Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth

Laurette Dubé, Pan Du, Cameron McRae, Neha Sharma, Srinivasan Jayaraman, and Jian-Yun Nie

> *"Collectively, we have only begun to scratch the surface of what is the biggest potential market opportunity in the history of commerce. Those in the private sector who commit their companies to more inclusive capitalism have the opportunity to prosper and share their prosperity with those who are less fortunate."*
>
> C. K. Prahalad & Stuart L. Hart (2002)
> In "The Fortune at the Bottom of the Pyramid"

Inclusive innovation has not yet reached societal scale due to a well-entrenched divide between wealth creation and social equity. Taking food as the initial test bed, we have proposed the convergent innovation model to address such challenges still facing 21st century society by bridging sectors and disciplines around an integrated goal on both sides of the social-economic divide for innovations that target wealth creation with an upfront consideration of its externalities. The convergent innovation model is empowered by two key enablers that integrate an advanced digital infrastructure with leading scientific knowledge on the drivers of human behaviour in varying contexts. This article discusses the structure, methods, and development of an artificial intelligence platform to support convergent innovation. Insights are gathered on consumer sentiment and behavioural drivers through the analysis of user-generated content on social media platforms. Empirical results show that user discussions related to marketing, consequences, and occasions are positive. Further regression modelling finds that economic consequences are a strong predictor of consumer global sentiment, but are also sensitive to both the actual price and economic awareness. This finding has important implications for inclusive growth and further emphasizes the need for affordable and accessible foods, as well as for consumer education. Challenges and opportunities inspired by the research results are discussed to inform the design, marketing, and delivery of convergent innovation products and services, while also contributing to dimensions of inclusion and economic performance for equitable health and wealth.

Introduction

Inclusive innovation has been proposed as a framework to reduce inequities that have oftentimes accompanied wealth creation and modern development since the onset of the first industrial revolution (Schillo & Robinson, 2017). Inclusive economic growth has been defined as "growth that not only creates new economic opportunities, but also one that ensures equal access to the opportunities created for all segments of society, particularly for the poor" (Ali & Son, 2007). More recently, the OECD has called for inclusiveness in "economic growth that creates opportunities for all segments of the population and distributes the dividends of increased prosperity, both monetary and non-monetary terms, fairly across society" (Planes-Satorra & Paunov, 2017). Yet, leading economists are reporting unprecedented increases in inequities within and across countries

around the world (Piketty, 2014) at the price of constraining human, social, and economic progress for all population segments in future generations (Stiglitz, 2016). As a sign of modern economies needing to modify paths to economic development, in the advanced industrial economy of the United States, overall life expectancy at birth has actually decreased for two years in a row and more so for poor and other disadvantaged population segments (Kochanek et al., 2016). Wealth is rising unequally and it is increasingly concentrated in fewer hands, with the benefits of innovation also shared unequally.

Considering the role that innovation has played in economic growth since the onset of the first industrial revolution (Beinhocker, 2007; Drayton & Budinich, 2010; Dubé et al., 2014), inclusive innovation holds significant promise for addressing social and economic inequities. Inclusive innovation projects typically aim to improve the welfare of lower-income and marginalized groups by enabling their full participation in the production and consumption of social and commercial goods, services, or programs (Chataway et al., 2014; Pansera & Martinez, 2017). Social entrepreneurs are bringing sophisticated technical solutions, business acumen, and increasing investment to address inequities in both industrialized and developing economies (Martin & Osberg, 2007, 2015). Commercial firms operating on different scales now place innovative supports for the most vulnerable and disadvantaged groups living in the communities where they operate as part of their corporate responsibility strategies (Campbell, 2007).

Pioneered by Prahalad and Hart (2002), bottom-of-the-pyramid and other forms of frugal and lower-cost commercial innovation have penetrated resource-poor markets in emerging economies and value-conscious markets in both developing and industrialized countries. As governments and civil society groups struggle for greater impact and longer-term viability from social supports that still often assume never-ending access to governmental or philanthropic funds, innovations targeted to bottom-of-the-pyramid markets have a high potential for economic growth in emerging economies for domestic and multinational businesses while also addressing the needs of disadvantaged populations (Prahalad & Hammond, 2002; Prahalad & Hart, 2002). However, such innovations occupy a limited share of national and global wealth-creation systems in both developing and industrialized countries, and the significance of their social and economic impact for individuals, organizations, and society remains limited (Dubé et al., 2012).

Constraints that still prevent the above instances of inclusive innovation from reaching societal scale are tied to a structural divide between pathways of poverty alleviation and those of wealth creation that have emerged from the linear and siloed features of Western-centric development since the first industrial revolution (Gillespie et al., 2013; Moodie et al., 2013). This divide is between the private sector – which typically focuses on technological innovation and economic growth that carefully caters to targeted customer needs – and the government and civil society sectors – which typically use a "one size fits all" approach to ensure acceptable conditions for health, education, and other social goods for all. This divide creates a disconnect between, on the one hand, the still predominantly rights-based human development approach deployed by governments and civil society to support the poor through social welfare and community mobilization to reach subsistence (Devaux et al., 2009), and, on the other hand, a precisely targeted economic focus driving wealth-creation activities in value chains and markets as industrialized urban societies develop (Reardon et al., 2012). Furthermore, two major negative externalities of existing partners of economic growth – namely healthcare and environmental costs – are now threatening the financial viabilities of governments in industrialized and developing countries alike. It is clear that investments and policies in current models of inclusive innovation will not suffice unless such "externalities" become mainstream in all industrial innovation.

With innovation accounting for 50% to 80% of all social and economic progress tied to modern development (Croitoru, 2012), it is only through *innovating the way we innovate* that we can go beyond what has been possible so far in simultaneously advancing paths of wealth creation and poverty alleviation in a way that fosters lasting human and environmental health. As information is key to transformation in complex dynamic systems (Hammond & Dubé, 2012), the most recent developments in artificial intelligence, big data, and digital technologies can serve as key catalysts for creating an adequate and lasting supply and demand for such innovation across the socio-economic spectrum and around the world. Such a catalyst requires a transformation of both our methods of innovation as well as the current practices of a broad spectrum of stakeholders, including consumers. In fact, individuals themselves often feel divided and conflicted between their expectations, intentions, and actual behaviours as consumers resulting in increased consumption for immediate gratification or to support long-term health goals and social or environmental causes. With modern society experi-

encing the fourth industrial revolution – where information and digital technologies are in the process of replacing fuel and other physical resources as drivers of both social and economic development – convergent innovation has been proposed as a next-generation approach to both inclusive and mainstream innovation that will bridge the social-commercial disconnect in both consumers' minds and innovation systems to build supply and demand for societal-scale solutions (Dubé et al., 2012; Dubé et al., 2014).

In this article, we offer a brief review of the convergent innovation approach that has taken food as an initial test bed, and we report on the early stages of a research program designed with the objectives to: i) develop the structure and methods for an artificial intelligence digital platform to support convergent innovation; and ii) generate consumer insights on sometimes conflicting demand drivers for convergent innovation, with a focus on user-generated content through social media. We review management research on user-generated content to inform our social media analysis for convergent innovation in food, then we report on the structure and methods used for the artificial intelligence platform. Next, we report the results from a first empirical analysis of user-generated content. We conclude by discussing the challenges and possibilities presented by the research results.

Convergent Innovation

Convergent innovation, which has been in development for more than a decade taking the food domain as a test bed (Figure 1), is an intersectoral translational framework that aims to *innovate the way we innovate* to address some of the most complex challenges and possibilities facing 21st Century society.

The convergent innovation framework combines technical, social, and institutional innovation and bridges science, policy, and action through a unique blend of digital technologies and social capital with human creativity and agency. The aim is to invent a 21st century intersectorality to improve lives, promote equity and health, and accelerate environmental sustainability – at the same time and through the same pathways where wealth is being created individually and collectively. The multifaceted intersectorality underlying convergent innovation combines: i) natural, life, social, and engineering sciences; ii) economic systems and the larger natural and social systems within which they reside; iii) public, private, and civil society actors in each of these systems; and iv) the various scales at which mechanisms, actors, and institutions are operating.

