

Article

Big Data and Their Social Impact: Preliminary Study

Miltiades D. Lytras ^{1,2,*}  and Anna Visvizi ^{1,3} 

¹ School of Business & Economics, Deree College, The American College of Greece, 153-42 Athens, Greece; avisvizi@acg.edu

² Effat College of Engineering, Effat University, Jeddah P.O. Box 34689, Saudi Arabia

³ Effat College of Business, Effat University, Jeddah P.O. Box 34689, Saudi Arabia

* Correspondence: mlytras@acg.edu; Tel.: +30-210-600-9800

Received: 26 July 2019; Accepted: 10 September 2019; Published: 17 September 2019



Abstract: Big data is the buzz-word of today, and yet their specific impact on individuals and societies remains assumed rather than fully understood. Clearly, big data and their use have already given rise to a number of questions, including those of how data can be collected and used in ethical and socially sensitive ways. Building on these points, the objective of this study was to explore how precisely big data and big data based services influence individuals and societies. This paper elaborates on individuals' perceptions of data, especially on how they perceive the actual sharing of their data. In this way, this paper defines a value space for the social impact of big data relevant to three factors, namely the intention to share personal data, individual's concerns, and social impact of big data. The main contribution of this study consists of the insights into the still nascent area of research that unfolds at the cross-section of social science and computer science. We expect that in the next years this area of research will gain prominence.

Keywords: social impact; big data research; information systems; analytics; decision making; social sciences

1. Introduction

Recent developments in data-driven information systems, set big data research and business analytics at the core computer science and social science. In computer science research, there is a consensus that big data and data analytics research will foster a new generation of information systems capable of managing the collective wisdom in human decision making and smart machines [1]. Emerging research areas like cognitive computing [2] combined with artificial intelligence and machine learning, permit advanced and sophisticated methods for processing data, including sentiment analysis, image processing, natural speech recognition and text mining. In parallel emerging technologies, including cloud computing, internet of things and virtual reality, the value proposition of application and services that process data in different formats such as text, images, videos, microcontents in social media is further enhanced [3,4]. The development of a huge data ecosystem around the globe, in which providers and users of data promote business value in terms of data and decision making, is a key development of our times. In this context, users of applications and services worldwide participate consciously, or unintendedly, to an integrated data dissemination and aggregation process with critical trust and privacy issues.

A great discussion on the real impact of big data research has been initiated. An interesting study [5] sets the significance of the human decision maker at the center of any type of big data information processing cycle. There is an agreement between different academics that big data can make a big impact [1,6].

Big data research is aligned with the evolution in emerging information technologies research. New information processing paradigms further promote the significance of big data, and have a great

impact on its volume and coverage. A number of application domains and industries already adopt big data research with significant success. Consider social networks research and the contribution of social networks to the big data ecosystem [2,3]. Other examples are artificial intelligence and machine learning applications in various domains, such as customers/clients of big data repositories for personalized and targeted services [3,4].

In the recent literature of big data research, an increasing section is dedicated to the capacity of big data to support social sciences research. There is the anticipation that big data is potentially a social good that must be secured and be used for the transparency of services, and for the evolution of a user-centric new culture for sustainable computing. In parallel, several concerns have been documented, mostly related to trust, privacy and the protection of personalities in the new technology-driven domain of services and applications. In Figure 1, below, we provide our initial framework for the investigation of the social impact of big data.

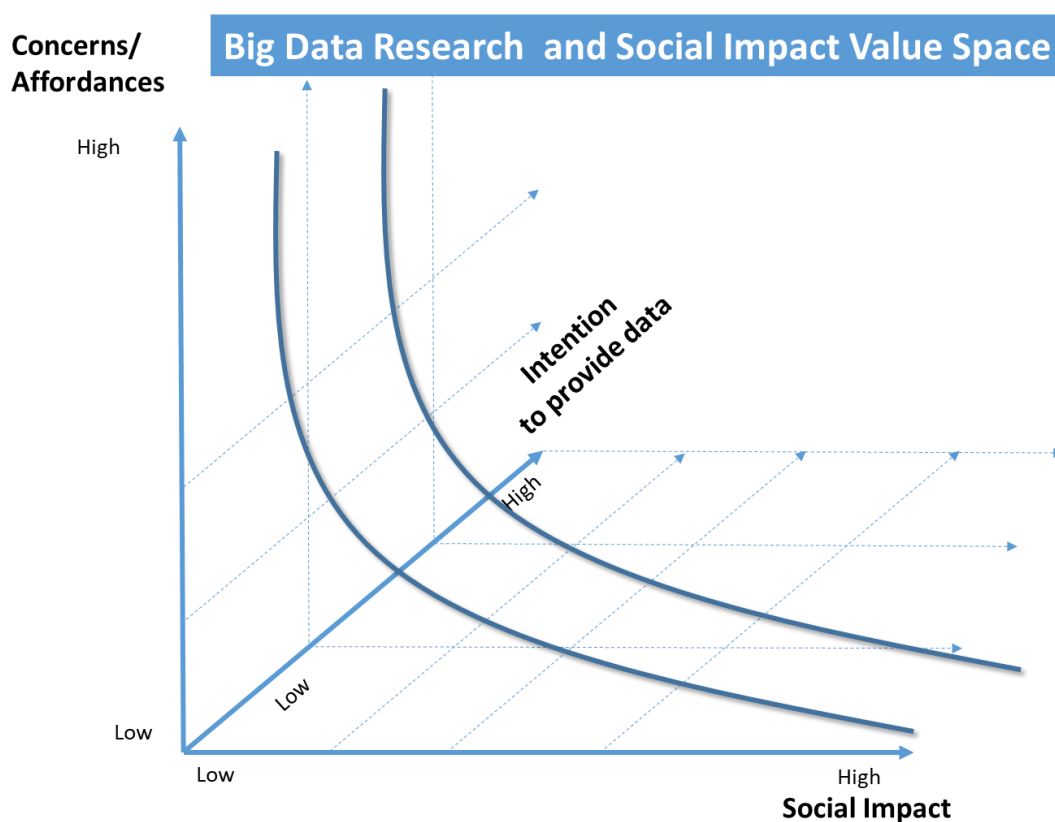


Figure 1. An initial framework for understanding the social impact of big data research. Source: The authors.

