

# Harnessing the Power of Data to Improve Agricultural Policy and Conservation Outcomes

Joshua D. Woodard, Bruce J. Sherrick, Jim Moseley, Collin O'Mara, Barry Gold, John Piotti, Sam Willett, Lynn Scarlett, Doria R. Gordon, Greg Fogel, William A. Salas, Carl Mattson, Kristin Weeks Duncanson, Robert Blair, Fred Yoder, and Deborah M. Atwood

*JEL Classifications: Q10, Q18, Q28, Q16*

*Keywords: Agricultural policy, Agricultural data, Agricultural technology, Government policy*

## Introduction

It is estimated that globally we will be collecting digital data at a rate of 40 ZB (40 trillion GB) per year by 2020 (Tien, 2013; Song et al., 2016). This increased availability of high-resolution biophysical, demographic, and economic data, coupled with dramatic increases in storage and computational capacities, is creating exciting new opportunities for data-driven scientific discovery. The term “big data” and its applications in agriculture have received particularly intense attention, and while no single standard has been developed, it is generally accepted that datasets that include high-frequency, high-resolution data with low error rates have been developing at accelerating paces; computational techniques designed to optimize the informational content extraction result in increased acceleration of the data-collection cycle (Coble et al., 2018). It is frequently noted that the population will soon approach 10 billion people and that changing standards of living will dramatically increase not only demand for caloric resources but continue to strain environmental and natural resource systems.

Farmers and food production systems are both threatened by and blamed for increased risks from climate change; given the realities of population location, trade, and international relationships, they will necessarily become even more critical elements in addressing the world's food needs. Overcoming these societal challenges and exploiting the opportunities at hand requires unprecedented levels of collaboration across scientific disciplines, including academia, industry, and government agencies. It is critical that researchers and policy makers be able to access the data needed to understand implications of changing production systems and make policy recommendations and explore new systems approaches. Remarkably, the data oftentimes exist but are collected in numerous disparate systems (such as on-farm vs. satellite) housed in multiple forms (survey, census, transactional records, market prices) or by different agencies (private firms, weather, commodity related, ERS, NRCS, FSA, etc.). In many cases, researchers simply need permission and a pathway to access and integrate disparate data from multiple systems.

Section 12618 of the Agriculture Improvement Act of 2018—otherwise known as the 2018 Farm Bill—provides a critical first step by directing the U.S. Department of Agriculture (USDA) to generate a report identifying available departmental data on conservation practices and the effect of such practices on farm and ranch profitability, including effects relating to crop yields and soil health. The original marker bill included in the original Senate-passed farm bill (S. 2487, the Agriculture Data Act of 2018, introduced by Senators Thune (R-SD) and Klobuchar (D-MN)) called for a full-fledged confidential data warehouse that would compile all the data across the various USDA subagency transactional, administrative, and survey databases into a single data system for analytical purposes, which academics would be allowed to access for research. In its final enacted form, the USDA will undertake a study and generate a report to summarize the process that the Secretary of Agriculture will use to provide access to a secure data warehouse containing these previously distinct data systems. The data will be made available to university researchers under privacy and access provisions to safeguard producer-specific information while

maximizing research and policy guidance benefits. Importantly, it directs the USDA to recommend to Congress any legislation needed to create this facility with the intent to unlock opportunities for data-driven scientific discovery supporting increased sustainable production.

Among the most immediate and critical applications, there is strong interest in linking conservation and crop insurance data to better incentivize and coordinate production and conservation goals while preserving accuracy in ratings and improving soil health outcomes. Broadly linked historic data, along with newly linked high-resolution and high-frequency data, offers the chance to improve the efficacy of soil health initiatives, better identify areas and activities to improve water quality and conservation initiatives, improve nutrient effluence management, adapt fertility and pesticide placement and usage, and better understand in-field zone management to enhance productivity (Woodard, 2016b,c; Garrett, 2013; Gustafson et al., 2016). These activities ultimately lead to greater yields and/or yield stability, which yield other positive externalities (Phalan et al., 2016). For example, with better data analytics and use of existing data resources, the gap between pollution mitigation and economic growth can be better identified, understood, and ameliorated (Obama, 2017).

