Challenging and Changing the Way Science is Conducted

Two recent Nature editorials point to activities that challenge the ways that we do science and have the potential to both significantly improve trust in and expand the impact of science. The first editorial highlights efforts of a group of diverse, young (i.e., under 40), internationally renowned scientists that created a Code of Ethics for Researchers under the auspices of the World Economic Forum. The Code of Ethics is based on the group’s deliberations with researchers and ethicists and is designed to shape individual behavior as well as the processes of scientific institutions that can promulgate cultural change. It includes seven key principles:

1) Engage with the public
   “Engaging with the public means having an open two-way communication about science and the implications of research, as well as its need for society. Such communication involves active listening, discussion and questioning by both parties to enable the transfer of scientific knowledge into public wisdom.”

2) Pursue the truth
   “Pursuing the truth means following the research where it leads, rather than confirming an already formed opinion. This is particularly challenging but necessary when questioning current beliefs. The discovered truth must be confirmed and verified by peers, which requires transparency and reproducibility in all steps of the research and publication, in the methods used and by providing access to raw data. Results must be represented accurately without over- or understatement, hiding facts and/or drawbacks, or misleading the reader in any way. Findings must be based on evidence and observations, rather than on preconceived truths or biases. Pursuing the truth is more than creating knowledge as it also entails fighting untruths and valuing negative results in an ethical way.”

3) Minimize harm
   “Minimizing harm means that research inevitably carries some risk and, while it may be impossible to eliminate it, researchers can minimize harm to science, to others, to the environment, to society and to themselves. … Every researcher must consider each experiment’s potential to cause harm, not only from the perspective of what can occur during the experiment itself, but also – in rarer cases – of whether the generated knowledge can be detrimental to society.”

4) Engage with decision-makers
   “Engaging with decision-makers means going beyond developing solutions, conducting experiments and publishing data. … This engagement may be at any or all stages of the research process as needed. Reasons to engage are manifold, but ultimately the involvement of decision-makers greatly facilitates the probability that scientific outcomes will be translated into positive societal change.”

5) Support diversity
   “Supporting diversity means providing an environment in which the ideas of all are evaluated equally, regardless of individual characteristics, on the basis of evidence. Diversity is not simply the representation of individuals and ideas but is actual inclusion, which can only be achieved by creating a culture of openness, and recognizing and addressing unconscious bias. A diverse and inclusive scientific workforce draws from the widest range of backgrounds, perspectives and experiences to maximize innovation for the benefit of society. Achieving this representation may require seeking out participation from under-represented groups, while ensuring that the research process and its outcomes do not negatively affect particular groups.”

6) Be a mentor
   “Being a mentor means trusting and empowering less experienced researchers, especially during the early stages of their careers, to help them reach their professional goals and realize their full potential. It means creating an environment of trust and respect for all individuals in the scientific workplace.”

7) Be accountable
   “Being accountable means taking responsibility for one’s actions when carrying out research. This duty is paramount when scientific research is funded by public sources. Indeed, scientists have a moral but also financial responsibility to answer questions raised by society, a core funder of research. … They must merit the trust of society and students by behaving responsibly at all times. They also have a duty to secure this trust and hold each other accountable for research results by engaging the scientific community through peer review, or by holding diverse positions on boards and evaluation committees.”

The document is concise and engaging—explaining why each element matters, providing practical approaches to implementation, presenting relatable stories, and including pointers to other pertinent resources. I plan to use the Code of Ethics in future leadership training workshops that I provide for new faculty and post-doctoral researchers. I encourage readers to visit http://wef.ch/coe and provide feedback; you may also formally endorse the Code.
The second Nature editorial[1] summarizes steps that are being taken by the journal to increase research transparency [see #2, “Pursue the truth”, in the aforementioned Code of Ethics]. Specifically, two new versions of a checklist document have been created to guide authors of papers in the [1] behavioral, and social sciences (available now); and [2] ecology, evolution and environment (EEE) research (available soon). The documents prompt authors to provide critical details about study design, data collection and analysis prior to sending the papers to reviewers. The EEE checklist is being adopted from a comprehensive Tools for Transparency in Ecology and Evolution [TTEE] checklist[2] that was created by EEE journal editors and transparency advocates following a 2015 workshop; the document includes a basic checklist for empirical studies, a checklist for papers that are meta-analyses, and a checklist for referees/editors. It will be interesting to watch how other journals respond to the precedent set by Nature and Conservation Biology, which previously adopted a rigorous checklist for authors. Regardless, the checklists represent a comprehensive guide for authors, reviewers, and journals that are interested in promoting research transparency.