In our approach three critical factors need further investigation:

- User concerns/affordances: The first factor, namely user concerns or affordances is related to all the psychological, social, personal or professional concerns of users in relevance to the use of applications and services that generate and share personal data or other kinds of data from individuals within the big data ecosystem
- Intention to share data/informed consent: This aspect of our research problem is related to the conscious agreement or the intrinsic motivation of users to share their data for the purposes of big data application. In our research, we are interested in the connection of social challenges and social problems to the intention of users to share their data. Furthermore, we wanted to understand if in some scenarios, users of applications and services share their data without formal agreement due to their interest in exploiting the added value of the service for themselves or for the society.

- **Social impact of big data:** The third critical factor also determines the value space in Figure 1. The measurement of the social impact of big data seems to require interdisciplinary approaches and metrics, thus we must deploy heuristics for the attachment of value contribution to the perception of users for the impact of big data research to their lives and to our society. This is the ultimate objective of our research, nevertheless, the requirements and the various research strategies we deployed exceed the length and scope of this paper.

The value space of big data is defined as an aggregation of three factors/forces. At the X-axis, the social impact of big data is presented in a spectrum of low to high value. Users of big data applications develop perceptions and have their own interpretation mechanisms for the impact of big data. On the Y-axis, concerns and fears of users of big data application develop an intrinsic motivation mechanism for the use of such application. They deploy different ways for the use of big data applications and they also express their concerns for various aspects of these applications. Furthermore, on the Z-axis in Figure 1, it is shown that users also execute a different degree of willingness to share their data, for the proper functioning of big data applications. Various studies in the literature mention these factors, and our previous research has tried to investigate these factors. The value space that is defined by these three axes, can be used as a model for discussing big data applications and services and for mapping such services in wider contexts e.g., smart cities research. From a practical point of view, this model can also be exploited by real users of big data applications for the customization of available services or the personalization for added value of such applications. Also, from a policy making view, such a model can guide public consultation and debate on how we protect the data and identity rights of citizens against big data applications without compromise of social value and impact.

Figure 1 is used as a metaphor to communicate the overall idea of our research, that somehow users, with their perceptions and intention to use big data applications, define their personal value space and maybe also a societal value space. We understand that in our approach some key assumptions are integrated. We do, however, believe that it is worthy to investigate this research problem which has many psychological and social aspects. In the next section we provide a critical review of the relevant literature towards the justification of our research model that will be presented in Section 3 of this research study.

2. Literature Review—Understanding the Debate on Big Data and their Social Impact

The agenda of big data research is quite wide and involved various multidisciplinary communities. From a computer science and information systems perspective issues related to standardization, data mining, aggregation of data, interoperability and recommendation systems are at the top of research priorities. From a social science perspective, data as a social construct affecting issues related to identity management, personality, privacy and security are the focus of social research. Furthermore, the concept of the digital self, that combines personal, professional, social, and other features of individuals is gaining more interest [1]. In an evolving way, big data that refer to human entities and communities of people are established with convenient computational methods that permit social analysis and reference.

The connection of big data research to social sciences as well as the big impact of data-intensive applications and processing methods to societal challenges provides a very interesting research challenge. From the one side we have the social actors, humans, decision makers that both provide and consume data available in diverse, interconnected information systems [5]. The quest for impact on big data platforms and big data [6] requires a detailed study of different factors and accordingly new metrics like analytics or KPIs (key performance indicators) [6]. Humans, from this point of view, realize a critical mental shift in their behavior. From data providers they are requested to perform a decision maker role, within the boundaries and across hi-tech socio-technical structures like smart cities [7].

From a different angle, the big data ecosystem requires distribution and aggregation of information in modes that were unforeseen in the past. The sophistication and the huge capacity of big data

services to process significant volumes of data, automatically, without human intervention, sets critical questions related to privacy, security and data protection [8].

Especially in the context of social networks and social media [9], the information diffusion has exceeded any prediction. The ease of sharing information as well as the increased openness of such data warehouses permits advanced data processing that leads to critical insights about the data providers. In this situation, big data applications serve as intermediaries, matching the gap between the providers and the consumers of data, allowing several innovative business models to appear [10]. There is a connection that needs further investigation. The power of big data applications as intermediaries and as unique business models for adding value to raw data with data processing data, like sentiment analysis and opinion mining [11]. The capacity of new information processing methods to conclude about sentiments, attitudes or opinions is directly linked to some forms of social impact for such applications [12].

Within this complex big data ecosystem, individuals, organizations as well as governments need to develop frameworks to measure their readiness for the integration of big data research for measurable individual and social objectives [13,14]. One direction for the exploitation of big data research is analytics. The exploitation of value through huge volumes of data, requires the development of big data analytics capabilities [14,15], aiming to provide visualizations and summaries of data that can promote enhanced decision making. From a social science perspective, this connection directly leads to a new era of smart urbanism, where human actors, e.g., citizens, exploit processed data in meaningful visual forms for the improvement of the quality of their lives [16,17].

Another key aspect of big data literature is related to the big data hype. The utilization of big data research for business or social purposes must identify opportunities, myths as well as risks [18]. It is necessary for our societies and for policy making purposes to ask various provocative questions related to the ownership, supervision, consumption and protection of big data [19]. Consider, for example, a system for social rating based on microcontent contributions of citizens on social media, capable of measuring sentiments, political beliefs etc. Smart cities and smart government research [20] must take into consideration, a number of delicate issues related to privacy, security, safety and social responsibility of individuals and groups. Without a focus on sustainability [21,22], social inclusive economic growth and social justice, any isolated, monolithic big data application in the long term will unfortunately fail to promote its social impact. Novel approaches are required in the management of big data and their interoperability, as well as the annotation of data and services for improved social services [23,24]

What seems to be less analyzed, is the social dimension, and the social dynamics of big data, that refer to groups of people, businesses or social constructs. In both cases, the ultimate objective of big data research is to provide useful insights for the personalization of services and the targeting of value adding services.

A key challenge of big data research is to justify and to develop value reference layers to big data. The usability of big data, for various purposes and targeted markets needs to be clarified. In our research, our focus is on the social impact of big data. The key research question is related to the capacity of big data to have a social impact, and to enable bold solutions and responsive actions related to social problems.

In Figure 2, below, we organized in a simple way, some complementary aspects of the big data literature from an information systems/computer science perspective. The list of topics is definitely not exhaustive, but rather representative of significant aspects of the research.

