A Survey on Context Awareness in Big Data Analytics for Business Applications

Loan Thi Ngoc Dinh, Gour Karmakar, Joarder Kamruzzaman
{l.dinh, gour.karmakar, joarder.kamruzzaman}@federation.edu.au
Federation University Australia, Australia

Received: 07 Jan 2018 / Revised: 28 Oct 2019 / Accepted: 11 February 2020

Abstract. The concept of context awareness has been in existence since the 1990s. Though initially applied exclusively in computer science, over time it has increasingly been adopted by many different application domains such as business, health and military. Contexts change continuously because of objective reasons, such as economic situation, political matter and social issues. The adoption of big data analytics by businesses is facilitating such change at an even faster rate in much complicated ways. The potential benefits of embedding contextual information into an application are already evidenced by the improved outcomes of the existing context-aware methods in those applications. Since big data is growing very rapidly, context awareness in big data analytics has become more important and timely because of its proven efficiency in big data understanding and preparation, contributing to extracting the more and accurate value of big data. Many surveys have been published on context-based methods such as context modelling and reasoning, workflow adaptations, computational intelligence techniques and mobile ubiquitous systems. However, to our knowledge, no survey of context-aware methods on big data analytics for business applications supported by enterprise level software has been published to date. To bridge this research gap, in this paper first, we present a definition of context, its modelling and evaluation techniques, and highlight the importance of contextual information for big data analytics. Second, the works in three key business application areas that are context-aware and/or exploit big data analytics have been thoroughly reviewed. Finally, the paper concludes by highlighting a number of contemporary research challenges, including issues concerning modelling, managing and applying business contexts to big data analytics.

Keywords: big data, context awareness, business applications, enterprise level systems

1 Introduction

Research on the integration of context into applications has been in existence for more than two decades. There exist a number of applications that use context in diverse disciplines including computer science and business to improve business activities or system performance. This underpins the importance of integration of context into business systems. Without context, the interpretation of a system, an event and even information may become meaningless. This is particularly true in data mining, especially in this information era, where without context, conclusions drawn from big data may be flawed [1].

Some techniques have been reported that consider contextual information in diverse areas (e.g., the context of a query and its users, and the query-driven context-aware recommendations in [2], [3]), and a number of papers have reviewed context-embedded methods in computer science, such as [4], [5], [6]. Some notable surveys on context-based methods and their brief discussion are presented in Table 1. In contrast, Table 2 presents a list of recent surveys on big data analytics applied to business applications, which shows none of them have included context-based methods for enterprise level application software. However, since context can play a significant role in data validation and meaningful information extraction, it has been identified as a major vehicle for big data analytics [7]. The tables and our analysis of the literature reveal that there is another aspect in this field that has not been reviewed to date: context-based methods supporting enterprise applications. Given the recent attention to big data analytics in business applications, this review will provide researchers in this field with an overview of the recent research in this area.

The aim of this paper is to sketch a picture of context-based methods for big data analytics in the business domain. Unlike reviews or surveys, which present comparisons among methodological and technical aspects of context-based methods, this paper provides an overview of what researchers have done, by considering context in big data analytics for applications supporting enterprise systems.

The structure of the paper is as follows: Section 2 summarises big data and the need of big data analytics in business organisations. The definitions of context, context models and context evaluation techniques for business applications are illustrated in Section 3. The main focus of this survey is highlighted in Section 4. This section presents three major enterprise level application software that exploit big data analytics and/or contextual information. The future research challenges in this field are summarised in Section 5 and the paper is concluded in Section 6.

Table 1. Selected recent surveys of context-based methods

Title	Year	Description/Content
A survey of context modelling and reasoning techniques [8]	2013	The requirements that context modelling and reasoning techniques should meet are discussed in this survey. These include the modelling of a variety of context information types and their relationships, high-level context abstractions describing real-world situations using context information facts, of histories of context information, and of uncertainty of context information.
A survey of context-aware workflow adaptations [9]	2008	Various existing approaches to adaptation in context-aware workflow are discussed.
A survey on context-awareness [10]	2011	The focus is context acquisition and sensing, context modelling and representation, context filtering and fusion, context storage and retrieval in context-aware computing. The development of context awareness and its applications is highlighted.

A survey of context data distribution for mobile ubiquitous systems [11]	2013	A unified architectural model and a new taxonomy for context data distribution, considering and comparing a large number of solutions, are presented.
A survey of context-aware recommender systems based on computational intelligence techniques [12]	2015	Surveys state-of-the-art context-aware recommender systems based on the computational intelligence techniques.

Table 2. Selected surveys of business applications in big data analytics

Title	Year	Description/Content	Inclusion of context-based methods
Big data and management [13]	2014	This survey explores the potentials and opportunities for new theories and practices that big data might bring. It also presents several conceptual foundations, as well as possible avenues for future research and applications in business management.	No
Big data: A survey [14]	2014	A review of big data, related technologies and several representative applications of big data, including enterprise management, online social networks, medical applications, collective intelligence and smart grids.	No
Big data and its applications: A review [15]	2015	Several techniques involving big data which can be applied to various fields of engineering, industry and medical science are reviewed.	No
Challenges in big data application: A review [16]	2015	Big data management techniques and their challenges for business applications are discussed.	No
The core enabling technologies of big data analytics and context-aware computing for smart sustainable cities: a review and synthesis [17]	2017	An overview of big data analytics and context-aware methods applied in smart city applications is presented.	Yes. But it focuses on the specific case of smart cities only, thus limiting its scope in business applications. Also it presents no discussion on enterprise level systems embedding big data and context.

2 Big data and big data analytics in business organisation systems

2.1 Big data

Big data encompasses a wide variety of data types, including text, image, audio and video, that are continuously generated from diverse sources such as the web, social media, mobile apps, sensor devices, networks, and data storage, as well as the internal data generated by the organisation itself. Due to the explosive growth in the volume of global data in recent times, many enterprises, not only large but also medium- and small-sized, are aware of the importance of extracting key information from big data. This can be seen not only in the way businesses are investing money in buying new technologies to extract precious information from big data, but also in their efforts to accelerate big data research and applications [4], [14], [18].

Big data is divided into four main categories: (i) structured (ii) unstructured (iii) semi-structured and (iv) mixed. According to Marcos et al. [18], a major proportion of data produced today is either unstructured or semi-structured. Traditional database management systems deal with structured data and thus cannot manipulate unstructured and other data types. Therefore, in parallel with database management systems, a number of big data analysis methods are being developed to deal with unstructured or semi-structured data.

There are seven main high-level dimensions of big data [19]: (i) variety – data types, (ii) velocity – data production and processing speed, (iii) volume – data size, (iv) veracity – data reliability and trust, (v) validity – data correctness and accuracy with respect to the intended usage, (vi) volatility – big data is volatile and destroyed when retention policies or warranty periods expire, and (vii) value – worth derived from exploiting big data. This is also known as 7V big data model. The two main goals of the development of effective methods for high-dimensional data analysis are to (1) predict accurately future observations in order to make decisions and (2) gain insight into the relationships among various features of the data and their impacts for many purposes in science as well as business.

However, the inherent complexity of big data derived from many sources poses a major challenge for analytics; for example, high dimensionality brings noise accumulation, spurious correlations and incidental homogeneity. In addition, high dimensionality combined with large sample size creates issues such as high computational cost and algorithmic instability. The massive samples in big data are typically aggregated from multiple sources at different time points using different technologies [18], [20], further contributing to the complexity.

As mentioned above, data come from many sources and are hidden in many different forms [14]. This leads to business organisations being increasingly dependent on big data analytics. Hence, the development of new techniques that improve the effectiveness and efficiency of big data analytics is extremely important, especially for business areas where information technology directly/indirectly affects businesses' activities and their income [14], [21], [22], [23]. The next section presents the impact of big data analytics on business organisations.

2.2 Big data analytics for business organisations

Recent studies show that the benefits of big data technology use in organisations are many-fold, including the production of more accurate results and cost savings [21], [24], [25], [26]. For example, an enterprise organisation can produce 61% more accurate results and save 56% of the cost. Other substantial benefits are the retention and analysis of more data, increasing data analysis speed and the reduction of manual processing [27]. The collection of valuable and relevant data and capturing their deep insights are not only a significant challenge for business organisations but also a major issue for the relevant research community.

According to another worldwide survey done by the Tech Republic Company [28], most businesses consider big data analytics as being mainly for revenue, customers and markets. Based on big data, companies can innovate and enhance their current products and services as well as develop and launch their new products and services [21], [24], [13], [29]. This adds substantial value. Moreover, many advantages and benefits can be derived by applying big data analytics in different areas, such as customer intelligence, supply chain intelligence, performance, quality and risk management and fraud detection. In addition, industries such as manufacturing, retail (refer to big data levers in the retail industry in Table 3), central government, healthcare, telecommunications and banking can also gain direct/indirect benefits from big data analytics [30].

Table 3. Big data levers in the retail industry (www.mckinsey.com)

Function	Big Data Levers
Marketing	Cross-selling. Location-based marketing. In-store behaviour analysis. Customer micro-segmentation. Sentiment analysis. Enhancing the multi-channel consumer experience.
Merchandising	Assortment optimisation. Pricing optimisation. Placement and design optimisation.
Operations	Performance transparency. Labour input optimisation.
Supply chain	Inventory management. Distribution and logistics optimisation. Informing supplier negotiations.
New business models	Price comparison services. Web-based markets.

Over the years, the process of making managerial decisions has become of paramount importance, and it has been a challenging research topic. In the view of most decision makers, the value of big data depends on the ability to deliver key information and knowledge to make better decisions. Better decisions can be made upon the subsequent knowledge derived from historical and real-time data generated through means like supply chains, customer behaviours and production processes [13], [21], [26].

