Review

Computational Social Science of Disasters: Opportunities and Challenges

Annetta Burger 1,2, Talha Oz 1,2, William G. Kennedy 1,2 and Andrew T. Crooks 1,2,3,
1 Department of Computational and Data Sciences, George Mason University, Fairfax, VA 4400, USA; aburger2@gmu.edu (A.B.); toz@gmu.edu (T.O.); wkennedy@gmu.edu (W.G.K.)
2 Center for Social Complexity, George Mason University, Fairfax, VA 4400, USA
3 Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 4400, USA
* Correspondence: acrooks2@gmu.edu; Tel.: +1-703-993-4640

Received: 20 March 2019; Accepted: 23 April 2019; Published: 26 April 2019

Abstract: Disaster events and their economic impacts are trending, and climate projection studies suggest that the risks of disaster will continue to increase in the near future. Despite the broad and increasing social effects of these events, the empirical basis of disaster research is often weak, partially due to the natural paucity of observed data. At the same time, some of the early research regarding social responses to disasters have become outdated as social, cultural, and political norms have changed. The digital revolution, the open data trend, and the advancements in data science provide new opportunities for social science disaster research. We introduce the term computational social science of disasters (CSSD), which can be formally defined as the systematic study of the social behavioral dynamics of disasters utilizing computational methods. In this paper, we discuss and showcase the opportunities and the challenges in this new approach to disaster research. Following a brief review of the fields that relate to CSSD, namely traditional social sciences of disasters, computational social science, and crisis informatics, we examine how advances in Internet technologies offer a new lens through which to study disasters. By identifying gaps in the literature, we show how this new field could address ways to advance our understanding of the social and behavioral aspects of disasters in a digitally connected world. In doing so, our goal is to bridge the gap between data science and the social sciences of disasters in rapidly changing environments.

Keywords: disasters; computational social science; crisis informatics; disaster modeling; Web 2.0; social media; big data; volunteered geographical information; crowdsourcing

1. Introduction

The frequency of disasters is on the rise [1], and projections suggest the risk will increase in the future [2]. However, progress in the field of disaster research continues to be challenged by a multifaceted context with psychosocial, socio-demographic, socioeconomic, and sociopolitical dimensions and associated shifting definitions of what qualifies as a disaster (e.g., [3–6]). These complexities lead to a broad range of questions pertaining to the social, psychological, cultural, political, and economic impacts. What are the interacting causal factors that lead to disasters? Who is vulnerable to disasters? Which factors contribute to their vulnerabilities and to what extent? How can practitioners apply lessons from research to prevent and mitigate disasters? In addressing these questions, traditional social science methods for the collection of social data garnered from interviewing and surveying during or immediately after disasters remains a challenge. Research conclusions in the field are limited due to a paucity (or “unobservability” [7]) of data (which is discussed more throughout this paper and especially in Section 5.1) and the fact that data gathered from disaster events are heavily context dependent and extremely heterogenous [8]. Critiques of the social science side of disasters include
Tierney et al.’s [9] identification of deficiencies in the knowledge base and a recommendation to find more evidence to support widely believed findings and Quarantelli’s [10] work noting the disturbing deficiency of empirical disaster studies and the broad acceptance of empirical generalizations that rely on small or weak datasets. However, fresh methods and data sources are emerging from new technologies in data analysis and computational modeling and the fields of crisis informatics and computational social science (CSS). In this context, the integration of traditional social science theories, innovations in data analysis, and developments in computational modeling offer notable approaches that can address current gaps in disaster research and provide opportunities to advance the field.

In addition to the inadequacy of large, empirically validated datasets, disaster research is affected by continuously adapting environments fed by ecological, social, cultural, political, and technological changes [11] and is addressed by multiple research disciplines. In this regard, some older research findings may no longer be relevant due to changes in society, culture, technology, etc. (as is discussed in Section 2), and traditional disciplines often do not address questions that investigate these adoptions at the intersections of various social sciences. We argue that, while the research literature of disasters is often structured by discipline, understanding the interacting social processes present in disasters is subsequently challenged by disciplinary stove pipes. The purpose of this paper is to review the existing state of the art in disaster studies and relevant disciplines, identify gaps and commonalities, and discuss how computational models and new forms of data analysis can cross over and break down the traditional disciplinary barriers of the social sciences. A literature review was accomplished by leveraging existing domain literature reviews, backwards snowballing, and extensive key-word searches in Google Scholar using English words: disaster, social science, psychology, anthropology, political science, economics, computational social science, and crisis informatics. Relevant literature in the form of books, journal articles, and conference papers were selected to represent the work of three areas relevant to disaster research, as depicted in Figure 1: social science (sociology, psychology, anthropology, political science, economics, computational social science, and crisis informatics). Relevant literature was further culled for exemplar questions of interest, methods used, and theories highlighted. Relevant findings were subsequently compared to identify gaps and commonalities.

As information and communication technologies (ICTs) such as the Internet of Things (IoT), smart mobile devices (including GPS and Bluetooth sensors), and advances in Web 2.0 pervade every aspect of daily life [12–14], they have also become ubiquitous in disaster events (e.g., [15,16]). Coinciding with this is the emergence of big data, innovations in data analysis that are providing us with new ways to explore disasters. Approximately a decade ago, informatics researchers (i.e.,
computer, information, and communication scientists) coined a term to address this aspect of disaster research—crisis informatics. Building upon Kling’s [17] definition of social informatics, we define crisis informatics as the study of the design, uses, and consequences of ICTs in times of crisis. Crisis informatics in this regard approaches behavioral data largely from a technology design perspective and not necessarily for the purpose of studying the underlying social theories that explain the processes leading to observed patterns in disasters. It is primarily interested in designing systems for better disaster management.

Researchers utilizing technological tools who are interested in expanding their work beyond the area of system design can shift their attention from the field of crisis informatics to computational social science (CSS: the study of social science through computational methods). In this domain, they can leverage additional themes and theoretical tools for studying social phenomena in disasters. These include: (1) social information retrieval and data mining, (2) modeling and simulation, (3) social networks and geospatial analysis, and (4) online crowdsourcing and experimentation [18–22]. Not only can CSS provide new data sources and methodologies with the growing availability of information through advances in Internet technologies and the proliferation of the IoT and mobile devices (as is discussed in Sections 3 and 4), it has the potential to bring new theoretical and methodological insights to disaster research (discussed in Section 5). Building on CSS while leveraging what we know of crisis informatics and disaster research, we introduce the computational social science of disasters (CSSD). We define CSSD as an approach to explaining the social dynamics of disasters via computational means by adopting the relevant parts of CSS, social sciences in disaster, and crisis informatics, as depicted in Figure 1. With this approach, researchers can take advantage of the new opportunities in CSSD to advance a better understanding of social phenomena in disasters through a new set of research questions.

In the remainder of this paper, we intend to provide a comprehensive description of computational social science of disasters (Section 5). However, we first provide some background on the three scientific fields with which it overlaps (as depicted in Figure 1). In Section 2, we briefly review the domains and the approaches of each of the traditional social science disciplines to disasters. In Section 3, we describe the other encapsulating field of computational social science. Following this, in Section 4, we discuss crisis informatics and its parent field social informatics, as there have been important developments in these fields that make use of “big crisis data” [23], e.g., social media. CSSD is basically the intersection of these three fields. In the following three sections, we discuss the social sciences in disaster research, computational social science, and crisis informatics that serve as the foundations of CSSD. In Section 5, we discuss the components of CSSD and highlight some exemplar studies that capture certain elements of CSSD along with the challenges and the opportunities it brings to the study of disasters. Finally, in Section 6, we provide a summary of the paper.

2. The Role of Social Science in Disaster Research

The study of disasters is part of many social science disciplines. Although sociology plays a leading role in disaster research and disaster-related policymaking, studies in this field leverage theories and methodologies from many disciplines (e.g., geography, medicine, industrial organization). Conceptually, rather than research derived from one discipline, research pertaining to disasters is popularly understood in terms of phases: preparedness, response, recovery, and mitigation. For example, in the preparedness phase, policymakers work alongside engineers and researchers to improve disaster planning and warning. Within the response phase, emergence is a core theme of disasters (and complexity science more broadly), and it has been a significant topic of research in disaster science from a variety of disciplines [24]. Disaster recovery, a long and multifaceted process, intersects with the domains of various disciplines, including psychology, economics, political science, tourism, and transportation. Finally, mitigation, which has received special attention since the 2000s, is studied by social geographers as well as environmental and sustainability scientists. How X is affected by disasters and what the impact might be if preparedness, response, recovery, and mitigation
factors were varied are questions that could be asked by any discipline studying X, whatever social phenomenon X may be. The century-long history of disaster research tells us that the trans-disciplinary nature of the field has kept evolving over time [25–27].

