Data-Driven Decision-Making in Digital Entrepreneurship

Abeba Nigussie Turi, Xiangming Samuel Li

Abstract—Data-driven business models are more typical for established businesses than early-stage startups that strive to penetrate a market. This paper provided an extensive discussion on the principles of data analytics for early-stage digital entrepreneurial businesses. Here, we developed data-driven decision-making (DDDM) framework that applies to startups prone to multifaceted barriers in the form of poor data access, technical and financial constraints, to state some. The startup DDDM framework proposed in this paper is novel in its form encompassing startup data analytics enablers and metrics aligning with startups' business models ranging from customer-centric product development to servitization which is the future of modern digital entrepreneurship.

Keywords—Startup data analytics, data-driven decision-making, data acquisition, data generation, digital entrepreneurship.

I. INTRODUCTION

AS a powertrain of the digital economy, data have proved to be one of the competitive advantages that businesses established on such resources excel. Accordingly, data science and practices have evolved, incorporating the dynamic business needs over time. With the increasing stock of big data from the digital footprints of users, integrating and exploiting the benefits of such information in decision making is proven to facilitate decision making and unlock business potentials [4]. Moreover, open data initiatives by the data hoarding tech companies, like Microsoft, can help businesses utilize the stock of digital footprints they hoard in their data empire [10]. With the fast pace of digitization, reinforced with COVID-19 and multiple other factors, by 2024, global data created, captured, copied, and consumed are predicted to reach 149 zettabytes from as low as 12.5 zettabytes a decade ago in 2014 [11]. In a world where such data can be used at a near-zero marginal cost, entrepreneurial DDDM strategies and robust analytics are vital. There is evidence that using data in decision-making has a significant cost reduction effect, specifically in relation to operational efficiency [1].

Startups face multilayer problems in today's dynamic business environment. Accordingly, informed decisions of this form in an entrepreneurial course of actions rather than mere reliance on intuition can improve business performance. However, in most cases, startups fail to incorporate DDDM in their core business lines. Some of the factors identified to limit startups from using data analytics include a mere focus on product development, team expertise in data analytics, amount of data at stock, a mere focus on core product delivery, attention to market penetration, and seeking customer validation in the short-run [3].

In this paper, by focusing on digital entrepreneurial practices, we will flesh out strategies to monetize the primary or secondary chain of data that businesses access in a way that boosts their potential growth opportunities. By identifying the value of data and examining the potential to exploit such values through smart business models, the paper presents a framework to unlock such potential in one of the digital economy's underutilized resources. Besides, by taking use cases and highlighting the respective on-demand business intelligence tools, the paper will flesh out the standard practices and strategies to accelerate growth through DDDM. Further, it will provide an extensive discussion on the potentials which open data holds for startup businesses and strategies to utilize such an open resource to boost business growth at a scale. Digital blueprints, data valuation, smart business models, on-demand business intelligence tools (services and software for processing digital footprints), DDDM, and data-centric value creation are the main topics covered in this paper.

The rest of this paper is organized as follows. Section II presents related works and highlights the research gap we intend to fill in this work; Section III briefly presents some data analytics tools, Section IV presents the practice of data-powered decision making in the startups' context, Section V presents data source, challenges, and opportunities in relation to startup data analytics, Section VI highlights the core data analytics application areas for startups; and, Section VII presents data analytics strategy and a conceptual framework to DDDM. Lastly, Section VIII concludes our key findings with direction for future works.

II. RELATED WORKS

Several studies have been conducted to unlock entrepreneurial business potential and identify key challenges such businesses observe [3], [5]. With this is the potential of digital footprints generating big data for predictive and visual analytics [6], [13]. Yet, there is evidence that early-stage startups face the challenges of skills, capital, market uncertainty, technological uncertainty, time management, and privacy issues to generate values that contribute to poor startup success rates [3], [5], [12]. Hartmann et al. [6] provided an extensive discussion on the taxonomy of data-driven business models used by startups, encompassing seven key business activities: free data collection and aggregation, analytics-as-a-

Abeba Nigussie Turi and Xiangming Samuel Li are with University Canada West, Canada (e-mail: abebanigussie@ucanwest.ca, samuel.li@ucanwest.ca).

