

---

Faculty of Engineering

Faculty Publications

---

DigitalCrust – a 4D data system of material properties for transforming research on crustal fluid flow

Y. Fan, S. Richard, R.S. Bristol, S.E. Peters, S.E. Ingebritson, N. Moosdorf, A. Packman, T. Gleeson, I. Zaslavsky, S. Peckham, L. Murdoch, M. Fienen, M. Cardiff, D. Taboton, N. Jones, R. Hooper, J. Arrigo, D. Gochis, J. Olson, D. Wolock

February 2015

*The Wiley Hindawi Partnership*

*This journal is published by Hindawi as part of a publishing collaboration with John Wiley & Sons, Inc. It is a fully Open Access journal produced under the Hindawi and Wiley brands. <https://www.hindawi.com/journals/geofluids/>*

This article was originally published at:

<http://dx.doi.org/10.1111/gfl.12114>

---

Citation for this paper:

Fan, Y. et al. (2015). DigitalCrust – a 4D data system of material properties for transforming research on crustal fluid flow. *Geofluids* 15(1-2), 372-379.

# DigitalCrust – a 4D data system of material properties for transforming research on crustal fluid flow

Y. FAN<sup>1</sup>, S. RICHARD<sup>2</sup>, R. S. BRISTOL<sup>3</sup>, S. E. PETERS<sup>4</sup>, S. E. INGEBRITSEN<sup>5</sup>, N. MOOSDORF<sup>6,7</sup>, A. PACKMAN<sup>8</sup>, T. GLEESON<sup>9</sup>, I. ZASLAVSKY<sup>10</sup>, S. PECKHAM<sup>11</sup>, L. MURDOCH<sup>12</sup>, M. FIENEN<sup>13</sup>, M. CARDIFF<sup>4</sup>, D. TARBOTON<sup>14</sup>, N. JONES<sup>15</sup>, R. HOOPER<sup>16</sup>, J. ARRIGO<sup>16</sup>, D. GOCHIS<sup>17</sup>, J. OLSON<sup>14</sup> AND D. WOLOCK<sup>18</sup>

<sup>1</sup>Rutgers University, New Brunswick, NJ, USA; <sup>2</sup>Arizona Geological Survey, Tucson, AZ, USA; <sup>3</sup>US Geological Survey, Denver, CO, USA; <sup>4</sup>University of Wisconsin-Madison, Madison, WI, USA; <sup>5</sup>US Geological Survey, Menlo Park, CA, USA; <sup>6</sup>University of Hamburg, Hamburg, Germany; <sup>7</sup>Leibniz Center for Marine Tropical Ecology, Bremen, Germany; <sup>8</sup>North Western University, Evanston, IL, USA; <sup>9</sup>McGill University, Montréal, QC, Canada; <sup>10</sup>San Diego Supercomputer Center, San Diego, CA, USA; <sup>11</sup>University of Colorado, Boulder, CO, USA; <sup>12</sup>Clemson University, Clemson, SC, USA; <sup>13</sup>US Geological Survey, Middleton, WI, USA; <sup>14</sup>Utah State University, Logan, UT, USA; <sup>15</sup>Brigham Young University, Provo, UT, USA; <sup>16</sup>CUAHSI, Boston, MA, USA; <sup>17</sup>NCAR, Boulder, CO, USA; <sup>18</sup>US Geological Survey, Lawrence, KS, USA

## ABSTRACT

Fluid circulation in the Earth's crust plays an essential role in surface, near surface, and deep crustal processes. Flow pathways are driven by hydraulic gradients but controlled by material permeability, which varies over many orders of magnitude and changes over time. Although millions of measurements of crustal properties have been made, including geophysical imaging and borehole tests, this vast amount of data and information has not been integrated into a comprehensive knowledge system. A community data infrastructure is needed to improve data access, enable large-scale synthetic analyses, and support representations of the subsurface in Earth system models. Here, we describe the motivation, vision, challenges, and an action plan for a community-governed, four-dimensional data system of the Earth's crustal structure, composition, and material properties from the surface down to the brittle–ductile transition. Such a system must not only be sufficiently flexible to support inquiries in many different domains of Earth science, but it must also be focused on characterizing the physical crustal properties of permeability and porosity, which have not yet been synthesized at a large scale. The DigitalCrust is envisioned as an interactive virtual exploration laboratory where models can be calibrated with empirical data and alternative hypotheses can be tested at a range of spatial scales. It must also support a community process for compiling and harmonizing models into regional syntheses of crustal properties. Sustained peer review from multiple disciplines will allow constant refinement in the ability of the system to inform science questions and societal challenges and to function as a dynamic library of our knowledge of Earth's crust.

Key words: data integration, deep crustal dynamics, earth system models, groundwater, groundwater-surface water interaction, permeability

Received 3 February 2014; accepted 4 September 2014

Corresponding author: Ying Fan, Rutgers University, New Brunswick, NJ, USA.

Email: yingfan@eps.rutgers.edu. Tel: 848-445-3437. Fax: 732-445-3374.

*Geofluids* (2015) 15, 372–379

## MOTIVATION

Fluid flow in the Earth's crust depends strongly on material permeability, which varies in space and through time. As data and knowledge accumulate, and as we increasingly tackle interdisciplinary questions (Bodnar *et al.* 2013), a georeferenced, time-evolving data system of crustal

structure and properties is needed to address a wide range of scientific and societal questions.

