Identifying and Evaluating Early Stage FinTech Companies:

Working with Consumer Internet Data and Analytic Tools

In partial fulfillment of the requirements for the Degree of Master of Science
At *Worcester Polytechnic Institute*, Worcester, MA

By:
Madhuri Surve
Sizhu Chen
Alexander Shoop
Khasan Dymov

Advisors:
Kevin Sweeney
Marcel Blais

Sponsor:
Vestigo Ventures, LLC. & Cogo Labs

January 24, 2018

wpigrad-fintech2017@wpi.edu

Abstract

The purpose of this project is to work as an interdisciplinary team whose primary role is to mentor a team of WPI undergraduate students completing their Major Qualifying Project (MQP) in collaboration with Vestigo Ventures, LLC. ("Vestigo Ventures") and Cogo Labs. We worked closely with the project sponsors at Vestigo Ventures and Cogo Labs to understand each sponsor's goals and desires, and then translated those thoughts into actionable items and concrete deliverables to be completed by the undergraduate student team. As a graduate student team with a diverse set of educational backgrounds and a range of academic and professional experiences, we provided two primary functions throughout the duration of this project. The first function was to develop a roadmap for each individual project, with concrete steps, justification, goals and deliverables. The second function was to provide the undergraduate team with clarification and assistance throughout the implementation and completion of each project, as well as provide our opinions and thoughts on any proposed changes. The two teams worked together in lock-step in order to provide the project sponsors with a complete set of deliverables, with the undergraduate team primarily responsible for implementation and final delivery of each completed project.

Special Thanks & Acknowledgements

We would like to extend a thank you to Ian Sheridan and Mike Nugent of Vestigo Ventures for the great opportunity of completing this project under their guidance and learning a vast amount about the world of venture capital. Their expertise and assistance was invaluable throughout the duration of the project.

We would also like to thank John Roland, Rob Fisher, and Matt Wiens of Cogo Labs for their patience and unwavering support throughout this project. Their willingness to provide recommendations and assistance, in particular with Cogo Labs' analytical tools, was exceptionally useful.

We would of course like to thank our project advisors, Professor Kevin Sweeney of the Foisie Business School and Professor Marcel Blais of the Mathematical Sciences department, for allowing us this opportunity and supporting us throughout the duration of the project.

Lastly, we would like to thank the undergraduate student team of Andrew Aberdale, Ben Sarkis, Rekik Tafesse, Zhi Hui, Juan Carlos, and Juan Pablo, for all of their hard work throughout the project, and for being a tremendous pleasure to work with.

Everyone's contributions were extraordinarily helpful and undoubtedly facilitated the success of this project.

Table of Contents

Abstract **2**

Special Thanks & Acknowledgements **3**

Table of Contents **4**

Executive Summary **5**

1. Introduction & Background **7**

1.1 Vestigo Ventures 8

1.2 Cogo Labs 9

1.3 Venture Capital 9

1.4 Fintech 10

1.5 The Projects 11

2. Methodology and Theory **11**

2.1 Project Planning and Management 12

2.1.1 Resources and Team Setup 12

2.1.2 Timeline 12

2.1.3 Work Progress Tracking 14

2.2 Theory behind Vestigo Project 1: Venture Decisions Measured Over Time 15

2.2.1 Cataloguing the deals 16

2.2.2 Analyzing the passed deals 17

2.3 Theory behind Vestigo Project 2: Deal Analysis and Industry Engagement 18

2.3.1 Market Segmentation and Business Model 19

2.3.3 Market Research (Survey) 20

2.3.4 Digital Transformation 21

2.3.5 User experience and Survey form design 22

2.3.6 Deal Evaluation 22

2.4 Theory behind Vestigo Project 3: Signal Search and Market Landscape Analysis 24

2.5 Theory behind Cogo Labs Fingerprinting Project 24

2.5.1 Terminology 25

2.5.2 Home IP 26

2.5.3 Probabilistic model 26

2.5.4 Algorithm model table processing workflow 29

3. Discussion and Conclusions **29**

Bibliography **33**

Appendix **38**

Executive Summary

This project was a collaboration during the Fall 2017 semester between Vestigo Ventures, Cogo Labs, and a total of ten students from Worcester Polytechnic Institute - four graduate students and six undergraduate students. Vestigo Ventures and Cogo Labs, both located in Cambridge, Massachusetts, had previously entered a partnership designed to leverage each firm's capabilities and expertise to advance both firms in their respective areas of operation. This partnership created the opportunity for a student team to step in and provide creative solutions to several projects in the pipeline for the firms.

Vestigo Ventures and Cogo Labs identified four specific projects for the student team to complete - three projects for Vestigo Ventures, and one project for Cogo Labs. The projects were centered around each company's area of operations.

The three Vestigo Ventures projects, titled "Venture Decisions Measured Over Time," "Deal Analysis & Industry Engagement," and "Signal Search & Market Landscape Analysis," were a mixture of administrative and infrastructure work, business environment research, and financial analysis. The business rationale - Vestigo Ventures' key objective - is to find and fund potential startups within the FinTech industry. In these projects, the student teams dove right into the business methodology of Vestigo Ventures, gaining valuable insight into the entire process driving the business, and ultimately provided a complete and useful set of deliverables to the Vestigo Ventures management team.

The Cogo Labs project, titled "Device Fingerprinting," was more technical and involved. This project required analyzing a minute portion of Cogo Labs' enormous data set by combining the use of mathematics and Cogo Labs' tools. As an analytics firm focusing on consumer internet behavior, Cogo Labs collects vast amounts of information about internet users. The issue arises in utilizing this data successfully. The student team, with insights from Cogo Labs analysts, dissected the provided dataset and constructed an algorithm that provides useful insights into the dataset that are in line with Cogo Labs' business strategies and goals - better consumer profiling.

The role of our graduate team was to serve as mentors to the undergraduate team throughout the duration of the project. As such, we had several key responsibilities. We needed to understand the projects proposed by Vestigo Ventures and Cogo Labs, and turn this understanding into a specific plan of action, timeline, and set of key deliverables for each project. This required iterative brainstorming and development for each project, communication with Vestigo Ventures and Cogo Labs management, and application of each graduate student's expertise and background knowledge. Our graduate team consisted of an Information Technology graduate student with former work experience as an Analyst at the National Stock Exchange of India (NSE) who served a Business Analyst role and was responsible for coordinating with Vestigo Sponsors in order to understand the business requirements and to convert them into technical ones with the help of the undergraduates, a Marketing & Innovation graduate student who brought a rich marketing project background and was responsible for market research and deal analysis reasoning, and two Financial Mathematics graduate students with statistical analysis backgrounds who brought general financial industry knowledge and were responsible in prototyping the algorithms for the probabilistic models.

The graduate team provided the undergraduate team with the resources necessary to successfully manage their workflow, as well as communicate their progress to us, Vestigo Ventures and Cogo Labs, as well as their individual academic advisors. This was done through the establishment of communication channels and more advanced tools, such as GitHub. Lastly, the graduate team served as a resource for clarification and assistance to the undergraduate team. The academic and professional backgrounds of each graduate team member facilitated this, as we were able to provide unique insights and tips to the undergraduate team. Additionally, we were able to provide recommendations and reviews of additional ideas that the undergraduate team had in addition to the goals and objectives set by us.

The role of the undergraduate team, who worked full-time on-site at Cogo Labs, was to complete the implementation and provide the final deliverables to Vestigo Ventures and Cogo Labs. Furthermore, the undergraduate team provided refinements and adjustments to the deliverables as deemed necessary by their own analysis. The results of their work can be found in their Major Qualifying Project final report (Aberdale et al., 2017)).