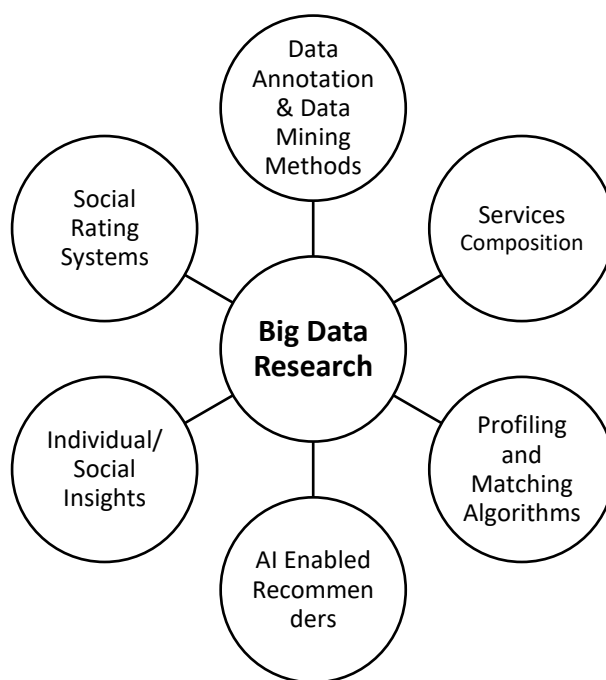


Figure 2. An overview of the key aspects of big data research literature for social sciences. Source: The authors.

Data annotation and packaging, as well as the emerging data mining methods such as sentiment analysis or social mining, set a new domain of research. Especially for social sciences, the capacity of methods like sentiment analysis to understand opinions, to analyze social behavior and the attitudes of individuals can be used extensively for sophisticated social sciences research.

Dynamic big data service composition and selection is also another very interesting area of research and literature. Big data, per se, have limited value without some services (clients) that consume these data for well-defined purposes. The design of socially aware value adding services that consume big data, will soon be a key trend in social sciences. From this perspective, we are going to realize a convergence of social sciences and computer science. Consider real time platforms that provide analysis and collective intelligence, over social media micro-contents (e.g., analysis of sexual harassment, bullying, anti-terrorism detection etc.).

Advanced user profiling [1,5] is also critical for the launch and management of social sensitive applications powered by big data research. The standardization of profiles is the first step toward interoperability of applications and social services. To this direction, latest developments in computer science as well as in policy-awareness frameworks, provide significant contributions. From a social impact perspective, a key question is how, within governmental institutions, and regulations, can we envision trustable, participatory and democratic platforms that exploit big data profiles for social good. Social rating systems, or social filtering platforms are key examples for this emerging area of research. Furthermore, from a social perspective, another key concern is about the ownership of the big data. However, this topic goes beyond the scope of our research articulated in this research study. Smart cities research is another example of critical integration for social sciences and computer sciences research [6–8]. In all these cases several research questions link big data research to critical social impact [11–17].

The availability of big data ecosystems offers numerous options for sophisticated services [6–8]. With the evolution of artificial intelligence e.g., machine learning approaches, we can have systems that are trained by the availability of big data. For special social problems, like poverty, exclusion, migration, we have a brand-new era of information services and recommender systems.

All the previous complementary aspects of research move the quest of our time forward for the provision of sophisticated social insights over individual's data or communities' data. Several ethical issues are involved, but it seems that the next thread of big data services and applications will materialize some of the aspects mentioned in this compact literature review.

In our approach the big data social impact research problem is part of a greater smart cities research approach [25–32]. In the next section, we provide our research methodology for this challenging research problem. From the beginning we have to communicate the limitations of our study and also the complexity of the research phenomenon. This study is based on a pilot survey, in which participants are academics and researchers from computer science and social sciences. The generalization of findings and conclusions should be analyzed within this context.

3. Research Methodology

In our research study we emphasize on an interpretive qualitative research, together with quantitative research. The understanding for the social impact of big data, is a very complex research problem because:

- The technical aspects of big data are complicated and continuously evolving.
- The social aspects of big data research involve human actors with complicated profiles and social references, and thus it is hard to generalize conclusions that are derived from a sample population.

In Figure 3 we provide a synopsis of the research model of our study. There are six critical research objectives:

- Research Objective 1. To understand the degree of awareness of big data research and the actual use of big data applications by our responders.
- Research Objective 2. To analyze, the perceived value of big data as it is realized by our responders.
- Research Objective 3. To understand the main concerns of users of big data services, as well as their affordances towards greater deployment of such applications.
- Research Objective 4. To clarify the impact of big data to individual's life and perceptions.
- Research Objective 5. To understand the determinants of social impact of big data in our societies in relevance to critical social problems and challenges.
- Research Objective 6. To develop guidelines for socially aware big data-enabled services.

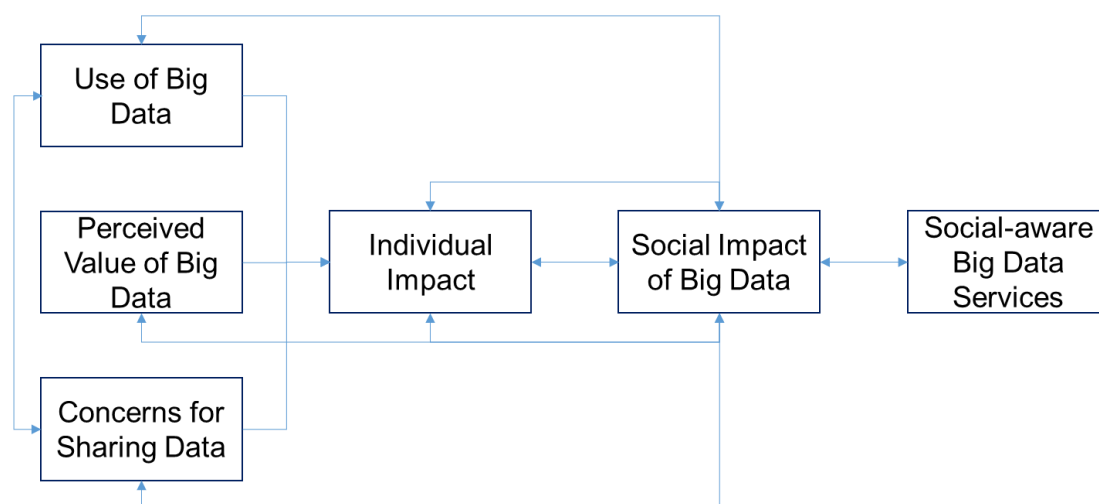


Figure 3. The research model for the study of social impact of big data. Source: The authors.

Two different directions of research have been deployed:

- a. A quantitative research based on a sample of responders for some initial insights related to research objectives 1–4. In the next section we present the results of this analysis. Briefly, we designed a research tool for the collection of responses based on a survey, and we collected 108 questionnaires in total from academics, students and researchers.
- b. A qualitative research design focused on targeted interviews with five heavy-users of big data applications and five non-users. The aim was to gain an in depth understanding of their mental models for the use and value of big data.

