

D2dprov: Statement of Needs 2022

Technology and policy requirements to fulfill Vision 2025's proposed approach to data-driven, science-informed climate resilience decisions

Author:	Brian Wee (ORCID: 0000-0002-0038-9381)
Organizational affiliations:	Founder and Managing Director, Massive Connections, LLC Senior Advisor for Data, Science, and Policy, Neptune and Company, Inc.
Date released:	2019-02-07
Version:	1.1
License:	CC-BY 4.0
Citation:	Wee, B. (2019). D2dprov: Statement of Needs 2022. Technology and policy requirements to fulfill Vision 2025's proposed approach to data-driven, science-informed climate resilience decisions. Washington, DC. https://doi.org/10.6084/m9.figshare.7692038
Team members:	William Teng (ORCID: 0000-0001-5642-6359) Nancy Hoebelheinrich (ORCID: 0000-0002-6797-7903) Tom Narock (ORCID: 000-0002-9785-4496) Arif Albayrak
ESIP Lab project title:	Enabling the encoding and visualization of provenance metadata for better discovery and understanding of climate resilience strategies for agriculture-related decision-making.
Acknowledgment:	This work is based on team collaboration tools provided by the ESIP Lab with support from the National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA) and the United States Geologic Survey (USGS).
Citation BibTeX:	@techreport{Wee2019,address = {Washington, DC},author = {Wee, Brian},doi = {10.6084/m9.figshare.7692038},institution = {Massive Connections},title = {{D2dprov: Statement of Needs 2022. Technology and policy requirements to fulfill Vision 2025's proposed approach to data-driven, science-informed climate resilience decisions.}},year = {2019}}

Table of Contents

1.0	Scope of the d2dprov project	1
2.0	Definitions	3
2.1	Provenance	3
2.2	Resilience through mitigation	3
3.0	Decisions	4
3.1	Anatomy of a decision	4
3.2	Multi-criteria Decision Analysis (MCDA)	5
3.3	Examples of agriculture-related decisions	8
3.4	Examples of flooding-related decisions	8
3.5	Examples of “decisions” in a Record of Decision	10
3.6	Examples of “decisions” that we shall not focus on using the flood-mitigation exemplar 11	
3.7	Are activities after a “decision” within the scope of d2dprov?	11
3.0	Policy Needs	12
3.1	Harmonized set of resilience planning protocols	12
3.2	Standardized vocabulary for publicly accessible documents	12
4.0	Technical Needs	13
4.1	Schemas for representing decisions	13
4.2	Knowledge graphs	14
4.3	Encoding of harmonized resilience planning steps into a reference protocol	17
5.0	Bibliography	18

Table of Figures

Figure 1: Scope of the d2dprov project.	1
Figure 2: Schematic of the d2d provenance space	3
Figure 3: US Climate Resilience Toolkit's "Steps to Resilience"	4
Figure 4: Decision components (Linkov & Bates, 2016).....	5
Figure 5: Example of Multi-criteria Decision Analysis (Linkov & Bates, 2016).....	6
Figure 6: Hypothetical community resilience objectives for a coastal town that ultimately inform decisions	7
Figure 7: Baltimore's Disaster Preparedness Project and Plan (IN1)	9
Figure 8: Baltimore's Disaster Preparedness Project and Plan (IN9)	9
Figure 9: Record of Decision for New Jersey Flood Mitigation Plan	10
Figure 10: Decision Ontology (Car, 2017)	13
Figure 11: Decision Making Ontology (Kornysheva & Deneckère, 2012)	14
Figure 12: Extraction of concepts from human-readable documents (Rotmensch et al., 2017)	16
Figure 13: Creating a knowledge graph after concept extraction (Rotmensch et al., 2017).....	16
Figure 14: Resilience Genome.....	17

1.0 Scope of the d2dprov project

Imagine that a coastal county finds itself at the start of a planning cycle to make the county “climate-ready”. These planners would greatly benefit from a fully traceable solution already implemented by another locality faced with similar socio-environmental challenges that can be used to bootstrap the local planning process. Just as one could download open-source scientific code from GitHub, imagine the benefits of a fully traceable resilience workflow (i.e. “recipes” for resilience planning) that is captured as a knowledge graph and shared as an open resource. That knowledge graph captures digital artifacts, or pointers to digital artifacts, including data, models, journal articles, integrated scientific assessments, applicable building codes, economic projections, risk assessment frameworks, formal decision-analysis, stakeholder priorities, and other relevant artifacts. Because each artifact is assigned a unique identifier and semantically linked to other relevant bodies of knowledge, the planning team is empowered to follow lines of inquiries that may otherwise have been difficult if not for the possibility of knowledge graph traversals.

The d2dprov (short for “data to decisions provenance”) project is aimed at assessing the technologies required to implement a “GitHub” for resilience planning as outlined above. Figure 1 provides a high-level overview of the project.

This document is an accompaniment to “D2dprov: Vision 2025. A transdisciplinary science, technology, and policy synthesis on data-driven, science-informed resilience planning for 2025 and beyond” (Wee, 2019b). This document outlines selected technological and policy components that will contribute to the fulfillment of the goals outlined in “Vision 2025”.

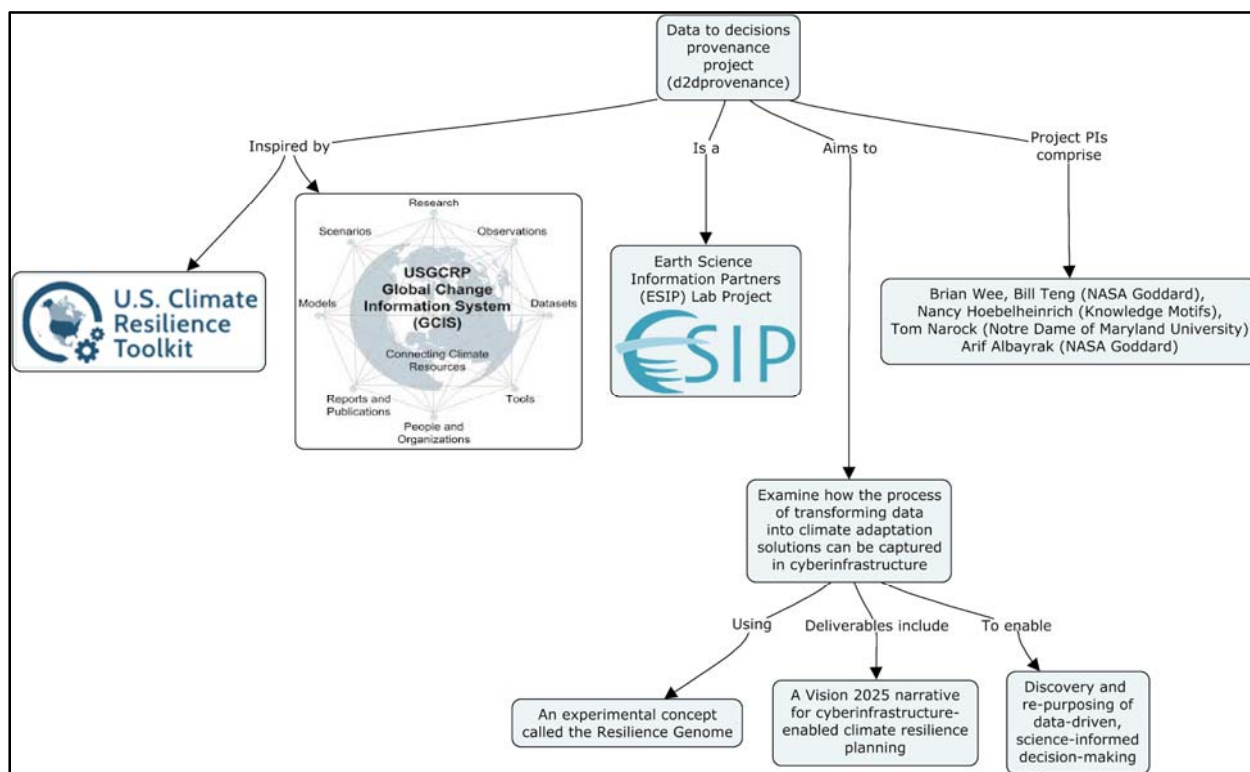


Figure 1: Scope of the d2dprov project.

