

Responsible Big Data Analytics for Enhancing E-Business Services

Xuehui Wei¹ and Zhaohao Sun^{2*}

¹Confucius Institute at Papua New Guinea University of Technology, Papua New Guinea

²School of Business, PNG University of Technology, Lae 411, Papua New Guinea

ABSTRACT

This article examines responsible big data analytics for e-business services and looks at how we can use responsible big data analytics to enhance responsible e-business services. It addresses why responsibility matters to big data analytics and e-business services. It reviews big data analytics and examines Google analytics as a data processing flow-oriented analytics and presents a data processing flow-oriented big data analytics. It examines big data analytics services as an e-business service and proposes four strategies for applying responsible big data analytics to enhance responsible e-business services. The proposed approach in this article might facilitate the research and development of big data, business analytics, responsible big data analytics, e-services, and digital services.

*Corresponding author

Zhaohao Sun, School of Business, PNG University of Technology, Lae 411, Papua New Guinea.

Received: July 22, 2024; **Accepted:** July 30, 2024; **Published:** August 05, 2024

Keywords: Big Data, Big Data Analytics, Responsible Big Data Analytics, e-Business Services, Artificial Intelligence (AI), Digital Services, Digitalized Society

Introduction

Big data and big data analytics have become one of the most important research frontiers in academia and industries [1]. Big data and its emerging technologies have been making big changes in the way e-business services operate [2]. Big data analytics is an emerging big data technology and has become a mainstream market adopted broadly across industries, organizations, geographic regions, and among individuals to facilitate big data-driven decision making and brought big value for organizations and individuals [3,4]. Big data and big data analytics are increasingly critical elements for nearly all industries and keys to a successful intelligent business, based on the following two predictions from Gartner researchers [5]. By 2026, more than a quarter of Fortune 500 CDAOs will have become responsible for at least one top-earning product that is based on data and analytics, 2). Through 2025, 90% of organizations that trade globally will have struggled to mitigate risks from sovereign data strategies and policies and incur costs that limit mission success.

The technical drivers for smart information systems (SIS) are artificial intelligence (AI), machine learning (ML), and big data analytics. the Internet of Things (IoT), autonomous machines, social media, location-based services, quantum technology, virtual/augmented reality, cloud computing, affective computing, and personal smart devices are enabling technologies for Information Communication Technologies (ICT) [6].

Gartner identifies responsible AI as one of its four trends driving near-term AI innovation [7]. Responsible AI become ever-increasingly important with AI's further development and its spreading applications in every part of our lives, industries, and

societies [8]. Considering the close relationship between big data analytics and AI, responsible big data analytics has also become an increasingly important topic for developing big data analytics [9,10]. Responsible big data analytics helps achieve fairness, gain trust, and improve transparency [7].

E-business including e-commerce and its services have benefited all the people in the world [11,1]. e-business services have become indispensable for billions of people around the world. However, responsible big data analytics for e-business services is a challenging issue for both academia and industries [12]. This issue can be detailed into the following research questions.

- Why does responsibility matter to big data analytics and e-business services?
- What is responsible big data analytics?
- How can we use responsible big data analytics to enhance responsible e-business services?

This article will address each of them by elaborating and extending our early work. To answer the first research question, we look at the data monetization of Facebook and Google. To answer the second research question, we review big data analytics by classifying big data analytics into four categories, look at Google analytics, and present a data processing flow approach to big data analytics considering a chain of responsibility. We also discuss responsible big data and big data analytics based on two case studies. To answer the third research question, we look at e-SMACS services as e-business services, examine big data analytics services as an e-business service, and propose four strategies for applying responsible big data analytics to enhance responsible e-business services.

The remainder of this article is organized as follows. Section 2 looks at why responsibility matters to big data analytics and e-business services. Section 3 reviews big data analytics, looks at

Google analytics and examines a data processing flow-oriented big data analytics. Section 4 discusses responsible big data analytics with two case studies. Section 5 examines e-SMACS services as e-business services and digital services and reveals their interrelationships. It also looks at big data analytics services as an e-business service. Section 6 proposes four strategies for applying responsible big data analytics to enhance responsible e-business services. The final sections discuss the related work and end this article with some concluding remarks and future work.

Why Responsible Big Data Analytics Matters?

This section addresses why responsible big data analytics matters from different perspectives.

In 2015, the General Assembly of the United Nations (UN) adopted the 2030 Agenda for Sustainable Development that includes 17 Sustainable Development Goals (SDGs) to realize that in 2030 all people enjoy peace and prosperity [13]. Responsible consumption and production are the 12th SDG of 17 SDGs. Big data analytics as a system and service is a part of responsible consumption and production [14]. Therefore, responsible big data analytics and their services are significant for realizing the 12th SDG of the UN.

Responsible big data analytics has drawn some attention in academia. Fothergill, et al. state that responsible data governance principles should include four responsible research and innovation features: anticipation, reflection, engagement, and responsiveness [15]. The inclusion of these principles would consider future consequences, integrate mechanisms for fostering reflexivity, and enable stakeholder engagement. Responsible research innovation principles can ensure that research in technologies, including big data and AI, are socially acceptable and responsible [6]. The underlying guidance for responsible big data analytics research is based on that most data impact people [10]. The privacy should be seen as contextual and depending on the nature of the data because privacy preferences differ across societies. Privacy is an important part of any responsible big data analytics as a service [16].