In this research paper we present the main findings of the first research direction, the quantitative analysis. From the beginning of our analysis we communicate the limitations of our study that refer to the limited sample and the complicated aspects of human behavior. We have completed the collection of the questionnaires and we are also currently designing the structured interviews for the heavy and light users of big data applications. Furthermore, we must state that this research on the social impact of big data is part of a greater research project related to smart cities research. In the latter, the objective is to explore end-users perceptions of smart cities' applications, especially as seen from the perspective of personal data sharing. We also understand that this is a complex research problem that needs further investigation and various contributions from different disciplines. Indeed, it is necessary to initiate a sound dialogue on these matters among researchers and decision-makers to create and explore synergies.

4. Analysis and Main Findings

Our effort to elaborate on the social impact of big data research is by default a very challenging research. The critical psychological and personal factors, affecting the adoption of big data applications goes far beyond traditional information systems or computer science research. From the other side, the technical sophistication of advanced applications and services also poses critical challenges to the protection of privacy, identity management and safety on the internet. From the beginning of this analysis, we have to declare the following facts:

- All the findings presented in this section refer to the “biased” population sample of our survey: Academics and researchers that are probably not the average users of big data applications.
- The generalization of findings and their interpretation must consider the previous fact.
- Beyond the previous two statements, the contribution of our study remains important: It is one of the first studies, that integrates social sciences and information systems research in the context of measuring the social impact of big data.
- We do not intend to discuss advanced statistics in this survey, but only descriptive statistics. Our intention is to provide interpretations of the main findings and to use these for the development of a global social impact of big data research study.

In this section we will present the main findings and their interpretation for the context of our research. As it was communicated in the previous section, we focus on the quantitative analysis of our survey. The presentation of key facts of our research will be accompanied by interpretations relevant to the key research objectives, clarified in Section 3. In Section 5, we provide an additional discussion. We start with an overview of the demographics of our survey.

4.1. Demographics

Our survey has a critical objective to understand the social aspects of big data and to interpret and measure the integration of social sciences and big data research.

The questionnaire used is available in Appendix A. We deployed SurveyMonkey software and we targeted users of applications, also familiar with the use of social networks. This survey is used as a pilot survey, since we are planning a global big data and social impact research to run in 2020. We circulated our survey to students and academics in universities of our scientific networks and we

received, within three months from July to October 2018, 108 responses. In Tables 1–3 we summarize the main demographics data from our study. In total 108 respondents; academics, researchers and students in management, international business, social sciences and information systems, in age clusters from 18 to 70 years old (Table 2). We admit that the sample is biased, since it is constituted by academics and researchers which are familiar with big data and analytics research. In the preliminary study, this special feature of responders was needed. We also have to declare that the generalization of the findings of this preliminary research study, should be done within the context of this limitation. Our purpose is to use the key findings of this research in order to populate a new research tool that will target broader clusters of users of big data applications.

Table 1. Age of responders.

Answer Choices.	Responses	Numerical
Under 18	0.00%	0
18–24	3.74%	4
25–34	18.69%	20
35–44	38.32%	41
45–54	26.17%	28
55–64	11.21%	12
65+	1.87%	2

Table 2. Discipline of responders.

Answer Choices	Responses	Numerical
Social Sciences	32.67%	33
Sciences (including Computer Science, Engineering)	67.33%	68
Total		101

Table 3. Awareness of big data concept by responders.

Answer Choices	Responses	Numerical
Yes	96.26%	103
No	3.74%	4
Total		107

Two thirds of the respondents were from a science discipline and one third from the social sciences. The balance is achieved since some respondents have joint expertise. Most of them were junior or experienced researchers in domains related to management, international business, social sciences and information systems. In Table 2 we summarize the relevant information.

Another key characteristic of our sample is that the vast majority of participants expressed their awareness about big data. Almost 97% of the responders claimed that they were aware of the big data phenomenon. This is important, also as a finding of our research, because big data for several communities is considered as an information system, or computer science research domain, but it seems that also social scientists are quite aware of it (Table 3).

In the next section of our survey, and in our relevant research model, we are interested in understanding the exposure of our responders to big data applications as well as their concerns and perceived value. A key motivation in our research is to understand the degree to which users of big data applications have concerns or feel ambiguity or danger in terms of trust or privacy.

4.2. Use of Big Data Applications and Perceived Value

The transparency of big data applications and their ubiquitous nature means that several times users of applications, or services that deploy big data, use them as black boxes. They do not care about the computational aspects of the information model of the application, but rather they want to enjoy the service. In our survey we discovered that 55% of applications users, have a deep knowledge of the big data applications that they use, that collect and aggregate data, which are stored in a distant server (Table 4). Also 45% of participants said that they are not users of big data applications, which also proves that currently, there are many people that do not intend to use advanced big data services. The question for sure is, how many of them dislike the use of smart cities services because they're afraid of the violation of their privacy or for other kinds of personal concerns. As a next question, we also focus on this.

Table 4. Do responders use big data services?

Answer Choices	Responses	Numerical
Yes	55.24%	58
No	44.76%	47
Total		105

Given the key findings that most of the participants of our survey are aware of big data applications and phenomenon, and also that more than half of the respondents are using big data applications extensively in their lives for various purposes, it is quite challenging to investigate, the perceived value they attach to the use of these applications and if it is relevant to their “quality of life” or just a “contribution” to their expectations and perceptions. While it is hard, even from a scientific way, or a statistical “correct” approach, to measure this value, we deployed a heuristic rule/approach. We asked our participants on a scale from 0–100 to attach a “numerical value” to the value of big data in their lives.

The result is summarized in Table 5. The average rating of all the responders for the value of big data in their lives is 66 out of 100. This numerical value seems to be overall “positive” in the sense that most responders attach a value greater than the average in the spectrum of low (0) to high (100) values. It is also evident from this value that responders seem to be skeptical about some aspects of big data. Thus, we need to understand the main concerns of users, specifically the features of big data that make them worried or concerned.

Table 5. Value of big data in the lives of the responders.

Answer Choices	Average Number	Total Number
	66	6.964
Total Respondents		106

Table 6, below, shows one of the most interesting findings of our survey, and deserves a more detailed analysis. Given the overall, rather high score (66 out of 100), for the perceived value of big data as provided by our respondents, we tried to understand some qualitative aspects of this positive effect.

Table 6. How responders perceive the value of big data. Personal ratings on the following statements.

