

## **Data Management and Analytics PDG**

**Attendees:** Daniel Weller (Vice-Chair), Shahram Ajamian, Juan Archila, Charles Bakin, Clifton Baldwin, Chris Boyles, Janet Buffer, Jack Burnett, John Kevin Castro, Sejin Cheong, Taejung Chung, Tyler Chandross-Cohen, Sitara Cullinan, De Ann Davis, James Doyle, Sofia Santillana Farakos, Connie Freese, Yupawadee Galasong, Alberto Garre, Steven Gendel, Gabriella Gephart, Stephen Grove, Abdul Aziz Al Hajeri, Olivia Haley, Roger Hancock, Eelco Heintz, Joe Heinzemann, Dave Hill, Allison Howell, Brian Izdepski, Basel Al Jaber, Megan Jensen, Chris Jordan, YeonJin Jung, Ahmed Al Kaabi, Paula Klassen, Barbara Kowalczyk, Michala Krakowski, Vijay Krishna, Bobby Krishna, Devin LaPolt, Xiyang Liu, Jinxin Liu, Nicolas Nguyen Van Long, Yuxiao Lu, Steven Mandernach, William Melnychenko, Caroline Motzer, Claire Murphy, Kefang Nie, Dana Hoffman-Pennesi, Brian Perry, Harmeen Prasher, Chenhao Qian, Erin Ramsay, Gustavo Reyes, Carrie Rigdon, Lisa Robinson, Todd Ruthman, Deepti Salvi, Mansour Samadpour, Elisa Guardado Servellon, Natalie Seymour, Julie Simcox, Matthew Stasiewicz, Kelly Stevens, Sriya Sunil, Abdullatif Tay, Corey Troutman, Xiaohong Wei, Linghuan Yang, Jiyeon Yi, Pam Young.

**Total Number of Attendees:** 72.

**Meeting Called to Order:** Sunday, 9:00 a.m. Eastern Time, July 16, 2023.

**Minutes Recording Secretary:** Daniel Weller and Elisa Guardado Servellon.

**Old Business:** We started the first annual, in-person PDG with an introduction of the PDG's mission and history (Barbara Kowalczyk). We then moved forward with brief introductions. After introductions, we reviewed the antitrust guidelines before moving on to IAFP Executive Board announcements (from Mark Carter). Key updates centered on the importance of *IAFP Connect*, and reminders about 2023 in Toronto (e.g., to check passport and visa requirements, that COVID19 vaccines are mandated for entry to Canada). The student liaison then provided updates from the Student PDG (Claire Murphy); if members are interested in getting involved with the Student PDG activities (e.g., mentoring programs contact Claire Murphy at cmarik@vt.edu). We then reviewed the events at IAFP 2022 that the DMA PDG co-sponsored (there were four total). Attendees were engaged in a discussion, through break-out groups, on data management and analytic needs for the IAFP community. Each break-out group then shared their ideas with the PDG. Most ideas focused on data integration; working with and addressing messy and dirty data; data sharing and ownership; and/or data translation and communication (e.g., storytelling). Other common themes included how to generate practicable outputs (visualizations and models) that can be used by companies with limited and/or by companies with ample resources, and how to identify data business needs and capabilities.

Afterward, we discussed how to best structure the PDG moving forward to best meet the PDG and wider IAFP needs. We first discussed how frequently we would like to meet. Suggestions were made for monthly, bimonthly, and quarterly. We settled on meeting bimonthly but recognized that this frequency could be increased or decreased as needed. We then discussed in break-out groups how and when to organize

webinars (e.g., as part of the bimonthly meetings) and potential webinar ideas; these ideas were recorded in the Google sheet (<https://bit.ly/3zM7OxE>). One point of discussion was if the webinars were to support analytical skill development of the PDG members or to support data management and use within the wider IAFP community. We determined that the webinars could support both aims. We then moved forward with idea generation for symposium and roundtable for IAFP 2023. We started with a Word Cloud which will be posted in *IAFP Connect*. Top words included data standardization, management, and literacy as well as predictive modelling, spoilage, successes, digital decision management, data communication and translation. We then split into small break groups and reported back with these symposium ideas. We discussed ideas on how to best establish collaborations with other PDGs or with speakers. Ways discussed to do this included sharing your ideas in other PDG meetings and reaching out to the speakers from or leadership of the other PDG in-person while at the conference. Ideas and submission progress will be recorded in <https://bit.ly/3zM7OxE>. We also discussed recommendations to the IAFP Executive Board and reminded attendees about symposium proposals and abstracts. Finally, it was announced that an election will be held for a new Secretary in spring 2023; at IAFP 2023 Barbara Kowalczyk will become Chair Emeritus, Sarah Murphy will become the new Chair and Daniel Weller will become the new Vice-Chair.