The following table lists resources developed by the d2dprov team that are freely accessible. The “Reference” column reflects entries in the bibliography that resolve to web resources. URLs are not listed in this column to facilitate maintenance and evolution of this document.

Type	Title	Description	Reference
Report	D2dprov: Vision 2025 A transdisciplinary science, technology, and policy synthesis on data-driven, science-informed resilience planning for 2025 and beyond	Outlines the vision for data-driven, science-informed, traceable resilience planning with an outlook to the year 2025 and beyond.	Wee, 2019b
Report (this report)	D2dprov: Statement of Needs 2022. Technology and policy requirements to fulfill Vision 2025’s proposed approach to data-driven, science-informed climate resilience decisions.	Outlines selected technological and policy components that will contribute to the fulfillment of the goals outlined in “Vision 2025”.	Wee, 2019a
Python code, documentation, and presentation slides	Prototype parser for concept extraction using community ontologies and natural language processing	Parses Environmental Impact Statements (EIS) and Records of Decisions (ROD) using Natural Language Processing (NLP) to identify terms from environmental and resilience ontologies.	Narock, Wee, Hoebelheinrich, Albayrak, & Teng, 2019
Documentation and presentation slides	Experimental modeling of workshop planning processes to W3C PROV	Describes the outcome of an experiment to ascertain how well W3C PROV could be used to capture the provenance of a planning process that used a National Park Service developed scenario planning process.	Narock, Wee, Hoebelheinrich, Albayrak, & Teng, 2019
Presentation slides	Strategy for advanced query and discovery of context-relevant resilience documents	Outlines a high-level approach for using machine learning to acquire the content structure of existing documents in order to return a set of documents that match the user’s query.	Narock, Wee, Hoebelheinrich, Albayrak, & Teng, 2019

2.0 Definitions

2.1 Provenance

One of the objectives of d2dprov is to enable the “discovery and re-purposing of data-driven, science-informed decisionmaking” (Figure 1) through capturing the provenance between data to decisions. The informatics community is well equipped to model and capture the provenance for data-products at the “data” end of the “data to decisions” pipeline. The provenance between data-products and decisions is not as well studied. This is depicted in Figure 2.

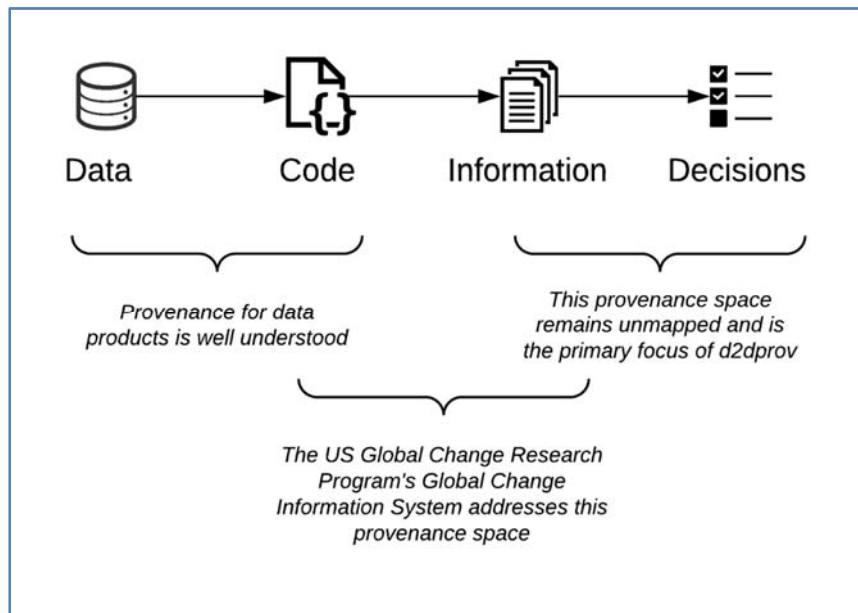


Figure 2: Schematic of the d2d provenance space

2.2 Resilience through mitigation

Section 3 of the Vision 2025 report (Wee, 2019b) outlined a Presidential Policy Directive for national preparedness requiring the creation of a series of integrated national planning frameworks covering prevention, protection, mitigation, response, and recovery. Following the strategy used for the Vision 2025 report, this document maintains the focus on mitigation.

3.0 Decisions

3.1 Anatomy of a decision

To better understand the nature of the provenance associated with climate resilience decisions, a definition of the end-point of that pipeline – the “decision” – is warranted. Having a clearer idea of what constitutes a “decision” helps determine the scope of the provenance problem. It also provides us a place to identify documents that may be described as “decisions” so as to devise ways to encode the information in “decisions” in machine-readable formats to enable the “discovery and re-purposing of data-driven, science-informed decisionmaking”.

Definitions of the term “decision” Merriam-Webster include “a determination arrived at after consideration” and a “report of a conclusion”.

The second definition aptly describes a type of document called a “Record of Decision (ROD)”. A ROD, described further in the Vision 2025 report (Wee, 2019b), is a document required under a number of federal regulations to fulfill specific requirements stipulated under the National Environmental Policy Act. **RODs may be used as the starting point by a concept extraction algorithm** to start assembling a “Resilience Genome”, further described in the Vision 2025 report. The “Resilience Genome” can be implemented as a knowledge graph (see Section 4.2).

The first Merriam-Webster definition above is too generic for the purposes of d2dprov. A more constrained definition of “decisions” that helps us determine what type of human-readable documents can be parsed by an algorithm to extract concepts related to “decisions” is warranted. For the purposes of d2dprov, **climate resilience related decisions include “decision points” as depicted in the US Climate Resilience Toolkit’s (USCRT) “Steps to Resilience”** (Figure 3).

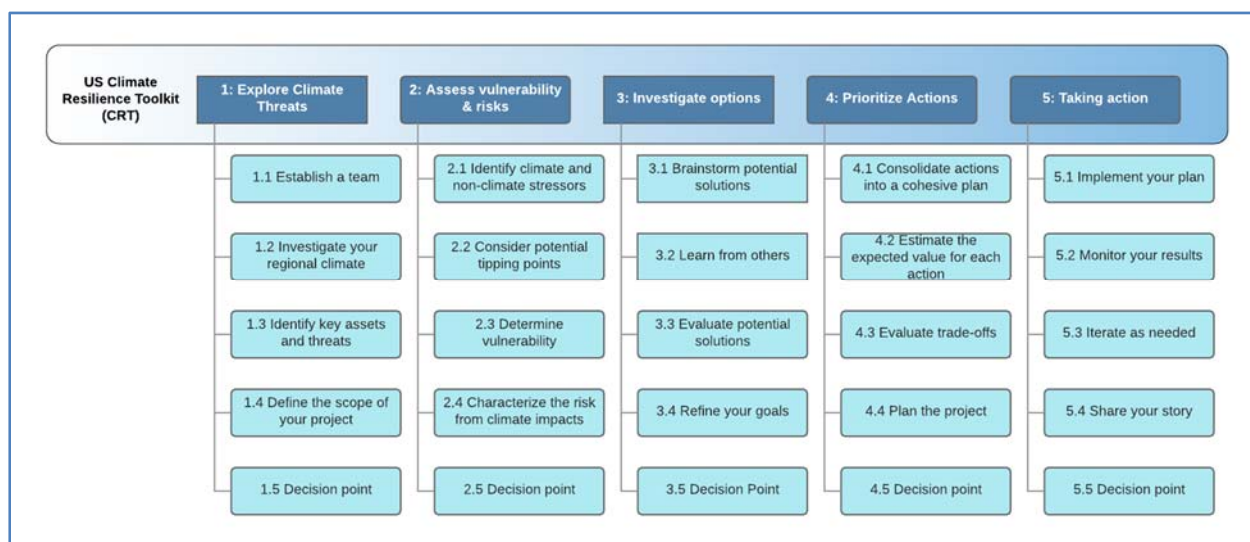


Figure 3: US Climate Resilience Toolkit's "Steps to Resilience"

3.2 Multi-criteria Decision Analysis (MCDA)

The types of decisions involved in resilience planning, such as the “decision points” in Figure 3 are likely to be complex decisions based on trade-offs between competing factors. Such decisions will also likely involve heterogeneous socio-environmental data. Multi-criteria Decision Analysis (MCDA) is one of many decision science techniques that can be used in resilience planning.