In the age of big data, analytics, and AI, big data and big data analytics should be responsible, because monetizing big data through the ever-increasing importance of global data trade and industry data trade using big data analytics can improve the business performance. Selling the internal data and big data analytics services to external consumers can be hundreds of billion-level profitable [17,18]. Data monetization of Facebook (or Meta) and Google (or Alphabet) has become a miracle in the past decade. They have no traditional natural resources like the iron core. They have only two artificial resources; one is big data from artificial donation and aggressive collection using intelligent techniques based on intelligent algorithms. Another is big data analytics which can transform big data into smart insights and big value [3,9]. The question arises: are the big data and big data analytics of Google and Facebook are responsible?

Gmail was launched in 2004; Google subsequently admitted that it has scanned private correspondence for personal information. In the same year, Facebook was founded, its business model is also based on the capture of and access to personal data and information [19].

Why does responsible data analytics matter to e-business services? There are many different answers to this question from different researchers. Big data analytics uses predictive algorithms and mathematically calculate human behaviors, and then data giants

sell certainty to business customers who would like to know with certainty what we do. The best way to make the predictions desirable to customers is to ensure that they come true: “to tune and herd and shape and push us in the direction that creates the highest probability of their business success” [19].

In the Cambridge Analytica (CA) scandal (also see Section 5), every aspect of CA’s operations was simply mimicking a day in the life of a surveillance capitalist” [19]. Trade of human futures using big data trade and big data analytics is irresponsible and illegal – like the slave trade was made illegal [20].

Big Data Analytics: A Data Processing Flow Approach

This section reviews big data analytics, examines Google Analytics as a data processing flow-oriented analytics, and presents a data processing flow approach to big data analytics.

What is Big Data Analytics?

Big data analytics is a science and technology about organizing and analyzing big data, and discovering knowledge, insights, and intelligence from big data, visualizing and reporting the discovered knowledge and insights for assisting decision making [21,3]. Big data analytics can facilitate better-informed decision-making and add significant value when applied to problems of specific domains, including healthcare, supply chain management, and marketing [22]. Big data analytics can be classified into big data descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics [21,23,24].

- Big data descriptive analytics is descriptive analytics for big data and is used to discover and explain the characteristics of entities and relationships among entities within the existing big data [2]. It addresses the problems such as what happened, and when, as well as what is happening.
- Big data diagnostic analytics is diagnostic analytics for big data [2]. It is used to examine diagnostic data or content to answer the question “Why did it happen?”, from the historical and current diagnostic data based on AI techniques such as drill-down, data discovery, data mining, and correlations [25]. For example, diagnostic analytics available on the cloud belongs to big data diagnostic analytics.
- Big data predictive analytics is predictive analytics for big data and focuses on forecasting trends by addressing the problems such as what will happen, what is likely to happen, and why it will happen. Big data predictive analytics is used to create models and insights to predict future outcomes or events based on the existing big data [2].
- Big data prescriptive analytics is prescriptive analytics for big data, which addresses the problems such as what we should do, why we should do it, and what should happen with the best outcome under uncertainty [9].

The fundamentals of big data analytics consist of mathematics, statistics, engineering, human interface, computer science, and data science [26]. The techniques for big data analytics include a wide range of mathematical, statistical, and modeling techniques [2]. Therefore, big data analytics can be represented below [21,24].

$$\text{Big data analytics} = \text{Big data} + \text{data analytics} + \text{DW} + \text{DM} + \text{SM} + \text{ML} + \text{visualization} + \text{optimization} \quad (1)$$

Where + can be explained as “and”; DW, DM, SM, and ML are the abbreviation of data warehousing, data mining, statistical modeling, and machine learning, respectively. This representation reveals the fundamental technological relationship between big data, data analytics, and big data analytics. It also shows that computer science and data science play a dominant role in the development of big data analytics by providing sophisticated techniques and tools of DW, DM, ML, and visualization [9,24]. SM and optimization still play a fundamental role in the development of big data analytics [9,23].

However, because human bias is removed from the systems, big data analytics can be used effectively in a variety of scenarios, such as predictive policing, where geographic areas with higher crime rates can be identified using a variety of data sources, such as prior crime data, reports, cameras, and weather [12]. On the other hand, organizations will need guidelines to assure the ethical use and application of big data analytics. Privacy, veracity, ownership, and accessibility are the four main difficulties that big data analytics faces. In this context, privacy entails ensuring that the information is kept private, while veracity entails assuring accountability for the information's legitimacy, integrity, and accuracy. The ownership components ensure that the information's owner is known and that access to the data is controlled and valued. The accessibility component guarantees that only individuals or organizations with the appropriate permissions are granted access, and data security protections are maintained.

Google Analytics as a Data Processing Flow Oriented Analytics

A data processing flow consists of four stages from left to right, from upstream to downstream, illustrated in Figure 1. In what follows, we look at each stage of the data processing flow and its corresponding Google Analytics services [27].