According to Elgendy and Elragal [30], business organisations not only focus on analysing internal data (e.g., sales, customer feedback via phone or email, shipments and inventory) but they also need to consider data from outside sources such as social media and forums. Since big data are generated from a variety of sources with different representations and formats, the question for any researcher or business is how important data relevant to a business context can be captured and analysed more accurately to represent deep and relevant business insights [13], [21], [24], [25], [31]. In literature, a model has been proposed for data mining process which is presented later in relevant to the importance of context for big data analytics used in business applications.

The capture of data and their values relevant to a business system demands the integration of the context in big data analytics. This is because of the utilisation of context in other application domains, such as information retrieval [2], [3], [12], [32], business process and workflow [9], [33], [34], and reasoning scenarios [35], has delivered significant benefits. The following section presents the evolution of the definition of context and context models over time.

3 Context definition, models and evaluation for business applications

In this section, context definition, context models and evaluation for business applications are presented. The progress in literature in each topic is discussed in the chronological order as shown in Figure 1.

3.1 Context definition

A number of recent research studies have shown that the use of context in interactive applications is very useful and important [36], [37], [38], [39], [40], [41]. By considering context in computer applications, it has been demonstrated that the number of irrelevant results is reduced and more accurate data are collected. For example, the application of a query context and user context in information retrieval improved retrieval precision substantially [3], [42]. In addition, the term “context” has been embedded in information technology, business and other application domains for a long time, and has improved the efficiency and effectiveness of these applications [3], [43].

Because of the richness of everyday language and the knowledge that people acquire during their life-time, humans find it easy to communicate and understand each other and they also react appropriately, based on the context surrounding them. However, a machine like a computer does not have the ability to convey ideas, transfer information well to humans or interact with them [44]. Therefore, the richness of communication in human-computer interaction will be increased by improving the computer's access to context, and as a result it is possible to produce more useful computational services [8].

To fulfil the purpose of using context successfully, people must understand what context needs to be integrated, how it can be used, and when context should be considered [8], [9], [44]. In [6], Abowd et al. he authors emphasise that an understanding of context would enable application designers to choose what context needs to be used in their applications, and provide them with an understanding of how context could be used. This will help application designers determine what context-aware behaviours support their applications.

There have been many definitions of contexts put forward in the literature as shown in Figure 1. However, the definition that has become the most popular is proposed by Dey as "Context is any information that can be used to characterise the situation of an entity" [44]. An entity can be a person and place or any object which is considered relevant to the interaction between a user and an application. Later, the definition was further updated by Coutaz et al. [45] in 2005 as "… not simply the state of a predefined environment with a fixed set of interaction resources. It is part of a process of interacting with an ever-changing environment composed of reconfigurable, migratory, distributed, and multi-scale resource".

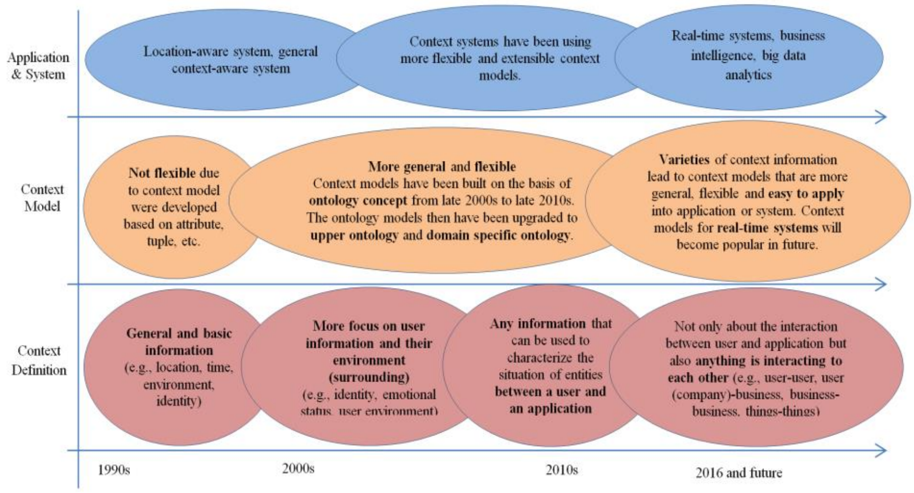


Figure 1. The development of context definitions and their applications

The rapid technological development and the advent of different data sources including big data have produced a large number of diverse applications and systems at an increasingly higher rate. For this reason, with the evolution of the context definition, context model and their application and system are being developed over the time (see Fig. 1). For example, the figure shows that context initially meant simply a location. However, over time it has been extended to represent any information about the situation of an entity. In the 1990s context-aware application and systems were limited to location-aware systems, but later in the 2010s they were extended to any system which uses a more flexible and extendable context model. Similarly, Figure 1 shows original context models in the early 1990s were less flexible, as they considered only the attributes or value tuples of location. However, over time, models have become more flexible and extensive since 2010 [32]. The next section presents some contemporary and widely-used models of context.

3.1.1. Contextual information for big data understanding and preparation

Figure 2 illustrates the six phases of the Cross-industry Standard Process for Data Mining (CRISP-DM), a data mining process model for describing approaches that data mining experts use to tackle problems. As per the model, ‘business understanding’ is the starting phase of a data mining process. Business understanding represents the overall business objectives and the strategies to achieve those objectives. Since the context of a business can represent its situational and physical environmental condition, Figure 2 indicates the importance of the consideration of the business context in a big data analytic process, as mentioned above.

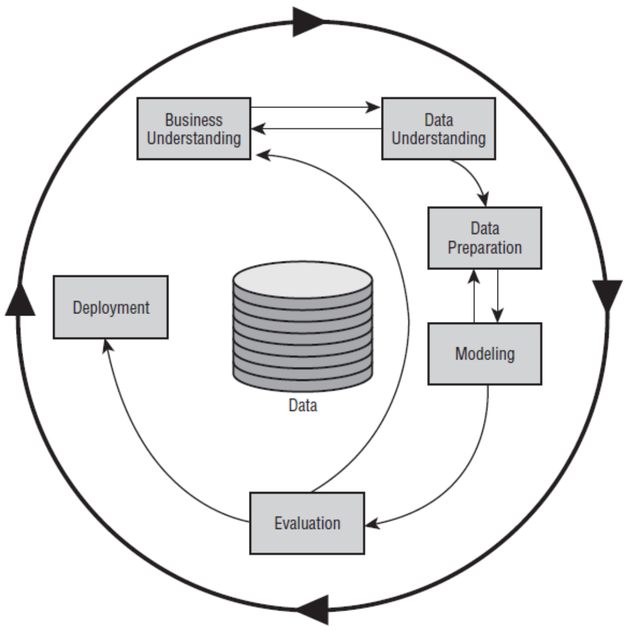


Figure 2. Phases of the CRISP-DM reference model [46]

The early phases of the CRISP-DM model, namely data understanding and data preparation, must consider data characteristics for mining purposes, and therefore these phases, when embedded within business contexts, have direct impacts on big data analytics. Context-aware methods have been found to be more effective when contextual information is considered in analytics. Table 4 describes the data understanding and preparation stages of CRISP-DM, considering the 7V big data model and contextual information.

Table 4. Big data models for data understanding and preparation of phases of crisp-dm in relation to context awareness

	Phases of the CRISP-DM reference model	
	Business/Data Understanding	**Data Preparation**
Variety	Since the context of a business organisation is developed considering its general business strategies, context can be applied in gathering different types of data from various sources that are essential to pursue the business strategies.	The amount of data is reduced by considering context to remove irrelevant data during data collection. For example, media data (e.g., text, audio and video) is transformed into another form of key-value pairs that is suitable to be handled by big data modelling tools (e.g., Riak, Oracle NoSQL, Cassandra, DyanmoDB, MongoDB, and Azure Table Storage).
Velocity/Volum	An up-to-date business context can be used to fine tune the sampling policy aimed at tackling high velocity and huge volume of data.	The amount of data is selected based on the relevance themselves to the context or the need of the organisation (e.g., business goals). This results in a significant reduction of volume and increase in relevancy in data collection [38], [47].
Veracity	Quality data is a pre-condition for acceptable analytical outcomes. In contrast, poor data quality because of problems of accessing reliable and trustworthy data from existing sources poses substantial hindrance to the improvement of business strategies.	Contextual information can be utilised to identify the inaccurate and missing data in the data cleansing task in this phase. One study [47] suggests this eventually improves the reliability and trustworthiness of big data.
Validity	As context reflects the data requirements for a business plan, it is highly likely that context focused data collection will be more relevant to what a business wants to achieve and more useful for its intended usage.	Data validation, one of the key recommendations for data construction, using a business context increases the likelihood of meeting all application-specific criteria for the intended use of the data [47].
Volatility	To indicate the duration for which data remain valid, the data expiry period can be added to the context description, and then it can be used to avoid collection of expired data.	A data expiry period, one of the pieces of contextual information, can be used to clean outdated data automatically from big data storage.
Value	By considering contextual information, the context-aware methods are capable of collecting valuable information towards fulfilling an organisation’s business objectives.	Similar to data understanding, since contextual information can be applied in solving a number of matters of other Vs in data preparation, it will increase the value of the data considerably.

3.2 Context models

Context definitions include only the abstract level of information for representational purposes. However, to embed context in a particular system, the context has to be designed (modelled), taking into account all its essential structural fine details so that it can meet the purpose of that system. Therefore, a context model is a set of information coming from external and/or internal systems that can have some effect on its relevant system [8]. Context models that have been applied in the business domain can be classified into three main groups: i) the n-gram model, ii) the tree structure model and iii) the onion model.

N-gram model: An n-gram context model is mainly used to present a user context and query context in information and document retrieval [3], [48]. Both user context and query context have been represented using language modelling in terms of n-grams where n represents the number of interested units/words. The query context represents the user query and considers both linguistic and semantic knowledge represented by n-grams. The user context indicates the user's domain/topic of interest and is constructed by an algorithm which considers a statistical language model built from a priori set of documents relevant to the topic of interest.

