Article

A Health SDI: The Benefits of and A Call-to-Action for Building a Global Health SDI

Ajay K Gupta, CISSP, MBA [1], Ram Peruvemba, MD[2], and Paul A Churchyard[3]

[1] Health Solutions Research; agupta@HSR.*health*
[2] Health Solutions Research, rperuvemba@HSR.*health*
[2] Health Solutions Research, paul@HSR.*health*
* Correspondence: agupta@HSR.*health*; Tel.: (+1 240-731-0756, AKG)

Abstract: Spatial Data Infrastructures (SDI) support the harvesting, curating, storage, and sharing of data along with providing access to development, analytic, and visualization tools that enable the building of innovative applications to address broad or specific challenges. SDIs can be especially powerful in bringing together data and tools supporting a particular theme – and this paper discusses and demonstrates the value of an SDI focused on Health. Many potential benefits of a Health SDI are proposed, and the case of supporting emergency response efforts is developed in detail. Leveraging a Health SDI, a Health Risk Index was created that provides emergency response personnel (both Emergency Operations Managers and Emergency Medical Responders) key insights into the unique health risks the impacted population faces due to the disaster. In order to establish the Health Risk Index, datasets from multiple national and global sources representing health data and social data that influences health outcomes – typically called social determinants of health – are harvested, merged, and republished to support further efforts at advancing the Health Risk Index. Visualizations of the Health Risk Index at the global, national, and sub-national levels down to the address level are presented along with demonstrations of its use.

Keywords: Spatial Data Infrastructure, Social Determinants of Health, Healthcare, Health, Geospatial Data Analytics, Geocoding, GeoHealth, GIS, Open Standards, Population Health, Disaster Response, Emergency Response.

---

1. Introduction

Through HSR's work in population health analytics, a distinct gap came to light with respect to the availability & accessibility of healthcare data and data related to health outcomes across the world. The gap is simply that a global repository does not exist that contains both health data and also data that influences health outcomes – commonly called Social Determinants of Health (SDoH). These are factors, typically outside of healthcare, that are known to be correlated with and even directly influence healthcare outcomes [1]. A few examples include:

- Educational attainment
- Income levels & distribution
- Availability and affordability of quality food choices
- Safe Housing

As researchers who develop solutions to address health challenges know, the key to solving wide-spread health issues generally will involve addressing the social issues in homes and communities that are a part of daily life. The role social issues play in influencing healthcare outcomes is becoming more and more recognized within and outside of the healthcare sector.

Over the course of curating and geocoding health and social data sets for various population health projects, it became clear that the collection of data was the beginnings of a Health Spatial Data Infrastructure or SDI. The potential of a Health SDI to address a wide range of population health challenges was also becoming clear. For example, food insecurity is defined as the lack of available and affordable healthy food options – *in an area*. Another example is that the lack of sufficient and accessible safe drug disposal resources contributes to the availability of potentially harmful prescription medications on the black market – *and specifically in that underserved communities*. These examples demonstrate that health and social data both share geographic context that is often overlooked. Leveraging an SDI can provide this geographic context.

Aggregating geocoded health and social data through a Health SDI allows for the inclusion of geospatial analysis into healthcare - addressing in innovative ways populations health challenges including the use cases discussed above. This paper presents the case of building an SDI to better serve healthcare goals at the global as well as national and local levels.

A layman's definition of a Spatial Data Infrastructure, or SDI, is that it is essentially a catalogued library that brings together all data that can be related to a particular theme and has a geospatial analytic platform featuring development, analytics, and visualization tools on which applications leveraging the data and tools can be built, hosted, and shared.

A more technical definition has been provided by Kuhn (2005) [2], as "An SDI is a coordinated series of agreements on technology standards, institutional arrangements, and policies that enable the discovery and use of geospatial information by users and for purposes other than those it was created for."

SDIs can be used to harvest, curate, and bring large data sets together in one place. As an example, GeoPlatform.Gov is a large-scale SDI that brings together data on natural disasters [3]. emPOWER is an application produced by the U.S. Department of Health and Human Services (HHS) that displays the number of individuals in the U.S. down to the individual ZIP Code level that are dependent on electrically-powered, life-saving medical devices [4]. The emPOWER application is accessible to the public through GeoPlatform.Gov. Alternately, applications can be made available to only select communities of interest based on user rights and security profiles. Figure 1 shows a small ample set of data sources leveraged by and available through GeoPlatform.Gov.

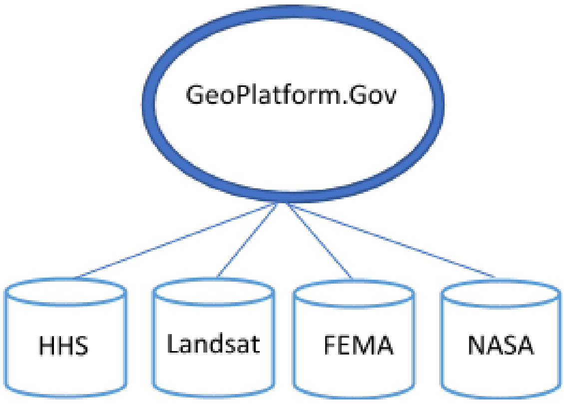

Figure 1. GeoPlatform.gov SDI Data Sources

The consequences of living in a world without a Global Health SDI are very real and impact many aspects of our lives, often in ways that are not immediately recognized.

1. Sourcing Data

   Medical research requires well defined data which can be difficult to source without a clear repository for such data. Medical datasets themselves are hard to find, especially in a geographical context and especially when analysis is at a specific geographical level.

2. Data sharing

   Resources are spent in recreating duplicative datasets, as there are limited mechanisms for sharing data sets of value across and among multiple communities within the health sector. Datasets that are created by medical researchers can often be duplicates of other datasets that have already been produced but were unknown or not accessible because of the lack of a widely recognized mechanism for the exchange and sharing of data among researchers and research efforts. In fact, a mechanism that allows for the communication of data set(s) needed and data set(s) available for sharing may itself not exist.

3. Semantic Mediation

Within the Healthcare space, there currently is limited semantic mediation overall. In other words, across the globe - and often also at local levels within individual health systems – the same medical conditions can be referred to differently. This does impact diagnosis and treatment to a certain extent, as well as the ability to use data in medical research. A great value of a Health SDI is that it can be an effective and powerful tool for supporting efforts towards achieving semantic mediation within the industry.

4. Standardization

Metadata and data formats should follow open standards to allow for interoperability among datasets – a boon to medical researchers across the globe. A health SDI can make data sets available following open standards. This applies to the data as well as to the metadata. Standardizing metadata can benefit healthcare. For instance, the definitions used for common terms in countries around the world can be different – which complicates efforts to identify trends in health outcomes and their potential costs. This variation in definition can be addressed at the metadata level. So, even though the definitions used for common terms may be different across the globe, by having metadata and file format in common the data can be easily adapted to the definitions used locally.

5. Emergency Response

It's not unusual that before data can be put to use in any way, time must be spent aggregating, standardizing, and preparing the data for analysis. In an emergency response scenario, time is simply not available, as response decisions must be made immediately. This essentially means that for health data to be considered and acted upon during emergency response planning process, it must be incorporated into the planning process a priori.