World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:16, No:4, 2022

service, data generation and analysis, free data knowledge discovery, data-aggregation as-a-service, and multi-source data mash-up and analysis (see also [7]). Kandel et al. [7] identified an enterprise's organizational feature as one of the key factors defining enterprise data analytics' efficiency.

The literature in the field is limited in providing theories and practices for early-stage startup data analytics. This paper will flesh out the principles of DDDM in the startup context. Besides, we extend the work on startup data analytics by providing a general conceptual framework of DDDM that applies to digital entrepreneurs.

III. THE PRACTICE OF ENTREPRENEURIAL DATA-POWERED DECISION-MAKING

Data analytics can play a crucial role at different stages and scales of business operation, starting from the inception stage of a product life cycle. Before launching a new business or introducing a new product, big data can be used to make informed decisions. This will allow product testing with an accurate picture before getting into the actual market. Relying on the core bubbles or a small circle of people to test a product might create some bias when bringing the product to general users. The digital economy has brought the world together into a virtual planet and this has made it easier for startups to have insights and tests to their new products from the crowd. For example, location analytics can help identify an appropriate business location that fits your business lines. Another example is property valuations of the real estate industry. A stock of data enables the analysis of markets and identifies non-economic and economic signals and trends of the real estate industry.

To help us look into the practice of data-powered businesses by startups, we picked the clients from the data analytics company, Mixpanel. The company tracks web and mobile usage that allows targeted product development. Here, we will pick and look at two sample companies that rely on its tool to make informed decisions using data.

One of companies is the transaction management and electronic Signature tech company, Docusign, which uses data-analytic tools from Mixpanel to boost its global reach. The system allows Docusign to enhance its basic and custom product metrics like customer retention, the number of documents customers successfully complete, and upgrades. For example, Mixpanel behavioral analytics enhanced Docusign's paid upgrades by exposing some premium features that incentivize users for the conversion. The analytical tool increased upgrade conversions by 5%, bringing about 130,000 new users per day [14].

Kast is another company using Mixpanel's data analytics tool in its service provision of a "virtual living room". The company uses content sharing technology to allow the traditional physical human interaction experience (e.g., watching movies or playing games in a shared space). The startup uses Mixpanel for its DDDM to understand its users. For example, the data analytics took user feedback to identify "Power Users" (active members in the last 15 days and a minimum of 30 time-frequency of partying through the app). This feedback is used for product and feature developments. The analytics allowed the startup to trace users' behavioral patterns of the app usage by tracking user data. Such an informed product decision allowed the company to increase its customer retention by about 50%. See also [12] for the feedback loops in the data-driven economy applied to the prosumers' blueprint towards data monetization in the supply chain.

IV. DATA SOURCE FOR STARTUPS: CHALLENGES AND OPPORTUNITIES

Data analytics applies to the notion of enabling technologies that capture relevant information for economic and non-economic value generation purposes. Such information can be revealed from massive volumes of data, big data, coming from the information society's digital blueprints. The massive blueprint of human-machine interaction comes from transactional mobile and internet records, user-generated content across social media, blogs, marketplaces, and various digital platforms, ad-hoc content collected through sensor networks, IoT, and other channels.

Many digital footprints left by users lay the basis for structured and unstructured data, which businesses can harness in their best interest. In most cases, specific data that apply to a given business come from the data businesses capture from their online store visitors across different sessions, platforms, and domains. This treasure contains potential customers of the product, which businesses pull for their marketing strategies. However, several barriers limit entrepreneurs from using these potentials [3].