## Understanding Earth's critical zone

The Earth's Critical Zone (CZ) is the region from the top of the terrestrial biosphere to the depth of active ground-

water circulation (NRC 2001). The CZ science effort focuses on understanding the physical, chemical, and biological processes regulating CZ evolution, determining its role in sustaining human society and terrestrial ecosystems, and predicting responses to anthropogenic, climatic, and tectonic forcing (Banwart *et al.* 2013). Fluid circulation plays a central role in CZ processes, regulating chemical weathering, soil formation, ecosystem evolution, and biogeochemical cycling (Berner & Berner 1996; Jones & Mulholland 2000; Brantley *et al.* 2011; Boano *et al.* 2014). Carbon cycle research has focused on the Earth's atmosphere and surface, but 99.9% of all carbon is stored in the lithosphere (Kempe 1979). Thus, even small changes in fluxes from the crust can have major consequences for the ocean-atmosphere system. Chemical weathering, a primary driver of global biogeochemical cycling, depends strongly on subsurface water residence times (Berner 1978; Maher & Chamberlain 2014), which is primarily controlled by 3D hydrological flow paths and material rock properties (McGuire *et al.* 2005). Weathering depth is unknown (West 2012) yet critical to understanding global biogeochemical fluxes. Existing predictions of material fluxes are based on 2D bedrock geological maps and therefore neglect deeper rock strata and geothermal waters (e.g., Becker *et al.* 2008). A major advance in overcoming these and many other limitations would be a 4D knowledge system for managing and synthesizing existing and newly acquired data on the Earth's crust.

### Assessing resource sustainability

Groundwater is the largest freshwater resource and primary source of drinking water for 2 billion people (Morris *et al.* 2003). It also plays a central role in agriculture (Foster & Chilton 2003; Giordano 2009) and sustains the health of many ecosystems (Alley *et al.* 2002). Nevertheless, groundwater is not adequately managed to ensure sustainability (Danielopol *et al.* 2003; Foster & Chilton 2003; Brunner & Kinzelbach 2005; Konikow & Kendy 2005; Fogg & LaBolle 2006; Gleeson *et al.* 2010; Sophocleous 2010), and nearly a quarter of humanity lives in areas of groundwater stress (Gleeson *et al.* 2012a). A key factor in sustainability is groundwater residence time related to the renewal rate which can be many millenia, well beyond the typical time horizon of human policies (Gleeson *et al.* 2012b). Residence time has been modeled assuming a consistent decrease in permeability with depth (Jiang *et al.* 2010), but a single low permeability layer can control groundwater age (Gassiat *et al.* 2013).

Fluid hydrocarbons in the upper crust also currently play a vital role in the energy budget for society. Knowledge of subsurface structures and properties is a prerequisite for addressing many of the energy issues surrounding energy resources, including harvesting of geothermal energy

(Mortensen & Axelsson 2013), carbon sequestration (Shrag 2007; Benson & Cole 2008), exploitation of unconventional oil/gas reservoirs, and fluid-injection-induced seismicity associated with all of these activities (Hitzman *et al.* 2012).

### Understanding deeper crustal dynamics

Hydrogeologists, geologists, and geophysicists have begun to actively explore the role of groundwater and other subsurface fluids in fundamental geologic processes, such as crustal heat transfer, ore deposition, hydrocarbon migration, seismicity, tectonic deformation, and diagenesis and metamorphism (e.g., Burns *et al.* 2015; Connolly & Podladchikov 2015; Howald *et al.* 2015; Micklethwaite *et al.* 2015; Miller 2015; Okada *et al.* 2015; Weis 2015). The permeability of the Earth's crust is of particular interest because it largely determines the feasibility of important physiochemical processes, such as advective solute/heat transport (Burns *et al.* 2015; Saffer 2015) and the generation of elevated fluid pressures by processes such as physical compaction, heating, mineral dehydration, and fluid injection (Connolly & Podladchikov 2015; Miller 2015; Weis 2015).

Current understanding supports a general distinction between the hydrodynamics of the brittle upper crust, where hydrostatic fluid pressures are the norm and meteoric fluids are common, and those of the ductile lower crust, where metamorphic reactions and internally derived fluids dominate hydrodynamic behavior. The brittle-ductile transition (BDT) between these regimes depends on temperature, strain rate, and rheology, but occurs at 10–15 km depth in typical continental crust. In tectonically active regions, high permeability episodically exists below the BDT (Connolly & Podladchikov 2015), such that fluid input from the ductile regime can be important to the cycling of some elements, and perhaps even to the balancing of the global water cycle over geologic time (Ingebritsen & Manning 2002).

This special issue of *Geofluids* highlights the historical dichotomy between the hydrogeologic concept of permeability as a static material property that exerts control on fluid flow and the perspective of other Earth scientists who have long recognized permeability as a dynamic parameter that changes in response to tectonism, devolatilization, and geochemical reactions. The dynamic view of crustal permeability is consistent with indications that fluid pressure is close to the lithostatic load during prograde metamorphism below the BDT (e.g., Fyfe *et al.* 1978); sufficiently overpressurized fluids cannot be contained in the crust, leading to fracturing and other processes that create permeability. More recently, it has been suggested that the permeability of the brittle crust may also be dynamically self-adjusting, responding to tectonism and external fluid

sources much as the deeper crust responds to the magnitude of internal fluid sources (cf. Cathles & Adams 2005; Rojstaczer *et al.* 2008; Weis *et al.* 2012). The temporal evolution of permeability can be abrupt or gradual: stream-flow responses to moderate to large earthquakes demonstrate that dynamic stresses can instantaneously change permeability by factors of up to 20 on a regional scale, whereas a tenfold decrease in the permeability of a package of shale in a compacting basin may require  $10^7$  years (Ingebritsen & Gleeson 2015). Thus, in the absence of seismicity, assuming that permeability is a static parameter can be reasonable for low-temperature hydrogeologic investigations with timescales of days to decades. Data compilations of deeper crustal material properties are likely to lead to a markedly better understanding of deeper crustal dynamics.