In these tumultuous times when well-accepted theories are deemed alt-facts by a vocal and ignorant few, it is important to promote and demonstrate strong ethical behavior and transparent science. The Code of Ethics and Tools for Transparency in Ecology and Evolution may contain principles or elements that fall outside the normal comfort zone for some scientists and institutions and I doubt that the documents will be wholly endorsed by all scientists and institutions and I doubt that the documents will be wholly endorsed some scientists and institutions and I doubt that the documents will be wholly endorsed ... would rectify any faults before exposing the system.

It is important to note that certificates have a finite lifetime and it is good practice to have in place a mechanism for reminding the operator when a certificate is expiring and providing several days lead time for ensuring a new certificate is available in a timely fashion. The LetsEncrypt Certificate Authority automates this process, which helps reduce ongoing maintenance tasks associated with operating a Member Node or any web server. A server certificate can be examined using a variety of tools, including the SSLabs test service above, a common web browser (e.g. in Google Chrome, click on the “Secure” notice to the left of the URL bar) or from the command line using a tool like openssl. For example, the following command would retrieve and check the certificate on search.dataone.org:

```bash
echo "Q" | openssl s_client -connect "search.dataone.org:443" | openssl x509 -text -noout
```

A slightly expanded version of the bash command is available for download[3].

### Client Certificates:

Client certificates are used in DataONE to identify an agent making a request to either a Member Node or a Coordinating Node. The agent could be an individual user making a request to a Member or Coordinating Node, a Member Node making a request to a Coordinating Node (e.g. to update the registered node capabilities), or a Coordinating Node making a request to a Member Node (e.g. during the synchronization process).

Client certificates used in DataONE are signed using the DataONE Certificate Authority or by CILogon[4]. Since certificates provided by CILogon are short-lived (18 hours), servers such as Member and Coordinating Nodes use a long-lived client certificate (3 years) signed by the DataONE Certificate Authority.

The client certificate for a Member Node is provided by DataONE administrators for either testing or production use. Note that a certificate issued for testing will not work in the production environment and vice versa. The client certificate is issued for a particular subject and is used to reliably ascertain by the server that the request cannot possibly be made by any other agent except exactly the agent identified by the subject in the certificate. For this reason, it is essential...
Building a Community of Scientific Data Repositories in an Open Science Landscape

Register now for the July 16th 2018 DataONE Users Group meeting, co-located with the Summer ESIP Federation Meeting at the Marriott University Park, Tucson, AZ. The DataONE Users Group (DUG) meeting will be a 1-day event featuring plenary presentations, topical breakout sessions and community-led discussions.

There is no registration fee to attend and participate in the DUG meeting
Information, registration and group hotel rates can be found at: bit.ly/DUG2018

Meeting Theme and Objectives
The 2018 meeting theme, “Building a Community of Scientific Data Repositories in an Open Science Landscape”, will bring together repository managers, users and other stakeholders to explore achievements and future work in the open science landscape. Community talks and posters that explore broad topics of interoperability, preservation, data discovery, reproducible research and sustainability are invited.

DataONE encourages DataONE Member Nodes, data scientists, researchers, scientists, students and others to submit abstracts for posters and talks.

Abstract Submission for Posters and Talks
Abstracts for talks and posters are solicited during the registration process. Talks will be approximately 10-20 minutes in duration, to be confirmed with development of the agenda. Submissions for talks will be accepted until June 10th, 2018. Oral presentations are not guaranteed. Those not accepted as oral presentations will be given the option to present a poster. Poster submissions will remain open until the close of registration.

Important dates
Oral Abstract Submission Deadline: June 10th 2018
Author Notification: June 20th 2018

For questions, contact dugchairs@dataone.org
See also bit.ly/DUG2018

DUG Steering Committee: Karl Benedict (co-chair), Robert Sandusky (co-chair), Amber Budden, Patricia Cruse, Devan Donaldson, Debora Drucker, Amy Forrester, Rebecca Koskela, Megan Mach, Laura Moyers, Shannon Rauch, Stephanie Simms, Heather Soyka, Richard Stephens, Gail Thornburg, Dave Vieglais.
that the private part of the certificate is kept secure.

A client certificate for a Member Node will be issued with the NodeId of the Member Node as the subject. For example, the client certificate for the ONEShare Member Node provides the following information:

```
$ openssl x509 -noout -text -in "urn:node:ONEShare-1.pem"
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            cf:87:45:f5:a1:6c:6c:fc
        Signature Algorithm: sha1WithRSAEncryption
        Issuer: DC=org, DC=dataone, CN=DataONE Production CA
        Validity
            Not Before: Jul 24 21:57:02 2015 GMT
            Not After : Jul 23 21:57:02 2018 GMT
(1)-->  Subject: DC=org, DC=dataone, CN=urn:node:ONEShare
        Subject Public Key Info:
        ... 
```

The line marked (1) shows the subject for the certificate. The DC elements indicate “Domain Components”, and the CN element indicates the “Common Name” of the certificate subject. Note that the value of the Common Name is the NodeId for the ONEShare Member Node. When this certificate along with the private component is specified as a client certificate in a request to a Coordinating Node, the Coordinating Node will examine the certificate subject, recognize that it was signed by DataONE (the Issuer entry), and therefore know that the agent making the request is the ONEShare Member Node.