Figure 4 depicts a schematic that reflects the anatomy of a complex decision. The final step in the process of decision making is represented as “Rank/Select final alternative(s)”. The “Tools” in the figure reflect the types of data and models that ESIP constituents care about.

Kornysheva & Deneckère, 2012 (see Section 4.1) propose an ontology for decisionmaking that reflects many of the elements in Figure 4.

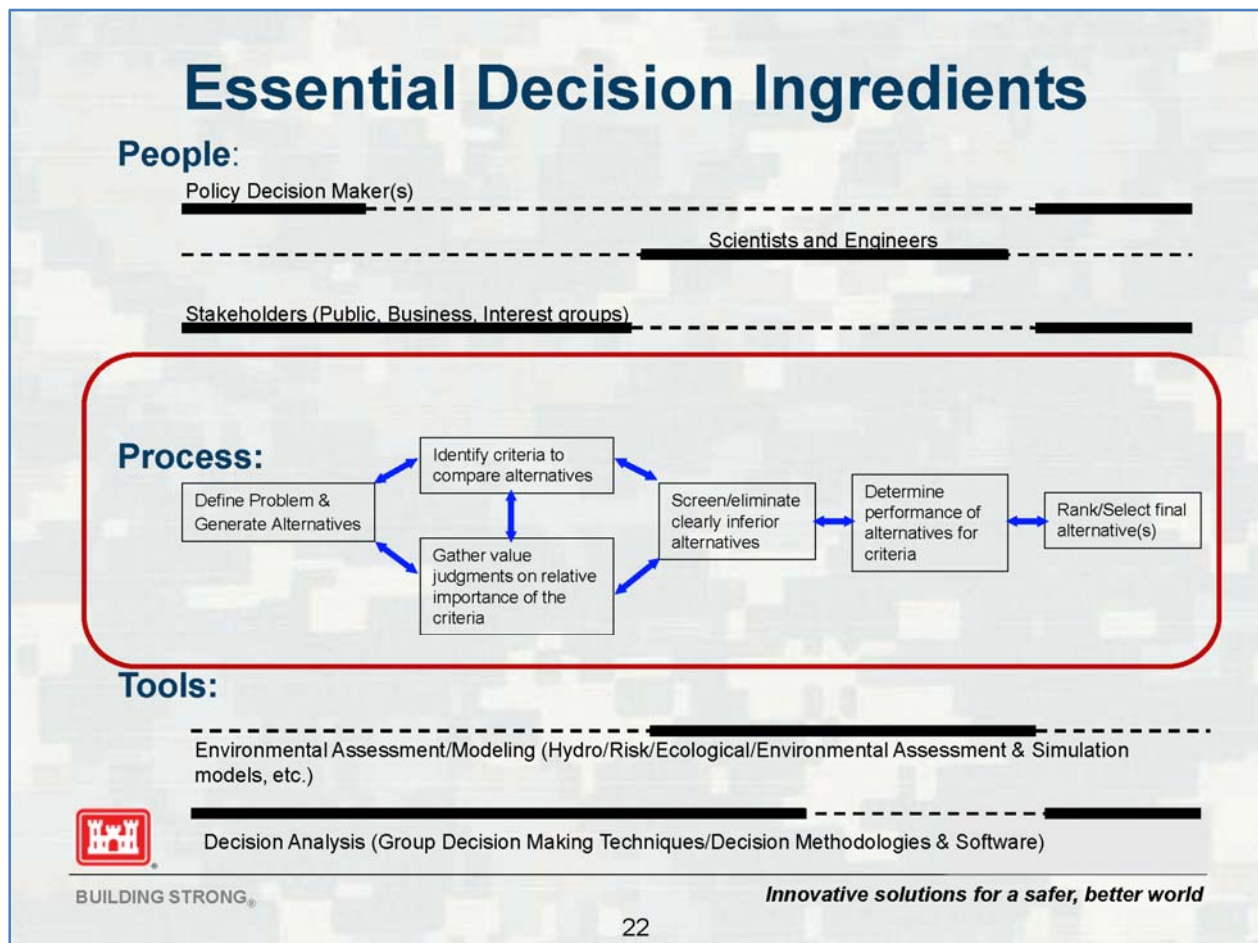


Figure 4: Decision components (Linkov & Bates, 2016)

Figure 5 illustrates how one might frame MCDA as it is applied to the process of buying a vehicle. In the case of resilience planning, the objectives may span economic, social, and environmental dimensions. Kornysheva & Deneckère, 2012 (see Section 4.1) propose an ontology for decisionmaking that reflects many of the elements in Figure 5.

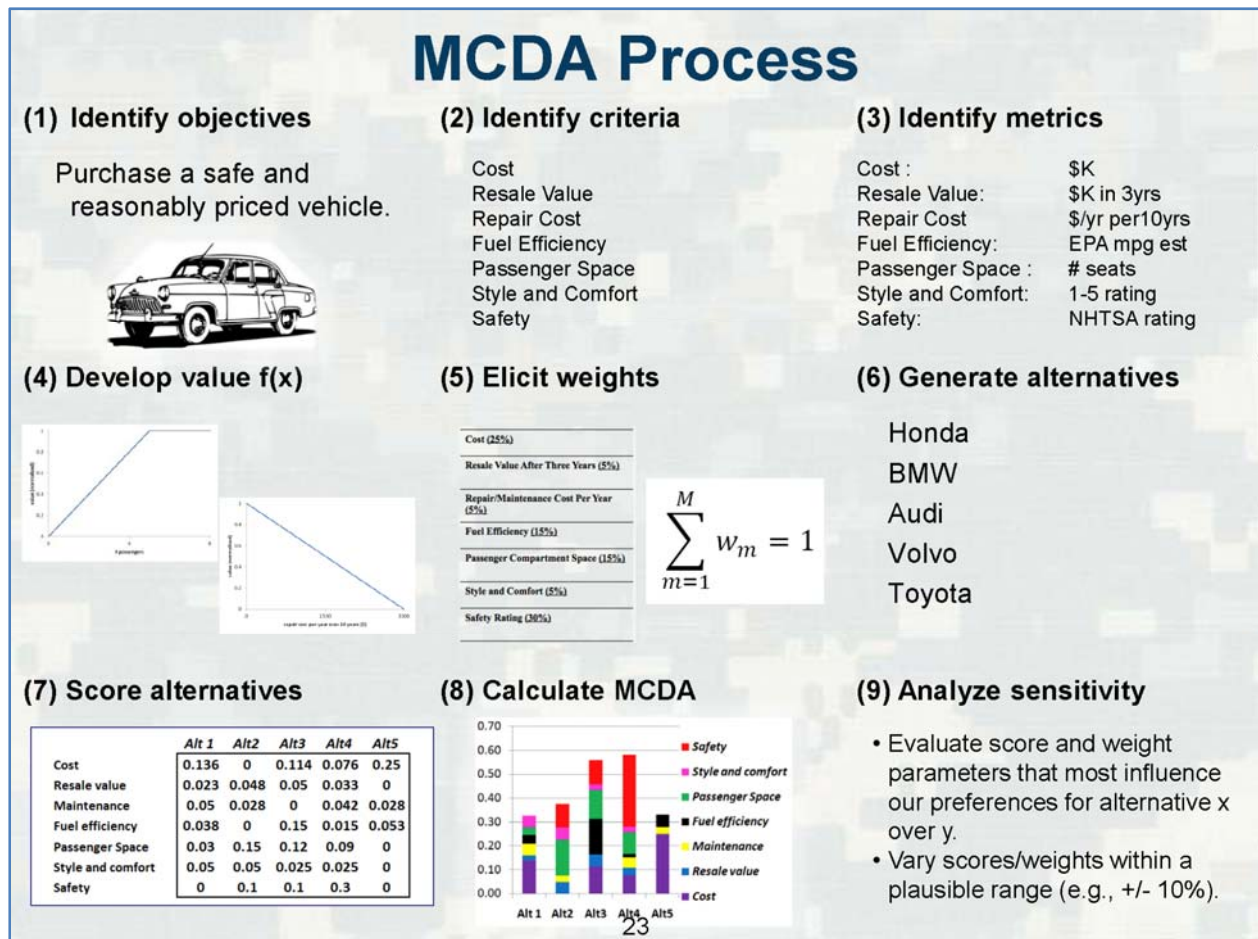


Figure 5: Example of Multi-criteria Decision Analysis (Linkov & Bates, 2016)

Figure 6 provides an example of community resilience objectives for a coastal town in the context of Structured Decision Making (SDM). SDM, a decision analysis method that uses an MCDA approach, is described in section 5.3 of the Vision 2025 report (Wee, 2019b).

