FEMA

Modeling and Data Working Group

Role of Data and Models in Supporting Planning and Response to an Improvised Nuclear Device (IND) Detonation

Phase III: Analysis of the IND Data and Models Inventory

TABLE OF CONTENTS

Executive Summary

During management of any large-scale emergency, whether an improvised nuclear device or a hurricane, federal, state and local entities must effectively leverage information from a wide range of sources to answer the fundamental questions of "What happened?"; "Who or what was affected?"; and "What needs to be done?" To address these three foundational questions and the corresponding critical information requirements, the information resources available to the federal interagency must be identified, collated, and made available so that they can be effectively utilized during all phases of emergency management to support operational decision making. These information resources include the iterative collection and processing of data through models and analysis tools that produce new, more useful data. These data can then be used directly to support informed decision-making or be further refined using additional models.

To capture this process of data collection and analysis, a conceptual flow of information framework to categorize the types of information, from raw data to operationally relevant information, has been developed. This categorization system contains seven basic types of resources: raw data, event characterization models and analysis tools, situational awareness data, consequence models, impact estimates, decision support tools, and mission specific requirements. Raw data collected are processed and analyzed iteratively to produce operationally relevant information that supports decision making across a wide range of missions. This framework provides a powerful method to parse the roles that data and models play and to identify the linkages between them.

This report presents the methods and findings from network and metadata analyses of the resource inventory with respect to improvised nuclear devices (INDs). As of this report, 185 interviews have been conducted with 243 people representing 54 federal agencies, divisions, or groups. These extensive and on-going interviews with emergency managers, subject matter experts, and high level decision makers have identified nearly 500 data and modeling resources, of which 138 have been identified as used in the context of emergency management related to IND scenarios and are included in the inventory presented here. Each resource is characterized by a series of metadata tags that describe its function, use, and availability. The list of the IND resources and the associated metadata can be found in Appendix 3.

The resources within the IND inventory have been analyzed to reveal trends in how information is processed. Two types of analyses have been performed: network and metadata analysis. The network analysis was a systems-level analysis to evaluate the robustness and interconnectedness of the IND network that considered the number of users, and the upstream and downstream connections for each resource. Metadata analysis characterized the types of resources used, and identified the major users and producers of data and modeling resources. The primary results from the analyses are summarized below:

- The IND relevant resources are generally well-connected within the network, though some resources, termed orphans, have no upstream or downstream connections.
- Although data and modeling resources are widely used by the federal interagency, only a few resources stand out as being heavily used.
- A qualitative analysis of the flow of information within the IND network shows a relatively clear progression from raw data to event characterization to consequence models. However, resources tagged later in the flow of information, such as decision support tools or mission

specific requirements, are not clearly organized within the network and are not well connected to upstream resources.

- Some resources are more highly linked to other resources in the network and are important information conduits within the network. These resources are central to the flow of information and play a key role in providing hazard- and mission specific information to support operational decision making.
- The bulk flow of information shows that there is limited information passed from consequence models to downstream resources and little flow of information into mission-specific requirements.
- The few impact estimates, decision support tools, and mission-specific requirements available to the federal interagency only support a narrow range of emergency management missions.
- FEMA, followed by DHS, DoD, and DOE, leads the federal interagency in using the available data and modeling resources to support operational decision during all phases of emergency management.

The final product of this project, which will be built in Phase IV, will be an interactive inventory of the data and modeling resources used by the interagency, accessible via a graphical user-interface. The Phase IV report will include an in-depth analysis of the network, identification of relevant gaps, and recommended courses of action to close the gaps. Ultimately, this effort will enable the entire emergency management community to identify and use the resources available to support operational decision making during all phases of emergency management for INDs.

Introduction

A nuclear terrorism incident, such as the detonation of an improvised nuclear device (IND), will have devastating large-scale consequences to public health and safety. At ground zero, the blast will cause mass casualties, destroy infrastructure, damage utilities systems, stall immediate emergency response activities, and will continue to present a challenge during all phases of response and recovery. The immediate emergency management challenges will be determining how to save the most lives and minimize the impact of the disaster, which will require that critical decisions be made quickly. These decisions cannot be made without timely, accurate, and well-coordinated information, which must be collected and analyzed rapidly and then disseminated to relevant stakeholders at all levels.

Data and models have been used for many years to support operational decision-making. The advent of readily-available, high capacity, mobile computing systems have enabled the federal interagency to collect and access an unprecedented amount of incident-relevant information from data, models, and analysis tools. This expansion of data and modeling resources provide a wealth of information that needs to be organized and made accessible to the emergency management community, not just during a response, but during all phases of emergency management. An IND detonation, having no historical precedent, presents a unique challenge because effective strategies to respond and recover from such a scenario must rely heavily on predictive modeling and extrapolation from first principles. The roles and responsibilities for data collection and modeling to characterize the event in the early phases of a response have largely been codified for nuclear detonations through ongoing interagency efforts. However, in the absence of experience, many of the data and modeling resources needed to inform response and recovery operations are less well-defined.

In recognition that informed decision-making is key to successful emergency management, the Emergency Support Function Leadership Group (ESFLG) established the Modeling and Data Working Group (MDWG) in August of 2012 to promote better collaboration between stakeholders across the interagency to identify and characterize the data and models used to support emergency management. The membership of the working group is chosen by the ESFLG and expanded upon request by current ESFLG or MDWG members. Current members include a wide range of emergency managers and subject matter experts from across the interagency, including members from each of the federal Emergency Support Functions as identified by PPD-8. The primary goal of the working group, as defined by the charter, is to identify and characterize the data and modeling resources available to support federal decision-makers during all phases of emergency management, particularly during the time-sensitive period of emergency response. The data collected were analyzed to determine when and how those resources are used in the context of emergency management. The resulting information has been collated into an inventory of the currently utilized resources. The resulting web-based tool will help ensure that decision makers have access to the information they need when they need it to support operational decision making for emergency management.

Defining Data and Models

Models and data are extensively employed across the interagency throughout all phases of emergency management. Given the breadth of information resources used and included in the inventory, the terms ‘model’ and ‘data’ are defined below, as used in the context of this work.

Data are defined as repositories of information used for emergency management. This definition of data encompasses tools that assist in the presentation or visualization of data without transforming the data itself (e.g., FEMA GeoPlatform, see Appendix 3). Data are classified as raw data, situational awareness data, impact estimates, or mission specific requirements. The data within these categories may be steady-state data describing features of the environment during normal operations. Alternatively, they may be event-specific assessment data collected as an event unfolds.

Models are defined as any program, algorithm, or computational tool that transforms or processes data to produce new information. Models are classified as event characterization models and analysis tools, consequence models, and decision support tools. Models accept, as inputs, data that are transformed to provide a new type of information (e.g., NARAC Modeling System, HPAC, NUEVAC, see Appendix 3).

Flow of Information

During management of any emergency, whether an IND or a hurricane, federal, state, and local entities must effectively leverage information from a wide range of sources to answer the fundamental questions of "What happened?"; "Who or what was affected?"; and "What needs to be done?" To address these three foundational questions and the corresponding critical information requirements, the data collected must be processed by models or analysis tools to generate new types of data and information. Once processed, these data can then be used directly for informed decision-making and be further refined using additional models.

To capture this iterative process of data collection and analysis, a conceptual flow of information framework to categorize the types of information, from raw data to operationally relevant information, has been developed and described in the Phase II report. This categorization scheme, shown in Figure 1, contains seven basic types of resources: raw data, event characterization models and analysis tools, situational awareness data, consequence models, impact estimates, decision support tools, and mission specific requirements. In brief, raw data, event characterization models and analysis, and situational awareness data together define the event itself. These data and models tend to be hazard-specific, focused on a subset of event types for which there are well-defined experts and authoritative sources of information. Consequence models and impact estimates provide information about the impact of that event to affected populations, infrastructure, the economy, and the environment, among others. These resources are designed to provide an understanding of the specific event to determine what happened and who and what was affected to what degree. Decision support tools and mission-specific requirements process information about the event and its consequences to determine what needs to be done in order to effectively respond. These data analysis tools and models tend to vary more by mission area than by hazard type.

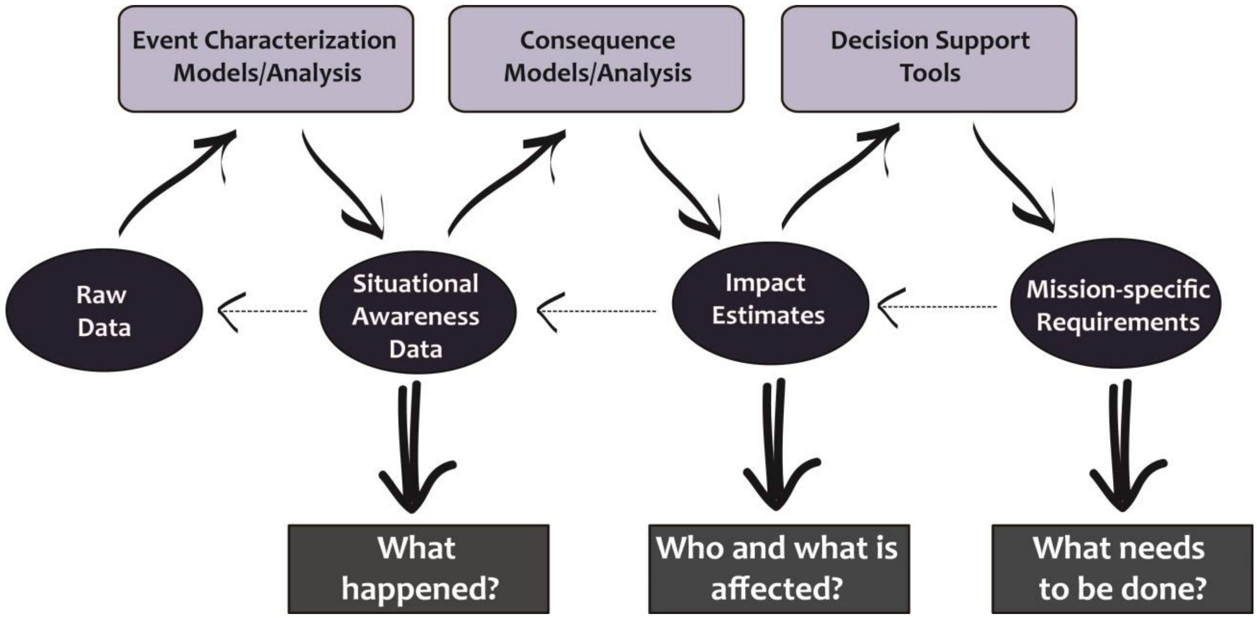

Figure 1. Framework describing the flow of information through iterative rounds of data and modeling. Information flows through modeling and data resource categories to produce answers to the fundamental questions of emergency management.

The flow of information between the categories defined in the framework above is not unidirectional. The data processing is iterative, and feedback loops serve as a mechanism for refining the information produced as new data become available. For example, the outputs of event characterization modeling of nuclear detonation scenarios produce post-event situational awareness data, which can in turn guide the collection and assessment of additional raw data. The Interagency Modeling and Atmospheric Assessment Center (IMAAC) uses the modeling system developed by the National Atmospheric Release Advisory Center (NARAC) and models developed and run by the Defense Threat Reduction Agency under the Department of Defense (including the Hazard Prediction and Assessment Capability, HPAC), to characterize the event and support planning efforts. Following a nuclear detonation, the Federal Radiological Monitoring and Assessment Center (FRMAC) is tasked with organizing and coordinating the collection of radiation assessment data by aircraft, on-the-ground radiation assessment teams, and first responders equipped with dosimeters that report continuous time-stamped and geo-tagged dose rate measurements. These raw data would then be used to refine the inputs for updated runs of the event characterization models managed by IMAAC. These modeling runs produce updated situational awareness data and can be used by FRMAC to refine the collection of additional radiation measurements.

