ANNUAL REVIEWS

*Annual Review of Food Science and Technology*

# How Can AI Help Improve Food Safety?

C. Qian,[1] S.I. Murphy,[2] R.H. Orsi,[1] and M. Wiedmann[1]

[1]Department of Food Science, Cornell University, Ithaca, New York, USA; email: mw16@cornell.edu

[2]Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, New York, USA

Annu. Rev. Food Sci. Technol. 2023. 14:517–38

First published as a Review in Advance on December 21, 2022

The *Annual Review of Food Science and Technology* is online at food.annualreviews.org

https://doi.org/10.1146/annurev-food-060721-013815



## Keywords

machine learning, public health, supply chain, computer vision, predictive analytics

## Abstract

With advances in artificial intelligence (AI) technologies, the development and implementation of digital food systems are becoming increasingly possible. There is tremendous interest in using different AI applications, such as machine learning models, natural language processing, and computer vision to improve food safety. Possible AI applications are broad and include, but are not limited to, (*a*) food safety risk prediction and monitoring as well as food safety optimization throughout the supply chain, (*b*) improved public health systems (e.g., by providing early warning of outbreaks and source attribution), and (*c*) detection, identification, and characterization of foodborne pathogens. However, AI technologies in food safety lag behind in commercial development because of obstacles such as limited data sharing and limited collaborative research and development efforts. Future actions should be directed toward applying data privacy protection methods, improving data standardization, and developing a collaborative ecosystem to drive innovations in AI applications to food safety.

**Artificial intelligence (AI):** a computer program that can sense, reason, act, and adapt

# 1. INTRODUCTION

Food safety remains a global concern, as supported by its estimated global impact on the economy and public health. Most recent estimates by the World Health Organization (WHO) indicate that globally every year, foodborne illnesses cause 600 million cases, resulting in 420,000 deaths and the loss of 33 million disability-adjusted life years (DALYs) (WHO 2015). Overall, progress in reducing foodborne illnesses globally has been slow and challenging. For example, foodborne illness rates for key foodborne pathogens in the United States (e.g., *Salmonella*, *Listeria monocytogenes*) have been relatively steady over the past decade (CDC 2021) and the United States has consistently failed to meet many of its "Healthy People" food safety–related public health goals. For example, although Healthy People (2020) targeted a reduction in *Salmonella* incidence from 15.0 cases to 11.4 cases per 100,000 people by 2020, *Salmonella* incidence reached 17.1 cases per 100,000 people in 2019 (Tack et al. 2020).

The challenges to reducing foodborne illnesses are multifaceted and complex and differ substantially by region, country, and the specific pathogen and foods of concern. In many countries, foodborne illness burdens are essentially unknown, primarily due to a lack of both public health surveillance systems and systematic food surveillance for foodborne pathogens. Although a focus on other public health issues perceived to be more urgent and impactful (e.g., malaria and HIV) may be one reason for the lack of surveillance, the costs and complexity of good foodborne disease surveillance also represent a barrier. Notably, the WHO recently recognized the public health impact of foodborne illness (including its impact on the health of children < 5 years), which may help enhance foodborne disease surveillance in some countries and regions (WHO 2021). However, analytical tools are still needed to support the (*a*) design of improved and efficient surveillance systems and (*b*) interpretation of surveillance data.

However, even in countries with relatively sophisticated surveillance systems (e.g., the United States, United Kingdom, some Western European countries, New Zealand), improved control of foodborne illnesses remains a challenge. Reasons for this include the extreme diversity of the food supply system (e.g., large agricultural and food operations as well as hyperlocal small operations) and different foodborne pathogens with distinct ecologies, transmission pathways, and contamination sources, ranging from pathogens where preharvest contamination is a key concern (e.g., *Salmonella* Enteritidis) to organisms where contamination from human sources, typically at the location of meal preparation, is the main concern (e.g., norovirus). However, the issue is not as simple as one may think; even with limited preharvest control, *Salmonella* Enteritidis could ideally be controlled effectively if all consumers would practice perfect food safety strategies (e.g., no cross contamination and no consumption of raw eggs). Although improved detection and testing may help reduce foodborne illnesses, testing cannot assure safety, as foodborne illnesses and contamination events are typically rare and heterogeneously distributed [e.g., a listeriosis outbreak in the United States involved one human illness per 339,200 servings (Pouillot et al. 2016)]. Hence, novel approaches and strategies are essential for improved control of foodborne pathogens. In particular, through improved data analytics and predictive capabilities, artificial intelligence (AI) shows tremendous potential for improving food safety, as further detailed below.

# 2. BACKGROUND ON ARTIFICIAL INTELLIGENCE

## 2.1. Overview

AI is conceptualized as the ability of a computer program or robot to perform human tasks (Copeland 2021). Moreover, AI programs are expected to have the ability of an intelligent being to interact with the environment and learn from the experience. Although not always visible, AI technology is ubiquitous in our lives in the form of software or embedded in hardware (Eur. Comm. 2018). Motivations to use AI are broad and generally include (*a*) mimicking human behaviors with the same level of competence and therefore reducing human labor or improving efficiency [e.g., an auto-driving car that can recognize obstacles while making appropriate driving decisions (Gupta et al. 2021)]; (*b*) performing complex tasks that require a high level of intelligence [e.g., AlphaGo, which outperformed humans in chess (Chen 2016)]; and (*c*) performing tasks beyond human capability or intuition [e.g., Netflix recommendation systems that automatically find and display movies that users will possibly like (Ranjan et al. 2019) and drug and vaccine discovery (Goodswen et al. 2021)]. The agri-food industry has also embraced the AI revolution by adopting different AI technologies for crop yield prediction (Sinwar et al. 2020), product quality control (Kondakci & Zhou 2016), improved traceability (Wang et al. 2017), and product development (Garver 2018). Although many review papers have already discussed AI applications in the food industry (Kakani et al. 2020, Mavani et al. 2021, Misra et al. 2020), this article focuses on AI applications in food safety, particularly microbial food safety.

**Computer vision:** a field of artificial intelligence that aims to analyze images and other visual information

**Natural language processing:** a field of artificial intelligence that aims to interpret human language

## 2.2. Categories of Artificial Intelligence Systems

AI systems in the food industry can support decision-making based on the data streams fed through various processing equipment, instruments, and devices. Depending on how an AI system processes data and generates decisions, AI applications can be divided into two groups: (*a*) rule-based AI and (*b*) data-driven AI. A rule-based AI has a list of behavioral rules preprogrammed in advance. These rules are usually based on previous knowledge or experience about the physical food system and are often determined by expert opinions. As an example, a rule-based system using fuzzy rule-based reasoning was designed to assess the risk of violation for a cross-border e-commerce commodity (Song et al. 2019). Although a rule-based AI can be useful in making real-time decisions, one of its weaknesses lies in the inability to improve itself automatically. For example, if there are changes to the system (e.g., changes in weather patterns or production processes) that might impact the rules governing the food system and thus management decisions, then a rule-based AI system needs to be reviewed or verified by experts to remain relevant. Another weakness is that due to the complexity of food systems, rules in such AI systems are often (*a*) transformed to qualitative or semiquantitative via fuzzy logic, which might limit the exploitation of the data's full potential, and (*b*) limited to those rules that can be defined with human intuition. In contrast, a data-driven AI system can overcome these drawbacks by (*a*) improving performance with the continuous collection of new data and (*b*) uncovering patterns beyond human knowledge. However, it is worth noting that sometimes an AI system can have components from both rule-based and data-driven applications, allowing it to be more flexible in handling different data sources. For example, a digital system named Supply-chain Pedigree Interactive Dynamic Explore (SPIDER) was proposed in a study (Wang et al. 2013) and has both a case-based reasoning engine and a neural network (NN) as part of the platform. The synergistic effects between these two techniques allow it to support and verify the implementation of Hazard Analysis and Critical Control Points (HACCP), which is a well-established management system for food safety. Despite this example, overall, AI is moving toward data-driven applications (Rajan & Saffiotti 2017), which are the focus of this review. Key subfields of data-driven AI relevant to food safety include (*a*) computer vision, (*b*) natural language processing, and (*c*) analytical tools (**Figure 1**), which are discussed in greater detail below, with examples from the (*a*) supply chain, (*b*) public health, and (*c*) microbial data collection (**Figure 2**).

