Innovating in the Big Data Ecosystem: *Public-Private Partnerships* for a Data-enabled World

Solving complex data challenges require innovative cross-border, multi-sector partnerships

BY MELISSA CRAGIN, PH.D
EXECUTIVE DIRECTOR,
MIDWEST BIG DATA HUB

Complex data challenges facing the Great Lakes region in the era of big data transcend industries, applications, and borders. While data is increasingly borderless, borders and barriers still present substantial problems to industry, academic, and government initiatives that are dependent on data policy and governance processes that structure access and use. These challenges require innovative cross-border, multi-sector partnerships that can leverage the benefits of shared high performance computing resources and cyberinfrastructure services.

In an effort to address these challenges, the Midwest Big Data Hub (MBDH) was launched in 2015 with initial funding from the U.S. National Science Foundation (NSF). As a regionally-facing organization that collaborates with a range of industry, academic, government, and NGO partners, the MBDH is focused on improving access to data and amplifying unique collaborative relationships. Similarly, the Council of the Great Lakes Region recognizes these challenges and is leading the Great Lakes 4.0 initiative to foster innovative multi-sector partnerships on both sides of the border.

To develop solutions, the MBDH and the Council have partnered to jointly engage on common regional big data grand challenges, including resilient agri-food systems, advanced materials for manufacturing, infrastructure, transportation and logistics, and workforce development.

Consider the complexity of modern agri-food production, where inputs such as fertilizer may be imported from Canada for U.S. crop production, and where outputs may be exported back to Canada for use in processing for food products. These products and services may subsequently be imported for sale in the U.S. and Canada.

In this scenario, each activity and border crossing generates data sets that may improve productivity and efficiency if shared between partners at various points along the supply chain. Data generation in the agri-food context results from the wide variety of activities and technologies that make up the cyber-physical infrastructure of modern agriculture, from basic plant genomics through crop production, processing, distribution, consumption, and waste management. Building a sustainable and resilient Food-Energy-Water (FEW) ecosystem requires attention to all of the elements and stakeholders of which it is comprised,

regardless of geographic location.

Agri-food has been a core focus area for the MBDH since its inception, with activities including workshops and webinars about machine learning, sensors, and unmanned aerial systems as sources of big data for real-time and predictive analytics.

The MBDH also generates strategic partnerships to improve access and use of heterogeneous data for managing the built and natural environments.

