Journal of Community Engagement and Scholarship

Volume 14
Issue 3 *Engaged Scholarship: Sharing Views of Possible Futures for Extension and University-Wide Engagement*

Article 6

January 2022

Open Science, Open Data, and New Opportunities for Cooperative Extension

Catherine E. Woteki
Iowa State University

Follow this and additional works at: https://digitalcommons.northgeorgia.edu/jces

Recommended Citation

Woteki, Catherine E. (2022) "Open Science, Open Data, and New Opportunities for Cooperative Extension," *Journal of Community Engagement and Scholarship*: Vol. 14 : Iss. 3 , Article 6.
Available at: https://digitalcommons.northgeorgia.edu/jces/vol14/iss3/6

This Research From the Field is brought to you for free and open access by Nighthawks Open Institutional Repository. It has been accepted for inclusion in Journal of Community Engagement and Scholarship by an authorized editor of Nighthawks Open Institutional Repository.

Open Science, Open Data, and New Opportunities for Cooperative Extension

Catherine E. Woteki

Abstract

The array of problems presented to Extension professionals is broad and growing in number and complexity. As Extension has demonstrated its adaptability in addressing these issues, new ways of working have emerged. Access to an expanding pool of scientific reports and data can potentially provide Extension professionals with the greater tools and knowledge they need to collaboratively engage with their communities. However, a key challenge and possible impediment will be access to rapidly emerging research and data. Open science and open data can broaden the evidence base available to Extension educators, and the emerging field of data science offers new tools to help Extension stakeholders make data-informed decisions. A new data-sharing partnership among Canada, Mexico, and the United States may serve as a model for other countries' rural advisory services and national Extension systems. Fully implementing this expanded role for Extension will require resources to establish a National Community Learning Network and a national data commons as well as advocacy for open access policies at all levels of government. As abstract as open science and open data may seem to local and regional Extension practitioners, equal access to scientific knowledge and underlying research data is not only imperative for local community engagement but also integral for locally appropriate decision-making. Widening access to research and data directly supports the democratization of science and Extension. Opening scientific research and providing effective access to publicly financed data will become essential platforms for university engagement and Extension. It is critical for Extension professionals to understand the analytic powers and emerging policies that easily-accessed research and data can bring to collaborative community engagement.

Throughout its history, land-grant university (LGU) Extension has demonstrated flexibility and innovation in response to the needs of the communities in which it works. Today, the array of problems being presented to Extension professionals is broad and growing in number and complexity, and the tools at their disposal have also multiplied. University engagement with communities is critical to the process of democratically setting priorities for resource allocation and responding to what stakeholders see as their highest needs. Among the new tools are *open science*, *open data*, and *data science*, an interdisciplinary as well as transdisciplinary approach that extracts insights and knowledge from available data. Science-based information is a hallmark of Extension programs, and analyzing locally generated data is inherent to what Extension professionals do. However, the information age and the advent of the "internet of things" have increased the types of data available and require novel analytical approaches and expertise beyond what currently may be tapped.

New ways of developing programs and engaging communities are emerging that will enhance Extension's primary directive for community engagement. Extension continuously demonstrates adaptability in its responses to community issues and problems. University Extension services have nimbly adopted new tools and sciences as they fulfill their missions domestically and build partnerships internationally. Presently, opportunities exist to involve a broader range of disciplines, such as computer science, data science, and statistics, in Extension work and to build a truly university-wide, interdisciplinary, and transdisciplinary Extension model. In North America, opportunities also exist for U.S. LGU services to engage more deeply with Canada and Mexico in a way that builds on existing programs and on the three countries' present commitments to open data policies. Such an arrangement could serve as a model for other countries' agricultural, rural, and community advisory services.

Open Science

During the past decade, scientists' widespread adoption of the principles of open science and governments' embrace of policies that broaden access to their administrative and research

data promise to unlock new opportunities for LGU Extension to work collaboratively with communities to address their self-identified needs. A chief motivation for governments to adopt open data policies has been the reasoning that any research conducted or data collected with the support of public tax money should be made publicly available—with suitable protections in place, of course, to guard individuals' privacy, protect national security, and prevent the release of trade secrets and other intellectual property. Within the scientific community, open science is viewed as a way to make research more efficient. Its goals are to make knowledge more widely available, to enhance collaboration among participants in a way that that amplifies their collective intelligence and creativity, and ultimately to lead to entrepreneurial breakthroughs that benefit society (National Academies of Sciences, Engineering, and Medicine, 2018). In some countries, open data is seen as a way of promoting democracy and accountability in government. Open science also plays a role in democratizing scientific knowledge by enabling citizen science; it opens a vast literature to anyone who wants to read it and eventually may produce a better-informed citizenry (Arza & Fressoli, 2017). Open science may also bring greater capacity to solving the needs of society by making local problems more visible, giving community members access to data and resources, and deterring the private appropriation of those resources (Arza & Fressoli, 2017).