The operational deployment of convergent innovation at scale may be made possible at this point in time by

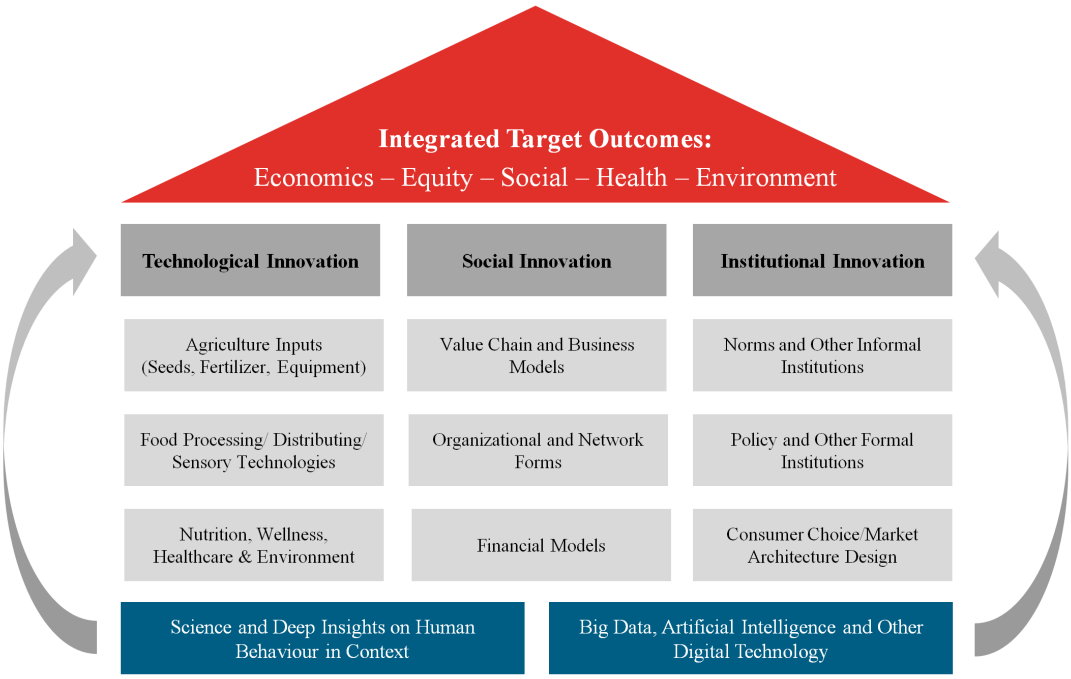

Figure 1. Convergent innovation: Behavioural change and ecosystem transformation solutions (Adapted from Dubé et al., 2014)

two key enablers. First, is the unique digital infrastructure that defines the 4th industrial revolution, including recent advancements in big data, artificial intelligence, and integrative analytics to map and bridge knowledge and its operational interfaces with policy and action. The second key enabler integrates cutting-edge scientific knowledge on complex drivers of human behaviour in varying contexts and their linkages to biological and social outcomes, accelerated by the conceptual and methodological development in genomics, neuroscience, and behavioural economics. The nomination of Richard Thaler, the father of behavioural economics, as the 2017 Nobel Laureate of Economics is a clear signal of the scientific significance of both rational and non-rational processes, and the importance of contexts, in our understanding of the drivers of real-world human behaviour. In terms of observed changes in the drivers of human behaviours, there are promising shifts from short-term gratifiers – such as pleasurable experiences, convenience, and status – towards longer-term normative considerations for oneself and society. However, discrepancies often remain between what one thinks, what one intends to do, and what one does (Dubé et al., 2008; Lin & Chang, 2012). This makes convergent innovation in food quite challenging. Creating a convergent innovation platform (Figure 2) requires deep insights into consumer behaviour empowered by advanced data and computer science capabilities to characterize individual and contextual diversity in the drivers of food choice and behaviour, as well as the corresponding characteristics of innovation, strategies, and operations.

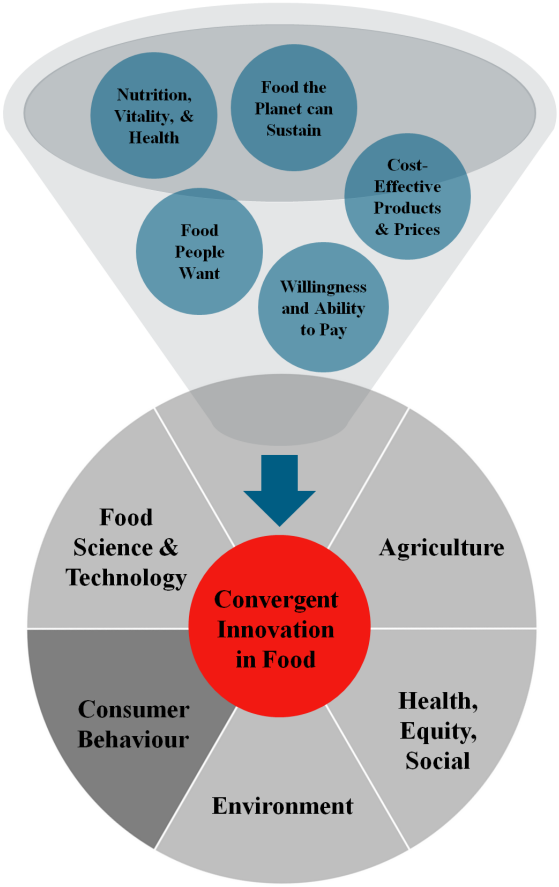

Figure 2. Platform for convergent innovation in food

Using Social Media and Artificial Intelligence in Management and Innovation

In this article, we focus on user-generated content from social media as a source of behavioural insights for convergent innovation. User-generated content refers to any forms of content, such as discussion posts, that are created by end consumers of an online system (e.g., Twitter) and are publically available. The proliferation and increasing availability of user-generated content have revolutionized industry in a new world that blurs the lines between the physical, digital, and biological spheres. Applications of user-generated content extend into a new realm of industrial innovation that takes consumer needs as the entry point for innovation. User-generated content can be used to gain meaningful insights from individuals in their often time-conflicting societal roles as consumers, patients, and citizens. Product attributes that tailor to one of these roles are often conflicting with those that are essential for meeting the needs for other roles. In the context of food, motives and product characteristics that typically please consumers (e.g., taste and convenience) often rate poorly when considering their role as a patient or citizen. Moreover, for food, consumer packaged goods have typically been the focus of marketing research and practice. Yet, consumer packaged goods are just one of many forms food can take. Big data obtained from user-generated content, in the context of retailing, opens immense opportunities along the dimensions of insights relating to consumers, markets, products, purchase/loyalty intent, and advertising at varying time and location-points across physical and digital channels (Bradlow et al., 2017).

Consumer emotion and experience

Innovators and marketers can enrich their understanding of consumers through user-generated content by capturing their behavioural complexity to inform the adaptive design of more competitive and consumer-centric value chains (Kwark et al., 2017). In this new

"market 4.0", information and consumer insights are connected with physical technologies to investigate consumer voices, opinions, and reactions to products and services (Pang et al., 2015; Thomopoulos et al., 2015). Thus far, consumer and marketing research on user-generated content has utilized both supervised and unsupervised text mining for the examination of word co-occurrence and sentiments, extracting product characteristics, quality dimensions (e.g., product cost and product extension), consumer opinions towards brands, and more broadly, general personality characteristics of the consumer (Culotta & Cutler, 2016; Golbeck et al., 2011; Willems & Top, 2015). These techniques allow researchers to better exploit the data and automate processes that traditionally relied on human intervention (Lee & Bradlow, 2011). By using technology to understand such drivers of behaviour, marketing and business strategies can be developed that contribute to society's journey towards sustainable development and affordable healthcare (Dubé et al., 2014; Hammond & Dubé, 2012) and reach the consumer diet and market.

Word of mouth and recommendations

New computational methods for the analysis of user-generated content allow researchers to dive deeply into the understanding of affective experiences through the dimensions of valence (i.e., attractiveness or adverseness) and activation (i.e., awareness and engagement) (De Choudhury et al., 2012). Electronic word-of-mouth communications through social media platforms have a significant influence on consumer behaviour (Babic Rosario et al., 2016). High variability and large volumes of electronic word-of-mouth communications have the largest impact on purchasing behaviour (Babic Rosario et al., 2016). Computational advancements in opinion mining and sentiment analysis allow for businesses to better understand consumer communications and recommendations in relation to their products and services (Pang & Lee, 2006).

Purchase/loyalty intent

Consumer purchase intent and loyalty can also be used by businesses to better understand their consumers. Purchase and loyalty intent can be measured through user-generated content (e.g., a tweet expressing a consumer's desire to purchase a product or service). Intent may be extracted through word- or phrase-based features, as well as through grammatical patterns (Reardon et al., 2014). Furthermore, user-generated content can be classified according to the four stages of the consumer decision journey (i.e., consideration, evaluation, purchase, and post-purchase) (Vázquez et al., 2014). Marketers and innovators can use insights gathered in relation to purchase and loyalty intent for personalized marketing efforts, demand planning, and market-level sensing, as well as to inform innovation and new product development (Reardon et al., 2014).

