

Effat University Repository

Crowdsourcing Research for Social Insights into Smart Cities Applications and Services

Item Type	Book chapter
Authors	Alhalabi, Wadee;Lytras, Miltiadis;Aljohani, Nada
Citation	Alhalabi, Wadee & Lytras, Miltiadis & Aljohani, Nada. (2021). Crowdsourcing Research for Social Insights into Smart Cities Applications and Services. Sustainability. 13. 7531. 10.3390/su13147531.
DOI	10.3390/su13147531
Download date	2025-08-06 18:00:09
Item License	https://creativecommons.org/licenses/by-nc-nd/4.0/
Link to Item	http://hdl.handle.net/20.500.14131/167

Article

Crowdsourcing Research for Social Insights into Smart Cities Applications and Services

Wadee Alhalabi ^{1,2}, Miltiadis Lytras ^{1,3,*} and Nada Aljohani ¹

¹ Department of Computer Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia; wsalhalabi@kau.edu.sa (W.A.); naljohani@effatuniversity.edu.sa (N.A.)

² Department of Computer Sciences, Dar Alhekma University, Jeddah 22246, Saudi Arabia

³ Effat College of Engineering, Effat University, Jeddah 21478, Saudi Arabia

* Correspondence: miltiadis.lytras@gmail.com or mlytras@acg.edu

Abstract: The evolution in knowledge management and crowdsourcing research provides new data-processing capabilities. The availability of both structured and unstructured open data formats offers unforeseen opportunities for analytics processing and advanced decision-making. However, social sciences research is facing advanced, complicated social challenges and problems. The focus of this study is to analyze the contribution of crowdsourcing techniques to the promotion of advanced social sciences research, exploiting open data available from the geographical positioning system (GPS) to analyze human behavior. In our study, we present the conceptual design of a device that, with the help of a global positioning system-data collection device (GPS-DCD), associates behavioral aspects of human life with place. The main contribution of this study is to integrate research in computer science and information systems with that in social science. The prototype system summarized in this work, proves the capacity of crowdsourcing and big data research to facilitate aggregation of microcontent related to human behavior toward improved quality of life and well-being in modern smart cities. Various ethical issues are also discussed to promote the scientific debate on this matter. Our study shows the capacity of emerging technologies to deal with social challenges. This kind of research will gain increased momentum in the future due to the availability of big data and new business models for social platforms.

Keywords: knowledge management; crowdsourcing; future smart cities research; cultural behaviors; moral values; individual behavior; GPS; data visualization; social impact; social sciences; social-aware platforms; social challenges; innovation



Citation: Alhalabi, W.; Lytras, M.; Aljohani, N. Crowdsourcing Research for Social Insights into Smart Cities Applications and Services. *Sustainability* **2021**, *13*, 7531. <https://doi.org/10.3390/su13147531>

Academic Editor: Jurgita Antuchevičienė

Received: 17 May 2021

Accepted: 2 July 2021

Published: 6 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The advent of the internet, social media and web-based applications provide a challenging environment for the design of applications and services that support human activity in modern rural and urban areas. The aggregation and mining of data for meaningful value-adding services is a key trend. Within the bioresearch area of knowledge management in recent years, crowdsourcing refers to the generation of data from individuals and its exploitation for user-centric services. In the context of smart cities research this is a new area of interest. The exploitation of structured and unstructured data to understand human profiles and offer personalization services has been well documented. On the other side, social sciences research on human behavior has started to exploit these techniques. Behavioral aspects of human personality, personality traits and opinion mining are just a few aspects of this emerging research field. In the intersection of computer science and social sciences, the design implementation and functioning of information systems for social impact raise critical ethical issues. The access to critical personal data, the prevention of anonymity, and privacy of data objects, the legislation framework for the protection of citizens from unwanted and unintended data processing of their critical data are only few examples of ethical issues. Furthermore, the storing and availability of crowdsourcing

data related to behavior require the design of procedures and protocols for the protection of human rights. Within this complicated and demanding environment and context, it is extremely important to conduct research that will bring light to various aspects of the phenomenon.

Our research study has a threefold objective: (i) to understand the feasibility and modern information and communication technologies to support the implementation of smart cities services that allow citizens to provide information about the behavior of other people towards better quality of life, increased security and safety; (ii) to set up a prototype for the further discussion of computing, social and ethical issues of crowdsourcing applications in the context of smart cities applications; and (iii) to elaborate on the recent development on big data, knowledge management and crowdsourcing technologies and their capacity and limitations for the provision of social impact. We must emphasize from the Introduction that our study presents empirical data and analysis related to a crowdsourcing prototype while the research phenomenon under study is extremely complicated and sophisticated.

Our focus is the exploitation of leading-edge, data-centric computational methods for the study of human behavior. There has been intense debate in recent years on the capacity of open-data approaches to provide insights into human behavior, and several ethical issues have been raised. Advanced data-mining methods, such as classification clustering, association rules and predictive models, together with methods such as text mining, sentiment analysis and opinion mining, have attracted the attention of both social scientists and information systems experts. The conclusion of studies on these methods and complementary approaches is that data science techniques can inform critical insights into human behavior but that several ethical issues need further consideration. In our study we focus on the human behavior and the social activity. Nowadays, citizens in modern cities use smartphone applications and location aware services to provide their reviews and experiences. Most of the time, the geographical positioning system allows users to provide microcontent on simple or more sophisticated applications and to attach their rating. In our effort we investigate the feasibility and impact of a service in which humans provide content related to the behavior of other people within a smart city.

In our study, we examine publicly available GPS traces to establish the capacity of techniques to update user profiles through processing these data. The interesting connection is analyzed between location-aware services and the traits, values and beliefs of human actors. Our approach is used as a test bed to investigate the efficiency of data-mining methods to process structured and unstructured data for the provision of value adding services within smart cities.

The user scenario that informs our research model is simple and ubiquitous today: humans are increasingly using mobile phones and portable devices and attach microcontents to location aware services.

A person's values are heavily influenced by their cultural background. Thus, to understand a culture we need to analyze its key values and, in this way, adapt any decisions on social traditions to its particularities. By having data on an ethical value, in-depth analysis can be undertaken, and its results can be used to inform decisions and actions that are beneficial to society. The idea of crowdsourcing systems that aggregate data related to human behavior or human activity has initiated an interesting scientific debate on the ethics of using personal data for social purposes, in other words, for the social good. The exploitation of big data for social good is another aspect of the same discussion. An individual's values are relevant to his or her cultural aspects, and many studies in recent decades have provided useful insights into certain societies. Our efforts are concentrated on the Kingdom of Saudi Arabia; the full research is funded by an institution dedicated to the study of moral values.

Various works have employed GPS technology in the study of behavior, health, environment, and so on; however, only a few have used it to analyze ethical values. This study aims to study ethical values in the context of the actual geo-politic environment and to offer useful insights that can be further used by key decision-making individuals. For

our research, we adopted a global positioning system-data collection device (GPS-DCD) that employs GPS technology to link the collected samples to their exact location on a world map.

This kind of research is not typical in data mining, yet it addresses the latest challenges in combined social and information systems research. The underlying logic is that humans, through their interaction with the real world, their mobility, and their actions, constantly update a data-rich environment that informs others about their behavior and preferences. In a way, we are trying to interpret GPS data with a focus on behavioral analysis.

Our research complies with data privacy and ethics regulations and serves as a test bed for experimentation with the ability of advanced data mining to offer insights into human behavior. Its special focus on ethical values relates to the funding organization and is in response to a request for a study of this kind. In accordance with the organization's policy on privacy in human behavior, the data were collected anonymously. All personal and privacy issues were protected. Users participated voluntarily and the study complied with the research ethics, with users' full consent. The GPS device was used on a voluntary basis by users without recording any personal details.

In order to communicate the justification of our research, we have to highlight the research context: (i) in modern cities, citizens are using mobile phones and they contribute microcontent related to their experience and rating; (ii) a special area of research related to computer and social science focusses on the analysis of human activity within smart cities; (iii) some aspects of human activity are more sensitive for the wellbeing, the quality of life, the security and the safety of the society; and (iv) the aggregation and the analysis of data related to human activity and behavior through information technologies are feasible but also raise critical ethical concerns related to privacy and security.

Thus, in our research, we present a prototype system that have four basic features. (i) It is a crowdsourcing platform in which volunteers, in their daily activity within their city, provide content related to their observations about human activity of others in an anonymous way. For example, a volunteer can use an application on their mobile phone and to report a red traffic violation in Jeddah without any reference to persons, only focusing on the event. Another example is that a volunteer can use the same application and add a case of charity activity to the system. (ii) The crowdsourcing application maintains a Datawarehouse with all the user-generated content about the human activity. For our prototype system, we focus on six different activities. We are using this as a test bed which can be extended in order to cover more human activities. (iii) The collected data are summarized in visual dashboards and maps allowing citizens to understand the intensity and the variety of human activities in their region or city. (iv) More advanced data-mining methods can be used in the future in order to perform more significant tasks, e.g., clustering of human activities or profiling of citizens that share their microcontent to the application.

We must also emphasize from this point that our approach has to face significant ethical challenges from a computer science point of view: (i) while modern crowdsourcing platforms permit a variety of services, it is critical to investigate their compliance with privacy and security regulations; (ii) the protection of human rights and personality must also be integrated to legislation frameworks for the provision of digital services in the smart cities context; (iii) in our research, we are focusing on user-generated content related to human behavior of others while also currently sophisticated artificial intelligence services and algorithms perform enormous data collection and categorization of human activities for social services.

The rest of our article is organized as follows. In Section 2, we summarize the state of the art of data mining in social sciences research and discuss the research problem. In Section 3, we discuss our research methodology and research design, and, in Section 4, we elaborate on the architecture and technical aspects of our unique value proposition for using data mining in GPS data analysis. Finally, in Section 5, we present the main findings and conclusions.

2. Literature Review—The State of the Art

The recent literature on big data and analytics research offers interesting insights into the integration of social sciences and computer sciences research. Handling big data in the future will involve several aspects of human behavior [1]. This trend to use sensitive behavioral data automatically raises various ethical concerns related to the depth and scope of artificial intelligence-based decision-making and information processing [2]. In our study, the widely available GPS traces and location-aware data are certainly accessible; the key question is whether we can exploit them to understand the cultural context of the human actors while respecting their personal data privacy and security.

In a recent study [3], the authors explain how to incorporate ethics in artificial intelligence. It is important for our exemplar study on the use of data mining for social sciences research to consider all ethical questions and concerns. Highly relevant to this direction of research is the incorporation of ethical aspects of the decision-making process in human and artificial agents, which has been investigated by various studies [4]. Thus, our research interest in associating values to data-driven analysis of GPS traces is an ethical and well-defined case study on the integration of social sciences and information systems research.

Another recent study [5] discusses the issues relating to the design of the ideal artificial ethical agent. In our research problem, the process of structuring and monitoring values concerning human behavior poses many interesting questions about an artificial intelligence agent able to impact on human actions and preferences. The current study does not expect to provide guidelines on creating such an agent but, rather, to prove the capacity of current data mining methods to promote advanced profiling of behavioral aspects of human life.

It is obvious that soon, numerous intelligent devices and agents will play a critical role in analyzing human behavior by using data that are available on the web or mobile phones. From this perspective, the “victimization” of such research is unscientific, and it goes against the value that the social impact could add [6]. Internet research and cognitive capabilities, as well as behavior analysis, will in the future be integrated to create smart machines with enhanced decision-making capabilities [7]. In the same way, advanced social network analysis of the data freely available from social media, to understand behavior, can address significant social challenges [8].