HSR proposes that a Health SDI can be used to solve these issues. The potentially most prominent and impactful being the case of emergency response efforts, as they can take place anywhere and at any time. According to the United Nations (UN) Office for Disaster Risk Reduction (DRR), financial damages and the cost in human lives due to natural disasters are tremendous [5]:

Table 1. UN statistics on global natural disasters, 2005-2014 [5].

Damages	People Impacted	Lives Lost
$1.4 Trillion	1.7 Billion	0.7 Million

These statistics pre-date the 2017 Atlantic Hurricane Season which was one of the most active in history with 17 named Hurricanes and with just three (3) storms - Hurricanes Harvey, Irma, and Maria - causing $265 billion in damages [6] (p. 1). Hurricane Maria, which made landfall in Puerto Rico on September 20, 2017 as a Category 4 hurricane caused $90 billion in damages, crippled the power grid for months, impacted 100% of the island's 3.8 million population and caused 2,975 deaths [7].

In the response to Hurricane Maria, there is anecdotal evidence that the lack of medical personnel early on in the response contributed heavily to storm-related fatalities and medical complications among survivors. These are lives that could have been saved and medical issues that could have been avoided if medical personnel were dispatched to the disaster impacted area sooner. The scope of the need for medical personnel wasn't known by Emergency Operations Managers until the response was underway.

A similar situation occurred during Camp Fire, which started on Camp Creek Road in Northern California's Butte County on November 8, 2018 and caused $16.5 Billion dollars in damages, forced 52,000 people to be evacuated, and cost 85 human lives, Table 2. For Camp Fire, there are numerous press reports of the elderly being essentially trapped in their homes as they were physically unable to move out of their homes [8]. With Camp Fire, the scope of the need for Search and Rescue teams to head *into homes* to evacuate people with limited mobility wasn't known in advance by Emergency Operations Managers until the response was underway – and in many cases, until even afterwards.

Table 2. Major United States natural disasters in 2017

Natural Disaster	Damages	People Impacted	Lives Lost
Hurricane Maria	$90 Billion [7]	3.8 Million [7]	2,975 [7]
Camp Fire	$16.5 Billion [9]	52,000 (evacuated) [10]	85 [11]

2. Materials and Methods

A generic data architecture for a Health SDI is shown in Figure 2, which is aligned with the architectural approach presented in The SDI Cookbook [12], developed with the support of the Global Spatial Data Infrastructure community. Both alternate architectures and dataflows are certainly possible to accomplish specific goals.

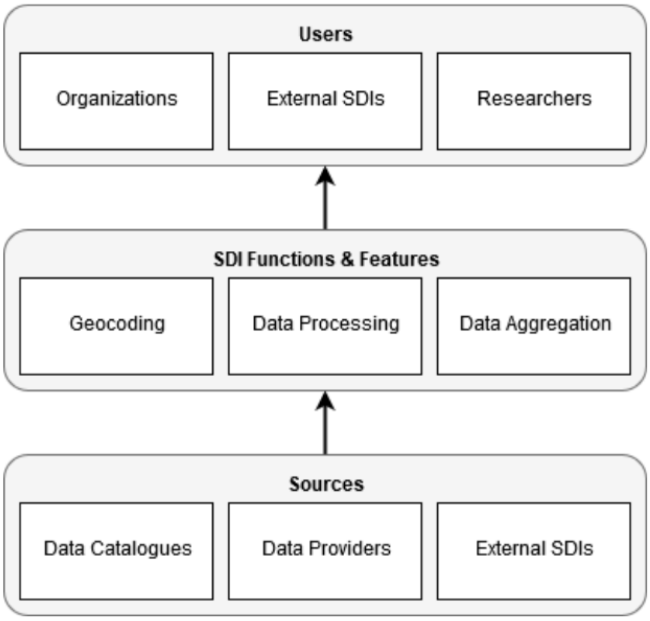

Figure 2. Simple Health SDI Architecture

There are cloud platforms available as a service that provide all features of this architecture, as well as a configurable workflow for a Health SDI. HSR's Health SDI is built as a GeoNode instance [13] on an Amazon Web Service server, which allows for both scalability and wide access to the SDI. Figure 3 displays a generalized overall workflow for supporting and hosting applications on the Health SDI.

The first step in this workflow is to identify and retrieve relevant health and social data. Data can be sourced from a variety of government, academic, non-profit, and other sources, while options can be kept open should new and better sources become discovered.

Geocoding, processing, and aggregating the curated and harvested data results in a Data Catalogue with a theme of health. The Data Catalogue acts as a centralized and searchable index of health and non-health data that is relevant to health outcomes and is certainly valuable in and of itself. It provides an invaluable resource aiding research into the causes of and potential remedies for health challenges facing populations.

A Health SDI can provide tremendous additional value by supporting applications leveraging data and geospatial tech to address specific health-related challenges.