1. Introduction & Background

This project was the result of a large collaborative effort between WPI's Foisie Business School (FBS), the Mathematical Sciences Department, the Global Projects Center, and the project sponsors, Vestigo Ventures and Cogo Labs. The intent of this collaboration was to provide a team of undergraduate students the opportunity to complete their Major Qualifying Project (MQP) with the WPI Wall Street & FinTech Project Center during the B-2017 term. However, it was decided by the project sponsors and the director of the Wall Street & FinTech Project Center, Professor Kevin Sweeney, that it would be beneficial to have a team of graduate students to begin the work prior to B-2017 term, and then oversee the undergraduate student team as they complete their MQP. Our team was formed as a result of this decision.

Given the nature of Vestigo Ventures' and Cogo Labs' business areas, it was decided that the graduate team should be comprised of two students from the Foisie Business School and two students from the WPI Mathematical Sciences Department. This would give the graduate team the necessary level of knowledge and expertise in order to complete the objectives set forth by the advisors and project sponsors. Our team consisted of the following students:

1. Madhuri Surve - M.S. Information Technology, Foisie Business School
2. Sizhu (Zoey) Chen - M.S. Marketing and Innovation, Foisie Business School
3. Alexander Shoop - M.S. Financial Mathematics, Mathematical Sciences
4. Khasan Dymov - M.S. Financial Mathematics, Mathematical Sciences

For the purposes of degree completion requirements, Professor Kevin Sweeney of the Foisie Business School managed the Business School students, and Professor Marcel Blais of the Mathematical Sciences department managed the Financial Mathematics students.

The objective of this collaboration was to complete a set of projects that Vestigo Ventures and Cogo Labs' currently had in the pipeline, with the thought that a student team with an outside perspective could provide new and interesting solutions to the current problems faced by both the companies.

1.1 Vestigo Ventures

Vestigo Ventures is a venture capital firm headquartered in Cambridge, Massachusetts. Vestigo invests solely in financial technology (FinTech) firms, and targets initial investment in seed and Series A companies. The company is the first major FinTech venture capital (VC) firm based in Massachusetts. Its founding general partners are Dave Blundin (co-founder of Vestmark, Cogo Labs, EverQuote, and a variety of other FinTech and general technology firms) and Mark Casady (retired CEO of LPL Financial). Vestigo leverages proprietary big data capabilities to identify investment opportunities, and also leverages its financial industry expertise, behavioral economics data, and a network of advisors to support portfolio companies (Vestigo Ventures, 2017).

Ian Sheridan is a co-founder and managing director at Vestigo Ventures, and Mike Nugent is a managing director. Together they oversaw the three Vestigo Ventures projects.

1.2 Cogo Labs

Cogo Labs is a startup accelerator, also headquartered in Cambridge, Massachusetts. Unlike Vestigo Ventures, Cogo Labs is not a venture capital firm. Instead, Cogo Labs is an incubator that focuses on big data analytics. Mira Wilczek is President and CEO of Cogo Labs and a Senior Partner at Link Ventures. David Blundin is Chairman, co-founder, and former CEO of Cogo Labs. With Cogo Labs, David Blundin took his past experience with TripAdvisor's search engine optimization mastery and spun into businesses in an array of different industries. In this regard, Cogo Labs differs as well from Vestigo Ventures - there is no FinTech specific focus. Instead, the company utilizes advanced statistical analysis and software engineering to identify startups or concepts with potential for growth. Any such startups then have the potential to come into Cogo's incubator.

Cogo Labs and Vestigo Ventures have come together in a partnership designed to leverage the expertise of both firms. This project is just one piece of that, and will give all of the participating students a unique look into the world of venture capital and data analytics.

1.3 Venture Capital

The term "venture capital" (VC) is generally defined as a type of "investment in the ownership element of new or fresh enterprise" (Merriam-Webster, 2017). In more depth, venture capital can go beyond just the act of providing financial assistance for startups - it can also provide business or growth expertise as well. As such, venture capital is extremely important to the business world, and is extremely prevalent in the United States. In fact, according to the National Venture Capital Association, in 2016 the United States' share of global venture capital investment was 54% (Veghte, 2017). Furthermore, in 2016 there were 898 venture capital firms in existence in the United States, and these firms raised a total of US $41.6 billion (Veghte, 2017). It is evident that the venture capital market is massive, and massively important.

As mentioned previously, Vestigo Ventures is a venture capital firm but focuses specifically on investing in "FinTech" companies. In fact, Vestigo "bills itself as the only VC firm in Greater Boston focused exclusively on financial services technology, or fintech, investments" (Ryan, 2016). This opens up a tremendous amount of opportunities for Vestigo, but consequently, results in a constant and large inflow of information and communication that the firm must work through. This fact became increasingly clearer to us as we progressed through the project, and helped give us direction.

1.4 Fintech

FinTech is short for financial technology, and is defined as the disruptive technology driving startups and revolutionizing banking, payments, and insurance (Mckinsey, n.d.). FinTech can also be defined as technology innovations applied to improve the financial service and activities. There are many technology applications included in this area, such as artificial intelligence, big data, Internet of Things (IoT), cloud computing, blockchain, cybersecurity, and many more. According to Forbes and KPMG, $3.5 billion was raised by financial technology

companies in the US in the first half of 2017, indicating a large and healthy market for FinTech (Gensler, 2017). In fact, the FinTech industry has had explosive growth in the past decade, particularly as a result of the 2008 Financial Crisis. According to Kantox, some of the reasons for this rapid rise of FinTech include "anger at the established banking system, widespread lack of trust with banks post-crisis, lack of lending post-crisis, the changing relationship people have with money as a result of the internet," and many more (Gelis, & WoodsT., n.d.). All of these factors provide an incredible amount of opportunity for Vestigo Ventures and Cogo Labs, and together, they are able to leverage their professional expertise in order to find and fund the future of FinTech.

1.5 The Projects

Vestigo Ventures identified three specific projects. The titles of these projects are "Venture Decisions Measured Over Time," "Deal Analysis & Industry Engagement," and "Signal Search & Market Landscape Analysis." The intent of these three projects is to tap into the entire work process at Vestigo Ventures in order to streamline it and provide recommendations for potential improvements.

Cogo Labs identified one specific project. The title of this project is "Device Fingerprinting." The intent of this project is to utilize Cogo's in-house analytical tools in order to analyze a specific database of collected internet-user information and provide an improved model for user-profile building.

Each project will be discussed in deeper detail in their respective sections later in this report. Based on the academic requirements and project logistics for each team, the undergraduate team worked full-time on these projects on-site at Cogo Labs during the B-2017 term, while the graduate team worked part-time and off-site for the most part, with weekly visits to the Cogo Labs Cambridge offices by at least one graduate team member.

2. Methodology and Theory

During the preliminary stages of discussions with our academic advisors, and after we received the full project descriptions from both Vestigo Ventures and Cogo Labs, it was up to us to learn, plan, and arrange the best strategy to accomplish all deliverables in the limited timeframe of B-2017 term. The following section describes our project planning decisions between our team and the undergraduate MQP team; it was the undergraduates (comprised of six members) who were primarily responsible for the implementation and development of the projects. We also explain our reasoning and theory for each respective project, as well as any algorithmic model.

2.1 Project Planning and Management

2.1.1 Resources and Team Setup

At first, in order for the graduate team members to serve as mentors to the undergraduates during their project development period, we had to learn all tools and software that were made available to us. It's important to note that all related computer accounts, proprietary programs, and tables are properties of Cogo Labs. During the first onboarding day and follow-up visits, we focused on understanding these tools and websites for our project purposes. Additional preliminary preparations included understanding the key skill sets that each individual student (both graduate and undergraduate) could bring to the projects.

Among the graduate team, each member carried a very specialized and expert mastery in their respective majors of studies. The Masters graduate team was comprised of an Information Technology student with experience in software and technologies while working at the National Stock Exchange of India, a Marketing Foisie Business School (FBS) student with a background in business decision making and market segments, and two Financial Mathematics students who had statistical knowledge, mathematical theory, and an understanding of data science techniques.