CSSD is proposed as a subset of the study of social sciences in disasters and is discussed in Section 5. In this section, we provide some background about the foundational findings and the methods of traditional social sciences, specifically sociology (Section 2.1), psychology (Section 2.2), anthropology (Section 2.3), political science (Section 2.4), and economics (Section 2.5). Instead of doing a detailed literature review in each of these disciplines (as it would be beyond the scope of any single paper to address them), we rely on broad reviews and supplement them with additional references as needed. Our summaries of each social science discipline cover an overview, questions of interest, methods used, theory highlights, and relevant findings in the context of this review. Interested readers are referred to the papers cited for more information about the disciplines’ long histories, approaches, and contributions to disaster research. It should also be noted that disasters are important topics in other disciplines such as geography, ecology, and medicine. These were not included in this paper for brevity, but a sample of key reviews is provided [28–32].

2.1. Sociology

How do individuals, groups, and societies behave in disasters and times of crisis? What are the underlying social processes? Under what conditions do behavioral patterns of social solidarity arise? How do these differ from those that lead to social conflict? What roles do gender, race, diversity, or economic inequality play throughout disaster planning, response, and recovery? These are just a small set of the questions sociologists in disaster research address [33], and a large body of related empirical work has been codified in works by Barton, Dynes, and Drabek (e.g., [8,25,33,34]). In 1994, Dynes [25] observed, “sociologists in the disaster area have had a much greater influence in the development of science and public policy than in any other [comparable] area.” The dominant approach of the sociology of disasters has been event-based and integrated systems theory with the realist assumption that disaster existed at the intersection of physical agents or “hazards”, such as earthquakes or tornados and vulnerable people and places [6,35].

The sociological methods of disaster research that have provided the basis of well-understood disaster theories are no different than those of any other sociological enterprise. Phillips [36] outlines four main methods: interviewing, observation, unobtrusive measures (items or traces left behind by people), and visual research through records. In a critique of disaster research methods, Quarantelli [10] recognized a reliance on retrospective and after-action interviewing, rather than systematic field observations, would lead to more reliable evidence. The context of disaster events creates unique methodological challenges, as noted by Stallings [37] and Mileti [38]. Ethical and operational considerations (1) are required to prevent physical and psychological harm to survivors and field researchers and (2) are compressed timelines that prevent adequate time to develop theory, hypothesis, and research instruments. The timetable of disaster events and research schedule is unforeseeable with a high degree of uncertainty surrounding potential subjects and behavioral events.

Two significant literature review papers on the sociology of disasters include Drabek [26], who examined the major contributions of sociology and its methodologies, and Tierney [39], who found traditional disaster research too applied and established that disasters were not distinct events but rather socially constructed by ongoing processes. Early significant findings of sociological research debunked the “disaster myths” that made up much of the cultural frames and media images of disasters, such as themes around social chaos such as panic, shock, ineffectiveness of local organizations, anti-social behavior, and low community morale (e.g., [35,40,41]). In an effort to create an inventory of sociological findings in disaster research, Drabek [8] discussed 146 themes and placed 654 major conclusions of the literature into a typology of system responses in which findings are classified into one of four disaster phases (preparedness, response, recovery, and mitigation) and six social system levels (individual, group, organizational, community, society, and international). Whereas early sociological
research focused on the “event,” describing disasters as a cycle of stability, disruption, and adjustment, current social constructivist approaches (e.g., [42,43]) shifted the concept of disasters towards social causation [39,44]. For example, Hurricane Andrew and the Chicago Heat Wave of 1992 were not isolated events caused by extreme weather; rather, they were socially constructed by social and economic processes that led to inequalities and created vulnerable populations.

Collective behavior, social control [45] symbolic interactionism [46], and emergent social behavior [47] have been among the popular theoretical orientations in disaster research. In his discussion of “social science research agenda for the disasters of the 21st century,” Quarantelli [10] found the earlier accounts narrow and suggested five formulations relevant to disaster research: attribution theory from social psychology, satisficing theory from organizational theory, diffusion studies, network theory, and social capital. Attribution theory and satisficing theory can be applied to decision-making in the context of disasters, and diffusion studies, network theory, and social capital could help provide explanations for behavior arising from social relationships.

2.2. Psychology

Just as sociology of disasters reflects the qualitative nature of sociological studies, the disaster research conducted by psychologists is mostly quantitative, as it is the common methodology in the field of psychology. The questions are formed to understand the human mind regarding preparation for and response to disasters. What leads some people to be better prepared for disasters than others? How can disaster preparedness be encouraged? How does disaster affect the mental health of individuals and their broader community? Psychology literature on disasters can be classified into two, preparedness for risk reduction and post-disaster psychopathology. The latter can be further categorized into four topics: (i) empirical predictive (predicts contributions of variables), (ii) empirical epidemiological (describes incidence at population level), (iii) clinical descriptive (identifies symptoms found in disaster victims), and (iv) clinical intervention (describes effectiveness of different intervention approaches) [48]. Methodology in psychology is aimed at identifying and testing the underlying mechanisms of people’s behavior and mental health. In disasters, these methods include a combination of screening and diagnostic reports and correlate a variety of psychosocial measures, such as insomnia, perceptions of safety, and changes in the ability to function. These are gathered through observation, interviews, and questionnaires, and they are integrated into structured experimental studies [49].

Several meta-studies reviewed this literature and highlighted the major findings. Rubonis and Bickman [48] examined the relationship between four sets of variables (the characteristics of the victim population, the characteristics of the disaster, the study methodology, and the type of post-disaster psychopathology) by reviewing 52 studies. In a similar effort, Norris [50] and Norris et al. [51] reviewed the post-disaster mental health problems and risk factors in 225 disaster samples (from 132 distinct events experienced by 85,000 individuals) quantitatively studied in the psychology literature. Rubonis and Bickman [48] found a small but positive relationship between disasters and psychopathology, and Norris [50] found post-traumatic stress disorder (PTSD) to be the most common problem occurring in post-disaster studies. Norris et al. [51] found that, among the adults they sampled, factors such as “more severe exposure, female gender, middle age, ethnic minority status, secondary stressors, prior psychiatric problems, and weak or deteriorating psychosocial resources most consistently increased the likelihood of adverse outcomes” [51], while for the youth, family factors had the greatest effect [51].

More recently, Ejeta et al. [52] identified the most common behavioral theories and models applied to disaster preparedness. Reviewing 33 articles on preparedness (including preparedness for disease outbreak, flood, and earthquake hazards), Ejeta et al. [52] found that the most common theories applied in the literature are the health belief model (HBM), the extended parallel process model (EPPM), the theory of planned behavior (TPB), and the social cognitive theories (SCT). In these studies, the main constructs of HBM (perceived susceptibility, severity, benefits, and barriers), EPPM (higher threat and higher efficacy), TPB (attitude and subjective norm), and SCT (cognitive, affective, emotional, and social influences) have been associated with disaster preparedness. However, they also noted the theories
were predominately applied to natural hazards and diseases, not man-made hazards. In dealing with the effects of disasters, the review of the resources by Norris et al. [51] found that theories on coping strategies (active outreach, informed pragmatism, reconciliation), beliefs (higher self-efficacy and optimism), social support (social embeddedness, received social support, and perceived social support), and conservation of resources (including objects, conditions, personal characteristics, and energies) help explain the moderators and mediators of psychological effects of disasters. Overall, the psychological literature, while focused on preparedness and post-disaster psychopathology, has also been limited for some uses due to its lack of application to man-made disasters.

2.3. Anthropology

By approaching disaster research with holistic and comparative perspectives, anthropologists study all aspects of human life—environmental, biological, and socio-cultural—as they relate to disasters. Their work focuses on the interconnections between cultural, social, political, economic, and environmental domains to provide explanations for cultural systems in disaster. Anthropologists ask questions such as how do people and cultures understand disaster? How does culture drive socio-cultural processes and responses to disasters? How do these processes interact with the corresponding physical and technical processes? Anthropological studies cross scales from the local to global and back; they explore not only the external physical relationship between human and environment but also the internal meaning that humans produce to understand and interpret their experience. They unravel long-term processes of cultural adaptation to changing social and physical environments as revealed in archeology and history, and reveal power dynamics in the social structures of individuals and groups.