The flow of information framework and categorization system is designed to capture the breadth of how data and models are used by those involved in federal emergency management. The inventory of resources on which the categorization is based includes both computationally-intensive scientific models built and run by technical experts; it also includes simple models or data visualization tools that are used to calculate resource requirements in the field in real time. The framework is flexible, such that one event characterization model can produce inputs for a second and another model or data analysis tool can be categorized as used both for event characterization and consequence modeling. Additionally, not all data or information enter the framework as raw data, nor are all the data the outputs of an upstream model. For example, the National Shelter System, categorized as providing situational awareness data, provides the locations and capabilities of open shelters in the United States, data that are furnished directly from the American Red Cross rather than from a model.

Although mission-specific requirements represent the most narrowly focused, actionable data in the flow of information, emergency managers also refer to situational awareness data and impact estimates when making decisions. These decisions usually feature a broad focus or serve as early estimates. For instance, the situational awareness data found in post-nuclear detonation briefing products generated by the NARAC modeling system greatly influence initial shelter-in-place and evacuation recommendations written into existing federal emergency management plans. As another example, FEMA Individual Assistance consults Preliminary Damage Assessment data (an impact estimate) collected during and after an event to determine the potential required size of its recovery programs. In this way, all categories of data with the exception of raw data are used as the basis for decision-making.

Methods

These methods were first developed and described in the MDWG Phase III report for Hurricanes and Earthquakes.[1] They are included in full detail in this report to delineate how the data was collected, processed and analyzed.

The workflow of analysis performed for this project is divided into three parts: data collection, data processing, and analysis. The workflow is depicted in Figure 2. Data collection was performed through interviews with members of the MDWG, other emergency managers, and subject matter experts to compile an inventory describing data and models that are used. Processing of these data included running algorithms to measure the interdependency of all the resources to aid in further analysis. Two types of analysis were completed: a network analysis based on the upstream and downstream connections of each resource, and a statistical analysis of the resource metadata. The network analysis makes use of network maps, which are visualizations of the resources and the data flow between them that enable trends and relationships to become apparent at a glance. The statistical analysis provides descriptive measures of the types and number of resources in the inventory and their characteristics. Each of the components of this workflow is described in more detail in the subsequent sections.

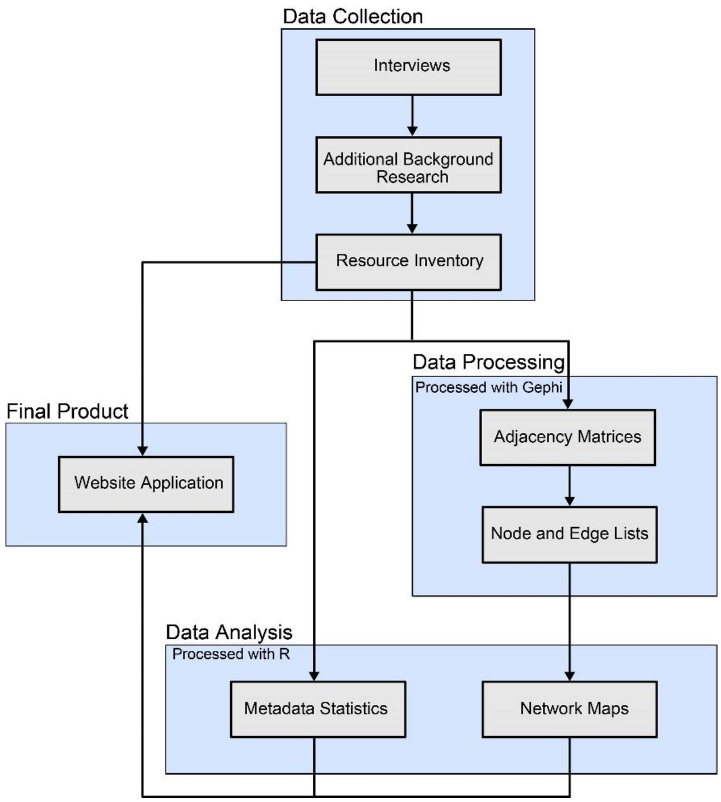

Figure 2. Analysis Workflow. A depiction of the sequence of work involved in producing quantitative analysis of the resource inventory.

[1] Graeden E (January 2014) MDWG Phase III Report Draft: Data and Models for Hurricanes and Earthquakes

Data Collection

Interviews

The information required to analyze the available data and modeling resources was collected through a series of in-person and phone interviews with the members of the MDWG and the subject matter experts they recommended. During these interviews, the users and producers of each resource identified and characterized the ways in which each resource is used to support planning and operational decision making. In most cases, the MDWG members were interviewed initially. Interviews with additional subject matter experts or leadership were scheduled upon recommendation to provide further breadth or depth of information depending on the size of the agency or division represented and the expertise of each interviewee.

As of this report, 185 interviews were conducted with 243 people representing 54 federal agencies, divisions, or groups. Interviews were opened with an introduction to the project. Throughout the project, interviewees have included those who are providers of data or are tool developers; those who are analysts and users of those data and tools; those who make operational decisions informed by data and modeling resources; and those who have roles that include a combination of tool-development, analysis, and decision making. Interviews are designed to capture an overview of the roles and responsibilities of each group and the ways in which data and data processing tools, including modeling, support those roles. The flow of the conversation varied widely based on the expertise of the interviewee and attempted to capture both the general and specific information requirements from each interviewee across the spectrum of emergency management missions and the phases of an emergency. A comprehensive list of the interviewees can be found in Appendix 1.

Resource Inventory

A comprehensive inventory of resources used across the federal interagency and the linkages between them was generated on the basis of the resources discussed during interviews, followed by background research to identify inputs and outputs of each resource. Only those resources meeting one or more of the following conditions were kept in the network:

1. The resource has direct federal users;
2. The resource feeds more than one other resource; or
3. The resource is fed by at least one resource and feeds another resource.

The first condition ensures that all resources known to be used by the interagency for emergency management are shown in the inventory. The second two conditions ensure that all those resources that directly contribute to the flow of information between resources are captured and can be analyzed as part of the resource networks. For instance, the North American Mesoscale Model (NAM) has no direct federal users, but two upstream resources feed into NAM, which in turn feeds 11 downstream resources. These rules focus the analysis on resources directly used by the interagency for emergency management and the resources that connect those resources. In this way the inventory can both act as an accessible catalog of useful resources for emergency management as well as a means to analyze the flow of information between resources and how data is processed into information useful to support federal decision making.

Those resources identified through interviews and research that were inputs for only one other resource and had no federal users of their own were excluded from the inventory and were "wrapped" into their downstream resources. For example, the Second-order Closure Integrated Puff (SCIPUFF) model is an input for the Hazard Prediction and Assessment Capability (HPAC), but it is only used for emergency management through HPAC. In this case, SCIPUFF was not included in the inventory, and a description of its role was included in the summary of HPAC.

Many resources under development or not currently used to support emergency management activities were identified, but not included in the inventory. Information about these resources has been retained and will be used in future iterations of the report to suggest mechanisms to fulfill any gaps identified in the network of resources.

Resource Inventory Metadata

The flow of information framework captures the functional, time-dependent, and mission-specific variation between resources used across the federal interagency. However, it does not describe other essential characteristics such as how those resources are accessed, used, and updated. These additional characteristics, or metadata, must also be collected to properly organize and analyze the resources to maximize effective usage during all phases of emergency management. These metadata will appear in the interactive inventory of resources upon completion of the project.

Metadata categories include: the resource's full name, abbreviation, model/data, owner, users, upstream resources, downstream resources, relevant hazards, core capabilities supported, emergency support functions (ESFs) supported, recovery support functions (RSFs) supported, key words, function tags, resource type, data collection method, phase specific utility, access information, access type, processing requirements, refresh rate, last known version, programming language, file type, contact information, contact during activation, website, and a brief summary of its function and use. Complete descriptions of each metadata tag are included in Appendix 2.

Data Processing

To build network maps describing the linkages between resources in the inventory, the metadata defining the upstream and downstream linkages for each resource was quantified in an adjacency matrix. An adjacency matrix is a mathematical method of representing a network that provides a simple way to calculate many network measures and statistics.[2] The adjacency matrix was then converted into separate node and edge lists. A node is a point on a network, and in this case, each node represents a single resource in the inventory. The nodes list contains the metadata of each node in the network, allowing that information to be visualized on the network map and analyzed in the context of the network. An edge is a line in the network that connects two nodes, and in this case, represents the transfer of information from one resource to another. The edge list contains a list of connections between nodes in the network. These node and edge lists were imported into Gephi,[3] an open source network visualization and analysis software, to create the network maps used in the analysis.

---

[2] A short, rigorous definition of an adjacency matrix: For a network of n nodes, the adjacency matrix A is an n x n matrix where the i,j^{th} entry in the matrix represents the number of connections from the i^{th} node in the network, to the j^{th} node in the network.

[3] Bastian M., Heymann S., Jacomy M. (2009). ***Gephi: an open source software for exploring and manipulating networks.*** International AAAI Conference on Weblogs and Social Media.

All data processing was performed using R, an open source, statistics-based programming language.[4] R was chosen because of its ease and efficiency in calculating basic and network-based statistics. An open source language, this coding language facilitates transfer of the analysis scripts to another party.

Data Analysis

To visualize the data contained in the resource inventory, network maps were generated from the node and edge lists in Gephi. In the network map, every node represents a single resource in the inventory and is sized proportionally to the number of users of the resource. Edges represent a flow of information from one resource to another and are graphically displayed as a clockwise arc, from the source node to the target. In this case, the source node is the upstream resource. A downstream resource is defined as the one that the source node feeds. Figure 3 illustrates an example of a simple network map. Unless explicitly stated otherwise, the nodes in each network are arranged by a Force-Direction algorithm that groups closely linked nodes. This algorithm treats each node as a charged particle that repels all other nodes, and each edge as a spring, pulling the nodes back together. Three attributes of the network described below — community structure, betweenness centrality, and resource connectivity—were explored using network maps.

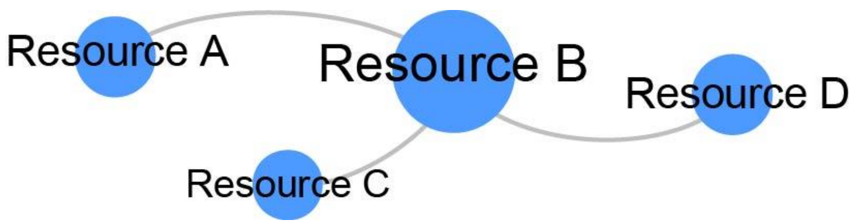

Figure 3. Example of a simple network map. Individual resources are represented by blue discs (nodes). Direct connections between resources are represented by gray curved lines (edges). The flow of information travels clockwise. In this example, information flows into Resource B from Resources A and D. Information from Resource B flows into Resource C. The size of each node can convey additional information; for the network maps presented in this report, nodes are sized relative to the number of users of that resource.