Among the undergraduate team, there were six students with a variety of majors of studies: Management Information System & Management Engineering (Double Bachelors),

Mathematical Sciences, Electrical Computer Engineering, and three Computer Science. We determined and agreed as a team that the MIS & MGE student, Andrew Aberdale, be set as the main manager of the undergraduates and responsible in assigning tasks to the development subteam (each subteam had a fair mix of CS and non-CS students). This way, we could communicate directly to Andrew and understand their status and any questions regarding our proposed models and theories.

2.1.2 Timeline

Early in the project planning process, we decided to create a timeline to visually represent what we envisioned, at least at first, to be the deadlines and expectations for the Vestigo and Cogo projects. This was helpful not only for us but also to the advisors and sponsors to understand the proposed workflow. We believed that the proposed deadlines were appropriate given the short-duration of the project period. Once the undergraduates arrived at the offices on October 23 and after their onboarding process, they immediately went to work and finished implementations for the first two Vestigo Venture projects. Regarding the Cogo Labs project, the undergraduates only delivered a recommended algorithmic model and test implementation; this was due to the Cogo Lab project's complexity and the project period's time constraints.

Figure 1: Initially proposed timeline for project workflow.

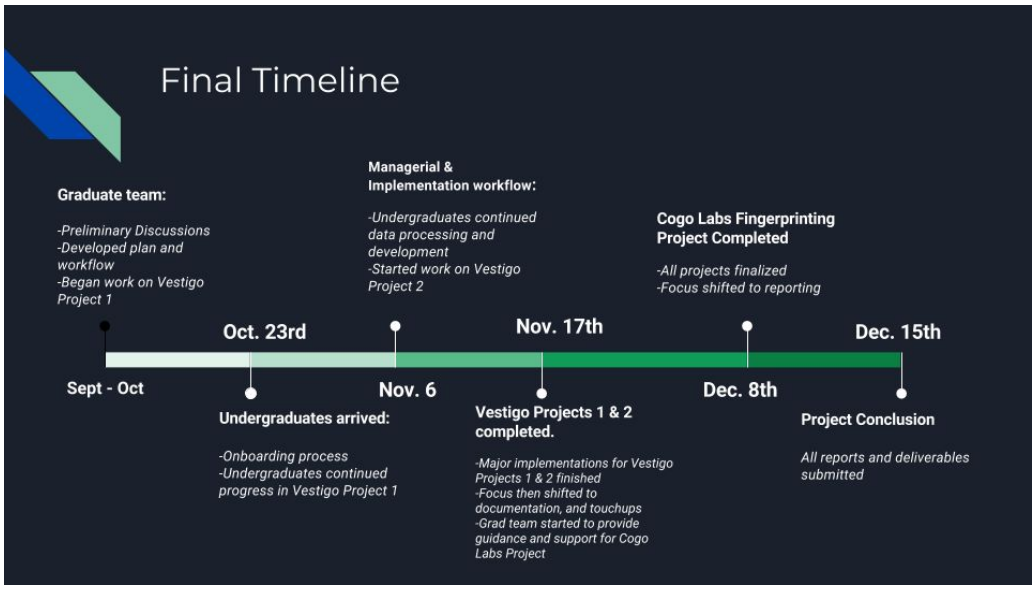

Figure 2: Final deliverable timeline for the projects.

2.1.3 Work Progress Tracking

In order to maintain the progress of each individual action item and tasks to accomplish, we decided to use the well-known GitHub website. While GitHub is known for its user-intuitive website interface of the Git version control feature, we discovered a relatively recent website addition that proved to be very useful for project management of this complexity. In September 2016, GitHub permanently implemented a new feature called Projects; the feature is described as a way to "manage work directly from your GitHub repositories … create cards from Pull Requests, Issues or Notes and organize them into custom columns, whether it's *In-Progress*, *Done, Never going to happen,* or any other framework your team uses" (GitHub, 2016). A comparable tool that other software engineers may be familiar with is called *Trello*, a "web-based project management application" that was bought by Atlassian in January 2017 (Trello, n.d.). All important members of Cogo, Vestigo, graduate team, and undergraduate/MQP team had access to the GitHub repository, to promote teamwork. This feature was crucial in fully understanding the set of tasks that were new/in-progress/completed as well as associated comments and notes for the tasks.

To further maintain order within each project, our managerial structural took the initial shape of the us leading and giving the first directions, while undergraduates followed along. The undergraduates and Andrew then evolved it a step further where each undergrad team member would take charge of multiple action items, and become the designated “specialist” or leader for that topic. It was then their responsibility to ensure the task got completed, with or without aid from the team, depending on their preference. If progress was made for an objective, the leading team member would update the GitHub and Task to reflect the updated development. Likewise, to keep the team on pace, Andrew Aberdale had fewer of “specialist” assignments when it came to bigger projects and instead took smaller yet critical roles. Anytime an objective was completed, Andrew would update the GitHub and kept the team updated with changes relevant to their development or the projects. A screenshot of an example view of our Projects arrangement can be found below:

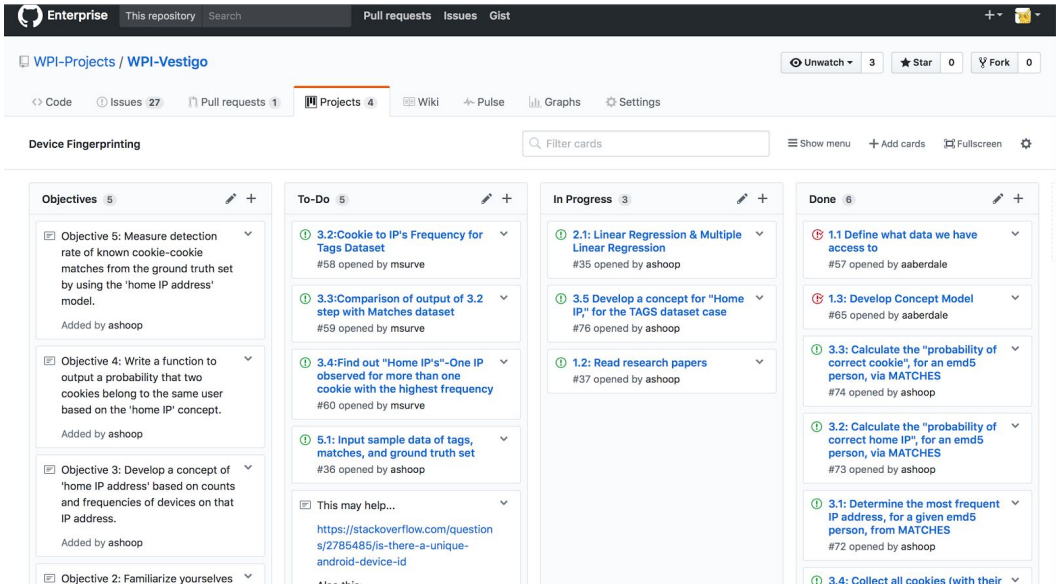

Figure 3: Screenshot of GitHub Projects progress tracking workflow.

2.2 Theory behind Vestigo Project 1: Venture Decisions Measured Over Time

Vestigo has received approximately 270 Fintech startup investment presentations since fund inception - by now it has likely increased dramatically. They reportedly were receiving an

average of 3 to 5 investment deals per week from entrepreneurs across North America seeking early stage investment. Based on the factors like startup's objective, fundraising requirement, valuation and virality, the Vestigo sponsors then decide to either invest or pass on a company. Vestigo also considers additional factors of a startup such as whether it has a strong and committed management team, whether it has a high growth potential in the market it seeks to serve and the return of investment. A "passed" deal in Vestigo's context means that they decided to leave the deal behind and not invest in a company as of now. The goal of this project was to analyze startup investment deals Vestigo had previously passed on using current data to determine how the deal did overtime.

For this project, we primarily performed a role of Business Analyst in which we coordinated with Vestigo team to understand their business and discuss the requirement. We decided ways to convert the business requirement to technical one. After this business analysis phase, we moved on to the data analysis stage in which we gathered all the required data and decided ways to clean-up, store and analyze this data in order to help Vestigo Sponsors in making informed decisions.

