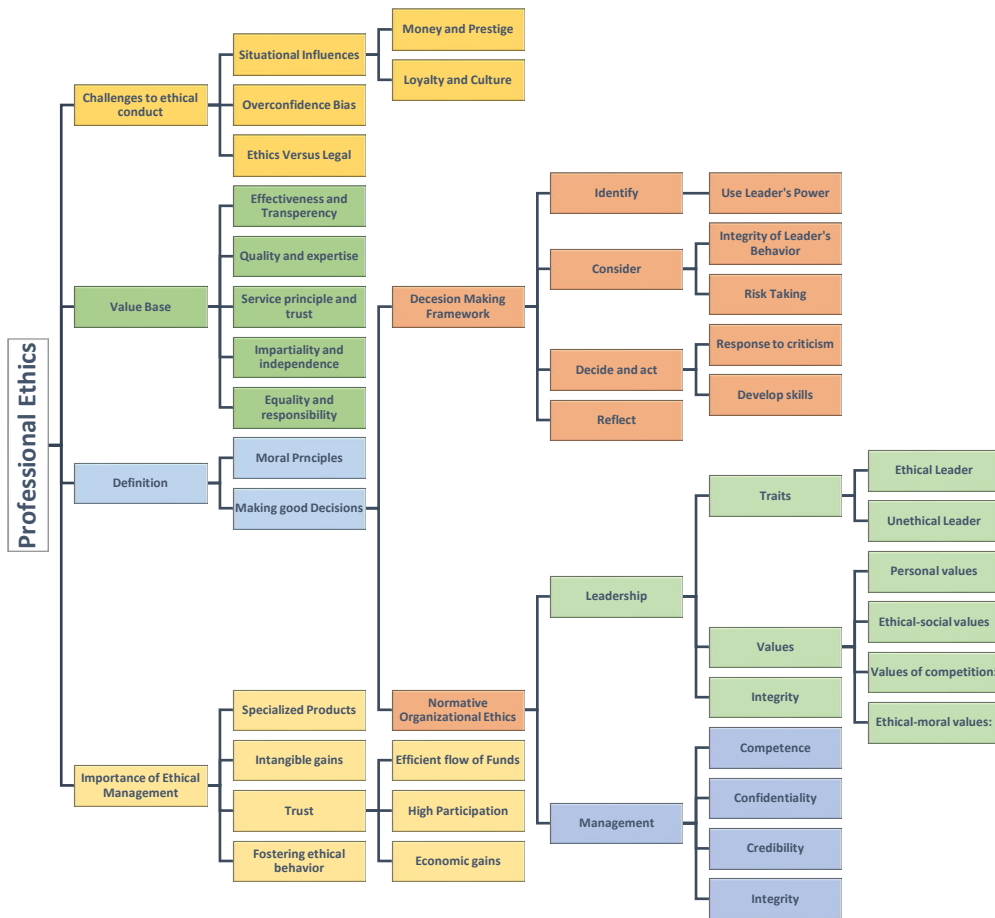


Source: Schwalbe (2015).

The ethical project management mind map is presented in the following figure. Concerning “professional ethics (of management and development)”, we focus on the

moral and ethical characteristics of the projects as well as the people related to the project that help the management and development of the projects.

Figure 18: Mind Map for Ethical Project Management



The main idea is to develop and manage the projects that benefits people.

Ethical concerns in this matter help building trust, spreading the network, and adding the value of the outcome. It concerns about the social benefit rather than the benefit of a single entity. Professional ethics build effectiveness and efficiency in the project as well as in the society. We mainly focus on equality while considering the project management. Equality brings the approach where all the citizens are treated equally.

No one is benefited more than the other citizen through a project. Hence, the new model aims to promote equality in every aspect in life including equality in decision making, equal distribution of wealth, equal employment opportunity and wages etc.

3.3.3. Project Time Management

Before starting the pilot project, the team focused on some of the issues concerning the time management. We focused on the steps and measures to be taken to run and maintain the project in a timely manner. We made a project plan scheduled where all the tasks are listed along with the time schedules for each tasks. We define the activities that needed to be done. We list the teams or the persons who are in charge of the task. While doing the sequence activity plan, we have concentrated on the network diagrams, project document updates etc. Before developing the whole schedule for the project, we also focused on the estimation of resources available versus resources needed for every team.