As a result of these analyses, anthropologists have uncovered complex interactions between physical, biological, and sociological systems [53,54] that involve people’s adaptations to and manipulations of their physical environment and construction of sociocultural institutions, beliefs, and ethos. As part of a social process, these interactions produce disaster, the event that involves a potentially destructive natural or technological agent and a population under varying conditions of vulnerability [4,53–55]. Anthropological work has shed light on the social production of disasters and the social structures that contribute to vulnerability and risk [4]. Theories of “embodiment” have contributed to a better understanding of how culture affects individuals experiences, along with how they comprehend and cope with traumatic experiences (e.g., [56–60]). Comparative work on multiple cultures has illuminated how different societies respond and adapt to environmental changes [61] and disasters with responsive belief systems and coping strategies [62]. Longitudinal studies have shown how societies cope and adapt through multiple disasters (e.g., [63]).

Since disasters affect every feature of society as well as its relations with the environment and its individuals and communities, anthropology’s holistic approach uniquely qualifies the field to study the processes of disaster and interactions that cut across domains. We can look at the findings of anthropological research temporally, studies explaining processes in pre-, (early) response-, and post-disaster phases. Anthropology has given special attention to structural conditions of pre-disaster vulnerability, such as gender inequality, global inequities, endemic poverty, racism, a history of colonial exploitation, imbalances of trade, and underdevelopment [56], and set them in the context of historical processes [64]. In responses to disasters, the themes that have been studied include changes occurring in cultural institutions (e.g., belief systems), within political organizations (i.e., power relations between individuals, the state, and international actors), and within economic systems (e.g., allocation of resources). For the post-disaster phase, a great deal of anthropological work criticizes various actors, including the relief programs, for their top-down, non-flexible strategies in which the affected populations are overlooked (which is in line with the sociological findings), or how media becomes a contested space in which actors try to control the narrative, especially in times of uncertainty [55]. As local and international communities wrestle with issues of environmental change, adaptation, and disaster mitigation, the work of anthropology provides examples of how indigenous
and local knowledge can substantially contribute to solutions for community disaster risk reduction and resilience [65–67].

What differentiates anthropology from other social disciplines in disaster research is not only its emphasis on cultural comparison, but also the qualitative, contextual data gathered in the ethnographic methods, such as from interviews, longitudinal participant observations, and linguistic analysis. These contexts of disaster reveal the heterogeneity of disaster experiences in multiple realities and decision-making rationalities. With its holistic approach, the field has the potential to fill methodological and theoretical gaps between the intersecting disciplines that study disaster. In practice, its bottom-up approach balances top-down biases in emergency management and enables the incorporation of local technical knowledge, insight, skills, and needs [56,68]. Conversely, the challenge of this field is that the complex and context-rich studies can become so specific in culture and context as to limit them for general application. Anthropology has also been criticized for privileging local knowledge and problematizing the dominant modes of relief efforts [56].

2.4. Political Science

Political scientists were not present at the foundation of the modern disaster research field, and many were reluctant to study disasters because they viewed disasters primarily as engineering problems, or they maintained the widely held moral stance that there should not be a “politics of disaster” [69]. Others such as Olson [69] argued that disasters are intrinsically political events. Do disasters foster cooperation or conflict? In which condition is one or the other manifested, and why? Although these questions were asked earlier by sociologists [70], more recently, political scientists in the conflict resolution and international relations fields started to investigate it with the greater amount of data that have been collected over the recent decades. Many of the political science studies in disaster research have been quantitative in methodology, and a typical study statistically analyzes decades of data on natural disasters, the incumbents’ preparedness and response, and election returns (e.g., [71,72]). Disaster research can be grouped in four subfields of political science: electoral behavior, conflict resolution, international cooperation and humanitarian aid, and political economy [73]. We discuss the first three here and review the political economy aspects of disasters in Section 2.5 under economics.

Elections are proxies for how voters judge incumbent politicians in preparedness and response to disasters, and they are an important factor in the field of electoral behavior (e.g., [74]). At times, politicians are either viewed as merely ineffective in coping with disasters or as causing the disasters. Connelley [75], in a character study of a senator, said the reason for his losing the election was his depiction of a natural disaster: “he couldn’t make it rain, and now we’ve got him down!” Attribution of responsibility is known to be a key issue in political decision-making as Iyengar [76]. Additionally, blame—which is likely to occur in response to disasters—carries far more weight in voting behavior than that of credit [74]. Gasper and Reeves [77] found a negative relationship in the U.S. between disaster damage and the share of incumbent votes for presidents and governors. Another study in this line, which reflects upon citizen competence and government accountability, shows that “voters reward the incumbent presidential party for delivering disaster relief spending, but not for investing in disaster preparedness spending” [78].

In addition to these theories of electoral behavior, theories on conflict resolution are also tested and developed by political scientists. One study on earthquakes argued that disasters increase scarcity of resources, which subsequently provoke frustrations that lead to anger and violence [71]. Some recent studies statistically showed a link between natural rapid-onset disasters and the likelihood of conflict and rebellion (e.g., [71,72]). Nel and Righarts [72], while investigating the impact of natural disasters on civil war, found that “natural disasters significantly increase the risk of violent civil conflict both in the short and medium term” [72]. In another study looking at the root causes of conflict in climate-related disaster, Peregrine [79] used archeological evidence and found an increase in conflict only when leaders tightly controlled access to political authority, such as when using violence to secure support. Others studied more basic dynamics behind conflict behaviors. In this respect, studying the repression
dynamics following rapid-onset natural disasters, researchers showed that repression is likely to increase after a disaster, but inflows of aid reduce its intensity [80].

The politics of humanitarian aid and disaster response in the international community involve both the political interests of particular governments, such as U.S. foreign disaster assistance [73], and the need to act cooperatively across traditional sovereign boundaries in international disaster assistance [81]. Political considerations may explain half of all federal disaster relief in the U.S. [82] and may determine whether a president decides to issue a disaster declaration [83]. The conflicting priorities of varying stakeholders often lead to aid policies that create subsequent disasters [84,85]. The need for international disaster cooperation and assistance can arise from civil war and failed states, such as famine in Ethiopia or Africa, or from natural disasters that cross boundaries, such as cyclones and drought. Given the challenges of human-caused climate change, it is arguable that much of today’s extreme weather disasters are the result of failed political and economic systems.

2.5. Economics

The economic impact of disasters and incentives for preparedness and response are two major areas of disaster studies. How do disasters affect state and local economies? What are the economic tradeoffs between instituting policies for economic growth versus those for disaster risk mitigation? How is the overall (economic) vulnerability of a population estimated? How can the macroeconomic resilience to disasters, i.e., the ability of an economy to cope with disasters, be measured? These are a small subset of the questions that economists ask regarding disasters and their economic impacts [86].

Traditional social research methodologies exploring the economic impacts of disasters include surveys, global, state, and local measures of GDP, and market and employment reporting. Examining basic economic indicators from a number of economic literature review papers, Kellenberg and Mobarak [87] found that natural disasters have significant impacts on short- and long-term gross domestic product (GDP), social and human capital, and the labor and real estate markets. A more recent 90-year study of U.S. disasters found that severe disasters do adversely impact economies, but milder disasters have little effect based on measures of out migration, housing prices, and poverty rates [88]. Economists looked at the impact of disasters and found varying effects on specific labor markets (see [89,90] for examples). Of interest to economists are the risk profiles of countries and which would most benefit from disaster risk management policies such as strengthening institutions and building standards, improving insurance markets, reducing corruption, and instituting more advanced warning and emergency response systems [87]. The risk insurance industry and derivative markets have been significant areas of study arising from hazards and risk reduction research (e.g., [91–94]). To find out the extent to which disasters affect the economy of a country, Albala-Bertrand [95] examined the effects of disasters on the growth rate of output of six countries by means of a quantitative macroeconomic model and found that “foreign and public disaster response may be better used to help actual victims and affected activities directly than to proceed on the rather unsound prima facie belief that the economy will be heavily affected by the disaster.” Other economists have observed the increasing costs of disasters and studied how improvements to international aid for disaster victims could help protect people or improve economic outcomes [85,96,97].