The drive to enable greater access to research and publicly financed data is not an abstract, 50,000-foot initiative of distant government policy-makers and academicians. It is absolutely essential to ensure the continued rapid adaptation of Extension and community programming success. In essence, the evidence base available to Extension professionals is dramatically expanded when research and government data are made public. But knowing how best to use and capitalize on this information requires skills and expertise that may be new to Extension professionals and that may require the help of specialists outside the current Extension faculty. Here is where a transdisciplinary, university-wide Extension model could facilitate transition to an extended role for Extension that builds on open science and open data.

Open science is characterized by the use of digital tools to share ideas, design and report experiments, disseminate results and supporting information, and preserve publications and data in archives that are open to all. In 2018, a committee convened under the auspices of the National Academies of Sciences, Engineering, and Medicine (2018) published a report describing "open science by design" and advocating that scientists build the six attributes of open science into their practice. The six attributes (listed in Table 1) begin when a scientist is exploring existing data and publications to conceptualize an idea and scope a research project. With today's online sources, much of

Table 1. Open Science by Design

Phases of Open Science by Design
Provocation: explore or mine open research resources and use open tools to network with colleagues
Ideation: develop and revise research plans and prepare to share research results and tools under FAIR principles (Findable–Accessible–Interoperable–Reusable)
Knowledge generation: collect data, conduct research using tools compatible with open sharing, and use automated workflow tools to ensure accessibility of research outputs
Validation: prepare data and tools for reproducibility and reuse and participate in replication studies
Dissemination: use appropriate licenses for sharing research outputs and report all results and support information (data, code, articles, etc.)
Preservation: deposit research outputs in FAIR archives and ensure long-term access to research results

this work is done through mining bibliographic databases and accessing existing open data sets. At each step of the research process—from developing research plans, collecting and validating data, and conducting analyses to publishing and archiving results and data—open science relies on the investigator's commitment to openness. The belief is that the research will benefit from transparency and enhanced rigor and that the results will have broader applicability and, therefore, greater impact. When incorporating the open science approach into their work, Extension professionals already have two assets in the U.S. Extension Foundation (which has become a good platform for Extension services to share and advance programs online) and the library science expertise located in campus libraries.

Those who favor open science argue that it offers many benefits (National Academies of Sciences, Engineering, and Medicine, 2018). Chief among them are enhanced rigor in the experimental designs submitted for review and comment and increased reliability of data collected using tools compatible with open sharing and automated workflow. Other benefits include faster and more inclusive dissemination of knowledge and broader participation in research. These upsides are part of the promise for Extension professionals and their collaborative cocreation of local programs. Laboratories and research teams that use digital tools to organize and record their work may use resources and perform research tasks more efficiently because such tools make their work more transparent to the entire team. And when science is conducted openly, there are public benefits as well. Open access to publications eliminates the price wall that prevents many people from reading research studies. Easy, open access to this information could foster innovation that would benefit everyone and could also help decision-makers at all levels of government make more informed and effective program and policy choices.

As attractive as open science may sound, several barriers are preventing its immediate adoption on a wide scale (National Academies of Sciences, Engineering, and Medicine, 2018). Research data sets become a resource in the open science ecosystem, and many unresolved questions about the life cycle of data, reproducibility, compliance, and sharing still need to be addressed. Adoption of digital workbooks and other technologies have costs and require new infrastructure. Many investigators are wary about privacy, security, and proprietary issues that may arise, and those worries may prevent them from committing to the full adoption of open science. Concerned investigators also point to a lack of incentives and training in their universities and disciplinary differences in how research is conducted as barriers to adopting all aspects of open science.

Open Science and Open Data Policies

Across Europe and North America and in other parts of the world, governments have adopted policies that require open access to scientific publications and their underlying data and that make administrative data collected by government programs publicly available. These seemingly abstract, distant policy actions are vastly extending the data available to Extension across its full range of programs and expertise. It is useful for Extension professionals and their community collaborators to understand the context of these national and international discussions. The origins of the international movement for open access to scholarly publications and underlying data can be traced to the Bethesda Statement on Open Access Publishing (Brown et al., 2003) and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (Max Planck Gesellschaft, 2003).