	1 (Totally Disagree)	2 (Rather Disagree)	3 (Neutral)	4 (Somehow Agree)	5 (Fully Agree)	Weighted Average
Big Data allow me to enjoy personalized services	2.83%	2.83%	21.7%	49.06%	23.58%	3.88
Big Data offer me access to unique services	0.94%	3.77%	29.25%	47.17%	18.87%	3.79
Big Data enabled services, save time for me and effort	1.90%	2.86%	20.00%	51.43%	23.81%	3.92
Big Data is about social security	13.46%	29.81%	30.77%	17.31%	8.65%	2.78
Big data protect my privacy	27.18%	41.75%	25.24%	4.85%	0.97%	2.11
Big Data promote the collective intelligence and this has an impact in my daily life	3.81%	2.86%	18.10%	53.33%	21.90%	3.87
Big Data promote interoperability of services worldwide—thus i enjoy integrated services	1.90%	6.67%	25.71%	48.57%	17.14%	3.72

From the answers of our sample, we found that one of the key features of *big data applications*, is that they save time and effort for their users. Thus, developers of big data applications or designers of smart cities services must know that users would be happier to use their applications if they realized that they would be saving time. The next most important value components, according to our findings are the interoperability that big data applications offer to users, irrelevant of country or place. Users like to enjoy the same services, worldwide, with the same quality and transparency. This is also one more extremely interesting finding. If users, want to enjoy mobility, and in parallel to have access to the same services, then the big data research community and industry must promote this ecosystem of services worldwide. From the other side, this request and wish of users, needs to comply with several local and/or global policy making requirements.

Uniqueness of services is another critical value component of big data applications. Users understand that several services that are big data enabled, are unique, so they are somehow happy to use them. From this perspective, another key characteristic of big data applications is their innovative nature. Users are happy to use innovative services, that save them time in their lives, and can be enjoyed locally or globally with the same, high quality standard. This also means that big data industry should be always in a progressive, evolutionary process for the launch of novel services and innovations to the market.

Personalization, also seems to be valued by our responders. Users understand that most big data applications exploit their data for the enhancement of their personal experience. For this finding, there is also a side effect. Users recognize that big data applications challenge the protection of their personal data. Somehow a compromise between user experience and privacy is understood by our responders.

Another important finding of our survey is also that users recognize that big data applications somehow aggregate the collective intelligence of humans and potentially this could improve the quality of life. A direct interpretation of this finding is that some big data applications must certainly promote collective intelligence, but at the same time they must secure the trust, and the protection of privacy.

Finally, in Table 6, one more user perception is recorded. There is a rather neutral understanding that big data research can potentially promote also social security. With the bold debate on social media, fake news, fake profiles, social networks, analytics scandals, it seems that social security is a big theme for social security and currently users are not convinced for big data contribution.

These key findings permit several interpretations for further investigation in future studies:

- If citizens and users of big data application recognize that several big data applications save time and effort for them, then the next research question is what is the cost they are willing to pay

for their use, in terms of money or indirect costs, for example partial loss of their privacy, or agreement from their side to offer their personal data under specific conditions.

- Also, if users are interested in personalization of services, the next research question is which are the clusters that categorize different users to different clusters, and how happy would they be for such personalization if greater openness and access to their personal data is required.
- The finding related to the uniqueness of services is also critical. If users recognize that some big data are unique, then the next question is how can they resist to the necessity to share their personal data with such applications. For the big data industry, how easy it is to keep a limit to the penetration of sensitive personal data for the parametrization of their services to different users features?
- The anticipation of interoperability of big data services across cultures and nations also needs to be understood in a social context. Recent examples of cool applications like FaceApp, prove the capacity of big data applications to generate spontaneously huge data bases of critical personal data e.g., faces.

In the next section we present a third level of analysis for our survey with an emphasis on the potential social impact of big data. One of the key assumptions of our research model is the following: If citizens use big data applications and if they attach a positive value to their behavior, then the next step is to understand if the individual behavior also have some social contracts and positive implications for the society. In our approach this is defined as the social impact of big data.

4.3. Social Value of Big Data

In our modern, complicated social environment, in order to respond to social challenges, we have to understand them. In our survey, with its given limitations and limited reach, we made a first effort to record the key societal challenges of our times. Given the focus of our research on the social impact of the big data, the direct connection between these concepts is the capacity of big data applications to promote bold actions or solutions to societal challenges or problems. In Table 7, we summarize the outcome of our effort. According to our survey, the top five societal challenges of our time are:

- Security
- Socially inclusive economic growth
- Access to education/quality of education
- Equal opportunities for all
- Job opportunities

Table 7. The main societal challenges of our times based on the own beliefs of responders.

Answer Choices	Responses	Numerical
security	74.77%	80
social inclusive economic growth	48.60%	52
equal opportunities to all	42.99%	46
fair justice	34.58%	37
poverty	31.78%	34
depression	14.02%	15
education quality	54.21%	58
Job opportunities	39.25%	42
happiness	27.10%	29
Total		107

If we accept, as a working hypothesis that our responders reflect a greater population, then a key point for our future research is to analyze the social impact of big data research in terms of its capacity to promote sustainable goals related to the societal challenges reflected in Table 7. For example, a direct interpretation could be the following:

- Can big data applications enhance the capacity of people to have access to high quality open education? (Can we develop a big data learning platform to offer free, open, personalized training modules to individuals that will enhance their skills and competencies?)
- Can we design big data enabled, advanced, sophisticated services that promote the feeling of security in modern societies? (e.g., can we build big data enabled antiterrorist detection systems over social media?)
- Is there a way to exploit big data research in order to promote socially inclusive economic growth by defining, for example, new markets, or new data-intensive industries or innovations?
- Can we deploy big data research in order to promote new and better jobs in our societies? Is there, for example, any possibility to “measure talents”, to codify skills and competencies and to match job profiles with candidates etc.?
- Can we integrate big data research with sophisticated computational methods like artificial intelligence and machine learning in order to investigate “personalities” and personal habits that are linked to critical social challenges e.g., antiterrorism detection, harassment etc.?

We understand, and we admit that this is an extremely significant objective that goes beyond the scope and the depth of this limited survey. Nevertheless, it offers a very good starting point for further analysis and integration to a forthcoming greater research in terms of scope and coverage. It is also a good context for skepticism and interdisciplinary understanding. Somehow, big data research needs new contributions from the social sciences that have for years developed research tools and theories for understanding human behavior and personality. In our understanding in the context of the virtual world, internet and social media, there are many more things to be done in this direction.