One of the main challenges early-stage digital entrepreneurs face is the limit in the amount of data they have in stock. With this limit in place, it will be difficult for startups to rely on internal data to make business decisions. In this regard, it is ideal for such startups to use relevant external data and tools that apply to their business needs. To mention one, leveraging social data mining of third parties can help in targeted ads. For example, data exchange platforms, vendors, and tools like Ocean Protocol, Gnip, Immersive Labs, Data Sift, BDEX, Quantcast, and iGrant.io facilitate and enable such real-time data for customer-centric businesses. Similar to these are the open-source big data analysis platforms and tools like Hadoop, MapReduce, GridGain, HPCC Systems, and Storm that have a significant innovation diffusion. Another way to excel on this is to make proper use of open-source data available on the web. An equally important aspect to discuss here is data optimization practice. Optimization of the data businesses capture from online store visitors across different sessions, platforms and domains can allow optimal business performance through improved business operations and accurate decision making.

V. DATA ANALYTICS APPLICATION AREAS FOR DIGITAL ENTREPRENEURSHIP

Data analytics, including data visualization and machine learning, can empower digital entrepreneurs to acquire new customers, create new value chains, and improve operational efficiency. Operational efficiency is one of the forefront value-add coming to businesses with data analytics. The following

World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:16, No:4, 2022

attempts illustrate some of the application areas of data analytics for startups.

A. New Product Design

Creating personalized new products for each customer can be challenging. However, the digital footprint of users has enabled this through aggregated data analytics. For example, Hellofresh (a meal-kit company) consumer data streams are massive, aggregated from social media, emails, and user impressions on their websites, allowing customized fresh ingredients and recipes for individual customers. Here, a special case of digital entrepreneurship takes the form of prosumers in which consumers of digitally enabled products and services are the producers themselves, thus adding to new product development and designs.

B. Online Platform Building

Data analytics tools can help digital entrepreneurs develop a professional website without HTML, Web 3.0, CSS, and JavaScript. For example, AI-powered website builders like Wit.ai, Wix, and Dialogflow. Content generation tools such as Articoolo, Wordsmith, and Quil can analyze and learn from other web data and generate human-like content to help startups engage online customers more. Besides, the Google Analytics tool allows to monitor, analyze and improve startup web business performances.

C. Reputation Management

Trust, as repeatedly coined, is the currency of the digital economic system. There are diverse rating and reputation schemes in the P2P business models. For example, Li [9] applied text analytics and machine learning approaches, including Naive Bayes, logistic regression, support vector machine (SVM), and long short-term memory (LSTM-AI), to more than three million online text reviews collected from the Airbnb platform. A large knowledge-based label (CORPUS) is built, and an innovative text review score (TRS) is constructed as a new online trust measure. Digital entrepreneurs can adopt a similar approach or take advantage of this sizeable online review CORPUS to rapidly build up new online reputation systems.

D. New Product Market Launch

Digital startups can use data analytics to find an optimal marketing strategy to launch their new products successfully. For example, various relevant data points from previous market launches can be collected from open data sources for similar products in a specific market; and data visualization and analytics tools can be employed to find the effectiveness of each marketing promotion and product pricing.

E. Digital Marketing and Customer Profiling

Data analytics can empower digital startups to know more about their customers regarding product features and pricing. Personalization is vital to delivering a satisfactory customer experience. Data analytics enables an efficient, personalized user experience enhancing customer conversion and retention. For example, Netflix traces users' movie or TV show preference or Amazon recommends related products based on transaction records and other users' digital footprints.

F. Customer Relationship Management through Chatbots

One of the powerful data analytics applications is online chatbots powered by artificial neural networks (ANN) as a deep learning technique. The chatbot is a conventional software agent to communicate with human users through a natural language based on ANN. Chatbots have the advantage of being self-service, including 24x7 full online care, global languages, cost-saving for call centers, as well as interactive and intelligent user experiences. Chatbots are expected to help enterprises to save about $8 billion annually by 2022 in customer-support costs, compared to only the $20 million estimated saving in 2017 [13]. One of the common chatbots is Amelia by IPSoft, as the market-leading digital employee and conversational AI application. Chatbots can offer high-volume and high-impact customer services such as Frequently Asked Questions (FAQs), service ticket management, account management, and customer care.

G. Fraud Prevention

Data analytics has been extensively applied for business fraud detection and prevention. Many fraud detections use UPL classification models, such as decision tree, logistics regression, SVM, and even ANN deep learning. In business practice, various unsupervised learning methods like profiling, clustering, anomaly detection, and co-occurrence association are also used for different fraud detection and prevention.