Kornysheva & Deneckère, 2012 (see Section 4.1) propose an ontology for decisionmaking that reflects the components of SDM and MCDA.

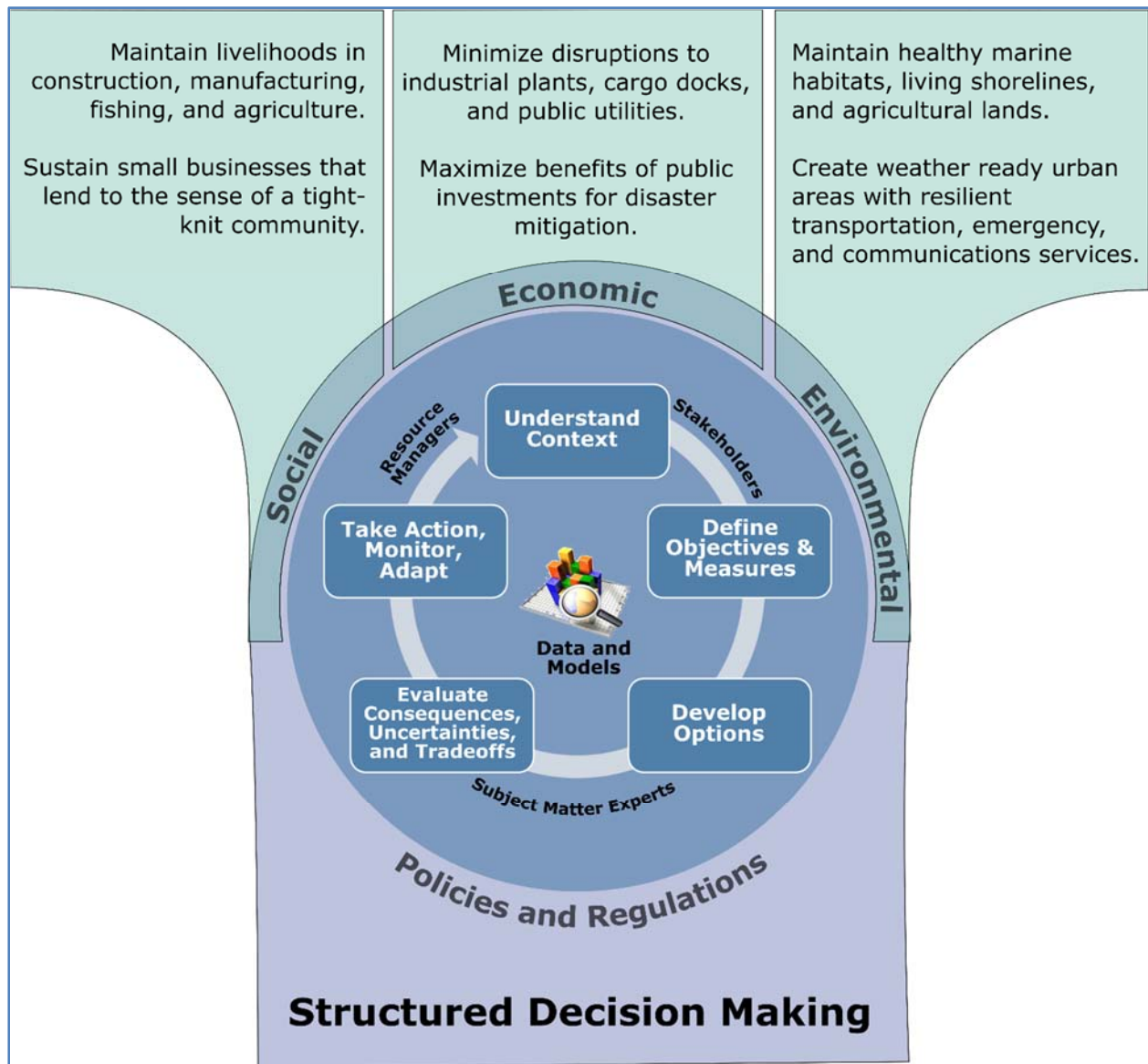


Figure 6: Hypothetical community resilience objectives for a coastal town that ultimately inform decisions

If we were to design an algorithm (e.g. supervised machine learning algorithm using tagged examples of text), what would such decisions look like? Given the interests and expertise of d2dprov team members, we focus on agriculture-resilience and flood-resilience related types of exemplars of text strings that resemble the type of decisions that are of interest to d2dprov.

3.3 Examples of agriculture-related decisions

- *“Transplant or cultivate traditional plant foods/medicines that are stressed to new more hospitable areas”* from the 2010 Xeni Gwet’in Community-based Climate Adaptation Plan (location: Canadian province of British Columbia) (Lerner et al., 2010).
- *“Mitigate invasive species and diseases”* from the Climate Change Adaptation Plan of the Blackfeet Nation (location: State of Montana) (*“Agriculture Sector in the Blackfeet Climate Change Adaptation Plan – Blackfeet Country and Climate Change,”* n.d.).
- *“Investigate alternative shellfish agriculture methods (e.g., suspended aquaculture; vertical sea gardens; clam gardens) to secure alternative food sources for the Tribe”* from the 2017 Stillaguamish Tribe of Indians Natural Resources Climate Change Adaptation Plan (location: State of Washington) (Binder et al., 2017).

3.4 Examples of flooding-related decisions

These “decisions” were excerpted from the City of Baltimore’s Disaster Preparedness Project and Plan (DP3). Examples of “decisions” include:

- IN 1, action #7: “Install external generator hookups for critical City facilities that depend on mobile generators for backup power” (Figure 7).
- IN 9, Action #1: “Prioritize infrastructure upgrades for roads identified at risk of flooding through the use of elevation data and Sea, Lake and Overland Surges from Hurricanes (SLOSH) model results” (Figure 8).
- IN 9, Action #2: “Raise streets in identified flood prone areas as they are redeveloped” (Figure 8).
- IN 9, Action #9: “Design and implement floodgates and barriers in transportation tunnels” (Figure 8).

Examples 1 through 4 above refer to “Generator hookups”, “infrastructure”, “streets”, and “floodgates” respectively. These four terms map to the ENVO ontology class “public infrastructure”. “Stormwater” is also defined in ENVO.

The SLOSH model is developed by NOAA’s National Weather Service.