Betweenness Centrality

The importance of specific nodes was also investigated using the betweenness centrality measure, which is the most common centrality measure that characterizes how often a node is between other nodes in the network.[5,6] Specifically, the betweenness centrality of a specific node is calculated as the number of times that node appears on the shortest path between any other two nodes in the network, and measures the degree to which a node acts as an intermediary between other nodes. With betweenness centrality, the most important nodes are those that act as “shortcuts” or “bridges” between different parts of the network. However, betweenness centrality only considers the shortest paths between nodes and therefore gives no weight to alternative paths over which information could be passed within a network. Nodes were colored on a gradient such that more central nodes were darker and less central nodes were lighter.

---

4 R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

5 Freeman LC (1977) A set of measures of centrality based on betweenness. *Sociometry***:** 35-41

6 Freeman LC (1979) Centrality in Social Networks Conceptual Clarification. *Social Networks* 1**:** 215-239

Resource Connectivity

Directed networks can further designate in-degree (the number of incoming edges) and out-degree (the number of outgoing edges). According to these measures, the most important nodes in a network are those that are connected to the largest number of other nodes, regardless of their position in context with the rest of the network. A node's in-degree is defined as the number of nodes feeding into it (in this case the number of upstream resources) and a node's out-degree is the number of nodes it feeds into (the number of downstream resources). A node's degree is the sum of its in-degree and out-degree, signifying the total number of connections that node makes to another node. These measures were used in an analysis of the flow of information to organize the nodes in space, comparing their relative in-degree and out-degree.

Results: Network Analysis

A network is defined as a system consisting of interconnected components where network analysis is the process of understanding the connections between those components. The individual components of the network are called nodes and the connections between them are called edges, with information moving through the network by a defined, or directed, flow. The analysis presented here describes the connections between the data and models used by the federal interagency in the context of emergency management. Over the course this project, nearly 500 modeling and data resources have been identified, researched, and vetted, of which 138 are included in the IND resource inventory (Appendix 3). This inventory includes a dataset of over 20 metadata characteristics describing each of the resources. Two metadata categories (upstream and downstream resources) describe linkages between the resources based on the flow of information between those resources. These linkages were used to build a flow-based network of the datasets and models collated in the inventory. This dataset, including the resources and their associated metadata, and the network based on this dataset, was used to perform a preliminary analysis of the IND resource inventory, as described in the following section.

Resource Network Overview

The network maps described in this report visually represent the flow of information between resources used by the federal interagency to support decision making during emergency management. In these networks, each dataset or model is a node in the network with each edge representing the flow of information and processing of data as it passes between those nodes. The size of a node and its label is directly proportional to the number of users of that resource, an indicator of the relative utility of each resource, which is defined by the number of federal agencies that directly use the resource in the context of their work.[7] The edges curve in a clockwise fashion, distinguishing which resource is the source and which is the target of the information. Both the inputs (upstream resources) and outputs (downstream resources) of each resource in the network were identified based on in-depth analysis of interview data and a review of the technical documentation of the resource, when available.

Network Overview

A network of resources used in the context of an IND is shown in Figure 3. Each resource (node) is shaded by where it falls in the flow of information: resources early in the flow of information, tagged as raw data or event characterization models are the lightest in color; resources tagged as decision support tools or mission specific requirements are darkest in color. A qualitative analysis of the network indicates that the resources used by the emergency management community are generally well-connected, with information flowing throughout the network. Disparate types of resources, such as weather, infrastructure, population, imaging, and other types, are interconnected. There is a relatively clear flow of information from weather-related resources (raw data and event characterization) and to atmospheric dispersion modeling (event characterization and situational awareness data) to consequence modeling. A few groups are clearly outliers: the cluster around EAGLE-I, the energy model owned by the Department of Energy, is, for example, only linked to the rest of the network through a single resource, MedMap (Department of Health and Human Services).

7 Note that users could also be calculated by including not only the number of direct users, but also those users of all resources that provide inputs for a given resources. We refer to this latter method as calculating "cumulative users", a method that significantly increases the number of users for resources that fall in the Raw Data and Event Characterization categories, for example.

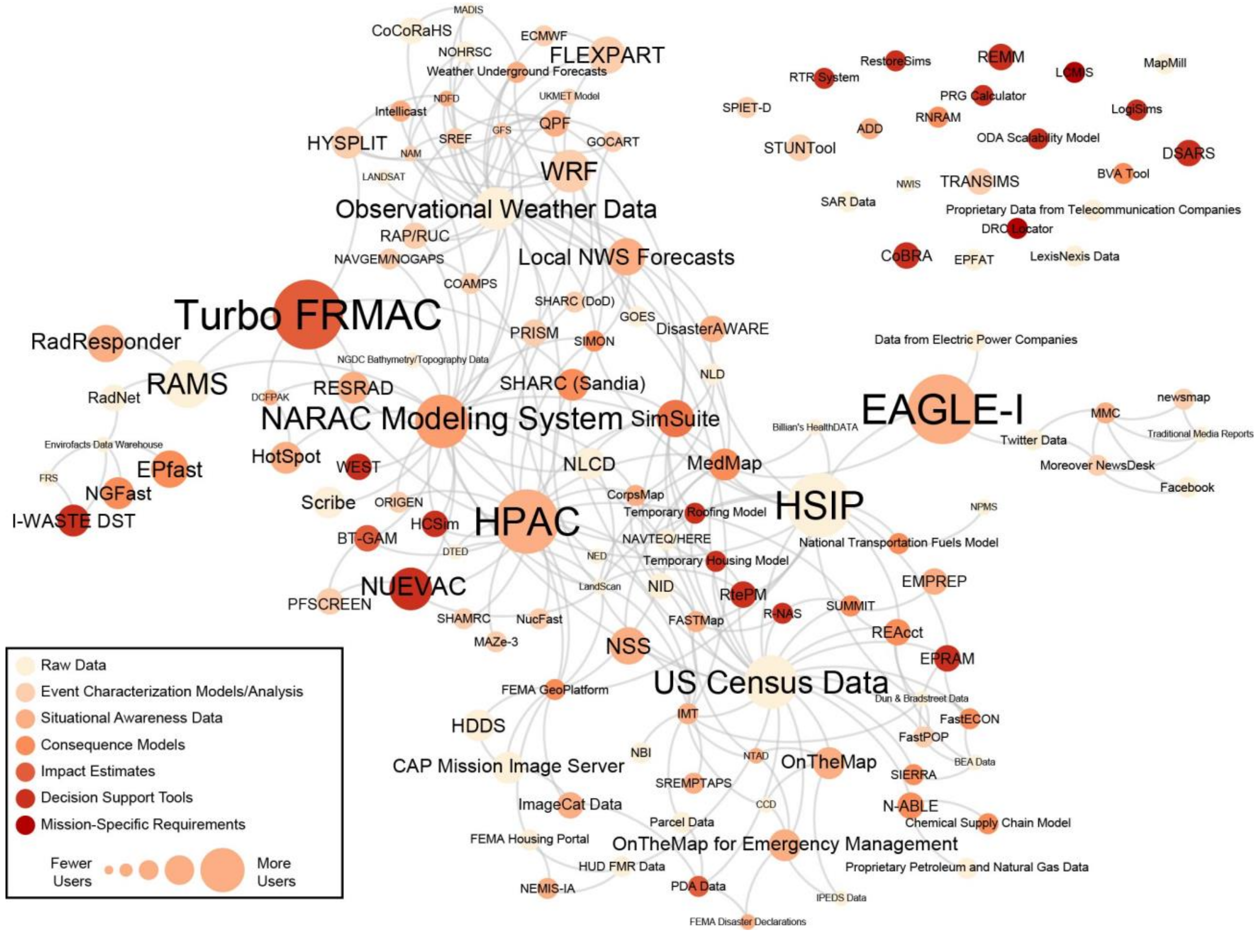

Figure 3. IND Resource Network. In this network, each node (circle on the graph) represents a resource in the inventory and is sized proportionally to the number of organizations that use that resource across the federal interagency. Edges, the curved lines connecting two nodes, represent information passing from one resource to another. The edges curve in a clockwise fashion, distinguishing which resource is the source and which is the target of the information.

Flow of Information within the Network

Although a relatively clear flow of information is seen from raw data to event characterization to consequence models, this trend is less apparent beyond consequence models. As seen in Figure 3, the darker colored resources, the decision support tools and mission specific requirements, are randomly distributed throughout the network. This suggests there is a breakdown in the flow of information from consequence models to downstream resources that can be used to inform the question as to what must be done (see additional analysis in Figure 6). The vast majority of darker colored resources are also not as heavily used (small nodes) as resources located earlier in the flow of information.

A large percentage of decision support and mission-specific resources are unlinked to the rest of the network (see Table 2 and Figure 4 for a more in-depth discussion of "orphan resources".) This lack of resource integration in the network suggests a breakdown in the sharing of information from consequence models to the more operationally-targeted resources that would be used by operations personnel prior to or during an event.

Not surprisingly, just as weather features prominently in the most used resources for hurricanes and earthquake scenarios, so is it widely used for IND scenarios.[8] In the case of an IND scenario, weather conditions are crucial to determine the direction of the radioactive plume, which is necessary to inform response activities, such as deployment of first responders, or determination of safe evacuation routes. This finding highlights and reinforces the role and impact of weather across hazards and the importance of maintaining a robust weather forecasting infrastructure to support all-hazards emergency management, not just hazardous weather events, like hurricanes.

Most Heavily-Used Resources

Although the network is well connected, only a few of the datasets and models are heavily used by the federal interagency in the context of emergency management. The most heavily used resources (6 out of 138 resources) with at least 7 users are listed in Table 1.[9] Of these resources, half are multi-tagged and four of six include some consequence modeling capability. These widely-used resources demonstrate that the information used most in the context of emergency management is both about the event itself and about the damage that cascading effects cause. This distribution may also reflect the fact that there are many more of these model types in the inventory, with many fewer resources categorized as impact estimates, decision support tools, and mission specific requirements.

The most heavily used resource in both networks is EAGLE-I, the recently-developed consequence model from the Department of Energy that provides real time information about electricity outages. EAGLE-I is heavily used in part because it provides information that was not previously available about the stability of a sector with consequences that have a large impact on all other sectors. However, this model is notably separated from the rest of the network (see Figure 3), and, while the majority of other resources on these lists are also highly central (see Figure 5), EAGLE-I is not yet well integrated into the interagency information networks. By contrast to most of the other heavily-used resources, EAGLE-I is a relatively young resource for which technical solutions to link to other resources have not yet been developed. This lack of integration suggests an area that may warrant further investment by the interagency.

8 Graeden E (January 2014) MDWG Phase III Report Draft: Data and Models for Hurricanes and Earthquakes

9 There are no resources with 6 users; an additional 12 resources have either 4 or 5 users.

Table 1. IND resources with the most federal users. Resources with at least 4 federal users are listed in decreasing order of number of users. Resources with the same number of users are listed alphabetically.

Resources	Users	Hazards	Resource Types	Descriptions
EAGLE-I	10	All-Hazards	consequence model	Models and monitors electric grid impacts
HPAC	9	IND	event characterization models/analysis; consequence model	Models CBRNE atmospheric dispersion and impacts
HSIP	9	All-Hazards	raw data	Critical infrastructure and key resource data
Turbo FRMAC	9	IND	consequence model; decision support tool	CBRNE assessment data analysis
NARAC Modeling System	7	IND	event characterization models/analysis; consequence model	Models CBRNE atmospheric dispersion and impacts
US Census Data	7	All-Hazards	raw data	Regional populations, demographics, and survey items

Orphan Resources

In the IND network, there are several resources that are not linked to any other resource (see Figure 3, 4). These "orphaned" resources (23 out of 138) do not receive or share information with any other resource in the network. While these resources are all identified as being used within the federal interagency, none of their results are linked electronically to other resources to be processed and analyzed, nor are their results derived from hazard-specific information produced or processed by upstream resources.