2.2.1 Cataloguing the deals

The first requirement of Vestigo was to catalogue information of all the deals received by Vestigo Sponsors. We decided to maintain a repository first in a spreadsheet and then to the database. Currently all the deal related emails come to Vestigo sponsor's (Ian/Mike/Mark) official email account which they forward to a shared email inbox account for consolidation; (referred to as "Dealflow" from here onward). Ian/Mike/Mark create a folder for every new company in the Dealflow and push the emails from Inbox to company's folder for future tracking. We were given an access to this Dealflow account for reading the deal related emails and gather the appropriate information. Whenever someone needs a funding from Vestigo for their startup, they send a pitch deck to Vestigo which contains brief information about their business and their funding requirement. We went through all these pitch decks to gather

information like company description, website, contact person, contact title and email-id, first date of contact and name of the Vestigo's sponsor who is handling the deal. We captured information about the current stage/round of funding of the startup. These rounds were Seed stage which is typically the very first stage of investment where money used for market research and developing product generally comes from personal and family savings), and Series A, Series B rounds which are Venture Capital Financing Rounds etc. (Hollis, n.d.). We also noted the amount required by the startup and its pre-money valuation. Pre-money valuation refers to the valuation of the company before the investment (Davie, 2017). It indicates what a startup company expects its valuation will be before the investment. Vestigo uses a pre-money valuation to analyze the deal and to decide whether the startup is worth the investment. We referred websites like Crunchbase, PitchBook, and FormDS for getting clear information about amount required, amount raised and pre-money valuation of the startups. Once we gathered all the data in a spreadsheet, we wanted to find out the a better option for cleaning, storing and accessing the current and future data. The manual cleaning and standardization of data which was entered in a spreadsheet required more efforts. Also, the storage requirements would increase in future with more number of records in a spreadsheet. As Vestigo has an access to Cogo's proprietary database, we decided to load the information in a database which can be easily accessed and analyzed using SQL. We created "wpi_deal_flow" table and handed over the spreadsheet to carry out further cleaning, formatting and loading process.

2.2.2 Analyzing the passed deals

As of now, Vestigo has invested in 3 companies and passed the rest of the deals. Vestigo wanted to analyze the performance of deals which they had passed in order to find out some good deals on which they can have a look again and may alter their decision. Cogo's large consumer dataset is connected to their proprietary tool called Apollo which has the ability to detect almost to the moment when a FinTech startup shows signs of "going viral". Apollo methodically detects viral growth and generates a real-time graphic which indicates confidence

levels of various domains to become viral. As the confidence level indicates the likelihood or the prominence of the domain to become viral, we decided to use these confidence levels to analyze the startups. Apollo captures the day wise proportion of early detection also known as the influential population proportion visiting the domain. This graph tends to rise at the early stage of the domain when the early detectors find about the new technology and visit the new domain. Apollo also captures the proportion of random population accessing the domain whose slope is generally low at the initial period and then it tends to increase as the general population starts to follow the things what early detectors are doing. The confidence level generally shows the aggregate of above two graphs which gives the overall picture of the virality of that domain.

Vestigo uses Cogo's Apollo interface as one of the key indicator for analyzing the deal. Whenever a new deal is received, Vestigo sponsors manually run that domain in Apollo to see the confidence level of that domain. A continuous increase in the confidence level is considered as a plus for that domain. We decided to build a robust method which Vestigo sponsors can use in future instead of manually looking at Apollo graphs to make a decision. The confidence level values are stored in Cogo's proprietary database. We designed an algorithm to capture the confidence levels, determine the percentage change and decide the best deals (See Appendix - Algorithm for Vestigo Projects). The undergraduate team implemented this part using a SQL and Python and automated the process to avoid Vestigo's time to manually check the Apollo graph. After deciding the best deals based on the confidence level, we recommended undergrads to look at the amount required/raised fields stored in wpi_deal_flow table to find out whether Vestigo still has a scope of investing in those companies. As a deliverable of this project, we presented the top 20 companies with the highest increase in confidence level and still having a scope of investment to Vestigo.

2.3 Theory behind Vestigo Project 2: Deal Analysis and Industry Engagement

Vestigo Project 2 aims at analyzing deals by considering both the deal itself and the industry situation. We considered several factors during the process of deal analysis, including market segmentation, business model type, market research, the need of digital transformation, and user experience.

2.3.1 Market Segmentation and Business Model

Wendell R. Smith (1956) raised a concept, market segmentation, several decades ago. In his paper published in the Journal of Marketing, dividing a heterogeneous market into a number of different and smaller markets by characteristics like customer preferences and demographics can benefit companies by bringing more profit (Smith, 1956). It also benefits customers by satisfying their various needs. For venture capital firms like Vestigo Ventures, different types of financial technology startups are raising funds from them.

In order to be more efficient on customer data analysis and comparison, it is necessary to conduct market segmentation. Generally, Business-to-Business (B2B) companies tend to segment its market by business types or countries, while Business-to-Customer (B2C) companies tend to segment its market by customer interest, customer behavior, demographic characteristics, or any other meaningful segment. To be more specific, B2B refers to commerce between business and business, B2C refers to commerce between business and customer. Since venture capital aims to invest in FinTech startups with high potential of growth, Vestigo segments use these startups' business types to make it more efficient and easy to be compared.

There was one confusion when deals were getting segmented. What if a startup could not be simply categorized as B2B or B2C? A new type of business, Business-to-Business-to-Consumer (B2B2C), drew our attention immediately. The concept B2B2C was first raised because of the emergence of Alibaba, a Chinese e-commerce website that provides both B2B, B2C, and even C2C services (Zhang & Zhang, 2011). B2B2C is a business model that combines B2B and B2C to complete the whole product/service transaction (Ragins & Greco, 2003). As an example of applying this theory, Nike Inc. pays Alibaba for using

its e-commerce website for providing a place to present their product and attracting potential customers. So far it is still B2B. Then, individual customers buy the Nike products sold on Alibaba's website and pay for these products directly on Alibaba. The business model between Alibaba and Nike's customer is B2C. Together, this type of business model groups product provider, distributor, and end user (customer) and thus provides an efficient and beneficial business ecosystem (Juan & Deixiong, 2013).

It is seemingly the B2B2C type of businesses that have contacted Vestigo Ventures for investment. However, they have a slightly different situation and thus could not be categorized simply as B2B, B2C, or B2B2C according to these business models' definition. For example, if company A is providing a 401K plan and retirement plan to both small business B and individuals, company A is targeting both end users (customers) and businesses. When employees' situations we considered when company B makes a transaction with company A, or when an individual purchases a product/service from company A, even though it sounds like two different types of business, we decided to group these two types of businesses into the same segment.

2.3.3 Market Research (Survey)

We identified the use of information as an important factor of a company's competitive advantage and developing direction (Moorman, Deshpandé, & Zaltman, 1993; Kohli & Jaworski, 1990). In this big data era, data is relatively easy to be obtained. How to organize data to make use of it, however, is a key question to be considered. As mentioned in this report, Vestigo Ventures received around 3 to 5 new deals each week. Based on the consideration of data security, only senior managements of Vestigo Ventures have the access to this deal data. This method slowed the investment progress and left room for typos and outdated information. Therefore, to make the deal data collection more efficient, we started to think about possible software, IT solutions, or any other recommendation to accelerate the data collection speed while keeping the data safe. In other words, we needed to conduct market research by creating a survey form for potential investors and collect deal data directly from it. Also, understanding the market

need, customer need, and any other business partners involved in the same investment are very important.

Compared to the Internet and other new industries, although the finance industry has more strict regulations, it still has a high potential to revolutionize. As Geoffrey G. Parker mentioned in his book *Platform Revolution* published in 2016 (Parker, Van Alstyne, & Choudary, 2016), increasing kind of technology invented in the finance industry such as blockchain and P2P lending and investment would make possible the revolution of financial service, which would lead to a new round of industry development.