In our analysis, processing GPS data can certainly inform us about individuals’ preferences and traces and promote so-called crowd intelligence or anticipatory computing [9]. Our objective is to study GPS data to update the user profiles with behavioral aspects. The concept of ethical values has been analyzed from a cultural analysis perspective by various studies, value systems and aspects of behavior that are linked to moral development. For example, in a study by [10], researchers carried out an in-depth analysis of shaping society’s cultural values by considering its individuals’ values. It was proposed to capture cultural values by data acquired collectively on day-to-day life events, and it was observed that this is flexible enough for a wide range of situations. In another study [11], researchers observed that the cultural approach is best suited when there are cross-cultural systems and showed how an individual’s behavior can be measured; however, for such studies, other researchers [12] recommend the use of smartphones and smart devices among graduates and undergraduates.

The above works point to the efficacy of using GPS systems to collect data that reflect values. The GPS report in another study [13] substantiates that its use in data collection is highly advantageous. GPS ensures that high-quality data are gathered at low cost. Two recent studies [14,15] have directly relied on GPS to investigate various possibilities of system development and to better understand economics and policies.

Using a GPS technique, Lara et al. [16] carried out an investigation into whether psychometric values could affect a person’s psychology about selecting a travel route. People driving from a single location to a destination along a variety of routes were observed and examined by means of a questionnaire. The study filtered and stored the

data in its survey database every four hours, and the data were automatically generated by the central server once every night. The data feedback to the survey database helps us to understand the distribution of ethical values on a world map. The association of destinations or places that are visited with ethical value is of questionable accuracy, yet it serves as an initial test bed for analysis.

The incorporation of GPS in modern scientific devices is now widespread in many apps. The accuracy of global positioning can be improved by processing the given input data through analysis [17], and the error can be calculated using satellite data. Nowadays, thanks to the efforts of Google and allied members, data from Android smartphones is increasingly reliable and can be correlated with GPS tags to analyze various aspects. For example, some researchers utilize GPS for tagging purposes [18], but most focus on tracking environmental characteristics [19]. Experiments have been conducted to combine other data with GPS to achieve a better understanding of the problem. For example, GPS data, together with measurements of certain environmental traits, such as oxygen level, have been used to determine the effect on transport-related physical activity in a specific area [20].

GPS is nowadays used even in healthcare [21]. For example, one study [22] proposed a smart real-time system to monitor patients using GSM/GPS technology. The authors aimed to monitor patients' vital signs and provide the necessary aid to healthcare workers in an emergency. In similar health studies, GPS data and surveys were used to devise strategies to encourage daily mobility in the neighborhood [23]. The neighborhood effect on coronary heart disease was studied through joint modeling of neighborhood determinants of participation [23]. The studies [24] clearly emphasized a neighborhood effect on health; however, in children, this was not well identified [25,26].

The wide variety of studies on the exploitation of GPS-driven data and location-aware services is a good indication that the relevant data ecosystem can be used for advanced analytics and decision-making. The challenge for our research is the exploitation of GPS data for sophisticated visualizations capable of informing significant social problems.

The challenges of a broad range of neighborhood behaviors and health outcomes for the same population have been investigated by several studies [27,28]. The main issue in randomly sampling neighborhood data is the difficulty in identifying those participants residing in a particular neighborhood and in distinguishing the individual and household from the neighborhood [29,30]. Another factor is that GPS data can be quickly and readily obtained from daily multi-place activity [31].

The association between objective methods in the constructed setting and participants' recreational physical movement was examined by a study on adults in Brazil [32]. It was observed that a cross-sectional study design limits casual inference and that limitations are inherent in self-reported measures. To overcome issues related to these neglected qualitative attributes, the study's model included residents from various areas. Each residence was geocoded to permit analysis of its environment. Another study reported [19] an association between momentary ecological awareness and physical movement that was slightly stronger in children in smart communities than in those from rural, low-density communities.

Another study [33] provides insights into the activity of children on weekdays and at weekends. It was conducted using GPS technology to investigate 9- to 10-year-olds' activity in a project known as the Sport, Physical activity and Eating behavior and Environment Determinants in Young children (SPEEDY). While the children participated in active movement on both weekdays and at the weekend, greater efficiency was noted on weekdays. Differences in the locations used were reflected by changes in the duration of activity [34]. Observations during the summer vacation and half-term holidays using an accelerometer and Garmin's forerunner GPS unit over consecutive four days and two weekends were used to undertake a detailed statistical analysis of 100 young children's behavioral values. This revealed significant activity during their leisure time. The four categories of activity

at matched data points from the accelerometer are sedentary, light, moderate or vigorous, recorded closest to the GPS location in terms of the date and timestamps.

Bond et al. [35] present a culture-level analysis of social axioms in individuals from five groups. This classified society's values onto several levels, from top to bottom.

So, the GPS approach can help to predict values. The role of GPS in promoting greater participation among youths and non-minorities was studied by the Oregon Household Travel Survey [36]. This showed that it is an effective technique for reducing the level of non-response in youths. Another study [37] conducted a time analysis of a large-scale and multi-country survey that can be used to develop alternative mid-range approaches, with a cultural and national application and explanation. Yet another study [38] discussed individual behavior in various cultural group and business operations in terms of gender, sexuality, and religion. This emphasized the consequences of each cultural activity, based on values. The aim of a further study [16] using GPS data was to reduce the burden of response in handling diary reports. It mainly focused on the design and performance of the system's set up. GPS points were allowed to identify the correct starting time and the duration of activities, which was also the objective of mapping all the points. Route choice may be estimated by data setting that incorporates attitudes to the general parameters of public transport routes.

A study [39] involved retrieving information related to diary and telephone data in the United States to establish travel patterns. The study used hand-held devices and GPS to obtain participants' daily travel and, by incorporating GPS, was able to collect data on both the route and speed. A study [40] identified two types of behavioral values: terminal and instrumental. The first presents life achievements in various cultures, while the latter refers to the preferred mode of behavior or method of achieving these.

Globalization of business is a specialized topic that finds expression in various zones of the world by neglecting how broader cultural values affect the way in which we live, as suggested by [41]. Based on events in people's lives, the authors [42] describe some ethical values in Israel and Germany. Derived from people's universal requirements, the basic cognitive representation of interpersonal coordination includes group welfare, survival and meeting biological needs. The motivational domains of values are security, enjoyment, achievement, social power, pro-social, self-direction, maturity, and restrictive conformity. Various countries, such as Hong Kong, Finland, Spain, and the United States, provide data on human values, and demonstrate that, together, these basic and motivational domains are ethical values that are seen to be held by individuals in the smallest regions.

The ease of usage of GPS data has superseded traditional travel techniques. A conceptual study, conducted at the Georgia Institute of Technology, [43] used GPS to provide essential trip information in terms of speed and location [44]: an application programming interface (API) locates any nearby school and sounds an alarm to limit drivers' speed in such an area. The API is capable of merging various services. Moreover, GPS data can be incorporated into insurance telematics, as used in existing or new insurance pricing structures [45]. A high correlation with accident rates represents values such as habits regarding mileage, peak-time travel, braking or start time: the risk factor most vital for human survival is speed.

Additional literature [46–57] is summarized in Tables 1 and 2, below, emphasizing on the crowdsourcing and knowledge management aspects of our research. Several research studies, elaborate on the capacity of crowdsourcing applications to set up sustainable knowledge management mechanisms for sophisticated smart cities applications with significant social impact.

Table 1. Critical literature review on crowdsourcing knowledge management systems.

Literature Review			
Author(s)	Title of Article	Key Contribution	Impact on Our Research Model
[46] Putrenko, Viktor. (2017)	Data mining of relationship in crowdsourcing projects and social activities of citizens	Focus: Investigation of relationships between various social activity of users in a region and its mapping in crowdsourcing project OpenStreetMap (OSM)	We want to investigate the capacity of citizens to contribute open anonymous public data for a crowdsourcing-based platform for quality of life.
[47] Chang Dann, Lee Carman. (2018)	A product affective properties identification approach based on web mining in a crowdsourcing environment.	Focus: The use of crowdsourcing consumers data for the effective identification of properties.	We are interested on the design of a knowledge management platform aiming to codify with the use of specific taxonomies and hashtags of social interaction content.
[48] Hoi Calvin SH, Khawaja Daniyal, Leung Carson K. (2017)	Constrained Frequent Pattern Mining from Big Data Via Crowdsourcing.	Focus: The design of an algorithm that permits crowds to rate their interesting patterns. Utilization of big data applications and services data streams	The user's involvement and engagement in the filtering of information towards meaningful services that enhance quality of life, well-being, and security.
[49] Guo Kehua, Tang Yayuan, Zhang Peiyun. (2017)	Crowdsourcing semantic fusion for heterogeneous media big data on the internet of things.	Focus: Utilization of the collective wisdom of social users and deployment of crowdsourcing service	The design principles for a knowledge management system that will aggregate social aspects of behavior in a crowdsourcing platform.
[50] Lytras Miltiadis D, Visvizi Anna. (2020)	Big data research for social science and social impact.	Focus: An integrated compilation of research studies on the use of big data research in social sciences toward social impact.	We want to set up a pilot system to understand its social impact at a limited scale before using big data for the same research phenomenon.
[51] Yoon Ayoung, Copeland Andrea. (2019)	Understanding social impact of data on local communities.	Focus: Exploitation of community data for social impact	The understanding of the motivation and incentives of citizens to contribute to community crowdsourcing platforms.
[52] Carpenter, Christopher J; Amaravadi, Chandra S. (2019)	A big data approach to assessing the impact of social norms: Reporting one's exercise to a social media audience.	Focus: The understanding of the contribution of online social to social norms and services	The integration of social research to big data and knowledge management applications. The sophisticated behavior mining through functional and quantifiable social generated data.
[53] West, Sarah Myers. (2019)	Data capitalism: Redefining the logics of surveillance and privacy.	Focus: Understanding the connections of data utilization, surveillance, and privacy	The analysis of surveillance and privacy issues related to a crowdsourcing prototype system.
[54] van der Schyff Karl, Flowerday Stephen, Furnell Steven. (2020)	Duplicitous social media and data surveillance: An evaluation of privacy risk.	Focus: Elaborating on the compromise of social media and privacy risks.	The analysis of prerequisites for the use of "private" data or data related to social behavior for social impact applications.
[55] Jahanbin Kia, Rahmanian Fereshte, Rahmanian Vahid, Jahromi Abdolreza Sotoodeh. (2019)	Application of Twitter and web news mining in infectious disease surveillance systems and prospects for public health.	This study aimed to develop a text-mining technique for extracting information about infectious diseases from tweets and news on social media.	The analysis of easy-to-use scenarios for citizen's involvement in knowledge management applications.
[56] Nouira Kaouther, Njima Nesrine Ben. (2020)	FluSpider as a new vision of digital influenza surveillance system: based on Big Data technologies and Massive Data Mining techniques.	Focus: A prototype system aiming to enhance digital surveillance	The understanding of critical surveillance aspects of crowdsourcing platforms.
[57] De Vreede Triparna, De Vreede Gert-Jan, Alawi Naif. (2021)	Achieving Success in Community Crowdsourcing: Lessons from the Field.	Focus: Analysis of success factors in community crowdsourcing applications	The development of design principles of social crowdsourcing applications that will promote social value of big data.

Table 2. Main areas of interest in the studied literature.