Market and competitive intelligence

The analysis and utility of user-generated content extend beyond understanding behaviour at the individual level and allows for scalable monitoring and analysis of broader markets, with applications for marketing intelligence and competitive intelligence. However, it is important to note that user-generated content is only one component of omnichannel retailing, and considerations for a broader research perspective are needed as consumers move through channels (physical and digital) in their buying process (Verhoef et al., 2015). User-generated content can be used to predict market trends and outcomes (Asur & Huberman, 2010) and gather competitor intelligence in regards to competing companies' products, promotions, sales, etc. from external sources (Dey et al., 2011). The value of user-generated content from social media sites, analyzed through text-mining and natural language processing technologies, are effective modalities to extract business value and inform strategy (He et al., 2013). Together, the monitoring of consumers, markets, and competitors through user-generated content can be used to inform product development and innovation pipelines.

Product design

Inclusive innovation and convergent innovation models can leverage technological platforms using artificial intelligence and natural language processing to design more consumer-centric products that benefit consumers and the broader health and economic systems. Empirical research that explores the uses of user-generated content from social networking sites in product development and innovation is sparse (Roberts & Candi, 2014). However, it is evident that user-generated content can be used for market research to better inform product development, engage with consumers to co-design new products, and better collaborate in the overall development process in an agile manner (Piller et al., 2012; Roberts & Candi, 2014). Although some researchers have started to use a methodology driven by data mining to analyze user-generated content for next-generation product design, it has yet to be applied to real-time, population-level user-generated content for product development or to predict consumer responses to new products and their respective features (Goel & Goldstein, 2013; Tuarob & Tucker, 2015).

Advertising

It is evident that the insights gathered through the analysis of user-generated content can be applied to many areas of business strategy and practice. In particular, these insights can be used by firms to market and advertise content in a precise way that resonates with consumers. Marketers can tailor advertising efforts to fulfill consumer needs for information, personal identity, and social interaction (Knoll & Proksch, 2017). At an individual level, user-generated content can also be leveraged to precisely advertise to consumers based on user profiles and the content they post (Tucker, 2014). By better-equipping marketers with computational tools that meet the needs and wants of consumers, innovators can better build demand for 21st century products and services that better bridge the divide between health and wealth.

Building an Artificial Intelligence Platform for Convergent Innovation in Food

To support convergent innovation, we have begun the development of integrated modular artificial intelligence platforms. The present article focuses on the social media platform that allows us to collect discussions from social media and to extract users' opinions and sentiments towards different aspects of food. The overall architecture, as shown in Figure 3, is broadly divided into three layers: i) the data collection and management layer, ii) the analysis layer, and iii) the application layer.

The functional stack covers the entire workflow for public opinion analysis towards food. The solution can be easily adapted to other domains with the support of corresponding domain knowledge. The main duty of the first data layer is two-fold: i) to acquire domain-related data from different social media platforms such as Twitter and Facebook; and ii) to manage the ever-increasing data to support efficient input/output operations for future processes. The second analysis layer is a complete text-mining workflow that also informs the construction of the food ontology and acquisition of data using intermediate results. The final layer includes various domain-specific applications built upon the analysis results to support decision making.

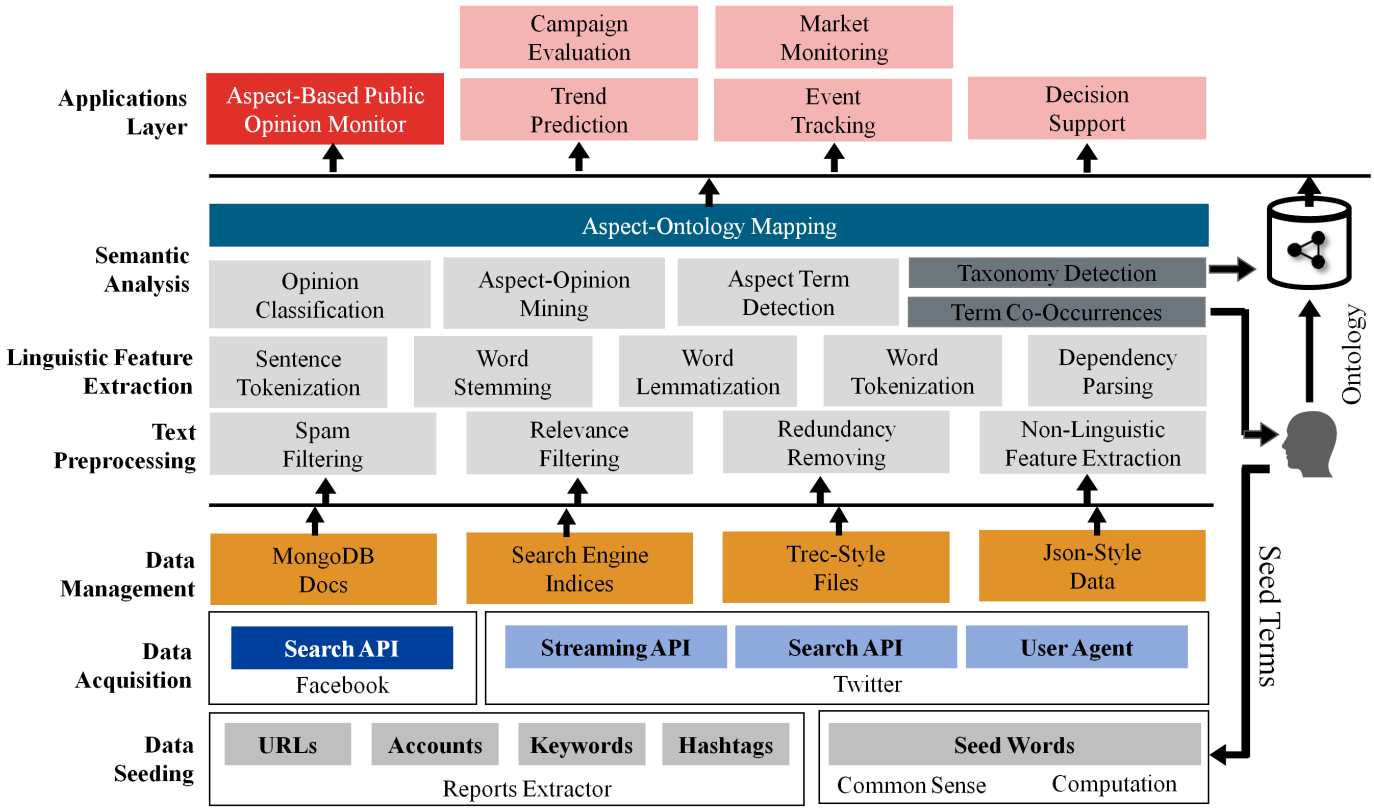

Figure 3. Overview of the artificial intelligence platform architecture

Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth

L. Dubé, P. Du, C. McRae, N. Sharma, S. Jayaraman, and J.-Y. Nie

Data layer

The data layer can be further divided into the functions of seed word acquisition, data collection, and data management. Seed words are used to form search queries submitted to the social media platforms (e.g., Twitter) to collect data. Seed words are domain-specific and are strongly related to the topics investigated (in our case, food). Industry reports, iterative text analysis, and common sense were used to acquire seed words. The candidate words from these sources are manually filtered. For our platform, 359 seed words were collected (including words, phrases, and hashtags) for food.

Given the selected seed words, data is collected in three different ways: i) searching historical data (up to 30 days) via the official application programming interface of Twitter; ii) searching streaming data from social media (Twitter and Facebook) application programming interfaces using the seed words; and iii) using a simulated user-agent to receive new posts on social media. Additionally, we also identify a set of known website URLs and Facebook accounts that are related to food. Data from the corresponding sites are collected automatically.

The data acquired from different social platforms are stored and managed in several different ways depending on the processing purpose. MongoDB is used for real-time input/output operations. The Trec-style data format is used for building search indices and conducting pseudo-relevance feedback searches for relevance filtering. Json-style files are used for intermediate analyzing of results.

Analysis layer

Five types of utilities compose the analysis layer and reflect the processing sequences to which the digital corpus is submitted – i) preprocessing, ii) feature extraction, iii) semantic analysis, iv) taxonomy extraction, and v) aspect-ontology mapping – as described below.

1. *Preprocessing:* The purpose of preprocessing is to filter out possible spam and to recognize the structure of the collected raw data. To filter spam and irrelevant posts, we used the Galago search engine to identify the top results using the seed words and expand search queries based on these top results. The expanded queries allow us to rank the data collected. We consider the low-ranked data as spams. This step filters the number of posts collected down to about 60%. The other 40% of the collection is more prone to spams and low-quality posts. Non-linguistic features such as URLs, time, geolocation, mentions, emojis, retweets, replies, and likes are also extracted in this step.