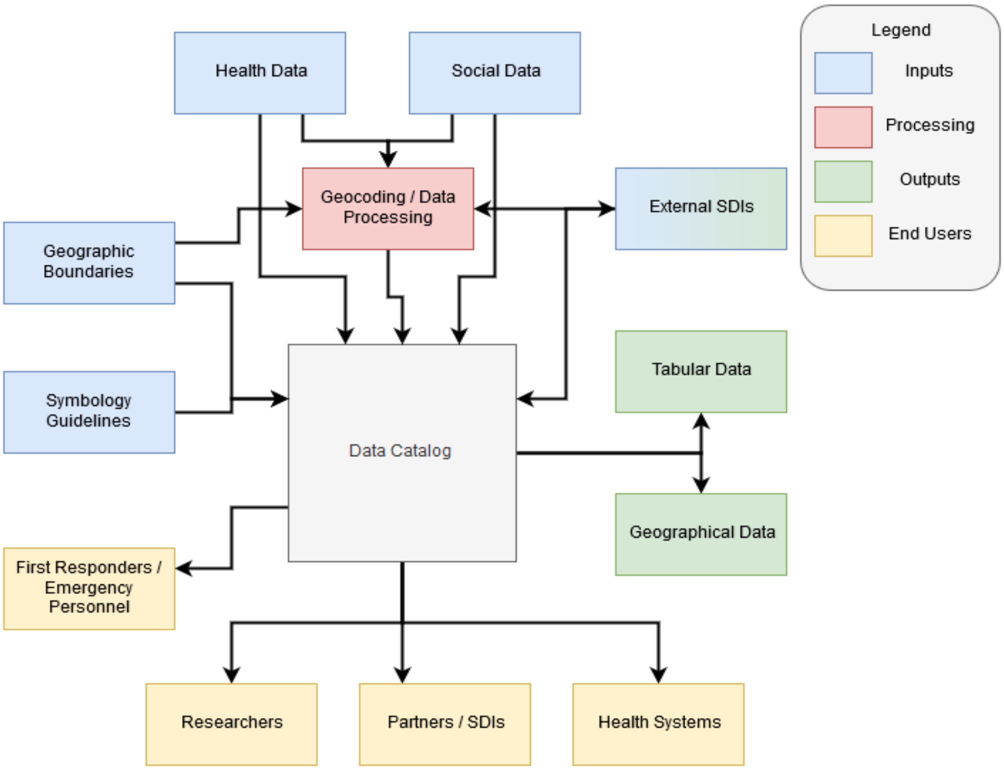

Figure 3. Health SDI Architecture & Workflow

Data from the Health SDI can be retrieved, in tabular or geographical data formats (when available) and can be geocoded (when and as necessary) and prepared for analysis. Analysis of the data can include leveraging analytical tools, including commercial tools such as Tableau, or open source tools such as R and Python. Symbology guidelines for specific outputs can be addressed in multiple ways. For instance, the National Alliance for Public Safety GIS Foundation produces symbology guidelines for use in emergency response applications, however, a Health SDI can support a wide variety of potentially symbologies. And ultimately, the analyzed data or layer can be published to the Health SDI. Finally, resources on the Health SDI can be made available to users in numerous ways, potentially including but not limited to:

- Linking to a data provider
- Direct download
- Through Open Geospatial Consortium (OGC) open standards including Web Feature Services (WFS) [14] and Web Map Services (WMS) [15]

Access can be open to the public or to private audiences through a username/password approach or through a variety of alternate security measures.

One value of an SDI, especially one utilizing open standards, is that it links to other SDIs, who themselves are engaged in aggregating, harvesting, and collecting data relevant to the specific theme or themes they support. Popular U.S. SDIs that include wide sets of data include GeoPlatform.gov and AmeriGEOSS. NextGEOSS is the European equivalent to AmeriGEOSS.

This interoperability provides immediate access to large and diverse sets of data. So, it may not be necessary to do a lot of building to build an expansive SDI – under the caveat that the SDI is built following open standards.

3. Results

3.1. How an SDI can Support Disaster Response Planning

A Health SDI can support disaster response planning through providing foresight to Emergency Operations Managers and Emergency Medical Responders so they can better ensure the medical and health needs of the impacted population are know in advance and can be addressed.

To provide this foresight, HSR is creating a Health-focused SDI, as a global Health SDI, with an initial application to allow for the inclusion of health information in the emergency or disaster response planning and execution processes. This application is named a Health Risk Index and merges data on social determinants of health as well as health data to identify an impacted population's risk – health risk – from a natural disaster. This information can be utilized by Emergency Operations Managers and the boots-on-the-ground Emergency Medical Responders to ensure their efforts address the unique health challenges caused by experiencing a natural disaster.

Currently, emergency response providers including the U.S. Federal Emergency Management Agency (FEMA) and their peers around the world as well as at State and Local levels within the country report having limited visibility into the health needs of disaster-impacted populations. Therefore, Emergency Operations Managers and Emergency Medical Responders may develop response & evacuation plans or may enter disaster struck areas unaware of the overall health and medical needs among civilians.

The Health SDI and its Health Risk Index application will provide the necessary, geocoded public health information that may prove vital to emergency response efforts and will align smoothly with emergency response systems in place across the globe. The Health Risk Index is currently being hosted on and utilizes data accessible through the Health SDI. Together, they form decision support tools whose capabilities are limited only by the imagination.

3.2. End Users and Stakeholders

As the near-term output of the Health Risk Index relates to disaster response efforts, the primary end users, or data consumers, include Emergency Operations Managers and Emergency Medical Responders. Emergency Operations Managers are those individuals who plan and oversee the overall disaster response effort. Emergency Medical Responders are the "boots-on-the-ground" who go into disaster-impacted areas and provide disaster relief services including needed medical care. The needs of these users are certainly related and both users can have visibility into the data provided to each other.

3.2.1. End User Needs

The identified end users are operating in a life-or-death environment. They must make split second, immediate decisions – often based on incomplete information and with human lives depending on the outcomes of those decisions. To help first responders improve their capability, it is essential that we provide accurate and actionable information that is synthesized into their current workflow and process flow.

The Health Risk Index aims to provide information that is useful to the identified end users, and in a manner which they can easily consume through their current practices.

This information includes:

1. Approach to prioritizing evacuations based on community-level health needs.

   Emergency Operations Managers develop evacuation plans responsive to specific nuances of each disaster event. The Health Risk Index provides information on the risk of adverse health impact due to the disaster. Emergency Operations Managers can use this information to design evacuation plans that best meet the impacted populations' health needs and limit the likelihood of adverse and life-threatening health events, including fatalities.

2. What healthcare assistance is needed, where, and for whom?

   To identify this information the Health Risk Index includes data on health conditions that may be worsened by or flair up during disasters such as allergies, asthma, heart disease; as well as health conditions that are caused by the disaster such as illnesses or trauma. In addition to Emergency Operations Managers and Emergency Medical Responders, this information is relevant to hospitals, healthcare providers, as well as pharmacy distribution personnel serving the disaster impacted area. The U.S. Centers for Disease Control and Prevention (CDC) may also be interested in this information for the purposes of disease outbreak monitoring and containment.

3. What health conditions is the population likely to present with?

   Related to 2, this will identify the medical equipment, medical supplies, and pharmaceutical medications first responders should bring to a disaster-impacted area. In addition to Emergency Operations Managers and Emergency Medical Responders, this information is relevant to hospitals, healthcare providers, as well as pharmacy distribution personnel serving the disaster impacted area.

   All information will be provided seamlessly and within the first responders current workflow.

3.2.2 Secondary Consumers & Partners

Secondary consumers of the Health Risk Index include partners to whom the Index is exported and include firms like Esri and Compusult. These partners can receive an export of the Health Risk Index for utilization in their platforms following open data sharing standards produced by the OGC such as the WFS and the WMS standards. An SDI following open standards simplifies the sharing of both data and applications, and this portability is a key benefit of an SDI.

In the long-term, the Health Risk Index can be used to track post-disaster recovery efforts, and additional consumers in this context can include providers (e.g., hospitals and health systems), economic development agencies given the relationship between the health and economic potential of communities, as well as departments and ministries of health at the national and local levels who are charged with ensuring the long-term health of populations.