Author	Data	Key Emphasis				
		Crowdsourcing	Social	Impact	Network	Surveillance
Putrenko, Viktor [46]	X Data mining	X Crowdsourcing projects	X Open street map Citizens social activities		x Internet	
Chang, Danni. et al [47]	X Data mining	x	X Product affective property		X Web	
Hoi, Calvin SH. et al. [48]	X Data mining Big data	x	X Crowds of users			
Guo, Kehua. et al. [49]	X Big data	X Crowdsourcing computing			X Social media Internet of things	
Lytras Miltiadis D. et al. [50]	X Big data research	X Social and humanistic computing	X Social sciences Social good	X Social impact	X Web sciences Innovation networks	
Yoon Ayoung. et al. [51]	X Data reuse Open data		X Community informatics	X Social impact		
Carpenter Christopher J. et al. [52]	X Big data		X Exercise		X Social networking sites	
West, Sarah Myers [53]	X Data mining			X privacy	X eCommerce	x
van der Schyff Karl. et al. [54]				X Information privacy	X Facebook apps Social media risk	X
Jahanbin Kia. et al. [55]	X Text mining		X Infectious disease		X Twitter	X Surveillance system
Nouira Kaouther et al. [56]	X Big data Data mining		X Diseases Social phenomena		X Web pages Google	X Digital influenza surveillance system
De Vreede Triparna. et al. [57]	X Filed study	X	X Social perspectives in collaboration research	X Phycological perspectives in collaboration research		

Putzenko et al. [46] set a significant basis for our research study since they discuss the integration of crowdsourcing and social activities in the context of a project called Open-Street Map. Our key intention in this research is to investigate the willingness of citizens to contribute open anonymous public data for a platform related to quality of life in a modern smart city. We do understand that such an intention is also related to significant ethical issues related to the data privacy and protection of personality.

In other studies, e.g., [47], researchers focus on the identification of properties that are exploited in crowdsourcing for meaningful scenarios. In our study we are interested on the classification of properties that are required for the annotation of human activities in modern smart cities. The next aim, closely related to the one stated is that such annotation will allow sophisticated visualization of human activities on dashboards that can be accessible by the citizens in a modern city. This capability will allow them to integrate their mental capability for enhanced decision making.

Sophisticated pattern mining [48] is also evident in the relevant literature in the research phenomenon under study. In our work, we are interested in understanding how citizens' involvement and contribution in a crowdsourcing platform can promote the well-being, the security, and the safety of the community. In our research we are aiming to understand how user-generated data related to human activity are leading to meaningful insights for the life and the behavior in the community. In the relevant literature, another

stream of discussion is related to the semantic annotation of multimedia content [49]. The modelling of annotation schemas of microcontent and the tagging with restricted properties of social activities and human activity is in the focus of our research. Such discussion automatically links our research problem to the debate on the social impact of big data applications and to topics related to privacy and surveillance [50–56]. The investigation of implications of big data and crowdsourcing applications to the data privacy and regulations is a very complicated topic and, in our study, we will not analyze it in detail. We intend soon to complement this study with new research exclusively focusing on the ethical issues associated to a big data crowdsourcing platform for human activities in a smart city.

Our research is an ambitious effort to investigate the way that big data research is creating social value in smart cities [58] towards innovative social impact applications [59] with a significant crowdsourcing component [60] that meets critical ethical, technical and social challenges [61].

The wide variety of the studies summarized in this section offers significant insights into our research problem:

- **Location-aware and GPS-driven data are used extensively for social research** purposes and to associate behavioral aspects with destinations and navigational patterns. It is significant to understand how these technological capabilities can be synthesized in modern cities for the provision of value-adding smart city services with an emphasis on the improvement of the quality of life, wellbeing and safety. In our research, we try to understand how location-based services, data mining and crowdsourcing technologies can be synthesized to allow analysis of human activity in modern cities. It is also important to investigate the different value layers for the processing of crowdsourcing data for different types of users including practitioners, citizens, policy makers and other stakeholders. For example, if you identify through a crowdsourcing platform that in a neighborhood of a city there is increased insecurity then the question is what are the implications for policy makers or social services?
- **Ethical and privacy issues arise when exploiting personal data for decision-making.** In our simple prototype, the following questions raise ethical issues and require extensive future research:
 - In a modern society, are citizens allowed to use smart services or applications in smartphones to report human activities of other people? If yes, under which circumstances, and if not, why?
 - What are the privacy and the data protection regulations for allowing crowdsourcing applications to maintain sensitive data related to human activity?
 - Is there a requirement for extensive debate and public consultation of regulations related to digital social services that involve human entities in the collection, processing and exploitation of personal data related to behavior, beliefs and opinions?
 - Under which circumstances can anonymity of both data subjects and objects be granted for the maintenance of social impact services?

How can we promote sustainable crowdsourcing platforms for social impact in a modern smart city with full protection of citizens' personal human rights and privacy? As a concluding remark, the literature review provided significant insights for the specification of the research problem of our study. Diverse aspects of the phenomenon include computing social and ethical parameters that need further investigation. We must communicate from the beginning that the deployment of information and communication technologies for addressing significant social challenges and problems is a complicated and multi-faced research context. Our main aim in this study is to provide the basis for a scientific debate on how crowdsourcing and big data technology can enable a trusted, ethical system for tagging human activity in a smart city environment.

In the next section, we summarize our research methodology.

3. Research Methodology

The integration of computer science in social research is a new trend: sophisticated computational methods are applied to well-defined social challenges. The availability of a plethora of big data in societies allows innovative business models and value-adding services. Our research has a key focus on exploiting open public data to analyze behavior.

In the previous section, we discussed various studies on the exploitation of GPS data and location-aware services for social analysis and behavior analysis. The main objective of this study is to use the latest research developments in data mining and artificial intelligence to study social phenomena. We have four critical research objectives:

- To establish how data mining and location-aware services can inform data-driven models for the study of human behavior. We exploit open public GPS data to construct a rating system for social behavior, based on a heuristic algorithm that we explain later;
- To identify the critical aspects of a social data-driven system for analyzing human behavior. We use GPS data available from mobile devices and our innovative app to analyze how citizens of a smart city tag locations and to associate behavioral aspects with their profile;
- To investigate how we can deploy advanced data-mining methods to understand individuals' ethical values through their GPS traces. The intention is to aggregate, through an innovative app, mobile phone users' GPS traces to construct a heuristic algorithm for analyzing humans' behavior in smart cities. Categories of human activity are recognized through GPS and then scored by a heuristic algorithm to update individuals' behavior profiles;
- To identify the ethical limitations and barriers to the conduct of computer systems research for social purposes, including privacy concerns. The availability of computational methods and services for aggregating, annotating and composing big data systems for social computing raises critical ethical questions and concerns. We intend to offer our basic understanding, based on the delivery of our research and social computing system.

In this study, a heuristic algorithm was used to enhance the abilities of a GPS device, and location services were used to trace individuals' preferences. A system of six ethical values was constructed to monitor individuals' navigation and behavior. Full user consent to use personal GPS data was secured, with voluntary agreement to participate in our survey.

The basic scenario for the social computing system that we developed is as follows:

A user provides consent for integrating their personal GPS traces into a web platform built for the purpose of our research. The system allows users with their mobile phones to tag six types of activities related to six ethical values that are listed by the research team. Then, a heuristic algorithm is used to attach scores and ratings to aspects of users' behavior. The sophistication of the system allows user reports to be generated and reports relating to behavioral aspects of human actors.

The participants in the study were students, researchers and citizens in the Kingdom of Saudi Arabia who agreed to use, on a voluntary basis, the prototype system presented in the next section. The selection of users was based on random criteria and is representative of the human factors in the Kingdom of Saudi Arabia. We plan to soon deploy the same system with a wider population of users and to integrate this study's key findings to produce an updated version of the system. Participants agreed to use it in the understanding that their anonymity and personal data were protected. The analytics generated by the system are analyzed in Section 5.

The study's research approach can be summarized as:

- Critical literature review of the domains of data mining and GPS-driven location-aware services for social sciences. The purpose is to understand complementary aspects of the phenomenon to enlighten key aspects of our research problem;

- Development of a heuristic approach to the GPS-driven set up of a behavioral profile of smartphone users, with special attention to navigational patterns and values. The selection of the research problem relates to a project supported by the funding organization, as well as to the sensitivity of personal data and their exploitation for decision-making purposes;
- Development of a research prototype based on a smartphone-enabled GPS system. Our aim is to prove that current advanced computational methods can promote data-driven insights into aspects of individuals' behavior. The intended impact on society and the capacity of big data analytics to serve the social good are investigated. One of the key dimensions is the integration of computer and social sciences to provide ethical and sustainable social computing applications;
- Visualization of and analytics for behavior analysis and decision-making. This computational approach investigates their ability to enhance decision-making capacity. The study proves, on the one hand, the feasibility of data mining and analytics and, on the other, raises important ethical issues;
- Generalization of findings on the potential of advanced information systems' research in the social sciences. The key objective is to interpret the pilot's findings in the wider context of the convergence of information systems and social sciences research.

Our study complies with all ethical standards for conducting social sciences research and respects the anonymity of those who participated in the pilot study.

In the next section, we summarize the main aspects of the pilot system that we designed and used for the research purpose of this study.

4. A Crowdsourcing-Enabled Prototype for Behavior Analysis: Associating Values and GPS Data

The intended contribution of this research is the integration of computer and social sciences. The main idea is to analyze public or personal data with users' consent to interpret aspects of human behavior.

The chief purpose is to understand how advanced information systems research can guide social sciences research. To collect and analyze the data, a prototype system was developed whose main component is a GPS device capable of understanding and monitoring individuals' navigation and preferences. A heuristic algorithm was used to analyze the GPS traces and provide insights into where users spend time in their daily schedule.

This research uses an android app to collect data related to behavioral values. GPS was used specifically to locate users' geographic distribution. The basic assumptions of our prototype are that:

- Analysis of human behavior involves detailed mapping and processing of human activities. Some can be concluded by GPS traces and by users tagging their locations and activities;
- A sophisticated social/behavior rating system can be associated with numerous aspects of behavior. In our pilot study, for research purposes we decided to focus on six "moral" values: honoring, respecting parents; providing food as charity; honesty; praying; traffic violation; and respect for market rights;
- We developed a prototype for social rating by providing smartphone users with an app with which they could tag and annotate their activities that relate to these six values. We decided on this set of six activities for research purposes and, in a future study, we intend to expand the list from which the activities can be selected;
- By aggregating all these activities and associated values, we created a visualization component to detect patterns of behavior and ratings;
- This basic infrastructure can be used for decision-making in various ways: examining behavioral patterns as a basis for the prediction of future behavior or for the analysis of "un-ethical" activities in each city or region. Additionally, advanced decision-making

can be delivered to reward human activity that is socially responsible or to motivate humans to promote ethical behavior.

We also have to state that our social computing app promotes collective intelligence. Users equipped with our web app can report and refer to ethical values associated with others' behavior. Consider the following three scenarios:

Scenario 1: A pilot user of our system is at a road intersection without traffic cameras and realizes that a car driver is jumping the traffic lights. The user reports this behavior, and an indication appears on the map for the relevant activity.

Scenario 2: A pilot user in a shopping mall reports that a citizen is disturbing someone, and another user in the mall is indicating that someone is jumping the queue. She or he uses the app to report these violation of market rights.

Scenario 3: Thousands of users provide items to the platform on the six indicative values, and policymakers across Saudi Arabia analyze regional differences so that they can adjust the social policies.