## Agricultural Analytics and Policy

The United States plants over 215 million acres of corn, soybeans, and wheat annually, more than any other country in the world (U.S. Department of Agriculture, 2019). Over the past two decades, it has become routine to associate geolocated records with field activities as a consequence of the broad adoption of tractor and combine RTK (real-time-kinematic) systems that allow auto-steer, input location and quantity tracking (fertilizer, pesticide, seed, tillage), condition tracking (moisture, compaction, soil characteristic), and yield monitoring. Rapid advancements in other georeferenced technologies—along with imagery, soils mapping, and weather data—offer unprecedented opportunities to use data to enable smarter production systems and habitat protection. Strongly held concerns about data privacy have appropriately inspired several careful conversations about best approaches to data integration, many of which are informed by other leading efforts from financial and medical system considerations, which have generally led on issues of privacy and data protection. Section 12618 of 2018 Farm Bill would ensure that the creation of any data facility would protect private data while creating a unified data access channel.

## Data Management Best Practices

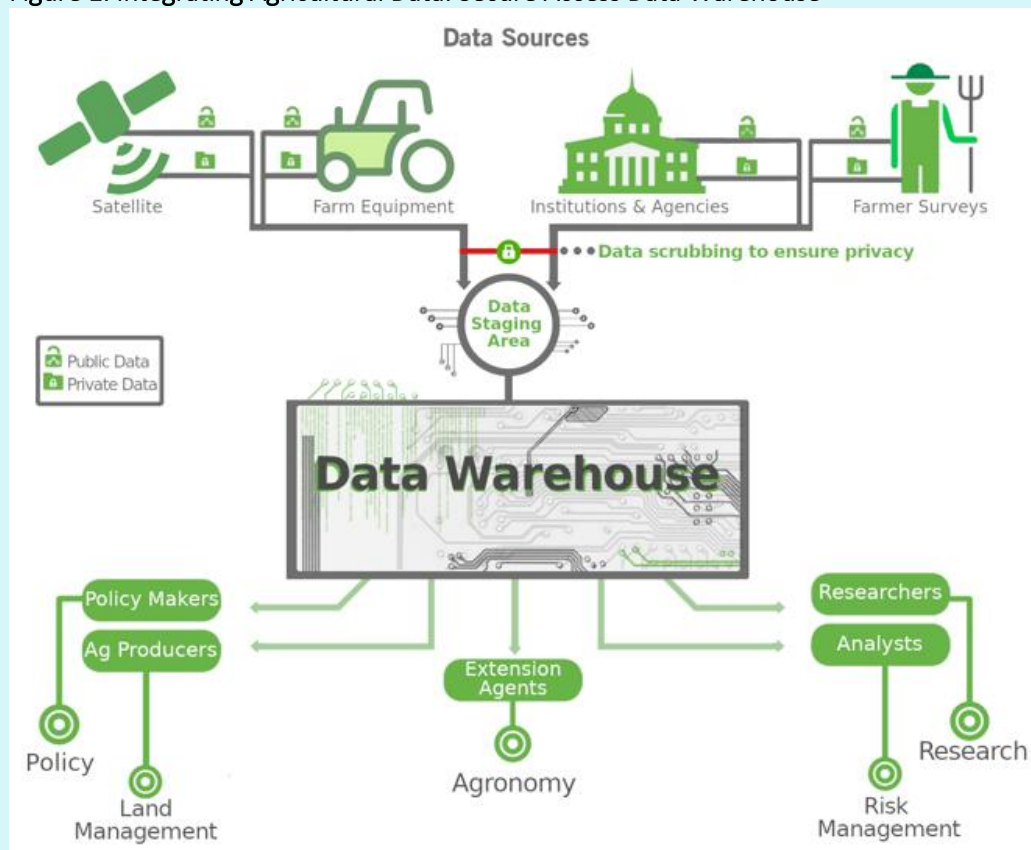
Issues related to data sharing and reproducibility are generally well recognized (McNutt et al., 2016). Despite policies regarding open data and frameworks for federal agencies to make their data public (Office of Science and Technology Policy, 2015), recent recommendations by the Commission on Evidence-Based Policy making (CEP, 2017), and other related efforts (NSF, 2015), the reality is that public data sources tend to be decentralized, often survey-based and summarized for privacy, offered in difficult-to-query forms, or simply not made available at all (Woodard, 2016a,c). As a result, the integration of public data sources with farm- or transaction-level databases remains severely limited. Viable funding models must also be associated with proposed policies (Berman and Cerf, 2013), but the reality is that such investments are likely to pay far greater dividends to society than the technical cost of construction. While many corporations (such as equipment vendors) maintain internal transaction databases, these data are seldom linked with other widely available data (such as high-resolution soil, weather, or spatial price data) and made available to policy researchers for broader learning. Likewise, many government agencies maintain comprehensive records on specific programs they administer—largely for administrative requirements.