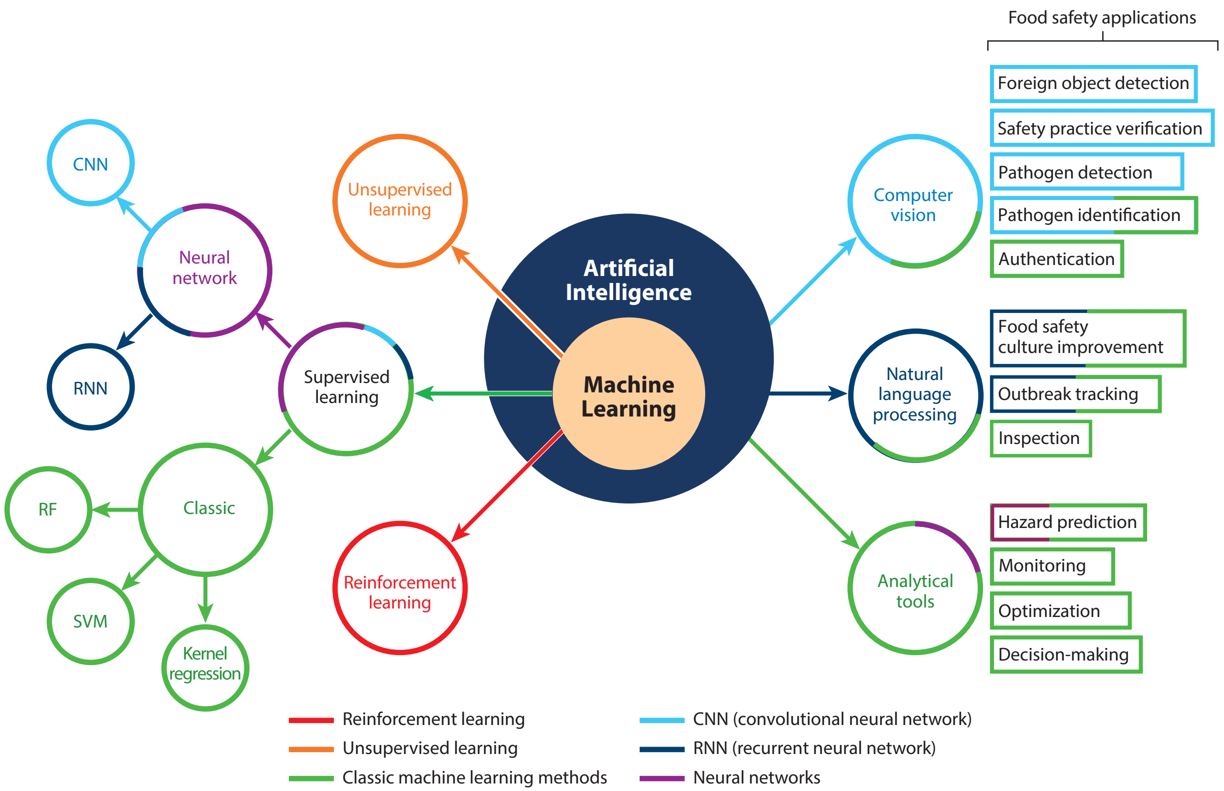


**Figure 1**

Conceptual illustration of key components within the realm of artificial intelligence (AI). Machine learning is one of the subfields of AI. Unsupervised learning, supervised learning, and reinforcement learning are learning methods within machine learning. Classic and neural networks are supervised learning algorithms. Random forest (RF), support vector machines (SVMs), and Kernel regression are algorithm subtypes within the classic algorithms, whereas recurrent neural network (RNN) and convolutional neural network (CNN) are algorithm subtypes of the neural network algorithms. Computer vision, natural language processing, and analytical tools are AI fields. Rectangular boxes depict examples of food safety applications in each AI field. The color percentage of each circle and rectangle (representing AI fields and food safety applications, respectively) represents a rough estimate of the proportion of algorithm types used in each AI field or for each food safety application type (based on the reviewed literature). Adapted from Mega Map of Machine Learning provided by While True: Learn() (**https://luden.io/wtl/**).

## 3. ARTIFICIAL INTELLIGENCE APPLICATIONS ACROSS AGRI-FOOD SUPPLY CHAINS

"Smart" technologies for the food industry have emerged and advanced rapidly in recent years, with numerous companies using AI technologies, especially at the production level (Kakani et al. 2020). Notably, there has been investment and a growing body of "smart" food industry applied research, including AI-based digital tools for data collection and analysis. AI applications in food safety across agri-food supply chains have the potential to support traceability, monitoring, inspection, and other purposes and processes. To date, however, there has been limited generalizability and commercialization of available AI technologies for food safety. Here, we present examples of opportunities for AI applications (**Table 1**) that support food safety that could be further developed and implemented in industry across all stages of agri-food chains, including raw materials, production, processing, packaging, storage, distribution, and retail.

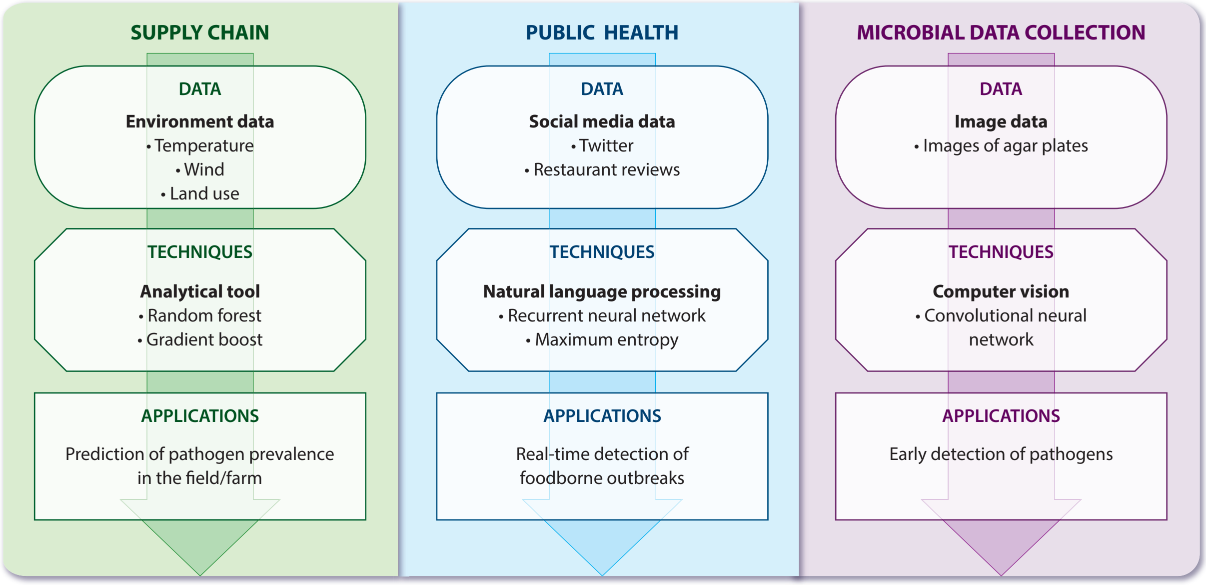


**Figure 2**

Workflow for AI applications in the areas of the supply chain, public health, and microbial data collection.

## 3.1. Raw Materials

Machine learning (ML) approaches can be used to support the inspection of foods, including raw materials. For example, a Bayesian network model based on data from the European Union's Rapid Alert System for Food and Feed from 2000 to 2013 was used to predict food fraud (Bouzembrak & Marvin 2016). In addition, since launching a pilot program in spring 2019, the US Food and Drug Administration (FDA) has been developing an ML-based screening tool using data from past seafood shipments to facilitate identification of high- and low-risk seafood shipments (FDA 2022a). ML-based approaches can also be used for supplier selection and procurement. For example, Mars, Inc., developed an aflatoxin predictive model that has been deployed internally to guide sourcing options (FDA 2022b). Another study used extreme gradient boosting (XGBoost) ML algorithms to predict risk factors associated with abundance of *Vibrio parahaemolyticus* in oyster farms in Taiwan (Ndraha et al. 2021). Future research should support ongoing efforts by industry and government to develop tools for food safety inspection and procurement of raw materials, which would ideally lead to near real-time prediction of high-risk raw materials streams with better specificity and speed compared to classical testing-based approaches to assuring low-risk raw materials.

**Machine learning (ML):** a subfield of artificial intelligence that leverages data and learning algorithms to improve its performance

## 3.2. Production

At the production level, AI applications have been developed to support data collection and processing, analytics, and management of crops, livestock, water, and soil (Astill et al. 2020, Kamble et al. 2020, Kamilaris & Prenafeta-Boldú 2018, Liakos et al. 2018, Martos et al. 2021, Sharma et al. 2020, Smith 2019). To date, production-level AI food safety research has been primarily focused on ML-based descriptive and predictive tools. For example, numerous supervised learning models have been developed for predicting foodborne pathogen presence and levels in agricultural water