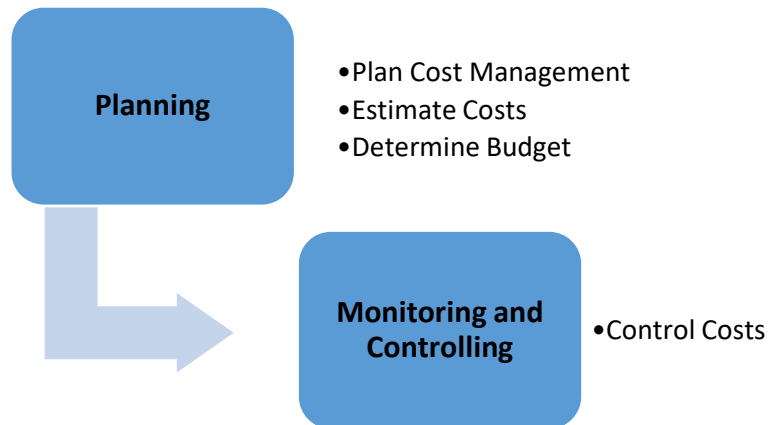
Time management also concerns about the activity duration and projection of the time for activities. We estimate the time required in every step of recording the tasks that are given by the clients. From uploading the task to the delivery of the task to the client, estimated time in every steps has been calculated. We are also storing the time duration in each step of the project so that we can predict the time required for future work. It also helps in cost benefit analysis or in maintaining the financial accounts too.

3.3.4. Project Cost Management

While designing the cost management of the project, we follow the strategy described in figure below. We first make a cost management plan concerning the estimated cost and allocate the budget for every department. We mainly focus on the cost in the pilot project, project funding and the forthcoming costs. We also make a plan to control the cost throughout the project. This includes the performance of the project, forecasting the cost, updating the project assets and resources.

In the updated model we also have to focus on the level of accuracy in the cost management, evaluation of possible risks and costs associated with the risks, and several cost controlling devices.

Figure 19: Project Cost Management Summary



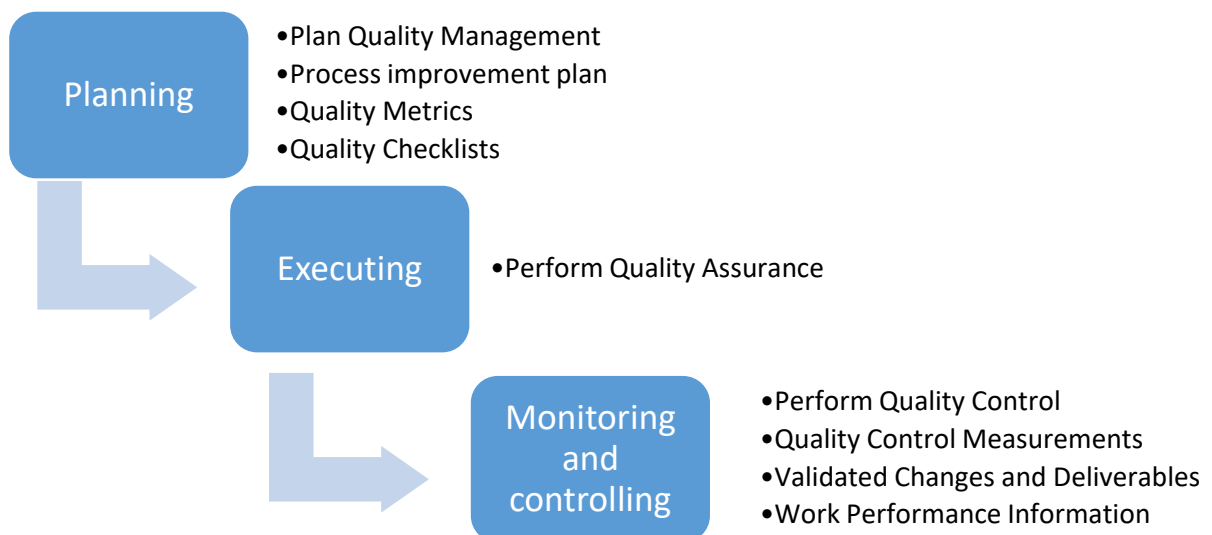
Source: Schwalbe (2015)

3.3.5. Project Quality Management

Our main focus of the project management relies on this section. We mainly focus on the project quality management. While are concerned about the service we

provide, we plan for the quality management for the project. We develop the quality metrics where the information of the products are listed. In our case it is the recording that we deliver to the clients. We also develop the quality checklists for the recordings. It is a very important step because it goes with the algorithm that we use in the online portal. The basic quality screening of the recordings are done automatically through the project portal. Throughout the whole process, we focus on the quality assurance of the product. We employ several members from the quality assurance team to monitor and control the quality. The project quality management system that we use for our project is described in figure below.