In a next step, we tried to cross-check and to integrate the perception of our responders with one more question, related to the value and the concerns of users about big data use. Our intention, from a research point of view, is to build a theoretical framework to be tested in a future research about the connection between the social impact of big data, concerns and intention of use. This will be presented in Section 5 of our paper.

In Table 8, the findings reflect key aspects of users concerns about big data use and their associated added value.

Table 8. How responders perceive the value of big data.

	1 (Totally Disagree)	2 (Rather Disagree)	3 (Neutral)	4 (Somehow Agree)	5 (Fully Agree)	Weighted Average
Big data applications violate my privacy	3.74%	11.21%	26.17%	45.79%	13.08%	3.53
Big data services require advanced computer knowledge	0.93%	21.50%	12.15%	46.73%	18.69%	3.61
Big data services promote the gap between computer literated and non computer literated	0.94%	18.87%	25.47%	39.62%	15.09%	3.49
I somehow feel that my data are used for unknown purposes	1.87%	5.61%	8.41%	37.38%	46.73%	4.21
Big data services do not use transparent methods for processing my data	1.87%	6.54%	14.02%	41.12%	36.45%	4.04
I do not like that companies use my data for getting customer insights about me	0.93%	10.28%	19.63%	28.97%	40.19%	3.97

The main concern is related to uncertainty. Users of big data applications have a fear that their data will be used for unknown purposes. From a policy making point of view it is an absolute requirement that regulatory and legislative frameworks provide protection to users. The feeling of our responders is that currently there is a significant gap in this area. In close relevance to this finding, users also believe that big data applications do not use transparent methods for data processing. Some initiatives, like the General Data Protection Regulation (GDPR) are headed in the right direction. Users must have the right to be informed about who uses their data, for which purposes, and under which methods. One more bold finding of our survey is that our responders are also not happy that companies use big data research in order to gain better customer insights about them. This is also another huge theme for further research, on which we will elaborate further in the conclusion of our research study.

Some interpretations of these research findings include the following:

- What is the role of governmental authorities and supervising bodies towards the design, implementation and well-functioning of big data awareness policies related to privacy and data protection?
- How can users and citizens have an increased awareness about the processing methods of their data. Is it possible for them to have access to an IT-service where all the “users” of their personal data appear and are analyzed further?
- Concerning the social impact of the big data research the key question is, is there a fair-justice approach in which social bodies or organizations can supervise and rate “behavioral” oriented big data of individuals or groups. Answers to these questions are not obvious. They need significant social agreement and consultation.
- Concerning the compromise of sharing personal data for the use of unique big data applications, another critical question is how can users resist in using unique services without letting third party organizations gain significant insights into their personalities?
- Also, if we promote it as socially-fair to offer supervising organizations access to personal data, then does the individual level of decision refer to the degree of declining such services? We have to investigate a rather increasing population of users that deny to use big data applications due to this fear.

The respondents answers to our questions related to their reluctance to offer customer insights to companies about themselves guided the next part of our survey. We need to understand how people and users of big data applications interpret the new trend in informatics and social computing about analytics research. For this purpose, we attached 3 more questions that are summarized in Tables 9–11.

Table 9. Are responders aware about the concept of data analytics?

Answer Choices	Responses	Numerical
Not at all familiar	1.87%	2
Not so familiar	9.35%	10
Somewhat familiar	30.84%	33
Very familiar	39.25%	42
Extremely familiar	18.69%	20
Total		107

Nine out of ten respondents in our sample are familiar with the analytics concept and consider analytics to be potentially beneficial for users. In a similar approach, as we did for big data research, our responders attach a value of 76 out of 100 to the value of analytics. Given the limitation of our numerical approach to value measuring, this is an indication that users consider analytics to be of greater value by ten units, than big data (76 versus 66 out of 100). This indicates that users indirectly attach and associate increased value to advance decision-making capabilities.

Table 10. What responders think about data analytics.

Answer Choices	Responses	Numerical
Not at all important	0%	0
Not so important	0.93%	1
Somewhat important	16.82%	18
Very important	50.47%	54
Extremely important	31.76%	34
Total		107

Table 11. How responders rate the value of data analytics for social purposes (related to the capacity of data analytics to provide insights to social problems or challenges).

Answer Choices	Average Number	Total Number
	76	8.172
Total Respondents		107

Given the rather high value attached by our responders, to the impact of data analytics research for social purposes, we asked our participants to clarify the key aspects of analytics research that have increased social impact. In Table 12, we present these responses. The most interesting finding is that our sample states that analytics enhance social aware services. According to our responders, the analysis and organized presentation of analytics (e.g., key performance indicators, or visual overviews of big data or advanced data mining methods) can lead significant socially aware responses to social problems. An interpretation of this finding is that designers of IT/IS services and social scientists have to collaborate to deliver fully functional big data and analytics platforms for social issues and problems.

Table 12. How responders perceive the value of data analytics for social impact.

	1 (Totally Disagree)	2 (Rather Disagree)	3 (Neutral)	4 (Somehow Agree)	5 (Fully Agree)	Weighted Average
Data Analytics provide usefull insights for understanding critical social problems/challenges	0.00%	1.89%	10.38%	58.49%	29.25%	4.15
Data Analytics techniques enable social-aware services	0.00%	2.83%	14.15%	59.43%	23.58%	4.04
Data Analytics promote solutions to social problems	0.95%	4.76%	28.57%	53.33%	12.38%	3.71
Data Analytics enhance social responsibility	1.89%	7.55%	41.51%	38.68%	10.38%	3.48
Data Analytics support civic engagement	0.00%	7.55%	40.57%	42.45%	9.43%	3.54
Data Analytics help the recognition and monitoring of new social problems	0.94%	3.77%	14.15%	55.66%	25.47%	4.01
Data analytics promote Social coherence	0.00%	10.38%	43.40%	37.74%	8.49%	3.44

In close relevance, our survey also concludes that based on the responses of participants, analytics can also help to clarify novel social problems and issues, not easily diagnosed with other methods. This finding can initiate, for sure, a social dialogue and a policy-driven open participatory procedure.

While our responders recognize the potential social impact of analytics research, more than one third (36%) of them still state that they are still not ready for sharing their data for social aware services

(Table 13). This proves that there is a great distance to be covered until a new era of social aware services commences. From the other side, the combined big data and analytics research and integration for social purposes increases the perceived value to 70 points (Table 14). Finally, 98% of the responders are confident that in the next few years social aware big data services will be a key trend (Table 15). This proves that most users nowadays feel that somehow their data are processed for such purposes and that soon they will be forced, or they will be happier to use such services.

Table 13. Would respondents be happy to share their data in order to promote social aware big data services?