2. *Feature extraction:* Each text goes through a series of linguistic analysis to recognize the part-of-speech (i.e., noun, verb, etc.) of words and the dependency relation between words (e.g., between verb and subject), to recognize phrases that are stored in our ontology and to transform (or lemmatize) a word into its stem (e.g., "computing", "computed", "computes", and "computation" to "comput"). The connection between the words and phrases in a discussion and the entities stored in our ontology will allow us to identify what aspect of food the discussion is about.

3. *Semantic analysis*: Our semantic analysis focuses on sentiment analysis – to extract the sentiment (or opinion) the user expressed about food or an aspect of food (target) in a discussion. We use our ontology to identify the aspect of food in the discussion, and a sentiment dictionary (SentiWordnet) to identify sentiment words. Target-opinion pairs are extracted by a classical approach based on grammatical rules: an opinion is assumed to be related to a target if they follow some grammatical pattern. For example, from the sentence "My dessert bar was so yummy at yesterday's event!!!", we can identify "dessert bar" as belonging to the "product" aspect, and "yummy" as a sentiment word. The two elements are connected in the sentence through a syntactic relation subject-predicative. Thus, they form a target-opinion pair <dessert bar, yummy>.

4. *Taxonomy extraction* and **5.** *Aspect-ontology mapping:* As mentioned, the ontology (or concept hierarchy of the application) is a key component to connect words in a sentence to food and food aspects. For our platform on food innovation, we performed a statistical analysis on word occurrences on all the raw data collected, with the most frequent words manually filtered and structured into a hierarchy.

Application layer

In the current work, the application layer was built to explore the aspects of food based on public opinion analysis. We consider five aspects of food to inform convergent innovation: behaviour, health, consequence, marketing, and characteristics. The application layer will include several tools, but only the first tool – aspect-based public opinion monitoring – is presently implemented.

Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth

L. Dubé, P. Du, C. McRae, N. Sharma, S. Jayaraman, and J.-Y. Nie

A New Approach to Semantic Analytics

This article presents a novel approach to the aggregation of population-level metadata to predict future market trends and support the development of products and marketing strategies. The early insights in key components of convergent innovation in food will serve as a springboard for articulating the formal knowledge structure that will enable different users to interact with the digital platforms and will define appropriate interfaces between diverse disciplinary and sectoral datasets, models, and rules (Figure 4).

The sentiment analysis (Abbasi et al., 2008; Feldman, 2013; Liu, 2012) follows the path from left to right in Figure 4. As mentioned earlier, natural language processing techniques (Ding et al., 2015; Maas et al., 2011; Nasukawa & Yi, 2003) such as sentence tokenization, word tokenization, stemming, lemmatization, dependency tree parsing, etc. are leveraged to acquire linguistic features for extraction of aspect terms and opinion terms as well as their relations. Rules (Liu, 2012) and automatic text-classification approaches (Mullen & Collier, 2004) were implemented for sentiment extraction. The aspects about food are identified using our ontology (Figure 5), and sentiment words are identified using external knowledge resources such as SentiWordnet (Baccianella et al., 2010; Miller, 1995).

The first task of sentiment analysis is to determine the polarity (positive, negative, or neutral) of a sentiment. In SentiWordnet, each synset (a set of synonymous

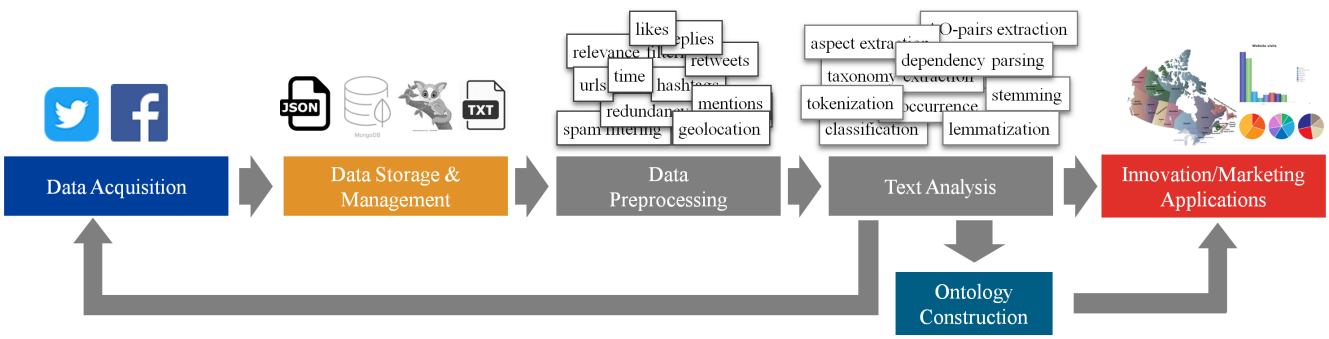

Figure 4. The digital platform architecture

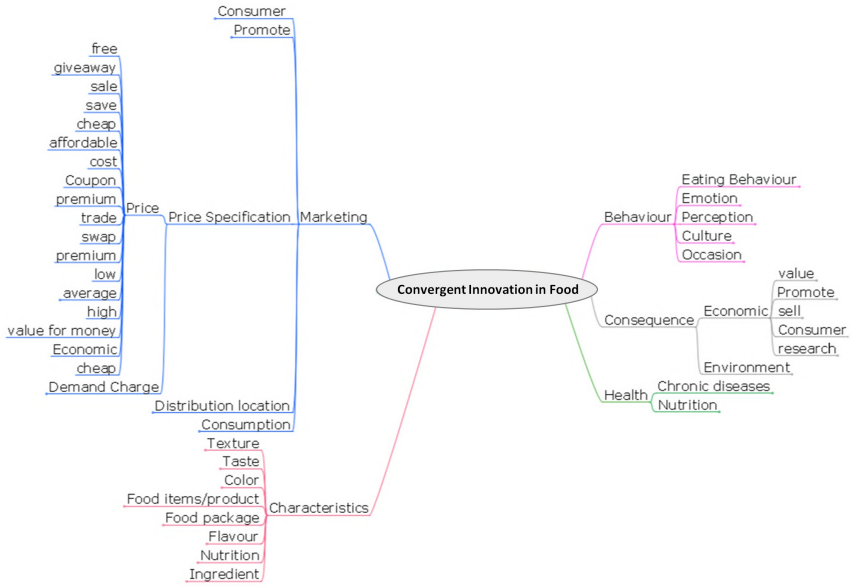

Figure 5. Vocabulary tree structure for convergent innovation in food

words) is assigned three scores: *Pos, Neg,* and *Obj,* describing their polarity distribution in general. For example, for a synset of "yummy", *Pos*(yummy)=0.75, *Neg*(yummy)=0.25, *Obj*(yummy)=0. We simply assign the strongest polarity to the opinion in our current implementation. Thus, our earlier example of <dessert bar, yummy> will be extended to the following triple <dessert bar, yummy, positive>. This method can be further extended in the future to take into account the context (e.g., considering the target the opinion word is describing). The triples extracted from sentences form the set of basic opinions recognized from user discussions.

The focus of the present exploratory study is to look at consumer behaviour in relation to food using data from social media. To this end, we use the basic opinions extracted for the analysis of sentiment scores, distributions, and influence on global sentiment (Figure 6). Results from the analysis uncover consumer likes/dislikes about food aspects, as well as the drivers of behaviour through regression modelling.

Results

Over 26 million posts about food from Twitter and about 1 million posts from Facebook were collected during the summer and fall of 2017. Most posts do not express an explicit sentiment or opinion. From this set of data, about 70,000 target-opinion pairs were extracted. The distribution of opinions on different aspects, and how the opinions on an aspect influence the global sentiment of the user and post, were the subjects of our analysis.

Sentiment probability distribution

From the extracted set of associations between the aspects and sentiments, we performed statistical analysis on the distribution of sentiments across aspects of the first- and second-order, concerning the aspects included in our ontology. The distribution of sentiments about the first level aspects (marketing, behaviour, health, consequence and characteristics) is summarized in Table 1a. The distribution on the sub-aspects is shown in Table 1b. Each value in the tables represents the joint probability. For example, *P* (Marketing, positive)=0.135, meaning that 13.5% of the sentiments detected are positive about marketing. In Table 1b, each number represents the joint probability of sub-aspect and polarity among the sentiments related to that aspect. For example, among all the sentiments expressed on sub-aspects of marketing, 30.7% are positive about the sub-aspect price specification. Notice that in Table 1b, we only count the sentiments expressed on the sub-aspects, while ignoring those that are expressed on the aspects directly.