### Supporting earth system modeling

There is an urgent need for large-scale data synthesis to support the development of integrated Earth System Models (ESMs), which account for material and energy fluxes and key abiotic–biotic interactions in the atmosphere, lithosphere, and hydrosphere. ESMs are critical tools for predicting future global environmental change, such as that addressed by the Intergovernmental Panel on Climate Change (IPCC). However, even well-understood ground-water–surface water interactions in the top tens of meters of the crust are poorly represented in current ESMs, and most do not include subsurface processes at depths  $>2$ – $3$  m. Efforts to extend ESMs deeper into the crust have been

hindered by deficiencies in subsurface data. Global, realistic 3D gridded permeability and porosity fields for continental crust do not yet exist, but recent efforts to map near-surface permeability and porosity (Gleeson *et al.* 2014) provide an important starting point.

### DATA INTEGRATION TO TRANSFORM SCIENCE

Table 1 is a partial list of ongoing data integration efforts that have impacted our views of Earth systems interactions in many different ways. One example is the Macrostrat database (Peters 2006), which integrates existing stratigraphic information and aims to represent the Earth's upper crust as surface polygons that extend from the surface downward as stacks of lithostratigraphic and chronostratigraphic units. Macrostrat has integrated more than 36,000 rock units in North America, New Zealand, and the deep sea and is being augmented with the DeepDive machine reading system (Peters *et al.* 2014). Interactions between biotic and abiotic processes leave signatures in the rock record, and Macrostrat puts these signatures back into stratigraphic context, allowing them to be quantified in a space-time framework. Fossil records in the Paleobiology Database and the GPlates paleogeographic reconstructions are integrated with these data to produce a 4D model of the evolving Earth. Global-scale, deep-time syntheses of biological, geochemical, and sedimentary data have allowed new quantitative tests of long-standing hypotheses. For example, large-scale compilations of sedimentary data have

**Table 1** Examples of ongoing data integration efforts and the starting point of DigitalCrust.

Data	Source	Format
World Topography, Bathymetry	CSDMS: <a href="http://csdms.colorado.edu/wiki/Topography_data">http://csdms.colorado.edu/wiki/Topography_data</a>	Gridded
FAO World Harmonized Soil Map	IIASA: <a href="http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/">http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/</a>	Global gridded, polygons for countries
Global Lithologic Map	University of Hamburg: <a href="http://www.clisap.de/research/b:-climate-manifestations-and-impacts/crg-chemistry-of-natural-aqueous-solutions/global-lithological-map/">http://www.clisap.de/research/b:-climate-manifestations-and-impacts/crg-chemistry-of-natural-aqueous-solutions/global-lithological-map/</a>	Surface polygons
World Geologic Maps	USGS WMS and ESRI map services: <a href="http://energy.usgs.gov/OilGas/AssessmentsData/WorldPetroleumAssessment/WorldGeologicMaps.aspx">http://energy.usgs.gov/OilGas/AssessmentsData/WorldPetroleumAssessment/WorldGeologicMaps.aspx</a>	Surface polygons
World Tectonic Stress Map	GFZ Potsdam: <a href="http://dc-app3-14.gfz-potsdam.de/pub/introduction/introduction_frame.html">http://dc-app3-14.gfz-potsdam.de/pub/introduction/introduction_frame.html</a>	Gridded, with points and lines
Global Sediment Thickness	UCSD: <a href="http://igppweb.ucsd.edu/~gabi/sediment.html">http://igppweb.ucsd.edu/~gabi/sediment.html</a>	Gridded
Global Map of Surface Heat Flow	Map: Cardiff U: <a href="http://onlinelibrary.wiley.com/doi/10.1002/ggge.20271/abstract">http://onlinelibrary.wiley.com/doi/10.1002/ggge.20271/abstract</a> , Point: IHFC: <a href="http://www.heatflow.und.edu/">http://www.heatflow.und.edu/</a>	Gridded and points
National Geothermal Data System, SMU Geothermal Map:	<a href="http://geothermaldata.org/http://www.smu.edu/Dedman/Academics/Programs/GeothermalLab/DataMaps">http://geothermaldata.org/http://www.smu.edu/Dedman/Academics/Programs/GeothermalLab/DataMaps</a>	Gridded maps and points, thermal profiles, thermal conductivity
Continental Stratigraphy	University of Wisconsin: <a href="http://macrostrat.org/">http://macrostrat.org/</a> . Dataset of polygons tessellating North America, with associated time-stratigraphy description.	Polygons with vertical sequence of layers
Global Aquifer Maps	BGR and UNESCO: <a href="http://www.whymap.org/whymap/EN/Home/whymap_node.html">http://www.whymap.org/whymap/EN/Home/whymap_node.html</a>	Surface polygons
US Aeromagnetic Survey	USGS: <a href="http://www.usgs.gov/science/science.php?term=18">http://www.usgs.gov/science/science.php?term=18</a>	Gridded, resolutions vary region to region
US Gravity Anomaly	USGS: <a href="http://mrddata.usgs.gov/geophysics/gravity.html">http://mrddata.usgs.gov/geophysics/gravity.html</a>	Gridded
Groundwater Atlas, 25 US Aquifers	USGS: <a href="http://pubs.usgs.gov/ha/ha730/gwa.html">http://pubs.usgs.gov/ha/ha730/gwa.html</a>	Surface polygons with thickness (isopachs)
Global Permeability and Porosity	McGill University (GLHYMPS): <a href="http://onlinelibrary.wiley.com/doi/10.1002/2014GL059856/abstract">http://onlinelibrary.wiley.com/doi/10.1002/2014GL059856/abstract</a>	Surface polygons