3.3.1 Discovering Data

To produce the Health Risk Index and provide actionable health information in a disaster response scenario, HSR sought health and social data that, when considered together, stratifies the health risks facing an impacted population due to a natural disaster. Boundary layers (Counties and ZIP Codes) [16] were discovered through searching the U.S. Census Bureau (Census) for the purpose of geocoding the data to geographies. Guidance on symbology was taken from the National Alliance for Public Safety GIS Foundation (NAPSIG) [17] and in addition, address data was discovered from the U.S. Department of Transportation's National Address Database (NAD) [18]. The following data sets were used in the construction of the Health Risk Index:

Table 2. Health Risk Index Data Sources

Data Provider	Data Catalogue	Name	Use
Homeland Infrastructure Foundation-Level Data	Geoplatform.gov	Health / Emergency Response Facilities [19]	Geographic Layers
National Alliance for Public Safety GIS Foundation	Symbol Library	Symbology Guidelines [17]	Symbology

National Oceanic and Atmospheric Administration	National Hurricane Center Data Archives	Storm Paths [20]	Geographic Layers
U.S. Census Bureau	American Fact Finder	American Community Survey 5-year Data Tables [21]	Social Determinants
U.S. Census Bureau	U.S. Census Bureau	TIGER Line U.S. Boundary Layer [16]	Geographic Boundaries
U.S. Centers for Disease Control	Agency for Toxic Substances and Disease Registry	Social Vulnerability Index [22]	Social Determinants
U.S. Centers for Medicare and Medicaid Services	AmeriGEOSS	Chronic Conditions [23]	Health Dataset
U.S. Department of Health and Human Services	Geoplatform.gov	emPOWER [4]	Health Dataset
U.S. Department of Housing and Urban Development	HUD User Portal	HUD USPS ZIP Code Crosswalk Files [24]	Geographic Scale Mobility
U.S. Department of Transportation	U.S. Department of Transportation	National Address Database [18]	Address Data and Attributes
The Dartmouth Institute for Health Policy and Clinical Practice	Dartmouth Atlas Data	Our Parents, Ourselves: Healthcare for an Aging Population Appendix Table 4 [25]	Health Datasets

3.3.2. Data Integration

The health and social data inputs are geocoded with crosswalk files from the U.S. Department of Housing and Urban Development (HUD) which allows for the information to be displayed at multiple geographic levels. Data from the health and social datasets are then used to calculate the Health Risk Index at each geographic level. Additionally, the health needs information relevant to the on-the-ground Emergency Medical Responders is displayed and exported. After the Health Risk Index is calculated for each geographic level the table is joined to the geographic boundary files from the Census, where symbology is then established, and the layer is published to the Health SDI. The layers available on the Health SDI can be retrieved using OGC standards. This allows organizations, such as Esri, Compusult, and others

to import and utilize the data in their platforms, allowing for wide use of the information by the First Responder community overall.

3.3.3 Republication of Data

All of the health data on which the Health Risk Index is comprised is republished as the HSR Medical Dataset on the Health SDI. Similarly, the social data components are republished as the HSR Social Determinants Dataset on the Health SDI. Additionally, the Health SDI republishes other data and resources related to disaster response efforts, including but not limited to a link to U.S. National Oceanic and Atmospheric Administration's (NOAA) National Hurricane Center data archive, Homeland Infrastructure Foundation-Level Data (HIFLD) datasets, and the NAPSIG's symbology guidelines.

Republishing data is one of the benefits of an SDI. It allows for data aggregated from multiple sources and processed or transformed as necessary for the purpose at hand to be made available along with the original data for those who may want to reconduct and validate the analysis. Republished datasets can be made freely available to the general public, under license, as well as under a paid license.

3.3.4 Displaying data with proper symbology

The symbology used in the Health Risk Index is based on the symbology guidelines produced by the NAPSIG [17]. This includes the symbols for hospitals, roads, the thematic coloring, as well as the legend. The legend, shown in Figure 4, is a 7-point scale ranging from very low risk, or a white color on the map, to an extremely high risk, or purple. The greater the risk, the more urgent the need to provide appropriate medical care and intervention, up to and including potential evacuation to a specific medical facility, to the population within the geographic region.

Description/Examples for Use	Color	RGB	HEX
No severity or risk; Normal Operations/Status	White	255,255,255	#FFFFFF
Low severity or risk; No Damage; Open Status	Green	0,172,58	#00AC3A
Low to medium severity or risk; Alert/Action Notice	Blue	35,122,207	#237ACF
Medium severity or risk; Moderate damage	Yellow	255,215,24	#FFD718
Medium to High severity or risk; Watch Notice;	Orange	255,137,24	#FF8918
High severity or risk; Warning Notice; Severe Damage; Closed Status	Red	255,24,30	#FF181E
Extreme severity or risk; Highest category possible.	Purple	237,26,252	#ED1AFC

Figure 4. NAPSG Symbology Guidelines for legend displaying risk levels [17]

Following NAPSIG's guidelines for the symbology is a choice. An SDI allows any symbology of interest to be used in visualizing applications. The Heat Risk Index can be replicated with entirely different color schemes for the legend and other symbols for hospitals, roads, building structures, and all other map features.

3.3.5 Public, Open Source Data vs Private, Licensed Data

At present, the Health Risk Index uses the most currently available data for the identified datasets that are in the public domain. An SDI can also support using private, licensed data and later phases of this effort will include discussion of licensing these data sets from providers to have access to either (or both where available) more regularly updated data and/or access to data at the Census tract level. These discussions will necessarily involve addressing and documenting policy and procedure issues, security and privacy issues, as well as issues surrounding the republication of such non-public data. However, for a proof-of-concept and in fact to offer a view into the background health posture of a region, publicly available data is certainly applicable and better than having no health insights at all.

3.4 Health Risk Index Output

The Health Risk Index is an application on HSR's Health SDI where relevant health and social data, related to a population's risk of adverse health events due to a natural disaster, is merged, analyzed, and exported in an accessible format. The Health Risk Index is produced at multiple geographic scales including the global level, Figure 5, the national level, Figure 6, and the sub-national level, including county and ZIP codes, Figure 7, for the United States.

Predicted storm paths produced by NOAA or other sources of disaster impact information can be overlaid onto the Health Risk Index to retrieve the index at all geographic levels as shown in Figure 8. Emergency Operations Managers can isolate the index for the impact region and extract health-related information for counties and/or ZIP Codes within the identified region as shown in Figure 9 at the ZIP Code level. The health-related information displayed identifies specific details on the local health posture of the disaster impacted population that can be provided to the On-the-Ground Emergency Medical Responders. This information includes, for instance, literacy levels, and the numbers of individuals experiencing mental health issues and critical health conditions that may require additional care of more urgent medical attention, can inform their delivery of medical services.