Answer Choices	Responses	Numerical
Yes	63.21%	67
No	36.79%	39
Total		106

Table 14. Responders rating of the impact of big data and analytics research to society (negative to positive).

Answer Choices	Average Number	Total Number
	70	7.360
Total Respondents		105

Table 15. Do respondents think that big data driven services for social impact will be a trend in the near future?

Answer Choices	Responses	Numerical
Yes	96.11%	104
No	1.89%	2
Total		106

5. Key Implications and Future Research Directions

The research problem under study, is complicated. The contribution of our research, within the given limitations, is significant: It provides interesting insights for the perception of big data users for the added value of these services to their lives and the potential social impact. We summarize, in the next paragraphs, the key response of our study to the research objectives we set out to address. We also elaborate on future research directions and policy making implications related to sustainability.

Research objective 1: To understand the degree of awareness of big data research and the actual use of big data applications by our responders.

- Key finding 1: Based on the qualitative analysis and interpretation of respondents' answers, we conclude that most responders, irrelevant if they are from a sciences or social sciences background, are aware of big data research. Of them, 55% also consciously use big data services due to unique features.
- Further research direction 1: It is interesting to investigate the people that do not use big data services, and under which conditions they would use big data applications for social purposes.
- Policy making and sustainability implication: The development of a policy aware regulation framework to deal with affordances and concerns of users of big data services.

Research objective 2: To analyze, the perceived value of big data as it is realized by our responders.

- Key finding 2: The majority of responders recognize the added value of big data services to their lives. They understand and value that applications save them time, and improve their life in some

respects. Furthermore, they recognize that most big data services offer unique services that can be enjoyed with the compromise of sharing personal data. Additionally, they feel that somehow big data applications aggregate collective wisdom and thus they improve the quality of their lives.

- Further research direction 2: The perceptions of individuals about the added value of big data services and big data research needs to be codified. An ontological approach to value components of big data should be promoted further. This will lead to a new measurement approach capable of codifying various aspects of value related to personality, ethics, social responsibility, social justice, open democracy etc.
- Policy making and sustainability implication 2: The evolution of big data research, from a policy making perspective needs to set all the required regulations about trust, transparency, security, anonymity as well as to protect the digital self of individuals. From a social perspective, it is necessary to integrate social insights in a well-defined context for socially inclusive economic and social growth. We need policies that will enhance the reusability and exploitation of big data for social purposes.

Research objective 3: To understand the main concerns of users of big data services, in regards to their use, as well as their affordances towards greater deployment of such applications.

- Key finding 3: Most of the responders recognize that there are three critical obstacles to big data applications: The unknown information processing methods of their data, and their purposes, as well as the fact that most of the time they consider that the ultimate destination of big data research is to gain more customer insights about them.
- Further research direction 3: Given the high degree of concerns of individuals about the use of their data, it is required to understand, through further research, the compromise between the shift of their willingness to share their data and to enjoy better services that improve the quality of their lives.
- Policy making and sustainability implication: One of the most significant directions for the future would be to develop policies that will protect individuals from being the victims of extreme third-party insights about their habits, preferences and social behavior. In a next step, there must be a sustainability regulation, through which all big data services should comply with general requirements and this must also be included in inform consent statements of user agreements.

Research objective 4: To clarify the impact of big data in the lives of individuals and to their perceptions.

- Key finding 4: The overall understanding of this survey is, that users have a positive impression of the value of big data research in their lives, and a rather positive perception. There is, however, a critical mass of users that are not willing to use big data services due to their concerns and fear that their data will be used for unknown purposes. The social impact of big data is also perceived as a potential under specific requirements.
- Further research direction 4: The measurement of the individual and social impact of big data will require advanced heuristic methods, and a lot of support is needed from the social sciences to adopt psychological, personality, mental, cognitive and other models to incorporate metrics and key performance indicators. It is also important to adopt studies and social research about the impact of big data applications in addiction or other psychometric directions.
- Policy making and sustainability implication: We need to develop more advanced social policies and regulations for the interconnection of information and communication technology with social sciences. We must recognize that all the modern big data applications are setting a huge socio-technical system where humans are actors and their interactions are more complicated than ever.

6. Conclusions

This paper calls for a revisit into the insights related to the use of big data research from individuals. The main point of departure is that users have an intrinsic motivation to protect their privacy and ownership of their data, but from the other side they are also happy to use unique services that improve the quality of their lives. In this required compromise, the contribution of our research is that it provides evidence from an empirical survey with objective interpretation of findings.

The research, research findings and discussion presented in this paper seek to initiate a debate on the diffusion and impact of big data on our lives. The ultimate contribution of this research work is the agreement of participants that the added value of big data towards responsive social aware services to critical social problems, is just the beginning of the journey and not its end. In the next few years this line of research will have a great impact on the development of a brand new research area related to social aware big data enabled social services and analytics.

Author Contributions: M.D.L. and A.V. contributed equally to the design, implementation, research and analysis of data and the delivery of the main findings. Both coauthors contributed equally to all phases of this research.

Acknowledgments: The authors would like to thank Effat University in Jeddah, Saudi Arabia for funding the research reported in this paper through the Research and Consultancy Institute.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. The Pilot Survey—Questionnaire

The research tool-questionnaire for our survey is available through this web link: <https://www.surveymonkey.com/r/LY2R3XV>.

