future generations to meet their own needs[12]. Resilient systems share qualities of sustainable systems because they are able to minimize the negative impacts of adverse events on societies and sustain or even improve their functionality by adapting to and learning from fundamental changes caused by those events.

In summary, risk analysis and risk management based on probabilistic quantitative methods have been widely adopted and have been useful for dealing with foreseeable and calculable stress situations. Benchmarks and thresholds for risk analysis are built into the regulations and policies of organizations and nations; however, this approach is no longer sufficient to address the evolving nature of risks in the modern world. Moreover, the increased complexity and interdependency of many of society's critical networks presents a fundamental challenge to even the most comprehensive and sophisticated risk analysis. Therefore, early integration of resilience into the design of systems and the regulatory structures of systems management is needed to address the emerging issues associated with complexity and uncertainty. An urgent need exists to complement the existing knowledge-base of risk analysis and management by further developing frameworks and models enabling system-wide and network-wide resilience analysis, engineering and management. Although research and development on methods and tools is progressing, establishing channels of communication for transparent dialogue on resilience management with stakeholders, such as industry associations and policymakers, is essential for the timely and broad acceptance of resilience concepts. ❒

Igor Linkov[1], Todd Bridges[2], Felix Creutzig[3], Jennifer Decker[4], Cate Fox-Lent[1], Wolfgang Kröger[5], James H. Lambert[6], Anders Levermann[7], Benoit Montreuil[8], Jatin Nathwani[9], Raymond Nyer[10], Ortwin Renn[11], Benjamin Scharte[12], Alexander Scheffler[13], Miranda Schreurs[14] and Thomas Thiel-Clemen[15] are at [1]United States Army Corps of Engineers — Engineer Research and Development Center, Environmental Laboratory, 696 Virginia Road, Concord, Massachusetts 01742, USA, [2]United States Army Corps of Engineers — Engineer Research and Development Center, Environmental Laboratory, 3909 Halls Ferry Road, Vicksburg, Massachusetts 39180, USA, [3]Mercator Research Institute on Global Commons and Climate Change, Torgauer Straβe 12–15, 10829 Berlin, Germany, [4]Embassy of Canada, Leipziger Platz 17, 10117 Berlin, Germany, [5]Swiss Federal Institute of Technology Zürich (ETH), Scheuchzerstrasse 7, 8092 Zürich, Switzerland, [6]University of Virginia, 151 Engineer's Way, Charlottesville, Virginia 22903, USA, [7]Potsdam Institute for Climate Impact Research, Telegrafenberg A 31, 14191 Potsdam, Germany, [8]Université Laval, 2325 Rue de l'Université, Québec G1V 0A6, Canada, [9]University of Waterloo, 200 University Ave W, Waterloo, Ontario N2L 3G1, Canada, [10]RNC Conseil and Ecole Centrale de Paris, 56 Rue Charles Laffitte, 92200 Neuilly-sur-Seine, France, [11]University of Stuttgart, Seidenstraβe 36, 70174 Stuttgart, Germany, [12]Fraunhofer Institute for High-Speed Dynamics, Eckerstraβe 4, 79104 Freiburg, Germany, [13]Hamburg University of Technology, Kasernenstraβe 12, 21073 Hamburg, Germany, [14]Free University of Berlin, Ihnestraβe 22, 14195 Berlin, Germany, [15]Hamburg University of Applied Sciences, Lohbrügger Kirchstrasse 65, 21033 Hamburg, Germany.*
**e-mail: igor.linkov@usace.army.mil*