Tree-structure model: The most popular tree structure context models are: i) user context model, ii) B2B (Business-to-Business) context model, iii) context tree model for a business processes and iv) business context model for big data collection and processing.

User context model: In [3], a user context model for information retrieval system assimilates to all the factors that describe a user's intentions and perceptions about the surrounding factors, as shown in A1 in Appendix. In this model, the user context is divided into five main sub-contexts – i) task, ii) social, iii) personal, iv) spacio-temporal and v) environment. The task context includes goals and task information, while the social context is information on social contacts such as friends and colleagues. Personal context is represented in terms of personal profile, physiological and mental contexts. The social-temporal context refers to location, direction and time, while the environmental context comprises light, services and people.

B2B context model: As an outcome of a number of research studies on B2B collaboration, the B2B context model shown in Figure 3 has been proposed to support supply chain applications [36], [49], [50]. In this model, information on the user or company, temporal and location is used. Since this model was developed particularly for B2B collaboration, the main focus was not on the business system processes. For this reason, it is not effective for business process modelling, and this paves the way for the development of a context model considering the main aspects of a business process called "context tree model for a business processes".

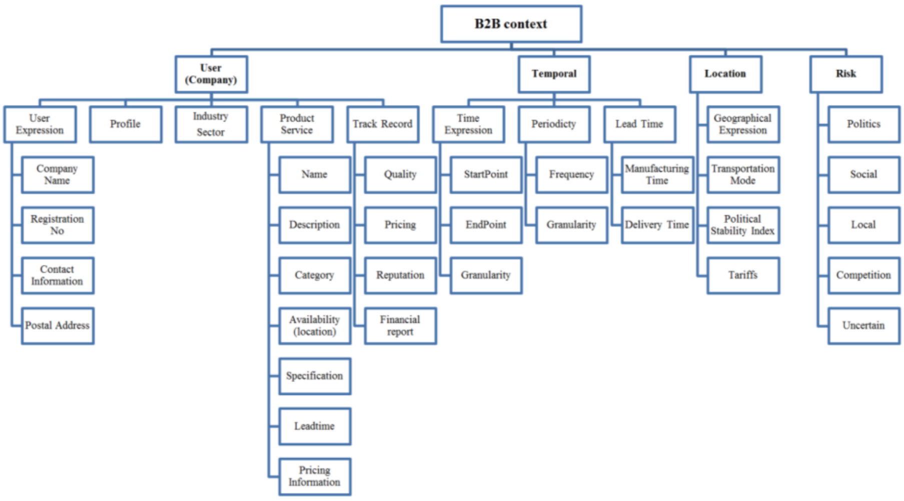


Figure 3. Part of B2B context model

Context tree model for a business processes: A tree structure context model based on the basic taxonomy of a business context which captures most common context- related knowledge has been developed by Saidani et al. [51]. This model consists of essential aspects of a business process, including resources and organisation unit. For example, resources include business objects and actors and organisational units (e.g., quality of communication, quality of relationships, actor's proximity) that belong to organisations in the context tree.

Business context model for big data collection and processing: Recently, the present authors introduced a business context model reflecting various aspects and activities in an organisation, such as the physical environment and product aspects of a business, suitable for data collection and processing in a big data scenario [38]. The contextual information of this model refers to the background nature of a business and its situational values (the values that change over time). Such a representative context of a grocery shop is based on the products it sells and the situational information of each product, as seen in A2 in Appendix.

The situational information of each product includes features such as shelf-life, retail price, season, discount, wholesale price, product condition, display frequency and country of origin in the fruit and vegetable sections, which are likely to change over time. Such information plays a significant role in determining the business strategies, and hence they need to be considered to represent the context of a business system.

The onion context model: Rapid economic and technological progress makes the expectations of customers ever-changing and diverse, resulting in changes in the context in which expectations are formulated. Therefore, there is a need to consider context- related knowledge in the behaviour of a business process [51]. As customers are the key output value, a context model for a business process must consider the business organisation and human beings, such as user context, environmental context, and the internal context, including organisational and goal contexts [52]. Considering the contextual information, a model has been proposed [52] which is called the onion context model as shown in Figure 4.

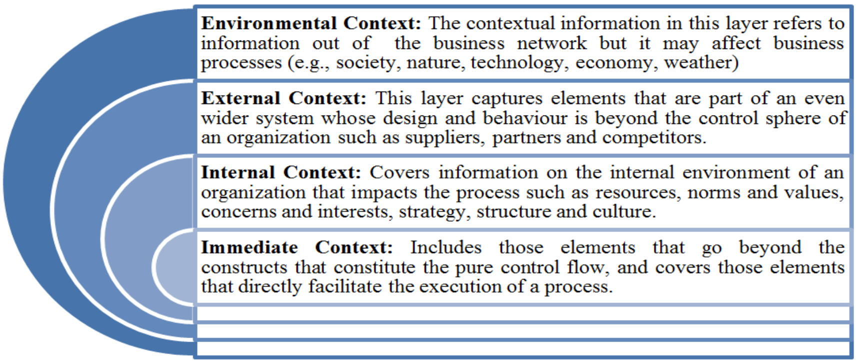


Figure 4. Onion context model

In the onion model, the innermost context ('immediate context' in Figure 4), is regarded as the core context and the contextual information from an outer layer is experienced by the immediate inner layer. There is no such core component and inter-layer experience in n-gram and tree structure context models with the exception of tree structure model where a child node inherits the information of its parent node.

Since context models are crucial for the understanding of the review topics, the comparison among the different context models is presented in Table 5.

Table 5: Comparison among the different context models

Context Model	Contextual Information	Targeted Application	Type	Purpose
Keyword context model [48]	Instances of keywords	Information and document retrieval	n-gram	For finding a best match sentence or document based on keywords.
User context model	User information	Information retrieval	Tree	For managing customer and keeping in track with them.
B2B context model	Supplier and their product information	Supply chain management	Tree	For evaluating and selecting prospective suppliers.
Context tree model for a business processes	Context related knowledge which adapts according to BPs' behaviours and stakeholders' requirements.	Role-driven BP modelling approach	Tree	To make BPs self-operating to minimise the requirement of an administrator assistance.

Business context model for big data collection and processing	Product information	Big data collection and processing for analysing customer feedback of a grocery shop	Tree	For understanding customer opinions for each product sold in shop
Onion context model	Description of regulatory systems and their surrounding environment.	Any type of application that includes business processes such as business process management, enterprise resource planning, and customer relationship management.	Onion	For preparing alternative plans to cope with changes which may come within regulatory system or from external environment of the system.

Since many different types of contextual models exist in the literature, this raises an important question: what types of contextual information are essential to context-based applications in the business domain and how can we ensure that they are appropriate for a business? The former requires the context modelling techniques, while the latter necessitates context evaluation. The next section addresses this issue.

3.3 Evaluation of contextual information

Generally, context evaluation can be performed in two ways – evaluation of (i) a context model and (ii) contextualised business process (i.e., a business process which embeds context into its bedrock). They are described in the following sections.

3.3.1 Evaluation of a context model

Thus far, we have presented various types of context model structures used in many different applications in Section 3.2, but how they were developed has not been discussed. Before evaluating and concluding the efficacy of context models for business applications, it is important to consider how a context model is usually developed. In developing context-based methods, the first and most important stage is to model contextual information. To do this, the basic and the most popular way is to build a model based on observing and examining the attributes of a business system and its users and physical environments. For example, what types of contextual information a business system needs, what contextual information including environmental attributes may influence its users, what information a user may need at the time they work on the system, and what information a business system needs to be aware of when it is going through some changes. Is there any relationship among these pieces of information?

In [53], there are five main sequential and cyclic steps presented in context modelling: (i) learning what information a system needs, based on observing the activities of the system and the behaviours of users, (ii) making a prediction of what contextual information is potentially required, (iii) deciding on what information needs to be considered in the context model, (iv) indicating the relationship among the types/pieces of contextual information based on practice and experience, (v) building the relationship among contextual information, users and systems. Although these steps have been mainly applied to context modelling in information retrieval, they have also been used in other context-based areas, such as business process, business intelligence and real-time analytics. By following these steps, context models can be developed more accurately [53]. Although a context model can be constructed following the above steps, an important question remains: how can this context model be evaluated appropriately in business applications? To address this issue, the model of a context can be divided into a number of fundamental contextual components. For example, the user context shown in Appendix A1 has been divided into five contextual components: (i) user expression, (ii) profile, (iii) industry sector, (iv) product service and (v) track record. For each of these components, similar to the unit or module test applying the white box testing technique in software engineering, some suitable test cases can be derived considering all possible, important and relevant contextual scenarios to ensure its validity for all aspects and assumptions, effectiveness and appropriate adaptation in a dynamic business process. Similarly, after testing and modifying all individual components, they can be combined. Some test cases can be derived for integration testing to verify whether the components work together. The generic nature of a context model can be tested by applying it in different types of applications. Therefore, to test whether a context model is generic, we need to test whether it is applicable in other types of applications with minor modification or extension [54]. Another way to evaluate a context model is to assess a context model and its components by applying the contextual analysis technique introduced in [33] for a particular contextual scenario. This is detailed in the following section.

3.3.2 Evaluation of contextualised business processes

In a business process, contexts are analysed and evaluated based on determining the context variants via truths (e.g., the facts that shoppers asked about a promotional campaign, that a public holiday was in the region of the store) and statements (e.g., a shopper showed interest in a promotion or lack of it for being in a hurry). A process contributor (e.g., process designer, marketer, or salesman) is able to verify a truth. However, they are unable to verify a statement. The use of context analysis enables reasoning about the context in the business process as well as the discovery of other contextual information that may relate to the current context.