H. Data Analytics for IoT Machine-to-Machine (M2M) Applications

M2M-based IoT data differ from traditional big data in terms of the growing "ocean" volume, heterogeneous formats, embedded noises, and real-time attention needed like medical care and animal surveillance. The wide-spreading 5G networks and IoTs applications will trigger a new round of Industry 4.0 innovations, including smart homes, smart cities, smart agriculture, and smart globe, or called Smart Xs (Smart Everything). Meanwhile, this brings a huge potential for new entrepreneurs to create various M2M-based IoTs applications using advanced data analytics tools.

VI. DATA ANALYTICS STRATEGY AND CONCEPTUAL FRAMEWORK FOR DIGITAL ENTREPRENEURSHIP

Based on the startup metrics of Kemell et al. [8] and the Pirate Metrics, or "AARRR" of McClure [10] and Hartmann et al. [6] taxonomy of data-driven business models for startups, we developed a conceptual framework for the entrepreneurial DDDM process.

Here, we aggregate the foundation for the entrepreneurial DDDM under two main pillars: (1) The enablers which allow the digital entrepreneur to make data-driven decisions, and (2) the business metrics which need to be identified to fit the decision-making process with the core business needs and goals. Kemell et al. [8] provided key startup metrics for tech entrepreneurs, including engagement metrics that measure how often people engage with a product. For example, for an

World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:16, No:4, 2022

uberfication business model, Daily Active Users (DAU) and Monthly Active Users (MAU) metrics measure customer engagement with the app. Another metric is the churn rate which measures customer stickiness to a product, like the rate of cancelation or non-renewal in a subscription-based business model. As is typical for startups to get customer endorsement and validation, customer engagement, product reviews, and customer feedback are essential metrics startups focus on for product development and enhanced services. As startups consider scaling their business, equally important are the growth metrics.

Under the first pillar, by extending the key big data analytics enablers indicated by Behl et al. [2] and Kandel et al. [5], we identified four main enablers which we identified them as people (encompassing the technical skills and entrepreneurial leadership a startup owns), startup's organizational features and process (referring to the startup's internal business process, relationships and diversity of data sources), infrastructure and platform (including analytical tools, cloud storage technology, security and privacy-preserving solution) and data (acquisition-external data-capture from data exchange platforms, data analysis platforms, open-source like social data for targeted ads and real-time data for customer-centric business, keyword performance and ranking, search behavior, page visits, tweets, likes, share comments, or generation- internal data capture from tracking (potential) customer or user data, transaction records, automated user tracking, customer feedback, customer engagement, business operations).

The second pillar of business metrics is the performance metrics under (i) Business and financial metrics, (ii) User/customer metrics that capture behavioral patterns of (potential) customers allowing for customer-centric business, (iii) Service metrics that capture the process and product metrics, and (iv) Social media metrics as a measure of business visibility for demand-pull.
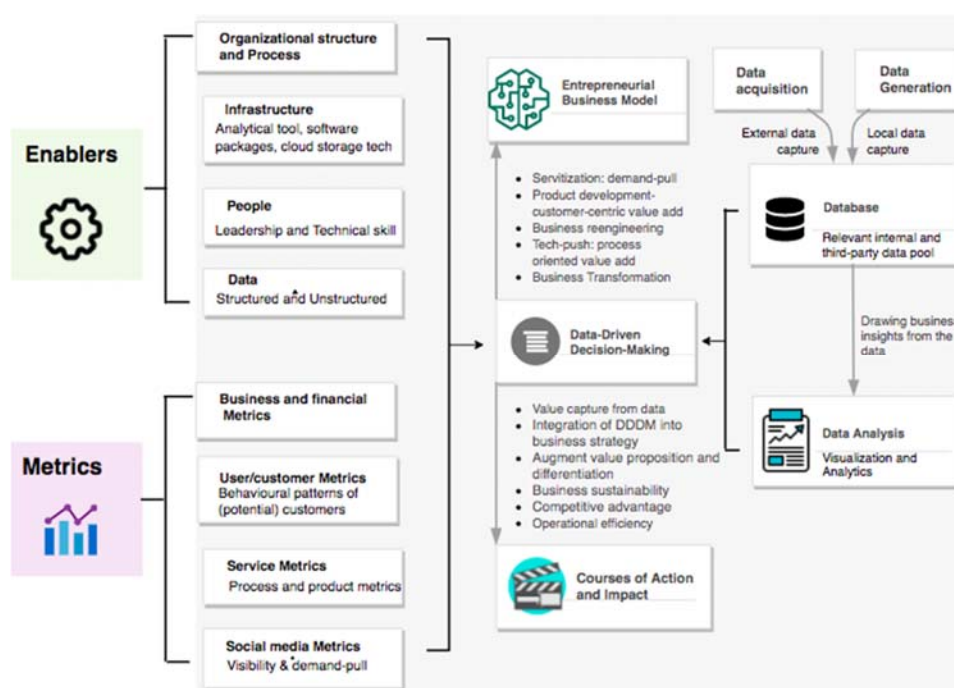