played an important role in modeling biogeochemical cycling (e.g., Ronov 1978; Berner 2004), and Macrostrat has been used to calibrate sulfate burial fluxes and better constrain the role of the sulfur cycle in regulating atmospheric oxygen (Halevy *et al.* 2012; Canfield & Kump 2013). Spatial-temporal patterns of sedimentation in Macrostrat have also been shown to quantitatively reproduce many major features in the macroevolutionary history of marine animals (Peters 2005, 2008; Finnegan *et al.* 2011) and planktonic foraminifera (Peters *et al.* 2013). Combined with stable isotopic proxy records of biogeochemical cycling, global temperature, and rates of volcanism and crustal weathering, it appears likely that the correlations between paleobiological and macrostratigraphic data reflect common biological and stratigraphic responses to Earth system changes (e.g., Peters 2005; Hannisdal & Peters 2011), a hypothesis that emerges from, and can only be adequately tested with, integrated data deriving from the Earth's crust.

A second example is the UN-FAO Global Harmonized Soil Database. Large amounts of soil-survey data from multiple nations and continents, often built using different soil taxonomies, horizon definitions and attributes, and compiled at different scales of resolution and with different formats, were harmonized through an international partnership, which defined a new set of soil attributes critical to agriculture and recommended methodologies for developing taxo-transfer rules. The result was a global dataset at 30 arc-sec grids with 20 soil physical, chemical, and biological attributes. This dataset (and predecessors) has been the sole basis for deriving soil hydraulic parameters necessary for calculating soil water fluxes in all global land models and servers as the primary resource for constraining global soil organic carbon stocks and fluxes (e.g., Batjes 1996; Hiederer & Köchy 2011).

## THE DIGITALCRUST VISION

We envision a 4D space-time (xyz-t) data infrastructure designed to accommodate the structure and properties of the upper crust, from the surface down to the BDT, which occurs at 10–15 km depth in continental crust with a geothermal gradient of  $\sim 25\text{--}30^\circ\text{C km}^{-1}$  but can be as shallow as 4–5 km in regions of high heat flow. In regions with adequate seismic networks, the BDT can be crudely mapped on the basis of the distribution of earthquakes with depth (e.g., Nazareth & Hauksson 2004; Tanaka & Ishikawa 2005).

The DigitalCrust must be a web-oriented, data-service enabled, and spatially and temporally referenced workspace where the geosciences community can contribute and register data and model outputs, visualize, explore, and synthesize existing data to test hypotheses across space-time in ways that account for uncertainties. This is a daunting task

and will require support from the broader Earth science community, including from initiatives like EarthScope, national and regional geologic surveys, and funding agencies. Below we describe some of the key elements required in the DigitalCrust.

## A geologic scaffolding

The foundation of DigitalCrust is a geologic scaffolding that describes the basic geologic fabric of the Earth's upper crust, from the Critical Zone to the BDT, and includes data spanning its full range of physical, chemical, and biological properties (Fig. 1). To accomplish this, the DigitalCrust must receive contributions from all disciplinary domains involving the lithosphere, the hydrosphere, and the biosphere. Thus, despite the fact that it was originally motivated by the need to better understand and model crustal fluid flow, it must be an integrative data infrastructure that spans multiple domains of expertise in the Earth sciences. This broad vision is an attempt to both express the actual level of Earth systems integration that we believe occurs in nature and to respond to a common scientific and data infrastructure need that has been expressed in many Earth science communities. Because the most relevant intersection for many different types of geoscientists is defined by the common field location and rocks that they work on, regardless of whether or not they share any scientific expertise or disciplinary knowledge, the DigitalCrust stands to promote both data discovery and interdisciplinary cross-fertilization by proactively connecting scientists on the basis of their intersection in the Earth's crust.

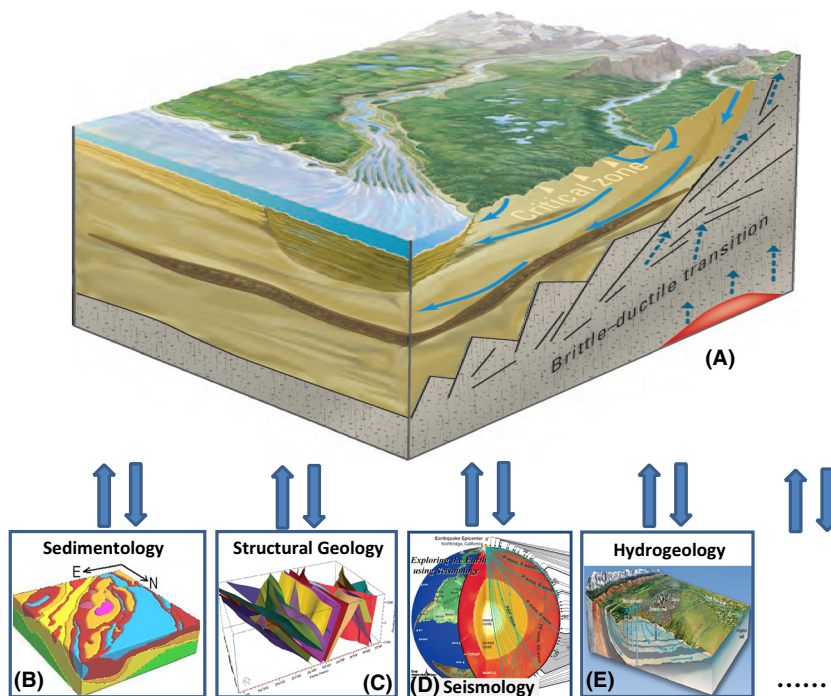
## Hydrogeologic properties as key data content and service

Within the foundational geologic scaffolding, the DigitalCrust will support multiscale integration of fluid-relevant properties. Improved description and synthesis of these properties, particularly permeability and porosity, has been a driving force behind the DigitalCrust. Although millions of soil and aquifer analyses and measurements have been made, the data are dispersed and unstructured in the scientific literature, government archives, and myriad online web pages and repositories. Scales, standards, and formats also vary. We face several major challenges, including discovering this vast amount of information and organizing it within the geologic scaffolding, and developing automated methods and algorithms for deriving meaningful hydrogeologic properties based on multiple data types.