Interestingly, the majority of resources are either tagged as raw data or as resources later in the flow of information framework, including decision support tools and mission specific requirements. Most likely, the raw data resources are those that have not yet been incorporated into event characterization or consequence models. These datasets, if linked to relevant downstream resources, may be useful to refine and improve the parameters of existing models. Effective decision-support tools and mission-specific requirements should be linked to upstream resources to ensure that the information provided is based on event-specific empirical data.

Of the 23 orphaned resources, seven are IND-specific resources. These resources are all models (event characterization models, consequence models, and decision support tools) that need to be linked to the rest of the network to ensure that their outputs are used appropriately and their inputs are informed by event-specific data. Over half of the orphaned resources are owned or housed by industry or the national laboratories, suggesting that when these resources are shared with the federal interagency, additional investment could help improve their integration with the larger information sharing network.

Table 2. Orphaned resources. These resources do not have any upstream or downstream linkages within the IND network. Resources are ordered by where they fall in the flow of information.

Resources	Hazards	Resource Types	Descriptions
Airport Facilities Database	All-Hazards	raw data	Airport facility locations and data
Communications Licensing	All-Hazards	raw data	Communications infrastructure

Database Extracts			locations
EPFAT	All-Hazards	raw data	Dataset of facility emergency power requirements
LexisNexis Data	All-Hazards	raw data	Census block-level insurance information from LexisNexis
MapMill	All-Hazards	raw data	Aerial imagery converted to maps by crowdsourcing
Proprietary Data from Telecommunication Companies	All-Hazards	raw data	Selectively shared, proprietary telecommunication data
SAR Data	All-Hazards	raw data	Synthetic Aperture Radar data describing the Earth's surface
ASPECT	All-Hazards	event characterization models/analysis	Airborne, real-time environmental sampling and data collection
NucFast	IND	event characterization models/analysis	IND explosion effects model
SPIET-D	IND	event characterization models/analysis	Post-IND charged particle formation model
STUNTool	IND	event characterization models/analysis	Tunnel explosion effects model
TRANSIMS	All-Hazards	event characterization models/analysis	Transportation Analysis and Simulation System for regional transportation modeling
ADD	All-Hazards	situational awareness data	Federal Emergency Management Agency automated database for personnel tracking
BVA Tool	IND	consequence model	Blast vulnerability analysis for water infrastructure
RNRAM	IND	consequence model	Radiological/nuclear threat risk assessment model
DSARS	All-Hazards	impact estimates; mission-specific requirements	Automated reporting system for Federal Emergency Management Agency disaster services
LogiSims	All-Hazards	decision support tool	Resource allocation decision support software
ODA Scalability Model	All-Hazards	decision support tool	Small Business Administration loan application volume model
PRG Calculator	IND	decision support tool	Preliminary Remediation Guides calculator for radiation clean-up
RestoreSims	All-Hazards	decision support tool	Resource allocation decision support software
RTR System	IND	decision support tool	System to establish post-IND radiation triage, treatment, and transport sites
DRC Locator	All-Hazards	mission-specific requirements	Locations and statuses of Disaster Recovery Centers
LCMIS	All-Hazards	mission-specific requirements	Federal Emergency Management Agency database for disaster relief supplies tracking

Resource Centrality

Betweenness Centrality

The degree of integration of each resource within the network can be quantified by betweenness centrality, a common centrality measure that characterizes how often a node is found between other nodes in the network.[10,11] Specifically, the betweenness centrality of a specific node is calculated as the number of times that node appears on the shortest path between any other two nodes in the network and measures the degree to which a node acts as an intermediary between other nodes. These nodes act as bridges between different parts of the network. As shown in Figure 5, of the 138 resources in the IND network, four stand out as highly central: NARAC Modeling System, Turbo FRMAC, HPAC and HSIP (see Figure 4).

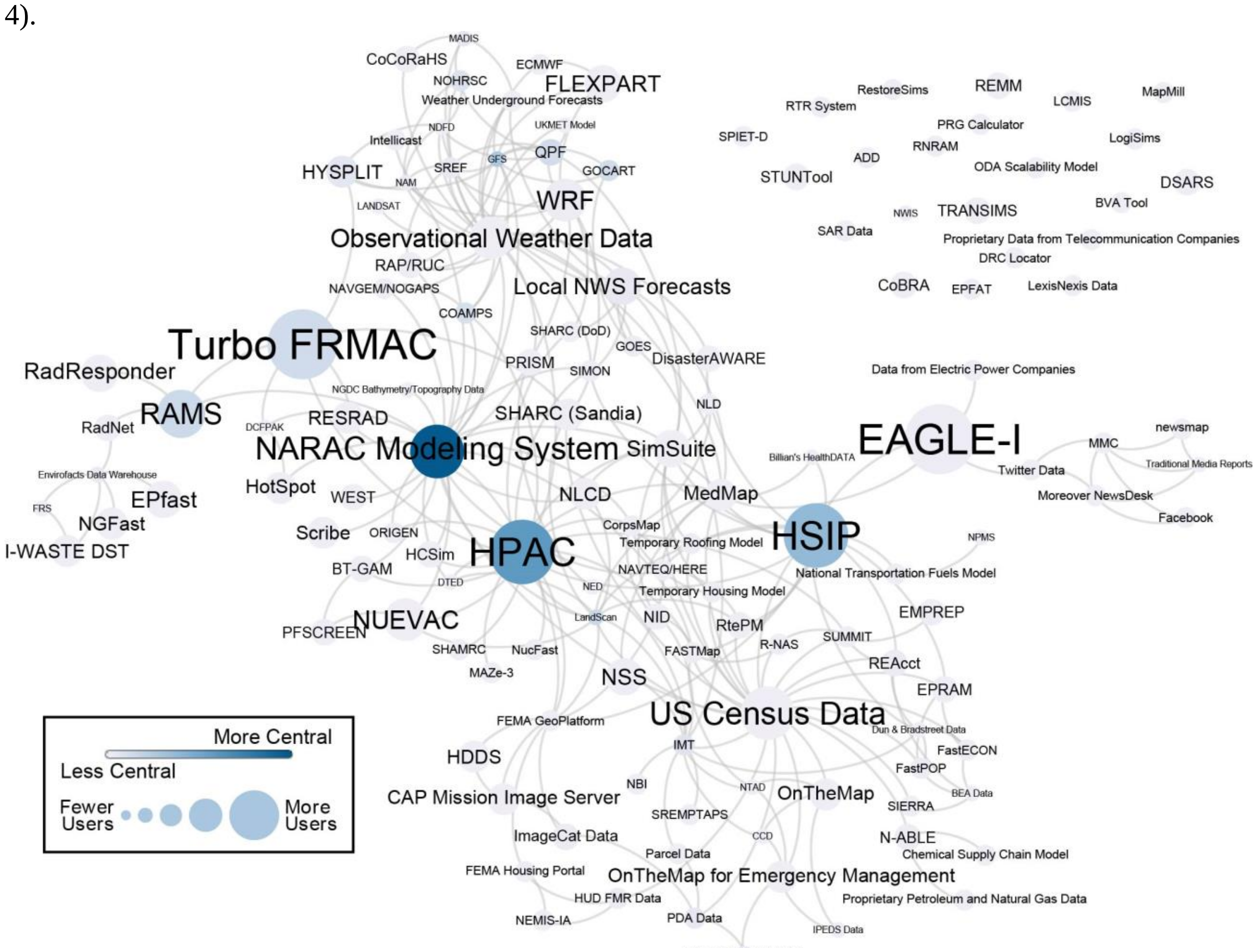

Figure 4. IND Betweenness Centrality. In the IND network, each node (circle on the graph) represents a resource in the inventory and is sized proportionally to the number of organizations that use that resource across the federal interagency. Darker blue represents more central resources, while lighter blue represents less central resources. Edges, the curved lines connecting two nodes, represent information passing from one resource to another. The edges curve in a clockwise fashion, distinguishing which resource is the source and which is the target of the information. Only IND and all-hazards resources from the inventory appear in the network.

10 Freeman LC (1977) A set of measures of centrality based on betweenness. *Sociometry*: 35-41
11 Freeman LC (1979) Centrality in Social Networks Conceptual Clarification. *Social Networks* 1: 215-239

Low-centrality Resources

There are a number of heavily used resources within the network that are not also central. However, a lack of centrality does not necessarily indicate a failure in the network. Raw data and mission specific requirements (early and late in the flow of information, respectively) are expected to have low centrality values because they are sources (raw data) or sinks (mission specific requirements) of information. For example, US Census Data, a source of raw data, is widely used, but has a low centrality measure. Therefore, while centrality is a useful indicator of resources critical for the integrity of the network, it is not the only indicator of resource value within the network.

Widely used resources that have low centrality values can, however, indicate network components that could be better integrated and linked. For example, the consequence model EAGLE-I has the largest number of users, but has a very low centrality value. Based on the flow of information, consequence models would be expected to be the most central resources because they act as an essential link between event characterization and decision support tools. EAGLE-I and EPFast, both consequence models, have very low centrality measures, although both would be expected to link upstream to event characterization models and situational awareness data and downstream to decision support tools and mission specific requirements. This lack of centrality indicates that even some of the most heavily-used consequence models are not well integrated into how information is processed, analyzed, or used by the interagency.

Linkages Between Resources within the Flow of Information

A simpler measure of centrality is the total number of upstream and downstream resources for each resource within the network. As described previously in brief, the degree of connectedness would be expected to correlate with the position of resources within the flow of information. Raw data resources, the source for feeds downstream, are expected to have downstream, but not upstream, linkages. Conversely, mission specific requirements would be expected to have upstream, but not downstream linkages.[12] Consequence models, the mid-point in the flow of information, would be expected to have a large number of both upstream and downstream linkages. The correlation between the number of upstream and downstream connections within the network and the position of resources within the flow of information is shown in Figure 5.

Markedly, the flow of information is not well-correlated to the number of upstream and downstream resources within the existing network. A couple of heavily used raw data resources (Observational Weather Data and US Census Data) have a very large number of downstream resources, as expected. Similarly, the NARAC Modeling System is a consequence model with the most upstream and downstream resources in the network. Moreover, while decision support tools and mission specific requirements would be expected to have upstream linkages, those resources are notably unrepresented on the X-axis of the graph: they appear to be generally unconnected from the rest of the resources in the network. This lack of connectedness suggests that they are not effectively pulling from existing data sets that can be used to define the event in real-time.

---

12 While there are some feedback loops within each category, these feedback loops do not represent the bulk flow of information in the system. See Figure 6 for a more complete discussion.