Fig. 1 Entrepreneurial DDDM Process: A Conceptual Framework

A data pool from diverse sources allows the startup to draw business insights through analytics and visualization using analytical tools in the decision process. This will feed into the entrepreneurial business model based on the revenue models defining the startups' potential and actual business lines. Here, the analytics can support customer-centric value-add; servitization to build new revenue streams through service business lines (demand-pull), business reengineering; product development; process-oriented value-add (tech-push and automation); or an overall business transformation and reengineering. Business analytics and visualization capture value from data and integrate data-powered decisions into the business strategy. Hence, we identified the major courses of action in the DDDM process.

Such courses of action include operational efficiency, augmentation of value proposition and differentiation for competitive advantages, and ultimately attaining business sustainability through robust data-supported measures prescriptive analytics. Fig. 1 depicts a conceptual framework for entrepreneurial DDDM. Summing up, based on the literature and best practices in data analytics, the main steps for entrepreneurial data analytics are discussed as follows:

A. Business Understanding

Digital entrepreneurs should start with compelling business problems in mind and look for robust data analytics tools to solve.

World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:16, No:4, 2022

B. Data Collection

At this stage, we identify internal or external data sources that will give an insight into the business and fuel growth for startup. In line with the business rules, another important thing to consider is to plan what you intend to acquire from the data. It should be noted that new goals can arise as one observes the trends and patterns in the data.

C. Data Understanding

It is crucial to understand the strengths and weaknesses of the data; since there is no exact match between the data and targeted problems to be solved. Because a historical dataset is often composed of unnecessary features which do not serve the purpose, it is vital that the analyst goes through the data and filter the required information in order to draw insight out of it. For example, a client database, product marketing response data, and business transactional records cover different intersectional populations with varying degrees of reliability.

D. Data Organization

This phase often is carried out with data understanding, where the data are manipulated and processed into proper forms for data visualization and data analytics tools to yield valuable results. At this stage, the data are organized to form a description and/or identify themes and trends in user behavior. Examples include converting raw data to a tabular format with rows and columns, dealing with missing data, changing data types, removing excess data, data value normalization, etc. Besides, data organization includes cleaning the data and getting an insight with further considerations for data compatibility issues in the data organization process. One of the principal reasons for this is that as the startup grows, the data analytics will be replicated at a scale as the business and data analytics systems evolve.

E. Modeling

It is the primary phase where data analytics techniques are applied to the data, including visualization, clustering, classification, and deep learning (see Section IV). The model has to specify key instruments and metrics on which the data analytics relies in order to address the core business problems, see Pirate Metrics, or "AARRR" [10].

F. Evaluation

Digital entrepreneurs can employ different evaluation metrics such as the confusion matrix and ROC/AUC curves to evaluate the model performance in terms of accuracy, precision, recall, cross-validation, and overfitting capacity. Model evaluation is a science rather than an "art". It includes a rigorous assessment of various models to select a robust model for a successful digital business.