This data-reporting capability serves as a data-ecosystem layer that allows the aggregation of user-generated geo-located tagging of human activities. Data that are aggregated and processed by the system have several decision-making aspects:

- At the individual level, users can understand a regional map of various behaviors. Through analytical processing of GPS data, they can recognize on a real-time basis the key facts revealed by this social rating system. The indicative system measures six ethical values, giving a good approximation of the collective set;
- At the policy-making level, the system offers interesting social behavior insights for use as input to revise policy initiatives;
- At the governmental level, the data processed by the system provide a sophisticated, neutral, and objective way to measure the evolution in ethical values over time and a sophisticated big-data ecosystem that justifies developing services that add value;
- At the business and industry level, selective ethical values reported in the system provide interesting facts to inform business trust and honesty and the preservation of market rights. This can be an indication of healthy economic activity in specific areas and regions.

In the next section, we provide technical details of the system and key aspects of the empirical analysis of testing. As mentioned earlier, our research does not involve questionnaires or any use of the user's identity. The study can be extended to any region of interest in the Kingdom of Saudi Arabia. The main function presented by this article is as a test bed. In a future implementation, we plan to add to the list of ethical values used by the system and employ further visual analytics. We need to communicate that this research responds to the demands of our times to implement services that relate to society's problems and social challenges.

4.1. Interface and User Model of the Prototype

The following steps were adopted to analyze ethical values:

- *Installing the app:* The developed app was made available in Google Play. Interested users installed it to participate in the study. During the installation, it was explained that only data pertaining to the device's location would be captured for the study; apart from that, no details would be recorded at any time. Certain permissions were requested so that the user's position could be used to supply data to a pool for further analysis and better understanding of ethical values;
- *Adding the moral value:* Users were prompted to select an ethical item, thereby deciding the parameter to be captured by the study. Next, they activated the geo-location function on the map. Each item marker represents a specific ethical value. Thus, upon completion, the procedure had recorded both the GPS location and the value, which are then used in the study;

- *Statistical analysis:* Each GPS marker indicates a corresponding ethical action. Both GPS and ethical value data were collected from the end sessions; these values represent those that maintain society's standards. The main statistical finding is the location, number of users and value data. When the user adds ethical value data through the app portal, each item is of a distinct color.

The main objective of the GPS-DCD is to facilitate data collection, at the same time adding the location of the individual ethical value. The app can be used by any individual with a smartphone.

Figure 1 shows the GPS-DCD's system architecture. In the first step, before he/she can utilize its functions, the user needs to be verified by the app as registered. This process is mandatory, as certain policies must be accepted so that the location can be used by the study. The services block represents the protected algorithm that checks user log-in activity and, upon successful log in, allows the user to connect to the app's database.

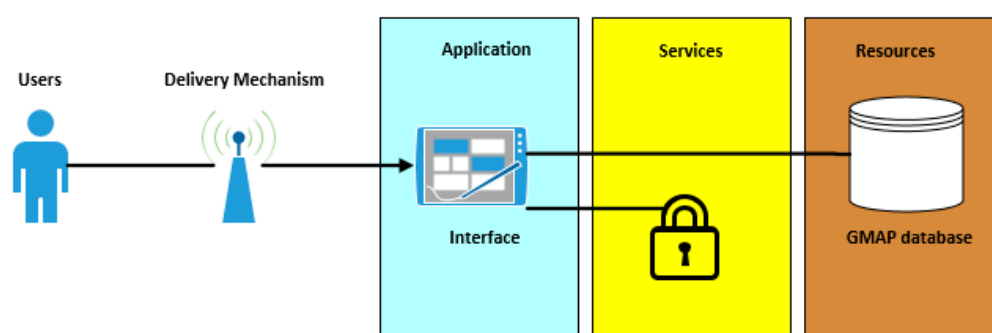


Figure 1. System architecture of GPS-DCD.

Figure 2 illustrates the conceptual structure of the database's construction. It shows various services, including GPS location, listing and feedback. The database's conceptual diagram shows the data flow from the user to the server end.

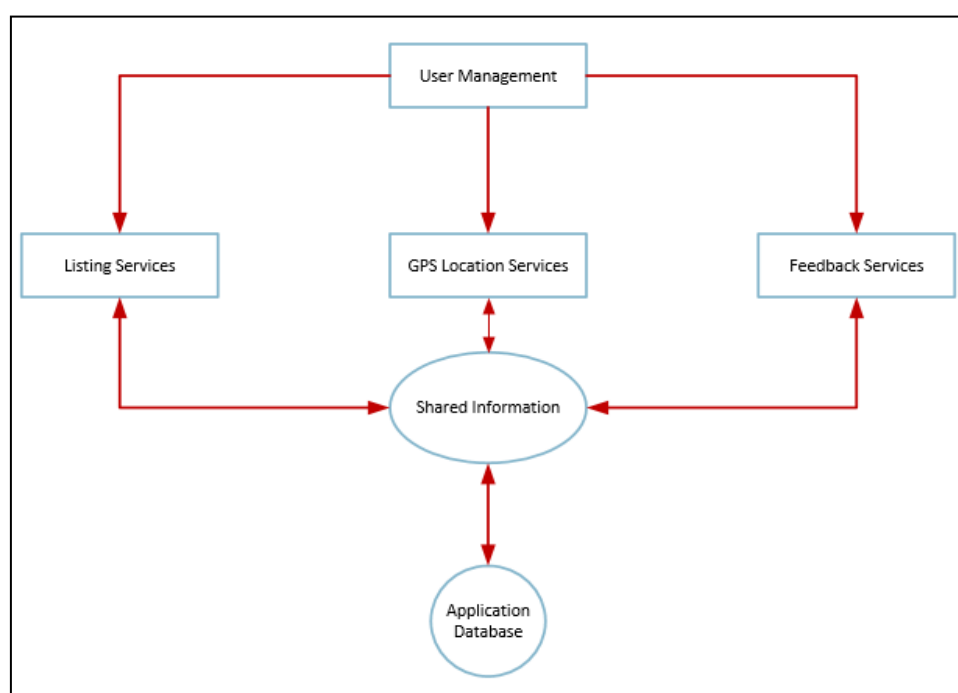


Figure 2. Database conceptual design.

Figure 3 shows the system's entity relationship diagram (ERD). The representation is as follows: the table "G-Map_users" contains multiple data items such as "uid: integer (10)". The integer 10 describes the datatype of the field's uid. The uid of the table "gmap_users" are linked with uid: "gmap_item_list". By sharing the same field as the table, data can be shared and arranged separately.

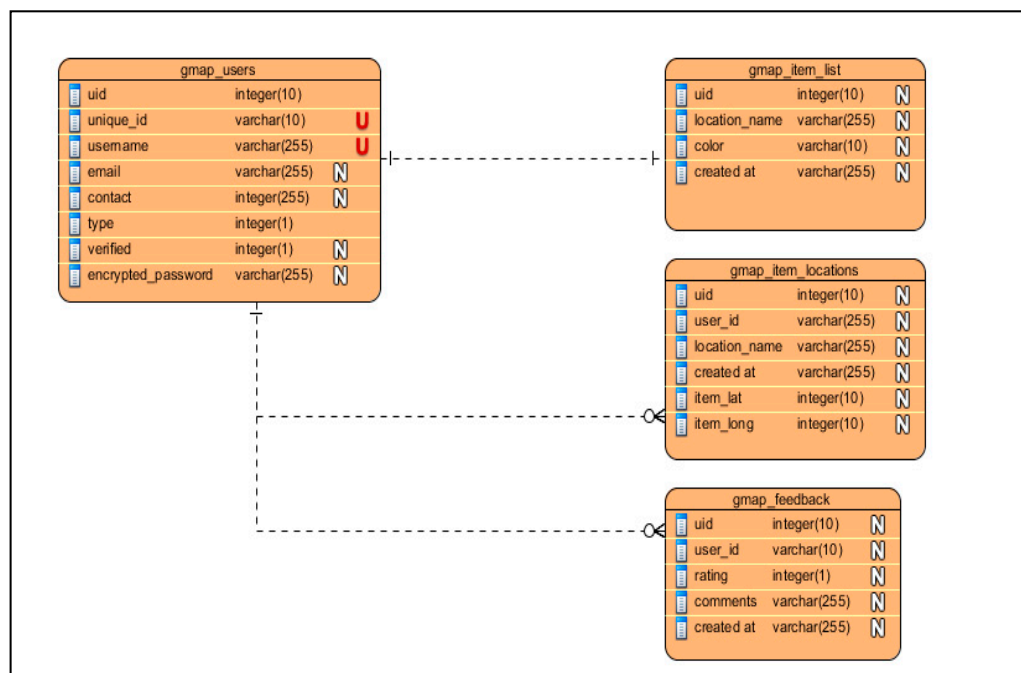


Figure 3. ERD diagram.

The block diagram in Figure 4 shows the app's working mechanism. The algorithm associated with the working mechanism is as follows:

- The user logs into the system;
- Then, the user is diverted to the Main Activity, which displays the menu in the dashboard.

From the dashboard, the user can perform the Barograph, select_item, Maps Fragment and Feedback activity, which allows him or her to view the statistics, select the corresponding item, view markers on Google Maps and leave feedback on the app, respectively.

- Once the specific activity is selected, the Data Object model loads the detailed data of that selected activity;
- When special credentials are used to log into the system, the automatic allocation and necessary permissions are granted to access the app as administrator. In this case, the dashboard lets the admin make changes to the app's items.

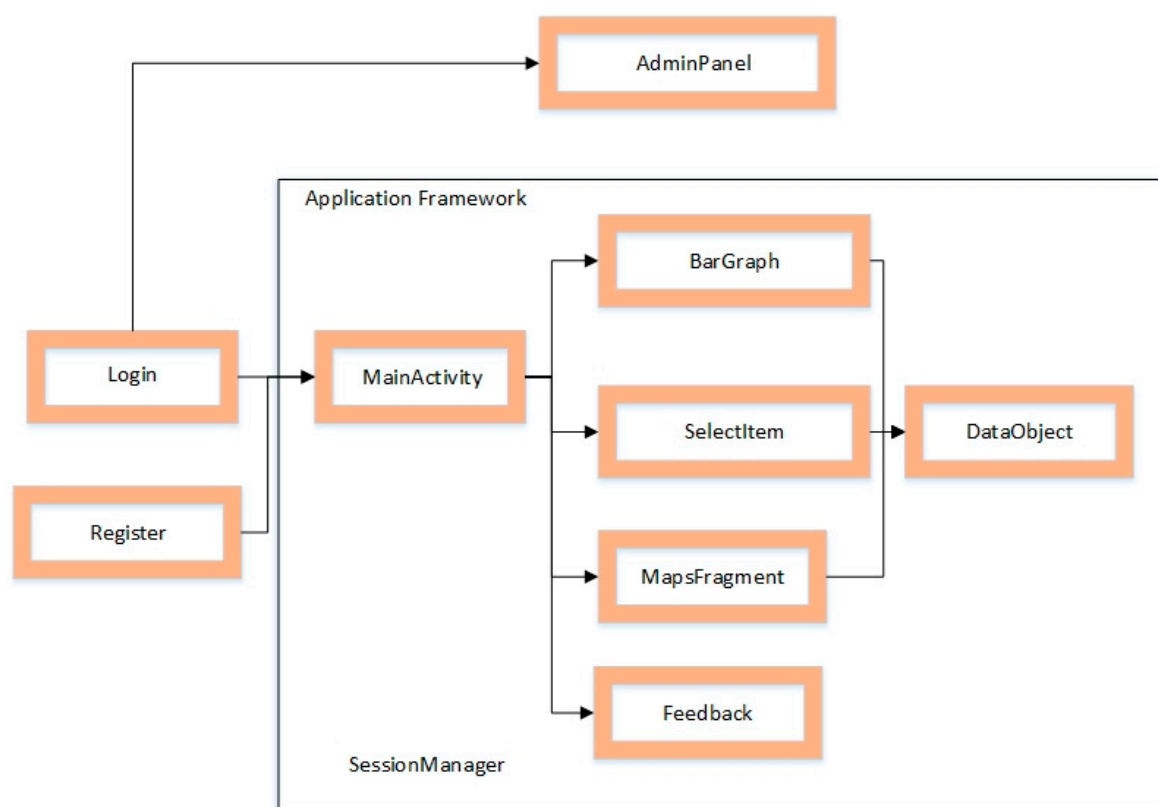


Figure 4. Working mechanism of application.