**Supervised learning:** learning performed on labeled data (i.e., data output is well-defined)

**Table 1** **Potential artificial intelligence (AI) applications in food safety to improve agri-food supply chains**

| Subfields | Specific applications | AI branch | Reference(s) |
|---|---|---|---|
| Raw materials | Alert system to predict food fraud | Analytical tool | Bouzembrak & Marvin 2016 |
| | Screening tool for high- and low-risk shipments | Analytical tool | FDA 2022a |
| | Supplier selection and procurement | Analytical tool | FDA 2022b |
| Production | Predicting pathogen presence in agricultural water | Analytical tool | Belias et al. 2021; Polat et al. 2020; Stocker et al. 2022; Toro et al. 2022; Weller et al. 2020a,b; 2021a,b |
| | Predicting indicator microorganisms' prevalence in poultry farm | Analytical tool | Golden et al. 2019a,b |
| Process and packaging | Optimizing quality management practices | Analytical tool | Murphy et al. 2021 |
| | Quality and safety assessment of foods with sensor technology | Analytical tool | Chen & Yu 2022, Feng & Sun 2012, Ropodi et al. 2016, Yang et al. 2010, Zhou et al. 2019 |
| | Digital twins to facilitate environmental monitoring programs and inform corrective actions | Analytical tool | Barnett-Neefs et al. 2022, Sullivan et al. 2021, Zoellner et al. 2019 |
| | Monitoring and optimizing equipment cleaning and sanitation | Analytical tool | Escrig et al. 2019, 2020a,b; Simeone et al. 2020; Úbeda et al. 2016; Wallhäußer et al. 2011, 2013 |
| Storage, distribution, and retail | Optimizing retail refrigeration temperature | Analytical tool | Onoufriou et al. 2019 |
| | Simulation-based food safety training using vision technology | Computer vision | Friedlander & Zoellner 2020 |
| | Predicting food hygiene compliance of food outlets | Analytical tool | Oldroyd et al. 2021 |
| | Identifying potentially unsafe food product online from review comments | Natural language processing | Maharana et al. 2019 |

and investigating associated risk factors (physiochemical, temporal, geospatial, weather, and anthropogenic) (Belias et al. 2021; Harrand et al. 2020; Polat et al. 2020; Stocker et al. 2022; Toro et al. 2022; Weller et al. 2020a, 2021a,b). Hence, there are opportunities for utilizing ML to support preharvest water management. For example, one study trained and compared advantages and disadvantages of various ML approaches, including support vector machines (SVMs) and Bayesian, tree-based, ensemble, instance-based, penalized regression, and rule-based learners to predict enteric pathogen (*Salmonella* and pathogenic *Escherichia coli*) presence in agricultural water sources; this study demonstrated that ML-based predictive models may be useful for identifying when and where pathogens are likely to be present in agricultural water (Weller et al. 2020b).

ML has also been applied for investigating foodborne pathogens in the broader produce field environment. For example, one study utilized tree-based modeling to identify landscape and meteorological factors associated with the presence of pathogens (*L. monocytogenes*, *Salmonella*, Shiga-toxin-producing *E. coli*) in soil, fecal, water, and drag swab samples from produce fields (Strawn et al. 2013). Another study utilized random forest (RF) to identify and rank factors associated with *L. monocytogenes* isolation from soil, drag swab, and agricultural water samples from produce fields, demonstrating that water-related factors were the most important (Harrand et al. 2020). An example of larger-scale data collection and modeling is the ongoing Western Growers Food Safety Data Sharing Project, which involves utilizing growers' data (inspections, product and water testing, location) as well as temporal, meteorological, and landscape data to train an ML model to support food safety risk assessment (FDA 2022b). Few studies have utilized supervised ML approaches in production environments other than produce (e.g., poultry farms) (Golden et al. 2019a,b). Unsupervised learning approaches have also been used at the production level for dimensionality reduction (e.g., principal component analysis) (Ivanek et al. 2009, Park et al. 2014, Strawn et al. 2013). Leveraging existing production-level ML applications, future work should focus on developing tools to support food safety management by producers, including implementation and validation of tools on the farm.

**Unsupervised learning:** learning performed on unlabeled data (i.e., only the input data are well-defined)

## 3.3. Processing and Packaging

At the processing and packaging levels, AI can be utilized for informed decision-making and task automation, facilitating optimization of food quality and safety management, reducing risks, and increasing quality and productivity. For example, conditional RF was used to identify and rank facility-level quality management factors associated with postpasteurization contamination of fluid milk (Murphy et al. 2021). There have been numerous applications of ML for anomaly or defect detection in foods (e.g., produce) (Nturambirwe & Opara 2020). ML has also been used with sensor-based devices (e.g., e-nose, spectroscopy, imaging) for food quality and safety assessment (Chen & Yu 2022, Feng & Sun 2012, Ropodi et al. 2016, Zhou et al. 2019). For example, one study developed an algorithm to detect fecal contamination on leafy greens using hyperspectral imaging (Yang et al. 2010). As previously shown (Haiminen et al. 2019), AI based on metagenomics of food and food ingredients can also help with food authentication, which is another relevant application to food safety, as fraudulent food or food with fraudulent ingredients may represent a higher food safety risk.

Environmental monitoring programs are implemented in food facilities to detect contamination and verify the effectiveness of control measures. Recent studies have developed agent-based models of *Listeria* contamination in food facilities, allowing for the simulation of sampling strategies and corrective actions (Barnett-Neefs et al. 2022, Sullivan et al. 2021, Zoellner et al. 2019). ML can be combined with agent-based modeling to improve model performance and decision-support (Zhang et al. 2021b). ML approaches can be also used to design and optimize environmental monitoring programs in facilities. For example, at the retail level, one study developed ML-based (SVMs and RF) classification models of *L. monocytogenes* persistence in retail delicatessen environments (Vangay et al. 2014). Building on this previous work on *Listeria*, future efforts to develop model-based tools to facilitate decision-support relating to environmental monitoring of foodborne pathogens in different types of facilities should be pursued. Additionally, the aforementioned agent-based models of *Listeria* in food facilities represent a first step toward development of digital twins (i.e., a virtual or digital representation of physical systems to simulate the behavior of the system) (Defraeye et al. 2021, Henrichs et al. 2022, Nasirahmadi & Hensel 2022); these types of tools will provide a tremendous opportunity to support and improve food safety management.

AI can also be employed for monitoring and optimizing cleaning and sanitation in the food industry. Previous studies have utilized ML with various sensor technologies (e.g., electrical, optical, acoustic, and ultrasonic) to monitor fouling and cleaning of food processing equipment (Escrig et al. 2019, 2020a,b; Simeone et al. 2020; Úbeda et al. 2016; Wallhäußer et al. 2011, 2013). For example, one study developed NN models able to predict area and volume of fouling during cleaning using data from ultrasonic sensors (Simeone et al. 2020). Another study developed a system to detect and diagnose anomalies in clean-in-place using historical data and ML (multiway principal component analysis) (Yang et al. 2018).

**Deep learning:** a subfield of machine learning that builds on deep neural networks

Preventive and corrective maintenance are important for maintaining food safety. Predictive maintenance is the concept of using predictive tools to determine when maintenance actions are necessary. Although ML methods applied to predictive maintenance across industries have been reviewed recently by Carvalho et al. (2019), to date, there have been no published AI-based applications for predictive maintenance in food processing facilities, presenting an opportunity for research and innovation.

### 3.4. Storage, Distribution, and Retail

At the storage, distribution, and retail levels, there are several AI applications for optimizing pricing and scheduling, identifying and navigating distribution, and personalizing advertising. However, there are few existing AI applications for food safety. Two main areas of opportunity for AI application are (*a*) supply chain management (especially cold chain) and (*b*) food safety management at retail.

Control of environmental conditions, especially temperature, in agri-food chains is important for food safety. Temperature control and monitoring can occur across all stages, including storage, distribution (e.g., refrigerated trucks), and retail (Chaudhuri et al. 2018, Mercier et al. 2017). Utilizing environmental data collected at various stages of supply chains, data analytics and modeling tools can enable traceability, risk management, and planning optimization. For example, AI tools can be developed to facilitate temperature control that maximizes sustainability and food safety, such as the deep learning–based tool termed Nemesyst, which optimizes retail refrigeration while ensuring temperatures stay within food safety limits (Onoufriou et al. 2019). AI can also be utilized for cold-chain break analysis, including temperature prediction and cold-chain break detection (Loisel et al. 2021).

Although there are few existing AI-based applications for food safety at the retail level, a recent article proposed many areas of opportunities for AI in retail (Friedlander & Zoellner 2020). For example, the authors proposed that simulation-based training modules utilizing vision technologies (such as virtual reality) should be developed to support retail food safety training (e.g., handwashing) (Friedlander & Zoellner 2020). ML-based approaches can also be used to support inspection of retail settings. For example, a recent study developed ML models based on neighborhood characteristics to predict food hygiene compliance of food outlets (including retail stores) (Oldroyd et al. 2021). Similar tools utilizing previously collected data can be developed by food companies with multiple locations to support their internal food safety inspections.

With the rise of online grocery shopping (i.e., e-commerce), there is an increased ability to track products and collect data, which can be leveraged for food safety AI applications. For example, one study (Maharana et al. 2019) developed ML- and deep learning–based models to identify potentially unsafe food products based on Amazon's customer reviews.

## 4. ARTIFICIAL INTELLIGENCE SYSTEMS TO ENHANCE PUBLIC HEALTH

AI technologies can be used to enhance public health by increasing our ability to predict (*a*) restaurants with poor hygiene practices, (*b*) the source of foodborne illnesses, and (*c*) etiological agents of foodborne cases. These predictions can help public health agents intervene more quickly, preventing the exposure of subjects to contaminated foods and preventing large outbreaks. Although these strategies have not been fully implemented in foodborne surveillance and outbreak investigations, they have shown promising preliminary results (**Table 2**). New implementations of these strategies in a format (e.g., apps) that can be easily used by public health investigators are needed before we can assess the practical usefulness of these methods in public health.