Figure 1: A Model for Data Processing Flow-Oriented Google Analytics

Store and Organize

The first stage consisting of “store and organize” is at the most left of the data processing flow. The corresponding Google analytics aims to store and organize website tags, which is accomplished by Google Tag Manager™. Tag Manager also allows running website surveys and multivariate optimization tests, tracking social media audiences [27,12]. Tag Manager can make changes to the website without requiring editing the site itself.

Collect and Clean

The second stage on the next left of the data processing flow consists of “collect and clean” [12]. The corresponding Google analytics aims to collect and clean data [27]. All the analytics tools in this stage have their APIs, but some external data sources like social media sites require data collection separately. The Google data cloud gathers big data and prepares it for analyzing and reporting. For example, we can collect social media engagement data from third-party vendors in a cloud MySQL database or

Google Sheets. We can extract ad data from third-party advertising systems and store it in Cloud SQL™.

Analyze and Optimize

The next stage on the right of the data processing flow consists of “analyze and optimize” [12]. The corresponding Google analytics are Google Optimize and Google AdWords aiming to analytically analyze big data and optimize the discovered value, knowledge, and insights [18,27]. Google analytics tells us what is happening on our website, which is a kind of descriptive analytics. Google Optimize can test out a variety of assumptions to understand what works best on our sites. Google AdWords then enables us to market to the audiences that we prove effective with analytics and Optimize [27]. Google analytics alone used to be the star of the show, but now it is one of the Google analytics platforms.

Visualize and Report

The rightmost stage of the data processing flow consists of “visualize and report” [12]. The corresponding Google analytics is Data Studio™. Data Studio can assemble and visualize big data and discover knowledge and insights [27]. Data Studio is where we transform our big data into analysis and insights, then make strategic recommendations about what to do next, which can be a part of predictive and prescriptive analytics [12]. We can make a basic but effective public relations reporting system out of Google BigQuery™ with the raw news feed from Google News.

Data Processing Flow-Oriented Big Data Analytics: A Unified Approach

The previous discussion demonstrates that Google Analytics is data processing flow-oriented. Google develops analytics as systems and services for each stage of the mentioned data processing flow, from upstream to downstream.

With the dramatic development of big data and big data technology in the past decade, big data has become hundreds of billion-level industries [28]. The data processing in the big data industry, more generally, consists of n stages of data processing flow from upstream to downstream, where n is a natural number, as follows.

$$DP_1, DP_2, \dots, DP_n, \tag{1}$$

Where DP is the abbreviation of data processing. Each of the above stages of the data processing flow corresponds to at least one intelligent analytics that is,

$$DP_1 \text{ analytics}, DP_2 \text{ analytics}, \dots, DP_n \text{ analytics} \tag{2}$$

Every DP_i ($i \in \{1, 2, \dots, n\}$) analytics might be renamed properly

The number n will be increased to very great with the further development of the big data industry. Therefore, more and more new intelligent analytics and their services will be emerging [3]. There are also more challenges for developing big data analytics and intelligent analytics as a system or service based on data processing flow. One of the challenges is the responsibility of these data processing flow-oriented big data analytics.

Data processing flow-oriented big data analytics should be responsible for every stage of the data processing flow. In other words, the responsibility of big data analytics should be divided into and aligned with every stage of the data processing flow. This is similar to the total quality control that should be carried out in every stage of a production line, rather than the final quality

check before leaving for selling to the customers. We will delve into it in Section 6.

In addition, a full analytic workflow is from data preparation to visual exploration and insight generation [3]. This corresponds to the proposed data processing flow. Then we have data processing flow-based analytics. One who controls the data processing flow-based analytics from upstream to downstream will dominate the data industry, just as Google has been doing.

Responsible Big Data and Big Data Analytics

This section looks at responsible big data and big data analytics. Based on Cambridge Dictionary, the term responsible means 1. “having control and authority over something or someone and the duty of taking care of it or them” 2 [29]. “Having good judgment and ability to act correctly and make decisions on you own”.

Therefore, big data analytics is responsible if it satisfies the following three conditions.

- It has control and authority over something or someone.
- It has the duty of taking care of it or them.
- It has good judgment and the ability to act correctly and make decisions properly based on the existing law and rules.

Briefly, responsible big data analytics is about human responsibility for the development and deployment of big data analytics along fundamental human principles and values to ensure human flourishing and well-being in a sustainable world, based on the definition of responsible AI [8]. Responsible big data and e-business services can be defined similarly. The responsibility of big data and big data analytics is relative in terms of individual, organization, community, society, and world. That is, for example, big data analytics should be responsible for individuals, organizations, people, community, country, and for international fairness. The following two cases will look at if the big data and big data analytics are responsible for individuals, communities, and societies.

Cambridge Analytica (CA) was a British analytics firm that aimed to use big data to build detailed and deeply personal, psychological profiles about people and then target them with what is essentially emotional manipulation in the form of ads [30]. CA had an app, collecting and exploiting a total of over 87 million Facebook users’ data before the profiling was exposed [20,31]. CA could have created psychological profiles of 230 million Americans according to estimates. At that time, Facebook’s platform did permit the apps to access the data from the users of the app and their friends, unless the friends explicitly prohibited the collection in their privacy settings [30].