While many economists have made analyses across multiple disasters, others have developed new economic measures for disasters or have been drawn to specific types of disaster or a case study of disaster, such as Hurricane Katrina. Zahran et al. [98], for example, developed quantitative measures of the mental health impacts of Katrina to explore the economics of disaster risk, social vulnerability, and mental health resilience. Yang [99] explored the impact of hurricanes on the global economy through changes in international financial flows (i.e., financial aid and migrant remittances). Hurricane Katrina has also been studied to find economic reasons for governmental failure in disasters and to measure the socioeconomic costs of disasters. For instance, Shughart [100] expanded the forms of empirical evidence used in identifying the political and economic failures that led to Hurricane Katrina: (1) maintaining existing infrastructure was cheaper than renewing the levees, (2) unlike
private corporations, politicians, and bureaucrats have weak incentives, and (3) public policies such as promises of grants, loans, tax breaks, low-interest loans, and insurances had unintended consequences. In another Katrina related study, using U.S. Center for Disease Control’s (CDC’s) Behavioral Risk Factor Surveillance System (BRFSS) database, Zahran et al. [98] investigated the relationship between individual exposure to hurricanes and poor mental health days and evaluated the economic costs of mental health days on focal populations. Their calculations showed that natural disasters regressively punish disadvantaged population strata [98]. To address pressing problems of dwindling resources arising from environmental change, economic resilience measures are also being developed and introduced for the study of disasters with works by Xie et al. [101] and Rose [102].

2.6. Summary

Traditional social science studies of disasters have provided the foundation of our understandings of disaster, and they continue to contribute research findings and increase our knowledge base. Specifically, we show how sociological research (Section 2.1) has been primarily qualitative, exploring social organizational behavior scaling from individuals to global institutions, and temporally ordered by four main phases of disasters (i.e., preparation, response, recovery, and mitigation). We show how psychological research (Section 2.2) skews more towards quantitative data and is focused on the individuals and the theories applied to preparedness, health, planned behavior, and psychological impacts of disasters. This later topic includes coping strategies, beliefs, social support, and the uses of resources to moderate psychological effects, i.e., post-disaster psychopathology. Anthropology (Section 2.3) confirms the sociological temporal phase approach to disasters with special attention to the structural conditions resulting in vulnerabilities and organizational responses to responses to disasters. Political science (Section 2.4) examines the local, state, and international politics of disaster and comparatively focuses less on the collective preparation for or mitigation of disaster effects on populations. Finally, the economic study of disasters (Section 2.5) focuses on the economic effects of disasters, examining basic economic measures such as GDP, risk management policies, global financial flows, and financial policies. Unfortunately, progress in these sciences has been constrained by their respective disciplinary approaches and methodologies that cannot manage the quantity of events and data available for collection and study in disasters nor fully address the social and physical interactions that cross scales and boundaries. Additionally, while approaches such as case studies allow for in-depth analysis of these events, they provide limited confirmation of theory and are not generalizable to all events. In the next session, we discuss computational approaches applicable to disasters, starting with the general field of computational social science.

3. Computational Social Science

Lazer et al. [18] characterized CSS as an emerging field “that leverages the capacity to collect and analyze data at a scale that may reveal patterns of individual and group behaviors.” Computational social scientists educate themselves in how to use and develop computational methods to address social science inquiries in the most effective ways, and CSS introduces new opportunities for collaboration to study the problems of social processes that cut across disciplines. We show in Section 2 how each social science field has its own sets of questions and preferred methods to address them. Paired with the foundational work of social science disaster research, the new methods in the computational social scientist’s toolbox, such as computational modeling (as is discussed below), coupled with new types of datasets and corresponding analysis techniques that are now available (e.g., social media, crowdsourced data, digital data at large, machine learning algorithms, and social network analysis) make CSS a uniquely valuable field in addressing the more complex problems of disaster research. We would argue it is the integration of the variety of new datasets and computational analysis tools and modeling under the umbrella of CSS that strengthens the processes of developing and testing social theories. We discuss CSS in four main areas: automated information retrieval and open platforms (Section 3.1), social complexity and simulations (Section 3.2), social networks and geospatial analysis.
Future Internet 2019, 11, 103

Section 3.3), and online crowdsourcing and digital field experiments (Section 3.4). For a greater discussion of CSS, readers are referred to [18–21,103,104].

3.1. Information Retrieval and Open Data Systems

Advances in processing technologies have made automated information retrieval standard practice in the social sciences, and these technologies can be used to detect social, behavioral, or economic patterns. In this area, information extraction algorithms are used to collect data from disparate sources, such as census records, economic data, newspapers, and social media, and to conduct data mining and content analysis of verbal data, such as interviews, speeches, and legislative testimony [19,105]. Information retrieval was traditionally defined as “finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers) [106]”. Computational social scientists engage in this activity by collecting and analyzing any digital traces that potentially address their social science inquiries (such as elections and international relations, e.g., [107–110]). There are technical challenges in this realm that include evaluation of item similarities, data scalability, and time sensitivity [111]. Salganik [112] comprehensively explores the characteristics of digital data, including its strengths (voluminous, always-on, and non-reactive) and its weaknesses (incomplete, inaccessible, non-representative, drifting, algorithmically confounded, dirty, and sensitive).

To leverage new capabilities in information retrieval, many governments and companies are adopting open data policies that allow researchers to access and study these social data. Prominent examples of such initiatives include Data.gov in the U.S. and OpenKenya [113] in the Republic of Kenya (see [114] for how such data can be used). The ever-increasing popularity of social media, enabled by Web 2.0 technology, is expanding the sources and volume of social data relevant to our daily lives through applications such as Facebook, Twitter, and Instagram, and these open sources are allowing researchers to explore a vast range of topics, including opinions during elections [115], opinions on public health [116], data on disease outbreaks [117], and studies of the connections between people and places [118,119]. The relevance and management of open-source data has become more important than ever, and they are well-positioned to support the quantitative study of disasters through the use of new computational methods, such as machine learning, natural language processing, sentiment analysis, and artificial intelligence [120].

Large-scale social data harvested from a variety of sources can be classified as part of a general category of “observational social data,” and these data vary depending on researchers’ interests and approaches. Their uses include identification of characteristics or patterns by quantitative and qualitative descriptions of individuals or groups, development of macro-level mathematical models of dynamics in data aggregates, identification of statistical relationships between variables and outcomes, examination of the emergent patterns on the aggregate level, calibration of parameters in computational simulations, inference of social events, and forecasting social phenomena (see [117,121–124] for a range of uses of such large-scale data).

3.2. Complexity and Simulations

CSS is primarily interested in better understanding social phenomena, and it builds on a foundation of existing social science paradigms. Two of the more salient of these paradigms are social complexity and social simulations. Social complexity is a conceptual framework for understanding the increasingly complex interactions of individuals and societies as they interact and adapt to each other and their environment [104]. A complex system is a system of subsystems (i.e., modules or components) whose intra-dependency is much stronger than inter-dependencies [125]. Complex systems can be characterized by distributional or statistical laws—in particular, power laws—and computational tools and new computer language packages, such as those in Python and R, that have made these systems tractable for analysis by a new generation of social scientists. Beyond simple description, power laws and computational tools provide new theoretical perspectives of social phenomena, including
self-similarity, scaling, fractal dimensionality, emergence, self-organized criticality, meta-stability, long-range interactions, and universality [104].

In investigating complex adaptive social systems, a promising set of modeling simulation tools known as social simulations has emerged and is often called the third way of conducting social science research [126]. Social simulations and computational models not only allow for discovery of the consequences of theories in artificial societies, but by enforcing formalization in terms of coherent programs, they play a similar role in social sciences as mathematics does in physical sciences [127]. Examples of this technique applied to disasters include [128–131]. Simulation is an alternative to common static modeling approaches in social sciences. Instead of modeling the interactions among variables, the social life is modeled by interacting adaptive agents [132] in an artificial world. In this area, the generative social scientist asks, “How could the decentralized local interactions of heterogeneous autonomous agents generate the given regularity? [133]”. Researchers in many social sciences use a wide variety of techniques in social simulations, including agent-based modeling, discrete event simulations, systems dynamics, microsimulations, and cellular automata (see [127,134] for reviews).