4.2. Experimental Set-Up

As described in the research methodology section, the participation of researchers, students, and citizens of the Kingdom of Saudi Arabia in this survey fully respected their privacy and anonymity. Participants agreed to use a web app installed on their smartphone and to share their GPS data for the purpose of analysis.

The selection of pilot users was based on random criteria and anonymity. In total, 150 users participated in the pilot. The main objective was to form a team reflecting the main elements of Saudi society. The small number of participants is certainly a limitation, yet it provides a compact and representative group. We intend in a second run of this study to increase the participants and update the functions with its key findings.

In the following paragraphs we discuss the experimental set-up of the pilot system.

All users were required to create an account in the app. He or she could choose an automatic sign-in process for rapid access without repeatedly keying in credentials. The app allows the administrator to select items and colors from the list. Note that each ethical value may be assigned a marker color only by the administrator, for ease of visualization.

The experimental set-up assumes that cultural behavior in any given society relates to the values of that society. The development of a mobile information system can fulfill the task of gathering its various values. The database of both positive and negative values is stored in the system to determine that society's human behaviors.

Figure 5 is a screenshot of an example of data collection based on location, detail, query and history. It shows the locations on the G-Map, with ethical values across the world.



Figure 5. Screenshot of the app for data collection.

The experiment was conducted with 150 participants. Each was requested to add 10 items by clicking on the button on the map, thus 1500 items were collected by the end.

The visualization capability of the system provides a user-friendly interface for the aggregation of instances demonstrating ethical values according to our own assumptions and the heuristic algorithm.

The overall options of the app are summarized in Figure 6. This shows the various options that a user can select. First, the location option enables the user to see all available locations on the app. The second option is only accessible to the administrator. Using the third option, feedback about the app can be provided for further improvement. This experiment was conducted to collect geographical data that relates to the following values:

1. Market rights;
2. Providing food as a charity;
3. Obeying traffic rules without the need for any form of camera monitoring;
4. Honoring one's parents;
5. Prayers;
6. Honesty.

The basic idea explored in this study is simplistic, yet it can be used to derive associations. Users share their GPS data and, based on their location, (whether they are visiting or it is a place that they prefer), the metrics for their “values” are aggregated and thus inform an advanced behavioral user profile. The use of a mobile app capable of reporting and monitoring these six indicative ethical values represents an adequate test bed for analysis of social computing capabilities. In a future implementation, these user-generated data will be enriched through automatic extraction of check-ins and activity records of users/human actors.

The system, in fact, exploits users' posts on the behavior of other individuals, in both parts anonymously. Thus, the collective intelligence aggregated by this system can be used in the analysis of social behavior and for benchmarks across regions, cities, neighborhoods and locations. We understand that the system-reporting capability raises critical questions

for social issues, as well as for computing ethics. Our main claim is that, nowadays, various technological capabilities enable cyber-physical systems to undertake extensive social computing, monitoring and reports.

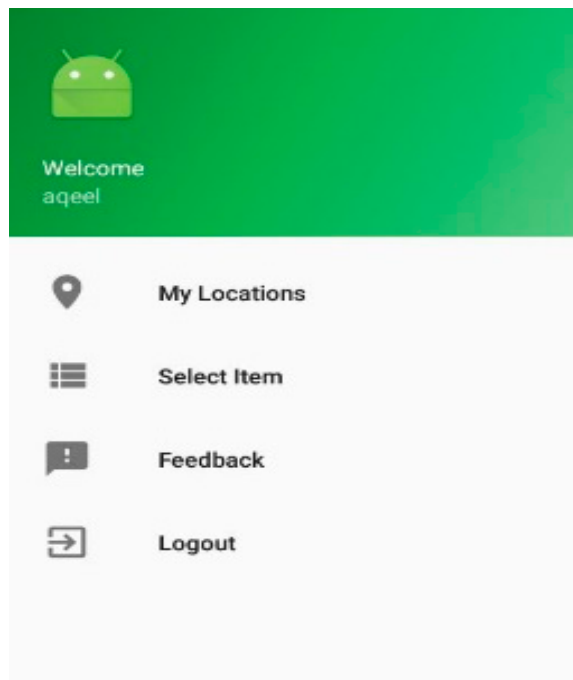


Figure 6. Application side bar.

The selection of ethical values by the system is indicative only, for use as a test bed to prove the capacity of a heuristic algorithm to match locations with a social rating system. Users can select any item that appears on their screen, as shown in Figure 7a,b. Details of their selected item and user records are supplied to the database, together with the GPS location and the user's identification. The items show the ethical values of the end user in the administration panel. This portrays the details of the number of users and locations listed on the G-Map.

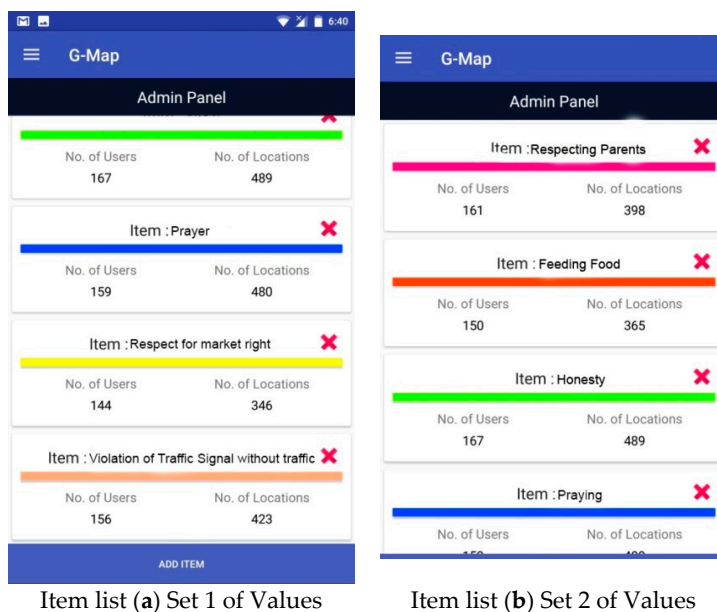


Figure 7. Item list (a,b).

5. Empirical Analysis and Testing

The overall purpose of this pilot study was to demonstrate that, currently, smartphones and other portable devices offer GPS location-aware services with significant crowdsourcing capabilities aiming to promote the social impact of big data applications. In fact, they deliver an entire data ecosystem for “consumption”. The data are available for further analysis.

We decided to build a pilot system in which users agree to share their location-aware data for scientific purposes. In fact, based on their navigation and the places that they visit, we developed a crowdsourcing application for recording and tagging significant social events and activities. We also developed a heuristic algorithm to aggregate their preferences. For example, if somebody visited a charitable organization that collects food for poor people, this was reported on the system through their current location. In our system, for research purposes we integrated analytics services capable of summarizing collective behavior; that is, participants’ social behavior. Potentially, our approach can be used in social rating systems or advanced behavioral analysis.

In the following paragraphs, we summarize the indicative results of our research. As mentioned, 150 users comprised our sample. Each was provided with the prototype app and used its functions to report on values.

The experiment’s results, as shown on the G-Map, are six ethical values of human life. The analysis summarized in this section is based on the color markers in the region of Saudi Arabia. The following results show the number of users, the values and the locations automatically retrieved from the GPS data.

For each value, a different-colored marker is displayed on the map, according to the number of instances.

Figure 8a shows the six values being studied and their distribution across Saudi Arabia. It also shows the location of a given item by its color. The app allows users to share data with other users. The value’s name and color are set by the administration only, who log into the app to do so. No standard was developed to decide the color, and the appropriate and conventional color chosen by the society in question can be used.

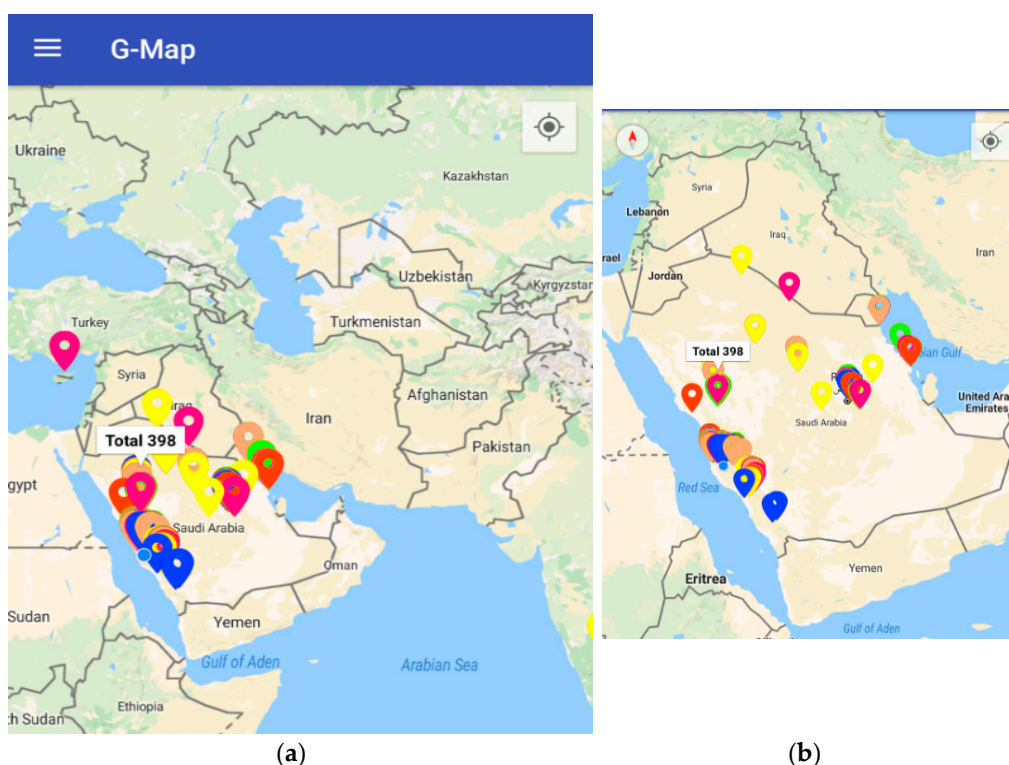


Figure 8. (a) Data (colored markers) on the map. (b) Data on the map for the whole country.

Figure 8b, on the basis of the colors shown on G-Map, indicates that 150 users reported their ethical values. It shows where those behaviors were located in Saudi Arabia and surrounding countries.

Figure 9 is a zoomed-in view of the distribution of ethical values. These were selected by people in Riyadh and west coastal areas. It shows where those behaviors are distributed in Saudi Arabia only.