In March 2018, CA was accused of engaging in enhanced micro-targeting and unethically using this information to support political campaigns in various countries [20]. CA used behavioral micro-targeting through its massive dataset to deliver targeted messages supporting the Trump campaign in the 2016 US Elections [30]. The Facebook-CA data and analytics case raised several ethical issues related to the responsible usage of big data by social networking sites, including user data collection and use practices [20,30].

Responsible big data and big data analytics are required at every stage along with the data processing flow from upstream and downstream, mentioned in Section 3.3. A new issue arises here: How to apply responsible big data and big data analytics to enhance responsible e-business services through monetization

of big data and big data analytics as an e-business service? We will address it in Section 6.

E-Business Services

Recently, e-business services have become indispensable for our lives, work, and society. The young generation has heavily relied on e-business services. This section looks at e-SMACS services as e-business services and big data analytics services as an e-business service. It also examines the characteristics of digital services.

E-Business Services as E-SMACS Services

E-business has experienced significant development since 1995 [1]. In the past about three decades, e-business has been transitioned from business online to e-business services, which have been evolved into e-services, social (online) networking services, mobile services, analytics services, cloud services, and security services [1]. All these can be considered e-SMACS services [32,24]. E-SMACS services can be called web services or digital services [2]. Almost all traditional services such as FREG (foods, resources, energy, and goods) services are fully or partially replaced by web services or digital services. In what follows, we will briefly look at each of them in some detail.

E-services

E-services are electronic services, corresponding to e-commerce introduced in the middle 1990s [1]. Generally, every service online is an e-service. Therefore, e-commerce services are one of the most important e-services. Other e-services are free services online. For example, many of us enjoy the Gmail as a free service. Most of us are expecting free Wi-Fi services [2].

Social networking services

Social networking services (SNS) correspond to social commerce and can be called social commerce services [1]. Social networking services include social media services [24]. Meta’s services, WeChat services, and TikTok services belong to this category.

Mobile Services

Mobile services correspond to mobile business and can be called mobile business services. Mobile services include location-based services and mobile banking services [11]. Mobile services have evolved rapidly over the past two decades [32]. The most popular mobile services include mobile financial or banking services, mobile commerce services (eg. Weixin pay and Ali Pay), mobile SNS, and the Internet of services (IoS).

Analytics Services

Analytics as a service (AaaS) or analytics service means that an individual or organization uses a wide range of analytics tools or services [33,2]. AaaS can turn a general analytic platform into a shared utility for an enterprise with visualized analytic services [34]. An analytics service can be available on the cloud or used by smartphones. Therefore, analytics services include e-analytics services, cloud analytics services, and mobile analytics services [12]. Tableau and Qlik are vendors of big data analytics services [3,2].

Cloud Services

Briefly, a cloud service is any service provided in the cloud [2]. More specifically, cloud services are designed to provide flexible, scalable applications, resources, and services, and are fully managed by a cloud services provider based on “pay-as-you-go” models [35]. Cloud services mainly comprise four different types of services: Infrastructure as a Service (IaaS), Platform as

a Service (PaaS), and Software as a Service (SaaS) as well as storage as a service (StaaS) [36,37]. The Internet of Services (IoS) together with the Internet of people and the Internet of things are important three components of the Internet of Everything (IoE) [9]. Amazon Web Services are well-known cloud services. IoT (the Internet of Things) as a service, IoP (the Internet of People) as a service, and IoS as a service have appealed to academia and industries [32,2].

Security as a Service

Security as a service is supported by security computing and technologies. Managed security service providers include Cisco (<https://tools.cisco.com>), IBM (<https://www.ibm.com/>), Huawei (<https://www.huawei.com/en/>) and Alibaba (<https://www.alibaba.com>).

Each of the above-mentioned services can be classified further. For example, as a kind of e-services, e-business services include e-commerce services, e-supply chain services, e-marketing services, e-procurement services, e-banking services. All these services can be considered digital services.

E-Business Services, E-SMACS Services, and Digital Services

The Google search result reflects the participation, attention, and penetration of industries, individuals, and societies (we used the data searched in 2022, and no data available now). The Google Scholar search results reflect the participation, attention, and penetration of academia. A Google search (retrieved on January 15, 2022) and Google Scholar search for “e-SMACS services” and “e-business service” have been summarized in the following table (retrieved on July 9, 2024).

Table 1: The Summary of Google Search and Google Scholar Search for e-SMACS Services and Digital Services

e-SMACS services	Google (all) Search (M)	Google Scholar Search (N)
e-service	6,800,000	245,000
Social Networking Service	4,300,000	4,500,000
Mobile Service	15,500,000	8,160,000
Analytics Service	7,810,000	4,120,000
Cloud Service	18,200,000	5,340,000
Security Service		5,160,000
Digital Service	41,100,000	7,820,000
e-Business Service	899,000	853,000

M and N in the parentheses are the found results of searches. The searched results in Table 1 demonstrates that:

- Based on the search number of Google (all), digital services and cloud services are most popular in society while mobile services and digital services are most popular in academia.
- E-business services are least popular in society, and e-service is least popular in academia.
- As modern e-business services, e-SMACS services become the most popular and inclusive in both society and academia. They are the most influential services of digital services

All the services mentioned in Table 1 can be considered as e-business services. The above preliminary statistical analysis holds:

$$e\text{-business services} \subseteq e\text{-SMACS services} \subset \text{digital services} \quad (2)$$

In other words, e-SMACS services are modern e-business services, and they are an important part of digital services. That is,

$$e\text{-SMCS services} + e\text{-business services} \subseteq \text{Digital services} \quad (3)$$

Finally, all these services lead to XaaS (anything as a service), Systems, platforms, clouds, tools and technologies delivered as a service also belongs to XaaS [38].