Agent-based modeling has become a dominant way of producing social simulations [135]. A distinct advantage of this technique is that it provides the ability to explicitly couple autonomous agents with geographical information when space is relevant to the simulation (e.g., [134,136]). Modeling people and their social systems is not without its challenges (such as dynamic systems, multivariate causation, and validation) [137], but agent-based models (ABMs) can operationalize the characteristics of social complexity, such as heterogeneity, autonomy, explicit space, local interactions, and bounded rationality, in controlled experiments [133]. Since agent-based modeling is well-suited to the object-oriented programming paradigm, they can be easily implemented in any object-oriented programming language, such as Java or Python. Moreover, several programming libraries and frameworks have been developed to facilitate the implementation of ABMs, such as NetLogo [138], MASON [139], GAMMA [140], and MESA [141]. In Section 5, we discuss in greater detail how simulation and these modeling techniques are also applicable to the field of CSSD.

3.3. Social Networks and Geospatial Analysis

Social network analysis and geospatial analysis are two promising computational methods for studying the structures of social and physical spaces within which humans live and interact, especially with the advent of crowdsourced information and the rise of social media [142,143]. Social network analysts see the social world as structured by a web of connected agents tied together by specific relationships [144]. Sociograms, which are graphical depictions of social networks, make social structure visible and tangible. Their representation in matrices and other visualizations on computers allow researchers to examine the properties of large networks. There are different relational types of networks, including directed (digraph), signed (valued), weighted, or multiplex, each of which reflect characteristics of social relations. The levels of analysis of networks at different levels (nodal, dyadic, triadic, n-adic, or at network-level) provide insights about the structure of the social system under study (e.g., on concepts, measures, and properties). Examples of these forms of data analysis are found in disaster studies such as [145–147], which explore the structure of organizational networks and the access and flow of information within networks. Other applications of social network analysis include human cognition and belief systems, decision-making models, organizations and meta-models, supply chains, diffusion, and international relations (for a brief review of these, see [104]).

Geography adds an important dimension to human interactions with their environments. Humans do not live in a spatial vacuum, and social reality is heavily dependent on spatial features. The gap between geography and social sciences is addressed by geographers [21,148]. Developments in geographic information systems (GIS), especially in the fields of spatial databases, positioning technologies, remote sensing, and geo-visualization, have made GIS a common tool in criminology, archaeology, public health, anthropology, economics, demography [21], and disaster research [149]. More importantly, we should note that GIS is not simply a set of technological tools; it brings “spatial
thinking” to the social sciences [148] in the form of geographical information science [150]. For example, Hu et al. [151] developed a technique for grid-based tessellation of space that provides a systematic approach for prioritizing areas needing to be mapped by digital volunteers based on information value theory [152]. In this regard, geography has both benefitted from and contributed to computational social sciences [21].

3.4. Online Crowdsourcing and Field Experiments

Internet technologies have opened new frontiers in social collective action and knowledge and the gathering of scientific data in the forms of online crowdsourcing and digital field experiments (see [153–157] for reviews). Crowdsourcing can be understood as the leveraging of information technologies for individual participation in collective processes [158], such as crowdfunding, mapping applications such as Waze, and citizen science data collection efforts (e.g., Christmas Bird Count [159] and Geo-Wiki [160]). Internet platforms such as Waze, Airbnb, and Ushahidi’s [161] crisis mapping applications aggregate individual action and knowledge with computational tools that enable individuals to address ongoing social problems. A significant new set of tools in the hands of computational social scientists are micro-tasking sites, such as Amazon Mechanical Turk, that provide a virtual environment for social science experiments.

Experimentation is the primary means for establishing causal relationships, and the cyber world is providing new opportunities and challenges for researchers to conduct large-scale experiments [156]. On the one hand, the “field” of the experiments, i.e., the Internet, narrows down the scope of the interventions only to those applicable in the cyber world, and it limits the ways subjects can be tracked. On the other hand, the increasing variety and prevalence of web applications in daily social life allow experiments with larger and more diverse subject pools in a shorter period of time and with greater participation. Researchers from different fields have conducted both field- and lab-like experiments in cyberspace to test the effects of controlled or natural interventions using various social computing platforms (for a recent survey of online field experiments, see [157], and for lab-like experiments, see [162]). In disaster studies, they have been used to look at crisis communication and emotions in Utz et al. [163], public behavioral responses to disaster information provided online in Liu et al. [164], and purchasing behavior post-Fukushima nuclear accident by Miyata and Wakayatsu [165]. Mao [166] provides examples of how experimental approaches in studying social computing systems can improve the design of such systems and advance our understanding of human behavior in crowd-tasking activities during crisis mapping. The online platforms available for “field” experiments include micro-tasking sites (e.g., Amazon Mechanical Turk [167]), question and answer sites (e.g., Stack Overflow [168]), collaborative encyclopedias (e.g., Wikipedia [169]), social networking sites (e.g., Facebook [170]), e-commerce sites (e.g., eBay [171]), massive open online courses (e.g., Coursera [172]), sharing economy sites (e.g., Airbnb [173]), dating sites (e.g., OkCupid [174]), massively multiplayer online games (e.g., World of Warcraft [175]), or other platforms over which experimenters can exert greater control (e.g., their own sites [176]). Technologies used for interventions in these experiments include emails with different contents [177], websites with different looks [178], bots with different strategies (server-side scripts) [169], and browser extensions with different pop-up behaviors (client-side scripts) [179].

In this section, we briefly review the computational methods developed and used to support social science inquiries with new techniques in computational models and data analysis. We also discuss how CSS provides open-source data, new theories of decision-making, social processes of aggregated behavior, complex adaptive systems, and spatial and network structure, and new experimental methods in online field experiments and social simulations. We now shift our focus to an area where computational methods intersect with disaster research and practice but do not necessarily address traditional social science research questions.
4. Crisis Informatics

The application of new computational methods to traditional fields of science has spawned numerous computational branches, such as digital anthropology, computational linguistics, and biometrics. As its name suggests, crisis informatics is a subfield of informatics, in particular, of social informatics, and can be defined as the study of the design, uses, and consequences of information and communication technologies in times of crisis [10]. As ICTs, IoT, and social media pervade every aspect of our lives, crisis informatics has increasingly become a critical tool in disaster preparedness, response, and recovery [180]. Additionally, information management problems and ineffective use of these technologies have been cited as major factors for failures in disaster management [181].

In the early days of crisis informatics research, many studies employed qualitative methods for both data collection and processing. Researchers manually monitored the ICTs, and they manually curated and classified the information. These studies could be considered mostly descriptive and used formative and interpretivist forms of inquiry (e.g., [182–184]). Recent advancements in computational and mobile technologies, the open-source culture, adoption of open data policies by companies and governments (e.g., [185]), and the popularity of social media platforms have made studies in crisis informatics both qualitative and quantitative, but the more recent studies of crisis informatics have been computational (e.g., [23,186,187]). Palen et al. [182] completed an ethnographic study of a human-induced crisis to understand what aspects of ICT were used, when they were used, and how they were used in the days following the 2011 Virginia Tech shooting event. In that study, Palen et al. [182] conducted 56 on-site, one-on-one, face-to-face interviews and manually monitored online activities of interviewees on social media sites including Facebook, Wikipedia, and Flickr. As a subset of the study, several Facebook groups as well as Wikipedia editors participated in an online, collective problem-solving task to build the list of victims before Virginia Tech officially released the names. The study found that no single online community group was able to come up with a complete list of victim names. Additionally, none of the online lists had false positives, i.e., people incorrectly listed as victims [182]. Another research effort [183] completed a qualitative longitudinal analysis of six disasters as documented by Flickr postings, which was the most popular photo sharing platform at the time of the study. Among the findings was norm development through finding group purpose or tagging nomenclatures as features of photographic contents were compared, categorized, and discussed. Panteras et al. [188] used triangulation techniques with place names paired with geo-location information in tweets and Flickr to delineate the extent of a wildfire, and Hagen et al. [189] used network analysis to identify distinct communities and influential actors from Zika-related tweets.

When we look at the definitions in the literature, we see that the focus of crisis informatics has been on the design and development of ICTs. Crisis informatics:

- “includes empirical study as well as socially and behaviorally conscious ICT development and deployment [182],”
- “strives for socially and behaviorally informed development of ICT for crisis situations [181],”
- “investigate[s] socio-technical interactions that occur during times of extreme crisis with an eye towards developing ways to support the mitigation of suffering [190],” and
- “is dedicated to finding methods for sharing the right information in a timely fashion during significant crises [191].”