**Table 2 Potential artificial intelligence (AI) applications in food safety to enhance public health**

| Subfields | Specific applications | AI branch | Reference(s) |
|---|---|---|---|
| Surveillance | Real-time detection of foodborne illness and outbreaks from surveillance data | Analytical tool | Sadilek et al. 2018, Teyhouee et al. 2017, Wang et al. 2021 |
| | Real-time detection of foodborne illness and outbreaks from social media data | Natural language processing | Du & Guo 2022, Kuehn 2014, Sadilek et al. 2017 |
| Source tracking and source attribution | Identifying likely sources of foodborne illness and outbreaks | Analytical tool | Barco et al. 2013, Harrison et al. 2021, Mikkelä et al. 2019, Pires et al. 2009, Ranta et al. 2011, Thépault et al. 2017, Tyson et al. 2016 |
| Food hazards prediction | Improving diagnostic accuracy | Analytical tool | Mei et al. 2020 |

## 4.1. Surveillance

Surveillance is a cornerstone of food safety and may include human foodborne disease surveillance as well as surveillance of foods for the presence of foodborne pathogens. An early example of AI application to foodborne disease surveillance is an ML model for real-time detection of foodborne illness using anonymous and aggregated web search and location data (termed FINDER) (Sadilek et al. 2018). When applied to data collected in two cities, FINDER improved the accuracy of health inspections with restaurants identified by FINDER being 3.1 times more likely to show critical violations during restaurant inspections. Previous studies have also developed ML-based models to detect foodborne illness based on customer restaurant reviews (Effland et al. 2018, Harrison et al. 2014).

AI has also been used to assign etiology to foodborne illnesses investigated through syndromic surveillance via an ML algorithm that can assign the most likely etiological agent based on symptoms, onset of disease, and geographical location (Wang et al. 2021). ML methods for outbreak detection need to overcome the issues of under-reporting and erroneous reporting (Zhang et al. 2021a). To address these issues, an ML approach based on hidden Markov models for syndromic surveillance monitoring and disease outbreak detection was developed and implemented using a smartphone-based app for tracing the location of food consumption and subclinical reporting; this approach showed promising preliminary results (Teyhouee et al. 2017). These and additional applications of ML approaches to foodborne disease outbreak detection have been reviewed recently by Deng et al. (2021). As detailed in this review, the development and application of AI tools to use disparate data types [from whole-genome sequencing (WGS) data to web searches and sales data] to detect foodborne disease outbreaks seem to be a promising future direction. A framework for AI application to foodborne disease outbreak detection could include two types of data: (*a*) case, outbreak, and other surveillance data and (*b*) data that might influence foodborne disease analysis, such as weather, holiday, and social media data (Du & Guo 2022, Kuehn 2014, Sadilek et al. 2017). These data should then be integrated into a single ML approach to detect outbreaks.

## 4.2. Source Tracking and Attribution

Source tracking and source attribution are important to help (*a*) identify likely sources of foodborne disease cases and outbreaks and (*b*) focus public health and regulatory actions on those sources that cause the largest public health impact. Although non-AI-based tools for source tracking and source attribution have been described (Barco et al. 2013, Harrison et al. 2021, Mikkelä et al. 2019, Pires et al. 2009, Ranta et al. 2011, Thépault et al. 2017, Tyson et al. 2016), new

AI-based tools have also been reported. For example, Zhang et al. (2019) applied an ML RF classifier for source prediction of *Salmonella* Typhimurium using genomic surveillance data collected in the United States. Two studies have used LogitBoost (Munck et al. 2020) and various ML algorithms (Arning et al. 2021) to assign sources to *Salmonella* Typhimurium and *Campylobacter* isolates, respectively. Another study (Duarte et al. 2021) reported a metagenomics-based approach to source attribution of antimicrobial resistance (AMR) determinants, which used different RF approaches to classify resistomes into corresponding reservoir classes. Overall, published works clearly support the rich opportunity that exists for using AI approaches to enhance foodborne disease source attribution efforts.

### 4.3. Food Safety Risk Prediction

AI tools also provide a powerful approach for the prediction of times and locations at which there is an increased risk of foodborne pathogen presence or where there is an increased risk of food contamination with foodborne pathogens (e.g., owing to the occurrence of factors that facilitate the transfer of pathogens from the environment or other vectors to finished product). The power of this approach has been particularly well demonstrated by several studies that used different predictor variables to predict the risk of pathogen contamination of agricultural water and the risk of pathogen presence in fields used for produce production (see Section 3.2 above). Although in the context of supply chain management, these AI tools can be useful to facilitate decision-making with regard to harvesting and water treatment, they can also provide value to public health. For example, model predictions can be used to enhance testing frequencies or interventions at locations and time periods when AI tools predict an increased risk of pathogen contamination.

AI has the potential to impact other aspects of microbial food safety risk prediction. For example, a study on qPCR (quantitative polymerase chain reaction)-based diagnosis of COVID-19 showed that AI-based tools can be used to combine PCR findings with data from completely different tests (in this case CT scans of patients) to improve diagnostic accuracy (Mei et al. 2020). In a food safety context, a similar approach could be used during production or in the finished product to analyze qPCR amplification curves, which are used for pathogen detection, along with other data from the same sample (e.g., water turbidity, pH) to improve identification of foodborne pathogens in raw materials.

## 5. ARTIFICIAL INTELLIGENCE–SUPPORTED TECHNIQUES FOR FOODBORNE PATHOGEN DETECTION, CHARACTERIZATION, AND IDENTIFICATION

AI applications have the potential to improve and enhance (*a*) detection of foodborne pathogens and indicator organisms (e.g., coliforms, *E. coli*); (*b*) classification, identification, and differentiation of bacterial isolates, including foodborne pathogens; and (*c*) characterization of bacteria in terms of AMR, host specificity, and virulence, (*d*) modeling of microbial growth, and (*e*) understanding of population dynamics (**Table 3**). In general, AI applications in these areas have the potential to reduce time, resources, and expertise required for detection and data analyses and will help reveal unintuitive patterns that could be valuable for food safety decision-making. Notably, there is a growing trend in applying image-based analysis to detect bacteria from environmental samples or identify bacterial isolates. Because of the speed of image data collection and maturity of image technology, some of these applications have features that facilitate commercialization and expansion, including (*a*) hardware and software integration, (*b*) large databases of images that support model training and validation, and (*c*) a platform for acquiring new data. Thus, some of these AI applications have potential to be integrated into food safety testing systems.

**Table 3 Potential artificial intelligence (AI) applications in food safety to facilitate microbial data collection**

| Subfields | Specific applications | AI branch | Reference(s) |
|---|---|---|---|
| Pathogen detection | Image-based early detection of pathogens in agar plates | Computer vision | Wang et al. 2020 |
| | Pathogen detection using paper chromogenic array | Computer vision | Yang et al. 2021 |
| Classification, identification, and differentiation of bacterial isolates | Identification and classification of bacterial isolates using Raman spectroscopy | Computer vision | Ho et al. 2019, Sil et al. 2021 |
| | Identification of bacterial isolates using HMI | Computer vision | Kang et al. 2021 |
| | Bacterial classification from microscopic images | Computer vision | Zieliński et al. 2017 |
| | Enhancing species identification via MALDI-TOF MS | Analytical tool | Singhal et al. 2015 |
| | Bacterial identification from digital HRM profiles generated from digital PCR | Analytical tool | Athamanolap et al. 2019 |
| Microbial characterization | Predicting HPI with end-to-end platform | Computer vision | Fisch et al. 2019, 2021 |
| | Predicting AMR | Analytical tool | Hyun et al. 2020, Jamal et al. 2020 |
| | Predicting MIC | Analytical tool | Nguyen et al. 2019, Pataki et al. 2020 |
| | Predicting host specificity | Analytical tool | Lupolova et al. 2017, 2019 |
| | Predicting virulence | Analytical tool | Allen et al. 2021 |
| Modeling microbial growth | Modeling the pathogen growth in various food categories | Analytical tool | Hiura et al. 2021 |
| | Predicting the growth status of pathogens in various media conditions | Analytical tool | Fernández-Navarro et al. 2010 |
| Characterization of bacterial population | Classifying microbial taxa and communities based on metagenomics data | Analytical tool | Ghannam & Techtmann 2021, Harris et al. 2019 |
| | Predicting the microbial interaction from the presence and absence of specific microbial characteristics | Analytical tool | DiMucci et al. 2018 |

Abbreviations: AMR, antimicrobial resistance; HMI, hyperspectral microscopic imaging; HPI, host–pathogen interaction; HRM, high-resolution melt; MALDI-TOF MS, matrix-assisted laser desorption/ionization and time-of-flight mass spectrometry; MIC, minimal inhibitory concentration; PCR, polymerase chain reaction.