## Community knowledge repository and management system

As a community knowledge repository, the DigitalCrust will integrate existing large-scale datasets (e.g., Table 1),





**Fig. 1.** The geologic scaffolding of the DigitalCrust from the Critical Zone to the Brittle-Ductile Transition (A), receiving contribution from and delivering service to a wide range of Earth science disciplines (B–E). Image source: (A) modified from Winter *et al.* (1998), (B) McInerney *et al.* (2005), (C) Hinz *et al.* (2012), (D) IRIS (<http://www.iris.edu/hq/>), and (E) Paschke *et al.* (2011).

and leverage current visualization tools to allow scientists to view what data already exist at given xyz-t coordinate and within a domain context, and what data/knowledge gaps remain to be filled. It will then allow scientists to contribute datasets to the growing knowledge base through a DigitalCrust node, with support for placing the data in an archival repository, obtaining an identifier for data, and releasing it for community use. Contributors can view how their new entries fit into or impact the framework and receive a response from the system with recommendations on related data that they may not be aware of, as well as recognition of their data/knowledge contribution.

As a knowledge management system, the DigitalCrust will index geoscience data sources from raw observation, through multiple levels of processing, interpretation, integration, and synthesis into models that are also incorporated into the repository. Linkage between observations and derived datasets through this chain should allow tracing provenance of information. The system should also include tools for social interactions such as review, discussion, correction, and updates to observations and interpretations at all levels. The resources in this system are accessed using simple web protocols and interchange formats that are documented, tested, and adopted by the DigitalCrust community. The data/information at a given geographic reference point will be delivered via an open application programming interface (API) that will support the development of specialized third party applications as well as the DigitalCrust online resource itself.

Central to the vision is the use of a branching and versioning system, such as ‘Git’ and ‘GitHub’ in software

development, which supports a common repository of best available data and most proven models, while allowing any researcher to create their own development fork. Formal peer review and community consensus will integrate branches back into the master DigitalCrust branch. Borrowing from the genomics community, which allows microcitation to unambiguously reference discrete data on organisms (Patrinos *et al.* 2012), the DigitalCrust will provide a capacity for citing and referencing data and data products.

Given the anticipated scope, the DigitalCrust must be governed by the community it intends to serve. It differs from many common crowd-sourcing models in that contributions will be attributed to specific members of the scientific community, allowing the community to regulate itself by, for example, trusting or not trusting the contributions based on individually demonstrated knowledge and expertise. A community governance model, to help sustain the integrity of the system as a whole, is being advanced as part of the NSF EarthCube initiative, which seeks to establish transformative cyber-infrastructure in support of the geosciences. Key features of organizational governance will likely involve standards for adoption and verification, organizational commitment through a membership process, impartial advisory boards, and other tested mechanisms to ensure system viability and sustainability.

#### Flexible information architecture

The DigitalCrust platform must provide a modular, configurable data storage and access component that is suffi-

ciently flexible to interface with existing databases and technologies, but also structured in such a way as to provide a useful synthetic resource. 'Standard' database designs have been developed for community use, but users inevitably find that there are missing entities and properties. The emergence of no-schema, document-type databases, such as CouchDb and MongoDB (e.g., Sadalage & Fowler 2012), provides technologies for hybrid fixed-schema and open-world information exchange models. The basic idea is to define a schema for common information items that are broadly shared, like geologic unit descriptions. Document-type databases allow unlimited addition of new properties to any entity as key-value pairs, or more complex multivalued data structures; thus, a standard schema can be readily extended by any individual or group. If the properties in new schema or schema extensions are mapped to properties in the existing information model, it then becomes possible to automate integration between data using the different schema. If new entities and properties emerge that many users find useful, they are documented and registered for consistent reuse and greater interoperability. This approach has been deployed in the National Geothermal Data System as a basis for information exchange using web services (Anderson *et al.* 2013) and in USGS ScienceBase as a method for continually expanding data capabilities with new access, analysis, and visualization parameters. The DigitalCrust will extend this concept, using content models as 'document templates' in a no-schema database that will provide the open-world flexibility and extensibility required by geoscientists, while also promoting standardization of commonly used entities and properties, such as lithostratigraphically defined local and regional rock units.

For geofluids applications, the DigitalCrust architecture needs to assimilate observational data and interpretations from all available sources, including geologic maps, cross-sections and structural contours, hydrogeologic unit delineations, soil tests, slug tests, aquifer pump tests, and indirect property estimates obtained through model inversions. The DigitalCrust architecture should be flexible such that researchers can upload any data and create products or models at any scales, choosing from a variety of automated methods, while supporting uncertainty propagation in derived products. Close disciplinary engagements are required to assure that data are used and interpreted properly in syntheses.

## AN ACTION PLAN

To ensure the success of DigitalCrust, we must reach out to the broader Earth sciences community, tapping common visions, synergistic efforts, and funding support to build the next generation of Earth science data infrastructure in a

distributed, loosely coupled architecture. The NSF EarthCube program, along with the USGS John Wesley Powell Center for Earth System Analysis and Synthesis, is poised to support these activities, bringing together Earth scientists and computer scientists to tackle some of the biggest data challenges. The first step to be taken is to use available collaborative mechanisms to engage additional disciplinary experts, data owners, and use case testers as we begin bringing together architectural and data components.