Characteristics of Digital Services

Digital services including e-business services are playing a crucial role in the digital economy. Digital services at least encompass six characteristics: granularity, instant, volume, ubiquity, variety, and cost reduction [39].

Granularity

Digital services vary from big services through microservices to nanoservices. In other words, digital services can be classified in a hierarchical structure from big to small to nano services.

Instant

The response time of a service request has been from a year-long to a nanometer short. For example, in the agriculture society, a farmer might sell his products such as wheat and rice to customers once a year. This means that the time from the request to the provision of the service is one year. In the digital age, the response time to request of services is instant. This is a machine-to-human service. In fact, in machine-to-machine (M2M) services, the response to the request of service is at a nanosecond level. Such services can be called nanoservices.

Volume

The number of service consumers is from 1 to billions because the service consumers of Facebook have surpassed 2 billion. All the services form the IoS.

Ubiquity

Services are ubiquitous. In particular, in the digital age, one can enjoy services like e-SMACS services and mobile commerce services at anytime and anywhere.

Variety

Services can be classified into a variety of categories. At a strategic level, services can be classified into human-to-human (H2H) services, human-to-machine (H2M) services, and M2M services. Services can be also classified into physical services and virtual services. The former includes offline services, and the latter includes online services. Furthermore, service providers and consumers are either humans or machines or software systems or apps. The customers do not care who and what is the services provider.

Cost Reduction

In the digital age, the cost of most digital services has been reduced. For example, in 1995, one of the authors called his friends in China from Germany through international call, the cost was 5 German Marks per minute. Now, he calls his friends in China using call functions provided by WeChat, the price is almost zero, considering the cost for mobile data. However, some traditional services become more expensive than two decades ago.

Big Data Analytics Services as an e-Business Service

Currently, big data and big data analytics are based on five cutting-edge technology pillars of electronic, social, mobile, cloud and security technologies. Each of these pillars corresponds to a special kind of e-business services, that is, electronic, social, mobile, cloud, and security services; all these constitute modern e-business services [21,3,24]. Each of these services has been empowered by sophisticated ICT technologies including big data and big data analytics as shown in Figure 2 [3,23].



Figure 2: Big Data Analytics as a Service and Technology

Currently, the core services and technologies of big data analytics include [2,3,12]:

- Support building, deployment, and management of analytics in the cloud based on data stored both in the cloud and on-premises through cloud analytics technology.
- Enable users to connect to, query, and ingest data, while optimizing for performance.
- Support for drag-and-drop, a user-driven combination of data from different sources, and the creation of analytic models through data preparation technology.
- Automatically generate and curate a searchable catalog of analytic content, and package and deliver analytic content in a compelling, easily understood form for presentation to decision-makers through insight generation and interpretation technology.
- Apply AI techniques to automatically generate findings or insights for end-users.
- Support for highly interactive dashboards and exploration of big data through data visualization technology.
- Create and distribute intelligent reports for end-users through visualization and reporting technology.

In addition, techniques of management, natural language query, processing, and generation have also been built-in big data analytics to empower e-business services [3].

Applying Responsible Big Data Analytics to Enhance Responsible e-Business Services

This section proposes four strategies for applying responsible big data analytics to enhance responsible e-business services.

Digital information sharing is indispensable for our work and lives in the digitalized society. Meta (formerly Facebook) is a digital giant with a set of e-business services on that we rely daily, just as we rely on food for our daily survival. The question is: are the e-business services of Meta responsible to ban news sharing to Australian people recently?

Making data available to external organizations is a key mechanism for business-to-business (B2B) e-business services in the age of big data [17]. Big data analytics has increased the use of this digital data exchange thanks to AI and cloud computing. This leads to open data initiatives, data marketplaces, and vendors who offer data sets are emerging in every aspect of business and society. The core services of e-business giants including Meta, Google, Amazon, and Alibaba are big data services and big data analytics services. Are their e-business services responsible? These issues require solutions to responsible e-business services. In what follows, we propose four strategies for applying responsible big data analytics to enhance responsible e-business services.

The first strategy is that big data analytics and its services should meet the basic ethic standard mentioned in the top nine ethical issues in AI, besides satisfying the three conditions mentioned in the definition of responsible big data analytics [40,41].

- **Unemployment:** What happens after the end of jobs?
- **inequality:** How do we distribute the wealth created by AI machines?
- **Humanity:** How do machines affect our behavior and interaction?
- **Artificial Stupidity:** How can we guard AI against mistakes?
- **Racist Robots:** How do we eliminate AI bias?
- **Security:** How do we keep AI safe from adversaries?
- **Evil Genies:** How do we protect AI against unintended consequences?
- **Singularity:** How do we stay in control of a complex intelligent system?
- **Robot Rights:** How do we define the humane treatment of AI?