## 5.1. Pathogen Detection

One emerging area of AI application to food safety is improved detection of pathogens and indicator organisms. One study used image-based detection methods to take advantage of the fact that bacteria have different morphology in agar plates (Wang et al. 2020). With a combination of image preprocessing and a deep NN algorithm for classification, this detection system was able to detect *E. coli* and total coliform bacteria at the detection limit of 1 colony forming unit (CFU) in less than 9 hours; this approach could easily be adapted to other media and microorganisms, allowing interpretation of plating-based tests substantially before colonies are easily visible, particularly if methods can be developed for different food matrices, media, and plates with high levels of background microflora. AI tools have also been applied to other microbial characterization methods such as paper chromogenic arrays (PCAs), which are composed of various chromogenic dyes that will change color after contact with volatile organic compounds produced by different bacteria. In a proof-of-concept study (Yang et al. 2021), the visual results of PCAs were used as input features to train a multilayer NN to detect the presence of specific pathogens on inoculated romaine lettuce. Although the trained model was able to detect *E. coli* and *L. monocytogenes* from inoculated

**Convolutional neural network (CNN):** a subtype of the deep learning algorithm specializing in processing images and progressively recognizing key visual features

romaine lettuce in the test data set, the detection limit was reported to be around 3 log CFU/g, which is too high for real-world use.

## 5.2. Identification and Classification of Bacterial Isolates

AI-based methodologies have also been applied to the analysis of complex and large data sets to facilitate improved classification, identification, and differentiation of microbial isolates (Ho et al. 2019, Kang et al. 2021, Sil et al. 2021, Weis et al. 2020). For example, several studies have described the use of AI-based image detection methods to analyze spectral information that has been generated for the identification of bacterial isolates, for example, through Raman spectroscopy (Ho et al. 2019, Sil et al. 2021) and hyperspectral microscopic imaging (HMI) (Kang et al. 2021). Kang et al. (2021) showed that using the center region of interest of living cells as the input data set, the HMI method coupled with a recurrent NN was able to classify and differentiate five common foodborne pathogens. Ho et al. (2019) reported that Raman-based methods were able to classify 30 selected pathogen isolates and predict the antibiotic treatment using recommended empiric treatment as the ground truth. In a different study (Sil et al. 2021), Raman spectra of bacterial DNA samples, which provide less noise than whole cell samples, were used to classify 15 species from *Brucella* and *Bacillus* genera. Because the context of pathogen identification is highly variable, efforts to develop larger databases for bacterial identification are essential, which would provide valuable input data for future AI tool development. For example, a database of 660 microscopic images representing 33 bacteria of different genera and species supported texture recognition (e.g., shape, size, spatial arrangement) in microscopic images through a convolutional NN (CNN) with subsequent identification of bacteria via ML algorithms (Zieliński et al. 2017).

AI has also been used to support pathogen characterization and identification by other analytical and diagnostic tools. Weis et al. (2020) summarized various studies that use ML algorithms to analyze the spectral features from matrix-assisted laser desorption/ionization and time-of-flight mass spectrometry (MALDI-TOF MS) (Singhal et al. 2015) to improve species identification. ML has also been used to assist bacterial identification by associating digital high-resolution melt profiles generated from digital PCR with specific bacterial species (Athamanolap et al. 2019). Similar to spectroscopy-based methods, these AI-powered analytical methods need a broader database covering common food pathogens to enhance their applicability to food safety. Currently, the studies reported (Barreiro et al. 2010, Böhme et al. 2010, Hazen et al. 2009) only focused on a few pathogens and products.

## 5.3. Microbial Characterization

When applied as an analytical tool to existing data (e.g., image data and genomic data), AI can also enhance our understanding of pathogen characteristics, including AMR, and host–pathogen interaction (HPI). Although AMR characterization can help inform treatment and control strategies, for example, for animal-associated pathogens (e.g., *Salmonella*), HPI characterization can inform pathogenicity, which may have implications for dose response. Hence, characterization of foodborne pathogens has indirect but relevant implications for food safety. One promising and advanced application in this area is the elucidation of HPI using computer vision. Although ML models (Sen et al. 2016) have been used to study the dynamics of bacterial infection, these models usually need rule-based features that are generated from manual labeling and identification of key phenotypic features from the image. Automated feature extraction in combination with image analysis based on CNN has been shown to reduce bias and improve performance and efficiency (Fisch et al. 2019, 2021). Apart from the prediction accuracy (e.g., in terms of predicting protein recruitment), this system exemplifies a mature AI technology in that it is automated, user-friendly, and applicable to other pathogens with an open platform to facilitate the data collection from users. It is important to note that the CNN algorithm used in this system is not novel, as it benefits from transfer learning that allows pretrained NN to be customizable to other applications. With this in mind, we believe that a similar research approach can be applied to other domains of pathogen characterization.

Although state-of-the-art deep learning algorithms revolutionized our analysis of image and spectral data, ML algorithms are still most commonly used to identify attributes of foodborne pathogens from genetic data (e.g., WGS data). Applications that leverage genetic data include prediction of AMR (Hyun et al. 2020, Jamal et al. 2020), minimal inhibitory concentration (Nguyen et al. 2019, Pataki et al. 2020), host specificity (Lupolova et al. 2017, 2019), and virulence (Allen et al. 2021). Tree-based methods, including XGBoost, AdaBoost, and RF, perform better with strong predictors and therefore are commonly selected algorithms because the input data (e.g., k-mer representation of genome and expression of specific genes) are categorical. Findings from these studies enhanced our understanding of associations between genotypic and phenotypic features, which can be used to characterize new foodborne pathogen isolates solely based on WGS data. Future development of automated systems integrating gene and genomic databases and ML algorithms will make this type of WGS-based pathogen characterization more accessible, including to individuals without coding knowledge.

### 5.4. Modeling Microbial Growth

Challenges with modeling of pathogen growth, as well as survival and die-off, are numerous and related to the fact that it is important to predict growth, survival, and die-off for several different pathogens (including possibly different pathogen subtypes) in a variety of different foods and under a variety of different environmental conditions (e.g., temperatures). ComBase (**https://www.combase.cc/index.php/en/**) is a database that collects fundamental research results about bacterial growth patterns under different food matrix and environmental conditions. Utilizing ComBase as a data source, Hiura et al. (2021) reported that ML models achieved accurate prediction of *L. monocytogenes* levels within 1 log CFU regardless of the food category, suggesting that more extensive usage of such models with other pathogens and food products is possible with the support of big data. AI can also be used to predict whether pathogens can grow under certain conditions (growth/no growth modeling), which is important for many food safety applications. Using a multiclassification model, a study was able to classify microorganisms in media under various growth conditions into growth, growth transition, and no growth (Fernández-Navarro et al. 2010). Overall, these examples illustrate how AI can be used to support decision-making with regard to pathogen control strategies, including through "food safety by design" approaches, where AI tools would be used to support development of formulations that minimize pathogen growth.

### 5.5. Studying Bacterial Population Dynamics

There are many examples of AI application to support modeling and characterization of microbial populations and population dynamics, even though most examples are from fields other than food safety. Not surprisingly, there are several publications that detail AI-based approaches to classify microbial taxa and communities based on metagenomics data (Ghannam & Techtmann 2021, Harris et al. 2019). Although 16S metagenomics tools, which until recently were the predominant metagenomics approaches used in food microbiology, are considered to have limited value for food safety (as they typically cannot reliably differentiate foodborne pathogens from closely related nonpathogens), increasing application of shotgun metagenomics, including AI-based analyses of the resulting data, is likely to broaden food safety relevant metagenomics applications. Another research direction to study population dynamics is to predict the interaction between microorganisms in a community from various traits. Using three microbial communities as experimental examples, DiMucci et al. (2018) predicted the microbial interactions of the microorganisms in each community from the presence and absence of specific genes or metabolic functions observed in the communities. This ML-assisted approach was suggested to have potential for discovering therapeutic interventions. In the context of food safety, this application might be useful to predict the effect of biopreservatives used as control strategies for foodborne pathogens.

## 6. CHALLENGES AND FUTURE DIRECTIONS FOR ARTIFICIAL INTELLIGENCE TO HELP IMPROVE FOOD SAFETY

Although in previous sections, we explored different opportunities AI technologies present for the improvement of food safety, application of AI tools in food safety still appears to lag behind compared to other food-related areas (e.g., marketing, agricultural production). The relatively low penetration of AI in food safety appears to be due to multiple factors. One key reason seems to be the limited availability of data that are needed to develop and implement AI tools for food safety applications with key limitations being (*a*) the low speed and high cost of microbial data collection and (*b*) limited sharing of microbial data, owing to industry concerns about data privacy and the business and reputational risk that may be associated with data sharing. In some cases, there may be a reluctance from food industry stakeholders to develop and adopt food safety AI technologies because of a fear that these tools might negatively impact their business interests. This concern is not just limited to the data sharing that may be required but may also extend to concerns that AI tools, if available to others, could be used to predict whether a specific company (e.g., the company that provided the data for the tool development) has an increased food safety risk. The lack of a clear legal and regulatory framework regarding AI applications and protection of sensitive data that are needed to feed these applications may compound these concerns. Even if food safety data can be shared, data from different sources often represent distinct formats and structures, which makes standardization and data use difficult. Also, data details, such as the test methods used and the sample size tested (e.g., 1 g, 25 g), are often either inconsistent or not available. Negative test results (i.e., results that do not detect a pathogen) may never be available at all or may show even less metadata.