Taken together, these factors greatly hamper researchers' ability to conduct big-data-oriented research targeting agriculture, the environment, nutrition, and food supply chains. As a result, many potential opportunities to improve production and profitability while improving soil health and environmental conditions may be missed. Developing large-scale, high-resolution, integrated data resources provides an important part of the solution to many of the problems that lie at the intersection between environmental systems and production activities.

## Data Solutions for Better Farm and Conservation Outcomes

While the idea of integrating producer data with broad-scale imagery data, soils data, topological data, weather and hydrological data, and governmental survey and program data seems at one level obvious (see Figure 1), in reality it represents a daunting practical challenge requiring many decisions and facets for design, including:

Figure 1. Integrating Agricultural Data: Secure Access Data Warehouse



Notes: Agricultural and environmental data come from various public and private sources. Data are extracted and then scrubbed and anonymized to ensure privacy. The data can be integrated, stored in a data warehouse designed by data scientists, and made available for use by policy makers, agricultural producers, extension agents, researchers, and analysts.

- **Metadata Standards:** There is a need to develop better metadata standards in agricultural data systems, particularly when linking to precision agriculture. There is currently no agreed-upon standard for the universe of different types of agricultural data, and existing industrial machine data standards are sometimes proprietary by brand or function and are not well-suited for the types of data collected by USDA and related agencies.
- **Legislation providing Authorization and Protections:** Section 12618 of 2018 Farm Bill directs the Secretary of Agriculture identify the steps needed for secure, confidential access to conservation and farm productivity data and holds promise for massive improvements in efficiency. Promulgating these standards will require time and co-operation among a large number of different entities with different starting points.

- Governance, Collaboration, and Funding among federal, state, and private sector institutions that produce data, fund science, and perform research: For example, the commercial incentive to process satellite imagery data into sets that could help augment or alleviate survey costs by the National Agricultural Statistics Service (NASS) and other agencies will require different models of funding.
- Open data and Secure access data: Despite headline calls for “open data” in a variety of realms (including medicine), there is growing recognition that not all contexts adapt well to the open-data paradigm; rather, open secure access protocols are more appropriate in cases where privacy is important (Woodard, 2016c; Garrett, 2013; CEP, 2017), and that mindfulness of who benefits should be considered (Serwadda et al., 2018).

## Applying Data Solutions to Agricultural Policy in the United States

In the United States, agricultural production and food handling systems are highly regulated, inspected, and also supported through varied channels. The crop insurance and conservation titles in the 2018 Farm Bill are very large and represent a substantial share of the overall direct expenditures outside of the nutrition assistance programs. The U.S. crop insurance program insures a vast number of crops with nearly \$10 billion in premiums and about \$100 billion in annual liabilities. Government pricing of the crop insurance is fairly coarse but designed to encourage broad participation and to be run in an actuarially fair manner over time. Currently, premiums are set by region and by production unit but do not consider specific field-level information—including soils, certain tillage and conservation practices, or many other production-level characteristics that can change across time or fields, or time within an insured unit. (Woodard, 2016c). Given today’s mapping and activity-tracking technologies, it would be straightforward to include soil characteristics and other features in premium ratings by integrating available data into a pricing model if data from separate systems were available. The accuracy improvements this simple idea could accommodate could have dramatic consequences both in terms of actuarial efficiency and in creating appropriate risk classifications. Moreover, the impacts from different practices over time (such as conversion in tillage systems, use of rotations and cover crops, etc.) would not be incidentally tied to data from different regimes in rating.

From this kickoff point, incorporating other data (such as historical practice data) that impact soil health would be straightforward. On the flip side, poorly priced insurance can lead to increased taxpayer costs and potentially even adverse incentives to produce on environmentally sensitive land (Woodard et al., 2012) and even affect producers’ choice of practices and what they choose to plant. The primary goals of crop insurance have historically been to maximize participation and minimize retroactive provisional disaster payments, which are likely to be far more expensive and less targeted. While the crop insurance program has been highly successful in reducing financial stress and leading to improved investment, program modifications could potentially promote and improve soil health and environmental outcomes while reducing errors in pricing individual coverage. This idea has been compared to modern car insurance policies, which allow drivers to track their driving habits and generate a specific policy quote tailored to the actual risks demonstrated rather than to a historic pool of cars and drivers with similar characteristics.

Accomplishing the goal of providing both effective crop insurance and conservation incentives for producers across a diversity of production systems is admittedly very complex. Fine tuning these massive programs requires data integration and smart design. The ag-data inclusions in the Agriculture Improvement Act are an important first step that will give qualified researchers at land-grant universities—whom farmers know and appropriately trust—the ability to analyze micro-data at scale and connect the dots between conservation practices, crop yields, soil health, and risk. Through this type of research, new policy designs can be pursued and achieved.