Contextualisation of a business process via context analysis can guarantee to output a better business process model. In addition, the incorporation of context analysis into a business process can validate that a contextualised business process fits its context and is effective [33] [55].

4 Major business application areas exploiting big data and contextual information

The survey of main business applications that use big data and the business ecosystem context has been conducted using a systematic literature review approach. The application of SLR in this survey is described in the following section.

4. 1 Review methodology

Recently, the systematic literature review (SLR) is being used in conducting the literature survey. This is because SLR provides a logical and systematic way to identify, evaluate and interpret existing research approaches associated with a particular research question, topic or issue. SLR reduces the likelihood of having bias and is regarded as secondary research as it forms the basis of primary research [56]. For this reason, in this paper, we chose to use SLR to present the state-of-the-art of major business application areas that use big data and context, which is the prime aim of this survey. This section presents the importance, implications and types of big data, including its 7Vs and contextual information for those applications. The requirement of context to leverage big data for a business application is also presented in this section.

To meet the above objective, the following research questions have been formulated:

RQ1: Which major business application areas have exploited big data, context and both?

RQ2: What are the implications of big data and contextual information on the selected applications?

RQ3: Why is contextual information essential to extract the value from big data?

For the research based on the questions mentioned above, the selection of search keywords, databases, and the inclusion and exclusion criteria are provided below:

Search keywords: The terms that are related to the targeted business applications and big data analytics and contextual information were used for searching the relevant information. The final search strings were: “business applications”, “enterprise systems”, “enterprise resource planning”, “ERP”, “customer relationship management”, “CRM”, “supply chain management”, “SCM”, “big data”, “big data analytics”, “context” and “contextual information”.

We also combined the associated search strings using OR operator such as "ERP OR big data", "ERP OR context", etc.

Databases: Since business applications are not only articulated in research articles, but they are also the main focus of discussion from a large number of commercial products. Therefore, the data sources for searching relevant topics include academic research databases such as Google Scholar, IEEE Xplore, Springer and Science Direct, and the websites/forums (nessi-europe.eu, dellemc.com, www.sap.com and intuital.com) for commercial products.

Inclusion and exclusion criteria: The search scope includes (1) the data sources in English, (2) papers and articles containing the keywords in their content, (3) peer-reviewed articles from academic sources including conference proceedings, journal papers and book chapters, and (4) articles and discussions from tech companies and magazines.

The search criteria exclude (1) non-English data sources, (2) the papers and articles mentioning the search keywords only in their references, (3) data sources published before 2015 (except outstanding articles with the significant number of citations).

Search results: The search results from Google Scholar returned about 11,300 articles that were relevant to the keywords. However, it reduced to approximately 7,880 articles after the time criterion “since 2015” was chosen. As per our aim (refer to the details provided in Section 4.2), the business application areas were tightened to three main areas: i) Enterprise resource planning (ERP), ii) Customer relationship management (CRM), and iii) Supply chain management (SCM). Moreover, the articles in these areas that mention big data analytics and contextual information were selected. The final number of primary articles including in the main section of this review was 23.

The selected primary articles were divided into three groups representing the three particular business application areas mentioned above. According to the search results, there are a considerable number of SCM approaches that use big data and contextual information. The results show the number of articles in SCM is higher than that of ERP and CRM.

4. 2 Review of major business applications that use big data and contextual information

The advancement of ICT innovation and the increased recognition of the potential of big data are progressively motivating business organisations to use more and more business applications exploiting big data. These applications, primarily in the form of enterprise level software, perform core business functions and hence enable business organisations to adopt cost-effective and innovative services, and increase productivity.

The most widely used enterprise level software that performs core business functions are: (i) ERP, (ii) CRM, and (iii) SCM. With respect to their central focus on business entities and running core functionalities, these three systems are unique. Besides that, an ERP system also includes several essential functions of CRM and SCM called CRM and SCM components. However, a medium to large enterprise system requires independent CRM and SCM systems with a full set of core functions. For example, as shown in Figure 5, the main focus of ERP, CRM, and SCM is on business entities like employees, customers, and suppliers, manufacturers and other B2B collaborators, respectively and their core functionalities.

Because of the high cost, duration and complexity of deploying these systems, usually medium and large business organisations use all three software. Since CRM boosts up sales, and its complexity and cost are comparatively less, small companies also use it. However, with the increased adoption of cloud computing and innovative ICT approaches like Internet of Things (IoT), the distinction between these three enterprise level software is disappearing [57].

Changes in business contexts associated with business organisations and entities, technology and environment also lead to unpredictable events. These events generate a massive amount of different types of data, including big data. The enterprise systems need to capture and leverage the value of these data. This value improves cost-effectiveness and introduces innovation in services.

In this section, we review the state-of-the-art business applications provided by ERP, CRM and SCM that have exploited big data and/or contextual information.

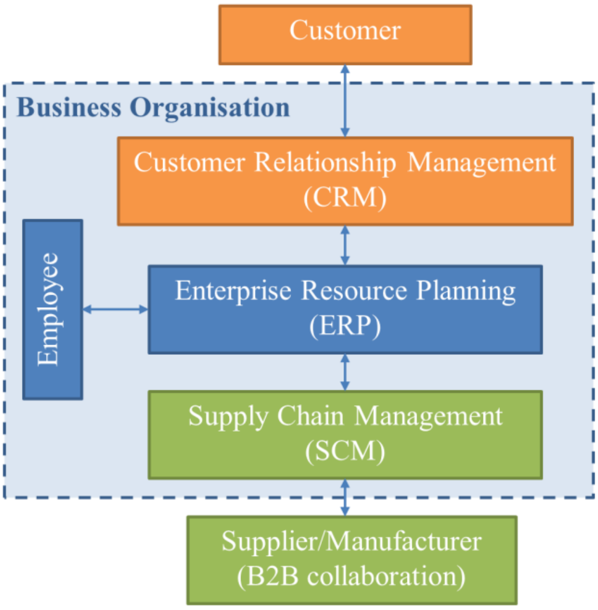


Figure 5. Enterprise level software systems showing their main entities and interactions

4.2.1 Enterprise resource planning (ERP)

An ERP system integrates and streamlines six core functions/features: i) CRM components, ii) SCM components, iii) Business intelligence, iv) Human resource management, v) Inventory management, and vi) Accounting management. These functions/features assist a business organisation in achieving cost-effective operations and providing innovative services.

In general, contextual information is mainly considered during the implementation stage of ERP systems ([58], [59], [60]). The effect of contextual factors including external contextual variables on the implementation and use of ERP systems has been analysed in [59] and [60]. These studies provided models of how to implement ERP systems considering contextual factors that can influence the benefits and satisfaction of users of ERP systems. For example, training programs and staff behaviours towards new ERP systems can be regarded as contextual factors. Exploration and management of these contextual factors help business organisations to achieve the goal of successful ERP implementation.

Table 6 shows the types of contextual information that have been used in ERP systems. Contextual information is mainly focused on the factors or variables that may affect the success of ERP adoption such as individual, organisational, technological and environmental influences. Individual context includes the information regarding human factors such as user behaviours, literacy and age group. Organisational context consists of the information that reflects the situation of project preparation for business organisations such as training programs, job assignment and scheduled deployment. The technological context shows the computational infrastructures that business organisations use and the budget they are willing to invest. Technological contexts have been highly recommended for cloud systems. Environmental context can be regarded as changes during ERP adoption such as over budget, economic crisis and bankruptcy.

A few studies have applied contextual information and big data analytics in business intelligence [61] and [62] and accounting management [63] components of an ERP system. Since business intelligence needs to leverage the value of different data types from various sources, it is expected that both big data and contextual information use will increase sharply over time. In contrast, an ERP's human resource management component has many approaches that utilise context and big data with a popular, but essential tool, namely talent analytics [64] [65] [66].

Big data solutions for ERP have attracted the relevant research community since the late 2000s. The review in [63] has emphasised that connecting traditional data with new unstructured data sources delivers powerful evidence for economic activity, especially accounting function. In [61] and [62], a new term of ERP called "big-data ERP" is presented. The big-data ERP framework processes not only transactional data coming from business operations but also the data that are collected from IoT networks and social media. This ERP framework helps the ERP systems acquire more data, including contextual information to process primarily for the function of business intelligence.

For collecting more data for an ERP system, several external devices such as RFID (“radio-frequency identification” tags or smart labels are encoded and attached on products) and GPS (global positioning system) have been attached with products, and relevant services have been developed to collect real-time information using those devices. With RFID being embedded with everyday products, such services will have to deal with a huge amount of data in the future. Khazeali et al. [67] discussed a solution for transforming traditional ERP to cloud-based ERP and highlighted the benefit of cloud-based ERP, especially for SMEs (small-medium enterprises).

Table 6. Contextual information used in key enterprise software

	ERP	CRM	SCM
Contextual	• Individual context • Organisational context • Technological context • Environmental context	• The physical behaviour of customers • Customer location • Traffic status	• Procurement: Supplier information and product details. • Manufacturing, logistics and transport, and demand management. Information includes consumer demand, product trend, traffic information, weather and disaster information

Apart from academic research, many ERP vendors have already developed their commercial ERP products to deal with big data. For example, SAP, one of the biggest ERP providers, has introduced a generation of ERP where big data analytics along with mobility services and IoT applications have been applied since 2006 (www.sap.com). Another well-known ERP vendor – Oracle, has also developed special tools for data integration from various sources (www.oracle.com). To collect and process big data, these tools need to integrate existing big data analytics platforms and tools such as Apache Spark, Hadoop and SAP HANA. Besides these, recently, most ERP vendors have focused on cloud-based ERP to reduce the cost of computing resources for their clients.