Another challenge to more widespread AI application is the lack of systematic efforts to assemble different AI tools. Given that many food safety AI tools are developed by research labs, they tend to fall short of delivering practical products that industry can adopt or that at least can be quickly developed into usable tools. Furthermore, the problems addressed by "research" AI tools are typically specific to a type of product, microorganism, or environmental condition; hence, the generalizability is limited and thus those AI tools need to be combined for holistic benefit.

To address the challenges mentioned above, we propose that key directions for future AI development in the field of food safety include (*a*) improving data sharing among different stakeholders through novel privacy preservation technologies, data standardization, and regulations, (*b*) encouraging multidisciplinary collaboration to encourage more commercial deliverables, and (*c*) implementing education in data science, software development, and design thinking among food safety professionals

Innovations in privacy protection methods have the potential to improve the utility and scalability of various AI tools by facilitating more widespread data sharing, as accessibility to big data is the key to AI applications. Two of the most promising approaches are differential privacy and federated learning. Differential privacy allows data providers or data processors to intentionally manipulate data by adding noise, which grants the data provider plausible deniability. Different levels of privacy can be customized depending on the sensitivity of the data and the trust level of data providers. Relevant to the area of food safety, differential privacy has been applied to protect personal health data to facilitate COVID-19 surveillance (Dyda et al. 2021). At the same time, federated learning is presented as the emerging tool in medical diagnostics (Rieke et al. 2020), which shares similar challenges regarding data sharing as the food safety field. Federated learning allows data providers to share local updates to the model on the cloud without sharing the data. Thus, the model applications can be trained without the need to store data at a central location. With the ability to utilize but not compromise patient data, federated learning is a promising approach that can aid in the transition of research models to practical tools in clinics. Future research and development efforts should apply these novel privacy-preservation methods to food safety data; these efforts could provide use cases for industry partners and insights into the sample size needed (i.e., a total of collectively shared data) to ensure acceptable utility of AI applications.

To address the industry concerns about the misuse of AI applications, laws and regulations need to keep up with the advancement of technology to reflect and avoid potential misuse. For example, templates of mutual agreements should be developed to better delineate the boundary for interpreting results from AI applications. A food company might be more inclined to adopt AI technology if the contamination predicted from the model is used for internal decision-making only and cannot be used by government and supply chain partners against the company's benefits. Digital solution providers for food safety should help food companies establish protocols and evaluate enterprise risks in the case of a data leak or model misuse. These strategies can help food companies choose the type of AI technologies they want to use, depending on the level of risk they are willing to accept.

Importantly, however, data standardization still needs to be developed and implemented across the industry to facilitate data sharing by limiting the time and effort to unify the data. This process could be facilitated with the help of a third-party organization, for example, Global Language of Business (**https://www.gs1.org/**), which is an organization that creates a common language for different systems, or through collaboration between different stakeholders, as exemplified by Coordinated Innovation Network from Dairy Brain (Ferris et al. 2020). Currently, many organizations, including the FDA, have already directed efforts toward standardizing metadata associated with WGS data (Balkey et al. 2021, Pettengill et al. 2021).

To address the lack of deliverable AI tools in food safety, researchers need to establish interdisciplinary collaboration with computer scientists, software developers, and IT professionals to develop user-centered application programming interfaces and data management systems to support commercialization or implementation of food safety AI analytical tools. Improved education and literacy in data, computer science, and design thinking are also needed to help food safety professionals better communicate across disciplines and express the specific industry needs. With an extensive library of ML models currently developed in academia, it is not difficult to envision the integration of these tools into commercial AI applications for food safety. For example, with a connection to a public geographical database and implementation of water sensors in irrigation water and rapid detection equipment for indicator microorganisms, an AI system can be powered by the Internet of Things to provide early warnings for increased risk of pathogen presence in agricultural water, fields, and products, informing different practices such as testing, irrigation source water selection, harvest timing, and harvest equipment sanitation frequency. Further governmental support is also needed to facilitate the development and implementation of AI technologies in food safety. Similar to the FDA's regulation of AI-based medical devices (Benjamens et al. 2020), clearer regulations with regard to clearance levels will help the commercialization of current existing food safety AI.

## 7. CONCLUSIONS

Although it is tempting to conclude that AI shows promise for application in food safety, so far there are only a few convincing examples that would support this statement. Most data and examples that provide strong evidence for successful application of AI to food safety are in the areas of public health surveillance, prediction of preharvest food safety risk factors, and, to a lesser extent, detection, identification, and characterization of foodborne pathogens. Key obstacles to wider commercialization and implementation of AI tools for food safety applications include limited availability of digitized food safety data (food safety data are still often collected and stored in hard copy), privacy concerns (most food companies hesitate to store food safety data on the cloud), a workforce that often has limited data literacy, and a limited food safety innovation ecosystem that could stimulate the collaborations needed to foster food systems' wide implementation of digital food safety tools. Overcoming these challenges will require considerable efforts, including an improved digital food safety infrastructure, implementation of innovative data sharing and privacy protection approaches, and improved digital training of food safety professionals. Importantly, industry as well as AI researchers and service providers will also need to be careful to not present AI as a panacea but to focus on those applications where AI will address specific needs and questions, using other more appropriate tools (including simulations, classical statistics, etc.) as needed. Finally, despite its alluring name, in the medium term, AI tools are likely to help provide decision-support tools, and final decisions will still be made by humans.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## LITERATURE CITED

Allen JP, Snitkin E, Pincus NB, Hauser AR. 2021. Forest and trees: exploring bacterial virulence with genome-wide association studies and machine learning. *Trends Microbiol.* 29(7):621–33

Arning N, Sheppard SK, Bayliss S, Clifton DA, Wilson DJ. 2021. Machine learning to predict the source of campylobacteriosis using whole genome data. *PLOS Genet.* 17(10):e1009436

Astill J, Dara RA, Fraser EDG, Roberts B, Sharif S. 2020. Smart poultry management: smart sensors, big data, and the internet of things. *Comput. Electron. Agric.* 170:105291

Athamanolap P, Hsieh K, O'Keefe CM, Zhang Y, Yang S, Wang T-H. 2019. Machine learning-assisted digital PCR and melt enables broad bacteria identification and pheno-molecular antimicrobial susceptibility test. bioRxiv 587543. **https://doi.org/10.1101/587543**

Balkey M, Batz M, Gopinath G, Gosal G, Griffiths E, Tate H. 2021. *Standardizing the isolation source metadata for the genomic epidemiology of foodborne pathogens using LexMapr*. Poster presented at FDA Science Forum, May 26

Barco L, Barrucci F, Olsen JE, Ricci A. 2013. *Salmonella* source attribution based on microbial subtyping. *Int. J. Food Microbiol.* 163(2–3):193–203

Barnett-Neefs C, Sullivan G, Zoellner C, Wiedmann M, Ivanek R. 2022. Using agent-based modeling to compare corrective actions for *Listeria* contamination in produce packinghouses. *PLOS ONE* 17(3):e0265251

Barreiro JR, Ferreira CR, Sanvido GB, Kostrzewa M, Maier T, et al. 2010. Short communication: identification of subclinical cow mastitis pathogens in milk by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Dairy Sci.* 93(12):5661–67

Belias A, Brassill N, Roof S, Rock C, Wiedmann M, Weller D. 2021. Cross-validation indicates predictive models may provide an alternative to indicator organism monitoring for evaluating pathogen presence in southwestern US agricultural water. *Front. Water* 3:693631

Benjamens S, Dhunnoo P, Meskó B. 2020. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digit. Med.* 3:118

Böhme K, Fernández-No IC, Barros-Velázquez J, Gallardo JM, Calo-Mata P, Cañas B. 2010. Species differentiation of seafood spoilage and pathogenic gram-negative bacteria by MALDI-TOF mass fingerprinting. *J. Proteome Res.* 9(6):3169–83

Bouzembrak Y, Marvin HJP. 2016. Prediction of food fraud type using data from Rapid Alert System for Food and Feed (RASFF) and Bayesian network modelling. *Food Control* 61:180–87

Carvalho TP, Soares FAAMN, Vita R, da P Francisco R, Basto JP, Alcalá SGS. 2019. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* 137:106024

CDC (Cent. Dis. Control). 2021. FoodNet Fast. *Centers for Disease Control.* **https://wwwn.cdc.gov/foodnetfast/**

Chaudhuri A, Dukovska-Popovska I, Subramanian N, Chan HK, Bai R. 2018. Decision-making in cold chain logistics using data analytics: a literature review. *Int. J. Logist. Manag.* 29(3):839–61

Chen JX. 2016. The evolution of computing: AlphaGo. *Comput. Sci. Eng.* 18(4):4–7

Chen TC, Yu SY. 2022. The review of food safety inspection system based on artificial intelligence, image processing, and robotic. *Food Sci. Technol.* 42:e35421

Copeland BJ. 2021. Artificial intelligence. *Encyclopedia Britannica*. **https://www.britannica.com/technology/artificial-intelligence**

Defraeye T, Shrivastava C, Berry T, Verboven P, Onwude D, et al. 2021. Digital twins are coming: Will we need them in supply chains of fresh horticultural produce? *Trends Food Sci. Technol.* 109:245–58