2.3.4 Digital Transformation

Digital transformation is inevitable for a business to remain competitive (ProQuest, 2015). As the CEO of Software AG said at Software AG's Innovation World 2015, "going digital is a new way of doing business and it is much more than just an technical project in a company". Digital transformation is the change in business that associated with the use of digital technology (Boolard, Singla, Sood, & Ouwerkerk, 2016). When Vestigo decided to transfer from manually recording deal data to looking for digital methods of doing this. This is actually the improvement of business process management, which is one of the digital business transformation areas raised in an article by Professor Michael Wade (2014). This digital transformation could change the dynamics of competition in finance industry.

The main method of collecting data changes over the history. Methods such as face-to-face interview, paper based questionnaire, and telephone based surveying in the 1970s are developing to internet based surveying in recent years (McDonald, Mohebbi, & Slatkin, 2012). After considering data quality, data collection cost, and data timeliness, we first considered Google Forms. Unlike other traditionally designed questionnaires, Google Form uses a personalized quiz or survey to collect information from targeted audiences (Google, 2016). The information collected is then automatically recorded in a spreadsheet. The advantages of using this form include its convenience, efficiency, custom themes, and it is easy to use (Google, 2016). Therefore, for purpose of data storing, Google Forms can be a good choice.

However, technology is changing, which leaves room for cyber security concern. Whether the information security issue of Google product could be well predicted and prevented is still unknown. In this stage, taking the security issue, issues related to the data accessibility, and the skill set that our Graduate team and Undergraduate MQP team possess into consideration, we decided to move away from Google platform to create a form from scratch and connect this form to Vestigo's CRM system.

2.3.5 User experience and Survey form design

Adding a survey form to Vestigo's website and creating a link to direct entrepreneurs to the survey web page have several advantages. First, it makes entrepreneurs and Vestigo much closer. Before this e-form was built, in order to raise funding, entrepreneurs need to first find a way to have Vestigo executives' contact information. Then, they would either call/meet those executives directly or send an email with succinct but limited information. This kind of approach is not efficient enough, especially in the fast growing FinTech industry. With a survey form and a web link, Vestigo can now collect both required information they need to assess a deal and optional information that could increase an entrepreneurs' chance of getting fundraising.

Then, we also considered how we can increase end-user's experience. User Experience (UX) refers to an end-user's emotion and attitude when he/she interacts with a product/service/company (Hassenzahl & Tractinsky, 2006). According to Dr. Soussan Djamasbi's user experience research, it is supported that Generation Yers, the generation of people born during the 1980s and 1990s, will show significantly less preference for a large amount of Web components compared to baby boomers (Djamasbi, Siegel, Skorinko, & Tullis, 2011). Since the majority of the entrepreneurs who approached to attract investment is Generation Yers, we decided to keep the survey form neat and simple.

In our survey form design, apart from simple text entry questions, we used a dropdown list as well to make this form more user friendly and attachment function for entrepreneurs to present more information before the first conversation was set up between Vestigo and them.

2.3.6 Deal Evaluation

As a major part of Venture Capitals' business, it is very important to analyze deals and be able to decide which deal to invest in and which deal to pass at the moment. There are several factors that need to be looked at closer while we are analyzing a deal. First, what sort of problem could be solved by using this company's product/service. Is this problem really hitting customer pain point?

Second, how large is the potential market for this opportunity? Before a product is officially launched in a market, it is relatively subjective to predict the market size. To do this, we decided to take business type (B2B or B2C), geographic area (regional, national, or international), the market share this company will be able to earn (competition analysis), and the penetration rate into consideration. Let's use a new golf ball manufacturer headquartered in Shanghai, China as an example. If we want to predict its market size in each year, we need to consider the population size of Shanghai, the need of golf balls from golf clubs (B2B business), the need of golf balls from hobbyists who don't join a golf club (B2C business), and the number of golf balls each golf club and hobbyist use/keep each year.

Apart from the market size and business model type, the market share this new golf ball manufacturer can earn is also an important factor. Since this golf ball market is not new, assuming that after market research we realize incumbents in this industry already take around 85% of market share, this new golf ball manufacturer might only be able to take under 15% of the golf ball market in Shanghai. Moreover, both golf clubs and hobbyists might not replace 100% of their golf balls each year. Assuming 50% of the ball would be replaced with newly purchased ones, the market size should be just 50% of the market size we predicted before taking this replace rate into consideration. After few years of running experience, the business life cycle and the industrial growth rate can also affect this business's performance. This golf ball case is a regional and traditional example. But, it is similar when a FinTech deal is being evaluated and should also be assumed and considered.

Third, team factors are also a key aspect when evaluating a deal (Bygrave & Zacharakis, 2009). For example, the management managing the company, their background, their

qualification to run a FinTech business, and also their experiences on delivering credible and receivable product/service to meet the market need.

Fourth, what is the business plan? For FinTech companies, their use of technology is one of the key success factors. Questions about using a newly developed technology and applying a certain technology's significance on improving financial activities, automated insurance, trading, and risk management (Härle, Havas, & Samandari, 2016) need to be considered.

Last but not least, we need to understand the financial meaning behind each deal (Roure, 2013). Both the outside market and the worthiness of investing in this FinTech company should be properly considered. More specifically, the scalability of business concept for investor to predict potential growing revenue, the possible exit for the investors to recover their investment with the expected return, the potential co-investors of a targeted deal, and consistency between the return provided by a deal and the level of risk implied to investors should be examined.

2.4 Theory behind Vestigo Project 3: Signal Search and Market Landscape Analysis

Vestigo Ventures' main focus is to find out the best-of-the-best Fintech companies for investment. Currently Vestigo uses Apollo's ability to predict the virality of the domain based on the confidence level. Cogo's large consumer database has the ability to capture how many early detectors and general population visit the particular domain which then integrated with Apollo to display the prominence of that domain. This logic fits well in the case of B2C companies as the end users in this case are the customers whose clicks details are captured by Cogo correctly. The capability of capturing Business to Business signals by Apollo is limited; according to us, the end user for B2B is a company itself, so it is actually impossible to capture the internet browsing data of a "company." We decided to look at the Apollo algorithm for B2C and build an algorithm which will be valuable for B2B. As this project was largely experimental in nature with an undefined final deliverable, Vestigo Ventures' Mike Nugent suggested to put this project on hold until completion of all other projects.

2.5 Theory behind Cogo Labs Fingerprinting Project

Cogo Labs has 3 petabytes of consumer profile data consisting of ~800 million profiles. 1 petabyte is approximately 1,000 terabytes which is about 1,000,000 gigabytes. To paint a clearer picture, one can describe 1 petabyte as 20 million four-drawer filing cabinets filed with text, 2000 years of listening MP3-encoded songs, and the size of local storage of the renderings of the 3D-CGI effects from the movie "Avatar." The dataset of Cogo Labs covers about 80% of the United States populace. Using this database and an extensive proprietary cookie match network, Cogo has a 30% accuracy rate to identify if a user hits their websites or their partners' websites. This allows Cogo Labs to tie more information together about a user and better personalize his/her experience. As most end-users have encountered already, some devices and browsers by default do not accept cookies or put limitations on their use, which makes targeted outreach and a personalized experience difficult. However, these devices and browsers can be uniquely identified (to a certain degree of confidence) using other browsing signatures such as user agent, IP address, browser version, etc. This is a technique known as "device fingerprinting," and originated about 10 years ago (Tadayoshi et al., 2005). More recent developments include a research project by the Electronic Frontier Foundation (P. Eckersley, 2010) and a Microsoft Research paper on "Host Fingerprinting" (Ting-Fang et al., 2012); We derived most of our understanding and theory for this Cogo Labs project from these research papers.