Table 7. Big data dimensions in relation to three key enterprise software

Big data dimension	Key enterprise software		
	ERP	**CRM**	**SCM**
Variety	- Internal data: structured data (e.g., financial, employee data), unstructured data (e.g., email). - Unstructured external data (e.g., social media and customer tracking information)	- Similar to ERP but more focused on the customer, for example, internal data like customer feedback and inquiries. External data include potential customer and customer trends.	- Data for SCM are mostly relevant to product information (e.g., product inventory, delivery information, innovation and development), product movement information such as product tracking information, transport and traffic situation, and weather forecast.
Velocity	- Speed is required mostly for cloud-based systems and tracking information.	- Can be very high as customer data are generated from various	- Very high as data are generated at high speed from external sources to support

		sources including online platforms.	on-time decision making, especially in manufacturing, and logistics and transport.
Volume	- Cloud-based ERP systems help business organisations to deal with data volume.	- Large organisations like Apple, IBM, Ford, and Walmart have a large volume of customer-related data. - A cloud-based CRM system is also a popular option for medium and large business organisations to deal with a huge volume of data.	- The amount of data processed in SCM systems is larger compared with the data processed in ERP and CRM systems as SCM systems cover a wide range of data starting from customer demand, product management, delivery to weather. - Cloud-based SCM systems are also a current trend to store and/or process huge volumes of data to reduce ICT costs.
Veracity	- ERP systems mostly focus on integrating and processing internal data; these data are usually filtered and stored based on the needs and data policy of business organisations. - Therefore, the trust of these data is very high.	- A large part of customer feedback is collected from online systems such as social media. Therefore, trust, correctness and worth of these data are crucial.	- Similar to CRM but the veracity in SCM needs more attention than CRM as data are relevant to not only the customer but also product trends and supply status.
Validity	- High validity.	- Filtering and capturing valid data are more challenging than in CRM.	- Capturing deep insight of data requires further research in SCM systems.
Volatility	- Depends on the nature of a business and its policy.	- Usually more volatile than ERP.	- Same as CRM but more volatile than ERP.
Value	- High value.	- Discovering value from CRM data is more challenging.	- Similar to ERP and CRM, SCM data are also valuable.

The description of the seven dimensions of big data and the type of contextual information has been presented in Sections 2 and 3, respectively. An overview of how seven dimensions of big data are considered in these enterprise level software systems is presented in Table 7. Similarly, Table 6 illustrates the types of contextual information considered in each of them.

Table 7 shows the research and commercial products for ERP systems have covered six dimensions of big data except “volatility”. This is because it depends on the nature of a business and its data retention policy and warranty period. As data processed in ERP systems are mainly produced within a business organisation, “variety” of big data in ERP systems includes structured internal data such as financial and employee data, and unstructured internal data like email and report. ERP systems can also process unstructured external data including customer information from social media and customer tracking information. However, processing unstructured external data is not the main objective of ERP systems. Those data are mainly analysed in CRM and SCM systems discussed later.

The speed of data processing, i.e., “Velocity” is mostly considered in cloud-ERP systems. However, none of the existing approaches for ERP systems have directly emphasised “veracity”, “validity” and “value”. As ERP systems focus on integrating and processing internal data, these data are usually structured, filtered and stored based on the need and data policy of business organisations. Therefore, trust, correctness and worth of these data should be very high.

For each of the seven dimensions, the significance of contextual information in terms of data understanding and preparation is presented in Table 4 in Section 3.1.1. Daneshgar [58] showed that without context it is difficult to extract deep insight or value of big data. The use of contextual information along with or without big data has also been proven to create a big impact on ERP systems. In [58], an object-oriented framework is presented which facilitates the sharing of contextual knowledge/resources that exist within ERP processes. Babu et al. [61] emphasised that big data helps ERP to have a contextual view of the automated decision-making process. A framework called “Application integration framework for big data analytics in ERP” was introduced to convert unstructured contextual data from many sources into a semantic representation that is usable for predictive models.

4.2.2 Customer relationship management (CRM)

A CRM system consists of three main functions: i) sale automation, ii) lead management, and iii) customer service and support. A CRM system helps business organisations facilitate and preserve strong customer relationships. CRM systems allow business organisations to categorise the customers based on their expectations and create strategies to attract, retain and track potential customers with an aim to increase revenue.

Contextual information and big data have been taken into account in these systems since the mid-2000s. As one of the main purposes of CRM systems is to improve the relationship between business organisations and their customers, contextual information is mostly related to the customer profile, action and location as presented in Table 6. This includes physical behaviour and customer location which are mainly collected via IoT networks. Besides this, traffic status like traffic congestions and roadworks that might impact customer order delivery is also one of the contextual information that has been considered in CRM systems.

Up to our knowledge, contextual information and big data analytics have been considered mostly in the two CRM functions: (i) lead management [68] and (ii) customer service and support [69]. As the primary purpose of the lead management is how to seek potential customers and keep track of them, analysing contextual information of potential customers helps a business to capture customer psychology. Similarly, to improve the quality of customer service and support, it is essential to understand the contextual information of customers so a business can build a strong bond with them.

An approach for the integration of service-based adaptation in a planning framework is presented in [69] so that compositional adaptation to the context-aware applications can be done. A scenario of context-aware CRM by using the proposed approach is introduced. In this scenario, context-aware CRM was installed on a user's smart-phone. By automatically detecting the current location of customers and their surroundings via the mobility services in ubiquitous and service-oriented environments, context-aware CRM enables its users to keep on track with customers. In [70], a company named Aspedia has claimed that its product is the world's first commercial product for context-aware CRM named CoCRM. In this product, contextual information is applied to CRM system so that it enables to detect what a user needs based on the business environment.

To collect these types of contextual information, there is a need for applying big data analytics in CRM systems through big data tools and IoT networks. Park et al. [68] introduced a framework for CRM system named uCRM. Unlike traditional CRM, uCRM can receive real-time information from customers such as their physical behaviour and current location by utilising ubiquitous computing technologies. This is done by attaching RFID tags with their products. When customers collect products with RFID tags, they are tracked and the relevant customer data are sent to the context data warehouse via intermediary networks such as a wireless local area or mobile network. This evidences the application of both contextual information and big data analytics to improve the effectiveness of CRM systems.

In practice, context and big data have been extensively considered in customer service and support. This vast use is evidenced from the trend of successfully deploying virtual staff to communicate with customers (e.g., AskJess from Jetstar (jetstar.com), virtual teams from Dell (dell.com) and Basecamp (basecamp.com)). To do this, not only a massive template of question-answer is made but also a large amount of contextual information is applied. The processing of big data plays an essential role in this area. The questions that arise from customers are generally unstructured but have the contextual information surrounding them. These inevitable and simultaneous questions are regarded as a source of big data whose value is to be perceived appropriately to improve customer satisfaction levels. Therefore, high-performing virtual information needs to be captured well and processed effectively.

Additionally, because people currently tend to share their thoughts, emotions, and feeling on social media, capturing more data from social media to get more valuable information strongly demands the implementation of big data analytics.

The "Variety" of CRM data is higher than that of ERP data (Table 7). This is because unlike the manipulation of structured internal data (primarily) in ERP systems, the aim of CRM systems is to consolidate the relationship between business organisations and their customers. Therefore, collecting and processing customer information from available external domains are essential tasks for CRM systems.

Similarly, "velocity" and "volume" of CRM big data are higher than those of ERP. This is a result of the focused attention on collecting real-time external data such as customer tracking information and customer trends in CRM systems. As alluded before, cloud-based CRM systems are highly recommended to deal with data volume and computational infrastructures to achieve cost-effectiveness and add innovative services. As types of data for CRM systems are different compared to ERP systems, "veracity", "validity" and "value" mainly depend on external data sources. This necessitates trust, correctness and worth of these data to be specially treated. However, to the best of our knowledge, no study exists in the literature that specifically discusses issues associated with the variety and validity of CRM big data. Therefore, this demands more research to devise a better solution for filtering and capturing relevant and reliable CRM big data.

4.2.3 Supply chain management (SCM)

Unlike ERP and CRM, the utilisation of big data analytics and contextual information is more prominent in SCM. A number of studies and commercial products which include big data analytics and context awareness have been introduced. Especially, a remarkable amount of them was released from 2014 to 2017 [71], [72].

A SCM system comprises five main functions: i) procurement [73], [74], ii) manufacturing [75], [76], [77], iii) warehousing [78], [79], iv) logistics and transportation [80], [81], [82], and v) demand management [83], [84]. Exploration of context awareness and applying big data analytics are mostly concentrated on manufacturing, and logistics and transportation. Besides, several studies have also focused on the implementation and adoption of SCM systems [85] [86].

As SCM not only manages the optimal inventory level and transportation but also takes care of product quality, the data processed in SCM systems are mostly related to raw materials, products, suppliers and partners. These generate a large amount of external big data. Similar to CRM, SCM data are also a combination of structured and unstructured internal and external data. In Table 7, therefore, "variety" and "volume" in SCM is considerably high compared with those in ERP and CRM.

The procurement in SCM primarily selects suppliers based on the database of product details available in suppliers' information systems [73], [74]. In addition to this, the manufacturing, warehousing, logistics and transport, and demand management in SCM also use contextual information collected from external data sources (Table 6). The example of this is the information on market demand and demand management [83], [84], warehousing, logistics and transports, and weather change [80], [81]. A number of big data analytic techniques and tools such as sentiment mining, predictive analysis and machine learning [78], [82] has been proposed for better demand management. The "Velocity" directly affects decision making in SCM such as just-in-time and lean supply chain. Additionally, SCM's dependency on external data sources means it requires more rigorous research on "veracity", "validity" and "value" to improve the effectiveness of applying big data analytics to SCM for better decision making.