Deng X, Cao S, Horn AL. 2021. Emerging applications of machine learning in food safety. *Annu. Rev. Food Sci. Technol.* 12:513–38

DiMucci D, Kon M, Segrè D. 2018. Machine learning reveals missing edges and putative interaction mechanisms in microbial ecosystem networks. *mSystems* 3(5):e00181-18

Du Y, Guo Y. 2022. Machine learning techniques and research framework in foodborne disease surveillance system. *Food Control* 131:108448

Duarte ASR, Röder T, Van Gompel L, Petersen TN, Hansen RB, et al. 2021. Metagenomics-based approach to source-attribution of antimicrobial resistance determinants: identification of reservoir resistome signatures. *Front. Microbiol.* 11:3447

Dyda A, Purcell M, Curtis S, Field E, Pillai P, et al. 2021. Differential privacy for public health data: an innovative tool to optimize information sharing while protecting data confidentiality. *Patterns* 2(12):100366

Effland T, Lawson A, Balter S, Devinney K, Reddy V, et al. 2018. Discovering foodborne illness in online restaurant reviews. *J. Am. Med. Inform. Assoc.* 25(12):1586–92

Escrig J, Woolley E, Rangappa S, Simeone A, Watson NJ. 2019. Clean-in-place monitoring of different food fouling materials using ultrasonic measurements. *Food Control* 104:358–66

Escrig J, Woolley E, Simeone A, Watson NJ. 2020a. Monitoring the cleaning of food fouling in pipes using ultrasonic measurements and machine learning. *Food Control* 116:107309

Escrig JE, Simeone A, Woolley E, Rangappa S, Rady A, Watson NJ. 2020b. Ultrasonic measurements and machine learning for monitoring the removal of surface fouling during clean-in-place processes. *Food Bioprod. Process.* 123:1–13

Eur. Comm. 2018. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial intelligence for Europe. *Eur-Lex*. **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN%0A**

FDA (US Food Drug Adm.) 2022a. Import screening pilot unleashes the power of data and leverages artificial intelligence. *FDA*. **https://www.fda.gov/news-events/fda-voices/import-screening-pilot-unleashes-power-data-and-leverages-artificial-intelligence**

FDA (US Food Drug Adm.). 2022b. TechTalk podcast episode 3: artificial intelligence in the new era of smarter food safety. *FDA*. **https://www.fda.gov/food/new-era-smarter-food-safety-techtalk-podcast/techtalk-podcast-episode-3-artificial-intelligence-new-era-smarter-food-safety**

Feng YZ, Sun DW. 2012. Application of hyperspectral imaging in food safety inspection and control: a review. *Crit. Rev. Food Sci. Nutr.* 52(11):1039–58

Fernández-Navarro F, Valero A, Hervás-Martínez C, Gutiérrez PA, García-Gimeno RM, Zurera-Cosano G. 2010. Development of a multi-classification neural network model to determine the microbial growth/no growth interface. *Int. J. Food Microbiol.* 141(3):203–12

Ferris MC, Christensen A, Wangen SR. 2020. Symposium review: Dairy Brain—informing decisions on dairy farms using data analytics. *J. Dairy Sci.* 103(4):3874–81

Fisch D, Evans R, Clough B, Byrne SK, Channell WM, et al. 2021. HRMAn 2.0: next-generation artificial intelligence-driven analysis for broad host-pathogen interactions. *Cell. Microbiol.* 23(7):e13349

Fisch D, Yakimovich A, Clough B, Wright J, Bunyan M, et al. 2019. Defining host-pathogen interactions employing an artificial intelligence workflow. *eLife* 8:e40560

Friedlander A, Zoellner C. 2020. Artificial intelligence opportunities to improve food safety at retail. *Food Prot. Trends* 40(4):272–78

Garver K. 2018. 6 examples of artificial intelligence in the food industry. *Food Industry Executive*. **https://foodindustryexecutive.com/2018/04/6-examples-of-artificial-intelligence-in-the-food-industry**

Ghannam RB, Techtmann SM. 2021. Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring. *Comput. Struct. Biotechnol. J.* 19:1092–107

Golden CE, Rothrock MJ, Mishra A. 2019a. Comparison between random forest and gradient boosting machine methods for predicting *Listeria* spp. prevalence in the environment of pastured poultry farms. *Food Res. Int.* 122:47–55

Golden CE, Rothrock MJ, Mishra A. 2019b. Using farm practice variables as predictors of *Listeria* spp. prevalence in pastured poultry farms. *Front. Sustain. Food Syst.* 3:15

Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L, Ellis JT. 2021. Machine learning and applications in microbiology. *FEMS Microbiol. Rev.* 45(5):fuab015

Gupta A, Anpalagan A, Guan L, Khwaja AS. 2021. Deep learning for object detection and scene perception in self-driving cars: survey, challenges, and open issues. *Array* 10:100057

Haiminen N, Edlund S, Chambliss D, Kunitomi M, Weimer BC, et al. 2019. Food authentication from shotgun sequencing reads with an application on high protein powders. *npj Sci. Food.* 3:24

Harrand AS, Strawn LK, Illas-Ortiz PM, Wiedmann M, Weller DL. 2020. *Listeria monocytogenes* prevalence varies more within fields than between fields or over time on conventionally farmed new york produce fields. *J. Food Prot.* 83(11):1958–66

Harris ZN, Dhungel E, Mosior M, Ahn TH. 2019. Massive metagenomic data analysis using abundance-based machine learning. *Biol. Direct* 14:12

Harrison C, Jorder M, Stern H, Stavinsky F, Reddy V, et al. 2014. Using online reviews by restaurant patrons to identify unreported cases of foodborne illness—New York City, 2012–2013. *MMWR* 63(20):441–45

Harrison L, Mukherjee S, Hsu CH, Young S, Strain E, et al. 2021. Core genome MLST for source attribution of *Campylobacter coli*. *Front. Microbiol.* 12:703890

Hazen TH, Martinez RJ, Chen Y, Lafon PC, Garrett NM, et al. 2009. Rapid identification of *Vibrio parahaemolyticus* by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl. Environ. Microbiol.* 75(21):6745–56

Healthy People. 2020. Healthy people 2020 objectives and topics: food safety. *Healthy People*. **https://wayback.archive-it.org/5774/20220415004842/https://www.healthypeople.gov/2020/data-search/Search-the-Data?nid=4478**

Henrichs E, Noack T, Piedrahita AMP, Salem MA, Stolz J, Krupitzer C. 2022. Can a byte improve our bite? An analysis of digital twins in the food industry. *Sensors* 22(1):115

Hiura S, Koseki S, Koyama K. 2021. Prediction of population behavior of *Listeria monocytogenes* in food using machine learning and a microbial growth and survival database. *Sci. Rep.* 11:10613

Ho CS, Jean N, Hogan CA, Blackmon L, Jeffrey SS, et al. 2019. Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning. *Nat. Commun.* 10:4927

Hyun JC, Kavvas ES, Monk JM, Palsson BO. 2020. Machine learning with random subspace ensembles identifies antimicrobial resistance determinants from pan-genomes of three pathogens. *PLOS Comput. Biol.* 16(3):e1007608

Ivanek R, Gröhn YT, Wells MT, Lembo AJ, Sauders BD, Wiedmann M. 2009. Modeling of spatially referenced environmental and meteorological factors influencing the probability of *Listeria* species isolation from natural environments. *Appl. Environ. Microbiol.* 75(18):5893–909

Jamal S, Khubaib M, Gangwar R, Grover S, Grover A, Hasnain SE. 2020. Artificial intelligence and machine learning based prediction of resistant and susceptible mutations in *Mycobacterium tuberculosis*. *Sci. Rep.* 10:5487

Kakani V, Nguyen VH, Kumar BP, Kim H, Pasupuleti VR. 2020. A critical review on computer vision and artificial intelligence in food industry. *J. Agric. Food Res.* 2:100033

Kamble SS, Gunasekaran A, Gawankar SA. 2020. Achieving sustainable performance in a data-driven agriculture supply chain: a review for research and applications. *Int. J. Prod. Econ.* 219:179–94

Kamilaris A, Prenafeta-Boldú FX. 2018. Deep learning in agriculture: a survey. *Comput. Electron. Agric.* 147:70–90

Kang R, Park B, Ouyang Q, Ren N. 2021. Rapid identification of foodborne bacteria with hyperspectral microscopic imaging and artificial intelligence classification algorithms. *Food Control* 130:108379

Kondakci T, Zhou W. 2016. Recent applications of advanced control techniques in food industry. *Food Bioprocess Technol.* 10(3):522–42

Kuehn BM. 2014. Agencies use social media to track foodborne illness. *JAMA* 312(2):117–18

Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. 2018. Machine learning in agriculture: a review. *Sensors* 18(8):2674

Loisel J, Duret S, Cornuéjols A, Cagnon D, Tardet M, et al. 2021. Cold chain break detection and analysis: Can machine learning help? *Trends Food Sci. Technol.* 112:391–99

Lupolova N, Dallman TJ, Holden NJ, Gally DL. 2017. Patchy promiscuity: machine learning applied to predict the host specificity of *Salmonella enterica* and *Escherichia coli*. *Microb. Genom.* 3(10):e000135