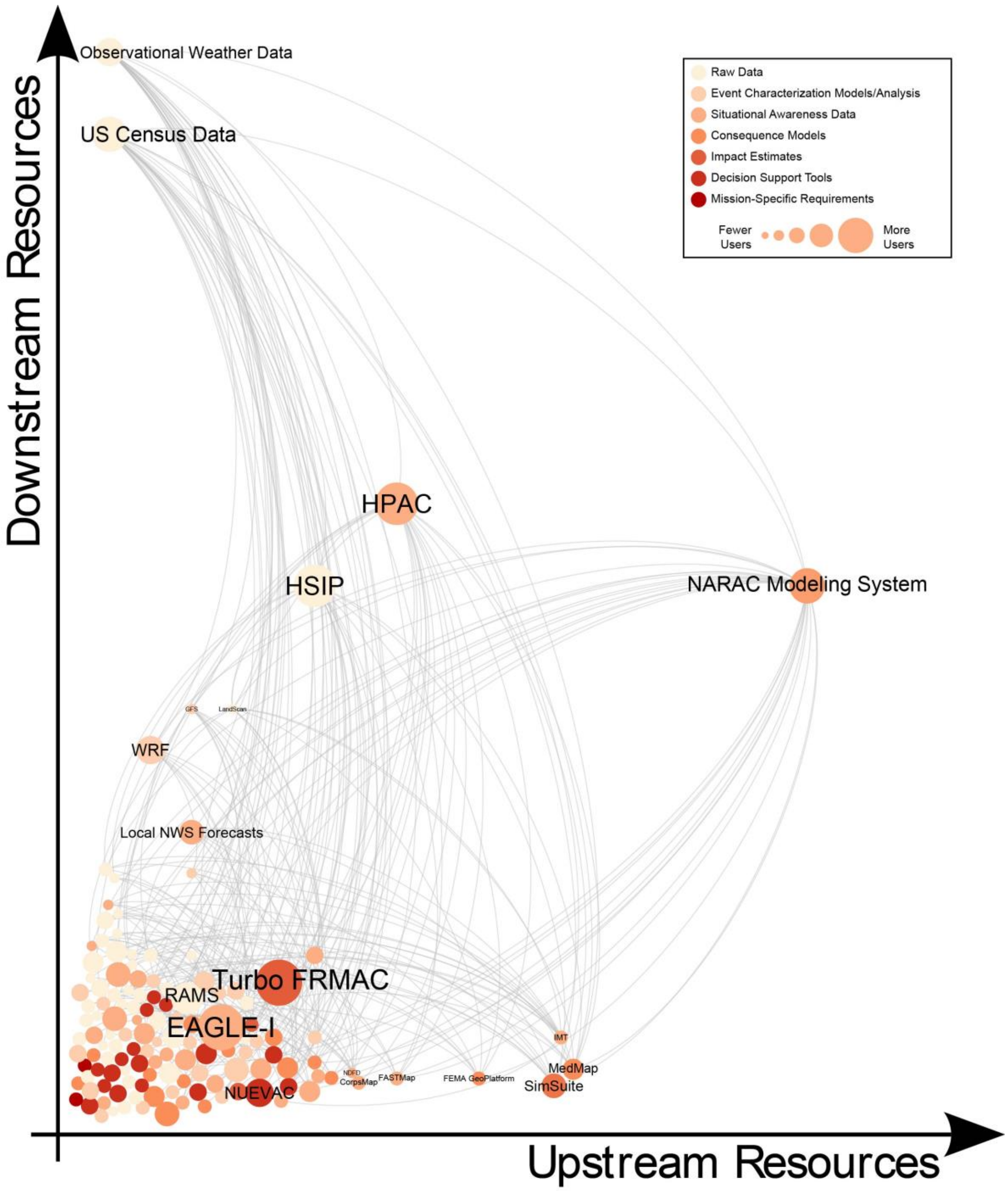

Figure 5. Centrality relative to the flow of information. In this figure, each node (circle on the graph) represents a resource in the inventory and is sized proportionally to the number of organizations that use that resource across the federal interagency. Darker red represents more central resources, while lighter red represents less central resources. The resources are graphed according to the number of upstream resources and the number of downstream resources. Node locations were adjusted slightly in order to display all resources in the network, and should not be interpreted absolutely but rather only relative to other nodes. Edges, the curved lines connecting two nodes, represent information passing from one resource to another. The edges curve in a clockwise fashion, distinguishing which resource is the source and which is the target of the information. Only IND and all-hazards resources from the inventory appear in the network.

Bulk Flow of Information

The flow of information from raw data to mission-specific requirements can also be analyzed by combining all resources within each category and visualizing the linkages between the categories, weighted by the number of resources in each category and the number of linkages between them (see Figure 6). Resources that were tagged as multiple resource types were duplicated and separated into each of those resource types to accurately represent how data is processed, even within a single resource. Not surprisingly, the primary sources of all information are raw data resources. Raw data feeds not only event characterization models directly downstream, but all other categories as well. This analysis highlights the non-linear aspects of the flow of information: information moves from upstream categories to all other categories. A relatively limited amount of information moves through feedback loops from resource categories later in the flow of information to resource categories earlier in the flow of information.

This method of analysis highlights a major gap in the flow of information: there is limited information passed from consequence models to downstream resources and very little flow of information into mission-specific requirements. This lack of connectedness suggests that these resources are unable to use data from upstream resources that provide real-time information. This lack of linkages also represents a failure of information sharing between those communities producing real-time event characterization and consequence modeling and those in the emergency management community tasked with performing mission-specific operations.

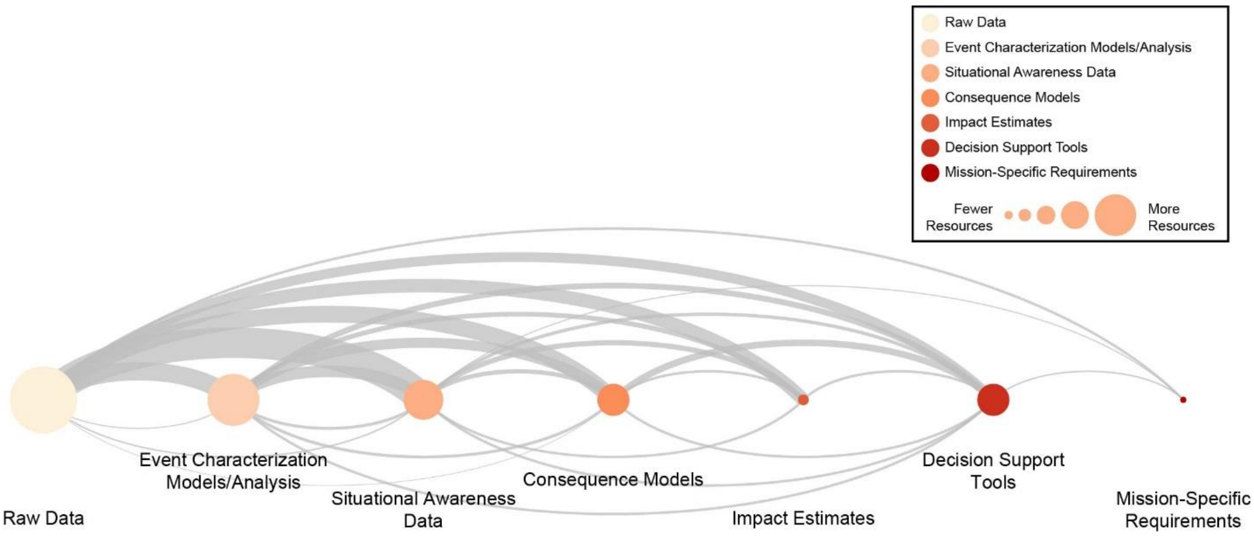

Figure 6. Bulk movement through the flow of information categories. Each node (circle on the graph) represents a resource in the inventory and is sized proportionally to the number of resources in the category. Nodes are graded by red to indicate progression through the flow of information. The resources are graphed according to their position in the flow of information. Edges, the curved lines connecting two nodes, represent information passing from one resource to another and are sized according to the number of connections between those resource types. The edges curve in a clockwise fashion, distinguishing which resource is the source and which is the target of the information. Only IND and all-hazards resources from the inventory appear in the network.

Results: Metadata Analysis

In addition to the network analysis, the metadata describing each resource were also analyzed. These preliminary analyses address the types of resources and information, such as agency-specific use of resources, that are available to support operational decision making in the context of emergency management.

Resource Type

Number of Resources Based on the Flow of Information

Resource types are defined by the categories described in the flow of information in Figure 1. While each resource type is required for efficient and informed disaster planning and response, the resources identified in the inventory are unevenly distributed between resource types (see Figure 7).

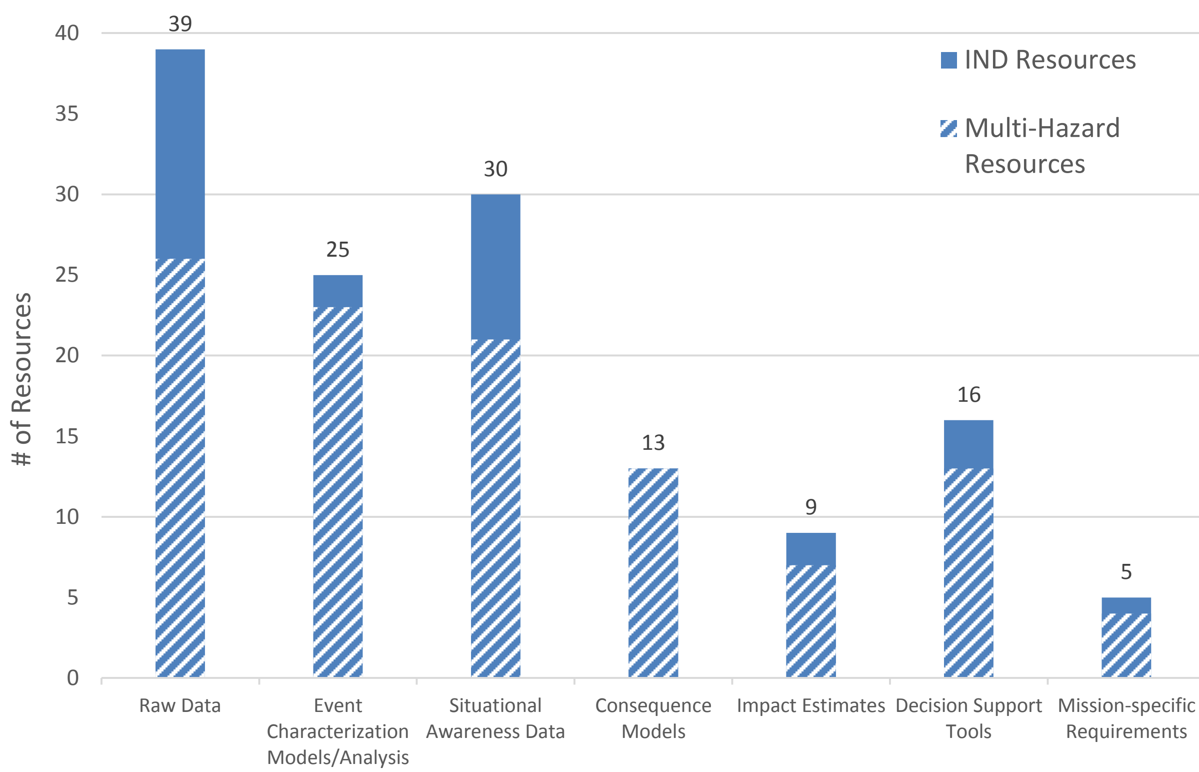

Figure 7. Number of Resources by Type. Each bar represents the total number of used resources that are relevant to INDs for each resource type from the Flow of Information. Resources only relevant to INDs are shown in blue. Resources tagged as relevant to multi-hazards are shown in white hatch marks. The total number of resources for each resource type is shown above each bar.

For IND scenarios, there are many more resources tagged early in the flow of information (raw data, event characterization, situational awareness, and consequence models), as there are later. To some degree, this trend is not surprising. It is intuitively obvious that a great deal of raw data from a wide range of sources is necessary to feed robust event characterization models. For example, many weather models and observational weather data feed into event characterization models, such as the NARAC Modeling System and HPAC. Both of these event characterization models can also provide estimates of consequences, such as casualties, and building damage. As such, these resources pull from raw data sources that provide infrastructure (e.g., HSIP) and population statistics (e.g., US Census Data).