The two most important key parts of SCM are – (i) how to manage inventory optimisation for providing on-time delivery and (ii) how to detect and prevent low/out of stock products. RFID tags attached to each consuming items are one of the feasible solutions for the above issues. The most updated product inventory status received through these tags can be regarded as contextual information. Because of the high cost of RFID tags and readers, sensors for detecting light or weight of products installed in supermarket shelves are used as an alternative solution. These sensors send warning of low/out of stock products to SCM systems when they reach the alert levels [84], [83]. Data collected from sensors along with other data from public domains (e.g., weather and traffic channels) using big data analytics tools are also integrated into SCM as contextual information [80], [81].

As alluded before, in general, an ERP system includes several basic functions similar to CRM and SCM systems. However, only ERP is not sufficient enough for business activities and management, especially for medium and large business organisations. Therefore, there is a need for utilising these three systems into current business organisations. ERP focuses on internal business processes, while CRM is on more targeted marketing, sales and customer services, and SCM mainly handles product distribution and supplier management.

4.2.4 Distribution of academic studies and commercial products for ERP, CRM and SCM systems that use the cloud, big data or context

For ERP, CRM and SCM systems, so far what we have presented show academic studies and commercial products that leverage big data, context or both are being evolved over time. This increasing usage of big data or contextual information represents their crucial role in adding value to business and significance in the introduction of innovative services for these types of business systems.

Table 8 shows the distribution of academic studies and commercial products that use the cloud, big data or context. Both CRM and SCM have a few approaches that use both big data and contextual information indicating the importance of context for harnessing the value from big data. Because these two systems require to handle dynamically changing information having different types of both structured and unstructured data.

Data are also being continuously generated from many various sources, especially from external sources like the data sources for sale automation; customer service and support; market and demand management; and logistics and transport management. Table 8 also presents the adoption of cloud services by ERP, CRM and SCM systems. Since ERP uses mainly internal data, it has more cloud services than CRM and SCM. Note, all commercial products for ERP, CRM and SCM systems are cloud-based. The main reasons for this are the enormous benefits of the cloud. Cloud significantly reduces the cost for software, hardware and maintenance; provides the illusion of availability of infinite storage, computing and I/O resources; dynamically scales those resources based on demands; and supports real-time data analytics and cognitive decision-making process for many applications from business, financial to healthcare.

Table 8. Research studies and commercial products for ERP, CRM and SCM systems that consider big data or/and context

Research study/vendor	**Business application**	**Product type**	**Cloud-based**	**Use**	
				Big data	**Context**
[58] [59] [60]	ERP	Academic studies	×	×	✓
[61] [62] [63]			×	✓	×
[67]			✓	✓	×
Acumatica Epicor Infor Microsoft Oracle-NetSuite SAP		Commercial product	✓	✓	×
[68]	CRM	Academic studies	×	×	✓
[69]			×	✓	✓
HubSpot Microsoft Oracle-NetSuite SAP Zoho		Commercial product	✓	✓	✓
Aspedia			×	✓	✓
[73] [74] [75] [76] [77] [78] [79] [80] [81] [82] [83]	SCM	Academic studies	×	✓	✓
[84] [85] [86]				✓	×
Epicor Infor GTNexus Manhattan Oracle SAP		Commercial product	✓	✓	✓

The applications and studies of big data analytics and context awareness for ERP/CRM/SCM systems have been proven to improve cost-effectiveness and introduce innovative services. As mentioned before, as ERP is mainly responsible for managing internal business processes, while CRM and SCM primarily deal with external factors, the research and applications on big data analytics and context awareness for CRM and SCM are substantially higher than those for ERP.

5 Future research challenges

As emphasised in the previous sections, the importance of embedding context into computing to support business management systems and the effectiveness of these methods/applications has been recognised by the relevant research community. Although many approaches/applications using context models have been developed, major challenges for business organisations and significant research issues for the research community remain. Future research challenges on context-aware systems for business applications are outlined below:

1. Once a business process is contextualised, it can be used in data mining for targeted collection of data and then process them accurately. This will collect more data but with increased relevancy through semantic manipulation, thereby enabling capture of deep insights. In the work reported in [38], we introduced a method that addressed this issue for text type data only, but it is not suitable for other data types (e.g., video, audio, image). Hence, it remains a challenge to model the business process to enable big data to handle diverse data types concurrently.
2. Current context models lack the use of image, audio and video data to represent themselves. However, in addition to text data, these multimedia elements can play a vital role in depicting a clearer picture of a context and its adaptation process over the time. To date, the techniques for metadata representation defining contexts in the domain of these elements and the scalability of those techniques to capture instantaneous context adaptation have not been explored. Therefore, challenges lie in developing innovative context-aware methods to extract comprehensive contextual information from these diverse data types.

   In addition, many metadata repository products are still designed by technical experts rather than business experts. Some of these products still have a cryptic metadata language, lack sophisticated reporting capabilities, are not context-sensitive, and require an understanding of the meta-model that describes the metadata objects and their relationships [87]. This indicates the pressing need for the development of a meta-data model focusing on a business view and context, in order to reduce the requirement to understand metadata objects and their relationships.

   On the other hand, the role of a metadata repository is extremely important in ERP. An ERP needs better metadata repository support to manage both structured and unstructured data, as well as methodologies or techniques that can analyse and process data more effectively.
3. A business organisation needs real-time processing of big data so that a user receives the required and desired information on time. The studies presented in Section 4 highlighted the importance of contextual information in real-time big data analytics, but the effort is still in an embryonic stage. The difficulty lies in the filtering and validation of relevant contextual information and processing being done within the time limits specified by respective applications. This motivates the

rigorous investigation of real-time context-aware big data processing techniques and the manipulation of contextual information of any business entity instantly, including internal and external entities.

4. As B2B collaborations are included in SCM, exchanging information/data between business organisations and their partners such as suppliers and retailers needs a stringent privacy and a lightweight and adaptive security policy. Rapid technological development is gradually making communication and computing devices very smart and tiny. As all of these devices are connected through a pervasive computing environment, they are always connected and business services are available on any device at anytime from anywhere across the world. This demands a security framework with pervasive security mechanisms (e.g., intrusion detection, security audit trails) for SCM enforcing stringent privacy among businesses and partners and a scalable and adaptive security policy.

6 Conclusions

In this paper, an overview of context-based methods in big data analytics for business applications is presented. A general background of big data analytics, the definition of context, context models and their evolution, and context evaluation techniques are also provided. We have presented three the state-of-the-art business applications, namely enterprise resource planning, customer relationship management and supply chain management, and showed how current research and commercial products of those applications, supported by enterprise level software, exploited big data and/or contextual information to their advantage. A number of further research challenges and their significance in business applications are emphasized. Of particular interests are how to model the business processes to enable big data to handle diverse data types and to devise methods for comprehensive extraction of contextual information from those data, perhaps by defining scalable metadata representation of context in various domains. Solution to these and other research challenges outlined in the previous section will see widespread use of context-aware business applications by small to large scale organisations in the next decade.

References

[1] L. Sokol and S. Chan, "Context-Based Analytics in a Big Data World: Better Decisions," An IBM Redbooks Point-of-View publication, 2013.

[2] N. Hariri, M. Bamshad and B. Robin, "Query-Driven context aware recommendation," in *ACM conference on Recommender systems*, 2013.

[3] R. Aknouche, O. Asfari, F. Bentayeb and O. Boussaid, "Integrating query context and user context in an information retrieval model based on expanded language modeling," in *Multidisciplinary Research and Practice for Information Systems. Springer Berlin Heidelberg*, 2012.

[4] K. Li, H. Jiang, L. T. Yang and A. Cuzzocrea, "Big data: algorithms, analytics, and applications," in *CRC Press*, 2015.

[5] W. Fan and A. Bifet, "Mining big data: Current status, and forecast to the future," *ACM sIGKDD Explorations Newsletter,* vol. 14, no. 2, pp. 1-5, 2013.

[6] G. D. Abowd, A. K. Dey, P. J. Brown, P. J. Davies, N. Smith and P. Steggles, “Towards a Better Understanding of Context and Context-Awareness,” in *Handheld and ubiquitous computing. Springer Berlin Heidelberg,* 1999.

[7] A. Lorentz, “With Big Data Context is a Big Issue,” April 2013. [Online]. Available: http://www.wired.com/insights/2013/04/with-big-data-context-is-a-big-issue/. [Accessed 5 May 2016].

[8] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan and D. Riboni, “A survey of context modelling and reasoning techniques,” *Pervasive and Mobile Computing,* vol. 6, no. 2, pp. 161-180, 2010.

[9] S. Smanchat, S. Ling and M. Indrawan, “A survey on context-aware workflow adaptations,” in *Advances in Mobile Computing and Multimedia (MoMM),* 2008.

[10] W. Liu, X. Li and D. Huang, “A survey on context-awareness,” in *Computer Science and Service System (CSSS),* 2011.

[11] P. Bellavista, A. Corradi, M. Fanelli and L. Foschini, “A Survey of Context Data Distribution for Mobile Ubiquitous Systems,” *ACM Computing Surveys (CSUR),* vol. 44, no. 4, pp. 24:1-24:45, 2012.

[12] A. Abbas, L. Zhang and S. U. Khan, “A survey on context-aware recommender systems based on computational intelligence techniques,” *Computing,* vol. 97, no. 7, pp. 667-690, 2015.

[13] G. George, M. R. Haas and A. Pentland, “Big data and management,” *Academy of Management Journal ,* vol. 57, no. 2, pp. 321-326, 2014.

[14] M. Chen, S. Mao and Y. Liu, “Big data: a survey,” *Mobile Networks and Applications,* vol. 19, no. 2, pp. 171-209, 2014.

[15] T. Rout, M. R. Senapati, M. Garanayak and S. K. Kamilla, “Big data and its applications: A review,” in *International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO),* 2015.

[16] S. Mishra, V. Dhote, G. S. Prajapati and J. P. Shukla, “Challenges in Big Data Application: A Review,” *International Journal of Computer Applications,* vol. 121, no. 19, pp. 42-46, 2015.