Figure 20: Project Quality Management system



Source: Schwalbe (2015)

iv. Conclusion

Due to huge demand of information technology (IT) based product, the concerned firms have been adopting the technology that faster the process and improves the quality of the products or services. This study concentrates on automation of a data collection model. Using the data and strategy approved by the concerned firm, we develop a new model that improves the whole data collection system. This study proposed an online platform to collect the data. The current model used to deal with a third party to collect the participants who can record the documents sent by the clients. The firm gets several different deals from several types of the firms who asks for voice recordings in many different languages. The participants who records the documents in different languages varies in ages. The clients specify the requirement of the gender and age of the participants for each task.

The current model gives the contract to a third party to make an advertisement about the task and collect the participants through their online portal. This model is very time consuming and lengthy because it involves several teams working in a chain. Every team must wait for the other team to complete their task and send to the next team. Moreover, the quality assurance team checks the quality of all the recordings while they are done. More than 20 percent of the recordings cannot pass the quality check. If this checking was done simultaneously while recording, this would save time both for the participants and the employees.

Considering the pros and cons of the current model, I proposed to get this new model where every step is automatic and done in a single online platform. This new online platform can be accessed by all the stakeholders' equality at the same time. Participants have their space to sign in and register for a task. The moderator and the recruiter then merge into one single team who provides supports for all other teams. Quality assurance team can run their work as soon as the recording is done. Hence the result on the quality check is updated very fast so that the participant can record the task again. The whole process is very short and less costly than the previous model.

The firm has adopted a pilot project to access the feasibility of the project and it has got very good feedback from all the parties involved. The cost is minimized with a higher quality service. Customer satisfaction and labor satisfaction is also very high. Further studies may concentrate on how to improve the proposed model with higher technology adoption.

References

- Damian, D., & Manea, C. (2019). Causal recipes for turning fin-tech freelancers into smart entrepreneurs. *Journal of Innovation & Knowledge*.
- Forbes.com. <https://www.forbes.com/sites/louiscolombus/2018/01/12/10-charts-that-will-change-your-perspective-on-artificial-intelligences-growth/#663b537c4758>
- Just, N. (2018). Governing online platforms: Competition policy in times of platformization. *Telecommunications Policy*, 42(5), 386-394.
- Needham, J., Sidwell, D., Frommolt, V., Barnewall, K., & Grafton, E. (2019). The implementation of a nursing student clinical assessment tool onto an online platform: An evaluation by clinical facilitators and academic staff. *Collegian*, 26(4), 422-427.
- Nidagundi, P., & Novickis, L. (2017). Introducing lean canvas model adaptation in the scrum software testing. *Procedia Computer Science*, 104, 97-103.
- Pocock, N. (1986). Opportunities in freelancing. *Data Processing*, 28(2), 82-86.
- Schwalbe, K. (2015). *Information technology project management*. Cengage Learning.
- Shevchuk, A., Strebkov, D., & Davis, S. N. (2015). Educational mismatch, gender, and satisfaction in self-employment: The case of Russian-language internet freelancers. *Research in Social Stratification and Mobility*, 40, 16-28.
- Süß, S., & Kleiner, M. (2010). Commitment and work-related expectations in flexible employment forms: An empirical study of German IT freelancers. *European Management Journal*, 28(1), 40-54.

- Sultana, R., Im, I., & Im, K. S. (2019). Do IT freelancers increase their entrepreneurial behavior and performance by using IT self-efficacy and social capital? Evidence from Bangladesh. *Information & Management*, 56(6), 103133.
- Wei, J., Lu, J., & Zhao, J. (2019). Interactions of competing manufacturers' leader-follower relationship and sales format on online platforms. *European Journal of Operational Research*.
- Womack, J. P., Womack, J. P., Jones, D. T., & Roos, D. (1990). *Machine that changed the world*. Simon and Schuster.
- Yoshimoto, H., & Nakabayashi, J. (2019). Search and resale frictions in a two-sided online platform: A case of multi-use assets. *Journal of Economic Behavior & Organization*, 162, 85-105.