The majority of the resources in each category are multi- or all-hazards, useful to support emergency management across a wide range of hazard types. Not surprisingly, the percentage of hazard-specific resources decreases as the information moves from raw data to mission-specific requirements. Those resources tagged early in the flow of information tend to be tailored toward the collection and processing of data specific to a hazard type (e.g., radiation readings performed by the Aerial Measuring System), whereas mission-specific requirements are more tailored to activities that need to be performed during any emergency (e.g., calculating resource requirements using FEMA Logistics' LCMIS tool).

Resources Tagged as Multiple Resource Types

The vast majority of the resources in the IND inventory are tagged by a single resource type. However, a small subset of the resources (16 of 138) is best described by more than one resource type (see Table 3).

Of the resources tagged by multiple hazard types, four are agency-specific situational awareness viewers that incorporate the outputs of consequence models. These resources are tagged as both situational awareness data and impact estimates: FEMA GeoPlatform (FEMA), MedMap (HHS), SIMON (State), and SimSuite (USACE).[13] SUMMIT (DHS/FEMA) is a tool designed to combine multiple model types to generate comprehensive modeling outputs with code that calculates results related to each category. DSARS (Red Cross) is both an impact estimate as well as mission specific requirement as it is used to track damage assessments, supply needs, and staffing needs for the American Red Cross.

Seven of the nine IND-specific resources with multiple resource types are tagged as both event characterization and consequence models. All these models not only characterize specific aspects of the event, but also estimate the consequences on the human health, building damage, infrastructure damage, and equipment damage, among others. BT-GAM (HHS) is used primarily as a decision-support tool to define the amount and type of resources needed to respond to an IND detonation, but can also be used to predict the health impacts to the population following such an event and so is tagged as both a consequence model and a decision-support tool. Turbo FRMAC (Sandia National Laboratories) is tagged as both a consequence model and a decision support tool as it accepts validated field sample data through RAMS and dispersion model outputs and provides actionable information on the basis of those data. Specifically, this tool can be used to address emergency management questions as to whether radiation doses exceed city, state, or federal limits; whether crops are safe for consumption or should be destroyed; and whether residents need to be evacuated or sheltered in place, among others.

[13] Parenthetical agencies are the owner of each resource.

Table 3. Resources with multiple resource types. Resources applicable to all-hazards are listed first and with boldface names. IND-specific resources are listed second and without boldface names.

Resources	Raw Data	Event Characterization Models/Analysis	Situational Awareness Data	Consequence Models	Impact Estimates	Decision Support Tools	Mission-Specific Requirements
DSARS					X		X
FEMA GeoPlatform			X		X		
MedMap			X		X		
SIMON			X		X		
SimSuite		X	X	X	X	X	X
SUMMIT		X		X		X	
BT-GAM				X		X	
EMPREP		X		X			
HotSpot		X		X			
HPAC		X		X			
NARAC Modeling System		X		X			
RESRAD		X		X			
SHARC (Sandia)		X		X			
SREMPTAPS		X		X			
Turbo FRMAC				X		X	

Agency-Specific Resources

Recognition of a need for rapid sharing of real-time information during emergencies and rapid technological advancements has led to an explosion of data and modeling tools in the last few years. Lists of available information resources within the federal emergency management community have been generated previously; none of these have, to our knowledge, identified which of those resources are used and by whom. To address that gap, interviewees across the federal interagency were asked not only which resources that they have developed or produced, but which resources they use in the context of their mission in emergency management. Only those resources identified as used directly were included in the inventory, and each resource was tagged by the federal agency identified as using the resource directly.[14] The results of this analysis are shown in Figure 8. Note that a resource is defined as used by an agency only if they use the resource directly; upstream resources or feeds of used resources are not included.

[14] Note that users can be defined as either individuals or agencies; for the purposes of our analysis, users are defined as federal government agencies or organizations explicitly included in the Emergency Support Functions, as described in the National Response Framework. Users can also be calculated by including not only the number of direct users, but also those users of all resources that provide inputs for a given resources. We refer to this latter method as calculating "cumulative users."

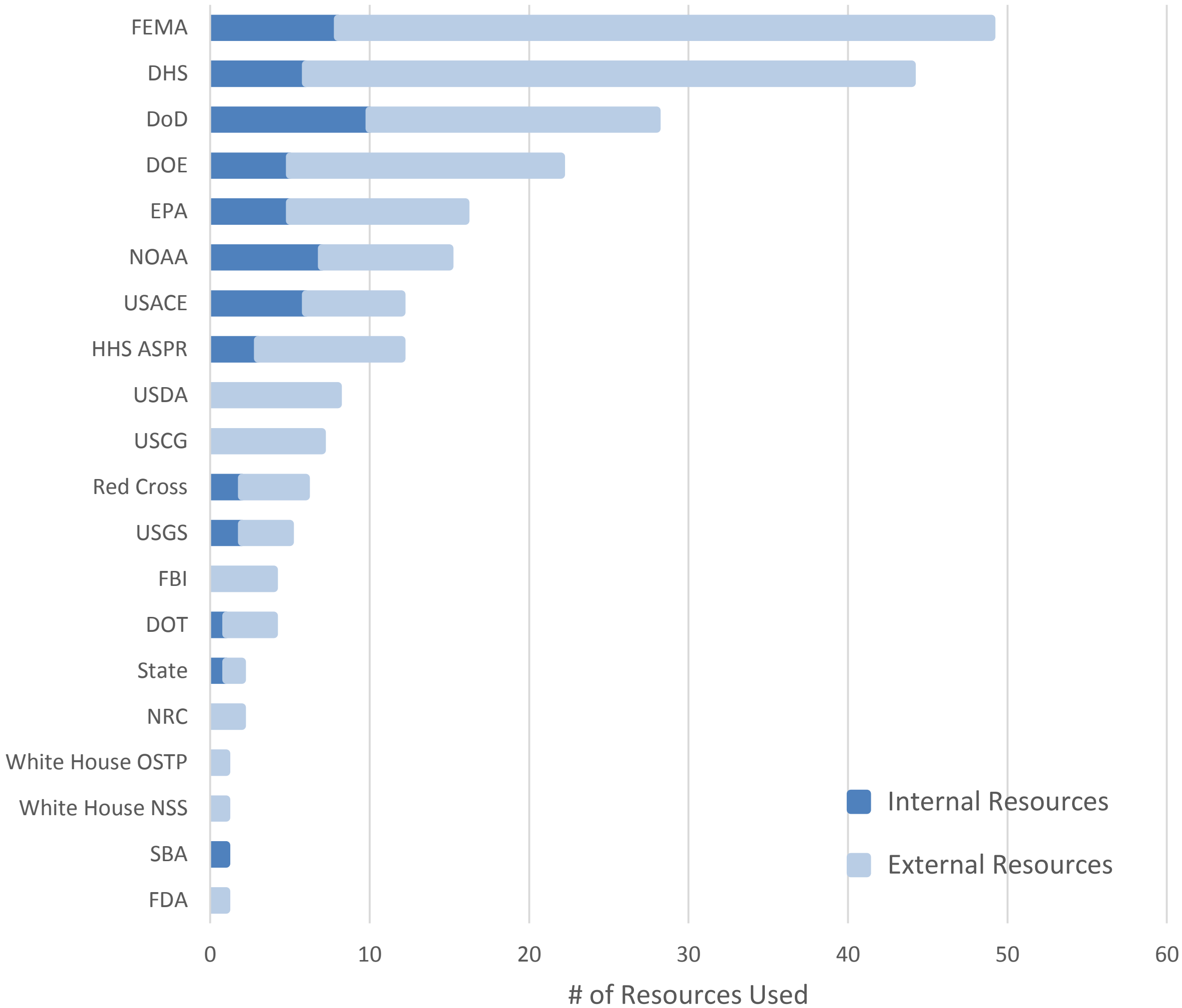

Figure 8. Number of IND Resources Used by Organization. The number of IND-applicable resources each federal organization, agency or department uses is shown. The number of used resources owned (internal) is shown in dark blue. The number of resources used, but not owned (external), is shown in light blue.

The Federal Emergency Management Agency (FEMA) is the largest user of interagency resources, followed by the Department of Homeland Security (excluding FEMA), the Department of Defense (including both the Defense Threat Reduction Agency and NORTHCOM), and the Department of Energy (including the National Laboratories). Because FEMA is tasked with coordinating efforts between all other agencies involved in emergency management, it is not surprising that they are heavy users of these resources from across the interagency. Other organizations have more specific missions and therefore use only a subset of resources relevant to that mission.

Almost all agencies use more resources owned by other organizations (external resources) than those they own themselves (internal resources). Most users use a mix of largely external resources and a smaller fraction of internal resources. Several agencies that use relatively few resources, like the Federal Bureau of Investigation (FBI), use only external resources. The SBA is the only exception that uses only internal resources.

Conclusions

Through interviews and background research, 138 IND and all-hazards data and modeling resources have been identified as being used to support the federal interagency and have been included in the inventory. Over 20 different metadata tags describe and organize these resources. Analysis of the IND network maps and the metadata tags are used to understand how information flows through the network to maximize effective usage of these resources during all phases of emergency management.

The network maps described in this report visually represent the flow of information between resources used by the federal interagency to support decision making during emergency management. Three attributes of the network — betweenness centrality and resource connectivity — were explored using network maps. From these analyses, it is apparent that the wide array of IND relevant resources used by the federal interagency are mostly well connected. There are, however, several resources, termed orphans, which have no upstream or downstream connections.

Although data and modeling resources are widely used by the federal interagency, only a few resources stand out as being heavily used. In most cases, as determined using betweenness centrality measures, these heavily used resources are also the most central resources, especially when they are used as consequence models, like the NARAC Modeling System or HPAC. In some cases, resources that would be expected to have high centrality measures based on the number of users and resource type, have low centrality values, like the consequence model, EAGLE-I. These resources that do not function as expected, like the orphan resources or EAGLE-I, highlight opportunities for better integration within the network, so they can be used more effectively during all phases of emergency management.

A qualitative analysis of the flow of information within the IND network shows a relatively clear progression from raw data to event characterization to consequence models. However, resources tagged later in the flow of information, such as decision support tools or mission specific requirements, are not clearly organized within the network and are not well connected to upstream resources. This breakdown in the flow of information framework suggests that the resources that directly support operational decision making are not as connected as they should be; forging better connections will ensure that these resources are using the same event-specific information to inform decision making.

An overview of the network map shows that there are far more lightly colored nodes that correspond to resources in categories early in the flow of information. This observation is supported by a quantitative analysis of the number of resources per resource type, and shows there are far more resources that are tagged as raw data, event characterization models, and situational awareness data. This finding is consistent with the notion that in order to provide a comprehensive picture of what happened and who was affected, a wide array of raw data are required as inputs into event characterization and consequence models. The analysis also reveals that the few impact estimates, decision support tools, and mission-specific requirements available to the federal interagency only support a narrow range of mission spaces. To ensure that the available resources are sufficient to inform emergency management decisions, critical information requirements from all ESFs will need to be identified, upon which the development of appropriate tools to support operational decision making will be based.

In addition to analyzing the network maps, and the number of resources by resource type, it is also important to determine which federal agencies are the users and producers of information. Collating and analyzing this information reveals which agencies are at the forefront of using data and modeling resources and which ones are the producers of these resources. Not surprisingly, FEMA, followed by

DHS, DoD, and DOE, leads the federal interagency in using the available data and modeling resources to support operational decision during all phases of emergency management. As the lead agency for coordinating efforts between agencies for emergency management, FEMA should be and is a major user of interagency resources.

