Science Societies

# Transforming the culture of data management in a federal science agency, one client at a time

February 23, 2021

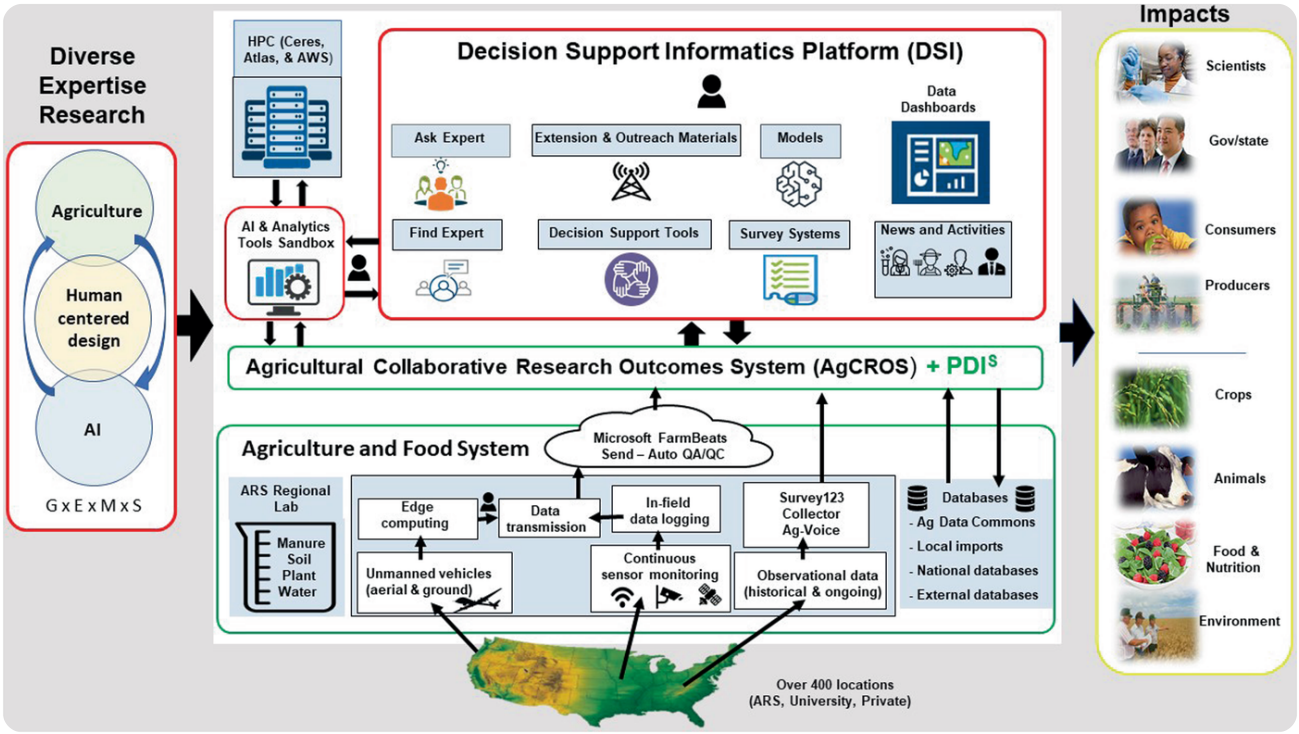

The Federal Corner is a bi-monthly column featuring writers from a diversity of federal agencies. Its purpose is to give funding agencies a chance to reach out directly to members.

---

The digital revolution is well upon us, enabling data-driven, evidence-based precision farming. Potential discovery and innovation await from new understandings of the

interactions of genotype, environment, management, and socio-economic factors (GxExMxS) to applications that optimize genomics, survey database information, and remotely sensed data, but these innovations are often buried in volumes of data. Standing between researchers and the possibilities of the digital revolution is the often mundane, but profoundly important, task of data management, which encompasses data collection, processing, analysis, storage, sharing, and maintenance. With data management, scientists can achieve new insights gained by the streaming of real-time sensor data, and then, with machine learning, scientists may optimize this data to integrate disciplines, creating systems for on-the-go precision adjustments, for example, for irrigation and drainage.