References

1. IPCC *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change* (eds Field, C. B. *et al.*) 582 (Cambridge Univ. Press, 2012).
2. Levermann, A. *Nature* **506,** 27–29 (2014).
3. Linkov, I. *et al. Env. Sci. Tech.* **47,** 10108–10110 (2014).
4. National Research Council *Disaster Resilience: A National Imperative* (The National Academies Press, 2012).
5. Park, J., Seager, T. P., Rao, P. C. S., Convertino, M. & Linkov, I. *Risk Analysis* **33,** 356–367 (2013).
6. Linkov, I. *et al. Environ. Syst. Decisions* **33,** 471–476 (2013).
7. Roege, P. *et al. Energy Policy* (in the press).
8. Dauphiné, A. & Provitolo. D. *Annales de Géographie* **654,** 115–125 (2007).
9. Lambert, J. H., Tsang, J. & Thekdi, S. *Am. Soc. Civ. Eng. J. Infrastr. Syst.* **19,** 384–394 (2013).
10. Bridges, T. *et al. Coastal risk reduction and resilience* (US Army Corps of Engineers, 2013).
11. Hollnagel, E. & Fujita, Y. *Nuc. Eng. Tech.* **45,** 1–8 (2013).
12. World Commission on Environment and Development *Our common future* (Oxford Univ. Press, 1987).

Acknowledgements

This paper resulted from discussions at a workshop on the use of risk and resilience assessment methodologies in guiding policy development, held at the Embassy of Canada in Berlin, Germany on 4 February 2014. Partial financial support for the workshop was provided by grant W911NF-13-1-0168 to RNC Conseil, Neuilly-sur-Seine, France. Additional funding was provided by the German National Academy of Technology and Engineering (Acatech), the Helmholtz Association, the Embassy of Canada and the Society for Risk Analysis. Permission was granted by the USACE Chief of Engineers to publish this material. The views and opinions expressed in this paper are those of the individual authors and not those of the US Army or other sponsor organizations.

COMMENTARY:

Capturing provenance of global change information

Xiaogang Ma, Peter Fox, Curt Tilmes, Katharine Jacobs and Anne Waple

Global change information demands access to data sources and well-documented provenance to provide the evidence needed to build confidence in scientific conclusions and decision making. A new generation of web technology, the Semantic Web, provides tools for that purpose.

The topic of global change covers changes in the global environment that may alter the capacity of the Earth to sustain life and support human systems[1]. This includes changes to climate, land productivity, oceans or other water resources, atmospheric composition and/or chemistry and ecological systems. Data and findings associated with global change research are of great public, government and academic concern and are used in policy and decision making, which makes the provenance of global change information especially important. In addition, because different types of decisions benefit from different types of information, understanding how to capture and present the provenance of global change information

Table 1 | A list of ontologies and schemas, their namespace prefixes and corresponding URIs that have been used in GCIS.

Full name	Namespace prefix	Namespace URI
The PROV Ontology	prov	http://www.w3.org/ns/prov#
The Bibliographic Ontology	bibo	http://purl.org/ontology/bibo/
The Dublin Core Metadata Initiative Metadata Terms	dct	http://purl.org/dc/terms/
The GCIS Ontology	gcis	http://data.globalchange.gov/gcis.owl
The Resource Description Framework Schema	rdfs	http://www.w3.org/2000/01/rdf-schema#
The eXtensible Markup Language Schema	xsd	http://www.w3.org/2001/XMLSchema#

is becoming more of an imperative in adaptive planning.

Provenance tracking

Many science issues in global change research are complex and have multiple implications. Take, for example, projections of sea-level rise. In some cases, including in academic research and for broad-scale policy discussion, it is useful to examine probabilistic estimates of future sea-level rise, such as the 'most likely' change over time. In other instances, such as for the risk-based framing of coastal management decisions, it is helpful to refer to a broader range of plausible sea-level changes. In 2013, the Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC)[2] examined four different greenhouse gas emissions pathways and associated projections of sea-level rise. The report concluded that it was 'likely' that global sea levels would be between 0.26 and 0.81 metres above current levels by the last two decades of this century. As a comparison, a report published in 2012 by the Climate Program Office (CPO) of the National Oceanic and Atmospheric Administration (NOAA)[3] in support of the National Climate Assessment[4], provided a wider range of plausible scenarios describing a sea-level rise of 0.2 to 2.0 metres above 1992 levels by 2100.