For a concrete example, the 508(h) process allows large-scale, research-based modifications to the Federal Crop Insurance program to be proposed. Oftentimes, the data that would be needed to structure or design a specific improvement may well exist in government databases but cannot be accessed by researchers, creating a data access–research catch 22. In addition to securing cost savings for America’s farmers, improved ratings and insurance designs will strengthen the crop insurance program and ensure that it remains defensible. Establishing a confidential, secure conservation and farm productivity data warehouse will enable more innovative research and analysis that advances meaningful policy and improved production practices. These recommendations are closely aligned with the USDA’s broader reorganization efforts and will enable the USDA’s Natural Resources Conservation

Service (NRCS), Risk Management Agency (RMA), and other agencies to analyze agricultural data with increased efficiency and precision. In all cases, producer privacy must remain a priority. The integrity and confidentiality of producer data can be protected by upholding strict privacy protocols already in place at the USDA and as strengthened by recommendations required to be considered by the provisions of the act.

## Conclusions

There is a well-recognized need to establish and quantify risk and conservation practices at scale and incorporate them into policy, insurance, lending, and conservation programs. Precision agriculture, government policy, and risk management markets have large roles to play in both intensification and conservation in agriculture. Improved agricultural productivity and improved sustainability are not mutually exclusive, but the design and execution of agricultural policy has a profound role in mediating the conservation versus intensification dynamic to benefit both farmers and the environment. Why would we leave our best tools to develop this policy in a toolbox we never open? Section 12618 and related provisions in the Agriculture Improvement Act of 2018 provide an important first step toward making these very valuable data available to the research community. We believe the data facility that will hopefully eventually result from this legislation and set of efforts will have a profound impact on academic and policy research in a wide range of scientific fields related to agriculture and the environment.

## For More Information

Agriculture Improvement Act of 2018. Pub. L. No. 115-334, 132 Stat. 4490 (2018).

Berman, F., and V. Cerf. 2013. "Who Will Pay for Public Access to Research Data?" *Science* 341(6146): 616–617.

Coble, K.H., A.K. Mishra, S. Ferrell, and T. Griffin. 2018. "Big Data in Agriculture: A Challenge for the Future." *Applied Economic Perspectives and Policy* 40(1): 79–96.

Commission on Evidence-Based Policymaking (CEP). 2017. *The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking*. Available online at <https://www.cep.gov/report/cep-final-report.pdf>

Garrett, K.A. 2013. "Big Data Insights into Pest Spread." *Nature Climate Change* 3: 955–957.

Gustafson, D., M. Hayes, E. Janssen, D.B. Lobell, S. Long, G.C. Nelson, H.B. Pakrasi, P. Raven, P.G. Robertson, R. Robertson, and D. Wuebbles. 2016. "Pharaoh's Dream Revisited: An Integrated US Midwest Field Research Network for Climate Adaptation." *Bioscience* 66(1): 80–85.

McNutt, M., K. Lehnert, B. Hanson, B.A. Nosek, A.M. Ellison, J.L. King. 2016. "Liberating Field Science Samples and Data." *Science* 351(6277): 1024–1026.

National Science Foundation (NSF). 2015. *Today's Data, Tomorrow's Discoveries: Increasing Access to the Results of Research Funded by the National Science Foundation*. NSF 15-52. Available online at <https://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf>

Obama, B. 2017. "The Irreversible Momentum of Clean Energy Private-Sector Efforts Help Drive Decoupling of Emissions and Economic Growth." *Science* 355(6321): 126–129.

Office of Science and Technology Policy. 2015. *Project Open Data*. Available online at <https://project-open-data.cio.gov/>

Phalan, B., R.E. Green, L.V. Dicks, G. Dotta, C. Feniuk, A. Lamb, B.B.N. Strassburg, D.R. Williams, E.K.H.J. z. Ermgassen, and A. Balmford. 2016. "How Can Higher-Yield Farming Help to Spare Nature?" *Science* 351(6272): 450–451.