Responsible big data analytics should provide rational and satisfactory answers to each of the above for the world, taking into account the existing regulations and laws worldwide. For example, a rational and satisfactory answer can be below, because big data analytics is a part of AI [12].

- **Unemployment:** What happens after the adoption of big data analytics?
- **Inequality:** How do we distribute the insights, foresights, and wealth created by big data analytics?
- **Humanity:** How do big data analytics affect our behaviors and interactions?
- **Artificial Stupidity:** How can we guard big data analytics against mistakes and irresponsibilities?
- **Racist Robots:** How do we eliminate the bias caused by using big data analytics?
- **Security:** How do we keep big data analytics safe from adversaries?
- **Evil Genies:** How do we protect big data analytics against unintended consequences?
- **Singularity:** How do we stay in control of complex intelligent big data analytics?
- **Robot Rights:** How do we define the humane treatment of intelligent big data analytics?

The second strategy is that big data analytics and its services should be accountable or explainable that is, every service of big data analytics should be explainable and auditable based on the new regulation of AI systems [8,42]. In such a way, we could avoid the scandal of Facebook and Cambridge Analytica or similar [30]. We can also use big data analytics and its services responsibly to enhance responsible e-business services.

Stoller claims that “Facebook was born, lives, and thrives in scandal” [43]. For example, Facebook’s violations of privacy led to a Federal Trade Commission consent decree in 2012. Mr. Mark Zuckerberg was shown explicitly predatory in his business methods that were supplemented by the Federal Trade Commission complaint filed in 2021. However, Facebook or Meta is still a globe-straddling monopoly over our information commons although its years of lawlessness and recklessness. The problem may not be Meta, it might be a policy regime that creates an incentive for monopolization, securities fraud, and surveillance advertising as well as user data collection and processing. This at least includes the opt-out policy used by the social media giant Meta and other social media and internet advertising firms in the USA.

Opt-in is a personal collection policy in which the company collecting the information does not use the information for any other purpose (e.g., sell or rent the information or data) unless the customer specifically chooses to allow that use [11].

Opt-out is a personal collection policy in which a business is allowed to collect personal information until the consumer specifically requests the data not to be collected [2].

Opt-in is the practice of EU countries. Opt-out is the practice of the USA. The opt-in model has appealed to privacy advocates and forced the data collection company to be responsible for the customer’s interest and satisfaction. However, the opt-out model is very aggressive, and the data collection company’s interest is a top priority. It is difficult for a data company to use the opt-out model to be responsible for the interest of customers. Therefore, a new user data collection policy should be introduced to allow the responsibility of big data analytics reasonably to consider both interests of data companies and customers. This implies that new policies and redevelopment of existing data collection policies are necessary for responsible big data and big data analytics. This is the third strategy for making big data analytics responsible.

The fourth strategy for responsible big data analytics is that we should look at the responsibility chain for data processing flow-oriented big data analytics mentioned in Section 3. More specifically, every data processing flow stage corresponds to a big data analytics system. Every big data analytics corresponds to its responsibility. Different data processing flow stages and corresponding big data analytics have different and special responsibilities. Every stage of data processing flow needs to use responsible big data analytics because the responsibility chain results from the data processing flow-based big data analytics. Therefore, a responsibility chain should be developed for a data processing flow-oriented big data analytics system or platform [8].

Related Work and Discussion

This section discusses the related work on responsible big data analytics and their incorporation into e-business services, based on big data-driven small data analysis [26]. It also examines the limitations of this research.

Google Analytics as a data processing flow-oriented analytics, discussed in Section 3.2 is a contribution to the academic community and industry. The proposed model is revised from the four layers from bottom to top: 1. Foundation, 2. analyze, test, and grow, 3. Collect and clean, and 4. reporting, proposed by Penn CS, based on the lifecycle of data mining [27,44,24]. The proposed model for data processing flow-oriented Google analytics is also motivated by analytic flow and the petroleum and chemical industry [3]. Therefore, the proposed model is a unified model of integrating data processing and analytic flow [12]. The extended form of the proposed model is data processing flow-based big data analytics, presented in Section 3.3. Based on the work of Dignum, we developed a chain of responsibility that corresponds to a data processing flow-oriented big data analytics system or platform [8].

Google Scholar search for “responsible data analytics” found 41 results (retrieved on July 10, 2024). This implies that responsible data analytics is still an emerging discipline in academia although AI and business analytics have become frenzied topics [2,8,26]. Among these 41 results, for example, Leonard considers trust, fairness, transparency, and discrimination as the concerns for responsible data analytics [45]. Clarke proposes the principles and business processes for responsible AI [46].

Big data analytics for e-business services has drawn increasing attention in the computing, business, and e-commerce community [2]. For example, de Véricourt and Perakis consider the management of data analytics services as one of the frontiers in service science [47]. However, they have not detailed the issues of responsible data analytics services.

A limitation of this research is that it should provide a deeper investigation into the case of Meta and Google in terms of big data and big data analytics services to provide more rationales for responsible big data analytics.