For research institutions, including federal agencies such as USDA-ARS, data management imposes tremendous barriers, but also opportunities, as researchers seek to take full advantage of the digital revolution in performing transdisciplinary science. Few scientists prioritize the FAIR principles—Findability, Accessibility, Interoperability, Reuse—or fully appreciate the value of their data as they contribute to a global common good. Old data management habits die hard, especially when replacing them requires a change in nearly every aspect of research: how we plan, collect, handle, manage, store, and access our data.




*USDA-ARS Partnerships for Data Innovations data management schematic with data flows. AgCROS and DSI are both cloud-based environments. All components within the green boxes represent the active/living data management workbench for agricultural researchers. The red box for DSI represents a one-stop-shop for stakeholder engagement. The DSI platform is still under development with the goal of creating a user-friendly environment were stakeholders find tools or information resources related to a wide range of agricultural research.*

## Partnerships for Data Innovation

For USDA-ARS, with 6,000 employees spread over 90 locations and 137 worksites, helping its staff to confront their data management habits requires cultural change and adoption of a data stewardship ethic. Enter the Partnerships for Data Innovation (PDI). Established in 2019, PDI is a partnership between agricultural research (ARS, university, and industry) and the private-sector technology companies that serves to equip ARS researchers with practical, innovative data management solutions that allow researchers to harness the possibilities of the digital revolution while shaping new habits that are consistent with FAIR principles. This national project started with a

handful of people with a common vision, who were able to pull together soft funds from more than 25 sources to develop a national cloud-based infrastructure and hire a talented technology team. This small group has greatly expanded to more than 300 people working on a highly diverse range of agricultural research topics.

The Partnerships for Data Innovation also serve industry partners, such as the cotton industry. Through the Dirt to Shirt Project, PDI, Cotton Incorporated, and ARS and university researchers from across the Cotton Belt partnered to: (1) generate a standard list of cotton data variables and units; (2) develop methods to incorporate data into PDI's cloud-based data network, Agricultural Collaborative Research Outcomes System (AgCROS); (3) produce field data collection apps; and (4) create data query and output tools (dashboards) in response to industry recognition of the scientific and economic value of increasing data management efficiency and data quality. For decades, the U.S. cotton industry has provided fiber quality data on each bale produced in the U.S. to its customers in partnership with the USDA Agricultural Marketing Service (USDA-AMS). Currently, AMS data are used in marketing bales to textile mills. Through the PDI process, these fiber quality data will be used to link GxExMxS factors to increase production and improve product quality while minimizing costs and environmental footprints.

## Ethic of Data Stewardship

In this and other projects (e.g., citrus greening, cattle fever tick eradication, long-term agroecosystem research, reservoir and dam management, and fertilizer recommendations), PDI has integrated its partners in the development of state-of-the-art digital solutions, applying off-the-shelf technologies from private-sector partners to ensure rapid adoption and application of their data management products. In doing so, PDI introduces its partners to the ethic of data stewardship, building good

habits and a culture of data management that ensure standardization, automation, and integration.

- ***Standardization***: Provides consistency in data management, ensuring that data are of the highest quality, and where possible, interoperable, discoverable, trackable, and transparent for maximum utility and impact.
- ***Automation***: Reduces inefficiencies, costs, and errors while preserving the integrity of the original data and supports standardization in processing.
- ***Integration***: Aligns potentially diverse objectives using comprehensive solutions that address the range of data management needs of PDI's partners.

A "Digital Research Workbench" was created by PDI as the foundation for data management, offering data handling, storage, analysis, and visualization solutions to partners. Operated by PDI's "Tech Team" of programmers and GIS analysts working with USDA-ARS data managers across the nation, the Digital Research Workbench houses applicable and customized digital solutions.

As customers have become partners, and partners have expanded across government, the private sector, and university institutions, PDI has sought to cultivate an ethic of responsible data management from collection to curation with the goal of establishing a culture of data stewardship within USDA-ARS and across the agricultural research community.

**More news & perspectives**

**Back to issue**

**Back to home**