Installation of the client certificate in a Member Node is dependent on the type of Member Node being deployed, and so the applicable Member Node documentation should be consulted to locate the client certificate for a particular installation. The Metacat documentation is available at the Metacat GitHub project and the GMN documentation is available at http://dataone-python.readthedocs.io/.

It is important to examine the certificate and ensure that the information contained therein matches expectations, and that the Member Node operator is aware of when the certificate expires. A new certificate should be requested at least two weeks in advance of expiration to ensure there is sufficient time to coordinate creation and installation of the new certificate with minimal service disruption. To obtain a new certificate for any reason, please contact support@dataone.org or visit us on Slack using the details below.

Contacts
Any questions about certificates, their use in DataONE, renewal for a Member Node or installation for use by a Member Node should be directed to support@dataone.org.

DataONE also provides an online discussion forum using Slack. To participate, please visit https://slack.dataone.org and request to join the DataONEorg.slack.com slack organization. Therein, general discussion topics appear in the default #dataone channel, and more technical topics are covered on the #ci channel.

Outreach UPDATE

The transition into summer marks the beginning of the annual DataONE Summer Internship Program. This year we are providing four research opportunities across cyberinfrastructure and community engagement projects, to support the current and future activities of DataONE. As always, we have been overwhelmed by the interest in our program and the quality of the applicants. Paperwork is being finalized, meetings are being coordinated and we look forward to introducing you to the 2018 cohort of summer interns shortly. As in previous years, projects will be openly documented at https://notebooks.dataone.org/.

We are also actively preparing for our summer season of conferences and training events. The Ecological Society of America meeting continues to be an important meeting for us and we will be representing DataONE at several sessions in addition to a co-organized workshop.
In each newsletter issue we will highlight one of our current Member Nodes. The full list of Member Nodes and summary metrics can be found on the DataONE.org site at bit.ly/D1CMNs.

The Organization for Tropical Studies NeoTropical Data Center
https://tropicalstudies.org/

The Organization for Tropical Studies NeoTropical Data Center (OTS-NDC) is a major generator of data on tropical and subtropical ecosystems. OTS-NDC joined the DataONE federation in April 2018 and by participating in the DataONE network, OTS-NDC now exposes more than 600,000 records of meteorological data to the broader DataONE community, contributing to a total of 38TB of data available through DataONE across 42 member nodes. OTS works to encourage researchers to share important resources, including long-term and new data sets. Over 300 scientists and graduate students from 25 countries have conducted research at the organization’s four research stations (three in Costa Rica and one in South Africa), resulting in over 5,000 scientific publications, books, and book chapters over the 50–year trajectory of the organization. OTS-NDC is supporting the development of some of the longest-running datasets in the tropics and we are delighted to support discovery of these data through DataONE.

Research at the Organization for Tropical Studies (OTS) stations has added significantly to the study of tropical ecosystems. The OTS-NDC is promoting public access to datasets that include seminal work from the last 50 years in tropical forest ecology and conservation, including related data on climate and systematics. The OTS-NDC accepts any type of research data generated at OTS Research Stations in Costa Rica, as well as related data from throughout the Neotropics. This includes ecological, biodiversity, evolutionary, meteorological, remotely sensed, and other related environmental data in text, tabular, raster, and vector formats. It also includes all data from the OTS biophysical monitoring system. As these datasets are transformed into scientific knowledge, they play a critical role informing important decisions that impact the future of tropical ecosystems.
DATA DISCOVERABLE THROUGH DATAONE

- **43 TB** of content
- **785 K** metadata
- **1.14 M** data

- **376,055** Unique downloads*  +12,825
- **2,236** Visitors to our search page*  +97
- **1,086** Searches conducted*  +59
- **99.99%** Uptime of Coordinating Nodes

*metrics are monthly averages; symbol denotes change since last quarter

OUR COMMUNITY

- **432** DataONE User Group members
- **5,000+** Users trained
- **41** Member Nodes

Education and Outreach

**Webinar Series**

- **30** Webinars
- **94** Average number of attendees
- **1599** Unique webinar attendees

**Education Resources**

- **18,964** Visits to the public webpage*  +304
- **214** Education Module downloads*  =

*metrics are monthly averages; symbol denotes change since last quarter

**Most Downloaded Resources**

1. Best Practices Primer
2. Data Management Plan Example from Manua Loa
3. Data Management Guide Public participation

**Most Visited Pages**

1. Best Practices: Create and document data backup policy
2. Data Management Planning
3. Best Practices: Define roles and responsibility

Metrics inclusive of April 2018