One of our partner projects, "Bridge-ing Big Data," was interested in testing new data tools for improving access to and use of a national dataset on physical bridges, to support management and maintenance operations and facilitate real-time public information. Through our formal agreement with the National Data Service (NDS), a pilot project was launched to mobilize a new database to enable public access to the National Bridge Inventory data, and to provide an analytics environment for training and use.
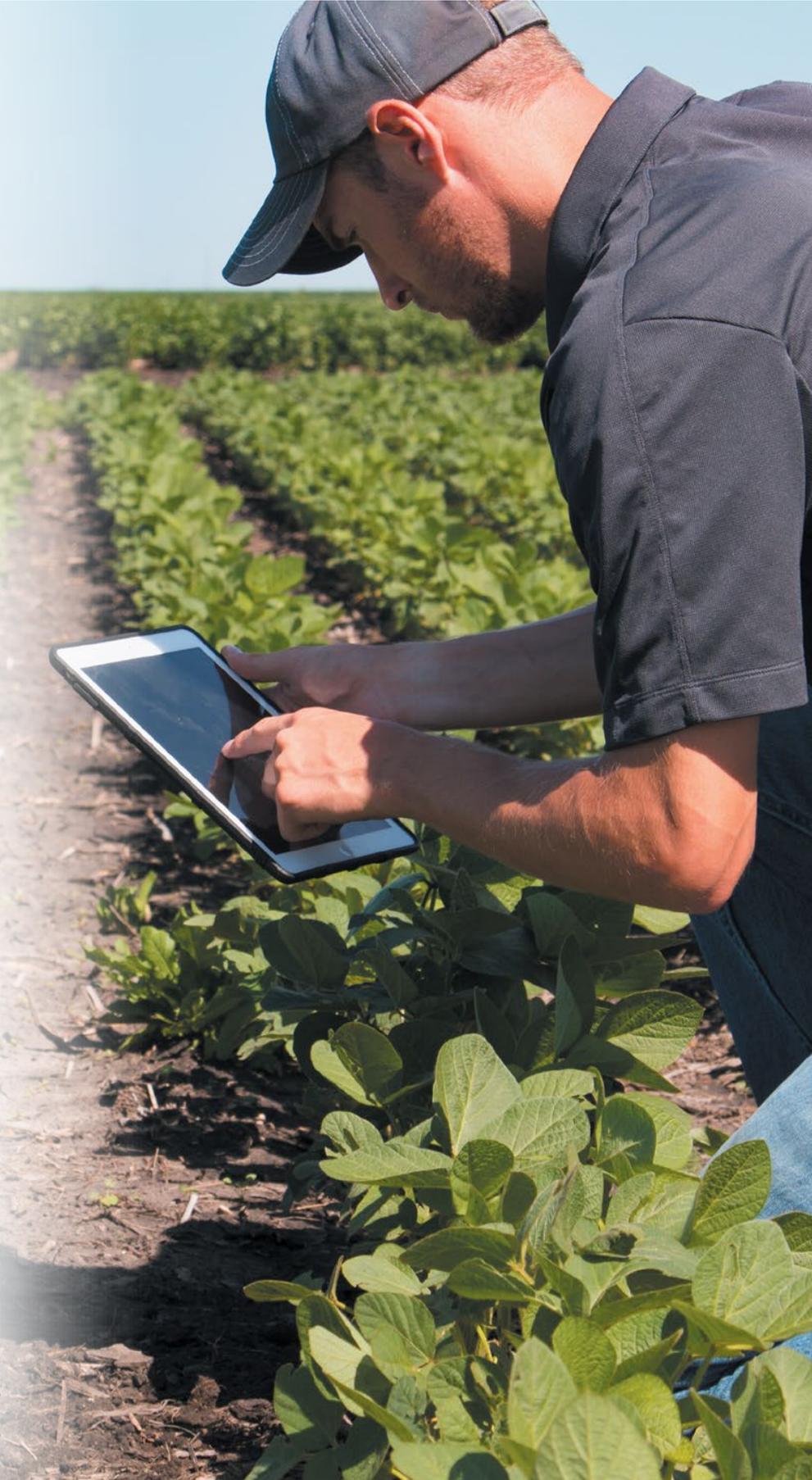
MIDWEST BIG DATA HUB

This strategic partnership with the NDS has granted MBDH members access to robust cyberinfrastructure, including testing "sandbox" services and a production cloud environment for data analytics.

Both examples above illustrate sector-level data challenges. At a higher level, we can see common themes emerging that cut across these sectors and others:

• large amounts of heterogeneous data generated by many different kinds of cyber-physical infrastructure devices; sensors, for example, on bridges, roads, and railways; in factories, cities, and fields; on trucks, trains, and agricultural equipment;

• complex systems and processes that cross organizational, geographic, and national boundaries, including integrated supply chain management and transportation ecosystems;

• a complex legal and regulatory environment, including local, state, provincial, and federal regulation; international treaties; and inter-organizational agreements that may involve shared data governance or cybersecurity.

Each of these themes - big data, complex systems, and uncertain regulation - require partners to collaborate, innovate, and share knowledge and resources. Uncertain times call for strong engagements around shared interests. ■

Melissa Cragin, Ph.D, was previously a Staff Associate at the National Science Foundation in the Directorate for Biological Sciences, and a AAAS Science & Technology Policy Fellow in Washington, DC.