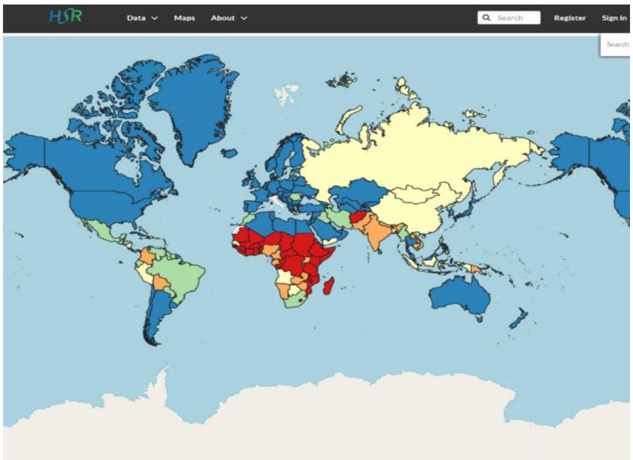

Figure 5. Global Level Health Risk Index

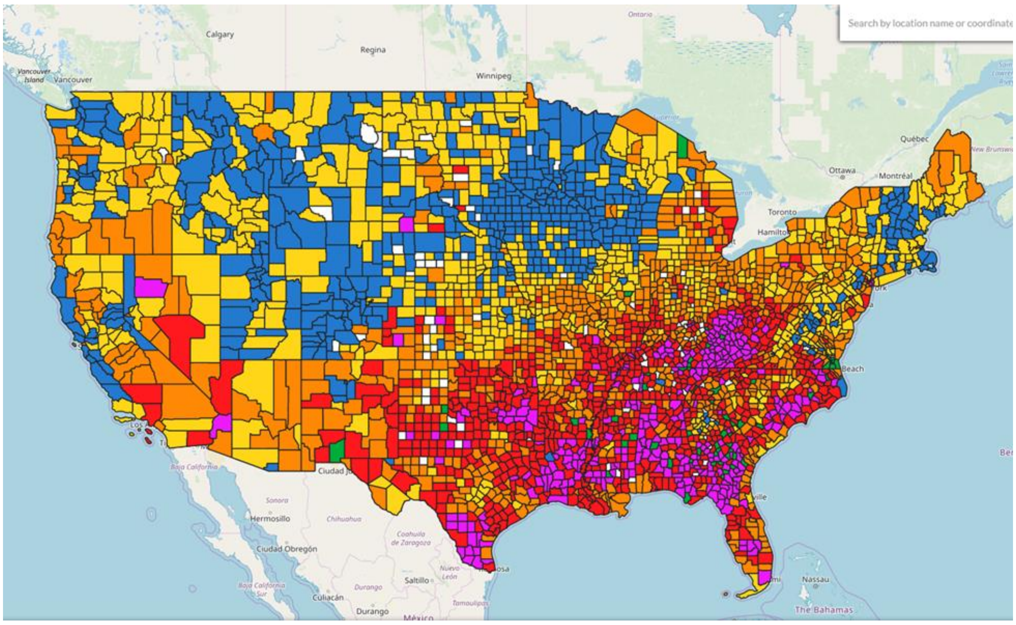

Figure 6. National County Level Health Risk Index

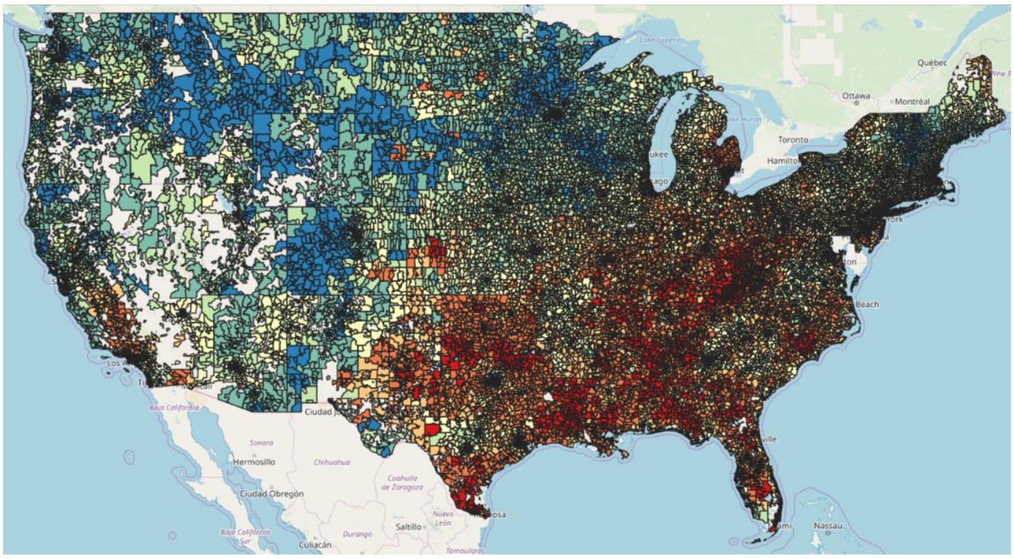

Figure 7. National ZIP Code Level Health Risk Index

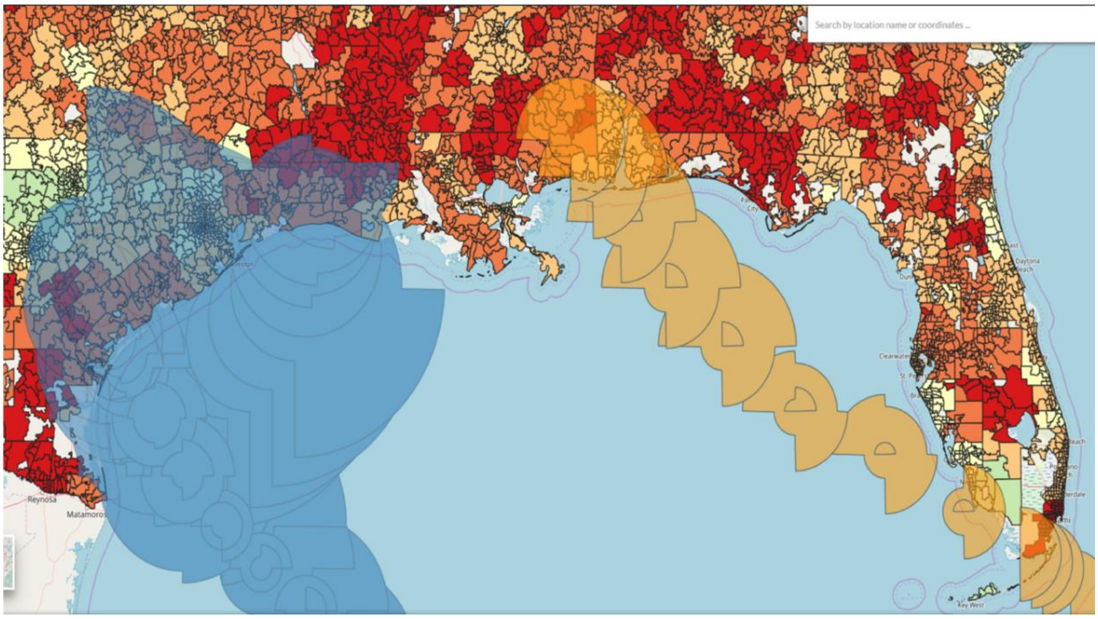

Figure 8. National ZIP Code Level Health Risk Index with Storm Tracks Overlaid

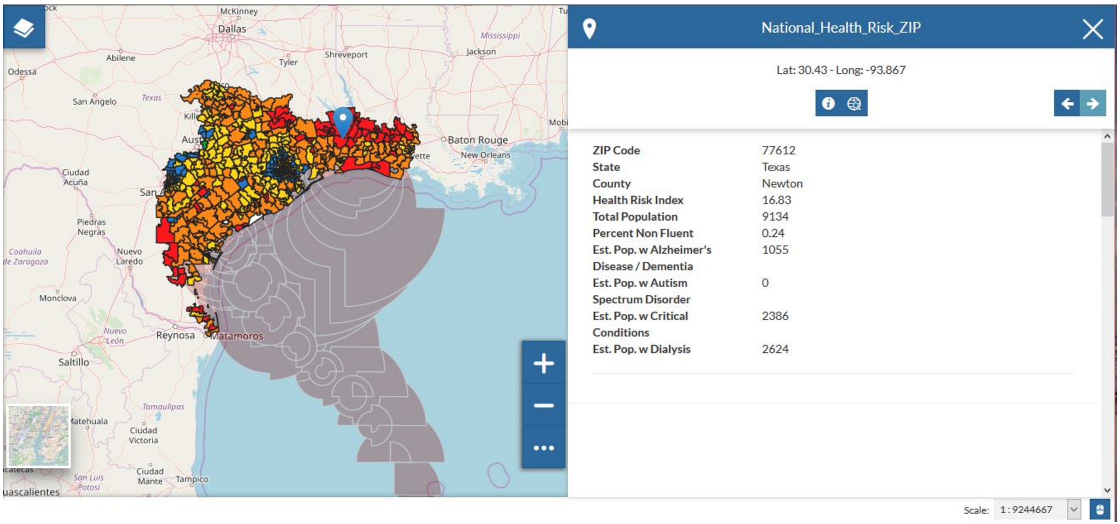

Figure 9. National ZIP Code Level Health Risk Index

4. Discussion

The Agency for Toxic Substance and Disease Registry (ATSDR) within the CDC maintains a Social Vulnerability Index (SVI) [22] that assess the "resilience of communities when confronted by external stresses on human health, stresses such as natural or human-caused disasters, or disease outbreaks". The Health Risk Index leverages the SVI in part. The purpose of the SVI is more broad than supporting disaster response efforts – which is a more narrow situational use case. However, the work done by the ATSDR and the CDC certainly advance the cause of identifying the health risk of disaster-impacted populations.

This underscores the value of an SDI. It allows for the SVI to be available alongside a dataset that is essentially a modified version of that same index. Future researchers can compare the analysis involved to validate the ultimate findings. Those with divergent views on what influences the risk of adverse health events in a disaster can use the original SVI or create an entirely separate version of this social component of health risk for their own testing, use, and purposes. A Health SDI allows for such flexibility. It allows for multiple data sources to be collected, merged, manipulated, and republished to address as specific purpose and make available for peer review.

An SDI is a living structure. There is always work to do in making it more valuable and this section provides a few areas where such improvement can be made. Outside of the specific yet critically important Disaster Response Scenario, a Health SDI has wide benefits to healthcare. Heath research efforts with a geospatial component can leverage the SDI as a library of existing data sources to inform the research, as well as to share results from the research – expanding the universe of knowledge publicly available. Efforts can be made to continue to search and determine optimal data sets and sources for inclusion in the Health Risk Index. This will include exploring data at the Census tract and the address levels.

Future work to improve the Health Risk Index will include formalizing the data sharing Policy and Procedure agreements with all data providers. Public data, as was used in establishing the Health Risk Index presented in this paper, especially data made publicly available from Government sources may not need a formal data use agreement – but it's always wise to check with the specific data providers.

Public data is sufficient for assessing the initial health risk of an impacted population to a specific disaster event - however, closer to real-time data will be more helpful for both response planning as well as on-the-ground delivery of emergency medical care. Discussions with data owners on formal data use agreements should address the licensing, policy, and procedural issues involved in accessing the most currently available data (hopefully closer to real-time), data availability at the Census tract and address levels, as well as security and privacy considerations.

The Health Risk Index will be refined through formal field tests with FEMA and their peers within and/or serving the First Responder community. Such field tests must validate the usefulness and efficacy of the Index as well as ensure the Index conforms to the First Responder community's planning and execution workflows.

Field tests of the Health Risk Index will also be an opportunity to develop a feedback loop into the system to gain information from Emergency Medical Responders and other on-the-ground personnel. Remote sensors can also be incorporated as part of a feedback loop that aids the overall planning and response effort. This can be an effective means of gathering, for instance, local road closures due to debris or flooding, as well as hospital status during the emergency. In addition, software and infrastructure, such as the OGC's GeoSMS [26] and SensorThings API [27] standards will also be explored to offer insights on the location of at-risk individuals.

Additional data sources and areas of application for the Health SDI and Health Risk index are as follows:

