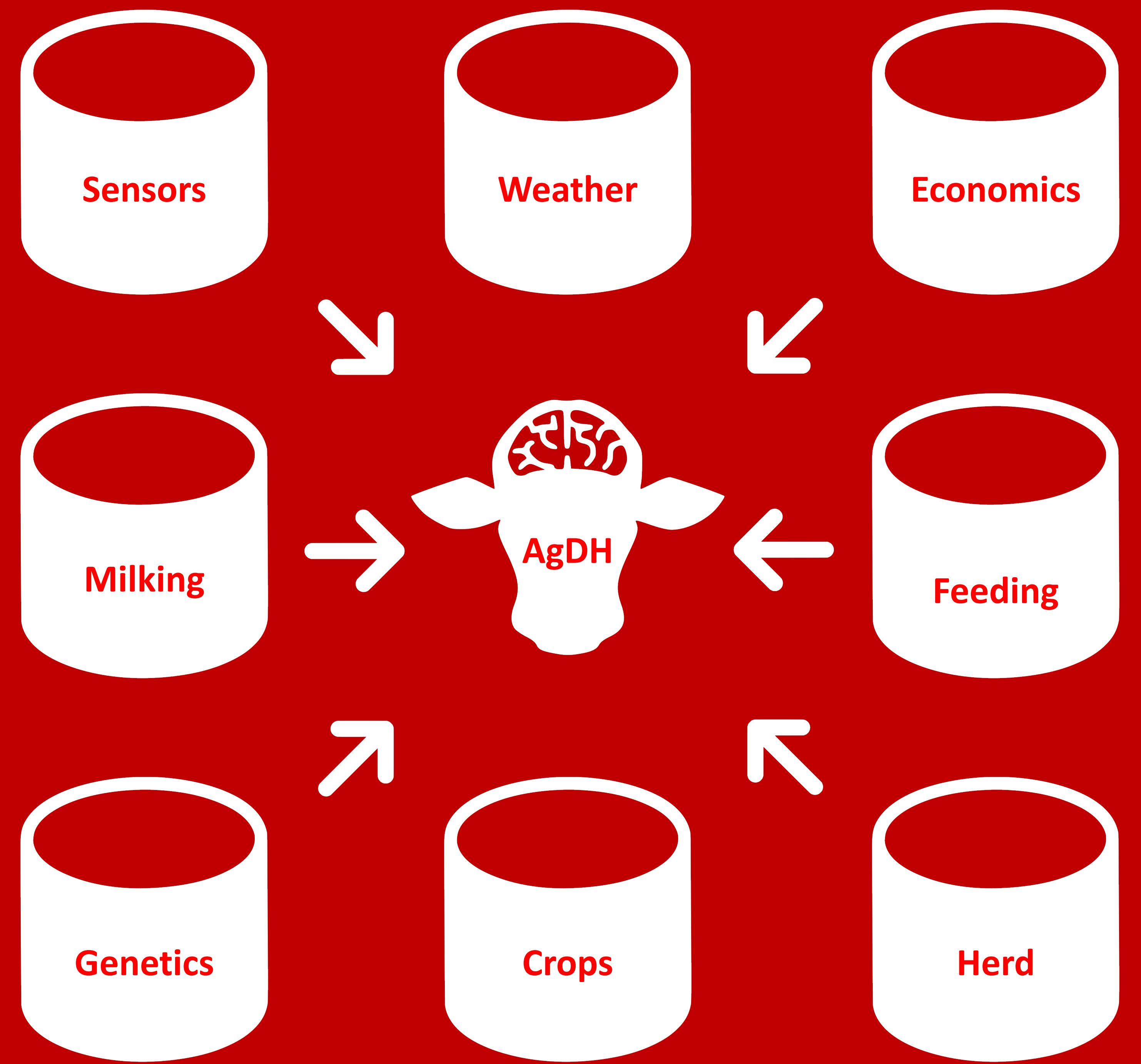


# Interpreting dairy data is hard for farmers. We at Dairy Brain made it easier by developing a centralized Data Hub.



## AgDH: A Distributed System for Gathering and Disseminating Dairy Data

### Introduction & Motivation

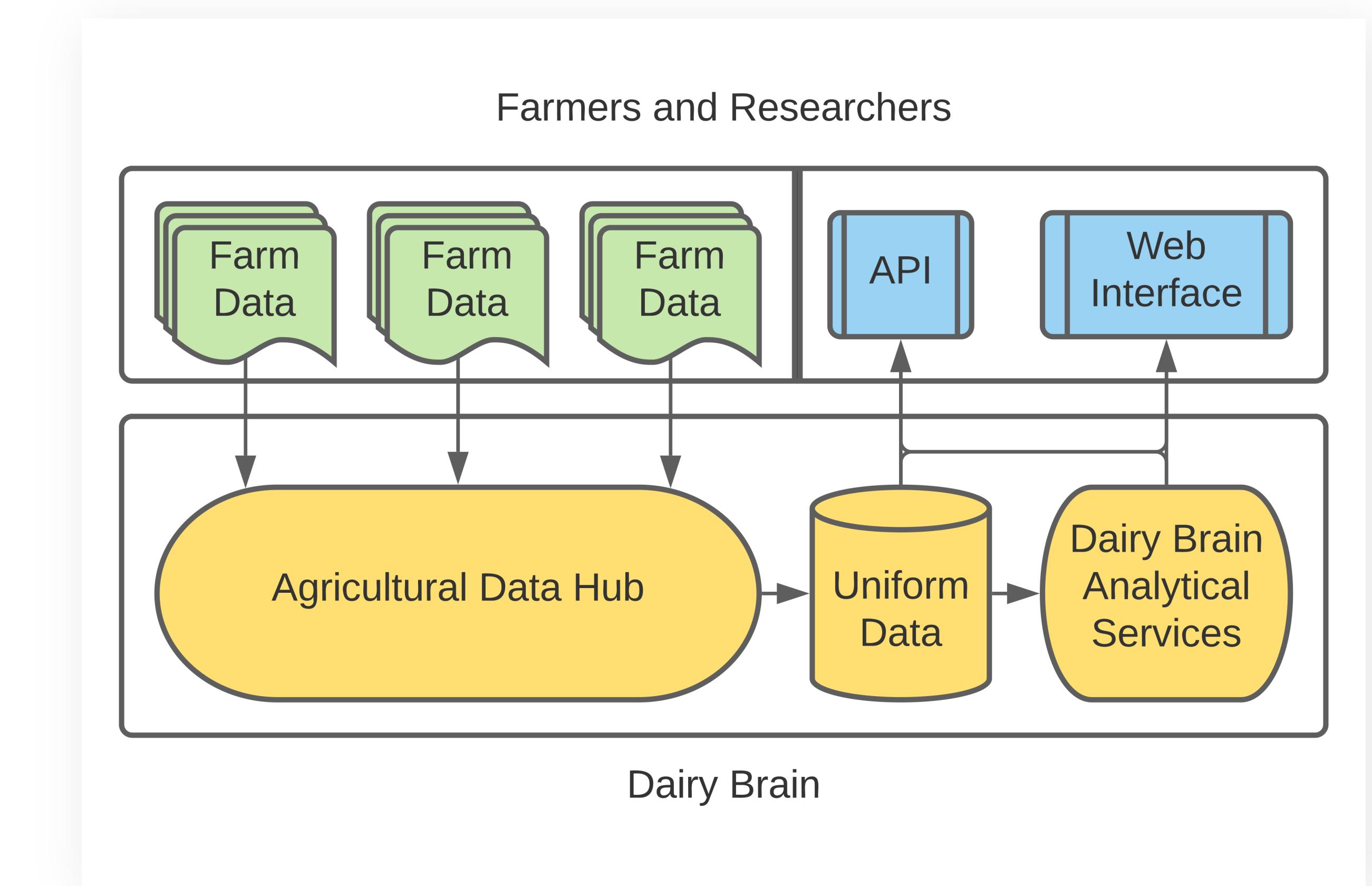
- Modern dairy farms use different software, generating vast amounts of data that are poorly linked.
- We present the **Agricultural Data Hub (AgDH)**, which integrates those data streams and establishes relationships between them.
- Data produced by AgDH is provided for easier future analyses.

### Gathering & Cleaning Data

- Currently, the software produces data that are stored in raw heterogeneous file formats (.csv, .xlsx, etc.)
- Raw dairy data goes through an **extraction, transformation, and loading (ETL)** process that prepares these data for the next step.
- Once the raw data is uniform, **entity matching** is performed to identify each individual cow and match them in different data sources. As a result, information in different sources of data is connected.
- Normalized, structured data will be available either via an **application programming interface (API)** or a **web-based user interface**. Analytical modules in the Dairy Brain services or third-party research scientists will use APIs whereas dairy farmers will benefit from the web-based interface.

