Abstract: Scientists are increasingly being called upon to publish data as well as conclusions. Computational science, in particular, often involves the creation of data products as a primary goal, rather than simply a means to an end. To have the same confidence in data that we have in peer-reviewed, appropriately-cited conclusions, we need mechanisms for capturing and conveying the origins and processing history -- the "provenance" -- of digital information. Geographic information presents particular challenges, both in the identification of locations and in the linking of information to locations. I describe recent work on automatic provenance capture and show how it might be applied in a distributed geographic context.

James Frew is an Associate Professor in the Donald Bren School of Environmental Science and Management at the University of California, Santa Barbara (UCSB), and a principal investigator in UCSB's Institute for Computational Earth System Science (ICESS). His research interests lie in the emerging field of environmental informatics, a synthesis of computer, information, and Earth sciences. He is interested in information architectures that improve the discoverability, usability, and reliability of distributed environmental information. Trained as a geographer, he has worked in remote sensing, image processing, software architecture, massive distributed data systems, and digital libraries. His current research is focused on geospatial information provenance, discovery, and curation, using remote sensing data products generated by his Environmental Information Laboratory as operational test beds.