1. Hospital Status During Emergencies

   Emergency Operations Managers learn of individual hospital and medical facility status during declared emergencies, such as inpatient bed availability and Emergency Department occupancy levels, through a manual process. It has been reported that an individual within the

Operations Command Center is tasked with calling hospitals to get their status. In certain states, an electronic application serves as a digital interface for this manual process. Hospital staff report their status into an electronic system, and that status is then available to the State's Emergency Operations Managers. However, there is both a lack of timeliness in this reporting as well as a great deal of subjectivity leading to inconsistencies among hospital reports. An automated and standards-based system to inform Emergency Operations Managers of hospital status throughout disaster scenarios could potentially optimize and improve the coordination of medical care delivery during disasters. A Health SDI can aid or serve as such a system.

2. Post-Disaster Recovery Needs

Assisting first responders address the needs of populations during a disaster is critical for saving lives. It is also important to ensure post-disaster recovery leads to full restoration of the pre-disaster health posture of the population. Disasters may cause harm to the environment, key infrastructure, as well as the healthcare service delivery capacity all of which can change the health posture of a community. The health risk index can be used to track overall recovery and adapting the index for such a purpose can certainly be an important and beneficial follow-on activity. This supports the concept of "Build Back Better" that according to the World Bank can reduce the impact of natural disasters by 31%.

Specifically, the Health Risk Index can identify the possible post-disaster health risks to the disaster-impacted region as well as the surrounding area and population. This can include, for example, a storm that enables a source of pollutant to contaminate ground water can have an adverse impact on crop yields, area plant life and wildlife, as well as human health. Natural disasters can also destroy or make unpassable roads, transportation and other infrastructure, damage or destroy hospitals and other medical service buildings and infrastructure, as well as damaged crops and impact the food supply. Efforts to purify the water supply, restore critical infrastructure, and rebuild safeguards can be tracked through the Health SDI to restore, or improve, pre-disaster population health risk levels.

This information will be of interest to the CDC for disease containment, state/local government, economic development representatives, as well as community groups.

3. Remote Sensing

The Health Risk Index leverages Remote Sensing Data for its base map. A number of remote sensing data is considered for future phases of this effort, partially including use of the SensorThings API [27]; the inclusion of Land Use and Land Cover Data to determine staging areas, accessibility issues, live situational awareness, status of transportation systems; real-time situational awareness such as through drones; as well as leveraging the distributed Internet-of-Things (IoT) and Internet-of-Medical-Things (IoMT) sensors of Smart Cities to expand situational awareness, resource availability, and service capacity to advance emergency response efforts in providing needed healthcare services.

4. Smart Cities

Smart Cities, and their large-scale application of distributed sensors, provide an exciting opportunity to continue and in fact drive further advancements in global health. Distributed sensors that are associated with smart cities can allow for new means of gathering data on the health risks present within communities at the population as well as at the individual level. These include both Internet-of-Things (IoT) sensors such as, for instance, those on electronic devices, appliances, and residential and commercial environmental systems such as for tracking weather and airborne pollutants, as well as Internet-of-Medical-Things (IoMT) sensors on medical devices, testing equipment, implants, and health-related wearables. Together, such smart and Internet-enabled sensors constitute both a larger and more real-time health data set that can help predict adverse health events, diagnose and treat patients, as well as monitor the success of prescribed treatments. And treatment plans that prove ineffective can be adjusted more quickly.