Here, we observe that social media users discuss marketing (*P*=0.135) and the consequences of food (*P*=0.133) more positively than other aspects. Marketing-related aspects include price, promotional activities, placement, and industrial sector, and are mostly discussed with a neutral sentiment with a slightly positive tendency. Similarly, environmental and economic consequences are discussed with a mostly neutral-to-positive sentiment. Food characteristics, including colour, texture, nutrition, packaging, and preparation method were also discussed with a mostly positive sentiment. In general, social media consumers tend to talk about food with neutral (57.6% of discussions) to positive (37.9% of discussions) sentiment. Only 4.5% of discussions included a negative sentiment towards food overall.

The results of second-order sentiment aspects are presented in Table 1b. Within marketing, discussions in

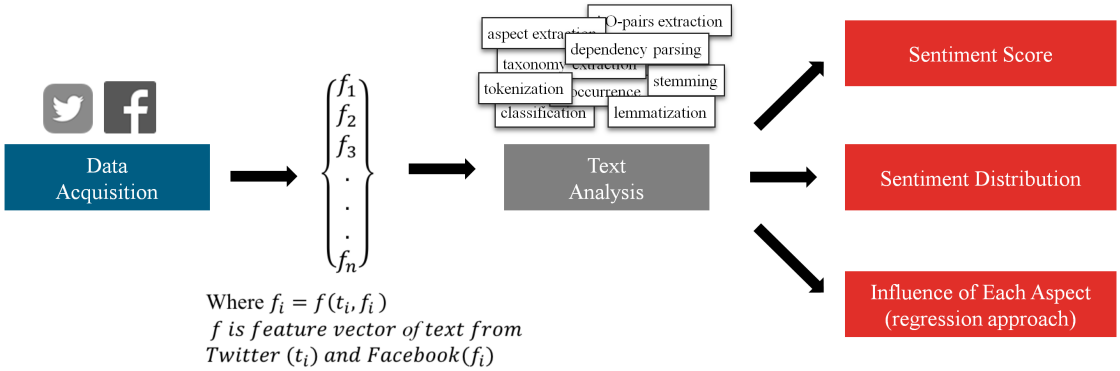

Figure 6. Flowchart for empirical analysis

Table 1. Sentiment distribution of aspects and sub-aspects

a. Sentiment distribution on aspects

Aspects	Positive	Negative	Neutral	Total
Marketing	0.135	0.014	0.278	0.427
Behaviour	0.018	0.000	0.002	0.020
Health	0.008	0.001	0.001	0.010
Consequence	0.133	0.012	0.278	0.423
Characteristics	0.085	0.018	0.017	0.120
Total	0.379	0.045	0.576	

b. Sentiment distribution on sub-aspects

Sub-Aspect	Positive	Negative	Neutral	Total
Marketing				
Price specification	0.307	0.029	0.651	0.987
Promotion	0.002	0.000	0.000	0.002
Consumer	0.002	0.002	0.001	0.005
Consumption	0.005	0.001	0.000	0.006
Total (marketing)	0.316	0.032	0.652	
Behaviour				
Emotion	0.000	0.000	0.005	0.005
Culture	0.019	0.000	0.010	0.029
Eating behaviour	0.002	0.000	0.000	0.002
Perception	0.021	0.000	0.022	0.043
Occasion	0.826	0.015	0.080	0.921
Total (behaviour)	0.868	0.015	0.117	
Health				
Nutrition	0.610	0.101	0.110	0.821
Benefit	0.000	0.006	0.003	0.009
Well	0.130	0.024	0.012	0.166
Total (health)	0.74	0.131	0.125	
Consequence				
Environment	0.002	0.000	0.000	0.002
Economic	0.312	0.029	0.656	0.997
Total (consequence)	0.314	0.029	0.656	
Characteristics				
Product	0.274	0.009	0.026	0.309
Packaging	0.001	0.000	0.000	0.001
Taste	0.009	0.000	0.002	0.011
Texture	0.002	0.000	0.001	0.003
Colour	0.002	0.000	0.001	0.003
Ingredients	0.424	0.140	0.108	0.672
Total (characteristics)	0.712	0.149	0.138	

relation to price are most probable to be neutral (P=0.651), or positive (P=0.307) to a lesser extent, hinting that consumers have a slightly positive association with food prices on a global scale. As for the behaviour sub-aspect, nearly all discussions are associated with positive eating occasions (e.g., birthdays or holidays). Consumers also have a tendency to discuss nutrition in relation to health. Most social media posts surrounding nutrition are positive (P=0.610) in nature (e.g., discussion of high-protein content with positive sentiment). Health benefits and wellness are less discussed by consumers. The economic consequences of food (i.e., economic impacts) are frequently discussed with a neutral (65.6%) or positive (31.2%) sentiment, whereas the environmental impact of food is less discussed. As for food characteristics, we evaluated aspects of packaging, taste, texture, colour, ingredients, and product. Ingredients and products are the most discussed aspects of food characteristics, according our analysis. Although a large proportion of discussions regarding ingredients are positive (42.4%), a noticeable proportion of discussion also expresses negative opinions (14%). The results of this analysis reveal that consumers discuss products and ingredients most on social media. The present analysis is limited in terms of depth (i.e., number of levels we can uncover below each aspect). Future studies will address data limitations to dive deeper into acquiring more meaningful insights (e.g., what type of texture is most positive or negative). The insights gleaned from this analysis and future iterations will inform the development and marketing of food products aimed at the convergent innovation sweet spot illustrated in Figure 2.

Influence of each aspect on sentiment

The aspects explored in this analysis may contribute different degrees to the overall sentiments expressed by a consumer expressed in a post or any unit of user-generated content, therefore impacting demand for food in general or for specific products or contexts. To determine the influence of each aspect, we evaluated the sentiments of an aspect in relation to its prediction of the global sentiment of user-generated content. We use linear regression modelling to calculate the regression coefficients for each aspect (Kutner et al., 2004; Seber & Lee, 2012), as shown in Table 2. The sentiment of an aspect valence is between -1 and 1: -1 (negative), 0 (neutral), or 1 (positive). The regression formula is as follows:

$$Sentiment_{post} = \sum_{aspect_i} \alpha_i * Sentiment_i + S_0 \quad (1)$$

where $Sentiment_{post}$ is the global sentiment valence of a post; $Sentiment_i$ is the sentiment valence of $aspect_i$; a_i is its coefficient, which reflects the importance of the aspect for the global sentiment; S_0 is a constant which captures the general trend of sentiment in tweets, independently from the aspects.

To perform the regression analysis, we have to detect the global sentiment of a post. Therefore, a trained classifier was used to analyze the social media data. Applying this classifier, each post was automatically assigned a sentiment valence between -1 and 1.

The regression task aims to reproduce the global sentiment polarities using the sentiments about food aspects observed in the post. Table 2 shows the coefficients obtained hierarchically in the linear regressions. In Table 2a, general sentiment predictions are made from the sentiments of the first-level aspects. In Table 2b, the sentiment of the aspect is predicted from those on the second-level aspects. In Table 2c, the sentiment of sub-aspects level 2 is predicted from those of the level 3 (sub-sub-aspects). Notice that when we move down to sub- and sub-sub-aspects, we have more and more data sparseness, where fewer sentiments are expressed on lower-level aspects. Therefore, the analysis cannot be done at a very deep level with the data collected.

The global sentiment results infer that the overall sentiment of the food is strongly correlated with marketing (a_i=0.032, p<0.01), behaviour (a_i=0.203, p<0.001), consequences (a_i=0.031, p<0.01), and characteristics (a_i=0.069, p<0.001) sub-aspects. Health is not found to be a significant predictor of food sentiment.

Further analysis was conducted on how the sentiments of second-order aspects influence that of an aspect (Table 2b). Again, a linear regression was performed using a sub-corpus for each of the aspects, where the tweets in the sub-corpus only contain tweets relating to the aspect and its sub-aspects. Most sub-aspects in level 2 were found not to be significant. Within marketing, price specification (a_i=0.288, p<0.10) and promotional activities (a_i=0.051, p<0.10) are significant predictors of positive global food sentiments, with the price having the strongest effect. Relatedly, the economic consequences (a_i=0.034, p<0.10) of food are also significant predictors of global sentiment. Sub-aspects found in the layers under behaviour, health, and characteristics were non-significant predictors of global sentiment.