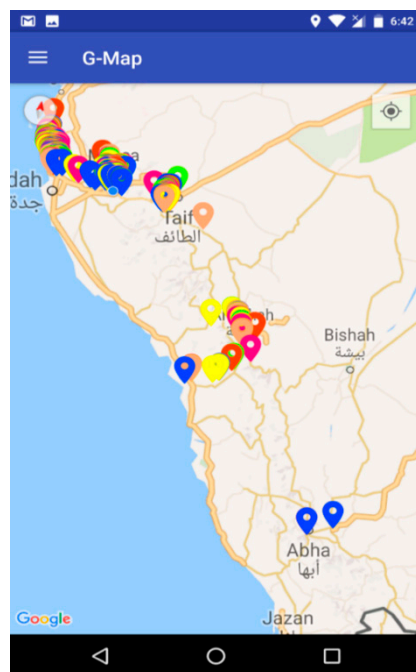


Figure 9. Data on the map of the west coast.

Figure 10 shows the data collected in the experiment. It clearly indicates that 150 users used the app to add 98 locations for Item 1, “Honoring parents”. For Item 2, “Providing food as a form of charity”, 150 users added 65 locations. “Honesty” is the value that attracted the highest number of entries of all the items. The “Honoring parents” instance shows that a total of 161 users used the app to add it at 398 locations; the instance of “Providing food as a charity” was recorded by 150 users, using the app at 365 locations. The instance of “Traffic without cameras” saw 156 users using the app in 423 locations. For “Honesty”, 167 users used the app at 489 locations. For “Prayers”, 159 users used the app at 480 locations. Finally, for “Market rights”, 144 users logged onto the app at 346 locations.

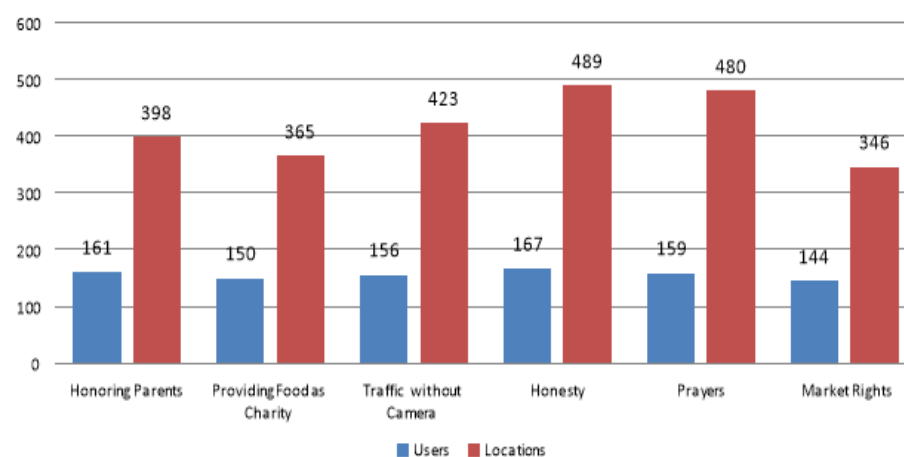


Figure 10. Data per item.

It is evident that most users logged in when they considered “Honesty”. Users have logged this same parameter into the app at a high number of locations.

Figure 11 illustrates that most data entries were made in the Makkah and Jeddah areas. It shows where those behaviors are distributed in the west of Saudi Arabia, showing just the Jeddah and Makkah areas.

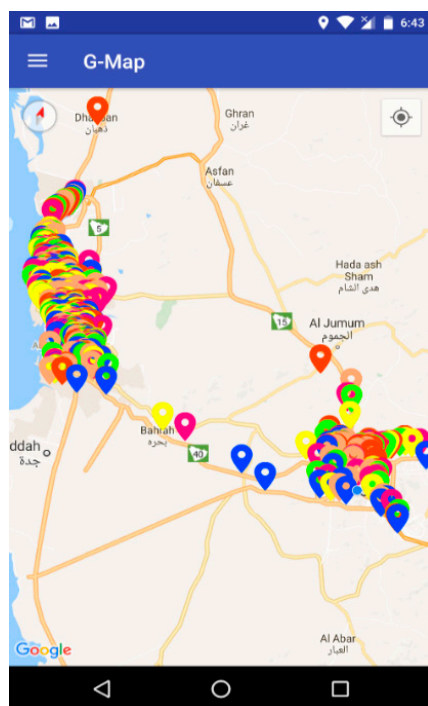


Figure 11. Data on the map for the Jeddah and Makkah areas.

Figure 12 illustrates how the six values are distributed in the Riyadh area. Figure 13 shows the area in a zoomed-out view, with a view of Alqasim. Figure 14 illustrates how honesty, as an ethical value, is distributed in Almadina.

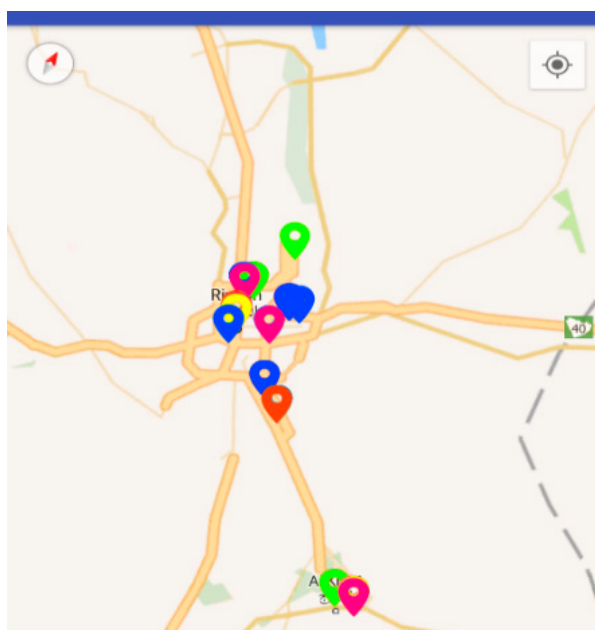


Figure 12. Data on the map for the Riyadh area.



Figure 13. Data on the map for the Riyadh and Qasim areas.

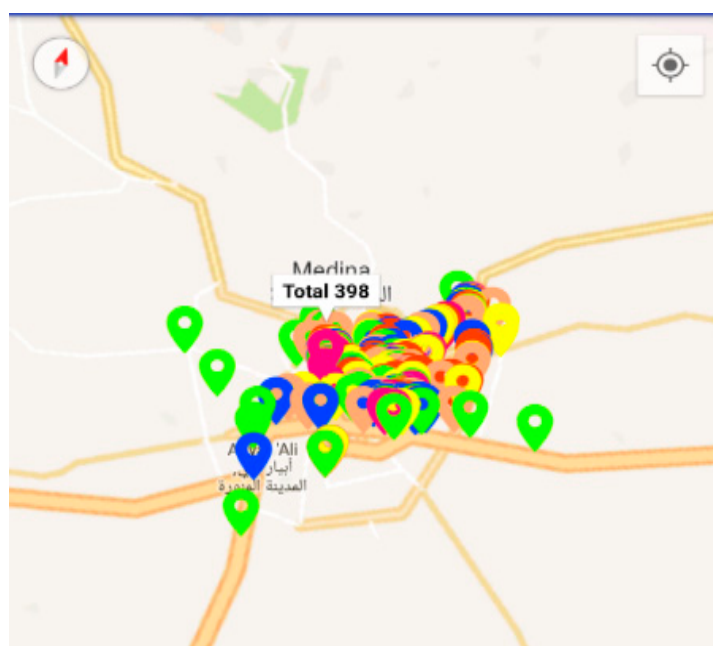


Figure 14. Data on the map for the Almadina area.

Figures 11–14 show significant values for the location and magnitude of each ethical value across the country. Zooming into an area shows its magnitude there, facilitating decision-making on actions (e.g., education, awareness).

Figure 15 shows all the data pooled from the application. It is evident that the highest number of user reports was for honesty as an ethical value; users logged into the app most often for this same value.

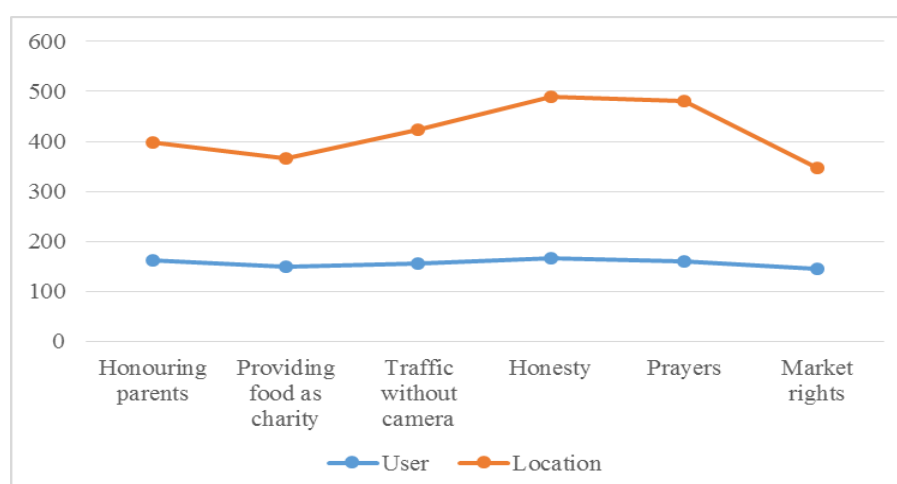


Figure 15. Scatter plot of data collected using the app.

It is evident that this kind of data-driven analysis requires the consent of all its participants. From a computer science point of view, it is evident that the sophistication of the system can be supported by current advanced data mining, analytics and location-aware services. The ethical side is more complicated. The quest for apps that promote the social good and increase social impact poses critical questions about the use of sensitive personal data for developing indexes that measure humans' behavioral aspects. In this context, our study remains neutral, even though we, as researchers, have personal opinions.

In Table 3 below, we provide an overview of participants in our pilot study as well as their crowdsourcing contributions from various locations.

Table 3. Number of samples, in terms of users and locations.

No.	Ethical Value	Users	Locations
1	Honoring parents	161	398
2	Providing food as charity	150	365
3	Traffic without cameras	156	423
4	Honesty	167	489
5	Prayers	159	480
6	Market rights	144	346

We must make some important clarifications of our research regarding the results previously presented.

First, our work contributes to the intersection of social science and computer science. Recent debate on the social value and social impact of big data and analytics has brought up various aspects of so-called social computing. Our work is a first effort to present a fully functional research tool that manages collective intelligence relating to a social challenge, namely rating ethical values in Saudi Arabia.

The model underlying this research presents a variety of aspects that require further future investigation. For example, a key question is under which circumstances it is ethical to analyze socially sensitive big data through GPS or other sources to gain better insights into social problems and challenges.

Another critical aspect of our research is its basic understanding that prototype apps, such as that in our study, provide a layer of critical data aggregation for sensitive social data. Further exploitation and analysis of these data for sophisticated decision-making is a matter of formulating strategies. From this point of view, there is a need to integrate computer scientists, social scientists, policymakers and government officials to devise

a long-term strategy for big data management for social purposes. In our research, we depicted how various instances of citizen-generated big data relating to people's behavior in the context of smart cities can be processed further through analytics and visualizations to offer key insights into life quality and socialization.