The theoretical foundations of crisis informatics can be found in social informatics and even earlier in socio-technical systems [192]. Social informatics goes back to the 1980s, when research interests were primarily focused on the impact of computerization on the quality of work [17,193]. Social informatics itself is a subfield of socio-technical systems and is concerned with the relations between social and technical systems [194]. The field of socio-technical systems originated in 1950 from interest in optimizing the productivity of postwar industries and at a time when organizations started to be seen not only as social systems but also as technical systems [194]. We can say that crisis informatics is a study of socio-technical systems that can be used in times of disasters.
Crisis informatics researchers develop new technology capabilities as information and communication technologies advance and needs in disaster preparation and response practices emerge. ICT for disasters can be developed for use by digital volunteers (on- and off-site citizens, see Section 3.1) to allow them to crowdsourcing (i.e., micro-task) productively, as well as by disaster managers (formal response agencies) to provide them with contextual information to improve decision-making. This aspect of crisis informatics also enables effective coordination and collaboration between emergency responders and digital volunteers [15,192]. For example, the lack of existing road information prior to the 2010 earthquake in Haiti complicated the disaster response, but it also motivated citizen volunteers to use crowd-sourcing applications to share and update road information, as it was encountered on the ground. The large number of citizen volunteers resulted in Haiti becoming one of the best mapped road networks [129]. To fully assess and realize the potential of these technologies in times of disasters, Hughes and Tapia [192] comment that crisis informatics researchers must first understand the ways individuals and organizations “collect, organize, manage, access, share, coordinate, and disseminate information within communities during crisis situations [192].” Understanding how victims, managers, and volunteers obtain and use information constitutes a significant part of crisis informatics.

Crisis informatics as an inherently digital method is continuously incorporating innovative computational methods. Castillo’s [23] book, “Big Crisis Data”, focuses on methods “for processing social media messages under time-critical constraints.” While Castillo [23] focused on computational methods, Meier’s [15] book, “Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response”, discussed crowdsourcing, the interplay between human curators (of satellite and aerial imagery, social media, and text messages), ICT, and the use of artificial intelligence in times of disasters. Imran et al. [187] also reviewed computational methods and applications for social media data retrieval and processing in the crisis informatics literature. Others (e.g., [186,195]) have discussed the history and the future of crisis informatics and provided a taxonomy of crisis analytics. The field is rapidly growing with continuing improvements to computational techniques (see [196–198] for examples). Of course, these new methods and new big data have not been immune to criticism. Spence et al. [199] addressed the challenges of social media use for collecting data related to a disaster event and the drawing of conclusive inferences from user generated content. The growing number of studies and reviews reflect an increasing interest in the application of crisis informatics to disaster events, particularly in preparedness and emergency response [200].

Interest in integrating crisis informatics with the implementation and use of decision support systems in times of crisis is reinvigorating the application of decision support system research to the disaster field. This is due not only to the growth of data availability and its near real-time nature but also to advancements in decision support systems that increasingly allow the application of knowledge management tools for tactical, operational, and strategic decision-making [201]. Such systems have a long history in urban planning and disaster management (see [202,203] for reviews). For example, systems have been built to aid decision-making during cyclones [204], floods [205], earthquakes [206], evacuations [207], disaster relief [208], and distribution [209,210] more generally. Advances in these areas of data collection and analysis linked to decision support science provide practitioners in the field of disaster and emergency management with not only basic real-time information but also actionable tactical, operational, and strategic knowledge for improved planning and response. As the field evolves, practitioners are beginning to promote necessary conversations among stakeholders to develop standards for best practice, tools, limitations, and ethics of using social media [211,212].

In crisis informatics, the emphasis is on technology, computational methodologies, and data applications rather than explanation and theory. While the field has provided a wealth of new data and analysis to the study of disasters and applications in disaster preparedness and emergency management, it has not put these advantages to use in the advancement of disaster theory. We now discuss how a new field of CSSD can close the gaps between the social sciences of disaster and the computational fields of CSS and crisis informatics.
5. Computational Social Science of Disasters

Empirical and theoretic understandings of disaster can be found at the intersection of social science, computational social science, and crisis informatics research in a combination of theories of social processes, complex adaptive systems, and the information and application of socio-technical systems, as is shown in Figure 1 and discussed in the previous sections. As a subset of computational social science, CSSD brings these domains together in the study of social and behavioral aspects of disasters and related phenomena via computational means. We can now formally define CSSD as the systematic study of the social behavioral dynamics of disasters utilizing computational methods. Computational social scientists and researchers in crisis informatics who are interested in disaster research should draw from and build upon the large body of work in sociology [8] and the other social sciences discussed in Section 2. As has been argued (e.g., [9,213]), there is a need for more evidence to support the social science findings from past quantitative studies of disasters. Through the lens of CSSD, disaster researchers can integrate new computational techniques, methodologies, and theory, which can then be used to test current understandings, develop new theories, and update policy recommendations with respect to disasters. In the remainder of this section, we sketch out how CSSD appears in practice and offer recommendations for paths forward.

The social science of disasters, CSS, and crisis informatics, as demonstrated in our review (Sections 2–4), have no distinctive ontologies, but each has a set of preferred methodologies to address their discipline’s research questions or goals. These methods are designed to bound each discipline’s research questions into tractable hypotheses for testing. However, these practices also isolate the disciplines into silos that are no longer able to address and test the interdependent, nonlinear processes that cross disciplinary domains. The social science of disasters remains largely dedicated to qualitative research and is thus unable to manage the wealth of new data in ICTs and big data or to quantitatively test the complex, nonlinear social processes evident in disaster events. Computational social science and crisis informatics theories and techniques provide data and tools for explaining underlying processes and predicting the outcomes of disaster events utilizing advances in ICTs; however, we would argue their work often provides only a superficial theoretical underpinning (or is disconnected) compared to that found in the traditional social sciences of disasters. In this respect, Palen and Anderson [200] find the marginalization of social science fields within the data science community as troubling. The marginalization of the social sciences is one of the key reasons we feel the need to define CSSD and highlight its potential for advancing disaster research. Unified around a common goal to understand disasters and provide knowledge and information for improved policy decision-making, the three fields contribute unique strengths to the study of disasters. The social science of disasters provides a deep background of theory and explanation for behaviors in disasters, CSS brings theories of complexity and tools for studying complex phenomena, and crisis informatics contributes new forms of data collection and analysis.

With the integration of these fields, we can fully implement our conceptualization of CSSD. As a data-driven, theoretically informed paradigm, CSSD leverages qualitative and quantitative approaches for gathering and analyzing data and developing and testing social theory throughout the stages of disasters, as shown in Figure 2. From the social sciences, theories and conceptual models should guide data collection and analysis and computer modeling and simulation. Computational techniques in CSS and crisis informatics such as digital tracing, online crowdsourcing, and aerial imagery provide the means for gathering data [e.g., crowdsourcing, volunteered geographical information (VGI, [214]), social media, and online field experiments] using information communication technologies and smart mobile devices. These techniques also provide artificial intelligence algorithms and visualization tools in social network analysis (SNA), GIS, machine learning, and deep learning to analyze the data and retrieve evidence to develop, support, or update social theory. Data from ICTs make up the observational components of CSSD that lead to new hypotheses for online experimentation and validation of computational models such as ABMs.
Data collection feeds data analysis, social theories, and computational models, all of which together form the main elements of CSSD. In the context of CSSD, these elements operate in continuous interactions, informing each other in cycles of discovery and explanation. Social theory and models provide us with the conceptual understandings of the processes in disaster, thus they can guide data collection. Data should also inform the models and theories, because the data provide the patterns of disasters (e.g., population displacement, extent of property loss, etc.). Digital data containing various kinds of information (e.g., “big crisis data”) are collected from online sources such as news reports and social media platforms and are integrated with more traditional qualitative and quantitative data. These new forms of data guide the formation of hypotheses, which are built upon the findings of social science disaster research. These hypotheses are then operationalized by identifying relevant information in the data and by finding ways to represent and integrate them into the models. The quantitative or computational models are then calibrated and run in simulations. To complete the process, the limitations, generalizability, and implications of the work are examined and inform the next cycle of data collection, theory formulation, modeling, analysis, and testing. This continuous loop of data collection and model refinement is necessary for understanding the processes and phases of disasters and their evolving nature. Although our conceptualization of CSSD has not been fully implemented in practice, we provide a few examples of the interactions among social theories, data collection and analysis, and computational modeling and simulation in Section 5.1. This is followed by Section 5.2, which outlines the challenges and opportunities arising from CSSD.