Table 2. Hierarchical estimation of contribution of aspects and sub-aspects to global sentiment

a. Level 1 coefficients

Aspects	α_i
Marketing	0.032†
Behaviour	0.203§
Health	0.015
Consequences	0.031†
Characteristics	0.069§

b. Level 2 coefficients

Sub-Aspects	α_i
Marketing	
Consumer	0.042
Price specification	0.288*
Promotion	0.051*
Consumption	0.107
Behaviour	
Emotion	0.046
Culture	0.164
Eating behaviour	0.578
Perception	0.081
Occasion	0.103
Health	
Nutrition	0.02
Benefit	0.001
Healthy	0.199
Wellness	0
Consequence	
Environment	0
Economic	0.034*
Characteristics	
Product	0.218
Packaging	0.198
Taste	0.064
Texture	0.101
Colour	-0.086
Ingredients	0.029

c. Level 3 Aspects coefficients

Sub-Sub-Aspects	α_i
Economic	
Promote.Awareness	0.278*
Research.Question	0.464
Sell.Demand	-0.051
Sell.Price	0.034*
Research.Report	-0.128
Promote.Target	-0.752
Price Specification	
Price.Affordable	-0.096
Charge.Value	0.232
Price.Cheap	-0.005
Price.Cost	0
Price.High	0.236
Charge.Tax	0
Price.Sale	0.197
Charge.Affordable	-0.096
Price.Low	0.426
Charge.Cost	0.035*
Charge.Profit	0
Price.Giveaway	0.066
Charge.Rate	-0.167
Charge.Bill	-0.199
Charge.Cheap	-0.005

Significance:
> 0.10 non-significant
* 0.10–0.05
** 0.05–0.01
† 0.01–0.001
§ < 0.001

Price specification and economic consequences may be further decomposed with a third level. Although most sub-sub-aspects in this level were non-significant, price (a_i=0.034, p<0.10) and promotional awareness (a_i=0.278, p<0.10) were significant predictors of positive global sentiments, with promotional awareness having the largest effect. Cost (a_i=0.035, p<0.10) was also a significant predictor of positive sentiment associated with price specification.

The above analysis constrains sub-aspects to impact global sentiment hierarchically, meaning that each sub-aspect is the predictor of the sentiment of the immediate superior sub/aspect, with ultimately the five first-level aspects (marketing, behaviour, health, consequences, and characteristics) being predictors of the global sentiments. To capture more of the richness of user generated content, we conducted a complementary analysis where all level 2 sub-aspects were used as direct predictors of global sentiments. Predictors of a global positive sentiment emerged as culture (a_i=0.205, p<0.001), emotion (a_i=0.031, p<0.05), and perception (a_i=0.033, p<0.01). Related to the characteristics of food products themselves, colour (a_i=0.165, p<0.05), packaging (a_i=0.100, p<0.10), taste (a_i=0.424, p<0.001), texture (a_i=0.126, p<0.001), and ingredients (a_i=0.044, p<0.05) were significant predictors of sentiment. Taste had the largest effect. Nutrition (a_i=0.083, p<0.05), under health, was also significant.

Table 3. Estimation of contribution of level 2 sub-aspect to global sentiment

Sub-Aspect	α_i
Behaviour.Culture	0.205§
Behaviour.Eating_Behaviour	0.805
Behaviour.Emotion	0.031**
Behaviour.Occasion	-0.043
Behaviour.Perception	0.033†
Characteristics.Food_Colour	0.165**
Characteristics.Ingredient	0.044**
Characteristics.Package	0.100*
Characteristics.Product	0.039
Characteristics.Taste	0.424§
Characteristics.Texture	0.126§
Consequence.Economic	0.168*
Consequence.Environment	0.107
Health.Benefit	0.221
Health.Nutrition	0.083**
Health.Well	0.007
Marketing.Consumer	0.369**
Marketing.Consumption	0.289**
Marketing.Price Specification	0.635
Marketing.Promote	0.232†

Significance: > 0.10 non-significant
* 0.10–0.05 ** 0.05–0.01
§ < 0.001 † 0.01–0.001

Conclusion and Future Research

Results show that positive and negative drivers of demands for convergent innovation in food, as expressed in this digital social media corpus, bear on their own belief systems, experiences, and culture, as well as the characteristics of the food they associate with and the expected consequences that motivate their behaviour. Environmental concerns did not emerge as salient and significant to consumer sentiments in this corpus, which may be tied to the present ontological structure as well as to the sample composition. Further research will explore this issue. In addition, although we report results from estimation for the overall user-generated content corpus, similar analyses could be performed for sub-samples formed on the basis of consumer segments, type of food, competitive products, geographical markets, etc.

From the perspective of inclusive innovation being concerned primarily with economic equity issues, the results presented in Table 2 revealed not only that economic consequence is a strong predictor of consumer global sentiment but that it is also sensitive to both the actual price and economic awareness. This underscores the importance of the complementary strategy to not only make food accessible at the appropriate price but also to inform and educate consumers of the value and ways to estimate price as it may be related to nutrition and other dimensions of what they need. This may be particularly relevant for disadvantaged population segments that are typically the target of inclusive innovation efforts.

An important limitation to the early-state results presented in this article is that, despite attempts to collect sufficient data, we have faced data sparseness problems that may lead to counterintuitive conclusions. The current corpus supports the analysis of the first-order aspects in the ontology hierarchy. However, to understand the lower-level concepts (aspects), a much larger corpus will be needed to provide sufficient support for each aspect, especially when the concept space keeps expanding. As a limitation of the current study, further exploration of the aspects of behaviour and health have been excluded and will be considered in future work.

In spite of these limitations, the results of the present study reflect the rich diversity of positive and negative drivers of consumer demand for food products covering the full spectrum from expected consequences to cultural, social, and emotional features of the experience, to characteristics of the consumption occasion, to actual product design features and the marketing mix. Although this first extraction and analysis do not allow us to capture the set of relationships between these influential factors, future work will combine relevant theoretical and empirical basis with deep learning and other artificial intelligence methods to trace the pathways through which a vibrant ecosystem can be created that supports the supply and demand of a portfolio of food that people need, want, and are willing and able to pay for, and that the ecosystem actors are able and willing to produce. These results add to the existing understanding of consumer behaviour for food most often theorized from an unhealthy/tasty bipolar view for both advantaged and disadvantaged populations, and provides insights on the complex systems of beliefs, motives, and goals encompassing familial and social bonds and norms, cultural meanings, and other considerations impacting consumer responses to food innovation or communication. Our results provide insights to

find or create a convergence path among systems of food beliefs, motives, and goals leading to individual healthy food behaviours that are sustainable from all these perspectives be they biological, psychological, cultural, economic, or environmental. In fact, research reporting results of geographical analysis of user-generated content in the future will provide geo-referenced information on the influence of food and food cues on food choice and suggest possibilities of fine-grained differentiation of consumer insights for better-targeted convergent innovation in food.

The future scope envisioned for the integrated digital architecture for convergent innovation in food is to combine the social media platform with others modules enabling the dynamic integration of past and present sectoral and intersectoral knowledge and metrics. We also plan to move toward predictive models that can link complex webs of relationships involved in specific innovation and marketing practices with their single and collective economic, social, and environmental outcomes that will benefit the firm and society.

Acknowledgements

Funding for the research programs reported in this article comes partly from an SSHRC operating grant (#435-2014-1964), a FRQSC team grant (#2015-SE-179342), and seed project funding from the International Development Research Centre (#107400-006) to Laurette Dubé. Additional support comes from the Global Pulse Innovation Platform at the McGill Centre for the Convergence of Health and Economics.

About the Authors

Laurette Dubé is a Full Professor and holds the James McGill Chair of Consumer and Lifestyle Psychology and Marketing at the Desautels Faculty of Management of McGill University in Montreal, Canada. Her research interest bears on the study of affects and behavioural economic processes underlying consumption and lifestyle behaviour and how such knowledge can inspire more effective health and marketing communications in both real life and technology-supported media. She is the Founding Chair and Scientific Director of the McGill Centre for the Convergence of Health and Economics (MCCHE). The MCCHE was created to foster partnerships among scientists and decision makers from all sectors of society to encourage a more ambitious notion of what can be done for more effective health management and novel pathways for social and business innovation.

Pan Du is a Research Associate in the Department of Computer Science and Operational Research at the Université de Montréal, Canada. Before that, Pan was an Assistant Professor at the Chinese Academy of Sciences. He received his PhD from the Institute of Computing Technology of the Chinese Academy of Sciences. His research interests lie in text mining, information retrieval, machine learning, and social network analysis. He has published academic papers in various conferences and journals. He is a recipient of the 2016 "Science and Technology Progress Award" of the Chinese Institute of Electronics for his contribution to a web-scale text mining system.