[17] S. E. Bibri and J. Krogstie, “The core enabling technologies of big data analytics and context-aware computing for smart sustainable cities: a review and synthesis,” *Journal of Big Data,* vol. 4, no. 1, p. 38, 2017.

[18] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. Netto and R. Buyya, “Big Data Computing and Clouds: trends and future directions,” *Journal of Parallel and Distributed Computing,* vol. 79, pp. 3-15, 2015.

[19] M. F. Uddin and N. Gupta, “Seven V's of Big Data understanding Big Data to extract value.,” in *Zone1 Conference of the American Society for Engineering Education (ASEE Zone 1),* 2014.

[20] J. Fan, F. Han and H. Liu, “Challenges of big data analysis,” *National Science Review,* vol. 1, no. 2, pp. 293-314, 2014.

[21] P. Russom, “Big Data analytics,” TDWI Best Practices Report, Fourth Quarter, 2011.

[22] A. Rajendra, Big data computing, CRC Press, 2013.

[23] S. Sagiroglu and D. Sinanc, “Big data: a review,” in *International Conference on Collaboration Technologies and Systems (CTS),* 2013.

[24] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A. H. Byers, “Big data: the next frontier for innovation, competition and productivity,” Mckensey Global Institute, 2011.

[25] A. Gandomi and M. Haider, “Beyond the hype: Big data concepts methods and analytics,” *International Journal of Information Management,* vol. 35, no. 2, pp. 137-144, 2015.

[26] D. Loshin, Big data analytics: from strategic planning to enterprise integration with tools, techniques, noSQL, and Graph, San Francisco, CA: Morgan Kaufmann Publishers Inc, 2013.

[27] B. R. Presentation, “The challenge of big data,” Ventana Research, 2012. [Online]. Available: http://www.ventanaresearch.com/uploadedFiles/Content/Landing_Pages/Ventana_Research_Big_Data_Benchmark_Research_Presentation.pdf. [Accessed 19 August 2015].

[28] Techrepublic.com, “Tech Republic Company”.

[29] A. Ghazal, T. Rabl, M. Hu, F. Raab, M. Poess, A. Crolotte and H.-A. Jacobsen, “Big Bench: Towards an industry standard benchmark for big data analytics,” in *The ACM SIGMOD International Conference on Management of Data (SIGMOD)*, 2013.

[30] N. Elgendy and A. Elragal, “Big data analytics: A literature review paper,” *Advances in Data Mining. Applications and Theoretical Aspects,* vol. 8557, pp. 214-227, 2014.

[31] J. Gantz and D. Reinsel, “The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east,” IDC iView: IDC Analyze the future, 2012.

[32] S. Lee, S. Park and S. G. Lee, “A study on issues in context-aware systems based on a survey and service scenarios,” in *Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing*, 2009.

[33] J. L. D. L. Vara, R. Ali, F. Dalpiaz, J. Sanchez and P. Giorgini, “Business processes contextualization via context analysis,” *Conceptual Modeling - ER,* vol. 6412, pp. 471-476, 2010.

[34] S. Boutanmina and R. Maamri, “A survey on context-aware workflow systems,” in *Intelligent Information Processing, Security and Advanced Communication*, 2015.

[35] D. Ejigu, M. Scuturici and L. Brunie, “An ontology-based approach to context modelling and reasoning in pervasive computing,” in *Pervasive Computing and Communications Workshops*, 2007.

[36] P. S. Tan, A. E. S. Goh and S. S. G. Lee, “An ontology to support Context-Aware B2B services,” in *Services Computing*, 2010.

[37] M. Leppanen, “A context-based enterprise ontology,” in *Business Information Systems,* 2007.

[38] L. T. N. Dinh, G. Karmakar, J. Kamruzzaman and A. Stranieri, “Business context in big data analytics,” in *International Conference on Information, Communications and Signal Processing (ICICS),* 2015.

[39] I. Kroschel, “On the notion of context for business process use,” in *ISSS/BPSC,* 2010.

[40] P. J. Brown, J. D. Bovey and X. Chen, “Context-Aware Applications: from the laboratory to the marketplace,” *Personal Communications,* vol. 4, no. 5, pp. 58-64, 1997.

[41] K. Ploesser, M. Peleg, P. Soffer, M. Rosemann and J. C. Recker, “Learning from context to improve business processes,” *BPTrends,* vol. 6, no. 1, pp. 1-7, 2009.

[42] J. Bai, J. Y. Nie, G. Cao and H. Bouchard, “Using query contexts in information retrieval,” in *The 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2007.

[43] H. Cao, D. H. Hu, D. Shen, D. Jiang, J. T. Sun, E. Chen and Q. Yang, “Context-aware query classification,” in *International ACM SIGIR conference on Research and development in information retrieval*, 2009.

[44] A. K. Dey, "Understanding and Using Context," *Personal and ubiquitous computing,* vol. 5, no. 1, pp. 4-7, 2001.

[45] J. Coutaz, J. L. Crowley, S. Dobson, S and D. Garlan, "Context is key," *Communications of the ACM,* vol. 48, no. 3, pp. 49-53, 2005.

[46] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in *International conference on the practical applications of knowledge discovery and data mining*, 2000.

[47] "Big Data - A New World of Oppotunities," December 2012. [Online]. Available: http://www.nessi-europe.eu/Files/Private/NESSI_WhitePaper_ BigData.pdf. [Accessed 15 January 2016].

[48] W. J. Turkel and A. Crymble, *Keywords in Context (Using n-grams) with Python,* The Programming Historian 1, 2012.

[49] P. S. Tan, A. E. S. Goh, and S. S. G. Lee, "A context model to support B2B collaboration," in *Enabling Context-Aware Web Services: Methods, Architectures, and Technologies*, CRC Press, 2010, pp. 243-271.

[50] P. S. Tan, S. S. G. Lee, A. E. S. Goh and E. W. Lee, "Context-enabled B2B Collaborations," in *International Conference on Services Computing (SCC)*, 2007.

[51] O. Saidani and S. Nurcan, "Towards context aware business process modeling," in *Workshop on Business Process Modeling, Development, and Support (BPMDS'07), CAiSE*, 2007.

[52] M. Rosemann, J. Recker and C. Flender, "Contextualisation of Business processes," *International Journal of Business Process Integration and Management,* vol. 3, no. 1, pp. 47-60, 2008.

[53] I. Ruthven, "Information Retrieval in Context," *Advanced topics in information retrieval,* vol. 33, pp. 187-207, 2011.

[54] G. K. Mostéfaoui and P. Brézillon, "A Generic Framework for Context-Based Distributed Authorizations," in *International and Interdisciplinary Conference on Modeling and Using Context, Springer Berlin Heidelberg*, 2003.

[55] R. Ali, F. Dalpiaz and P. Giorgini, "A goal-based framework for contextual requirements modeling and analysis," *Requirements Engineering,* vol. 15, no. 4, pp. 439-458, 2010.

[56] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," *EBSE Technical Report,* vol. 2, no. 3, 2007.

[57] K. B. Hendricks, V. R. Singhal and J. K. Stratman, "The impact of enterprise systems on corporate performance: A study of ERP, SCM, and CRM system implementations," *Journal of operations management,* vol. 25, no. 1, pp. 65-82, 2007.

[58] F. Daneshgar, "Context-Aware Framework for ERP," in *Encyclopedia of Information Science and Technology,* vol. 27, 2005, pp. 105-117.

[59] C. A. Rajan and R. Baral, " Adoption of ERP system: An empirical study of factors influencing the usage of ERP and its impact on end user," *IIMB Management Review,* vol. 27, no. 2, pp. 105-117, 2015.

[60] M, Bradford and J. Florin, "Examining the role of innovation diffusion factors on the implementation success of enterprise resource planning systems," *International journal of accounting information systems,* vol. 4, no. 3, pp. 205-225, 2003.

[61] M. S. P. Babu and S. H. Sastry, "Big data and predictive analytics in ERP systems for automating decision making process," in *IEEE 5th International Conference on Software Engineering and Service Science,* Beijing, 2014.

[62] Z, Shi and G, Wang, "Integration of big-data ERP and business analytics (BA)," *The Journal of High Technology Management Research,* 2018.

[63] M. A. Vasarhelyi, A. Kogan, A and B. M. Tuttle, "Big Data in accounting: An overview," *Accounting Horizons,* vol. 29, no. 2, pp. 381-396, 2015.

[64] D. Angrave, A. Charlwood, I. Kirkpatrick, M. Lawrence and M. Stuart, "HR and analytics: why HR is set to fail the big data challenge," *Human Resource Management Journal,* vol. 26, no. 1, pp. 1-11, 2016.

[65] N. Jain, "Big Data and Predictive Analytics: A Facilitator for Talent Management," *Data Science Landscape. Studies in Big Data,* vol. 38, 2018.

[66] F. Liu, W. Guo, H. Wang and X. Li, "Data Science and Big Data Technology Professional Talent Demand and Training System Construction," in *9th International Conference on Education and Social Science (ICESS 2019),* 2019.

[67] M. Khazaeli, L. Javadpour and G. M. Knapp, "ERP adoption in enterprises with emerging Big Data," in *IIE Annual Conference, Institute of Industrial and Systems Engineers (IISE),* 2015.

[68] S. C. Park, K. H. Im, J. H. Suh, C. Y. Kim and J. W. Kim, "Ubiquitous customer relationship management (uCRM)," in *International Conference on Rough Sets and Knowledge Technology, Springer,* Berlin, 2007.

[69] K. Geihs, R. Reichle, M. Wagner and M. U. Khan, "Modeling of context-aware self-adaptive applications in ubiquitous and service-oriented environments," *Software engineering for self-adaptive systems,* pp. 146-163, 2009.

[70] *Architecture for a context-aware CRM,* http://www.intuital.com/.