The above two reports took different approaches for different reasons. The CPO report included higher and lower estimates of sea-level rise because they were considered useful in risk-based decision making even though the IPCC considered them less 'likely'. In other words, awareness of low-probability and high-impact futures is useful for coastal decision making, but arguably less valuable for broad-scale policy framing. For decision makers to understand (and feel confident in their understanding of) the value and nature of the different approaches, scenarios and projections, it is essential for them to be able look into the data, models and analytical sources behind the reports. And they need to be able to access this information easily and in an interpretable format — a goal that provenance tracking can support.

The literal meaning of provenance is the origin or source of something. As discussed above, understanding the origins of global change information is not only critical for understanding its use, but also facilitates the integration of knowledge across disciplines. This kind of interdisciplinary assessment is important for managing the risks associated with global change. For the above example, the AR5 and the CPO report provided textual descriptions for tracking the provenance of their projections, but the descriptions are relatively technical and not easy to interpret in terms of what is 'best' or most appropriate in any given circumstance. A more useful approach, in the long term, is to provide search options that facilitate access to data and allow traceability back to original sources, authors, programs and even observing systems.

Although provenance tracking is useful and important, as shown by the above example, capturing and presenting provenance is not an easy task. In scientific works, documenting provenance includes linking a range of observations and model outputs, research activities, people and organizations involved in the production of scientific findings with the supporting data sets and methods used to generate them[5]. It requires significant effort to identify, extract, link and assemble pieces of information from accumulated documents, codes, data sets and so on.

Categorization and annotation

The root of provenance-capture grows from the soil of metadata collection. Metadata are data about data. Traditionally seen in library catalogues, metadata have received significant attention in the past decade and several metadata standards have been developed to address the recent data deluge. One widely used standard is the Dublin Core metadata schema[6], which consists of 15 core elements — such as title, creator, subject, description, publisher and so on — that can be used to describe a resource. The schema is simple, which makes its use convenient. However, the schema also has a weakness because it does not define categories of resources and the provided core elements do not accommodate the rich annotation of a resource, such as the geographical and temporal location or resolution of a data set that could be useful for narrowing down its applicability. Even with the extended metadata terms[7] released by the Dublin Core Metadata Initiative (DCMI) — an open organization supporting shared innovation in metadata design and best practices — such an issue still exists. For example, the resource type 'Bibliographic Resource' in Table 2 is a term in the DCMI extended terms[7], whereas the specific type 'Figure', which refers to Fig. 1, is not and has to be added by the metadata curator.

Other metadata standards further extend the functionality for annotating the production process of a resource, such as the data lineage model in ISO 19115 (ref. 8) and ISO 19115-2 (ref. 9). Two initiatives, the Proof Markup Language[10] and the Open Provenance Model[11], have developed enriched categories and properties for representing and capturing provenance. Three top categories — entity, process and agent — arise from those works. They can be used to describe a process such as the generation of a figure, the source entities such as data sets and models used in the process, and the agents such as people and/or organizations that participated in the process. Those categories became the core of the recent World Wide Web (WWW) Consortium (W3C) PROV Data Model (PROV-DM)[12] that was derived from earlier efforts.

Another significant advancement in those metadata schemas and provenance models[7,10–12] is built on the principles of linked data — using Uniform Resource Identifiers (URIs) as identifiers of resources, rather than literal values. For example, the third reference record in Table 2 may be replaced by a URI (http://data.globalchange.gov/article/10.1080/01490419.2010.491031) that points to a web page with more information about that document, such as the title, document type, source journal, publication year and so on. The use of URIs is just one of the many features that are enabled by the Semantic Web, a new generation of the WWW[13]. The Semantic

Web presents a 'web of data' compared to the traditional 'web of documents'. It adds machine-readable meanings such as more specifically defined categories and annotations to data by using ontologies (specifications of concepts and relations among them) and vocabularies (arranged terms of certain topics) encoded in the Resource Description Framework (RDF) format[14]. For instance, the web page corresponding to the above URI is just the front end of data stored in an online database. Whereas humans read the textual description on the web page and know that it is about a journal article, machines can read data from the back end and recognize the resource type 'bibo:Article' (Table 1) by tracking the ontologies used in the data.