**New Business:** The meeting started with a welcome from the Board and reminders. Each member of the PDG introduced themselves. It was announced that *JFP (Journal of Food Protection)* is available as open access, notifications must be activated to receive the journal publications. It was also reminded that members can sign up for the IAFP mentorship program.

Barbara Kowalczyk welcomed the PDG attendees. All attendees accepted the previous minutes. Barbara Kowalczyk is an outgoing chair, and Sarah Murphy is the current chair. The committee is accepting nominations for the secretary position for this PDG. Overview of roundtables and symposia, the PDG cosponsor these roundtables alongside other PDGs. Data sharing was mentioned as an important topic that would be discussed at a roundtable.

#### **Roundtable about Data Management:**

The presenters of the roundtable were introduced. The roundtable had presenters from academia, FDA, and organizations. The presenters gave an insight into the current challenges of data sharing. Attendees were encouraged to ask questions and participate. After the introduction, the roundtable started with the main question: "How to convince growers/people that investing in data management is important?" all the presenters agreed that people do see a return on investment despite the challenges. It was noted that there is an opportunity in risk factors associated with food safety, it is necessary to share across academics and companies to get involved in data sharing. Retail inspections have a high-performing sector, when looking at the data, they realized there are many things that could be performed much better. This data allowed them to rethink how to improve.

There was participation from a manufacturer's point of view, manufacturers want the data to be a means to solve a problem, the trustworthiness of the data, and how to prioritize the data. Data aggregation is how to

separate data in an effective way. Discussion started around this topic, inspectors may have long inspection times. There are many challenges when it comes to working with data in a controlled environment. An attendee mentioned that this is an important topic, in the food industry, there has been a lot of predictive model data, thousands of food companies have had issues with people outsourcing their data, and there is a concern about how data is protected and handled. On Wednesday morning there will be a meeting discussion about cybersecurity. It was mentioned that another challenge with data management is how much of the information can be shared. It is possible to design a platform that can conform to these points. Another participant mentioned how to handle quantitative and qualitative data, from FDA regulators and mentioned that it is possible to apply multicriteria models or semi-quantitative models. There was participation raised another question, "has there been a need to change this or improve data sharing" A second question was "Is there a value that is seen in sharing data in the first place" Government data systems have a lot to be desired, there is an interest on these models, alongside a return for them. A participant that works with providing software technology for data analysis mentioned that it is interesting how there is a focus on who is going to benefit from the data, how is it going to be shared, and who verifies it. A question that was raised "What are the biggest hurdles seen within companies or organizations when it comes to data sharing" One of the challenges is showing the importance and usefulness of data. It was mentioned that for FDA it takes a significant amount of time to gather data, making data comparable, it is necessary to create data dictionaries, transparency is key to capture differences in information. It was mentioned that one of the obstacles to data sharing is "what is in it for me. Another challenge that was mentioned is that data can be stolen easily, if there is a good data system setup, if possible for it to be used. As an answer during the roundtable, most data of organizations is voluntary. A question that was raised was "How to go about working with massive amounts of variables in data sets" It was answered that they prioritize what the customer is asking for. Another question of a participant is that companies could be interested in how to transfer data in a way that can be analyzed and understood. From an academia perspective, is to prepare data that is significant for food safety. A participant asked "How to engage users to act" Engagement is tracked in FDA websites, this is focused on data analysis, not solely on the user. A comment that was mentioned is that it is necessary to define the problems or challenges for organizations in data sharing and food safety.

### **Symposia and Webinar Ideas:**

The deadline for submitting symposium proposals to IAFFP is October 3, 2023. Daniel Weller will keep track of webinars and symposium ideas. One proposal was "Defining a Food Safety Data Standard for Laboratories." It was commented that it would be necessary to engage in ways that people can be active. It was commented it might be good to target an industry for data standards. Another idea is to segregate this topic into groups. Another suggestion was "Leading Indicators in Data Sharing in Food Safety." "Traceability and International Perspective" was another suggestion. "Needs of Publics and Private Sectors: How to train people in data sharing." Another discussion was "Food Safety Data Standard in Water." Symposium on using "Correlated Data." A suggestion was "How to Take a Process and How to Capture the Data Correctly." A suggestion was "Differences in Mechanisms in Date Sharing from an International Perspective." Another idea was "Date Digitalization." "How to Define Data Metrics" and "What is Define as a Good Food Safety Performance." "How

to Deal with Weak Signals in Data Sharing in Regulatory Areas.” “What Does Data Science Must Know About Food Science.” “How to Start Data Sharing and Know about the Legal Issues Involved.” More than eighteen ideas were collected during the meeting. A webinar series of ideas were suggested “Learning How to Organize Data in RStudio.” “How to Translate Food Safety in Date Science” could be set up as a technical forum.

**Recommendations to the Executive Board:** None

**Next Meeting Date:** July 14, 2024, Long Beach, California.

**Meeting Adjourned:** 11:00 a.m. Eastern Time.

**Chairperson:** Sarah Murphy.