Cogo uses the MD5 hash of an individual's email address (appropriately called "emd5") in their process of matching cookies with users. This way, there should be no personally identifiable information (PII) in regards to privacy protection. In the cases that an emd5 cannot be traced to a cookie, other browsing signatures facilitate detection. Such signatures include user agent (the physical device used), IP address, and browser version. The objective of this Cogo Labs project was to increase Cogo's match accuracy rate from 30% to 45%, and to make the system compatible with Safari browser detection. It was up to us to formally write-out the proposed algorithm and model for the undergraduates to then develop. Much of the following material was derived from the methods outlined within the research paper from Microsoft. The

undergraduate created a final output table named “wpi_matches” to calculate the probability that a specific cookie matches to a specific user. Firstly, we define some background terminology.

2.5.1 Terminology

The Internet Protocol (IP) address is a numeric address (a typical IPv4 example address would be 130.215.15.XX) that is assigned to every single hardware device that is part of a TCP/IP-based internet network (Techopedia, n.d.).

A cookie, more formally an HTTP cookie, are small files that are stored on personal devices such as computers and smartphones. Cookies are designed to “hold a modest amount of data specific to a particular client and website, and can be accessed either by the web server or the client computer” (Cookies, n.d.). These allow the server to deliver a more tailored/personalized page for each particular user; cookie details can also contain carried information from one website to another.

The User-Agent (UA) is one of the text headers contained in HTTP requests that holds a lot of useful information including, but not limited to: computer operating system, internet browser type and version (User-Agent, n.d.). A cookie can hold this UA information.

2.5.2 Home IP

One of the first steps to create the fingerprinting probabilistic model was to design a concept that would tie a user to a particular cookie. Only about 14% of the given database had a user tied to a cookie, hence the initial goal was to tie the rest of the data to a particular user already in the database. We understood that all users and cookies in the database correspond to a particular IP address, so we decided to tie users to cookies through IP addresses. Most importantly, we developed the concept of a “home IP address.” We proposed the following definition:

Home IP address : The most frequently occurring IP address for a particular cookie

2.5.3 Probabilistic model

After choosing a Home IP address for each cookie to analyze, then came the challenging task of figuring out an approach to create a probabilistic model to determine the probability that a specific cookie matches to a specific user (whether emd5 or Home IP). The key concepts used here are independent and dependent probability. Cogo Labs supplied us the available source tables; we will call them "Matches" and "Adrequests" for the remainder of this paper.

For our project, we were under the assumption that we can always accurately identify someone based on the cookie. We wanted to use the concepts from the Microsoft Research paper and its associated "precisions," ie. the frequency probability calculations. These were developed and put in place incase Cogo Labs decides there is a different associated percentage chance that can accurately identify a user based on a specific field. Therefore, the "probabilistic coefficients" used in our model can be found below:

	Cookie Precision (C)	User Agent Precision (U)	IP Precision (I)
Assumed Accuracy (Current implementation)	1	1	1
Microsoft Paper	0.8235	0.6201	0.4899

After gathering all entries with the same Home IP, we calculate the "cookie independent home IP probability" (shortened as cookie_ind_hip_prob). This is the probability that a certain IP address belongs to the designated Home IP:

$$\text{cookie\_ind\_hip\_prob} = \frac{(\text{Frequency count where the specific IP}{=}\text{HomeIP}) \times \text{C} \times \text{I}}{\text{Total count of Home IP}}$$

Upon assigning each cookie a Home IP address, we decided to develop the final cookie probability concept, that would express how confident the model is that a given cookie belongs

to the user (Home IP) specified. We derived the formula below to calculate cookie probability for cookies with the same Home IP address:

$$Cookie\_Prob = 1-(1-X)*(1-Y)*(1-Z),$$

Where X, Y, and Z are calculated by:

X = C * Cookie_Frequency, Y = U * UA_Frequency, Z = I * IP_Frequency

The "variable_Frequency" related formulas are the relative frequency probabilities for each variable, across a specified Home IP:

$$\text{Cookie\_Frequency} = \frac{\text{Frequency count where the specific Cookie=Cookie}}{\text{Total count of HomeIP}}$$

$$\text{UA\_Frequency} = \frac{\text{Frequency count where the specific UA=UA}}{\text{Total count of HomeIP}}$$

$$\text{IP\_Frequency} = \frac{\text{Frequency count where the specific IP=IP}}{\text{Total count of HomeIP}}$$

To further explain the reasoning behind the above final Cookie_Prob formula, say we are focusing on the Z variable. By definition, this means "the probability that a specific IP belongs to a given Home IP." Therefore, 1-Z would be the probability that an IP address does NOT belong to a given Home IP. We gathered the three factors of "probability of NOT matching to a Home IP" (cookie identity itself, UA, and IP address), multiplied them together, and took the "1 - sum" result of this probability to get the final probability result.

All of the above implementation, development, and sample test results can be found in the undergraduate's MQP paper (Aberdale et al., 2017) as well as their explanation for their variable "cookie dependent home IP probability (cookie_dep_hip_prob)."

2.5.4 Algorithm model table processing workflow

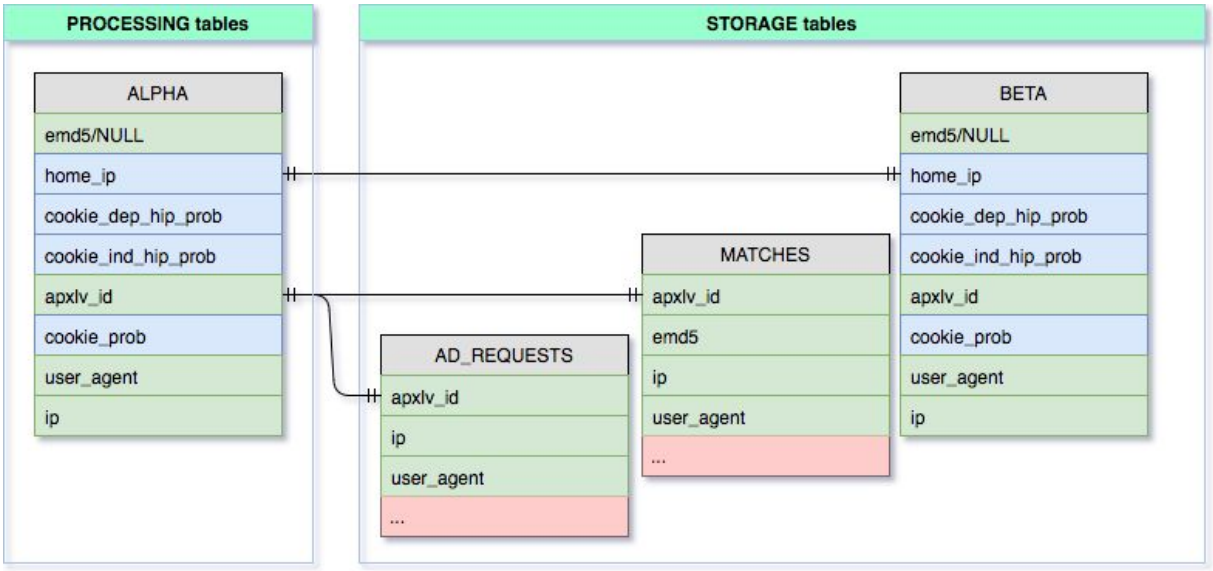

Figure 4: Diagram of proposed fingerprinting project algorithm model (credit to Zhi Hui).

The diagram above shows the data source tables (Matches, Adrequests), Home IP and probability calculation processing table (here called Alpha), and the final storage table (here called Beta). There was a limitation why we could not process and store at the same time in the same type of database table. This complexity and further technical explanations can be found in the undergraduate's MQP paper (Aberdale et al., 2017).

3. Discussion and Conclusions

Ultimately, this project covered venture capital fundraising, big-data analytics, teamwork, and identifying best project management practices, and we all learned an incredible amount in all of these areas. For all of us, it was the first time being exposed to the venture capital and big-data analysis world, as these things cannot really be taught in the classroom, and we are really grateful for this opportunity. Besides getting a very interesting glance into the world of venture capital, we had many additional lessons. We learned the importance of clean & seamless communication between fundraisers and entrepreneurs, and understood where and when

breakdowns in communication can happen and how to mitigate those issues. We also learned about the implications and stories that one can deduce from looking at the results via virality search and segment analysis.