Examples include the use of sensors to: Monitor current bed occupancy and medical device utilization levels at hospitals, skilled nursing centers, and other medical facilities to enable seamless patient movement throughout the healthcare system; Monitor levels of medical supplies and other critical resources - to ensure treatment and care capacity is sufficient to meet demand and population health needs.

This information can address current challenges in assessing hospital and medical facility status during emergencies. Overall, the analytics that a smart city can provide will improve the ability to assess the overall medical care needs of the population. A Health SDI can be an ideal and central location for such data to be stored – where the data and information can be available to emergency planners, public health agencies, and the overall provider community.

5. Data at the Address Level

In order to develop the Health Risk Index, publicly available health and social data at the National, State, County, and ZIP Code levels are sufficient. The Health SDI greatly benefits from the high-quality data produced by the various and numerous data providers that have been leveraged in this effort. However, in practice, First Responders will be further aided with data down to the most granular level possible – including at the Census tract and even at the address level. If an Emergency Medical Responder is tasked with evacuating a Dialysis patient to a medical treatment center, they will need an address for that individual.

As a demonstration, a percentage of addresses from a coastal ZIP Code in North Carolina that sat within the Hurricane Dorian anticipated storm track have been pulled from the U.S. Department of Transportation's National Address Database (NAD) [18] and shown in Figure 10 below. This percentage is the percentage of individuals with critical health conditions in that ZIP Code.

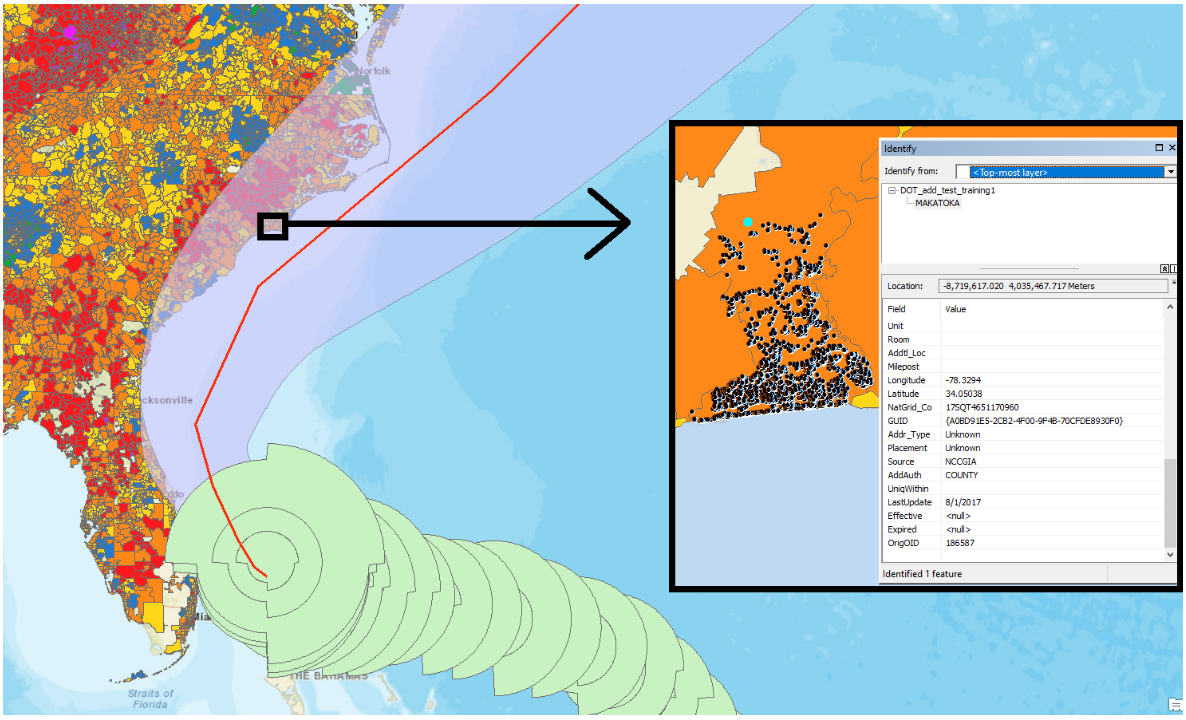

Figure 10. Address Data and Attributes Shown at the ZIP Code Level

The value of the NAD is that it contains longitude and latitude data as well as (at times) landmark data that can help actually find an address, especially in a disaster scenario when roads, etc., may be flooded, and debris may complicate finding a particular home or residence. HSR commits to working with the data providers on licensing, policy and procedure, as well as security and privacy issues to access and provide the Health Risk Index at the Census tract and address level.

5. Conclusions

The Health SDI can advance the state of healthcare globally by addressing data sourcing, data sharing, semantic mediation, standardization, and other health industry challenges. This allows for more geohealth analysis including as an example, providing actionable health insights to the emergency response planning and execution process.

A Health Risk Index, leveraging the health and social data sets and development, analytic, and geospatial visualization tools housed in the Health SDI, is capable of assessing a population's risk of adverse health events due to a natural disaster. And of communicating this information to Emergency Operations Managers and Emergency Medical Responders, as well as to other potential users, who can take the appropriate actions based on this information.

The Health Risk Index is a response to a recognized gap in overall disaster response planning and execution process in the U.S. – that is the lack of health information of disaster-impacted population. The responses to Hurricane Maria and Camp Fire demonstrated the consequences of that gap: If health and medical needs are not known in advance, efforts to address those needs may not be executed in the timeliest fashion to have the greatest impact. This can lead to increased economic costs in addition to costs in human lives and suffering.

HSR is able to demonstrate that a Health SDI hosting a Health Risk Index can fill that gap by leveraging publicly available health and social data from numerous government sources, research organizations, educational institutions, and other sources. Open data standards established by the OGC can then be leveraged to provide that data to Emergency Operations Managers and Emergency Medical Responders directly, as well as to partner organizations and other potential users. While further efforts are warranted as discussed in the last chapter, the end result is that a Health SDI can led to the timely delivery of essential medical care and intervention resulting in fewer lives lost to and medical complications caused by natural disasters.

Beyond emergency response, the benefits to healthcare can be widespread. A repository of global health data and data on factors impacting health outcomes can speed the identifications of mechanisms to reduce the adverse impact of those factors, to address the health complications they may cause, and also to potentially direct future medical & social research towards addressing those issues more holistically.

Imagine a world in which the world's health community openly collaborates to solve both our shared health issues, as well as those health issues that impact only unique populations or geographies. A truly open, glorious world. A Health SDI is the missing component that can enable such world-wide collaboration.

6. Patents: The work discussed in this article is not subject to patents, however, HSR has filed multiple patents on our work in geospatial analysis of health and social data to develop risk stratification of certain adverse health outcomes, such as the risk of opioid overdose and the risk of maternal and infant mortality and other pregnancy and labor and delivery related complications.