5.1. Interactions Among the Components of CSSD

In this section, we discuss the interactions among data, theory, and modeling. With respect to the interactions between data informing theory, a preponderance of the CSS and crisis informatics literature only explores the application of computational techniques and new technologies (as discussed in Sections 3 and 4). However, there are limited examples of traditional scientific methodology with researchers gathering data in the interest of testing theory and producing analytic results. One such example of this disconnect is that of Olteanau et al. [215], who used crowdsourcing to label 1000 tweets from 26 different crisis situations that took place between 2012 and 2013. Their findings identified six broad categories for information communicated over Twitter during disasters (affected individuals,
infrastructure and utilities, donation and volunteers, caution and advice, sympathy and emotional
support, and other useful information). The observational data were used both to inform the kind of
computational analysis performed on experimental datasets and to explore what types of crises elicit
specific Twitter user behaviors. This work demonstrates the potential of how data could be used to test
theories of human behavior in crises and disasters, such as how role theory explains individual and
group behavior, but data analysis studies do not go this far [216].

Data have been used less frequently to develop computational models for social experimentation
in simulation. For example, in Jumadi et al. [128], evacuation data were used to improve the social
simulation of populations escaping from volcanic eruptions. The data performed multiple functions
in the research—calibrating the model, verifying model dynamics, and validating the final model
build. As a result of the data used in the modeling process, the researchers were able to refine the
model and improve the prediction of the locations where people would evacuate to. Crooks and
Wise [129] demonstrated how ICT data in the form of VGI can provide similar functions. Through
data analysis of crowd contributed information, they studied the response of populations to variations
in aid distribution and subsequently used social simulations to explore how rumors relating to aid
availability propagated through the population.

Similar to data informing theory, the published literature showing how social science theory of
disasters informs the collection of new forms of data (e.g., social media, VGI) for analysis is relatively
scarce; however, we can find some examples. In the tradition of qualitative descriptive research in
disaster, Lin and Margolin [217] examined inter-communal emotions and expressions tied to theories
such as the social amplification of risk [218]. During the 2013 Boston bombings, they found that the
people who had visited Boston or were within close proximity to it had the most predictive power for
raising the level of fear, sympathy, and solidarity to Boston. Wen and Lin [219] studied the factors
(geographic proximity, media exposure, social support, and gender) of distress (anxiety, sadness, and
anger) after the 2015 Paris terror attacks, and compared the immediate acute responses and the ones
before the attacks. Glasgow et al. [220] compared the expressions of gratitude for social support
received after the 2011 Alabama tornado and the 2012 Sandy Hook school shooting (Newtown, CT),
and found that, despite the Alabama victims suffering from a more severe disaster (quantitatively),
they received proportionally fewer expressions of support. By examining the microblog posts (i.e.,
from Twitter) about the Flint water crisis, Oz and Bisgin [221] studied the attribution of responsibility
and blame in a man-made disaster. Classic scientific methodologies use theories and observations to
develop hypotheses for testing in further observations and experiments. Mao et al. [222] developed
an online experiment to test the relationship between team size and productivity (e.g., [223]) in a
realistic crisis mapping task. Not only did their work use collected data to update the existing theory
on complex tasks, but it also used the existing theory to inform their data collection.

Turning to how computational models can be informed by theory, models are often built to test
specific theories, and although they are not yet applied in the area of disasters, theory-based models
are prevalent in the field of CSS and conflict crises. In an agent-based model of conflict in Sierra Leone,
Pires and Crooks [224] tested Le Billon’s [225] theory that the spatial dispersion of a resource (in this
case, diamonds) leads to warlordism, secession, and mass rebellion. The theory dictated the selection
of data collection from a variety of spatial data sources, including OpenStreetMap [226]. Their model
subsequently provided confirmation of the theory using basic bottom-up processes operationalized in
the model by enabling agents to choose whether to mine, rebel, or do nothing when varying the spatial
dispersion of diamond mines and areas under government control. Traditional social science models
often simplify the complex interactions of socio-economic dynamics with linear representations, e.g.,
increasing education and employment will improve the quality of life in a community. Although there
is ample confirmation of these theories, real-world data paint a more complex picture of confounding
relationships, such as global economic trade or group unrest. Exploration of these complex interactions
and nonlinear relationships has been done in another agent-based model [227] that uses identity and
social influence theory [228] to inform data collection and exploration of the emergence of riots. Other
applications related to disasters and conflict apply the theory of planned behavior [229] to agent-based models of social-ecological systems, for example, Kniveton et al. [230] and Schwarz and Ernst [231], and these are currently being fit into a framework for mapping and comparing behavioral theories [232].

Traditionally, models were built to test a specific theory. However, with advancements in computation, models are now also used to develop theory, changing the relationship between theory and modeling [126,135]. This is specifically evident in the utilization of agent-based models. Agent-based models have been shown to be suitable for capturing the heterogeneity and the complex interactions between agents and their environments and have made progress in the development and study of theories of complexity in disaster. Perhaps the most well-studied area of agent-based models of complex human behavior in disasters is that of evacuation during fires [122,233,234], earthquakes (e.g., [235–237]), tsunamis (e.g., [238]), hurricanes (e.g., [239]), volcanic eruptions (e.g., [128]), and floods (e.g., [240]). At a finer level of spatial resolution in predicting human behavior, Helbing et al. [241] modeled escape panic in the spirit of self-driven many-particle systems while [242] focused on more behavioral rich agents that lead to collective egress in evacuations. These and other agent-based modeling frameworks (see [243] for a review) have informed the development of theory on evacuation dynamics in emergencies and disaster. More current work builds on these efforts with further explorations of the complex dynamics in evacuation relating to second-order relationships. Wang et al. [238] simulated evacuation for a near-field tsunami in an ABM, investigating the impacts of decision-time, modes of transportation, and the availability of evacuation paths on mortality rates.

Just as models can inform theory, models can also inform data analysis. Social simulations in the form of computational models have used data from ICTs and created simulated data for analysis. Dawson et al. [240] developed an agent-based model and simulated data to analyze the risks of flooding in different scenarios and provide new insights about flood incident management. Wise’s [122] agent-based model of wildfire evacuation demonstrates how, when given a set of parameters for specific scenarios, social simulations in computational models can be used to create data for analysis, predict human behavior, and provide data for policymakers. To assess the longer-term welfare impacts of urban disasters, Grinberger et al. [244,245] made several simulations spanning the three years after an earthquake. They simulated the urban dynamics (residential and non-residential capital stock and population dynamics) using both bottom-up (locational choice for workplace, residence, and daily activities) and top-down (land use and housing price) protocols and analyzed the resulting data to find that low-income groups lose both housing and embedded social support systems. Realistic population synthesis is another important aspect of social simulations of disaster response. Burger et al. [246] proposed a model to synthesize agents using public open data sources such as the U.S. Census Bureau’s demographic profile dataset, business patterns dataset, and workflow (LODES) dataset. In this case, the model informs the creation of a synthetic population from census data in social simulation. This review of the interactions among components of CSSD identifies some gaps in current practice and brings us to the opportunities and challenges that need to be overcome to fully operationalize this new field of CSSD.