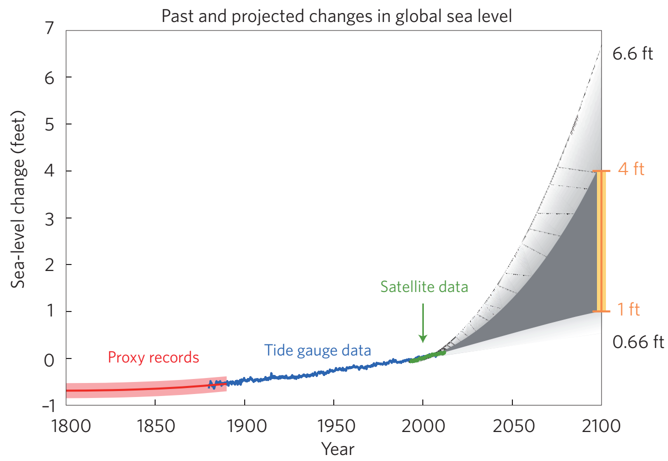

Figure 1 | Global sea-level rise figure associated with the collected metadata records in Table 2. Figure reproduced from ref. 19.

Linking for tracking

The aforementioned categorization and annotation focuses on the description of individual entities, processes and agents. The Semantic Web allows links to be established among those individual instances, such that provenance tracking permits one to retrieve not only the literal description of a data set but also an accessible or downloadable version of the data set itself. In the Global Change Information System (GCIS)[15] — under development through the U.S. Global Change Research Program (USGCRP) — those Semantic Web technologies have been used to capture and present provenance information. The initial focus of GCIS is to support the third United States National Climate Assessment (NCA3). It will present the NCA3 report and also incorporate integrated access to interlinked resources supporting that report. This significantly enhanced transparency also improves the ability of decision makers to understand the conclusions of the report and to use the underlying data for their own purposes.

The interlinks among provenance information in the GCIS are realized by using properties defined in a number of ontologies, including the PROV Ontology (PROV-O)[16] — an ontology for representing and interchanging provenance information from the W3C PROV-DM. For example, Fig. 2 shows a part of the provenance information captured from the third reference in Table 2. Most of the properties in Fig. 2 are from the PROV-O. Properties from a few other domain-specific ontologies, such as the Bibliographic Ontology and the GCIS Ontology (Table 1), are also used. Those specific properties better describe a few relationships — such as those between instruments and sensors — than the general properties from the PROV-O.

By categorizing, annotating and linking provenance information, the finished GCIS will be capable of answering provenance-tracking questions for the final NCA3, such as: (1) which data sets were used in the analysis and projection of global sea-level rise? (2) which agencies and

Table 2 | Metadata records of Fig. 1 collected using Dublin Core metadata elements.