In April 2003, 24 people representing libraries, publishers, scientific societies, and research sponsors met in Maryland under the auspices of the Howard Hughes Medical Institute and issued a short statement defining open access publication and the commitments that each sector was making to move toward open access. Later that fall, more than 500 signatories to the Berlin Declaration wrote that their mission of disseminating scientific knowledge was not complete if "the information is not made widely and readily available to society," and they encouraged researchers to publish in accordance with the principles of open access, interoperability, and open access archiving (Max Planck Gesellschaft, 2003). More recently, the European scientific community has revised and updated its Science Europe Principles on Open Access to Research Publications (Science Europe Working Group on Open Access, 2015). Plan S, as it is called, was launched in September 2018 by the Open Access Envoy of the European Union and pledges that, beginning in 2021, scientific publications from public grants will be published in open access journals or platforms without embargo. Plan S is funded by Science Europe in

concert with 16 national science agencies, five charitable and international organizations, and the European Research Council.

In the intervening years in the United States, the National Institutes of Health began implementing aspects of open access. In 2013, President Barack Obama issued an executive order entitled "Making Open and Machine Readable the New Default for Government Information." The order articulated some general principles:

> Openness in government strengthens our democracy, promotes the delivery of efficient and effective services to the public, and contributes to economic growth. As one vital benefit of open government, making information resources easy to find, accessible, and usable can fuel entrepreneurship, innovation, and scientific discovery that improves Americans' lives and contributes significantly to job creation. (Exec. Order No. 13,642, 2013)

Obama's executive order and an accompanying directive to the federal science agencies from his science advisor, Dr. John Holdren, led to a government-wide effort to release vast amounts of scientific and administrative program information in digital formats. To help the public find the available information, all the agencies were directed to use a single portal: Data.gov. Today, over 190,000 different data sets from federal as well as state governments and municipalities are accessible through Data.gov.

Across the world, many countries have signed on to the principles of open government and open data. In North America, Canada and Mexico have adopted open data policies. Thirteen of the G20 countries have adopted open data policies. Open Data Charter, an international nongovernmental organization, advocates for governments at all levels to make data "open and freely available while protecting the rights of people and communities" (Open Data Charter, n.d.), and 79 national and local governments now endorse its principles.

Undergirding all of these efforts is the commitment to "FAIR principles." When data are made public, they must be Findable, Accessible, Interoperable, and Reusable (Wilkinson et al., 2016). In practice, FAIR means that data sets must have clear license, carry appropriate metadata, have persistent provenance, and be machine readable. As discussed earlier with respect to the barriers that university scientists encounter in implementing open science, commitment to implementing FAIR principles can be a high bar for governments, especially if their data are only available in spreadsheets and PDF files. Converting them to appropriately documented, machine-readable formats requires time, effort, and resources that may be beyond the capacity of local governments, even those in high-income countries and emerging economies.

New Opportunities for LGU Extension

Extension has many opportunities to build on these open platforms, and many Extension professionals are already experimenting with using easily accessible data to better serve the communities they work with. In the United States, enthusiasm is growing for forming a National Community Learning Network that would use data analytics and applied statistics techniques to help communities both access their own data and merge it with other local, state, and federal information to form more comprehensive evidence bases for decision-making. A new organization is now forming among the United States, Canada, and Mexico that aims to promote innovation, knowledge utilization, and information sharing in the agricultural sector across the continent. There are also opportunities to share innovations with and learn from the experiences of other countries through the Global Forum for Rural Advisory Services (GFRAS) and the Global Open Data for Agriculture and Nutrition (GODAN) initiative. We will discuss each of these organizations in turn.