The second step is to set up the basic system architecture and integrate the existing community data systems, such as those listed in Table 1. This will allow us to demonstrate the concept immediately and expose data availability and gaps. Some simple visualization capabilities will be developed leveraging the development in other Earth science communities. This will prepare us to develop community-sourcing capabilities that allow data uploading, indexing, and editing, as well as a discussion forum for testing multiple interpretations or models.

The third step, of immediate interest to the geofluids community, is to develop and test the capabilities of the system to generate 3D gridded datasets of crustal permeability, porosity, and other relevant properties, integrating multiple data types, scales, and levels of uncertainty. Research is needed to define models, standards and rules of data harmonization, and a working group will be formed to help guide technical development of these standards. This will connect DigitalCrust with the science and society motivations discussed in Section 1 and facilitate the longer term process of building a coherent data system in support of crustal fluid investigations.

## CONCLUDING REMARKS

The current need for Earth system-level syntheses related to crustal fluid dynamics, the explosion of information on crustal structure and material properties, and rapid advances in computing and information science and technology have all converged to both enable and require the development of the DigitalCrust. It is a nontrivial task, one that requires transdiscipline, transcommunity, and transagency collaboration in a sustained effort. The NSF EarthCube program presents one opportunity to construct the DigitalCrust, primarily because both are aligned by their need to engage a much broader swath of the geoscience community than typically routinely collaborates. The potential utility of the DigitalCrust as a community resource for hydrogeologists to better understand fluid flow in the Earth system and its role in Earth's material and energy cycles at multiple scales and to more broadly reach the geoscience community, does, however, provide ample motivation.

## ACKNOWLEDGEMENTS

This project is supported by the joint NSF-USGS John Wesley Powell Center for Earth System Analysis and Synthesis working group and an NSF EarthCube Geo-Domain Community Workshop grant (EAR-1251557). We thank Jeanne DiLeo at the USGS for graphic support. The authors declare no conflict of interests. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## REFERENCES