G. Deployment

Deployment is the final phase of putting data analytics to real use to achieve business results. It involves implementing the best-evaluated model from the previous step, such as predicting the likelihood of customer churn and sending a special offer to highly possible "churn" customers to retain them in business to maximize the return on per customer. A new fraud detection model can be trained and integrated into a customer relationship management (CRM) system to monitor accounts and identify potential transactional fraud to minimize business loss.

VII. CONCLUSION

Digital entrepreneurs often fail to incorporate DDDM in their businesses during their startup incubation, contributing to a very high startup business failure rate. This work fleshed out the core state-of-the-art data analytics tools and application areas for early-stage startups, including new product design, online platform building, digital marketing, customer care, operational efficiency, and IoT-based machine-to-machine automation. Further, we offer practical strategies for entrepreneurial data-powered decision-making, such as data acquisition in the face of limited internal data generation. The conceptual framework for the entrepreneurial DDDM process developed in this work relies on two key pillars of enablers that allow the digital entrepreneur to make data-driven decisions and the entrepreneurial business metrics that fit the decision-making process with core business needs and goals. In this regard, future work direction will be to look into the dynamics of data analytics techniques and startup business performance under the agile digital entrepreneurship environment. Another avenue of research related to this is the empirical evidence and quantitative analysis of data-powered decision-making in the startup business context, including the sectoral analysis.

REFERENCES

[1] Bean R. (2017) How Companies Say They're Using Big Data in Harvard Business Review- Analytics And Data Science. Retrieved February 21, 2022, from https://hbr.org/2017/04/how-companies-say-theyre-using-big-data

[2] Behl, A., Dutta, P., Lessmann, S., Dwivedi, Y. K., & Kar, S. (2019). A conceptual framework for the adoption of big data analytics by e-commerce startups: a case-based approach. Information Systems and e-Business Management, 17(2), 285-318.

[3] Berg, V., Birkeland, J., Pappas, I. O., & Jaccheri, L. (2018). The Role of Data Analytics in Startup Companies: Exploring Challenges and Barriers. In Conference on e-Business, e-Services and e-Society (pp. 205-216). Springer, Cham.

[4] Caulkins, J. P., Bao, Y., Davenport, S., Fahli, I., Guo, Y., Kinnard, K., ... & Kilmer, B. (2018). Big data on a big new market: Insights from Washington State's legal cannabis market. International Journal of Drug Policy, 57, 86-94.

[5] Giardino, C., Bajwa, S. S., Wang, X., & Abrahamsson, P. (2015, May). Key challenges in early-stage software startups. In International Conference on Agile Software Development (pp. 52-63). Springer, Cham.

[6] Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data–a taxonomy of data-driven business models used by startup firms. International Journal of Operations & Production Management.

[7] Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2012). Enterprise data analysis and visualization: An interview study. IEEE Transactions on Visualization and Computer Graphics, 18(12), 2917-2926.

[8] Kemell, K. K., Wang, X., Nguyen-Duc, A., Grendus, J., Tuunanen, T., & Abrahamsson, P. (2020). Startup Metrics That Tech Entrepreneurs Need to Know. In Fundamentals of Software Startups (pp. 111-127). Springer, Cham.

World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:16, No:4, 2022

[9] Li SX (2021), Short or Long Review? Text Analytics and Machine Learning Approaches to Online Reputation, International Journal of Business and Management Research (IJBMR), 9 (1), pp 28-40, e-ISSN: 2347-4696.

[10] McClure, D. (2007). Startup metrics for pirates: AARRR. Startup Metrics for Product Marketing & Product Management.

[11] NCES. (n.d.). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025(in zettabytes). In Statista - Telecommunications. Retrieved February 21, 2022, from https://www.statista.com/statistics/871513/worldwide-data-created/

[12] Oliva, F. L., & Kotabe, M. (2019). Barriers, practices, methods and knowledge management tools in startups. Journal of knowledge management.

[13] Turi, A. N. Technologies for Modern Digital Entrepreneurship Understanding Emerging Tech at the Cutting-Edge of the Web 3.0 Economy.

[14] Mixpanel- service performance and efficiency available at, https://mixpanel.com/customers/docusign-accelerates-growth/ Retrieved Apri1 2, 2022