The implications of our work for computer science are also significant. Significant enhancements are required for the standardization of tagging and annotation schemas for behavioral aspects of human activity, as they are recorded and stored in social media applications. Furthermore, sophisticated algorithms for the automatic annotation and tagging of human behavior based on publicly available data can be promoted further. These developments in big data and crowdsourcing research set significant ethical concerns. It is critical to promote the scientific debate on the protection of human personality, privacy and identity and also to develop systematic, trusted and transparent procedures for the user consent on the use of their personal data.

In the context of the smart cities research, the evolution of human-centric, citizen-aware applications will require new standards and collaboration of diverse technical communities and social institutions and associations for trusted ways for processing personal and community data.

The chosen narrow focus of our approach to a limited set of values represents a pilot study that can be expanded further. The overall idea is that GPS applications can be used to monitor a huge variety of social aspects of human activity and behavior. In fact, this is our intention in a further round of research soon. Our intended contribution in the intersection of the human social and semantic web [62] and social multimedia systems [63] for social value and impact. Another critical dimension for our applied data science experiment is the formulation of a "knowledge creation" process [64], towards value adding layers for advanced decision making [65] in the context of smart cities [66] for social insights.

6. Conclusions

This article shows the viability of gathering geographical data relating to values by using GPS systems on hand-held devices such as smartphones. The app was used to collect data relating to six ethical values in society: honoring parents; providing food as charity; observing traffic regulations without camera monitoring; honesty; prayers; and market rights. The efficacy of this tool has been proven through the experiments, and it can be used to gather various forms of data.

This work emphasizes data collection via GPS and enhancement by data visualization. These can easily be accommodated and interpreted for decision-makers. Only six values were considered for the purpose of this evaluation, but in future the number of value parameters can be increased. Another issue is the location of study. This was the Kingdom of Saudi Arabia, yet the same study can be applied to other regions.

While, for a naïve reader, such an app and research seem to fall outside the typical remit of information systems and computer science, it is evident that in future years values and other socially aware variables will initiate significant research. The adoption of technologies such as artificial intelligence, machine learning, sentiment analysis and advanced visual analytics will launch a new research field at the intersection of information and computer science and the social sciences.

One of the most critical aspects of this research is related to the ethical issues associated to the data management and data utilization for applications that potentially can have a social impact. In the context of smart cities, the quest of value adding services for improvement of the quality of life, wellbeing, safety and security, is closely related to advanced processing of personal or community data. From this point of view, the feasibility of crowdsourcing platforms needs a thorough justification of data processing regulations, including ethical considerations.

In our research, we combined leading edge technologies of our times to design and implement a crowdsourcing application in which volunteers with confirmed consent ob-

serve human activity around them in the context of a smart city and annotate microcontent linking human activities and locations. The technical aspect of our application, presented in detail, allows for several comments about the capacity of various technological components to support services. However, the social aspect of our application must be discussed further.

The first important fact is related to the social impact of this kind of application. The adoption of crowdsourcing platforms in the context of smart cities, can serve as the basis for the collection and aggregation of big data with references to human activity, preferences and the rating of experiences. However, when it comes to human behavior, the use of big data and data mining to maintain sophisticated dashboards that inform the citizens and the community about sensitive social data opens a big discussion about the advantages and the consequences of these services to the privacy and security of data objects and subjects. According to recent research [50], while citizens are concerned about their privacy in big data applications, they could consider the use of such services if the social value and impact is well justified.

Our research study aimed to communicate the use of big data, data mining and crowdsourcing technologies for the delivery of smart cities applications for tagging human activity in the context of modern cities. It is evident from our prototype and empirical testing that humans can potentially use such applications and contribute big data on data repositories that can be processed further. The directions and the depth or sophistication of such analysis are related to desired strategies and scope of data processing. Our main intention was to provide an initial testbed for further elaboration with the various computing, social and ethical aspects of such applications.

In such a demanding context, the limitations of our study are also significant. In our service, we deployed anonymity for the subjects and the objects of the crowdsourcing platform. This means that, when a volunteer uses our application, they contribute microcontent for the human activity of other people without emphasis on their identity, they simply report the relevant activity. This selection automatically limits the capacity of sophisticated data mining to further process these data.

The second limitation of our study is also related to our decision to deal with only six types of human activities. We intended to provide empirical evidence for our prototype and to deliver the key technical aspects of our crowdsourcing strategy and algorithm. In a future research deployment, several other activities can also be implemented, permitting the population with related social big data.

Finally, one of the most critical limitations for our study is related to the detailed analysis and study of ethical and social concerns for such a system in which humans comment on human activities of other people. We plan to deliver a thorough study on these aspects, and thus this discussion is limited in the current research work.

The future work and research related to our study can be related to technical, social and ethical aspects. From a technical point of view, the next phase of enhancement in our approach includes the design of a more sophisticated dashboard and a better interface. We also intend to adopt machine learning for the understanding of human activity patterns and prediction of events in the context of a smart city. From a social sciences point of view, we intend to investigate in future research the social impact of such applications and to study the compromise on data privacy that citizens are willing to allow to enjoy crowdsourcing applications that allow them to have better quality of life and security in a modern city. From an ethical point of view, we also plan to run a qualitative research with key stakeholders to further analyze and map the key ethical concerns and the required actions in terms of regulations and policy making. Some more directions of future research are highlighted below. Another future direction of our research is to adopt sentiment analysis through social media to study ethical values in a wider context. Processing the micro-content available in social media offers critical insights into people's beliefs, opinions and values in a wider context of human expression.

In the same direction, our plan for future research is a pilot study on a social dashboard and analytics system based on human activities. The analytics and visualization of this

study will support interesting key performance indicators (KPIs) to analyze aspects of human behavior. It is our belief that this is a worthwhile case study of the ability of modern information processing methods to enhance the study of social phenomena. We look forward to additional contributions to our study from both computer scientists and social scientists and to international collaboration.

The recent significant funding on smart cities applications for quality of life, wellbeing and safety together with the evolution of emerging technologies, such as virtual reality [67] and social media [63], set a more demanding context for crowdsourcing applications with enormous capacity for monitoring human activity and behavior. This will set new ethical questions for the new generation of crowdsourcing applications.

Author Contributions: All authors contributed equally to all phases. All authors have read and agreed to the published version of the manuscript.

Funding: The project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (DF-581-611-1441). The authors, therefore, gratefully acknowledge DSR technical and financial support.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Wadee Alhalabi institution.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data for this study are available upon request from Wadee Alhalabi.

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

References

1. Anagnostopoulos, I.; Zeadally, S.; Exposito, E. Handling Big Data: Research Challenges and Future Directions. *J. Supercomput. Int. J. High-Perform. Comput. Des. Anal. Use* **2016**, *72*, 1494–1516. [\[CrossRef\]](#)
2. Khalil, O.E.M. Artificial Decision-Making, and Artificial Ethics: A Management Concern. *J. Bus. Ethics* **1993**, *12*, 313–321. [\[CrossRef\]](#)
3. Etzioni, A.; Etzioni, O. Incorporating Ethics into Artificial Intelligence. *J. Ethics Int. Philos. Rev.* **2017**, *21*, 403–418. [\[CrossRef\]](#)
4. Wallach, W.; Franklin, S.; Allen, C. A Conceptual and Computational Model of Moral Decision Making in Human and Artificial Agents. *Top. Cogn. Sci.* **2010**, *2*, 454–485. [\[CrossRef\]](#)
5. Wiltshire, T.J. A Prospective Framework for the Design of Ideal Artificial Moral Agents: Insights from the Science of Heroism in Humans. *Minds Mach. J. Artif. Intell. Philos. Cogn. Sci.* **2015**, *25*, 57–71. [\[CrossRef\]](#)
6. Schwitzgebel, E.; Garza, M. A Defense of the Rights of Artificial Intelligences. *Midwest Stud. Philos.* **2015**, *39*, 98–119. [\[CrossRef\]](#)
7. Heersmink, R. The Internet, Cognitive Enhancement, and the Values of Cognition. *Minds Mach. J. Artif. Intell. Philos. Cogn. Sci.* **2016**, *26*, 389–407. [\[CrossRef\]](#)
8. Nohuddin, P.N.E.; Sunayama, W.; Christley, R.; Coenen, F.; Setzkorn, C. Trend Mining in Social Networks: From Trend Identification to Visualization. *Expert Syst.* **2014**, *31*, 457–468. [\[CrossRef\]](#)
9. Tricco, A.C.; Zarin, W.; Lillie, E.; Jeeblee, S.; Warren, R.; Khan, P.A.; Robson, R.; Pham, B.; Hirst, G.; Straus, S.E. Utility of social media and Crowd-Intelligence Data for Pharmacovigilance: A Scoping Review. *BMC Med. Inform. Decis. Mak.* **2018**, *18*, 38. [\[CrossRef\]](#)
10. Vauclair, C.-M. Measuring Cultural Values at the Individual-Level: Considering Morality in Cross-Cultural Value Research. *RAM Rev. Adm. Mackenzie* **2009**, *10*, 60–83. [\[CrossRef\]](#)
11. Hills, M.D. Kluckhohn and Strodtbeck's values orientation theory. In *Online Readings in Psychology and Culture*; Center for Cross-Cultural Research, Western Washington University: Bellingham, WA, USA, 2002.
12. López-Yáñez, I.; Yáñez-Márquez, C.; Camacho-Nieto, O.; Aldape-Pérez, M.; Argüelles-Cruz, A.-J. Collaborative learning in postgraduate level courses. *Comput. Hum. Behav.* **2015**, *51*, 938–944. [\[CrossRef\]](#)
13. Kohler, J.; Neumann, T.A.; Robbins, J.W.; Tronstad, S.; Melland, G. ICESat Elevations in Antarctica Along the 2007–09 Norway–USA Traverse: Validation with Ground-Based GPS. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1578. [\[CrossRef\]](#)
14. Gibson, J.; McKenzie, D. Using global positioning systems in household surveys for better economics and better policy. *World Bank Res. Obs.* **2007**, *22*, 217–241. [\[CrossRef\]](#)
15. Hu, Y.; Yang, Y.; Huang, B. A Comprehensive Survey of Recommendation System Based on Taxi GPS Trajectory. In Proceedings of the 2015 International Conference on Service Science (ICSS), Weihai, China, 8–9 May 2015; p. 99.
16. Montini, L.; Rieser-Schüssler, N.; Axhausen, K.W. Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area. In Proceedings of the 13th Swiss Transport Research Conference, Ascona, Switzerland, 24–26 April 2013.
17. Huang, Y. Research on global positioning system in mobile communication equipment based on android platform. In Proceedings of the International Conference on e-Education, e-Business, and Information Management (ICEEIM 2014), Shanghai, China, 17–18 April 2014.