- Alley WM, Healy RW, LaBaugh JW, Reilly TE (2002) Flow and storage in groundwater systems. *Science*, **296**, 1985–90.
- Anderson A, Blackwell D, Chickering C, Boyd T, Horne R, Mackenzie M, Moore J, Nickull D, Richard S, Shevenell L (2013) National Geothermal Data System (NGDS) Geothermal Data Domain: Assessment of Geothermal Community Data Needs. In *Proceedings of the Stanford Geothermal Workshop 2013: Stanford University, California, paper SGP-TR-198*, p. 14. <http://www.geothermal-energy.org/pdf/IGAstandard/SGW/2013/Anderson.pdf>.
- Banwart SA, Chorover J, Gaillardet J, Sparks D, White T, Anderson S, Aufdenkampe A, Bernasconi S, Brantley SL, Chadwick O, Dietrich WE, Duffy C, Goldhaber M, Lehnert K, Nikolaidis NP, Ragnarsdottir KV (2013) *Sustaining Earth's Critical Zone Basic Science and Interdisciplinary Solutions for Global Challenges*, p. 47. University of Sheffield, Sheffield.
- Batjes NH (1996) Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, **47**, 151–63.
- Becker JA, Bickle MJ, Galy A, Holland TJB (2008) Himalayan metamorphic CO<sub>2</sub> fluxes: quantitative constraints from hydrothermal springs. *Earth and Planetary Science Letters*, **265**, 616–29.
- Benson SM, Cole DR (2008) CO<sub>2</sub> sequestration in deep sedimentary formations. *Elements*, **4**, 325–31.
- Berner RA (1978) Rate control of mineral dissolution under Earth surface conditions. *American Journal of Science*, **278**, 1235–52.
- Berner RA (2004) *The Phanerozoic Carbon Cycle: CO<sub>2</sub> and O<sub>2</sub>*, p. 160. Oxford Univ. Press, New York, NY.
- Berner EK, Berner RA (1996) *Global Environment: Water, Air, and Geochemical Cycles*. Prentice-Hall, Upper-Saddle River, NJ.
- Boano F, Harvey JW, Marion A, Packman AI, Revelli R, Ridolfi L, Wörman A (2014) Hyporheic flow and transport processes: mechanisms, models, and biogeochemical implications. *Reviews of Geophysics*, accepted, doi: 10.1002/2012RG000417.
- Bodnar RJ, Azbej T, Becker SP, Cannatelli C, Fall A, Severs MJ (2013) Whole earth geohydrologic cycle, from the clouds to the core: the distribution of water in the dynamic Earth system. *The Geological Society of America Special Paper*, **500**, 431–61.
- Brantley SL, Megonigal JP, Scatena FN, Balogh-Brunstad Z, Barnes RT, Bruns MA, Van Cappellen P, Dontsova K, Hartnett HE, Hartshorn AS, Heimsath A, Herndon E, Jin L, Keller CK, Leake JR, McDowell WH, Meinzer FC, Mozdzer TJ, Petsch S, Pett-Ridge J, Pregitzer KS, Raymond PA, Riebe CS, Shumaker K, Sutton-Grier A, Walter R, Yoo K (2011) Twelve testable hypotheses on the geobiology of weathering. *Geobiology*, **9**, 140–65.
- Brunner P, Kinzelbach W (2005) Sustainable groundwater management. In: *Encyclopedia of Hydrological Sciences* (ed. Anderson MG), Part. 13. Wiley.
- Burns E, Williams CF, Ingebritsen SE, Voss CI, Spane FA, DeAngelo J (2015) Understanding heat and groundwater flow through continental flood basalt provinces: insights gained from alternative models of permeability/depth relationships for the Columbia Plateau, USA. *Geofluids*, **15**, 120–138.
- Canfield DE, Kump LR (2013) Carbon cycle makeover. *Science*, **339**, 533–4.
- Cathles LM, Adams JJ (2005) Fluid flow and petroleum and mineral resources in the upper (<20 km) continental crust. *Economic Geology*, **100th Anniversary Volume**, 77–110.
- Connolly JAD, Podladchikov YY (2015) An analytical solution for solitary porosity waves: implications for dynamic permeability and fluidization of nonlinear viscous and viscoplastic rock. *Geofluids*, **15**, 269–92.
- Danielopol DL, Griebler C, Gunatilaka A, Notenboom J (2003) Present state and future prospects for groundwater ecosystems. *Environmental Conservation*, **30**, 104–30.
- Finnegan S, Peters SE, Fischer WW (2011) Late Ordovician–Early Silurian selective extinction patterns in Laurentia and their relationship to climate change. In: *Ordovician of the World* (eds Gutiérrez-Marco JC, Rábano I, García-Bellido D), pp. 155–9. Cuadernos del Museo Geominera 14, Madrid, Spain.
- Fogg GE, LaBolle EM (2006) Motivation of synthesis, with an example on groundwater quality sustainability. *Water Resources Research*, **42**, W03S05.
- Foster SSD, Chilton PJ (2003) Groundwater: the processes and global significance of aquifer degradation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **358**, 1957–72.
- Fyfe WS, Price NJ, Thompson AB (1978) *Fluids in the Earth's Crust*. Elsevier Scientific, New York, NY.
- Gassiat C, Gleeson T, Luijendijk E (2013) The location of old groundwater in hydrogeologic basins and layered aquifer systems. *Geophysical Research Letters*, **40**, 3042–7.
- Giordano M (2009) Global groundwater? Issues and solutions. *Annual Review of Environment and Resources*, **34**, 153–78.
- Gleeson T, Van der Steen J, Sophocleous MA, Taniguchi M, Alley WM, Allen DM, Zhou Y (2010) Groundwater sustainability strategies. *Nature Geoscience*, **3**, 378–9.
- Gleeson T, Alley WM, Allen DM, Sophocleous MA, Zhou Y, Taniguchi M, Van der Steen J (2012a) Towards sustainable groundwater use: setting long-term goals, backcasting, and managing adaptively. *Ground Water*, **50**, 19–26.
- Gleeson T, Wada Y, Bierkens MFP, van Beek LPH (2012b) Water balance of global aquifers revealed by groundwater footprint. *Nature*, **488**, 197–200.
- Gleeson T, Moosdorf N, Hartmann J, van Beek LPH (2014) A glimpse beneath Earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophysical Research Letters*, **41**, 2014GL059856, doi: 10.1002/2014gl059856.
- Halevy I, Peters SE, Fischer WW (2012) Sulfate burial constraints on the Phanerozoic sulfur cycle. *Science*, **337**, 331–4.
- Hannisdal B, Peters SE (2011) Phanerozoic earth system evolution and marine biodiversity. *Science*, **334**, 1121–4.
- Hiederer R, Köchy M (2011) Global Soil Organic Carbon Estimates and the Harmonized World Soil Database. European Commission Joint Research Center Scientific and Technical Report, EUR 25225 EN. Publications Office of the European Union. [http://eusoils.jrc.ec.europa.eu/esdb\\_archive/eusoils\\_docs/other/EUR25225.pdf](http://eusoils.jrc.ec.europa.eu/esdb_archive/eusoils_docs/other/EUR25225.pdf).
- Hinz N, Siler D, Faulds J (2012) 3D Geologic mapping – structural studies of geothermal systems in the Basin and Range, Digital Mapping Techniques 2012 Proceedings. <http://ngmdb.usgs.gov/Info/dmt/DMT12presentations.html>.



