****

# 4 Summary of R&D Projects

The company has developed a unique solution for both customers and brands by providing them with unadulterated information regarding consumer needs and patterns. This highly unique platform enables brands to use consumer authenticated and provided, browser, mobile, and productivity data alongside large social media feeds, to analyse and identify trends in customer needs and desires and incorporate those into their products and services.

Significant research and development activities have been, and are continuing to be, undertaken. These activities continue to show solid research based results, which is evident in the continued development of the products.

For users, they have built a user dashboard and insights panel to help consumers understand patterns in their online footprint through metrics and graphs.

For brands they are in the process of developing products that can perform thorough cross-examination of consumer behaviours and provide important analytics related to their trends from data collected through various social media platforms, devices and technologies.

The company are also using various Application-Programming-Interfaces (APIs), building browser extensions and mobile apps with proprietary algorithm in order to further expand their data collection sources. Their core analytics engine can make any user data anonymous while extracting insights from it that are utilised by brands for creating desired products and developing effective marketing strategies.

# 5 Baseline Technology

The company conducted a thorough analysis of the competitive marketplace and studied many different offerings that were available for both the customers and the brands.

Existing competitors relied primarily on indirect methods for collection of large amounts of consumer data through focus groups, survey, website cookies and web scraping. This comprise an inefficient method for analysing consumer behaviour since only a small number of consumers can be contacted / interviewed in order to gain insights for developing products and strategies for larger audiences. Their target market required large volumes of online consumer data through social media that is live, natural and unaltered; however, the existing baseline technologies were unable to meet such demands.

Competitors are also looking at intrusive technology to collect consumer data with an “all data” or “no data” approach, where the user has to either provide permission to everything on their devices which makes their solution only viable for a small portion of consumer.

# 6 Scientific / Technological Uncertainties

There are a number of scientific/technological uncertainties within these projects where knowledge of whether or not the project/activity was technologically feasible, or how to achieve it in practice, was not readily known by a relevant competent professional.

The main uncertainties are detailed below:

1. If the analytics system and algorithms can be designed to handle the large volume of data - both scale and variety of data sources to capture more facets of people’s digital lives.
2. Whether the system will be able to handle mass volume queries for different types of data. Will the system will be able to handle increasingly complex and nested, unstructured documents as there’s more data logged for users every year?
3. Will it be possible to process large volumes of data in a timely manner in a simple interface for the company’s services to be made available for small and large businesses?
4. Also, can server power be used in a more efficient way in order to only use what is necessary thereby reducing costs incurred on third party cloud computing servers?
5. Would we will be able to analyse browsing behaviour, mobile app usage, activity on sites to tag consumer interests, shopping habits, search topics with context?
6. Whether it is feasible to capture details of activity on specific sites by user?
7. Whether it would be possible to identify duplicates and detect fraud documents in the database.
8. Whether the platform can be scaled by global partnerships with existing players in a seamless way.
9. Whether the system would be compliant with EU data regulations?
10. Will it be possible to create a personal identity for the digital world which will enable users to port their data as an asset class?
11. Would it be possible to enhance user behaviour data on iOS and Android apps consistently (apps used, connected time, browsing habits), while minimizing impact to battery life?

# 7 How the Uncertainties Were Overcome

The scientific/technological uncertainties outlined above required substantial research & development work and were addressed as follows:

1. The company is continually developing and adapting their data structure with a noSQL database (MongoDB). They’re working to scale analytics with Python and considering new platforms for analytics GUI. They have done research and development for their SaaS analytics offering with a beta project now in progress.
2. Significant and ongoing research & development has been undertaken with regards to establishing and maintaining reliable communication between the various APIs that they are integrating with as they are changed by the owners, especially Facebook as they have changed the way they provide data with a long term data deprecation and API changes strategy they’ve rolled out in 2016. They’re adapting the data schema to make it more efficient for queries along with efforts to optimize data processing by testing out various indexes on each data collection within MongoDB to speed up processing.
3. The company is still researching how best to make the data and analytics available real-time for both users and clients. Amongst their advancements over the past year, they have moved to a queue based processing system to streamline the processing steps for data extraction and insights processing. This also added transparency to the process and has helped reduce issues in the insights generation process, however adjustments are being tested to find the best order of data extraction to speed up insights generation for users. They are conducting research & development work to generate and enhance complex analytics using a graphical drag & drop interface for both user types, starting with the client platform.
4. The company has moved to a cloud server setup with Amazon Web Services and Microsoft Azure services to test the optimal balance for computing power and cost. The setup is being tested each quarter for optimal performance as well so there’s limited to no impact to data processing speeds required as they scale the user offering.
5. The company has devised the algorithm to categorize URLs, mobile apps with a 3-level taxonomy. Further developments have also been made in how they parse URLs to analyse consumer journey on specific sites, search keywords, phrases. Lastly, new algorithms have been put in place using permissions based on Accessibility services to capture in-depth behaviour on app usage for Android devices.

Further effort is being undertaken to: (1) tag interests based on content they are reading or viewing especially on key news sites and platforms such as YouTube, Netflix, etc. (2) enhancements to capturing e-commerce habits in a sustainable way as the e-commerce sites change their webpage designs (3) analysis of the consumer journey from path-to-purchase (4) capturing mobile usage and browsing behaviour for iOS devices

1. The company has enhanced their browser extensions across key browsers (Google Chrome, Mozilla Firefox, Apple Safari accounting for over 90% of internet browsing) with ways to track activity within a website. The first implementation of this process has been tested on Google Chrome. The enhancements also allow the company to capture shopping habits in detail from basket value to products purchased on e-commerce sites Further effort is being undertaken to capture e-commerce habits in a sustainable way as the e-commerce sites change their webpage designs
2. The company is continuing to build algorithms to identify duplication and detecting incorrect data across data sources and has implemented a monthly process to clean-up this data.
3. The company has tested the widget to onboard users and is in final discussions to gain agreement with pilot partners for UK user onboarding. They are testing an additional ways to onboard users at scale using (1) utm tags on desktop or laptop devices, and (2) using deep-linking via Branch for mobile devices. The goal is to run a test starting January 2017
4. Their systems are compliant with all EU regulations currently and they are monitoring the progress on EU GDPR implementation guidelines. At the moment, they are compliant with any published guidelines, including not storing data for Russian users.
5. The company has setup the data schema around each user and have a process in place for users to request their data. They are now researching implementation of an OAuth process that can allow them to provide an API for users to use their data with other services as well. The research is currently in progress.
6. The company has continued research to enhance their iOS and Android apps. Due to changes in permissions available on both platforms, the company has had to research and test out algorithms to keep up with changes in policy by the respective OS owners.

On Android they have developed their own algorithm to capture music habits and now also capture browsing behaviour with user permission.

On iOS they have enhanced their methods for location and usage tracking further to have less than 5% battery usage and processing power required.