Another limitation of this research is that it should provide more practical strategies for applying responsible big data and big data analytics to responsible e-business services because responsible big data and big data analytics for enhancing responsible e-business services is the real expectation from most customers. We will delve into this challenge through theoretical investigation and experiential study as the future work.

Conclusion

This article examined responsible big data analytics for enhancing e-business services. It looked at why responsibility matters to big data analytics and e-business services. It reviewed big data analytics and looked at Google analytics. It examined responsible big data and big data analytics based on two case studies. It also looked at big data analytics services as an e-business service and examines eSMCAS services as e-business services. This article also discussed the interrelationship between responsible big data analytics and e-business services. This research has at least three main contributions: 1. It presented a data processing flow approach to big data analytics; 2. It revealed the relationships between e-business services, e-SMACS services, and digital services and presented the six characteristics of digital services; 3. It proposed four strategies for applying responsible big data analytics to enhance responsible e-business services. The proposed approach in this research might have theoretical and technological implications for business analytics, big data analytics, e-business services, digital services, and e-society.

In future work, we survey information technology managers to collect data concerning the acceptability of responsible big data analytics for e-business services. Based on this planned survey analysis, we will further investigate how responsible big data analytics could be used for empowering responsible e-business services. We will also delve into the responsibility chain for a data processing flow-oriented big data analytics platform with intelligent process technology.

Acknowledgement: Dr Narasimha Rao Vajjhala is appreciated for his reviews and comments on this article.

This research is supported partially by the Papua New Guinea Science and Technology Secretariat (PNGSTS) under the project grant No. 1-3962 PNGSTS.

References

- Schneider G (2017) Electronic Commerce. Boston, MA, USA: Cengage Learning <https://erkashif.wordpress.com/wp-content/uploads/2015/12/ecom-schneider.pdf>.
- Sun Z (2024) Data Analytics for Business Intelligence: A Multi-industry Approach. Florida, USA: CRC press | Routledge | Taylor & Francis Group <https://www.routledge.com/Data-Analytics-for-Business-Intelligence-A-Multi-Industry-Approach/Sun/p/book/9781032585321>.
- Richardson J, Schlegel K, Sallam R, Kronz A, Sun J (2021) Magic Quadrant for Analytics and Business Intelligence Platforms. Gartner <https://www.gartner.com/doc/reprints?id=1-254T1IQX&ct=210202&st=sb>.
- Holsapple C, Lee-Postb A, Pakath R (2014) A unified foundation for business analytics. Decision Support Systems 64: 130-141.
- James S, Duncan AD (2024) Over 100 Data and Analytics Predictions Through 2028. Gartner <https://www.gartner.com/en/doc/over-100-data-and-analytics-predictions-through-2028>.
- Sharda BC, Wright D (2018) Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation. IEEE Security & Privacy 16: 26-33.
- Four Trends Driving Near-Term Artificial Intelligence Innovation. Gartner <https://www.gartner.com/en/newsroom/press-releases/2021-09-07-gartner-identifies-four-trends-driving-near-term-artificial-intelligence-innovation>.
- Dignum V (2019) Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Cham, Switzerland: Springer <https://link.springer.com/book/10.1007/978-3-030-30371-6>.
- Sun Z, Stranieri A (2021) The Nature of Intelligent Analytics. In Z. Sun, Intelligent Analytics with Advanced Multi-industry Applications. Hershey: IGI-Global 1-22.
- Zook M, Barocas S, Boyd D (2017) Ten simple rules for responsible big data research. PLOS Computational Biology 13.
- Laudon KG, Laudon KC (2020) Management Information Systems: Managing the Digital Firm (16th Edition). Harlow, England: Pearson <https://www.amazon.in/MANAGEMENT-INFORMATION-SYSTEMS-MANAGING-DIGITAL/dp/0135191793>.
- Sun Z, Vajjhala NR (2021) Responsible Big Data Analytics for e-Business Services. The 5th Interl Conf on Big Data Research (ICBDR), Tokyo Japan.
- (2021) THE SDGS IN ACTION. United Nations Development Programme (UNDP) <https://www.undp.org/sustainable-development-goals>.
- (2024) What is XaaS? IBM <https://www.ibm.com/topics/xaas>.
- Fothergill BT, Knight W, Stahl BC, Ulmiche I (2019) Responsible Data Governance of Neuroscience Big Data. Frontiers in Neuroinformatics 13.
- Sun Z, Strang KD, Pambel F (2020) Privacy and security in the big data paradigm. Journal of Computer Information Systems 60: 146-155.
- Fattah A (2014) Going Beyond Data Science Toward an Analytics Ecosystem. IBM <https://www.ibmdatahub.com/blog/going-beyond-data-science-toward-analytics-ecosystem-part-3>.
- (2017) Top 10 Big Data Trends. Tableau <https://www.tableau.com/resource/top-10-big-data-trends-2017>.
- Kavenna J (2019) Shoshana Zuboff: Surveillance capitalism is an assault on human autonomy. The Guardian <https://www.theguardian.com/books/2019/oct/04/shoshana-zuboff-surveillance-capitalism-assault-human-autonomy-digital-privacy>.
- Hu M (2020) Cambridge Analytica's black box. Big Data & Society 7: 1-6.
- Sun Z, Sun L, Strang K (2018) Big Data Analytics Services for Enhancing Business Intelligence. Journal of Computer Information Systems 58: 162-169.
- Mikalef P, Boura M, Lekakos G, Krogstie J (2019) Big data analytics and firm performance: Findings from a mixed-method approach. Journal of Business Research 98: 261-276.
- Sharda R, Delen D, Turba E (2018) Business Intelligence and Analytics: Systems for Decision Support (10th Edition). Boston, MA: Pearson.
- Ghavami P (2020) Big Data Analytics Methods: Analytics Techniques in Data Mining, Deep Learning and Natural Language Processing (2nd edition). Boston/Berlin: de Gruyter <https://www.amazon.in/Big-Data-Analytics-Methods-Techniques/dp/1547417951>.
- (2020) Diagnostic Analytics. Gartner-diagnostic analytics <https://www.gartner.com/en/information-technology/glossary/diagnostic-analytics>.
- Sun Z, Huo Y (2021) The spectrum of big data analytics. Journal of Computer Information Systems <https://www.tandfonline.com/doi/abs/10.1080/08874417.2019.1571456>.
- Penn CS (2016) Understanding the Google Analytics Ecosystem. Christopherspenn <https://www.christopherspenn.com/2016/10/understanding-the-google-analytics-ecosystem/>.
- Howarth J (2023) 80+ Amazing IoT Statistics (2024-2030). Explodingtopics <https://explodingtopics.com/blog/iot-stats>.
- (1995) Cambridge International Dictionary of English. Bath, UK: Bath Press. Cambridge.
- Wood M (2018) Cambridge Analytica, Facebook and the new data war. Marketplace <https://www.marketplace.org/2018/03/19/tech/cambridge-analytica-facebook-and-new-data-war>.
- Isaac M, Frenkel S (2018) Facebook Security Breach Exposes Accounts of 50 Million Users. The New York Times <https://www.nytimes.com/2018/09/28/technology/facebook-hack-data-breach.html>.
- Sun Z, Wu Z (2021) A Strategic Perspective on Big Data Driven Socioeconomic Development. The 5th Interl Conf on Big Data Research (ICBDR), Tokyo Japan.
- Delena D, Demirkanb H (2013) Data, information and analytics as services Decision Support Systems 55: 359-363.
- Demirkan H, Delen D (2013) Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. Decision Support Systems 55: 412-421.