DP3 INFRASTRUCTURE		Still Pending	Very Early Stages	Early Stages	Mid-Stages	Advanced Stages	Implemented/ Ongoing
IN 1 Protect and enhance the resiliency and redundancy of electricity system							
1	Work with the Maryland Public Service Commission (PSC) to minimize power outages from the local electric utility during extreme weather events by identifying and protecting critical energy facilities and located within the City	○	●	○	○	○	✓
2	Evaluate the City of Baltimore utility distribution system, and identify "underground utility districts" using BGE's May 2014 short term reliability improvement plan	○	●	○	○	○	✓
3	Support BGE's collaboration with the Maryland Public Service Commission to implement various smart grid solutions that will provide the City with real-time access to data during events	○	●	○	○	○	✓
4	Identify, harden, and water seal critical infrastructure relative to electrical, heating, and ventilation hardware within the flood plain	○	●	○	○	○	✓
5	Increase resiliency in our energy generation system by encouraging the development of decentralized power generation and developing fuel flexibility capabilities	○	●	○	○	○	✓
6	Develop a comprehensive maintenance and training program for City employees at facilities with backup generators to ensure proper placement, hook-up and function during hazard events.	●	○	○	○	○	✓
7	Install external generator hookups for critical City facilities that depend on mobile generators for backup power	●	○	○	○	○	✓
8	Partner with utility to evaluate protecting power and utility lines from all hazards	○	○	●	○	○	✓
9	Determine low-laying substation vulnerability and outline options for adaptation and mitigation	○	○	●	○	○	✓
10	Evaluate and protect low laying infrastructure - switching vaults, conduit and transformers	○	●	○	○	○	✓

Figure 7: Baltimore's Disaster Preparedness Project and Plan (IN1)

IN 9 Alter transportation systems in flood-prone areas in order to effectively manage stormwater							
1	Prioritize infrastructure upgrades for roads identified at risk of flooding through the use of elevation data and Sea, Lake and Overland Surges from Hurricanes (SLOSH) model results	●	○	○	○	○	✓
2	Raise streets in identified flood prone areas as they are redeveloped	●	○	○	○	○	✓
3	Encourage development of Green Streets in flood prone areas and throughout the City	○	○	●	○	○	✓
4	Encourage use of permeable pavement in non-critical areas – low-use roadways, sidewalks, parking lots and alleys where soils permit proper drainage	○	●	○	○	○	✓
5	Add pumps or other mitigation alternatives to streets as they are redeveloped (if needed)	●	○	○	○	○	✓
6	Assess need for new culvert capacity and identify where upgrades are needed	●	○	○	○	○	✓
7	Conduct an in-depth analysis of the impacts of drain fields that feed the harbor	○	●	○	○	○	✓
8	Expand and reinforce existing stormwater education programs	○	○	●	○	○	✓
9	Design and implement floodgates and barriers in transportation tunnels	●	○	○	○	○	✓

Figure 8: Baltimore's Disaster Preparedness Project and Plan (IN9)

3.5 Examples of “decisions” in a Record of Decision

In October 2017, the U.S. Department of Housing and Urban Development (HUD) authorized \$230 million for the Hudson River flood-resilience project. One month earlier, in September 2017, New Jersey’s Department of Environmental Protection’s (NJDEP) Bureau of Flood Resilience released a ROD entitled “Rebuild by Design. Resist. Delay. Store. Discharge.” (*Record of Decision. Rebuild By Design: Resist, Delay, Store, Discharge Project. Cities of Hoboken, Weehawken, and Jersey City Hudson County, New Jersey*, 2017) The contents page of the ROD is shown in Figure 9.

Record of Decision	
Rebuild By Design: Resist, Delay, Store, Discharge Project Cities of Hoboken, Weehawken, and Jersey City Hudson County, New Jersey	
TABLE OF CONTENTS	
	Page
1.0 INTRODUCTION	1
2.0 PROJECT SUMMARY	3
3.0 DECISION	5
4.0 OTHER ALTERNATIVES CONSIDERED	9
5.0 ALTERNATIVES ANALYSIS.....	11
6.0 SUMMARY OF ENVIRONMENTAL IMPACTS	14
7.0 PROGRAMMATIC AGREEMENT	16
8.0 MITIGATION MEASURES TO AVOID OR MINIMIZE HARM.....	17
9.0 MONITORING/ENFORCEMENT AND ONGOING COORDINATION	18
10.0 COMMENTS ON FINAL EIS.....	20
11.0 APPROVAL	25

Figure 9: Record of Decision for New Jersey Flood Mitigation Plan

Excerpts of Section 3 of the ROD are reproduced below. These excerpts are useful for automated concept extraction because of the need to explicitly link these decisions to data products.

1. “NJDEP approves the selection of Alternative 3 as the Project identified in the Final Environmental Impact Statement (FEIS) for the RBD-HR Project. The flood-resist structure selected for construction as the Project will provide flood risk reduction for the City of Hoboken, parts of Jersey City and Weehawken and for critical infrastructure located in those communities, including three fire stations, one hospital and the North Hudson Sewerage Authority (NHSA) wastewater treatment plant. The Project provides coastal flood risk reduction to approximately 85 percent of the population residing within the Study Area’s 100-year floodplain.” (Page 5)
2. “Resist structure heights (also known as the "Design Flood Elevation" or "DFE") were developed for all segments of the Resist infrastructure for the EIS and Feasibility Study. The DFE’s were developed using the criteria stated in 44 CFR 65.10 and by incorporating sea level rise. The DFE's were based on the FEMA Base Flood Elevation (BFE) for the one percent annual chance flood

(100-year flood) plus an additional 2.34 feet in elevation to account for possible sea level rise by 2075, based on NOAA's intermediate-high projections, as well as one foot of freeboard. Depending upon the location (i.e., waterfront or inland), the DFE values are different. For locations along or near the waterfront where wave action would be expected during a coastal surge event (such as along Weehawken Cove and Lincoln Harbor), the criteria stated in 44 CFR 65.10 required the use of additional structure height to accommodate for wave run-up to prevent potential overtopping of the structure by wave action.” (Page 5)

3.6 Examples of “decisions” that we shall not focus on using the flood-mitigation exemplar

These are decisions that lack a level of specificity where a course of action is unlikely to be informed by data and models. Examples include:

- IN 1, action #1: “Work with the Maryland Public Service Commission (PSC) to...” (Figure 7). This is clearly the result of a decision. However, successfully engaging the PSC is not obviously contingent on the use of quantitative or qualitative scientific analyses. One could argue successful PSC engagement requires Social Network Analyses (which would require data and models), but that would be a stretch.
- IN 1, action #6: “Develop a comprehensive maintenance and training program...” (Figure 7). This is clearly another action. However, it is not evident that you would use a scientific analysis to inform the development of the maintenance and training program. In such cases, a financial and programmatic analysis might be more relevant, which is not the focus of d2dprov.

3.7 Are activities after a “decision” within the scope of d2dprov?

What happens after a “decision” is made? Projects often employ formative evaluation or summative evaluation techniques to monitor and evaluate (often referred to as “M&E”) the efficacy of actions arising from decisions.

Providing full traceability from data-to-decisions enables adaptive management. If decisions are traceable, decision-makers should be able to monitor and evaluate the results of decisions and re-visit the entire data-to-decisions workflow and tweak processes accordingly. The concepts developed in d2dprov should, therefore, be highly applicable for M&E professionals who specialize in the practice of project performance improvement.

3.0 Policy Needs

3.1 Harmonized set of resilience planning protocols

Section 4 of the Vision 2025 report (Wee, 2019b) outlines the role of resilience planning frameworks for the modeling of the provenance for data-driven, science-informed resilience planning, and provides examples of such frameworks used by the US Climate Resilience Toolkit (Figure 3), the USAID, and the United Kingdom's UKCIP risk framework. All these frameworks provide a means to describe the sequence of steps that resilience planners should undertake.

Such frameworks are largely interoperable: steps in one planning framework can be approximately mapped to a step in another framework. For example, the components of the US CRT framework can be easily mapped to the UKCIP framework. The definition of equivalence of steps between frameworks facilitates interoperability between the frameworks.

Creating a small library of cross-walked, interoperable resilience planning protocols would facilitate better discovery of resilience plans that share the same overall sequence of planning.

3.2 Standardized vocabulary for publicly accessible documents

Federal agencies that disburse grants for mitigation planning should encourage **grant recipients** to produce publicly accessible documents that are **structured into sections that are named using vocabulary that reflect that agency's preferred mitigation planning framework**. Section 4.3 explains the advantage of using such a standardized vocabulary.