The project structure and layout for this Independent Study Project was unique in that there was no dedicated company correspondent or manager who would guide us through the planning, analysis, and implementation of the projects. Our team was responsible in performing this liaison role. Although it was difficult to initially figure out the best workflow and set of steps to achieve all deliverables, after our first onboarding day and several visits to the office, we developed a plan for both teams to follow throughout the remainder of the project term. As time went on, this plan was refined and adapted, and ultimately paved the way to successfully complete our defined objectives. Working through this process of adapting and refining a project plan in such a unique environment was a remarkable experience for all of the graduate team members, and will undoubtedly help us all in any future project management work.

In Vestigo Ventures project 1, we knew that since the initial email contacts and deal flows were manually recorded by hand that this project would have plenty of room for automation. After setting up the large data spreadsheet and handing it off to the undergraduates, we asked that they not only finish cataloguing the deals but to also develop a front-facing survey or online form that gathers the data and automatically collects it all in a digestible database. We felt that this would significantly improve the deal flow process at Vestigo Ventures, by drastically improving the speed and efficiency of deal flow data collection, as well as greatly reducing the potential for human error. The undergraduate team followed our guidelines but refined the project goal along the way, and ultimately exceeded all expectations and impressed the Vestigo Ventures sponsors.

For Vestigo Ventures project 2, we went through a whole process of deal analysis and industrial engagement. By transforming a traditional method of data recording to an ideal IT solution specifically for Vestigo Ventures, we had a chance to understand the importance of digital transformation and the unique need of Vestigo Ventures. In addition, we started to think of a product/deal in the context of other disciplines such as tech team, market size, and so forth.

During our work, it became clear to us the importance of taking into account different perspectives - technical and business oriented - and different stakeholders.

Due to the complexity of the work and the short project duration, the sponsors at Vestigo Ventures deemed it appropriate to not pursue work on Vestigo project 3, but rather focus time and efforts into fully flushing out and completing Vestigo projects 1 & 2 in addition to the Cogo Labs Fingerprinting project.

Throughout the entire project period, we kept close contact with the undergraduate development team; this is especially true during the brainstorming and implementation work for the Cogo Labs Fingerprinting project. During the development of the algorithm, we worked closely with the undergraduate team, several Cogo Labs employees, and Professor Jon Abraham of the Mathematical Sciences department to ensure that we had a solid mathematical foundation for the probabilistic model. As a result of this fingerprinting work, we learned the importance of understanding all of the variables in a dataset, particularly in those as large as the ones used at Cogo Labs. Additionally, we learned the importance of providing sound justifications for algorithm development, most notably in environments as open and collaborative as Cogo Labs. In order to ensure that our work will be understandable to future readers, we created in-depth documentation detailing all of the steps in our model. This took several iterations and rounds of discussion, and was ultimately a good lesson in real-world algorithm development.

This project was an invaluable collaboration opportunity between WPI and Vestigo Ventures and Cogo Labs that we hope will continue moving forward. We were very impressed with the professional results that the undergraduate MQP team provided, and we are very grateful for the unique educational experience we received through this collaboration. We hope our work and the results of the undergraduate MQP team have laid the foundation for future project opportunities with Vestigo Ventures and Cogo Labs.

Bibliography

Aberdale, A., Guerrero, J. C., Hui, Z., Lima, U. P., Sarkis, B., & Tafesse, R. (December 17, 2017). Big Data Analytics and Tactical Decision Making [Scholarly project].This report will be published on WPI's online project repository in the future.

Bollard, A. (2016, October). Transforming operations management for a digital world. Retrieved January 17, 2018, from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/transforming-operations-management-for-a-digital-world

Browne, R. (2017, October 04). Everything you've always wanted to know about fintech. Retrieved January 17, 2018, from https://www.cnbc.com/2017/10/02/fintech-everything-youve-always-wanted-to-know-about-financial-technology.html

Business to Business to Consumer (B2B2C) - Definition from Techopedia. (n.d.). Retrieved January 17, 2018, from https://www.techopedia.com/definition/23169/business-to-business-to-consumer-b2b2c

Bygrave W., Zacharakis A. (2009). The Portable MBA in Entrepreneurship. Wiley, 4th Edition, ISBN 978-0470481318

Cookies - Computer Cookies Explained. (n.d.). Retrieved January 17, 2018, from http://www.whatarecookies.com/

Crunchbase - Funding Rounds defined. (n.d.). Retrieved January 20, 2018, from https://www.crunchbase.com/search/funding_rounds

Davie, A. J. (2017, May 16). Pre-Money Valuation vs. Post-Money Valuation. Retrieved January 20, 2018, from https://www.strictlybusinesslawblog.com/2017/05/16/pre-money-valuation-post-money-valuation/

Djamasbi, S., Siegel, M., Skorinko, J., & Tullis, T. (2011). Online Viewing and Aesthetic Preferences of Generation Y and the Baby Boom Generation: Testing User Web Site

Experience Through Eye Tracking. International Journal of Electronic Commerce,15(4), 121-158. doi:10.2753/jec1086-4415150404

Edna, J. R., & Greco, A. J. (2003). Customer relationship management and E-business: More than a software solution. Review of Business, 24(1), 25-30. Retrieved from http://ezproxy.wpi.edu/login?url=https://search.proquest.com/docview/220966899?accountid=29120

Fintech. (n.d.). Retrieved January 17, 2018, from https://www.mckinsey.com/global-themes/fintech

Gelis, P., & Woods, T. (n.d.). The Rise of Fintech in Finance. Retrieved from http://cdn2.hubspot.net/hub/310641/file-1445626583-pdf/Rise_of_Fintech_in_Finance/Fintech_DEF.pdf

Gensler, L. (2017, October 04). The 10 Biggest Fintech Companies In America. Retrieved January 17, 2018, from https://www.forbes.com/sites/laurengensler/2017/08/08/biggest-us-fintech-companies/#5d741c3f59d8

GitHub. (2016, September 14). A whole new GitHub Universe: announcing new tools, forums, and features. Retrieved January 17, 2018, from https://github.com/blog/2256-a-whole-new-github-universe-announcing-new-tools-forums-and-features

Google Forms. "Create a survey using Google Forms". Docs editors Help. Google. Retrieved October 30, 2016, from https://support.google.com/docs/answer/6281888?co=GENIE.Platform%3DDesktop&hl=en

Google Forms. "Custom themes in Google Forms". G Suite Updates. Google. September 2, 2014. Retrieved December 12, 2016, from https://gsuiteupdates.googleblog.com/2014/09/custom-themes-in-google-forms.html

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. Behaviour & Information Technology,25(2), 91-97. doi:10.1080/01449290500330331

Harle, P., Havas, A., & Samandari, H. (2016, July). The future of bank risk management. Retrieved January 17, 2018, from https://www.mckinsey.com/business-functions/risk/our-insights/the-future-of-bank-risk-management

Hollis, M. (n.d.). Accounting Resources. Retrieved January 20, 2018, from https://www.indinero.com/resources/startup-funding-rounds-roadmap

IP address - Definition from Techopedia. (n.d.). Retrieved January 17, 2018, from https://www.techopedia.com/definition/2435/internet-protocol-address-ip-address

Kohli, A. K., & Jaworski, B. J. (1990). Market Orientation: The Construct, Research Propositions, and Managerial Implications. Developing a Market Orientation,1-18. doi:10.4135/9781452231426.n2

Li, Z., & Xiaolong, L. (2013). The B2B2C Integrated E-Commerce Model Design Based on SaaS. Liss 2012,1123-1127. doi:10.1007/978-3-642-32054-5_158

Link Ventures: The Link Up - Monthly Newsletter. (2017, October 3). Retrieved January 17, 2018, from https://www.abovo42.com/sean@abovo42.com/52326