Cameron McRae is a Senior Research Analyst at the McGill Centre for the Convergence of Health and Economics in Montreal, Canada. Since joining the centre in 2014, he has led many translational research projects to support innovation in the agricultural, food, and health sectors. Cameron has strong interdisciplinary training at the nexus of science, technology, and management, with a Bachelor of Science in Pharmacology from McGill University, a Graduate Certificate in Business Administration from John Molson School of Business, and a Master of Health Informatics from the University of Toronto. Previously, Cameron has worked in both the public and private sectors to support strategy and practice in the areas of governance, business development, and business/market intelligence related to life sciences and digital health.

Continued on next page...

Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth

L. Dubé, P. Du, C. McRae, N. Sharma, S. Jayaraman, and J.-Y. Nie

Neha Sharma is currently pursuing her PhD at the Department of Bioresource Engineering at McGill University in Montreal, Canada. She completed her Master's degree in Biochemical Engineering from Harcourt Butler Technical University, India. The title of her Master's research project was "Optimization of Process parameters for Bacterial solid-state fermentation of Nattokinase to prevent myocardial infarction", which culminated in principles of food processing, microbiology, and bioprocessing. Her Bachelor's degree in Biotechnology is from IMS Engineering College, India, where she took various courses in molecular biology, genetic engineering, bioprocess engineering, fermentation biotechnology, food biotechnology, and environmental biotechnology, etc. In her final year, her Bachelor's project was based on the study of plant extracts and their antimicrobial properties.

Srinivasan Jayaraman is a Research Associate/Visiting Scholar at the Desautels Faculty of Management, at McGill University in Montreal, Canada. He obtained his Bachelor's degree in Electronics and Instrumentation Engineering from Bharathidasan University, India, his MTech degree in Biomedical Engineering from SASTRA University in Thanjavur, India, and his doctorate from the School of Bioscience at the Indian Institute of Technology Madras in Chennai, India. Previously, he has held roles at TCS Innovation Labs, the University of Nebraska Lincoln, the New Jersey Institute of Technology, and INRS-EMT Canada. His research interests include human behavioural and performance modelling, ontology, ergonomics, personalized diagnosis systems, wearable devices, biosignal processing, and human-machine interfaces. In 2011, he won the MIT-TR35 young innovator award Indian edition and was recognized as one among the Top 50 most impactful social innovators (global listing) by World CSR Congress & World CSR Day at 2016.

Jian-Yun Nie is a Professor in Computer Science at the University of Montreal, Canada, and is associated with the IVADO institute. He obtained a PhD degree from Université Joseph Fourier of Grenoble, France. He specializes in information retrieval, natural language processing, and artificial intelligence. He has been doing research in these areas for 30 years and has published many papers on these topics. He has served as general chair and PC chair for several conferences in the area of information retrieval. He is on the board of several international journals, including *Information Retrieval Journal*. He has been an invited researcher at several institutions (Tsinghua University, Peking University) and companies (Microsoft Research, Baidu, and Yahoo!).

References

Abbasi, A., Chen, H., & Salem, A. 2008. Sentiment Analysis in Multiple Languages. *ACM Transactions on Information Systems,* 26(3): 1–34. https://doi.org/10.1145/1361684.1361685

Ali, I., & Son, H. H. 2007. Measuring Inclusive Growth. *Asian Development Review,* 24(1): 11–31.

Asur, S., & Huberman, B. A. 2010. Predicting the Future with Social Media. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT):* 492–499. https://doi.org/10.1109/WI-IAT.2010.63

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. 2016. The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research,* 53(3): 297–318. https://doi.org/10.1509/jmr.14.0380

Baccianella, S., Esuli, A., & Sebastiani, F. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining SentiWordNet. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10),* May 19–21, Valletta, Malta.

Beinhocker, E. D. 2007. *The Origin of Wealth: The Radical Remaking of Economics and What It Means for Business and Society.* Boston, MA: Harvard Business Press.

Bradlow, E. T. T., Gangwar, M., Kopalle, P., & Voleti, S. 2017. The Role of Big Data and Predictive Analytics in Retailing. *Journal of Retailing,* 93(1): 79–95. https://doi.org/10.1016/j.jretai.2016.12.004

Campbell, J. L. 2007. Why Would Corporations Behave in Socially Responsible Ways? An Institutional Theory of Corporate Social Responsibility. *Academy of Management Review,* 32(3): 946–967. http://doi.org/10.5465/AMR.2007.25275684

Chataway, J., Hanlin, R., & Kaplinsky, R. 2014. Inclusive Innovation: An Architecture for Policy Development. *Innovation and Development,* 4(1): 33–54. https://doi.org/10.1080/2157930X.2013.876800

Croitoru, A. 2012. Schumpeter, J.A., 1934 (2008), The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle. *Journal of Comparative Research in Anthropology and Sociology,* (2): 137–148.

Culotta, A., & Cutler, J. 2016. Mining Brand Perceptions from Twitter Social Networks. *Marketing Science,* 35(3): 343–362. https://doi.org/10.1287/mksc.2015.0968

De Choudhury, M., Counts, S., & Gamon, M. 2012. Not All Moods are Created Equal! Exploring Human Emotional States in Social Media. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media:* 66–73.

Devaux, A., Horton, D., Velasco, C., Thiele, G., López, G., Bernet, T., Reinoso, I., & Ordinola, M. 2009. Collective Action for Market Chain Innovation in the Andes. *Food Policy,* 34(1): 31–38. https://doi.org/10.1016/j.foodpol.2008.10.007

Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth

L. Dubé, P. Du, C. McRae, N. Sharma, S. Jayaraman, and J.-Y. Nie

Dey, L., Haque, S. M., Khurdiya, A., & Shroff, G. 2011. Acquiring Competitive Intelligence from Social Media. In *Proceedings of the 2011 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data,* September 17, 2011, Beijing, China. https://doi.org/10.1145/2034617.2034621

Ding, X., Liu, T., Duan, J., & Nie, J. 2015. Mining User Consumption Intention from Social Media Using Domain Adaptive Convolutional Neural Network. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence:* 2389–2395. January 25–29, 2015, Austin, TX.

Drayton, B., & Budinich, V. 2010. A New Alliance for Global Change. *Harvard Business Review,* 88(9): 56–64.

Dubé, L., Bechara, A., Böckenholt, U., Ansari, A., Dagher, A., Daniel, M., DeSarbo, W. S., Fellows, L. K., Hammond, R. A., Huang, T. T.-K., Huettel, S., Kestens, Y., Knäuper, B., Kooreman, P., Moore, D. S., & Smidts, A. 2008. Towards a Brain-to-Society Systems Model of Individual Choice. *Marketing Letters,* 19(3–4): 323–336. https://doi.org/10.1007/s11002-008-9067-9

Dubé, L., Jha, S., Faber, A., Struben, J., London, T., Mohapatra, A., Drager, N., Lannon, C., Joshi, P. K., & McDermott, J. 2014. Convergent Innovation for Sustainable Economic Growth and Affordable Universal Health Care: Innovating the Way We Innovate. *Annals of the New York Academy of Sciences,* 1331(1): 119–141. http://doi.org/10.1111/nyas.12548

Dubé, L., Pingali, P., & Webb, P. 2012. Paths of Convergence for Agriculture, Health, and Wealth. *Proceedings of the National Academy of Sciences,* 109(31): 12294–12301. https://doi.org/10.1073/pnas.0912951109

Feldman, R. 2013. Techniques and Applications for Sentiment Analysis. *Communications of the ACM,* 56(4): 82–89. https://doi.org/10.1145/2436256.2436274

Gillespie, S., Haddad, L., Mannar, V., Menon, P., & Nisbett, N. 2013. The Politics of Reducing Malnutrition: Building Commitment and Accelerating Progress. *The Lancet,* 382(9891): 552–569. https://doi.org/10.1016/S0140-6736(13)60842-9

Goel, S., & Goldstein, D. G. 2013. Predicting Individual Behavior with Social Networks. *Marketing Science,* 33(1): 82–93. https://doi.org/10.1287/mksc.2013.0817

Golbeck, J., Robles, C., & Turner, K. 2011. Predicting Personality with Social Media. In *Proceedings of CHI'11: Extended Abstracts on Human Factors in Computing Systems:* 253–262. https://doi.org/10.1145/1979742.1979614

Hammond, R. A., & Dubé, L. 2012. A Systems Science Perspective and Transdisciplinary Models for Food and Nutrition Security. In *Proceedings of the National Academy of Sciences,* 109(31): 12356–12363. https://doi.org/10.1073/pnas.0913003109

He, W., Zha, S., & Li, L. 2013. Social Media Competitive Analysis and Text Mining: A Case Study in the Pizza Industry. *International Journal of Information Management,* 33(3): 464–472. https://doi.org/10.1016/j.ijinfomgt.2013.01.001

Knoll, J., & Proksch, R. 2017. Why We Watch Others' Responses to Online Advertising–Investigating Users' Motivations for Viewing User-Generated Content in the Context of Online Advertising. *Journal of Marketing Communications,* 23(4): 400–412. https://doi.org/10.1080/13527266.2015.1051092

Kochanek, K. D., Murphy, S. L., Xu, J., & Arias, E. 2016. *Mortality in the United States, 2016.* NCHS Data Brief No. 293. Atlanta, GA: National Center for Health Statistics (NCHS).