The focus here is on **publicly accessible documents**, that include:

- Documents that are required by law, like Records of Decision and Environmental Impact Statements;
- Documents that are associated with legally mandated opportunities for public input. Even though such documents themselves may not be required by law, there are instances where the public is legally required to be involved in the decisionmaking process, presumably by being provided access to relevant documents. For example, FEMA's Hazard Mitigation Grant Program provides funding for communities to submit a disaster mitigation plan after a Presidential disaster declaration. In relation to the process for submitting a grant to FEMA for disaster planning, 44 CFR § 201.6(b) states that an

"open public involvement process is essential to the development of an effective plan. In order to develop a more comprehensive approach to reducing the effects of natural disasters, the planning process shall include: An opportunity for the public to comment on the plan during the drafting stage and prior to plan approval".

4.0 Technical Needs

4.1 Schemas for representing decisions

Further prototyping is required to develop an ontology that sufficiently models decisions of interest. Useful publications in this regard include:

1. “Using an ontology for modeling decision-making knowledge” (Kornysheva & Deneckère, 2012). This paper reflects many of the elements of Structured Decision Making (SDM: see Section 3.2) including concepts like “preference”, “alternative”, “criterion”, “consequence”, “goal”, “stakeholder”. The schema is shown in Figure 11.
2. “Modelling causes for actions with the Decision and PROV ontologies” (Car, 2017). An OWL implementation exists for this ontology. This ontology may be an adequate light-weight alternative to the approach proposed in Kornysheva & Deneckère 2012, because the latter is likely to require very nuanced parsing of human-readable documents if we are to successfully represent complex decisions in machine-readable equivalents. By comparison Car 2017 is more “forgiving” in terms by virtue of modeling decisions at a higher conceptual level (Tom Narock, personal communications). The ontology proposed by Car 2017 is shown in Figure 10.
3. “The multi-entity decision graph decision ontology: A decision ontology for fusion support” (Locher & Costa, 2017). this paper incorporates decisionmaking under uncertainty and bears some similarities with SDM.
4. “Framework for ontology-driven decision making” (Baclawski et al., 2017). This paper proposes an ontology for decisionmaking that incorporates PROV.

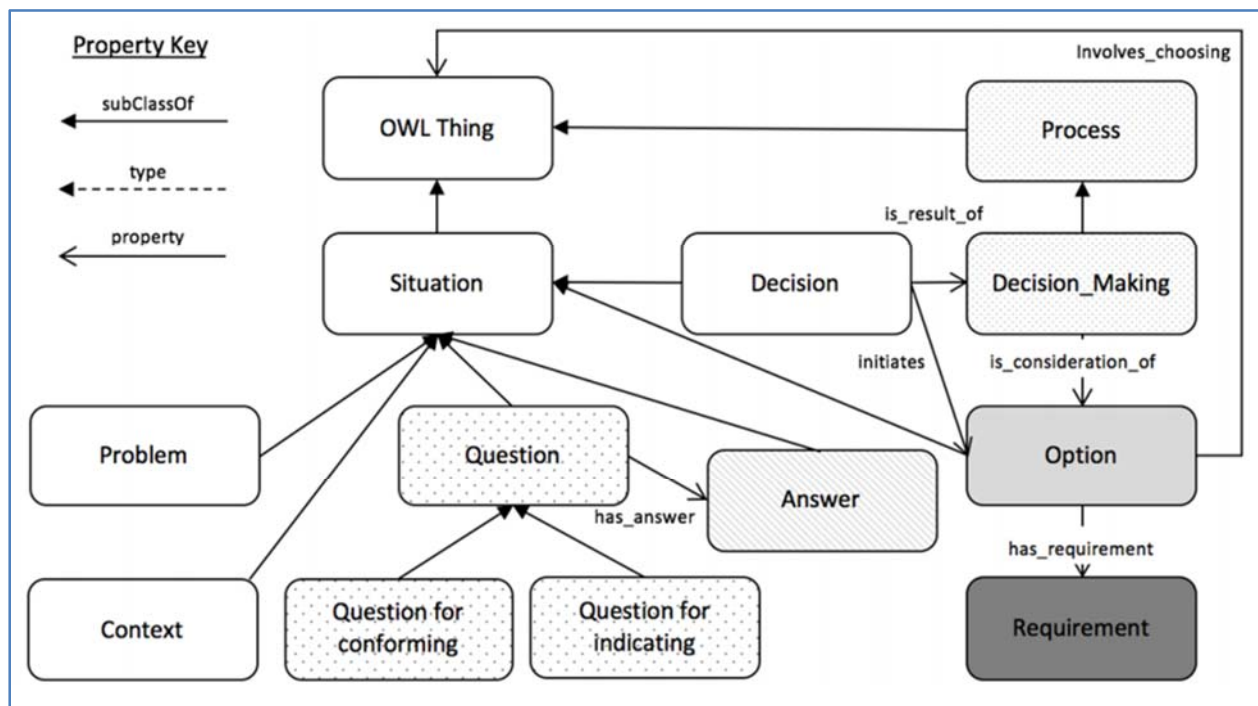


Figure 10: Decision Ontology (Car, 2017)

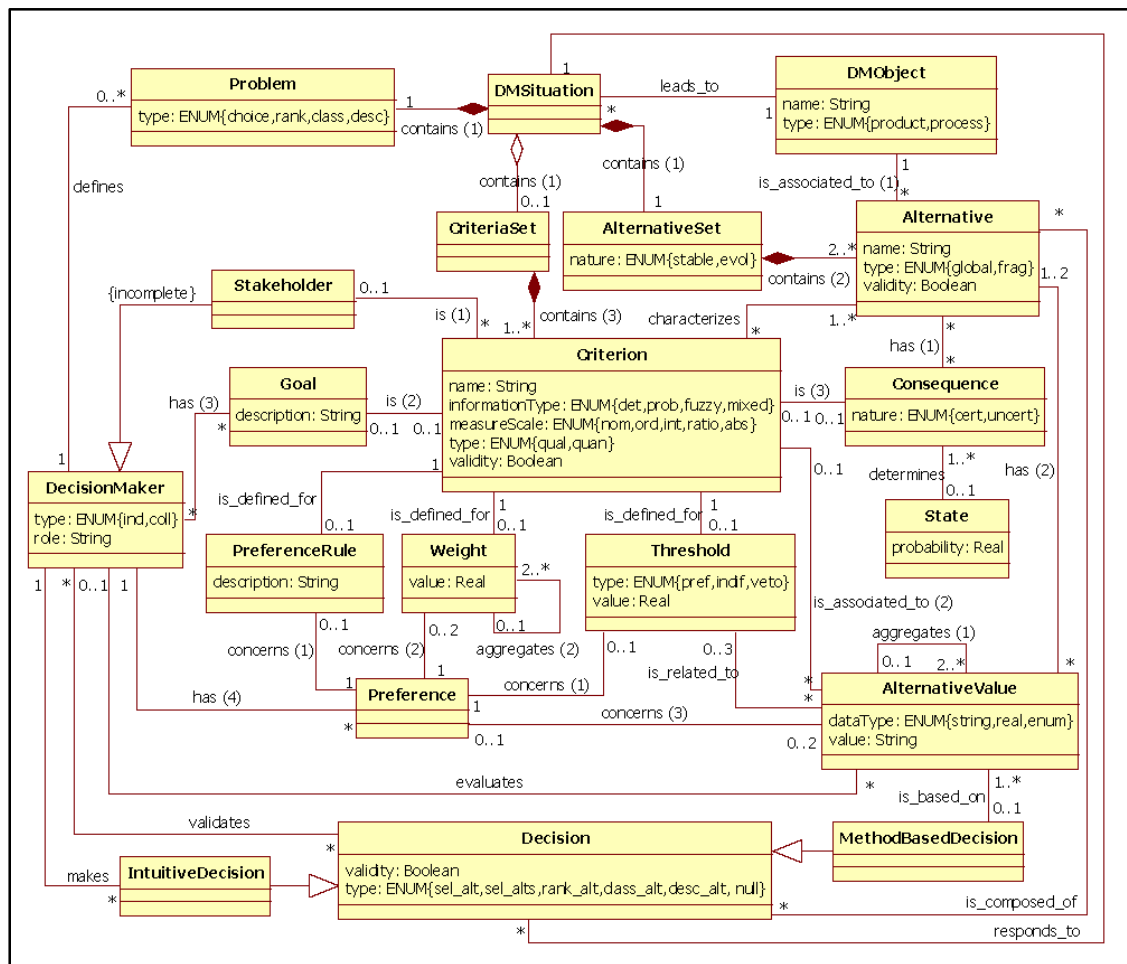


Figure 11: Decision Making Ontology (Kornysheva & Deneckère, 2012)

4.2 Knowledge graphs

Document associated with decisions that are informed by scientific data and models (such as those described in Sections 3.3, 3.4, and 3.5) could be annotated using knowledge graphs. The knowledge graphs could utilize concepts in ontologies like ENVO, like the example from the City of Baltimore given in Section 3.4.