- Serwadda, D., P. Ndebele, K.M. Grabowski, F. Bajunirwe, and R.K. Wanyenze. 2018. "Open Data Sharing and the Global South—Who Benefits?" *Science* 359(6376): 642–643.
- Song, M.L., R. Fisher, J.L. Wang, and L.B. Cui. 2016. "Environmental Performance Evaluation with Big Data: Theories and Methods." *Annals of Operations Research* 270(1–2): 459–472.
- Tien, J.M. 2013. "Big Data: Unleashing information." *Journal of Systems Science and Systems Engineering* 22(2):127–151.
- U.S. Department of Agriculture. 2019. *World Agricultural Production*. Circular Series WAP 9-19. Washington, DC: U.S. Department of Agriculture, Foreign Agricultural Service, September.
- Woodard, J.D. 2016a. "Big Data and Ag-Analytics: An Open Source, Open Data Platform for Agricultural & Environmental Finance, Insurance, and Risk." *Agricultural Finance Review* 76(1): 15–26.
- Woodard, J.D.. 2016b. "Data Science and Management for Large Scale Empirical Applications in Agricultural and Applied Economics Research." *Applied Economic Perspectives and Policy* 38(3): 373–388.
- Woodard, J.D. 2016c. "Integrating High Resolution Soil Data into Federal Crop Insurance Policy: Implications for Policy and Conservation." *Environmental Science & Policy* 66: 93–100.
- Woodard, J.D., A.D. Pavlista, G.D. Schnitkey, P.A. Burgener, and K. Ward. 2012. "Government Insurance Program Design, Incentive Effects, and Technology Adoption: The Case of Skip-Row Crop Insurance." *American Journal of Agricultural Economics* 94(4): 823–837.

### Author Information

Joshua D. Woodard ([woodardjoshua@gmail.com](mailto:woodardjoshua@gmail.com)) is CEO and Founder, Ag-Analytics, Ithaca, NY.

Bruce J. Sherrick ([sherrick@illinois.edu](mailto:sherrick@illinois.edu)) is Professor and Director, TIAA Center for Farmland Research at University of Illinois at Urbana-Champaign, Urbana, IL.

Jim Moseley ([jmose74@gmail.com](mailto:jmose74@gmail.com)) is AGree Co-Chair, Former USDA Deputy Secretary.

Collin O'Mara ([collin@nwf.org](mailto:collin@nwf.org)) is President, National Wildlife Federation, Reston, VA.

Barry Gold ([bgold@wffmail.com](mailto:bgold@wffmail.com)) is Environment Program Director, Walton Family Foundation, Washington, DC.

John Piotti ([jpiotti@farmland.org](mailto:jpiotti@farmland.org)) is President, American Farmland Trust, Washington, D.C.

Lynn Scarlett ([lscarlett@TNC.org](mailto:lscarlett@TNC.org)) is Co-Chief External Affairs Officer, the Nature Conservancy, Arlington, VA.

Doria R. Gordon ([dgordon@edf.org](mailto:dgordon@edf.org)) is Lead Senior Scientist, Environmental Defense Fund, Washington, DC.

Greg Fogel ([gfogel@sustainableagriculture.net](mailto:gfogel@sustainableagriculture.net)) was Policy Director, National Sustainable Agriculture Coalition, Washington, DC.

William Salas ([wsalas@appliedgeosolutions.com](mailto:wsalas@appliedgeosolutions.com)) is President, Applied GeoSolutions LLC, Durham, NH.

Carl Mattson ([carl@mattsonfarms.com](mailto:carl@mattsonfarms.com)) is Farmer, Mattson Farms Inc., Chester, MT.

Kristin Weeks Duncanson ([kwdjist@hickorytech.net](mailto:kwdjist@hickorytech.net)) is Owner and Partner, Highland Family Farms, Mapleton, MN.

Robert Blair ([robert@threecanyonfarms.com](mailto:robert@threecanyonfarms.com)) is Owner, Blair Three Canyon Farms, Kendrick, ID.

Fred Yoder ([fredyoder4510@gmail.com](mailto:fredyoder4510@gmail.com)) is Chairman, North American Climate Smart Agriculture Alliance; Former President, National Corn Growers Association, Columbus, OH.

Sam Willett ([Willett@ncga.com](mailto:Willett@ncga.com)) is Senior Director, Farm, Risk and Tax Public Policy, National Corn Growers Association, Washington, DC.

Deborah M. Atwood ([DATwood@merid.org](mailto:DATwood@merid.org)) is Executive Director, AGree, Transforming Food and Ag Policy, Washington, DC.

**Acknowledgments:** We thank the members of the AGree Conservation and Crop Insurance Taskforce for helpful comments and discussions. We thank Rex Raimond, Todd Barker, and Annie Shapiro (Meridian Institute), and Irene Bradley (Ag-Analytics®) for editorial and coordination assistance and graphic design of the data warehouse figure.

**Funding:** Meridian Institute received funding from the Walton Family Foundation (grant number 2017-340) to gather input on the draft article, which is part of Meridian's ongoing support to the AGree Conservation and Crop Insurance Taskforce.