References

1. Lytras, M.D.; Raghavan, V.; Damiani, E. Big-data and data analytics research: From metaphors to value space for collective wisdom in human decision making and smart machines. *Int. J. Semant. Web Inf. Syst.* **2017**, *13*, 1–10. [CrossRef]
2. Lytras, M.D.; Aljohani, N.R.; Hussain, A.; Luo, J.; Zhang, X.Z. Cognitive Computing Track Chairs' Welcome & Organization. In Proceedings of the Companion of the Web Conference, Lyon, France, 23–27 April 2018.
3. Lytras, M.D.; Mathkour, H.I.; Abdalla, H.; Al-Halabi, W.; Yanez-Marquez, C.; Siqueira, S.W.M. Enabling technologies and business infrastructures for next generation social media: Big-data, cloud computing, internet of things and virtual reality. *J. Univers. Comput. Sci.* **2015**, *21*, 1379–1384.
4. Lytras, M.D.; Mathkour, H.I.; Abdalla, H.; Al-Halabi, W.; Yanez-Marquez, C.; Siqueira, S.W.M. An emerging—Social and emerging computing enabled philosophical paradigm for collaborative learning systems: Toward high effective next generation learning systems for the knowledge society. *Comput. Hum. Behav.* **2015**, *5*, 557–561. [CrossRef]
5. Arnaboldi, M. The Missing Variable in Big Data for Social Sciences: The Decision-Maker. *Sustainability* **2018**, *10*, 3415. [CrossRef]
6. Wamba, S.F.; Akter, S.; Edwards, A.; Chopin, G.; Gnanzou, D. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *Int. J. Prod. Econ.* **2015**, *165*, 234–246. [CrossRef]
7. Calzada, I. (Smart) Citizens from Data Providers to Decision-Makers? The Case Study of Barcelona. *Sustainability* **2018**, *10*, 3252. [CrossRef]
8. Edwards, L. Privacy, Security and Data Protection in Smart Cities: A Critical EU Law Perspective. *Eur. Data Prot. Law Rev.* **2016**, *2*, 28–58. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2711290 (accessed on 17 August 2019). [CrossRef]
9. Guille, A.; Hacid, H.; Favre, C.; Zighed, D.A. Information diffusion in online social networks: A survey. *Sigmod. Rec.* **2013**, *42*, 17–28. [CrossRef]
10. Grabowicz, P.A.; Ramasco, J.J.; Moro, E.; Pujol, J.M.; Eguiluz, V.M. Social features of online networks: The strength of intermediary ties in online social media. *PLoS ONE* **2012**, *7*, e29358. [CrossRef]
11. Maynard, D.; Gossen, G.; Funk, A.; Fisichella, M. Should I care about your opinion? Detection of opinion interestingness and dynamics in social media. *Future Internet* **2014**, *6*, 457–481. [CrossRef]

12. Liu, B. *Sentiment Analysis and Opinion Mining*; Morgan & Claypool Publishers: San Rafael, CA, USA, 2012.
13. Olszak, C.M.; Mach-Król, M. A Conceptual Framework for Assessing an Organization's Readiness to Adopt Big Data. *Sustainability* **2018**, *10*, 3734. [CrossRef]
14. McAfee, A.; Brynjolfsson, E. Big Data: The Management Revolution. *Harv. Bus. Rev.* **2012**, *90*, 60–668. [PubMed]
15. Mikalef, P.; Pappas, I.O.; Krogstie, J.; Giannakos, M. Big data analytics capabilities: A systematic literature review and research agenda. *Inf. Syst. e-Bus. Manag* **2017**, 1–32. [CrossRef]
16. Kitchin, R. The Real-Time City? Big Data and Smart Urbanism. *GeoJournal* **2014**, *79*, 1–14. [CrossRef]
17. Kent, P.; Kulkarni, R.; Sglavo, U. Finding Big Value in Big Data: Unlocking the Power of High Performance Analytics. In *Big Data and Business Analytics*; Liebowitz, J., Ed.; CRC Press Taylor & Francis Group, LLC: Boca Raton, FL, USA, 2013; pp. 87–102. ISBN 9781466565784.
18. Davenport, T. Big Data at Work. In *Dispelling the Myths, Uncovering the Opportunities*; Harvard Business Review Press: Boston, MA, USA, 2014; ISBN 9781422168165.
19. Boyd, D.; Crawford, K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Inf. Commun. Soc.* **2012**, *15*, 662–679. [CrossRef]
20. Edelenbos, J.; Hirzalla, F.; van Zoonen, L.; van Dalen, J.; Bouma, G.; Slob, A.; Woestenburg, A. Governing the Complexity of Smart Data Cities: Setting a Research Agenda. In *Smart Technologies for Smart Governments. Public Administration and Information Technology*; Rodríguez Bolívar, M., Ed.; Springer: Cham, Switzerland, 2018; Volume 24, pp. 35–54. ISBN 978-3-319-58576-5.
21. Kharrazi, A.; Qin, H.; Zhang, Y. Urban big data and sustainable development goals: Challenges and opportunities. *Sustainability* **2016**, *8*, 1293. [CrossRef]
22. Wielki, J. The Opportunities and Challenges Connected with Implementation of the Big Data Concept. In *Advances in ICT for Business, Industry and Public Sector*; Mach-Król, M., Olszak, C.M., Pelech-Pilichowski, T., Eds.; Springer: Cham, Switzerland, 2015; pp. 171–189, ISBN 978-3-319-11327-2.
23. Spaletto, J. An Investigation of Strategies for Managing Exponential Data Growth in the Enterprise. *J. Leadersh. Org. Eff.* **2013**, *1*, 4–14.
24. Alkmanash, E.H.; Jussila, J.J.; Lytras, M.D.; Visvizi, A. Annotation of Smart Cities Twitter Microcontents for Enhanced Citizen's Engagement. *IEEE Access* **2019**, *7*, 116267–116276. [CrossRef]
25. Angelidou, M.; Psaltoglou, A.; Komninos, N.; Kakderi, C.; Tsarchopoulos, P.; Panori, A. Enhancing sustainable urban development through smart city applications. *J. Sci. Technol. Policy Manag.* **2017**. [CrossRef]
26. Bi, S.; Liu, Z.; Usman, K. The influence of online information on investing decisions of reward-based crowdfunding. *J. Bus. Res.* **2017**, *71*, 10–18. [CrossRef]
27. Yin, C.; Xiong, Z.; Chen, H.; Wang, J.; Cooper, D.; David, B. A literature survey on smart cities. *Sci. China Inf. Sci.* **2015**, *58*, 1–18. [CrossRef]
28. Visvizi, A.; Lytras, M.D. Editorial: Policy Making for Smart Cities: Innovation and Social Inclusive Economic Growth for Sustainability. *J. Sci. Technol. Policy Mak.* **2018**, *9*, 1–10.
29. Visvizi, A.; Lytras, M.D. Rescaling and refocusing smart cities research: From mega cities to smart villages. *J. Sci. Technol. Policy Mak.* **2018**. [CrossRef]
30. Van de Voorde, T.; Jacquet, W.; Canters, F. Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landsc. Urban Plan.* **2011**, *102*, 143–155. [CrossRef]
31. Carrasco-Sáez, J.L.; Careaga Butter, M.; Badilla-Quintana, M.G. The New Pyramid of Needs for the Digital Citizen: A Transition towards Smart Human Cities. *Sustainability* **2017**, *9*, 2258. [CrossRef]
32. Smart Cities Research: Contemporary Issues in Smart Cities Research. 2018. Available online: https://docs.google.com/forms/d/e/1FAIpQLSciGvQE78AWnUcHqt4Q_MLFC6FT52UQskZSSNIZzOBNQTDkA/viewform (accessed on 10 June 2018).