18. Maryam, A.; Siddiqui, M.; Qadeer, A.; Umar, M.S. Travel management system using GPS & geo tagging on Android platform. In Proceedings of the 2016 Thirteenth International Conference on Wireless and Optical Communications Networks (WOCN), Hyderabad, India, 21–23 July 2016; pp. 1–6.
19. Almanza, E.; Jerrett, M.; Dunton, G.; Seto, E.; Pentz, M.A. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. *Health Place* **2012**, *18*, 46–54. [CrossRef] [PubMed]
20. Duncan, M.J.; Badland, H.M.; Mummery, W.K. Applying GPS to enhance understanding of transport-related physical activity. *J. Sci. Med. Sport* **2009**, *12*, 549–556. [CrossRef]
21. Rodriguez, D.A.; Cho, G.H.; Evenson, K.R.; Conway, T.L.; Cohen, D.; Ghosh-Dastidar, B.; Pickrel, J.L.; Veblen-Mortenson, S.; Lytle, L.A. Out and about: Association of the built environment with physical activity behaviors of adolescent females. *Health Place* **2012**, *18*, 55–62. [CrossRef]
22. Kahtan Aziz, K.; Tarapiah, S.; Ismail, S.H.; Atalla, S. Smart real-time healthcare monitoring and tracking system using GSM/GPS technologies. In Proceedings of the 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, Oman, 15–16 March 2016; pp. 1–7.
23. Wheaton, B.; Nisenbaum, R.; Glazier, R.H.; Dunn, J.R.; Chambers, C. The Neighborhood Effects on Health and Well-being (NEHW) study. *Health Place* **2015**, *31*, 65–74.
24. Kawachi, I.; Berkman, I. *Neighborhoods and Health*; Oxford University Press: New York, NY, USA, 2003.
25. O'Campo, P. Invited commentary: Advancing theory and methods for multi-level models of residential neighborhoods and health. *Am. J. Epidemiol.* **2003**, *157*, 9–13. [CrossRef]
26. Diez Roux, A.V.; Mair, C. Neighborhoods and health. *Ann. N. Y. Acad. Sci.* **2010**, *1186*, 125–145. [CrossRef]
27. Stockdale, S.; Wells, K.B.; Tang, L.; Belen, T.; Zhang, L.; Sherbourne, C. The importance of social context: Neighborhood stressors, stress-buffering mechanisms, and alcohol, drug and mental health disorders. *Soc. Sci. Med.* **2007**, *65*, 1867–1881. [CrossRef]
28. Meyer, O.L.; Castro-Schilo, L.; Aquilar-Gaxiola, S. Determinants of mental health and self-rated health: A model of socioeconomic status, neighborhood safety, and physical activity. *Am. J. Public Health* **2014**, *104*, 1734–1741. [CrossRef]
29. Sastry, N. *The Design of a Multilevel Survey of Children, Families, Communities, the Los Angeles Family and Neighborhood Survey*; Rand Labor and Population Working Paper Series; 2003; pp. 3–21. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.571.226&rep=rep1&type=pdf> (accessed on 14 June 2021).
30. Mujahid, M.; Diez Roux, A.V.; Morenoff, J.; Raghunathan, T. Assessing the measurement properties of neighborhood scales: From psychometrics to ecometrics. *Am. J. Epidemiol.* **2007**, *165*, 858–867. [CrossRef]
31. Zenk, S.N.; Schulz, A.J.; Matthews, S.A.; Odoms-Young, A.; Wilbur, J.; Wegrzyn, L.; Gibbs, K.; Braunschweig, C.; Stokes, C. Activity space environment and dietary and physical activity behaviors: A pilot study. *Health Place* **2011**, *17*, 1150–1161. [CrossRef] [PubMed]
32. Hino, A.A.; Reis, R.S.; Sarmiento, O.L.; Parra, D.C.; Brownson, R.C. The built environment and recreational physical activity among adults in Curitiba, Brazil. *Prev. Med.* **2011**, *52*, 419–422. [CrossRef] [PubMed]
33. Klinker, C.D.; Schipperijn, J.; Toftager, M.; Kerr, J.; Troelsen, J. When cities move children: Development of a new methodology to assess context-specific physical activity behavior among children and adolescents using accelerometer and GPS. *Health Place* **2015**, *31*, 90–99. [CrossRef] [PubMed]
34. Coombes, E.; van Sluijs, E.; Jones, A. Is environmental setting associated with the intensity and duration of children's physical activity? Findings from the SPEEDY GPS study. *Health Place* **2013**, *20*, 62–65. [CrossRef] [PubMed]
35. Bond, M.H.; Leung, K.; Au, A.; Tong, K.K.; De Carrasquel, S.R.; Murakami, F.; Yamaguchi, S.; Bierbrauer, G.; Singelis, T.M.; Broer, M.; et al. Culture-level dimensions of social axioms and their correlates across 41 cultures. *J. Cross-Cult. Psychol.* **2004**, *35*, 548–570. [CrossRef]
36. Bricka, S.; Zmud, J.; Wolf, J.; Freedman, J. Household travel surveys with GPS: An experiment. In Proceedings of the 88th Annual Meeting of the Transportation Research Board, Washington, DC, USA, 11–15 January 2009.
37. Earley, P.C. Leading cultural research in the future: A matter of paradigms and taste. *J. Int. Bus. Stud.* **2006**, *37*, 922–931. [CrossRef]
38. Hofstede, G. *Culture's Consequences: International Differences in Work-Related Values*; Sage: Beverly Hills, CA, USA, 1980.
39. Murakami, E.; Wagner, D.P. Can using global positioning system (GPS) improve trip reporting? *Transp. Res. Part C Emerg. Technol.* **1999**, *7*, 149–165. [CrossRef]
40. Rokeach, M. *The Nature of Human Values*; Free Press: New York, NY, USA, 1973.
41. Schwartz, S.H.; Bilsky, W. Toward a universal psychological structure of human values. *J. Personal. Soc. Psychol.* **1987**, *53*, 550–562. [CrossRef]
42. Schwartz, S.; Bilsky, W. Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *J. Personal. Soc. Psychol.* **1990**, *58*, 878–891. [CrossRef]
43. Wolf, J.; Guensler, R.; Bachman, W. Elimination of the travel diary: An experiment to derive trip purpose from GPS travel data. In Proceedings of the 80th Annual Meeting of the Transportation Research Board, Washington, DC, USA, 7–11 January 2001.
44. Whipple, J.; Arensman, W.; Starr Boler, M. A public safety application of GPS-enabled smartphones and the android operating system. In Proceedings of the 2009 IEEE International Conference on Systems, Man and Cybernetics, San Antonio, TX, USA, 11–14 October 2009.
45. Ma, Y.-L.; Zhu, X.; Hu, X.; Chiu, Y.C. The use of context-sensitive insurance telematics data into auto insurance rate making. *Transp. Res. Part A* **2018**, *113*, 243–258. [CrossRef]

46. Putrenko, V. Data mining of relationship in crowdsourcing projects and social activities of citizens. In Proceedings of the 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 29 May–2 June 2017; pp. 1060–1065.
47. Chang, D.; Lee, C. A product affective properties identification approach based on web mining in a crowdsourcing environment. *J. Eng. Des.* **2018**, *29*, 449–483. [\[CrossRef\]](#)
48. Hoi, C.S.; Khowaja, D.; Leung, C.K. Constrained frequent pattern mining from big data via crowdsourcing. In Proceedings of the International Conference on Big Data Applications and Services, Tashkent, Uzbekistan, 15–18 August 2017; Springer: Singapore, 2017; pp. 69–79.
49. Guo, K.; Tang, Y.; Zhang, P. CSF: Crowdsourcing semantic fusion for heterogeneous media big data on the internet of things. *Inf. Fusion* **2017**, *37*, 77–85. [\[CrossRef\]](#)
50. Lytras, M.D.; Visvizi, A. Big data research for social science and social impact. *Sustainability* **2020**, *12*, 180. [\[CrossRef\]](#)
51. Yoon, A.; Copeland, A. Understanding social impact of data on local communities. *Aslib J. Inf. Manag.* **2019**, *71*, 558–567. [\[CrossRef\]](#)
52. Carpenter, C.J.; Amaravadi, C.S. A big data approach to assessing the impact of social norms: Reporting one’s exercise to a social media audience. *Commun. Res.* **2019**, *46*, 236–249. [\[CrossRef\]](#)
53. West, S.M. Data capitalism: Redefining the logics of surveillance and privacy. *Bus. Soc.* **2019**, *58*, 20–41. [\[CrossRef\]](#)
54. van der Schyff, K.; Flowerday, S.; Furnell, S. Duplicitous social media and data surveillance: An evaluation of privacy risk. *Comput. Secur.* **2020**, *94*, 101822. [\[CrossRef\]](#)
55. Jahanbin, K.; Rahmania, F.; Rahmanian, V.; Jahromi, A.S. Application of Twitter and web news mining in infectious disease surveillance systems and prospects for public health. *GMS Hyg. Infect. Control* **2019**, *14*. [\[CrossRef\]](#)
56. Nouira, K.; Njima, N.B. FluSpider as a new vision of digital influenza surveillance system: Based on Big Data technologies and Massive Data Mining techniques. In Proceedings of the 2020 International Multi-Conference on: “Organization of Knowledge and Advanced Technologies” (OCTA), Tunis, Tunisia, 6–8 February 2020; IEEE: New York, NY, USA, 2020; pp. 1–10.
57. De Vreede, T.; De Vreede, G.-J.; Alawi, N. Achieving Success in Community Crowdsourcing: Lessons from the Field. In Proceedings of the 54th Hawaii International Conference on System Sciences, Maui, HI, USA, 5–8 January 2021. Available online: <https://scholarspace.manoa.hawaii.edu/handle/10125/70695?mode=full> (accessed on 14 May 2021).
58. Cappa, F.; Oriani, R.; Peruffo, E.; McCarthy, I.P. Big Data for Creating and Capturing Value in the Digitalized Environment: Unpacking the Effects of Volume, Variety and Veracity on Firm Performance. *J. Prod. Innov. Manag.* **2021**, *38*, 49–67. [\[CrossRef\]](#)
59. Natalicchio, A.; Messeni Petruzzelli, A.; Garavelli, A.C. Innovation problems and search for solutions in crowdsourcing platforms—A simulation approach. *Technovation* **2017**, *64*, 28–42. [\[CrossRef\]](#)
60. Iandoli, L.; Quinto, I.; De Liddo, A.; Buckingham Shum, S. On online collaboration and construction of shared knowledge: Assessing mediation capability in computer supported argument visualization tools. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 1052–1067. [\[CrossRef\]](#)
61. Jin, X.; Wah, B.W.; Cheng, X.; Wang, Y. Significance and Challenges of Big Data Research. *Big Data Res.* **2015**, *2*, 59–64. [\[CrossRef\]](#)
62. Vossen, G.; Lytras, M.; Koudas, N. Revisiting the (machine) Semantic Web: The missing layers for the human Semantic Web. *IEEE Trans. Knowl. Data Eng.* **2007**, *19*, 145. [\[CrossRef\]](#)
63. Zhuhadar, L.; Yang, R.; Lytras, M.D. The impact of Social Multimedia Systems on cyberlearners. *Comput. Hum. Behav.* **2013**, *29*, 378–385. [\[CrossRef\]](#)
64. Naeve, A.; Yli-Luoma, P.; Kravcik, M.; Lytras, M.D. A modelling approach to study learning processes with a focus on knowledge creation. *Int. J. Technol. Enhanc. Learn.* **2008**, *1*, 1–34. [\[CrossRef\]](#)
65. Spruit, M.; Lytras, M. Applied Data Science in Patient-centric Healthcare: Adaptive Analytic Systems for Empowering Physicians and Patients. *Telemat. Inform.* **2018**, *35*, 643–653. [\[CrossRef\]](#)
66. Lytras, M.D.; Visvizi, A.; Torres-Ruiz, M.; Damiani, E.; Jin, P. IEEE Access Special Section Editorial: Urban Computing and Well-Being in Smart Cities: Services, Applications, Policymaking Considerations. *IEEE Access* **2020**, *8*, 72340–72346. [\[CrossRef\]](#)
67. Zhang, X.; Jiang, S.; Ordóñez de Pablos, P.; Lytras, M.D.; Sun, Y. How virtual reality affects perceived learning effectiveness: A task–technology fit perspective. *Behav. Inf. Technol.* **2017**, *36*, 548–556. [\[CrossRef\]](#)