Next Steps

During Phase IV of this project, the interactive inventory of IND data and modeling resources, accessible via a web-based graphical user-interface, will be built and delivered. The user interface will allow end users to easily interact with the inventory to learn what resources are available and used by the federal interagency. In addition, an in-depth analysis of the resource network will be conducted to identify gaps and recommend courses of action. The resulting inventory and analyses will help ensure that decision makers have access to the information they need, when they need it, to support operational decision making during all phases of emergency management.

Appendix 1: Interviewees

NAME	AGENCY
Buikema, Ed	Argonne National Laboratory
Folga, Steve	Argonne National Laboratory
Gunn, Julia	Boston Public Health Commission
Demarais, John	CAP
St. John, Courtney	Columbia University, Center for Research on Environmental Decisions
Alexander, David	DHS
Billado, William	DHS
Briggs, Kevin	DHS
Chacko, Betsie	DHS
Cole, Ray	DHS
Coller Monarez, Susan	DHS
Cotter, Dan	DHS
Danielson, Glen	DHS
Franco, Crystal	DHS
Klucking, Sara	DHS
Langhelm, Ron	DHS
MacIntyre, Anthony	DHS
Mapar, Jalal	DHS
Maycock, Brett	DHS
Moe, Mathew	DHS
Shepherd, Dave	DHS
Valentine Davis, Victor	DHS
DeCroix, Michele	DHS
Berscheid, Alan	DHS NISAC/HITRAC
Chatfield, Catherine	DHS NISAC/HITRAC
Norman, Mike	DHS NISAC/HITRAC
Stamber, Kevin	DHS NISAC/HITRAC
Aeschelman, Jeremiah	DoD DTRA
Basiaga, Dariusz	DoD DTRA
Blandford, Michael	DoD DTRA
Blandford, Mike	DoD DTRA

Cooper, Charles	DoD DTRA
Grouse, Andy	DoD DTRA
Kahn, Todd	DoD DTRA
Leong, Timothy	DoD DTRA
Lowenstein, Eric	DoD DTRA
Mazzola, Tom	DoD DTRA
Meris, Ron	DoD DTRA
Phillips, Michael	DoD DTRA
Baron, Thomas	DoD NORTHCOM/NORAD
Danaher, Leo	DoD NORTHCOM/NORAD
DeGoes, John	DoD NORTHCOM/NORAD
Friedman, Andy	DoD NORTHCOM/NORAD
Jackson, Mike	DoD NORTHCOM/NORAD
Wireman, Jody	DoD NORTHCOM/NORAD
Allen, Gary	DoD Office of the Secretary of Defense
Gerrig, Dan	DoD Office of the Secretary of Defense
Greenberg, Brandy	DoD Office of the Secretary of Defense
Miller, Brian	DoD Office of the Secretary of Defense
Mullen, Frank	DoD Office of the Secretary of Defense
Sorden, Caryn	DoD Office of the Secretary of Defense
Yu, Leigh	DoD Office of the Secretary of Defense
Blumenthal, Daniel	DoE
Cedres, Stewart	DoE
Clark, Jamie	DoE
Corredor, Carlos	DoE
Favret, Derek	DoE
Fernandez, Steve	DoE
Hsu, Simon	DoE
Lippert, Alice	DoE
Lucas, Anthony	DoE
Rollison, Eric	DoE
Scott, Margaret	DoE
Willging, Pat	DoE

Schilling, David	DoT
Stuckey, Bill	DoT
Vanness, Jeffrey	DoT
Howard, Jeffrey	Dun & Bradstreet
Clark, Steve	EPA
Haxton, Terra	EPA
Hudson, Scott	EPA
Irizarry, Gilberto	EPA
Lemieux, Paul	EPA
Magnuson, Matthew	EPA
Mosser, Jen	EPA
Snead, Kathryn	EPA
Woodyard, Josh	EPA
Almonor, Niclaos	FEMA
Anderson, Lindsey	FEMA
Bahamonde, Marty	FEMA
Bausch, Doug	FEMA
Bellamo, Doug	FEMA
Bennett, Gerilee	FEMA
Berman, Eric	FEMA
Bonifas, Michelle	FEMA
Boyce, Carla	FEMA
Brierly, Mick	FEMA
Brown, Cliff	FEMA
Crawford, Sean	FEMA
Daigler, Donald	FEMA
Decker, K.C.	FEMA
Demorat, David	FEMA
Ewing, Melvin	FEMA
Faison, Kendrick	FEMA
Farmer, Bob	FEMA
Gilmore, Lance	FEMA
Gorman, Chad	FEMA

Griffith, David	FEMA
Harned, Rebecca	FEMA
Hewgley, Carter	FEMA
Hinkson, Tasha	FEMA
Hodge, Craig	FEMA
Huyck, Charles	FEMA
Ingram, Deborah	FEMA
Jackson, Liz	FEMA
Jacques, Richard	FEMA
Juskie, John	FEMA
Kazil, Jacqueline	FEMA
Lawson, David	FEMA
Legary, Justin	FEMA
Longenecker, Gene	FEMA
Lumpkins, Donald	FEMA
McDonald, Blair	FEMA
Pollock, Marcus	FEMA
Preusse, Paul	FEMA
Rabin, John	FEMA
Ransom, Darrell	FEMA
Roberts, Nikki	FEMA
Rogers, James	FEMA
Rozelle, Jessee	FEMA
Sanderson, Bill	FEMA
Schlossman, Mikhail	FEMA
Scott, Kara	FEMA
Sonhaus, Daniel	FEMA
Stanfill, Derek	FEMA
Stuart, James	FEMA
Truax, Wayne	FEMA
Vaughan, Chris	FEMA
Wolfgul, Gus	FEMA
Woodhams, Katrina	FEMA

Wright, Roy E.	FEMA
Wycoff, Kristen	FEMA
Zohn, Ashley	FEMA
Zuzak, Casey	FEMA
Butgereit, Richard	Florida Division on Emergency Management
Baker, Jay	Florida State University
Gabriel, Edward	HHS ASPR
Koerner, John	HHS ASPR
Lant, Tim	HHS ASPR
Lurie, Dr. Nicole	HHS ASPR
Olsen, Jennifer	HHS ASPR
Shankman, Robert	HHS ASPR
Wright, Sue	HHS ASPR
George, David	JHU APL
Taylor, Steven	JHU APL
Waddell, Richard	JHU APL
Alai, Maureen	Lawrence Livermore National Laboratory
Buddemeier, Brooke	Lawrence Livermore National Laboratory
Goforth, John	Lawrence Livermore National Laboratory
Glascoe, Lee	Lawrence Livermore National Laboratory/NARAC
Homann, Steve	Lawrence Livermore National Laboratory/NARAC
Nasstrom, John	Lawrence Livermore National Laboratory/NARAC
Pobanz, Brenda	Lawrence Livermore National Laboratory/NARAC
Simpson, Matthew	Lawrence Livermore National Laboratory/NARAC
Sugiyama, Gayle	Lawrence Livermore National Laboratory/NARAC
Tuttle, Benjamin	NGA
White, Greg	NGA
DiMego, Geoff	NOAA
Draxler, Roland	NOAA
Feyen, Jesse	NOAA
Heffernan, Robyn	NOAA
Knabb, Richard	NOAA
Lapenta, Bill	NOAA

McQueen, Jeff	NOAA
Mitchell, Daisy	NOAA
Mongeon, Albert	NOAA
Roohr, Peter	NOAA
Sokich, John	NOAA
Tallapragada, Vijay	NOAA
Tolman, Hendrik	NOAA
Collins, Andy	Old Dominion University
Jordan, Craig	Old Dominion University
Myer, David	Old Dominion University
Robinson, Mike	Old Dominion University
Tune, Greg	Red Cross
Bynum, Leo	Sandia National Laboratories
John, Charles	Sandia National Laboratories
Jones, Dean	Sandia National Laboratories
Kimura, Margot	Sandia National Laboratories
Knowlton, Robert	Sandia National Laboratories
Kraus, Terry	Sandia National Laboratories
Mahrous, Karim	Sandia National Laboratories
Miller, Trisha	Sandia National Laboratories
Pennington, Heather	Sandia National Laboratories
Pless, Daniel	Sandia National Laboratories
Teclemariam, Nerayo	Sandia National Laboratories
Vurin, Eric	Sandia National Laboratories
Dial, Patrick	SBA
Valliere, John	SBA
O'Neill, Ed	State
Dowell, Earlene	US Census
Pitts, Robert	US Census
Diaz, Steve	USACE
Harris, Dewey	USACE
Hendricks, Joel	USACE
Irwin, Bill	USACE

Keown, Patrick	USACE
Markin, Chad	USACE
Nye, Bill	USACE
Schargorodski, Spencer	USACE
Schuster, Michael	USACE
Town, Patrick	USACE
Gleason, Joe	USCG
Gunning, Jason	USCG
Hunt, Michael	USCG
Landry, Mary	USCG
Lundgren, Scott	USCG
McGlynn, Matt	USCG
Moore, Brian	USCG
Carpenter, Ryan	USDA
Li, Yun	USDA
Collins, Brian	USFS
Erickson, Rod	USFS
Hill, Laura	USFS
Triplett, Sean	USFS
Applegate, David	USGS
Blanpied, Michael	USGS
Gallagher, Kevin	USGS
Haines, John	USGS
Hammond, Steve	USGS
Ludwig, Kris	USGS
Lyttle, Peter	USGS
Mandeville, Charles	USGS
Mason, Robert	USGS
Perry, Sue	USGS
Driggers, Richard	White House NSS

Appendix 2: Metadata Tags

Model/Data

All resources are categorized as either models or data. Models are defined as programs, algorithms, or sets of calculations which may be used for emergency management. Many models accept as input a type of data which they transform into another type to provide new information (e.g., NARAC Modeling System, HPAC). Other models collate individual data resources to yield a new dataset with enhanced utility (e.g., Moreover NewsDesk, which identifies and collates news items of interest from media sources). Data are defined as repositories of information that may be used for emergency management. This definition of data encompasses tools which assist in the presentation or visualization of data without transforming the data itself (e.g., FEMA GeoPlatform). Resources that have both modeling capabilities and a repository of their output, or some other data feed, are tagged as both a model and data.

Hazard

Resources are tagged based on the hazards during which they can be used to inform operational decision making. One or more hazards can be tagged for each resource. Resources can be tagged as: hurricane, earthquake, tsunami, inland flood, tornado, chemical release, contagious outbreak, non-contagious outbreak, nuclear detonation, explosion, and radiological release. Resources may be tagged with a single hazard (e.g., NUEVAC is tagged only with nuclear detonation) or multiple hazards (e.g., HotSpot is tagged with nuclear detonation, radiological release, explosion, and fire). Additionally, resources that support emergency planning and response for any hazard type are tagged as All-Hazards.

Cascading effects were not considered when tagging hazards. Users interested in the cascading effects of a given hazard (e.g., a radiological release from a power plant damaged during an explosion) would instead search the inventory for the secondary hazard directly.