One of the precursor activities that led up to d2dprov was the ESIP “Data to Decisions for Climate Resilience” cluster. One of the products of the cluster was a loosely-constrained prototype knowledge graph (Wee, 2017). The graph was instantiated based on an existing multi-billion dollar, multi-decadal large watershed-scale climate adaptation project called the Yakima River Basin Water Enhancement Project (YRB). The graph was “loosely-constrained” because the graph creation was not meant to be machine-readable. Very little attention paid to properly designing the predicates between graph nodes such that those predicates could be used for light-weight machine inferencing. Moreover, the graph utilized concepts that were not mapped to any ontology.

Team member Tom Narock developed Python code that extracted concepts from human-readable documents (Narock et al., 2019). Tom’s code used the Open Biological and Biomedical Ontology (OBO) Foundry’s Environmental Ontology (ENVO) and the United Nation’s Sustainable Development Goals

Interface Ontology (SDGIO). The code also used the Python based NLTK library which a natural language toolkit.

Tom also reported that the predicates (i.e. relationships that connect two concepts in a knowledge graph) used in SDGIO and NASA's Semantic Web for Earth and Environmental Terminology (SWEET) ontology could not be utilized for d2dprov's purpose (Narock et al., 2019).

Examples of initiatives that appear to have established formal, tightly-constrained graph technologies include:

- The Cyc project (<http://www.cyc.com/kb/>) which has been running for more than two decades has a list of more than 10,000 predicates and hundreds of thousands of concepts. Those predicates could be used to inform how knowledge graphs for resilience planning. Although the Cyc database and an application that functions as a query interface to the database comes with a free license for academic use, there is an application process to get the license.
- Google uses its proprietary knowledge graph technology for search.
- Kbpedia's (<http://kbpedia.org/>) knowledge structure "combines seven 'core' public knowledge bases — Wikipedia, Wikidata, schema.org, DBpedia, GeoNames, OpenCyc, and UMBEL — into an integrated whole".

Jovanović & Bagheri, 2017 provide a review of biomedical ontologies and semantic annotators that utilize ontologies with the objective of "improved clinical decision making". Rotmensch, Halpern, Tlimat, Horng, & Sontag, 2017 discuss how concepts extracted from clinical records (Figure 12) can be used to construct knowledge graphs (Figure 13).

Further work is needed to replicate these techniques for extracting concepts from resilience planning decisions, and then relating the concepts into knowledge graphs.

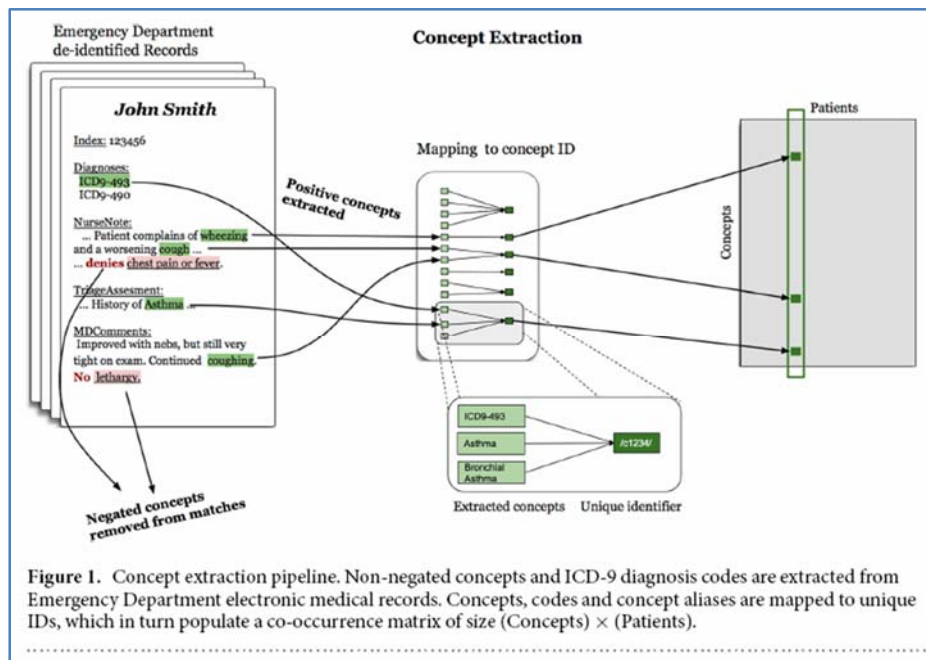


Figure 1. Concept extraction pipeline. Non-negated concepts and ICD-9 diagnosis codes are extracted from Emergency Department electronic medical records. Concepts, codes and concept aliases are mapped to unique IDs, which in turn populate a co-occurrence matrix of size (Concepts) $\times$ (Patients).

Figure 12: Extraction of concepts from human-readable documents (Rotmensch et al., 2017)

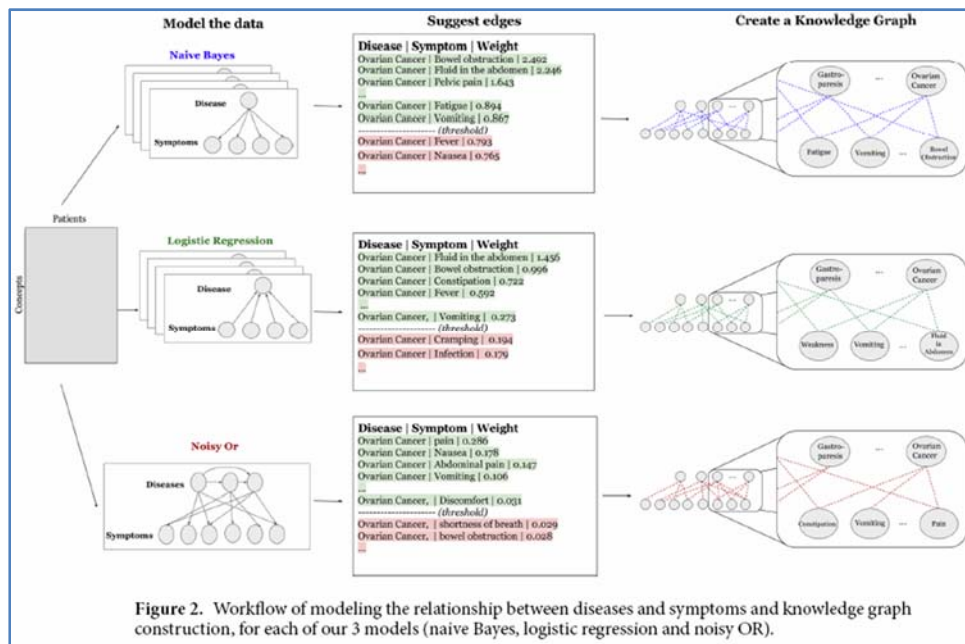


Figure 2. Workflow of modeling the relationship between diseases and symptoms and knowledge graph construction, for each of our 3 models (naive Bayes, logistic regression and noisy OR).

Figure 13: Creating a knowledge graph after concept extraction (Rotmensch et al., 2017)

4.3 Encoding of harmonized resilience planning steps into a reference protocol

Section 3.1 calls for the need to create a small library of cross-walked, interoperable resilience planning protocols. Once the equivalence of terminologies between planning protocols is determined, harmonized terms can be used to tag nodes in knowledge graphs to relate nodes in graphs to resilience planning steps (Figure 14).