Metadata element	Record
Type	Bibliographic Resource/Figure
Identifier	2.26
Title	Global sea-level rise
Description	Estimated, observed and possible amounts of global sea-level rise from 1800 to 2100. Proxy estimates (Kemp *et al.* 2012) (for example, based on sediment records) are shown in red (the pink band shows uncertainty), tide gauge data in blue (Church and White 2011a) and satellite observations are shown in green (Nerem *et al.* 2010). The future scenarios range from 0.66 feet to 6.6 feet in 2100 (Parris *et al.* 2012). Higher or lower amounts of sea level rise are considered implausible, as represented by the gray shading. The orange line at right shows the currently projected range of sea-level rise of 1 to 4 feet by 2100, which falls within the larger risk-based scenario range. The large projected range reflects uncertainty about how glaciers and ice sheets will react to the warming ocean, the warming atmosphere and changing winds and currents. As seen in the observations, there are year-to-year variations in the trend.
Creator	Josh Willis, NASA Jet Propulsion Laboratory
Source	Chapter 2 in report of the third National Climate Assessment
Publisher	U.S. Global Change Research Program
Date	01/11/2013
References	Kemp, A. C. *et al.* Climate related sea-level variations over the past two millennia. *Proc. Natl Acad. Sci. USA* **108,** 11017–11022 (2012).
	Church, J. A. & White, N. J. Sea-level rise from the late 19th to the early 21st century. *Surveys in Geophysics* **32,** 585–602 (2011)
	Nerem, R. S., Chambers, D. P., Choe, C. & Mitchum, G. T. Estimating mean sea level change from the TOPEX and Jason altimeter missions. *Marine Geodesy* **33,** 435–446 (2010).
	Parris, A. *et al. Global sea Level Rise Scenarios for the United States National Climate Assessment* (NOAA, 2012).

The metadata element 'References' is taken from the DCMI Metadata Terms[7] and the other elements are taken from the Dublin Core Metadata Element Set[6].