Kutner, M. H., Nachtsheim, C. J., & Neter, J. 2004. *Applied Linear Regression Models* (4th Ed.). New York: McGraw-Hill/Irwin.

Kwark, Y., Chen, J., & Raghunathan, S. 2017. User-Generated Content and Product Design of Competing Firms. *Management Science,* Forthcoming. https://doi.org/10.1287/mnsc.2017.2839

Lee, T. Y., & Bradlow, E. T. 2011. Automated Marketing Research Using Online Customer Reviews. Journal of Marketing Research, 48(5): 881–894. https://doi.org/10.1509/jmkr.48.5.881

Lin, Y.-C., & Chang, C. A. 2012. Double Standard: The Role of Environmental Consciousness in Green Product Usage. *Journal of Marketing,* 76(5): 125–134. https://doi.org/10.1509/jm.11.0264

Liu, B. 2012. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies,* 5(1): 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. 2011. Learning Word Vectors for Sentiment Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies:* 142–150. Portland, Oregon, June 19–24, 2011.

Martin, R. L., & Osberg, S. 2007. Social Entrepreneurship: The Case for Definition. *Stanford Social Innovation Review,* Spring 2007: 29–39.

Martin, R. L., & Osberg, S. 2015. *Getting Beyond Better: How Social Entrepreneurship Works.* Boston, MA: Harvard Business Review Press.

Miller, G. A. 1995. WordNet: A Lexical Database for English. *Communications of the ACM,* 38(11): 39–41. https://doi.org/10.1145/219717.219748

Moodie, R., Stuckler, D., Monteiro, C., Sheron, N., Neal, B., Thamarangsi, T., Lincoln, P., & Casswell, S. 2013. Profits and Pandemics: Prevention of Harmful Effects of Tobacco, Alcohol, and Ultra-Processed Food and Drink Industries. *The Lancet,* 381(9867): 670–679. https://doi.org/10.1016/S0140-6736(12)62089-3

Mullen, T., & Collier, N. 2004. Sentiment Analysis Using Support Vector Machines with Diverse Information Sources. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing:* 412–418. July 25–26, Barcelona, Spain.

Nasukawa, T., & Yi, J. 2003. Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In *Proceedings of the 2nd International Conference on Knowledge Capture:* 70–77. October 23–25, Sanibel Island, Florida.

Pang, B., & Lee, L. 2006. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval,* 2(1-2): 91–231.

Pang, Z., Chen, Q., Han, W., & Zheng, L. 2015. Value-Centric Design of the Internet-of-Things Solution for Food Supply Chain: Value Creation, Sensor Portfolio and Information Fusion. *Information Systems Frontiers,* 17(2): 289–319. https://doi.org/10.1007/s10796-012-9374-9

Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth

L. Dubé, P. Du, C. McRae, N. Sharma, S. Jayaraman, and J.-Y. Nie

Pansera, M., & Martinez, F. 2017. Innovation for Development and Poverty Reduction: An Integrative Literature Review. *Journal of Management Development,* 36(1): 2–13. https://doi.org/10.1108/JMD-02-2015-0013

Piketty, T. 2014. *Capital in the Twenty-First Century.* Boston, MA: Harvard University Press.

Piller, F., Vossen, A., & Ihl, C. 2012. From Social Media to Social Product Development: The Impact of Social Media on Co-Creation of Innovation. *Die Unternehmung,* 66(1): 7–27.

Planes-Satorra, S., & Paunov, C. 2017. *Inclusive Innovation Policies: Lessons from International Case Studies.* OECD Science, Technology and Industry Working Papers, 2017(2): 1. Paris: Organisation for Economic Co-operation and Development (OECD).

Prahalad, C. K., Hart, S. L. 2002. The Fortune at the Bottom of the Pyramid. *Strategy+Business,* 26: 54–67.

Prahalad, C. K., & Hammond, A. 2002. Serving the World's Poor, Profitably. *Harvard Business Review,* 80(9): 48–59.

Reardon, T., Chen, K. Z., Minten, B., Adriano, L., Dao, T. A., Wang, J., & Gupta, S. D. 2014. The Quiet Revolution in Asia's Rice Value Chains. *Annals of the New York Academy of Sciences,* 1331(1): 106–118. http://doi.org/10.1111/nyas.12391

Reardon, T., Timmer, C. P., & Minten, B. 2012. Supermarket Revolution in Asia and Emerging Development Strategies to Include Small Farmers. *Proceedings of the National Academy of Sciences,* 109(31): 12332–12337. https://doi.org/10.1073/pnas.1003160108

Roberts, D. L., & Candi, M. 2014. Leveraging Social Network Sites in New Product Development: Opportunity or Hype? *Journal of Product Innovation Management,* 31(S1): 105–117. http://doi.org/10.1111/jpim.12195

Schillo, R. S., & Robinson, R. M. 2017. Inclusive Innovation in Developed Countries: The Who, What, Why, and How. *Technology Innovation Management Review,* 7(7): 34–46. https://timreview.ca/article/1089

Seber, A. J., & Lee, G. A. F. 2012. Linear Regression Analysis (2nd Ed.). Hoboken, NJ: John Wiley & Sons.

Stiglitz, J. E. 2016. An Agenda for Sustainable and Inclusive Growth for Emerging Markets. *Journal of Policy Modeling,* 38(4): 693–710. http://dx.doi.org/10.1016/j.jpolmod.2016.05.012

Thomopoulos, R., Croitoru, M., & Tamani, N. 2015. Decision Support for Agri-Food Chains: A Reverse Engineering Argumentation-Based Approach. *Ecological Informatics,* 26(P2): 182–191. https://doi.org/10.1016/j.ecoinf.2014.05.010

Tuarob, S., & Tucker, C. S. 2015. Quantifying Product Favorability and Extracting Notable Product Features Using Large Scale Social Media Data. *Journal of Computing and Information Science in Engineering,* 15(3): 31003. http://doi.org/10.1115/1.4029562

Tucker, C. E. 2014. Social Networks, Personalized Advertising, and Privacy Controls. *Journal of Marketing Research,* 51(5): 546–562. https://doi.org/10.1509/jmr.10.0355

Vázquez, S., Muñoz-García, Ó., Campanella, I., Poch, M., Fisas, B., Bel, N., & Andreu, G. 2014. A Classification of User-Generated Content into Consumer Decision Journey Stages. *Neural Networks,* 58: 68–81. https://doi.org/10.1016/j.neunet.2014.05.026

Verhoef, P. C., Kannan, P. K., & Inman, J. J. 2015. From Multi-Channel Retailing to Omni-Channel Retailing. Introduction to the Special Issue on Multi-Channel Retailing. *Journal of Retailing,* 91(2): 174–181. https://doi.org/10.1016/j.jretai.2015.02.005

Willems, D. J. M., & Top, J. L. 2015. *Social Media: Embedding in the eFoodLab Framework.* Report 1532. Wageningen, Netherlands: Wageningen UR Food & Biobased Research.

Citation: Dubé, L., Du, P., McRae, C., Sharma, N., Jayaraman, S., & Nie, J.-Y. 2018. Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth. *Technology Innovation Management Review,* 8(2): 49–65. http://doi.org/10.22215/timreview/1139

Keywords: convergent innovation, artificial intelligence, user-generated content, social media, food

Technology Innovation Management Review

timreview.ca

Academic Affiliations and Funding Acknowledgements

Canada

The Federal Economic Development Agency for Southern Ontario (FedDev Ontario; feddevontario.gc.ca) is part of the Innovation, Science and Economic Development portfolio and one of six regional development agencies, each of which helps to address key economic challenges by providing regionally-tailored programs, services, knowledge and expertise.

- *The TIM Review receives partial funding from FedDev Ontario's Investing in Regional Diversification initiative.*

Carleton UNIVERSITY

TIM

Technology Innovation Management (TIM; timprogram.ca) is an international master's level program at Carleton University in Ottawa, Canada. It leads to a Master of Applied Science (M.A.Sc.) degree, a Master of Engineering (M.Eng.) degree, or a Master of Entrepreneurship (M.Ent.) degree. The objective of this program is to train aspiring entrepreneurs on creating wealth at the early stages of company or opportunity lifecycles.

- *The TIM Review is published in association with and receives partial funding from the TIM program.*