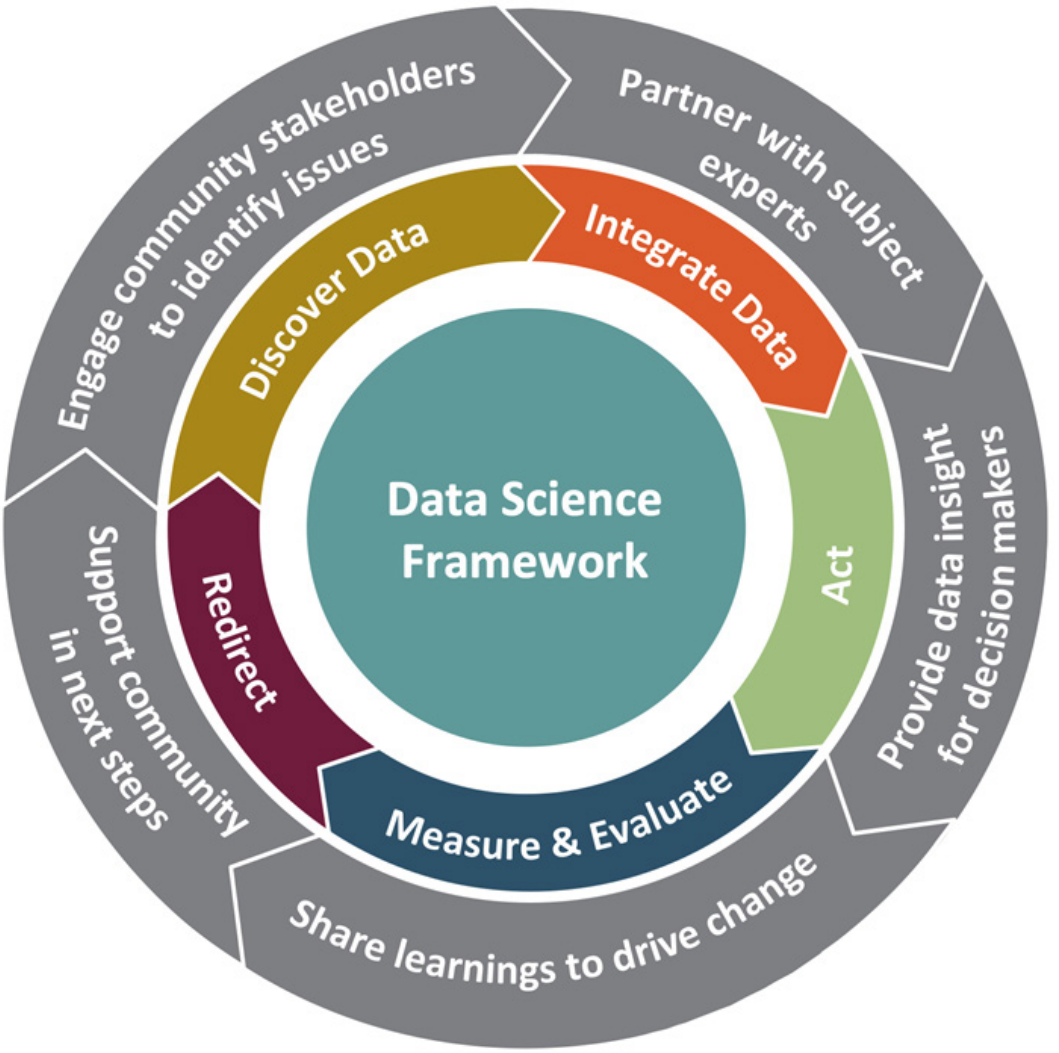
The idea for a National Community Learning Network (Keller et al., 2018) grew out of the experiences of Extension professionals in three states (Virginia, Iowa, and Oregon) working together through the Data Science for the Public Good (DSPG) program. Funded by a grant from the National Institute of Food and Agriculture and later supplemented with funding from the Bill & Melinda Gates Foundation, LGUs in these states pilot tested a model that first engages with communities to identify issues of concern and then works iteratively with them to seek insights from data analyses to inform future decisions. After a problem is identified, work begins to discover what data the community has. Then, relevant data from state and federal governments and other open data sources are assembled, and the data are cleaned up, analyzed, and curated to shed light on the identified problem. The DSPG program (University of Virginia Biocomplexity Institute,

n.d.-b) provides experiential learning for students and Extension professionals as they perform the data discovery, cleaning, and analyses to address the community's identified problem. The process (shown in Figure 1) derives from Cooperative Extension's long history of community engagement. The pilot test demonstrated that three public goods for communities, especially small and rural ones, could be scaled nationally:

1. Local communities were empowered to use their own data-rich and human-centered capacity for ongoing, community-directed problem-solving.
2. Extension professionals learned the skills and knowledge needed to effectively engage with communities in identifying and applying data-driven insights.
3. A data commons was developed that can be utilized to accelerate the advancement of community-based projects.