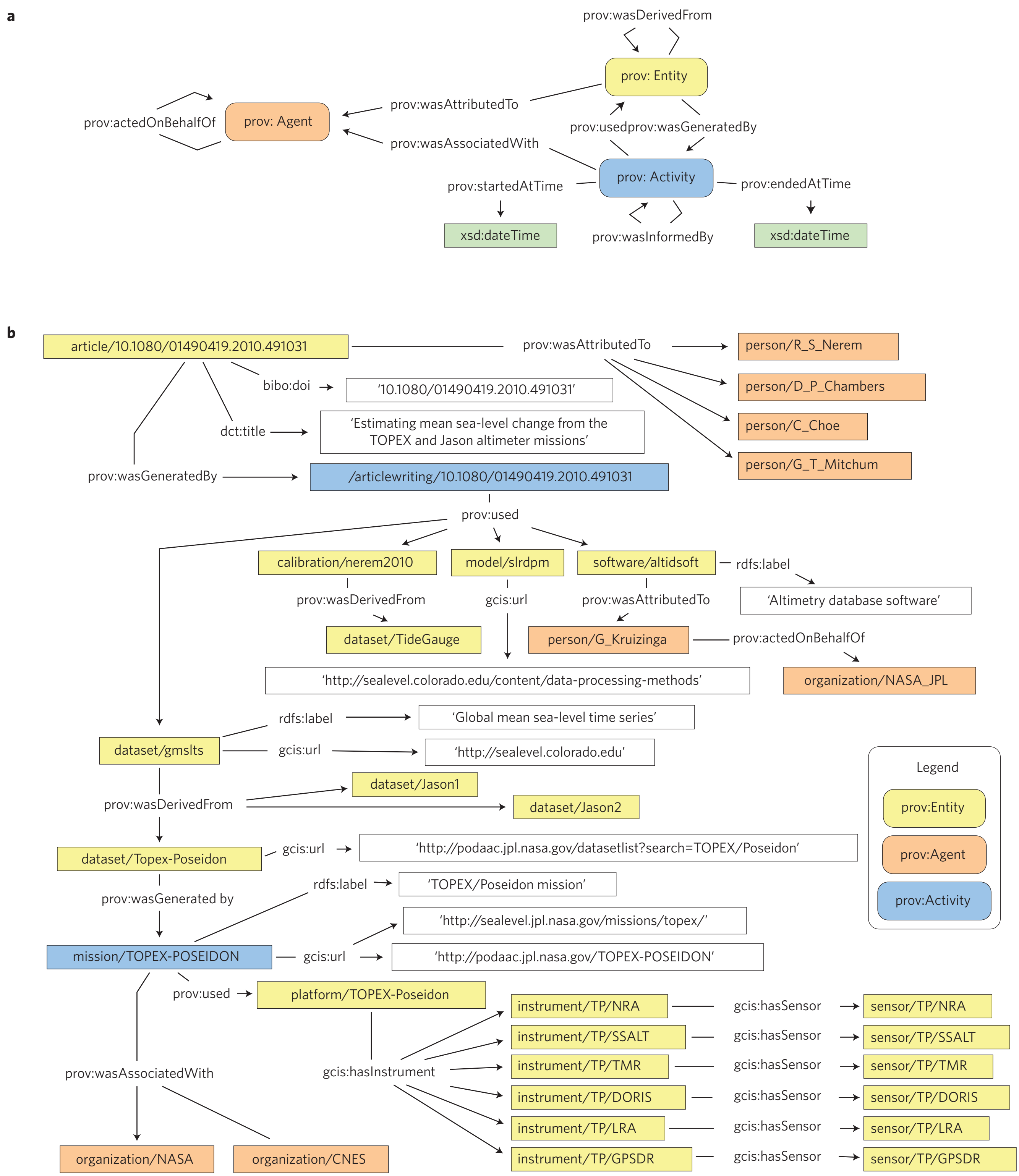

Figure 2 | A part of the provenance information of a journal article depicted with PROV-O and other ontologies. Classes/categories are represented by rounded boxes; instances by rectangular boxes. The first term of the record in each coloured rectangular box represents the class/category of the instance, and the colour of that box represents the corresponding super class/category in the PROV-O. A colourless rectangular box represents a literal record or a web address. The namespace prefix of a class/category or property represents its source ontology; details are listed in Table 1. **a,** Three starting point classes/categories in the PROV-O and the properties that relate them. **b,** A conceptual map of the provenance information. In **b** the record in each coloured rectangular box is a Uniform Resource Identifiers (URI) and the common root address 'http://data.globalchange.gov/' is omitted from all the records to save space. The provenance information of 'dataset/TOPEX-Poseidon' is shown in detail and the information of 'dataset/Jason1' and 'dataset/Jason2' is omitted.

individuals are working on projects related to the societal impacts of extreme weather events? The GCIS is intended to be a web-based source of authoritative, accessible, usable and timely information about global change for use by scientists, decision makers and the public. The Semantic Web technology will help make the GCIS a part of the Web of Data, such that other tools and services are also able to interact with data and information in the GCIS. This should enhance approaches applied to address socio-economic, physical, ecological and other intellectual challenges.

Persistent and universally resolvable identifiers, such as DOI (Digital Object Identifier), are widely accepted for research articles and increasingly also for data. ResearcherID and ORCID (Open Researcher & Contributor ID) make literature and data easily accessible and citable, and global change research increasingly benefits from open-access literature and data sets[17]. We argue that the global change research community should take one step further with the curation of provenance information — following the example of the GCIS. These works promote meaningful eScience[18] — the digital or electronic facilitation of science — and wider participation from the global change research community is desired.

Concluding thoughts

As global change information becomes both more abundant and increasingly important, our need to know more about what, how, when, where and why information is produced is becoming ever more necessary. Well-curated provenance information makes scientific workflows transparent and improves the credibility and trustworthiness of their outputs. It also facilitates informed and rational policy and decision-making based on the outputs of global change research. For all these reasons, work on provenance is timely and foundational, and is now an embedded component of the GCIS and a sustainable approach to climate assessment. ❐

Xiaogang Ma[1], Peter Fox[1], Curt Tilmes[2,3], Katharine Jacobs[4,5] and Anne Waple[6,7] are at the [1]Tetherless World Constellation, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180, USA, [2]NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA, [3]US Global Change Research Program, Suite 250, 1717 Pennsylvania Avenue, NW, Washington, DC 20006, USA, [4]Center for Climate Adaptation Science and Solutions, Institute of the Environment, University of Arizona, Tucson, Arizona 85721, USA, [5]Department of Soil, Water and Environmental Science, University of Arizona, Tucson, Arizona 85721, USA, [6]Second Nature Inc., 18 Tremont Street, Suite 930, Boston, Massachusetts 02108, USA, [7]Waple Research and Consulting, Weaverville, NC 28787, USA.*
**e-mail: max7@rpi.edu*