5.2. Challenges and Opportunities

The field of CSSD encompasses a cycle of interactions in data analysis, computational models, and social theories in the scientific process, and as discussed above, we find examples of this in the literature. In reviewing the work of these areas, we hope it has become clear to the reader that, while each element of CSSD is present, the full conceptualization of CSSD has not been achieved. Two examples of this are data not being used to inform models and few applications of disaster theory in agent-based models. More critically, there is no significant published disaster research that completes a full cycle of interactions in data analysis, computational models, and social theories. Future work needs to close these gaps to take advantage of the opportunities and address the challenges in this new field of study.
As with any emerging field, there are many challenges ranging across a wide spectrum of topical, technical, methodological, and ethical issues. Topically, we see an overreliance on case-study analysis in agent-based models, the disaster sciences, and crisis informatics research, caused in part by the context-dependent nature of disasters and the challenge of sharing data and models. Open science practices can support open exchange of research and allow for generalization to larger theory and replication with platforms sharing data (e.g., Dataverse project [247]) and models (e.g., OpenABM, the Computational Modeling in Social and Ecological Sciences (COMSES, [248]), and GitHub [249]). The nature of most new sources of data (e.g., public polls and social media) is short-term and post-event, and they contribute to understanding the processes of preparedness and response in disaster. CSSD needs to develop strategies for obtaining longitudinal sources of data for the long-term processes evident in mitigation and recovery stages of disaster. From a technical and methodological perspective, challenges in the forms of data, the collection techniques, and the machine learning algorithms all create biases in the data (as discussed in Section 3.1 and [250]). Verification and validation are also problematic. For example, deep learning algorithms suffer from a lack of human interpretability because their machine learning processes operate in a black box and do not have intermediate measures to verify whether they are performing as intended. Validation (demonstrating that results align with real-world outcomes) in this area also suffers from a number of issues, including a lack of high-quality, real-world datasets for comparison to model outputs. The complex subject matter of disasters presents a challenge with the requirement of analyzing heterogeneous variables and multiple interacting processes that prevent the isolation and evaluation of specific actors and processes.

The greatest opportunity in CSSD is the wealth of data sources now available to researchers. Big data and ICT are providing new data sources in the forms of online data collection, social media, and VGI. These sources enable the quick mapping of roads and geographic terrain of disaster event areas, individual reports of events on the ground, and more sophisticated online data collection applications and organizations, such as Ushahidi [161], Missing Maps [251], and Humanitarian OpenStreetMap Team (HOT [252]), that can now be implemented during disaster events. These data opportunities can be expanded with decision support science for improved decision-making in agent-based models such those used for wildfire training, incident command, and community outreach [253]. For instance, SimTable was used in the 2016 Sand Fire in California [254]. Not only are these platforms being used to inform policy decision-making on aid, but they also provide easier post-event data collection using the footprints of digital activities. Because the collection of disaster information can be undertaken post-event and far from the event’s location, researchers can help address a major limitation in disaster research, “unobservability” [7]. ICT has also opened up a new frontier in social science experimentation through the use of Internet platforms for online field experimentation; examples include Survey Monkey [255] and Amazon Mechanical Turk [256]. Due to the inherent unpredictability of disasters’ effects, crisis informatics and other disaster studies are often vulnerable to an overreliance on post-event data. Pre-event data is necessary to establish baselines of social phenomena and event causation. VGI and ICT could be leveraged to gather these data with less cost and conduct digital tracing backwards from the time of any event. Beyond ICT, there are opportunities in the use of new data analysis tools and packages widely available to data scientists that have made problems subject to multivariable causation and complex nonlinear processes both tractable and feasible on individual computer platforms. Artificial intelligence techniques for machine learning (a statistical method for describing a set of data features) and deep learning (a statistical learning that extracts features from raw data) can now be used to create knowledge and have moved beyond the domain of computer scientists into that of social scientists.

Finally, the data science community at large has yet to develop ethical standards for the collection and the handling of human subject data. Current work leverages existing standards in the social sciences, but there are risks and consequences of aggregating this information into big data. Privacy issues arise when analysis from data collected through ICT and social media (e.g., [257–259]) reveal more than what was intentionally provided, such as the identification of vulnerable individuals from the
aggregated information. The level of detail available in big data increases the risk of de-identification to human subjects and requires mitigation with privacy and security controls in the use and protection of the data [112,257,260].

6. Summary and Conclusions

In this paper, we explore three research domains that contribute to the modern understanding of disasters—the social sciences of disasters, computational social science, and crisis informatics. Social science lines of inquiry contribute to our fundamental understanding of the social processes and interactions at work in disasters (Section 2). However, disciplinary structures in academic research have prevented analysis of the complex social process that cross traditional boundaries, such as scaling, long-range interactions, and tipping points. In addition, they have not been able to fully take advantage of the increasingly available sources of new data generated by the proliferation of Internet technologies and mobile devices (e.g., social media, volunteered geographical information, digital news, open data, etc.). We introduce CSS (Section 3), the exploration of social science questions through advanced computational techniques, to show how new forms of data analysis and computational models are providing a new lens with which to study the world around us. Moreover, CSS provides new theoretical underpinnings to explore the complexities and the interacting processes seen within disaster studies. We do this because our goal is to close the gap between the crisis informatics (Section 4) and the social sciences of disasters. The idea of CSSD is introduced to merge traditional social science research with advances in CSS and crisis informatics (Section 5), and we discuss the interactions among the data, theory, and modeling components of CSSD in Section 5.1, along with the opportunities and the challenges of this new avenue of disaster research in Section 5.2. CSSD provides a trans-disciplinary approach to the study and management of disasters and moves beyond simply looking at disasters from a technical or social disciplinary perspective.

Closing the gaps separating the social science of disasters, CSS, and crisis informatics, CSSD’s foundation in quantitative data collection, processing, simulation, and analysis provides new knowledge at a deeper level with new forms of data and longitudinal evidence. It is important to integrate these lines of inquiry. Techniques such as computational modeling allow us to explore the patterns of and responses to the various phases of disaster, especially in the era of big data, but they do not enable us to explain the processes behind them. The social sciences of disaster contribute theory and explanation for the complex social and environmental processes involved in the construction, mitigation, response, and recovery from disasters, but they do not have the data tools to collect, compute, and analyze the immense volume and potential interactions in disaster data. Together, these lines of inquiry allow for more thorough investigations of the interactions among the development of social theory, data collection and analysis, and computational modeling of disasters. Through CSSD, we are able to leverage advantages from these domains, go beyond disciplinary boundaries, and gain a deeper understanding of the social and behavioral aspects of disasters in a digitally connected world.

Author Contributions: T.O. and A.T.C. developed the topic, and T.O. along with A.B. prepared the initial draft; A.T.C., A.B. and W.G.K. prepared the final draft for submission and all authors approved its content.

Funding: This work has been partially sponsored by the Defense Threat Reduction Agency (DTRA) grant DTRA1–16–1–0043. The Center for Social Complexity at George Mason University also supported this work.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

References

17. Kling, R. What is social informatics and why does it matter? Inf. Soc. 2007, 23, 205–220. [CrossRef]
21. Torrens, P.M. Geography and computational social science. Geoforum 2010, 41, 133–148. [CrossRef]
39. Tierney, K.J. From the margins to the mainstream? Disaster research at the crossroads. *Annu. Rev. Sociol.* 2007, 33, 503–525. [CrossRef]
44. Rodríguez, H.; Barnshaw, J. The social construction of disasters: From heat waves to worst-case scenarios. *Contemp. Sociol.* 2006, 35, 218–223. [CrossRef]


62. Oliver-Smith, A. Anthropological research on hazards and disasters. Annu. Rev. Anthropol. 2016, 46, 151–166. [CrossRef]


64. Barrios, R.E. What does catastrophe reveal for whom? The anthropology of crises and disasters at the onset of the anthropocene. Annu. Rev. Anthropol. 2017, 46, 151–166. [CrossRef]


79. Peregrine, P.N. Reducing post-disaster conflict: A cross-cultural test of four hypotheses using archaeological data. Environ. Hazards 2019, 18, 93–110. [CrossRef]


82. Garrett, T.A.; Sobel, R.S. The political economy of fema disaster payments. Econ. Ing. 2003, 41, 496–509. [CrossRef]
90. Kirchberger, M. Natural disasters and labor markets. *J. Dev. Econ.* 2017, 125, 40–58. [CrossRef]
96. Strömberg, D. Natural disasters, economic development, and humanitarian aid. *J. Econ. Persp.* 2007, 21, 199–222. [CrossRef]


118. Mearns, G.; Simmonds, R.; Richardson, R.; Turner, M.; Watson, P.; Missier, P. Tweet my street: A cross-disciplinary collaboration for the analysis of local Twitter data. *Fut. Internet* 2014, 6, 378–396. [CrossRef]


121. Wang, F.; Cui, P.; Sun, G.; Chua, T.S.; Yang, S. Guest editorial: Special issue on information retrieval for social media. *Inf. Retr.* 2012, 15, 179–182. [CrossRef]


228. Stets, J.E.; Burke, P.J. Identity theory and social identity theory. Soc. Psychol. Q. 2000, 63, 224–237. [CrossRef]