[71] T. Nguyen, L. Zhou, V. Spiegler, P. Ieromonachou and Y. Lin, "Big data analytics in supply chain management: A state-of-the-art literature review," *Computers & Operations Research,* vol. 98, pp. 254-264, 2018.

[72] D. Mishra, A. Gunasekaran, T. Papadopoulos and S. J. Childe, "Big Data and supply chain management: a review and bibliometric analysis," *Annals of Operations Research,* vol. 270, no. 1-2, pp. 313-336, 2018.

[73] Y. Choi, H. Lee and Z. Irani, "Big data-driven fuzzy cognitive map for prioritising IT service procurement in the public sector," *Annals of Operations Research,* vol. 270, no. 1-2, pp. 75-104, 2018.

[74] M. H. Tan and W. L. Lee, "Evaluation and improvement of procurement process with data analytics," *International Journal of Advanced Computer Science and Applications,* vol. 6, no. 8, p. 70, 2015.

[75] Y. Zhang, S. Ren, Y. Liu and S. Si, "A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products," *Journal of Cleaner Production, 142, 626-641.,* vol. 142, pp. 626-641, 2017.

[76] P. Helo and Y. Hao, "Cloud manufacturing system for sheet metal processing," *Production Planning and Control,* vol. 28, no. 6-8, pp. 524-537, 2017.

[77] J. Krumeich, D. Werth and P. Loos, "Prescriptive control of business processes," *Business & Information Systems Engineering,* vol. 58, no. 4, pp. 261-280, 2016.

[78] J. Li, M. Moghaddam, and S. Y. Nof, "Dynamic storage assignment with product affinity and ABC classification—a case study," *The International Journal of Advanced Manufacturing Technolog,* vol. 84, no. 9-12, pp. 2179-2194, 2016.

[79] B. Li, E. Ch'ng, A. Y. L. Chong and H. Bao, "Predicting online e-marketplace sales performances: a big data approach," *Computers & Industrial Engineering, 101, 565-571.,* vol. 101, pp. 565-571, 2016.

[80] G. Walker and A. Strathie, "Big data and ergonomics methods: a new paradigm for tackling strategic transport safety risks," *Applied ergonomics,* vol. 53, pp. 298-311, 2016.

[81] S. L. Ting, Y. K. Tse, G. T. S. Ho, S. H. Chung and G. Pang, "Mining logistics data to assure the quality in a sustainable food supply chain: A case in the red wine industry," *International Journal of Production Economics,* vol. 152, pp. 200-209, 2014.

[82] R. Mehmood, R. Meriton, G. Graham, P. Hennelly and M. Kumar, "Exploring the influence of big data on city transport operations: a Markovian approach," *International Journal of Operations & Production Management,* vol. 37, no. 1, pp. 75-104, 2017.

[83] A. Y. L. Chong, B. Li, E. W. T. Ngai, E. Ch'ng and F. Lee, "Predicting online product sales via online reviews, sentiments, and promotion strategies: A big data architecture and neural network approach," *International Journal of Operations & Production Management,* vol. 36, no. 4, pp. 358-383, 2016.

[84] M. Salehan and D. J. Kim, "Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics," *Decision Support Systems,* vol. 81, pp. 30-40, 2016.

[85] K. J. Wu, C. J. Liao, M. L. Tseng, M. K. Lim, J. Hu and K. Tan, "Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties," *Journal of Cleaner Production,* pp. 663-676, 2017.

[86] T. Papadopoulos, A. Gunasekaran, R. Dubey, N. Altay, S. J. Childe and S. Fosso-Wamba, "The role of Big Data in explaining disaster resilience in supply chains for sustainability," *Journal of Cleaner Production,* vol. 142, pp. 1108-1118, 2017.

[87] L. T. Moss and S. Atre, Business Intelligence Roadmap: The complete Project Life cycle for decision-support applications, Addison-Wesley Professional, 2003.

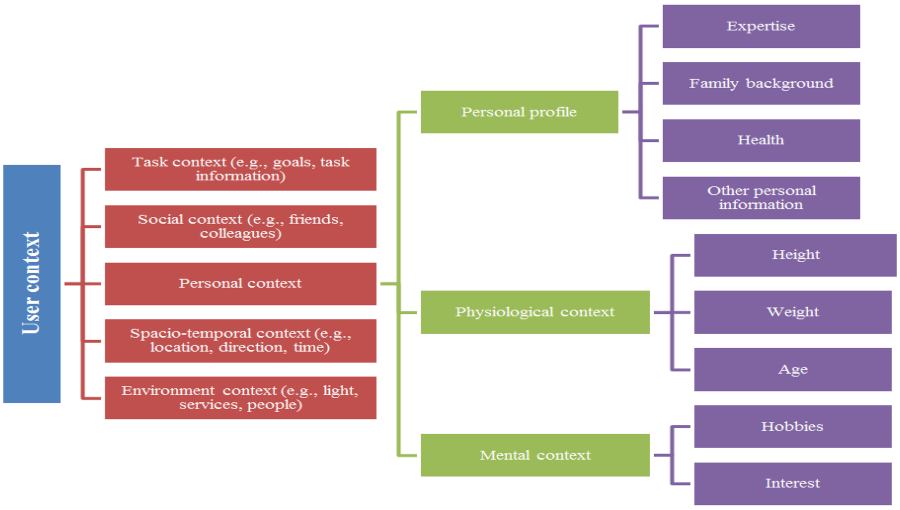
Loan Thi Ngoc Dinh received Master of Information Systems Management from Swinburne University of Technology, Australia, in 2011 and Doctor of Philosophy from Federation University Australia, in 2019. Her research interests include information systems and big data analytics.

Gour Karmakar received a B.Sc. degree from CSE, BUET in 1993 and Masters and Ph.D. degrees in Information Technology from the Faculty of Information Technology, Monash University, in 1999 and 2003, respectively. He is currently a senior lecturer at Federation University Australia. He has published over 150 peer-reviewed research publications including 33 international peer-reviewed reputed journal papers and was awarded six best papers in reputed international conferences. He received a prestigious ARC linkage grant in 2011. He has successfully supervised 14 Ph.D. and four Masters by research projects from the

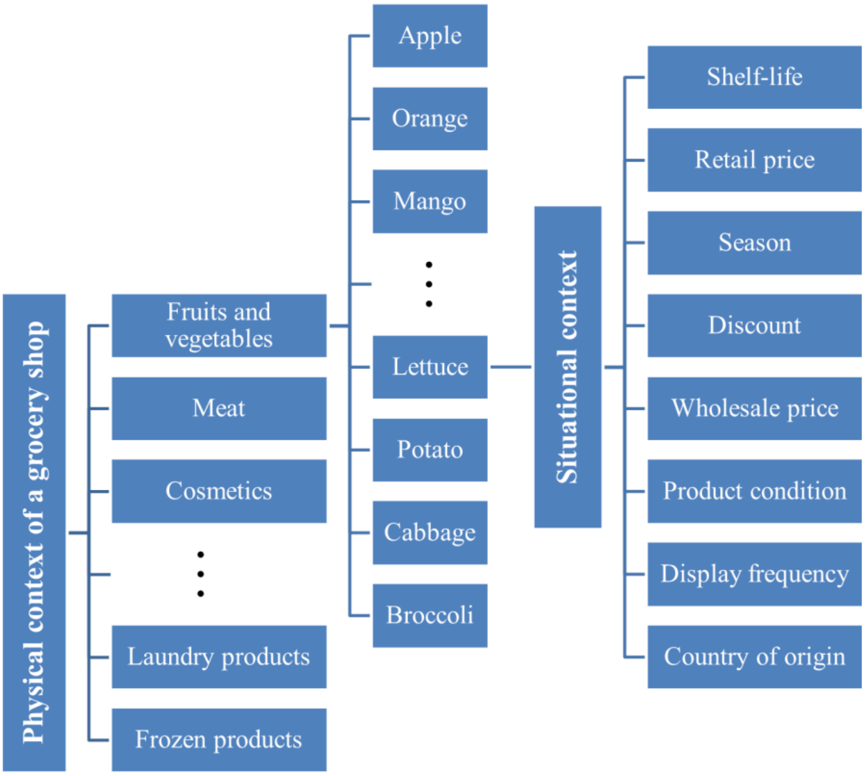
commencement to completion. One of his Ph.D. students received both Mollie Holman Doctoral and Faculty of Information Technology Doctoral Medals in 2009. His research interest includes multimedia signal processing, big data analytics, Internet of Things and cybersecurity.

Joarder Kamruzzaman received the BSc and MSc degrees in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology, Dhaka, and the PhD degree in Information Systems Engineering from Muroran Institute of Technology, Hokkaido, Japan. He is currently a Professor in the School of Science, Engineering and Information Technology, Federation University Australia.
His research interests include distributed computing, Internet of Things, machine learning and cyber security. He has published 250+ peer-reviewed publications which include 75 journal papers, 160+ conferences, 11 book chapters and two edited reference books. He is the recipient of Best Paper award in four international conferences: ICICS’15, Singapore; APCC’14, Thailand; IEEE WCNC’10, Sydney, Australia and in the IEEE-ICNNSP’03, Nanjing, China. His publications are cited 2750 times and have h-index 24, and g-index 43. He has received nearly A$2.3m competitive research funding, including prestigious ARC (Australian Research Council) grant and large CRC (Collaborative Research Centre) grant. He was the founding Program co-Chair of the first International Symposium on Dependability in Sensor, Cloud, and Big Data Systems and Applications (DependSys), China in 2015. He has served 32 conferences in leadership capacities including Program co-Chair, Publicity Chair, Track Chair and Session Chairs, and since 2012 as an Editor of the Elsevier Journal of Network and Computer Applications, and had served as the lead Guest of Elsevier Journal Future Generation Computer Systems and guest editor of Journal of Networks.

Appendix



A1. An example of a user context model



A2. A business context model for big data collection and processing