Supplementary Materials: The following demonstration video is available online at https://www.youtube.com/watch?v=z49CQfyV1bk&t=48s, Video title: Health SDI - Facilitating Emergency Health Response During a Disaster (HSR). The video describes the creation and use of the Health Risk Index discussed in the paper.

Author Contributions: Conceptualization, Gupta, Ajay K; Methodology, Gupta, Ajay K, Peruvemba, Ram; Software, Churchyard, Paul A; Validation, Gupta, Ajay K, Peruvemba, Ram; Formal Analysis, Gupta, Ajay K, Peruvemba, Ram, Churchyard, Paul A; Investigation, Gupta, Ajay K, Churchyard, Paul A; Resources, Gupta, Ajay K; Data Curation, Churchyard, Paul A; Writing-Original Draft Preparation, Gupta, Ajay K; Writing-Review & Editing, Gupta, Ajay, Peruvemba, Ram, K Churchyard, Paul A; Visualization, Gupta, Ajay K, Churchyard, Paul A; Supervision, Gupta, Ajay K; Project Administration, Gupta, Ajay K; Funding Acquisition, Gupta, Ajay K.

Funding: The Health SDI received no direct external funding. The Health Risk Index has been funded to the prototype stage through the Open Geospatial Consortium.

Acknowledgments: Technical support in the creation of the Health Risk Index was provided by the U.S. Geological Survey (USGS), the Federal Geographic Data Committee (FGDC), as well as by the U.S. Federal Emergency Management Agency (FEMA).

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

References

1. Blas E.; Kurup A.S. Introduction and methods of work. In Equity, social determinants and public health programmes, 1st ed; Blas E., Kurup A.S., Eds.; World Health Organization: Switzerland, 2010; Volume 1, pp. 3-10
2. Kuhn W. Introduction to spatial data infrastructures. Presentation, March 14th 2005.
3. Geoplatform.gov. Available Online: https://www.geoplatform.gov/ (accessed on October 14, 2019)
4. HHS emPOWER Map 3.0. Available Online: https://empowermap.hhs.gov/ (accessed on October 14, 2019)
5. The Economic and Human Impact of Disasters in the last 10 years. Available Online: https://www.flickr.com/photos/isdr/16111599814/in/album-72157628015380393/ (accessed on October 14, 2019)
6. FEMA. 2017 Hurricane Season FEMA After-Action Report. Available Online: https://www.fema.gov/media-library/assets/documents/167249 (accessed on October 14, 2019)
7. Milken Institute of Public Health. Ascertainment of the Estimated Excess Mortality from Hurricane Maria in Puerto Rico. Available Online: https://publichealth.gwu.edu/sites/default/files/downloads/projects/PRstudy/Acertainment%20of%20the%20Estimated%20Excess%20Mortality%20from%20Hurricane%20Maria%20in%20Puerto%20Rico.pdf (accessed on October 14, 2019)
8. Bizjak T.; Yoon-Hendricks A.; Reese P.; Sullivan M.; Many of the dead in Camp Fire were disabled, elderly. Could they have been saved? Available Online: https://www.sacbee.com/news/california/fires/article222044970.html (accessed on October 14, 2019)
9. Reyes-Velarde A.; California's Camp fire was the costliest global disaster last year, insurance report shows. Available Online: https://www.latimes.com/local/lanow/la-me-ln-camp-fire-insured-losses-20190111-story.html (accessed on October 14, 2019)
10. McBride A.; Gutierrez M.; Asimov N.; 29 dead in Camp Fire as firefighters make gains on Butte County blaze. Available Online: https://www.sfgate.com/california-wildfires/article/Butte-County-fire-only-grows-slightly-as-13381900.php (accessed on October 14, 2019)
11. Associated Press; List of Missing in Camp Fire down to 1. Available Online: https://fox40.com/2019/08/02/one-still-missing-in-camp-fire/ (accessed on October 14, 2019)
12. Global Spatial Data Infrastructure Association; The SDI Cookbook. Available Online: http://www.gsdiassociation.org/images/publications/cookbooks/SDI_Cookbook_from_Wiki_2009.pdf (accessed on October 14, 2019)
13. GeoNode; Open Source Geospatial Content Management System. Available Online: http://geonode.org/ (accessed on October 14, 2019)
14. Open Geospatial Consortium; Web Feature Service. Available Online: https://www.opengeospatial.org/standards/wfs (accessed on October 15, 2019)
15. Open Geospatial Consortium; Web Map Service. Available Online: https://www.opengeospatial.org/standards/wms (accessed on October 15, 2019)
16. United States Census Bureau; Geographies TIGER/Line Shapefiles. Available Online: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2018.html (accessed on October 15, 2019)
17. NAPSG Foundation; Incident Symbology Framework, Guideline, and Operational Implementation Guidance Version 2.0. Available Online: https://www.napsgfoundation.org/wp-content/uploads/2015/10/IncidentSymbol_Guideline_20160830_v2.0_PDF.pdf (accessed on October 15, 2019)

18. Transportation.gov; National Address Database. Available Online: https://www.transportation.gov/gis/national-address-database/national-address-database-0 (accessed on October 15, 2019)
19. Homeland Infrastructure Foundation-Level Data (HIFLD). Available Online: https://gii.dhs.gov/hifld/ (accessed on October 15, 2019)
20. NHC Data Archive (Text). Available Online: https://www.nhc.noaa.gov/data/?text (accessed on October 15, 2019)
21. United States Census Bureau; American Fact Finder. Available from: https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml (accessed on October 15, 2019)
22. Agency for Toxic Substances and Disease Registry; CDC's Social Vulnerability Index (SVI). Available Online: https://svi.cdc.gov/ (accessed on October 15, 2019)
23. Centers for Medicare and Medicaid Services; Chronic Conditions. Available Online: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Chronic-Conditions/CC_Main.html (accessed on October 15, 2019)
24. Office of Policy Development and Research (PD&R); HUD USPS ZIP CODE CROSSWALK FILES. Available Online: https://www.huduser.gov/portal/datasets/usps_crosswalk.html (accessed on October 15, 2019)
25. Dartmouth Atlas Data; Atlas Special Topics: Healthcare for an Aging Population (2016). Available Online: https://atlasdata.dartmouth.edu/static/atlas_special_topics#aging (accessed on October 15, 2019)
26. Open Geospatial Consortium; Open GeoSMS Standard – Core. Available Online: https://www.opengeospatial.org/standards/opengeosms (accessed on October 15, 2019)
27. Open Geospatial Consortium; OGC SensorThings API. Available Online: https://www.opengeospatial.org/standards/sensorthings (accessed on October 15, 2019)
28. Zaleski, John R, PhD; Peruvemba, Ram, MD; The Analytic Link Between Population Health and Leading Hospital and Ambulatory Patient Safety Considerations, Chapter 6 in Analytics, Operations, and Strategic Decision Making in the Public Sector, Evans, Gerald W, Biles William E, Bae, Ki-Hwan G editors, 1st ed; IGI Global; 2019, pp. 111-136