In the pilot work (Shipp et al., 2021), LGU Extension services in each state pioneered ways of working with community stakeholders to access and leverage their own data alongside other relevant data to inform local and state government programming and policy decisions. The pilot

Figure 1. The Community Learning Through Data Driven Discovery process.



Note. The outside wheel represents continuous communication and interaction throughout the process life cycle. The middle wheel is the data-driven learning process. The frontier between the outer and middle wheels connotes active collaboration among all stakeholders. The inner circle codifies the rigorous research framework to guide the data science.

Table 2. Examples of Case Studies Conducted by Extension, University Students and Faculty, and Community Stakeholders in Three States as Proof of Concept for Community Learning Through Data-Driven Discovery

Iowa	Oregon	Virginia
Expand Iowa State University Extension Community Helpline Services across the state	Regulatory challenges and impact on economic development in Eastern Oregon border region	Address barriers to health care access and use in Patrick County
Identify communities in greatest need of excessive alcohol–prevention efforts	Forecasting tools for cost analysis of water and wastewater facilities in small towns and cities statewide	Understand incarceration and recidivism in Halifax County
Pilot "systems of care" data infrastructure to inform a health information platform	Water quality requirements for fresh produce growers	Measure economic and social infrastructure: intergenerational poverty in Page County
	Create an economic mobility baseline for the South Wasco County area	Measure regional food insecurity and the role of a Loudon County food hub

test was conducted during the summer of 2020, when the COVID-19 pandemic had closed many organizations. Each state took a slightly different approach to engaging the community. The Oregon and Virginia teams built on their Extension models and had local stakeholders identify problems and data. In Virginia, for example, the Patrick County Cooperative Extension Office worked with the Virginia Department of Health and the Healthy Patrick County coalition to identify and prioritize the county's health challenges through a listening session and review of previously conducted health assessments and situation analyses. In Iowa, the state health department decided that the pandemic required changes to statewide programs and asked Extension to work with them on priority projects. The projects then became the focus of a 10-week program of learning and data analysis involving undergraduate students, graduate students, and faculty at the participating public universities. Table 2 lists some of the projects undertaken in each state as proof of concept for the approach. The range of projects was wide, spanning barriers to health care, regional food insecurity, the challenges of attracting and retaining workforce in a rural county, water quality requirements for fresh produce growers, and regulatory challenges and their impact on economic development. The nature of these problems required bringing expertise from across disciplines and across universities to inform the framing of the projects, the data discovery, and the interpretation. In addition to these projects, the three-state team also created the prototype for a national data commons (University of Virginia Biocomplexity Institute, n.d.-a), employing the community capitals framework for advancing economic mobility.

Discussions are underway to seek authorization in the next U.S. Farm Bill to allocate additional resources to the Extension system for (a) creating a community of practice for this newly expanded role and (b) developing a national data commons to capture and curate processes in support of data discovery, sharing, access, analytics, and evaluation for data-driven decision-making.

Recently, the United States, Canada, and Mexico joined together to form the North American Agricultural Advisory Network (NAAAN), creating another opportunity for learning between and among agricultural extension leaders, practitioners, users, and policy-makers. The NAAAN will be a discussion forum that expands on the relationships that already exist between and among the three countries'

agricultural extension programs. The NAAAN will also be the regional representative to GFRAS's global network and will link the North American Extension communities of practice to their counterparts across the world.

GFRAS is an international nongovernmental organization headquartered in Switzerland that provides a forum for agricultural advisory services to share learning and information. As the single global voice for agricultural Extension programs and service providers, GFRAS has regional networks that cover nearly all countries in the world. Each regional network works with agricultural Extension stakeholders at the country level (in many countries, these stakeholders have formed GFRAS country forums) to inform and improve the effectiveness of agricultural Extension programs at national and local levels.

Lastly, Extension can find other partners for its pioneering work with open data, data science, and new models for helping communities and rural governments make evidence-based program and policy decisions through GODAN. GODAN came into existence in 2015, with the United Kingdom and the United States as its initial supporters. The organization now has over 1160 partners that span governments, United Nations organizations, private sectors, and academic sectors who are committed to a statement of purpose and to collaborating across existing agriculture, nutrition, and open data activities to solve food- and agriculture-related problems.