35. Coronel C, Morris S, Rob P (2020) Database Systems: Design, Implementation, and Management (14th edition). Boston: Course Technology, Cengage Learning https://books.google.co.in/books/about/Database_Systems.html?id=6EbczgEACAAJ&redir_esc=y.
36. (2018) Final Version of NIST Cloud Computing Definition Published. NIST <https://www.nist.gov/news-events/news/2011/10/final-version-nist-cloud-computing-definition-published>.
37. Buyya R, Broberg J, Goscinski AM (2010) Cloud Computing Principles and Paradigms. Hoboken: John Wiley & Sons, Inc http://dphoto.lecturer.pens.ac.id/lecture_notes/internet_of_things/CLOUD%20COMPUTING%20Principles%20and%20Paradigms.pdf.
38. Flinders M, Smalley I (2024) What is XaaS? IBM <https://www.ibm.com/topics/xaas>.
39. Sun Z (2022) Demand-driven Services. In L. Rubalcaba, RESER-Elgar Encyclopedia of Service Studies (p. In Press). Cheltenham, United Kingdom: Edward Elgar Publishing.
40. Bossmann J (2016) Top 9 ethical issues in artificial intelligence. World Economic Forum <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/>.
41. Bostrom N, Yudkowsky E (2018) The Ethics of Artificial Intelligence. Cambridge Handbook of Artificial Intelligence. New York: Cambridge University Press.
42. Turek M (2020) Explainable Artificial Intelligence (XAI). DARPA <https://www.darpa.mil/program/explainable-artificial-intelligence>.
43. Stoller M (2021) Facebook was born, lives and thrives in scandal. It's been lawless for years. The Guardian <https://www.theguardian.com/commentisfree/2021/oct/26/facebook-scandal-mark-zuckerberg-frances-haugen>.
44. Kantardzic M (2011) Data Mining: Concepts, Models, Methods, and Algorithms. Hoboken, NJ: Wiley & IEEE Press <https://ieeexplore.ieee.org/book/5265979>.
45. Leonard P (2017) Emerging Concerns for Responsible Data Analytics -Trust Fairness Transparency and Discrimination: Paper for the NSW Data Analytics Centre Showcase. Data Synergies https://commsalliance.com.au/__data/assets/pdf_file/0018/58104/Peter-Leonard-Emerging-Concerns-for-Responsible-Data-Analytics_-Trust-Fairness-Transparency-and-Discrimination-Paper-for-the-NSW-Data-Analytics-Centre-Showcase-12-Jul.pdf.
46. Clarke R (2019) Principles and business processes for responsible AI. Computer Law & Security Review 35: 410-422.
47. de Véricourt F, Perakis G (2020) Frontiers in Service Science: The Management of Data Analytics Services: New Challenges and Future Directions. Service Science 12.

Copyright: ©2024 Zhaohao Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.