To facilitate this strategy, the section-headings of human-readable documents should be named using terms that are prescribed by the funding agency as per the agency's preferred planning framework (see Section 3.1). These terms are ultimately relatable to different planning frameworks that may be used by other agencies through the harmonized reference protocol. This facilitates the searching of knowledge graphs regardless of which agency funded the resilience planning effort.

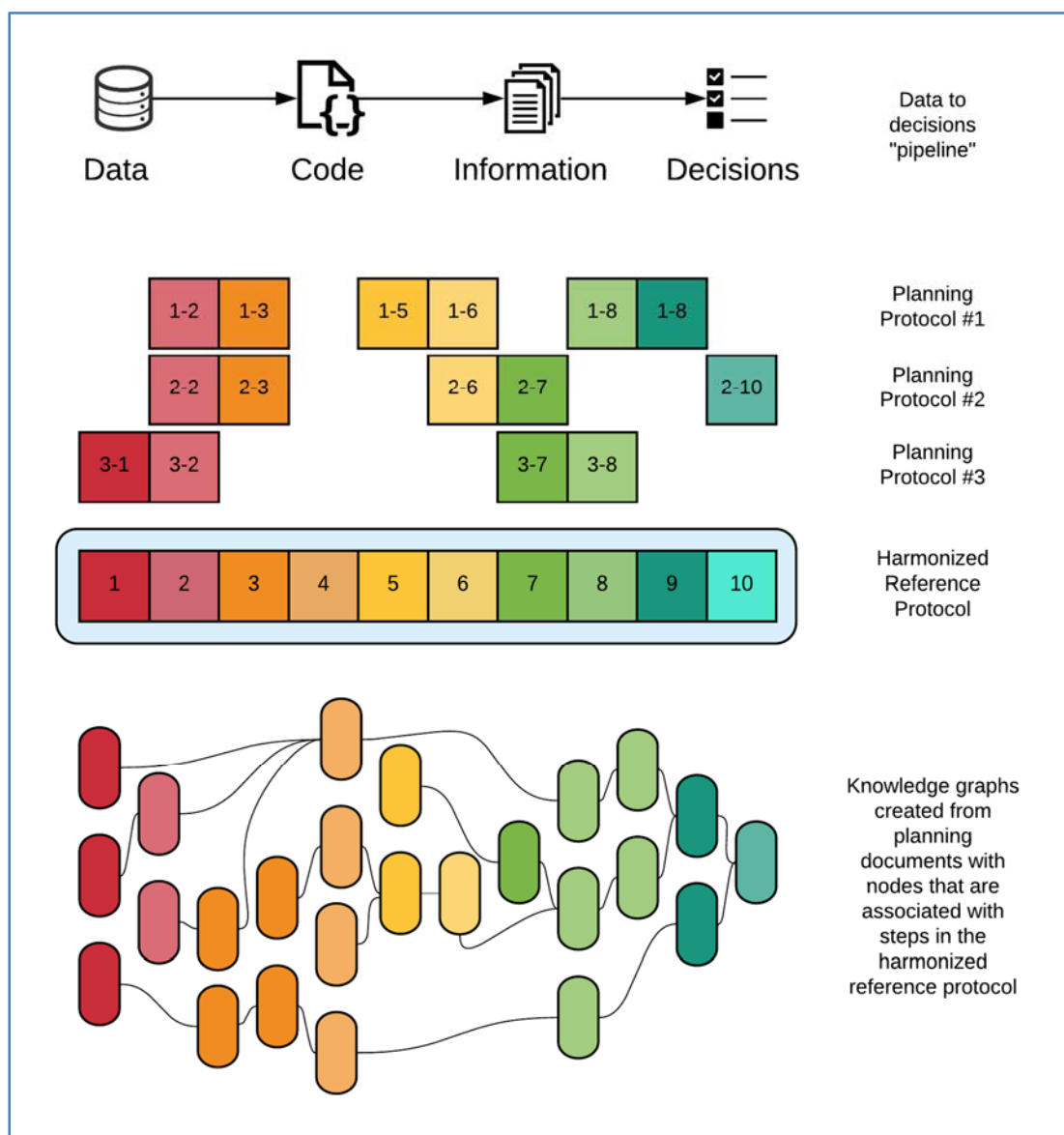


Figure 14: Resilience Genome

5.0 Bibliography

- Agriculture Sector in the Blackfeet Climate Change Adaptation Plan – Blackfeet Country and Climate Change. (n.d.). Retrieved February 6, 2019, from <https://blackfeetclimatechange.com/our-environment/climate-change-adaptation-plan/agriculture/>
- Baclawski, K., Chan, E. S., Gawlick, D., Ghoneimy, A., Gross, K., Liu, Z. H., & Zhang, X. (2017). Framework for ontology-driven decision making. *Applied Ontology*, 12(3–4), 245–273. <https://doi.org/10.3233/AO-170189>
- Binder, L. W., Morgan, H., Krosby, M., Sevigny, J., Summers, A., & Neuffer, T. (2017). *Stillaguamish Tribe of Indians Natural Resources Climate Change Adaptation Plan*. Retrieved from <https://cig.uw.edu/wp-content/uploads/sites/2/2014/11/SNRD-Adaptation-Plan.pdf>
- Car, N. J. (2017). *Modelling causes for actions with the Decision and PROV ontologies*. Retrieved from <http://promsns.org/def/decprov>
- Jovanović, J., & Bagheri, E. (2017). Semantic annotation in biomedicine: the current landscape. *Journal of Biomedical Semantics*, 8(1), 44. <https://doi.org/10.1186/s13326-017-0153-x>
- Kornysheva, E., & Deneckère, R. (2012). Using an ontology for modeling decision-making knowledge. *Frontiers in Artificial Intelligence and Applications*, 243(September 2012), 1553–1562. <https://doi.org/10.3233/978-1-61499-105-2-1553>
- Lerner, J., Rossing, T., Delong, D., Holmes, R., McCrory, W., Mylnowski, T., & Opperman, N. (2010). *Xeni Gwet'in Community-based Climate Change Adaptation Plan*. Retrieved from <https://www.cakex.org/sites/default/files/documents/XENI GWETIN CBA PLAN 2010.pdf>
- Linkov, I., & Bates, M. (2016). *Multi-Criteria Decisional Analysis: Methodology & Case Studies*. Retrieved from <http://el.erdc.usace.army.mil/riskdecision/index.html>
- Locher, M., & Costa, P. C. G. (2017). The multi-entity decision graph decision ontology: A decision ontology for fusion support. In *Information Fusion (Fusion), 2017 20th International Conference on* (pp. 1–8).
- Narock, T., Wee, B., Hoebelheinrich, N., Albayrak, R., & Teng, B. (2019, February 20). ESIPFed/d2dprovenance: Data to Decisions Provenance. <https://doi.org/10.5281/ZENODO.2574103>
- Record of Decision. Rebuild By Design: Resist, Delay, Store, Discharge Project. Cities of Hoboken, Weehawken, and Jersey City Hudson County, New Jersey.* (2017). Hoboken, NJ. Retrieved from <https://www.nj.gov/dep/floodresilience/docs/rbdh-rod-201709.pdf>
- Rotmensch, M., Halpern, Y., Tlimat, A., Horng, S., & Sontag, D. (2017). Learning a Health Knowledge Graph from Electronic Medical Records. *Scientific Reports*, 7(1), 5994. <https://doi.org/10.1038/s41598-017-05778-z>
- Wee, B. (2017). Data to Decisions for Climate Resilience Knowledge Graph Prototype. Retrieved February 7, 2019, from <https://cmapscloud.ihmc.us/viewer/cmap/1RR3CGPBQ-1BQMS3Z-C11>
- Wee, B. (2019a). *D2dprov: Statement of Needs 2022. Technology and policy requirements to fulfill Vision 2025's proposed approach to data-driven, science-informed climate resilience decisions*. Washington, DC. <https://doi.org/10.6084/m9.figshare.7692038>

Wee, B. (2019b). *D2dprov: Vision 2025. A transdisciplinary science, technology, and policy vision for data-driven, science-informed resilience planning for 2025 and beyond*. Washington, DC.
<https://doi.org/10.6084/m9.figshare.7591238>