### Building the Ground for Future Analyses

- With uniform data, analytical modules and decision supporting tools (which utilize modern ML frameworks like PyTorch) in Dairy Brain can be applied, enabling more powerful analyses which assists with **data-informed whole-farm decision makings** (predicting diseases, grouping cows by nutritional needs, etc.)



### Acknowledgments

- We thank Oscar Zhang, Bingshu Li, Yanzheng Li, Liliana Fadul-Pacheco, Adam Christensen, Miron Livny, Jignesh Patel, Derek Cooper, Andrew Maier, Ben Huebner, Heather White, and Kent Weigel for their contributions to the Dairy Brain project.
- References:
  - Ferris M., Christensen A., Wangen S.. Symposium review: Dairy Brain--Informing decisions on dairy farms using data analytics, Journal of Dairy Science, Volume 103, Issue 4, 3874 – 3881. DOI: 10.3168/jds.2019-17199.
  - Cabrera V., Barrientos-Blanco J., Delgado H., Fadul-Pacheco L.. Symposium Review: Real-time continuous decision making using big data on dairy farms, Journal of Dairy Science, Volume 103, Issue 4, 3856-3866. DOI: 10.3168/jds.2019-17145.



Rui Pan, Shu Liu, Steven R. Wangen, Michael C. Ferris  
University of Wisconsin-Madison, Wisconsin Institute for Discovery