What This Means for LGU Extension Services and Other Public University Engagement

In the short term, LGU Extension has opportunities for accessing open data and applying data science in all aspects of its work. There is enormous potential to expand the range and influence of Cooperative Extension in rural and urban settings where data-driven insights can help in making better decisions. To capitalize, Extension professionals must acquire new expertise in data analytics and engage more broadly with data science faculty experts on campus, regionally, nationally, and globally. Acquiring these skills and harnessing the potential of research and programmatic data will become essential for effective and collaborative community engagement. The new transborder relationships among Canada, Mexico, and the United States can build on the three countries' existing programs and open data commitments and serve as a model for other countries' advisory services.

Fully implementing this expanded role for Extension will require some advocacy for resources to build the National Community Learning Network and a national data commons as well as advocacy for open access policies at all levels of government. Hesitancies about embracing open science, fully engaging in transdisciplinary work, and developing international partnerships are all hurdles to be overcome. The upcoming reauthorization of research and education programs in the Farm Bill's Title 7 is an excellent place to start (United States Department of Agriculture, n.d.).

Open science and open data will become essential policy platforms for future Extension programming and for communities partnering with Extension organizations. This is equally true for all public university engagement. What may seem to be an abstract debate among distant policy "experts" is in fact of great importance for public university engagement, research, and teaching, as is moving Extension programming from safe and familiar disciplinary programming toward inter- and transdisciplinary solutions that address the complexities of local social, economic, and ecological conditions. Effective programming for climate change, pandemics, food systems, and the vast array of transdisciplinary problems will require equally involved local collaboration among communities and their university partners. Opening doors to knowledge that is presently hard to access and analyzing data to better inform local decisions and programming are important roles for Extension professionals and their university and local partners.

References

Arza, V., & Fressoli, M. (2017). Systematizing benefits of open science practices. *Information Services & Use*, *37*(4), 463–474. https://doi.org/10.3233/ISU-170861_

Brown, P.O., Cabell, D., Chakravarti, A., Cohen, B., Delamothe, T., Eisen, M., Grivell, L., Guédon, J.-C., Hawley, R.S., Johnson, R.K., Kirschner, M.W., Lipman, D., Lutzker, A.P., Marincola, E., Roberts, R.J., Rubin, G.M., Schloegel, R., Siegel, V., So, A.D., . . . Watson, L. (2003, June 20). *Bethesda statement on open access publishing*. http://legacy.earlham.edu/~peters/fos/bethesda.htm

Exec. Order No. 13,642, 3 C.F.R. 244 (2013). https://www.govinfo.gov/content/pkg/CFR-2014-title3-vol1/pdf/CFR-2014-title3-vol1-eo13642.pdf

Keller, S., Nusser, S., Shipp, S., & Woteki, C.E. (2018). Helping communities use data to make better decisions. *Issues in Science and Technology*, *34*(3), 83–89. https://issues.org/helping-communities-use-data-to-make-better-decisions

Max Planck Gesellschaft. (2003, October). *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. https://openaccess.mpg.de/Berlin-Declaration

National Academies of Sciences, Engineering, and Medicine. (2018). *Open science by design: realizing a vision for 21st century research*. The National Academies Press. https://doi.org/10.17226/25116

Open Data Charter (n.d.). *Who we are*. https://opendatacharter.net/who-we-are/

Science Europe Working Group on Open Access. (2015). *Science Europe principles on open access to research publications*. https://www.scienceeurope.org/media/4kxhtct2/se_poa_pos_statement_web_final_20150617.pdf

Shipp, S., Thurston, J., Lambur, M.T., & Dorius, C. (2021, February 23). *Evaluation – Towards a community learning network to advance economic mobility* (TR# 2021-021). University of Virginia Biocomplexity Institute. https://doi.org/10.18130/p4np-qc43

United States Department of Agriculture. (n.d.). *Farm bill*. https://www.usda.gov/farmbill

University of Virginia Biocomplexity Institute. (n.d.-a). *Community insights*. https://datascienceforthepublicgood.org/economic-mobility/community-insights

University of Virginia Biocomplexity Institute. (n.d.-b). *Social and decision analytics: Data science for the public good*. https://biocomplexity.virginia.edu/institute/divisions/social-and-decision-analytics/dspg

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., . . . Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, Article 1600018. https://doi.org/10.1038/sdata.2016.18

About the Author

Catherine E. Woteki is currently a professor at Iowa State University, where she previously was Dean of Agriculture, and Visiting Distinguished Institute Professor at the University of Virginia's Biocomplexity Institute. She served as Chief Scientist and Under Secretary for Research, Education, and Economics in the U.S. Department of Agriculture from 2010-2016, and as Under Secretary for Food Safety from 1997-2001.