Marchand, D. A., Wade, M., Sands, L., & Plueckebaum, A. (2014). Digital Business Transformation. Insights@IMD,41. Retrieved from https://www.imd.org/research/publications/upload/41-Digital-Business-Transformations-03-09-14-NEW.pdf

McDonald, P., Mohebbi, M., & Slatkin, B. (n.d.). Comparing Google Consumer Surveys to Existing Probability and Non-Probability Based Internet Surveys. Retrieved from https://www.google.com/insights/consumersurveys/static/consumer_surveys_whitepaper.pdf

Meisler, L., Rojanasakul, M., & Diamond, J. S. (2016, May 25). Who Gets Venture Capital Funding? Retrieved January 13, 2018, from https://www.bloomberg.com/graphics/2016-who-gets-vc-funding/

Moorman, C., Deshpandé, R., & Zaltman, G. (1993). Factors Affecting Trust in Market Research Relationships. Journal of Marketing, 57(1), 81-101. doi:10.2307/1252059

Network Asia. (October 14, 2015). Digital transformation is imperative for a business to remain competitive. Retrieved January 17, 2018, from https://www.networksasia.net/article/digital-transformation-imperative-business-remain-competitive.1444789571

Parker, G., Alstyne, M. V., & Choudary, S. P. (2017). Platform revolution: how networked markets are transforming the economy - and how to make them work for you. NY, NY: W.W. Norton.

P. Eckersley, "How Unique Is Your Browser?" in Proceedings of the 10th Privacy Enhancing Technologies Symposium (PETS), 2010. https://panopticlick.eff.org/static/browser-uniqueness.pdf

Roure, J. (2013, March 12). Ten Key Questions when Evaluating an Investment Opportunity: the Business Angel Approach. Retrieved January 17, 2018, from http://blog.iese.edu/entrepreneurship/2013/12/03/ten-key-questions-when-evaluating-an-investment-opportunity-the-business-angel-approach/

Ryan, G. (2016, December 8). New fintech VC firm Vestigo invests in Boston software maker. Retrieved January 17, 2018, from https://www.bizjournals.com/boston/news/2016/12/08/new-fintech-vc-firm-vestigo-invests-in-boston.html

Smith, W. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. Journal of Marketing, 21(1), 3-8. doi:10.2307/1247695

Staff, I. (2017, September 11). Venture Capital. Retrieved January 17, 2018, from https://www.investopedia.com/terms/v/venturecapital.asp

Startup Funding Rounds: Funding Life Cycle. (n.d.). Retrieved January 20, 2018, from https://fundingsage.com/startup-funding-rounds-and-the-funding-life-cycle/

Tadayoshi Kohno, Andre Broido, K. C. Claffy. (June 20, 2005). Remote physical device fingerprinting. *IEEE Transactions on Dependable and Secure Computing,* vol. 2, pp. 93 - 108. http://ieeexplore.ieee.org/abstract/document/1453529/?reload=true

Trello. (n.d.). Retrieved January 17, 2018, from https://trello.com/about

Ting-Fang Yen, Yinglian Xie, Fang Yu, Roger Peng Yu, Martin Abadi. “Host Fingerprinting and Tracking on the Web: Privacy and Security Implications.” Microsoft, February 2012. https://www.microsoft.com/en-us/research/wp-content/uploads/2012/02/ndss2012.pdf.

User Agent - definition from What-is-my-browser (n.d.). Retrieved January 17, 2018, from https://www.whatismybrowser.com/detect/what-is-my-user-agent

Veghte, B. (2017, March 6). 2017 NVCA Yearbook Highlights Busy Year for Venture Industry and NVCA. Retrieved January 17, 2018, from https://nvca.org/pressreleases/2017-nvca-yearbook-highlights-busy-year-venture-industry-nvca/

Venture Capital. (n.d.). Retrieved January 17, 2018, from https://www.merriam-webster.com/dictionary/venture%20capital

Vestigo Ventures. (n.d.). Retrieved January 17, 2018, from https://www.vestigoventures.com/

Vestigo Ventures Portfolio Company, Vestmark, Named to Boston Business Journal's Best Places to Work 2017. (2017, May 4). Retrieved January 17, 2018, from http://www.vestigoventures.com/2017/05/04/vestigo-ventures-portfolio-company-vestmark-named-to-boston-business-journals-best-places-to-work-2017/

Zhang, L., & Zhang, C. (2013). Engineering Education and Management Vol 2, Results of the 2011 International Conference on Engineering Education and Management (ICEEM2011). Berlin: Springer Berlin.

Appendix

Appendix - Algorithm for Vestigo Projects

Following are the key steps which are decided by grad team and are assigned in Git for tracking.

A. Cataloguing:

1. Create a Spreadsheet
   1.1 Create a spreadsheet with following columns:

   Company Name: Name of the start-up company requiring fund
   Website: Company website
   Description of Company: Brief description of start-up
   Segment: One of the four Market segments (Vestigo proprietary segment)
   Contact Name: Contact person from the company who contacted Vestigo
   Contact Title: Designation of the contact person
   Contact Email: Email-id of the contact person
   Contact Phone #: Phone number of the contact person
   Date of first contact: First contact date
   Round (Stage of financing): Current stage/round of financing of start-up
   Raising: Targeted fundraising amount
   Amount Raised: Amount raised by the company
   Pre-$ Valuation: Pre money valuation of the company
   VV Sponsor: Contact person from the Vestigo handling the deal
   Status: Status of the deal
   Apollo/Minerva: Indicates whether the company domain is present in Apollo/Minerva
   Initial Confidence Level: Initial confidence level detected by the Vestigo proprietary tool
   Initial Confidence Level Date: Company domain's first detection date in a tool
   Latest Confidence Level (on 09.30.2017): Company domain's latest detection date in a tool
   Note: Any special note that a Vestigo's sponsors want to make

   1.2 Read the pitch deck from Dealflow company's folders emails and fill the spreadsheet for all the companies.

2. Load Spreadsheet to Database

The next step in the cataloguing is to load the spreadsheet data into Cogo's proprietary database. This proprietary database needs a tab separated csv(tsv) file as an input for loading.

2.1 Clean the spreadsheet

Clean the spreadsheet to ensure there are no tabs inside columns. Write a python script for automation of the cleanup activity.

2.2 Write a single query for fetching values for Apollo-Apollo(Y/N),Initial date,Initial confidence,Latest date,Latest Confidence,Percentage Change in Confidence

2.3 Query for loading values of point C in Quake

2.4 Write an automation script for performing cleaning, fetching and loading at one go.

B. Analysis Based On Confidence Level & Raised Amount Fields

1. Find out top 20 Companies with highest increase in the confidence level

2. Check whether Vestigo still has the scope for investment in those companies.

2.1 Check whether the amount required field for those companies is greater than the amount raised till now.

2.2 Find out next best companies in a row if the amount criteria of few companies is already fulfilled.

2.3 Write down a white paper summary based on the outcome of 2.1 and 2.2

2.3.1 If the company is satisfying the criteria of 2.1, give a recommendation to Vestigo saying these are the best passed deals in which Vestigo still has a scope to invest.

2.3.2 If the company is not satisfying 2.1 ie the company has raised all the required amount, give a recommendation to Vestigo saying these were some of the good deals in which Vestigo should had invested.

C. Market segmentation and landscape Analysis

1. Decide appropriate segment for each company based on the propriety segmentation standards given by Vestigo.

2. Conduct discussion meetings with Vestigo Sponsors for clarification/confirmation on market segmentation.

3. Draw a landscape analysis graph- Count & Percentage for each market segment

D. Rate the deals

1. Follow the proprietary rating standards provided by Vestigo to rate the deals.

2. Conduct discussion meetings with Vestigo Sponsors for clarification/confirmation on ratings.

E. Create a Survey/form for collecting and loading the deal information

1. Create a template for survey/form with all the desired fields.

2. Enhance the GUI, add the security features and finalize the form with Vestigo sponsors.

3. Create an automation script to extract the data from survey and load it to wpi_deal_flow table.