****

# 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 personal insights and data monetization tool to help consumers understand patterns in their online footprint through metrics and graphs.

For brands we’ve built research products that can perform thorough cross-examination of consumer behaviours and provide important analytics related to their trends from data collected through the company’s proprietary technology and algorithms, various digital platforms, devices and technologies including social media, music apps, fitness tracker, productivity tools. This year they have further innovated to bring about instant data and instant analytics offering to help business decision makers keep at-pace with the continually changing digital world.

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. This type of technology is now also being barred from the approved mobile app stores.

Lastly the time to get actionable insights from these data sets make it difficult for brands to keep up with the pace of change in the digital landscape. Specifically, the creation of the instant data and instant analytics services has been a huge step forward in research technique built by the company.

# 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:

a) 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?

b) Will it be possible to extend the instant data capture to mobile devices to enable real-time data sharing with permission, specifically around app usage data as clients are looking to understand mobile use behaviour instantly?

c) How can we architect the instant data solution to not interact with user data when the panellist does not belong to our own platform, i.e. it’s the client's panellist

d) How can we bring our instant browsing data offering to browsers beyond Chrome and Firefox?

e) How do we build a new sales & marketing model, and related processes?

f) Would it be possible to create a dynamic model to categorize websites?

g) Will it be possible to create a tool that vets the audience based on their actual browsing activity rather than their declared or claimed behaviour?

h) How can we improve the iOS and Android browsing apps to allow users to browse easier, faster and safer on their devices?

# 7 How the Uncertainties Were Overcome

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

a) Continuing from the research findings of the previous year, we have now designed, defined the implementation plan and started the move over to a micro-services based infrastructure to (1) help deliver a portable 'personal identity cloud' to consumers (2) optimize to take advantage of the serverless infrastructure that is increasingly being rolled out in cloud based infrastructure. The key to delivering the 'personal identity cloud' is starting with a secure authentication method so we started the implementation with authentication protocols.

Authentication is needed when an application needs to know the identity of the current user. The most common authentication protocols are SAML2p, WS-Federation and OpenID Connect – SAML2p being the most popular and the most widely deployed. Although OpenID Connect is the newest of the three, but is considered to be the future because it has the most potential for modern applications. It was built for mobile application scenarios right from the start and is designed to be API friendly and OAuth2 is a protocol that allows applications to request access tokens from a security token service and use them to communicate with APIs. This delegation reduces complexity in both the client applications as well as the APIs since authentication and authorization can be centralized. Based on this research we have implemented our authentication processes using the OAuth2 protocol. Since we are also moving to centralized authentication, we have divided our Monolithic application to Microservice base API, which ensures high scalability and resilience i.e. failure in one service does not impact other services. Other benefits of using Microservices are it's easy to maintain, enhance and deploy.

b) We researched and launched an instant data product for the Android market. With user permission, the app is able to capture Android device information along with a list of the installed apps, and the date the user installed the app which is really valuable to understand user's interests, trends in app categories and correlate user lifestyle to their behaviour. However, after extensive research, we were not able to device a way to create a similar proposition on the iOS platform as the App Store guidelines will not allow apps to access this type of detail even if the user provides permission

c) Having successfully developed the instant data solution on our panel, we had to determine a way to make the technology plug into any user community or panel rather than just our own. In keeping with GDPR regulations, there are limitations to what a panel company can share with a technology provider such as ourselves, and vice versa. So, we need to design the solution to be GDPR compliant and keeping privacy in mind for the user experience as well. We've designed the solution to be available for open deployment with any community or user group. We've made the configuration up front simple so that everything from the client name, terms, incentives, scope of study can all be customized and the links are dynamic so that the user’s identity can be anonymous when they come over to using Snapshot. The solution is GDPR compliant.

d) The company has found great client interest in the instant data product, Snapshot, made available on Chrome and Firefox browsers. Though the two browsers have captured large share of the browser market there are many other browsers that are coming up the ranks, to name a few Safari, Edge, Opera, Brave, Vivaldi. The company needs to build Snapshot functionality for these browsers as well in order to continue leading the market in passive data tech and be able to address client needs. So, the company researched the core tech behind these browsers to see how the current could be adapted. In their research they discovered that Opera, Brave, and Vivaldi browsers use Chromium, open-source web rendering engine. This is valuable for the company as it makes adapting the current extension to the other Chromium based browsers very easy.

e) We looked to a subcontractor for expertise on developing new sales and marketing models to find new ways to communicate and sell our products. With his time for 2 months we were able to discover key findings on sales methodology that we are now employing as we work across so many different client types to how best to reach them too.

f) Analysing big data about people's browsing and mobile usage behaviour is key to piecing together the consumer journey, understanding triggers for behaviours as well as implications and most importantly being able to paint a comparable picture of people's behaviour so brands can understand key differences that set their audience apart. In the past we had devised a manual, 3-tier algorithm for categorization of sites and apps. However, over the past few years as the data availability has increased, there's increasingly too many URLs and Apps that wouldn't fit neatly into the tiered model. This led us to start researching ways to redesign the categorization. After assessing different versions of the tiered model, we actually got to the conclusion that the best way to use the categorization in the future would be to make it a flat model with juts 1 level, i.e. tags. We have designed a Machine Learning algorithm that will automatically be able to tag a site with the relevant tags and weighting on how much the tag matches the site's purpose/description. After numerous iterations, our model can now automatically categorize over 350k sites. We're improving that further by developing algorithms to capture more than just meta descriptions to be able to categorize another 250k sites that we've seen browsed to date, and will be adapting the Machine Learning algorithm next year to mobile app categorization

g) Leveraging our instant data tech, we wanted to go further in developing tools that allow consumers to be vetted based on their behaviour rather than sharing all of their data to have the vetting step happen only as a consequence. Clients using the instant data tech also expressed a keen interest in being able to improve their sample quality for research so we researched and have now devised a way for users to be screened based on their browsing behaviour or apps and device details. This method allows the client to launch a study confidently screening for search terms, sites visited, accounts held, apps they have installed for how long, device identifiers and more. We also did some additional research to rearchitect the solution so that in getting vetted no data is shared with us or the Client, in fact the product has been designed with privacy-first approach and is GDPR compliant. All the algorithms to vet the consumer are checked on the consumer's device and no data is shared in qualify the research participant

h) We have spent a considerable effort this year also to enhance the mobile browsers for users so that they can continue to benefit from the latest improvements in blocking tech, and are now also able to use their browsing data to be matched to research studies. The browser has been revamped entirely on Android to make the app lighter for the user's device as we realize the breadth of Android devices out there. We also carried out a feasibility study on porting over the functionality to do a detailed basket capture when user is shopping online via the mobile browser (similar to our desktop browser extensions). We've devised the pilot and used it for user testing with a small group and know that this is feasible. We will be incorporating this feature into the browsers next year so that users are able to further enterprise on their data footprint.