References

1. *US Code Global Change Research Act of 1990* (US Public Law 101–606, 1990).
2. IPCC *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) (Cambridge Univ. Press, 2013).
3. Parris, A. *et al. Global Sea Level Rise Scenarios for the United States National Climate Assessment.* (NOAA, 2012).
4. http://nca2014.globalchange.gov
5. Tilmes, C. *et al. IEEE Trans. Geosci. Remote Sensing* **51,** 5160–5168 (2013).
6. *The Dublin Core Metadata Element Set 15836* (ISO, 2003).
7. *DCMI Metadata Terms* (DCMI Usage Board, 2012).
8. *Geographic Information — Metadata* (ISO, 2003).
9. *Geographic Information — Metadata — Part 2: Extensions for Imagery and Gridded Data* (ISO, 2009).
10. Moreau, L. *et al. Future Gener. Comp. Sys.* **27,** 743–756 (2011).
11. Pinheiro da Silva, P., McGuinness, D. L. & Fikes, R. *Inform. Sys.* **31,** 381–395 (2006).
12. Moreau, L. & Missier, P. (eds) *PROV-DM: The PROV Data Model* (2013).
13. Hendler, J. *Science* **299,** 520–521 (2003).
14. *Resource Description Framework* (W3C, 2004).
15. *Global Change Information System Facts Sheet* (USGCRP, 2012).
16. Lebo, T., Sahoo, S. & McGuinness, D. L. (eds) *PROV-O: The PROV Ontology* (2013).
17. Overpeck, J. T., Meehl, G. A., Bony, S. & Easterling, D. R. *Science* **331,** 700–702 (2011).
18. Fox, P. & Hendler, J. in *The Fourth Paradigm: Data-Intensive Scientific Discovery* (eds Hey, T., Tansley, S. & Tolle, K.) 147–152 (Microsoft Res., 2009).
19. Melillo, J. M., Richmond, T. C. & Yohe, G. W. (eds) *Climate Change Impacts in the United States: The Third National Climate Assessment* (USGCRP, 2014).

Acknowledgements

This work was supported by a National Science Foundation grant through the University Corporation for Atmospheric Research to Rensselaer Polytechnic Institute under contract S13-94358. The authors thank Jin Guang Zheng, Justin Goldstein, Linyun Fu, Patrick West, Steven Aulenbach, Stephan Zednik, Ana Pinheiro-Privette and Brian Duggan for collaborations in the research project Global Change Information System: Information Model and Semantic Application Prototypes (GCIS-IMSAP). The authors also thank Tim Lebo and Deborah McGuinness for their comments on PROV-O, OPM and PML.

Author contributions

All authors contributed to the planning of the paper. X.M. led the work. P.F. contributed the use-case-driven Semantic Web application method which generated the provenance graph in the global sea-level rise use-case. C.T. participated in use-case analysis and provided suggestions on ontologies to be used. X.M. prepared the figures. All authors contributed to the writing of the paper.

COMMENTARY:

Arctic shipping and marine invaders

A. Whitman Miller and Gregory M. Ruiz

The emergence of new Arctic trade routes will probably change the global dynamics of invasive species, potentially affecting marine habitats and ecosystem functions, especially in coastal regions.

With striking reductions in Arctic sea-ice coverage in recent years[1,2], a long-anticipated opportunity for modern interocean shortcuts is being realized. The first commercial bulk carrier loaded with British Columbian coal successfully transited the Northwest Passage in September 2013[3]. Perhaps more importantly, ships in larger numbers are already navigating the icy waters of Norway and Russia through the Northeastern Passage, also known as the northern sea route (NSR) — a 3,000 mile passage along Russia's northern coast that connects the Barents and Bering seas. The Russian Federation's Northern Sea Route Administration, which issues permits, provides icebreaker escort and regulates