Core Capabilities, ESFs, and RSFs

The Core Capabilities are designations that represent a list of critical elements within the five mission areas (Prevention, Protection, Mitigation, Response, and Recovery) necessary for Emergency Management.[15] The Core Capabilities are used to assess both the capabilities and gaps over the entire federal interagency emergency management community. In order to facilitate this effort, resources are tagged based on which Core Capabilities they support. Each resource may be tagged as supporting one, more than one, or no Core Capabilities. However, it is not always clear which Core Capabilities a given resource may support. This is especially true of raw data, event characterization models, and situational awareness data. For example, it is not immediately obvious which Core Capabilities rely on Digital Terrain Elevation Data (DTED) because this data must first be processed before it is useful for emergency management. To address this problem, each resource was tagged with any Core Capabilities it directly supports, in addition to any downstream resources supported. For instance, DTED feeds into HPAC. As HPAC can be used for Situational Assessment, Environmental Response/Health and Safety, Public Health and Medical Services, and Planning, DTED is also tagged with those Core Capabilities.

The Emergency Support Functions (ESFs) and Recovery Support Functions (RSFs) provide a coordinating structure for the key functional areas that are most frequently needed in response and

15 (2011a) National Preparedness Goal. Department of Homeland Security

recovery, respectively.[16, 17] Identifying the resources that directly support each ESF and RSF will allow emergency managers to ascertain which resources can be used to support their specific missions. Like the Core Capabilities, each resource may be tagged as supporting one, more than one, or no ESFs and RSFs. Resources were only tagged with RSFs if they were also tagged with the Recovery phase (see the 'Phase Specific Utility' subsection). Unlike the Core Capabilities, the ESFs and RSFs are directly used in coordination of federal disaster response and recovery. Therefore, it is only necessary to know which resources directly support each ESF and RSF, and these tags are not inherited from downstream resources as Core Capabilities tags are.

As described in their Framework documents, each ESF and RSF has one Coordinating Agency and one or more Primary Agencies chosen on the basis of authorities and resources. These agency assignments were used in ESF and RSF tagging to help users identify inventory resources useful for their missions. First, based on information from interviews and research, resources were tagged depending on whether those resources were expected to support ESF or RSF missions. In addition, resources were automatically tagged with the ESFs and RSFs for which their federal users were Coordinating and/or Primary Agencies. This approach ensured that the ESF and RSF tags were informed by both interview data and existing policies for emergency management.

Keywords and Resource Functions

In addition to the Core Capabilities, ESFs, and RSFs, resources are further characterized based on their function. Keywords are simple titles designed to describe the resources independently of the flow of information. Each resource may be tagged with one or more keywords. For example, observational weather data, a raw data resource, the Global Forecast System (GFS), an event characterization model, and the Local National Weather Service (NWS) Forecasts, a situational awareness data resource, are each tagged with the keyword 'Atmospheric,' even though they have different flow of information categories. Additionally, resources may be tagged with more than one keyword. For instance, the National Bridge Inventory (NBI) is tagged with both 'Infrastructure' and 'Transportation.'

In order to provide an even higher level of resolution for the functions of resources included in the inventory, the keywords are further split into categories based on the flow of information. Each resource may be tagged with one or more resource functions. These tags provide a succinct description of the utility of a resource, both with regards to situations for which the resource is relevant and how it is incorporated into the flow of information. In the previous example, observational weather data is tagged as 'Atmospheric Raw Data,' the GFS is tagged as 'Atmospheric Event Characterization,' and the Local NWS Forecasts are tagged as 'Atmospheric Situational Awareness Data.'

Resource Type

Resource types are directly drawn from the flow of information categories. As outlined in the phase II report, data are categorized as raw data, situational awareness data, impact estimates, and/or mission specific requirements, while models are categorized as event characterization models/analysis, consequence models, and/or decision support tools. Each resource may be tagged as one or more resource types. Modeling resources that are useful as multiple resource types can also have multiple tags. Multi-tagged modeling resources represent models that perform multiple, successive steps of data processing.

[16] (2008) National Response Framework. Federal Emergency Management Agency
[17] (2011b) National Disaster Recovery Framework. Federal Emergency Management Agency

Similarly, data resources that are useful as multiple resource types can have multiple tags. For example, MedMap is an interactive mapping system that contains a diversity of datasets and so is tagged as situational awareness data, impact estimates, and mission-specific requirements.

Data Collection Method

There are three primary methods of data collection: instrumentation, reporting, and the use of social media and crowd-sourced data. Data that are collected, aggregated, and processed directly (i.e., not generated as the output of models) fall into one or more of these three categories. It is important to specify the methods used to collect the data within a resource because collection methods can influence the availability, accessibility, and error associated with the resource.

Instrumentation data are obtained through the use of instruments that are capable of recording repeated observations. Often, data collected by instrumentation is raw data and requires processing by event characterization models or analysis tools before it can be used in support of decision making. Successful collection and aggregation of instrumentation data requires investment in a data collection infrastructure, which must be developed and deployed before an event occurs in order to collect and transmit the data in real time.

Data collected through human observation or non-automated data entry are considered reporting data. These data include damage assessments, hospital records, and critical infrastructure locations. While many types of instrumentation data can be continually collected without the need for large numbers of personnel during an event, reporting data generally take longer to collect and aggregate, and they demand larger personnel investments. Thus, reporting data are typically available at a lower resolution and after a longer delay than instrumentation data.

Crowd-sourced data are also used to inform and validate operational models and decision support tools, though much less frequently than the other two types. There is considerable interest across the interagency to develop methods to use social media data to support decision-making in a way that accounts for the data's inherent uncertainty. Particularly in instances where traditional data feeds are unable to address a question, social media has the potential to serve as a valuable resource. For example, the Department of Energy's EAGLE-I tool is fed in part by data obtained through the Twitter accounts and webpages of private electric power companies.

Owner

The agency, division, or group responsible for updating, maintaining, and validating a given resource is identified. As specific contact information and organizational structures may change over time, specifying the entity in control of a given resource will ensure that it continues to be accessible, regardless of personnel changes or reorganization within agencies. If a resource has more than one organization that is in control of the resource, both organizations are listed as an owner.

User

Resources are tagged with known members of their user communities. Here, users are defined as federal level organizations who directly apply information from the resource in order to answer a policy- or operations-related question in support of their missions for emergency management. Therefore, for the purposes of this project, state and local governments as well as private sector or academic organizations were not considered users (with the one exception of the Red Cross).

It is necessary to note that, while it is informative to tag resources with their known users, this is not the only way to judge the utility or reliability of a resource. New or recently updated resources may be underrepresented due to a lack of familiarity within the emergency management community. Similarly, it is also useful to consider the quality control methods used to verify and validate a given data resource. In any case, identifying the existing user communities who regularly use specific information resources in support of decision making allows both users and producers of these resources to work together in a process of ongoing development, evaluation, and maintenance.

Upstream and Downstream Resources

Based on the understanding that data collection, analysis, and modeling is an iterative process, the data and models that lie upstream of a given resource (i.e., those that serve as inputs for that resource) are defined. Complementary to the upstream resources category, downstream resources list the data and models that are fed by a given resource. This information indicates the datasets and models that use the resource as an input. It is important to identify the data and modeling resources that are interdependent, as validity of any model relies heavily on the accuracy of its inputs.

Phase Specific Utility

To assist users in determining which inventory resources are most relevant to their missions, the resources are tagged with the phases of emergency management for which they are useful. The phase tags are planning, pre-event preparedness (only for advance-notice events), immediate response (within approximately 36 hours following the event), deployment, sustained response, and recovery. Resources are phase-tagged based on their potential uses, not only their known ones. Thus, a resource which has been used for planning but which could likely be used in the immediate response phase would carry both tags. A resource may be tagged by one or more of the listed phases.

Summary

A brief summary of each resource is provided to capture key usage and feature information.

Access Information

The procedures or credentials necessary to view, use, or update a resource are also defined. Resources can either be open access (immediately available to anyone or only requiring a free, automatic registration) or limited access (which can include proprietary data, classified data, or data that requires permission to access). Each resource may only be tagged as limited access or open access. These two tags are mutually exclusive. If possible, specific instructions on how to access the resource are included.

Access Type

There are three primary ways a model can be run: Standalone, through a Reachback Capability, or through interaction with a Subject Matter Expert. Every model is tagged as one or more of these three access types. If a resource can be run through multiple sources, then it is tagged appropriately. For example, the ORIGEN model is run as a Standalone model when used independently, but it is run as a Reachback Capability when used when requested through NARAC. Therefore, ORIGEN is tagged as Standalone as well as Reachback Capability.

A model tagged as Standalone describes any resource that can be run by any individual with access on a personal computer. A model tagged as a Reachback Capability is accessed through a reachback facility. This tag refers to resources run and managed by specific organizations and accessed through formal Requests for Information. A model tagged as Subject Matter Expert is defined as any model that can only be accessed through personal interactions with the model developer or owner. Often, the outputs from these models can be accessed by the public online but the model itself is restricted for use by the subject matter expert. This also includes models run on a schedule, based on computing limitations, that precludes additional runs of the model outside the set schedule.

Processing Requirements

The processing requirements for viewing a data resource or running a model are given in relatively broad terms. Rather than detailing the exact system specifications needed to use each resource, their processing requirements were generalized into three categories: supercomputer, desktop/laptop, and mobile device. Resources are only tagged with 'mobile device' if they have a dedicated mobile application. Likewise, an Internet-based resource that could be accessed with a mobile browser is not tagged with 'mobile device' unless its website is optimized for mobile viewing. In certain cases, a resource may be tagged with two of the three processing requirements. For instance, a weather model that can be run on a desktop computer but is often run on a supercomputer would be tagged as both 'desktop/laptop' and 'supercomputer.' Similarly, a resource run on a desktop application with the same capabilities would be tagged with both 'mobile device' and 'desktop/laptop.'

Refresh Rate and Last Known Version

During all phases of emergency management, frequently updated resources are necessary to inform all levels of decision making. If the information is available, resources are tagged based on their refresh rate (how often they are updated). For data resources, this category specifies how often new information is uploaded into the dataset. For models, it indicates whether the model is routinely run, and if so, how frequently. For example, EAGLE-I updates its information on electric power grid status every 15 minutes. On the other hand, the GFS weather model is re-run approximately every 6 hours.

Not all data used to support decision making during emergency management can or should incorporate real-time data. While observational weather data must be updated every few minutes to reflect current conditions, data regarding the locations of critical infrastructure or residential building codes do not require the same update frequency to be operationally relevant. For datasets that do not consist of real-time data, the last known version of the dataset (often a release date) is indicated.

Similarly, not all models can or should be automatically run. While automatically refreshing weather forecasts are required for up-to-date situational awareness, many of NOAA's weather forecasting systems are run on a predetermined schedule because of the processing limitations of their supercomputers. This means that many of these models can only be run on their predetermined schedule and cannot be run more frequently during activation. As with datasets, the last known version of the model is indicated to ensure users are aware of the most recent release.

Programming Language

When possible, the programming language in which a resource is coded is given. This metadata category is not only important for developers interested in updating, modifying, or adapting a resource, but it may also provide essential compatibility information, indicating whether or not a resource can operate on a certain computer platform.

File Type

If relevant, the file type for a data resource or the file type for the output of a model is given. This information can be used by a model developer or analyst when determining data compatibility or other technical issues. It can also be used to indicate software requirements. If resources are capable of outputting multiple file types, then every file type it is capable of creating will be listed.

Contact Information and Contact during Activation

The contact information for the group or individual responsible for updating, maintaining, or granting access to each resource is provided. When possible, coordinates for specific individuals are listed. Contact information always contains the organization or agency and, if applicable, the division of the contact in case of personnel changes. Where applicable, an additional contact is listed for use during activation.

Appendix 3: Data and Models Resource Catalog

The IND resource inventory is electronically attached as an Excel file.