- Hitzman MW, Clarke DD, Detournay E, Deiterich JH, Dillon DK, Green SJ, Habiger RM, McGuire RK, Mitchell JK, Shemeta JE, Smith JL (2012) *Induced Seismicity Potential in Energy Technologies*, p. 300. The National Academies Press, Washington, DC.
- Howald T, Person M, Campbell A, Lueth V, Hofstra A, Sweetkind D, Gable C, Banerjee A, Luijendijk E, Crossey L, Karlstrom K, Kelley S, Phillips F (2015) Evidence for long-time scale (>103 years) changes in hydrothermal activity induced by seismic events. *Geofluids*, **15**, 252–68.
- Ingebritsen SE, Gleeson T (2015) Introduction to the special issue on crustal permeability. *Geofluids*, **15**, 1–10.
- Ingebritsen SE, Manning CE (2002) Diffuse fluid flux through orogenic belts: implications for the world ocean. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 9113–6.
- Jiang X-W, Wan L, Cardenas MB, Ge S, Wang X-S (2010) Simultaneous rejuvenation and aging of groundwater in basins due to depth-decaying hydraulic conductivity and porosity. *Geophysical Research Letters*, **37**, L05403.
- Jones JB, Mulholland PJ (eds) (2000) *Streams and Ground Waters*. Academic Press, San Diego, CA.
- Kempe S (1979) Carbon in the rock cycle. In: *The Global Carbon Cycle* (eds Bolin B, Degens ET, Kempe S, Ketner P), pp. 343–75. Scientific Committee on Problems of the Environment (SCOPE), Old Woking.
- Konikow LF, Kendy E (2005) Groundwater depletion: a global problem. *Hydrogeology Journal*, **13**, 317–20.
- Maher K, Chamberlain CP (2014) Hydrologic regulation of chemical weathering and the geologic carbon cycle. *Science*, **343**, 1502–4.
- McGuire KJ, McDonnell JJ, Weiler M, Kendall C, McGlynn BL, Welker JM, Seibert J (2005) The role of topography on catchment-scale water residence time. *Water Resources Research*, **41**, W05002, doi: 10.1029/2004WR003657.
- McInerney P, Guillen A, Courrioux G, Calcagno P, Lees T (2005) Building 3D Geological Models Directly from the Data? A New Approach Applied to Broken Hill, Australia, USGS Open-File Report 2005-1428. <http://pubs.usgs.gov/of/2005/1428/mcinerney/>.
- Micklethwaite S, Ford A, Witt W, Sheldon H (2015) The where and how of faults, fluids, and permeability – insights from fault stepovers, scaling properties, and gold mineralization. *Geofluids*, **15**, 240–51.
- Miller S (2015) Modeling enhanced geothermal systems and the essential nature of large-scale changes in permeability at the onset of slip. *Geofluids*, **15**, 338–49.
- Morris BL, Lawrence ARL, Chilton PJC, Adams B, Calow RC, Klinck BA (2003) Groundwater and its Susceptibility to Degradation: A Global Assessment of the Problem and Options for Management, Nairobi, Kenya, United Nations Environment Programme (UNEP) Early Warning and Assessment Report Series, RS 03-3 126 p.
- Mortensen AK, Axelsson G (2013) Developing a Conceptual Model of a Geothermal System: presented at “Short Course on Conceptual Modelling of Geothermal Systems”, Santa Tecla, El Salvador, February 24 – March 2, 2013, accessed at <http://www.os.is/gogn/unu-gtp-sc/UNU-GTP-SC-16-29.pdf>.
- National Research Council (NRC) Committee on Basic Research Opportunities in the Earth Sciences (2001) *Basic Research Opportunities in the Earth Sciences*. National Academies Press, Washington, DC.
- Nazareth JJ, Hauksson E (2004) The seismogenic thickness of the southern California crust. *Bulletin of the Seismological Society of America*, **94**, 940–60.
- Okada T, Matsuzawa T, Umino N, Yoshida K, Hasegawa A, Takahashi H, Yamada T, Kosuga M, Takeda T, Kato A, Igarashi T, Obara K, Sakai S, Saiga A, Iidaka T, Iwasaki T, Hirata N, Tsumura N, Yamanaka Y, Terakawa T, Nakamichi H, Okuda T, Horikawa S, Katao H, Miura T, Kubo A, Matsushima T, Goto K, Miyamachi H (2015) Hypocenter migration and crustal seismic velocity distribution observed for the inland earthquake swarms induced by the 2011 Tohoku earthquake in NE Japan: implications for crustal fluid distribution and crustal permeability. *Geofluids*, **15**, 293–309.
- Paschke SS, Banta ER, Dupree JA, Capesius JP, Litke DW (2011) Groundwater Availability of the Denver Basin Aquifer System, Colorado, USGS Professional Paper 1770.
- Patrinos GP, Cooper DN, van Mulligen E, Gkantouna V, Tzimas G, Tatum Z, Schultes E, Roos M, Barend M (2012) Microattribution and nanopublication as means to incentivize the placement of human genome variation data into the public domain. *Human Mutation*, **33**, 1503–12.
- Peters SE (2005) Geologic constraints on the macroevolutionary history of marine animals. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 12326–31.
- Peters SE (2006) Macrostratigraphy of North America. *The Journal of Geology*, **114**, 391–412.
- Peters SE (2008) Environmental determinants of extinction selectivity in the fossil record. *Nature*, **454**, 626–9.
- Peters SE, Kelly DC, Fraass A (2013) Oceanographic controls on the diversity and extinction of planktonic foraminifera. *Nature*, **493**, 398–401.
- Peters SE, Zhang C, Livny M, Ré C (2014) A machine-compiled macroevolutionary history of Phanerozoic life. ArXiv Preprint:1406.2963.
- Rojstaczer SA, Ingebritsen SE, Hayba DO (2008) Permeability of continental crust influenced by internal and external forcing. *Geofluids*, **8**, 128–39.
- Ronov AB (1978) The Earth's sedimentary shell. *International Geology Review*, **24**, 1313–63.
- Sadalage PJ, Fowler M (2012) NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence: Addison-Wesley Professional, p. 192.
- Saffer D (2015) The permeability of active subduction plate boundary faults. *Geofluids*, **15**, 193–215.
- Shrag D (2007) Preparing to capture carbon. *Science*, **315**, 812–3.
- Sophocleous M (2010) Review: groundwater management practices, challenges, and innovations in the High Plains aquifer, USA—lessons and recommended actions. *Hydrogeology Journal*, **18**, 559–75.
- Tanaka A, Ishikawa Y (2005) Crustal thermal regime inferred from magnetic anomaly data and its relationship to seismogenic layer thickness. *Physics of the Earth and Planetary Interiors*, **152**, 257–66.
- Weis P (2015) The dynamic interplay between saline fluid flow and rock permeability in magmatic-hydrothermal systems. *Geofluids*, **15**, 350–71.
- Weis P, Driesner T, Heinrich CA (2012) Porphyry-copper ore shells form at stable pressure-temperature fronts within dynamic fluid plumes. *Science*, **338**, 1613–6.
- West AJ (2012) Thickness of the chemical weathering zone and implications for erosional and climatic drivers of weathering and for carbon-cycle feedbacks. *Geology*, **40**, 811–4.
- Winter TC, Harvey JW, Franke OL, Alley WA (1998) Ground water and surface water: a single resource. *U.S. Geological Survey Circular*, **1139**, 79.