Lupolova N, Lycett SJ, Gally DL. 2019. A guide to machine learning for bacterial host attribution using genome sequence data. *Microb. Genom.* 5(12):e000317

Maharana A, Cai K, Hellerstein J, Hswen Y, Munsell M, et al. 2019. Detecting reports of unsafe foods in consumer product reviews. *JAMIA Open* 2(3):330–38

Martos V, Ahmad A, Cartujo P, Ordoñez J. 2021. Ensuring agricultural sustainability through remote sensing in the era of agriculture 5.0. *Appl. Sci.* 11(13):5911

Mavani NR, Ali JM, Othman S, Hussain MA, Hashim H, Rahman NA. 2021. Application of artificial intelligence in food industry—a guideline. *Food Eng. Rev.* 14(1):134–75

Mei X, Lee HC, Diao K-Y, Huang M, Lin B, et al. 2020. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat. Med.* 26(8):1224–28

Mercier S, Villeneuve S, Mondor M, Uysal I. 2017. Time-temperature management along the food cold chain: a review of recent developments. *Compr. Rev. Food Sci. Food Saf.* 16(4):647–67

Mikkelä A, Ranta J, Tuominen P. 2019. A modular Bayesian *Salmonella* source attribution model for sparse data. *Risk Anal.* 39(8):1796–811

Misra NN, Dixit Y, Al-Mallahi A, Bhullar MS, Upadhyay R, Martynenko A. 2020. IoT, big data and artificial intelligence in agriculture and food industry. *IEEE Internet Things J.* 9(9):6305–24

Munck N, Njage PMK, Leekitcharoenphon P, Litrup E, Hald T. 2020. Application of whole-genome sequences and machine learning in source attribution of *Salmonella* Typhimurium. *Risk Anal.* 40(9):1693–705

Murphy SI, Reichler SJ, Martin NH, Boor KJ, Wiedmann M. 2021. Machine learning and advanced statistical modeling can identify key quality management practices that affect postpasteurization contamination of fluid milk. *J. Food Prot.* 84(9):1496–511

Nasirahmadi A, Hensel O. 2022. Toward the next generation of digitalization in agriculture based on digital twin paradigm. *Sensors* 22(2):498

Ndraha N, Hsiao HI, Hsieh YZ, Pradhan AK. 2021. Predictive models for the effect of environmental factors on the abundance of *Vibrio parahaemolyticus* in oyster farms in Taiwan using extreme gradient boosting. *Food Control* 130:108353

Nguyen M, Wesley Long S, McDermott PF, Olsen RJ, Olson R, et al. 2019. Using machine learning to predict antimicrobial MICs and associated genomic features for nontyphoidal *Salmonella*. *J. Clin. Microbiol.* 57(2):e01260-18

Nturambirwe JFI, Opara UL. 2020. Machine learning applications to non-destructive defect detection in horticultural products. *Biosyst. Eng.* 189:60–83

Oldroyd RA, Morris MA, Birkin M. 2021. Predicting food safety compliance for informed food outlet inspections: a machine learning approach. *Int. J. Environ. Res. Public Health* 18(23):12635

Onoufriou G, Bickerton R, Pearson S, Leontidis G. 2019. Nemesyst: a hybrid parallelism deep learning-based framework applied for internet of things enabled food retailing refrigeration systems. *Comput. Ind.* 113:103133

Park S, Navratil S, Gregory A, Bauer A, Srinath I, et al. 2014. Farm management, environment, and weather factors jointly affect the probability of spinach contamination by generic *Escherichia coli* at the preharvest stage. *Appl. Environ. Microbiol.* 80(8):2504–15

Pataki BÁ, Matamoros S, van der Putten BCL, Remondini D, Giampieri E, et al. 2020. Understanding and predicting ciprofloxacin minimum inhibitory concentration in *Escherichia coli* with machine learning. *Sci. Rep.* 10:15026

Pettengill JB, Beal J, Balkey M, Allard M, Rand H, Timme R. 2021. Interpretative labor and the bane of nonstandardized metadata in public health surveillance and food safety. *Clin. Infect. Dis.* 73(8):1537–39

Pires SM, Evers EG, Van Pelt W, Ayers T, Scallan E, et al. 2009. Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathog. Dis.* 6(4):417–24

Polat H, Topalcengiz Z, Danyluk MD. 2020. Prediction of *Salmonella* presence and absence in agricultural surface waters by artificial intelligence approaches. *J. Food Saf.* 40(1):e12733

Pouillot R, Klontz KC, Chen Y, Burall LS, Macarisin D, et al. 2016. Infectious dose of *Listeria monocytogenes* in outbreak linked to ice cream, United States, 2015. *Emerg. Infect. Dis.* 22(12):2113–19

Rajan K, Saffiotti A. 2017. Towards a science of integrated AI and robotics. *Artif. Intell.* 247:1–9

Ranjan AA, Rai A, Haque S, Pradeep K, Kushwaha BPL. 2019. An approach for Netflix recommendation system using singular value decomposition. *J. Comput. Math. Sci.* 10(4):774–79

Ranta J, Matjushin D, Virtanen T, Kuusi M, Viljugrein H, et al. 2011. Bayesian temporal source attribution of foodborne zoonoses: *Campylobacter* in Finland and Norway. *Risk Anal.* 31(7):1156–71

Rieke N, Hancox J, Li W, Milletarì F, Roth HR, et al. 2020. The future of digital health with federated learning. *npj Digit. Med.* 3:119

Ropodi AI, Panagou EZ, Nychas GJE. 2016. Data mining derived from food analyses using non-invasive/non-destructive analytical techniques; determination of food authenticity, quality & safety in tandem with computer science disciplines. *Trends Food Sci. Technol.* 50:11–25

Sadilek A, Caty S, DiPrete L, Mansour R, Schenk T, et al. 2018. Machine-learned epidemiology: real-time detection of foodborne illness at scale. *npj Digit. Med.* 1:36

Sadilek A, Kautz H, Di Prete L, Labus B, Portman E, et al. 2017. Deploying nemesis: preventing foodborne illness by data mining social media. *AI Mag.* 38(1):37–48

Sen R, Nayak L, De RK. 2016. A review on host-pathogen interactions: classification and prediction. *Eur. J. Clin. Microbiol. Infect. Dis.* 35(10):1581–99

Sharma R, Kamble SS, Gunasekaran A, Kumar V, Kumar A. 2020. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Comput. Oper. Res.* 119:104926

Sil S, Mukherjee R, Kumbhar D, Reghu D, Shrungar D, et al. 2021. Raman spectroscopy and artificial intelligence open up accurate detection of pathogens from DNA-based sub-species level classification. *J. Raman Spectrosc.* 52(12):2648–59

Simeone A, Woolley E, Escrig J, Watson NJ. 2020. Intelligent industrial cleaning: a multi-sensor approach utilising machine learning-based regression. *Sensors* 20(13):3642

Singhal N, Kumar M, Kanaujia PK, Virdi JS. 2015. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* 6:791

Sinwar D, Dhaka VS, Sharma MK, Rani G. 2020. AI-based yield prediction and smart irrigation. In *Internet of Things and Analytics for Agriculture*, Vol. 2, ed. PK Pattnaik, R Kumar, S Pal, pp. 155–80. Singapore: Springer

Smith MJ. 2019. Getting value from artificial intelligence in agriculture. *Anim. Prod. Sci.* 60(1):46–54

Song B, Yan W, Zhang T. 2019. Cross-border e-commerce commodity risk assessment using text mining and fuzzy rule-based reasoning. *Adv. Eng. Inform.* 40:69–80

Stocker MD, Pachepsky YA, Hill RL. 2022. Prediction of *E. coli* concentrations in agricultural pond waters: application and comparison of machine learning algorithms. *Front. Artif. Intell.* 4:768650

Strawn LK, Fortes ED, Bihn EA, Nightingale KK, Gröhn YT, et al. 2013. Landscape and meteorological factors affecting prevalence of three food-borne pathogens in fruit and vegetable farms. *Appl. Environ. Microbiol.* 79(2):588–600

Sullivan G, Zoellner C, Wiedmann M, Ivanek R. 2021. In silico models for design and optimization of science-based *Listeria* environmental monitoring programs in fresh-cut produce facilities. *Appl. Environ. Microbiol.* 87(21):e0079921

Tack DM, Ray L, Griffin PM, Cieslak PR, Dunn J, et al. 2020. Preliminary incidence and trends of infections with pathogens transmitted commonly through food—foodborne diseases active surveillance network, 10 U.S. Sites, 2016–2019. *MMWR* 69(17):509–14

Teyhouee A, McPhee-Knowles S, Waldner C, Osgood N. 2017. Prospective detection of foodborne illness outbreaks using machine learning approaches. *Lect. Notes Comput. Sci.* 10354:302–8

Thépault A, Méric G, Rivoal K, Pascoe B, Mageiros L, et al. 2017. Genome-wide identification of host-segregating epidemiological markers for source attribution in *Campylobacter jejuni*. *Appl. Environ. Microbiol.* 83(7):e03085-16

Toro M, Weller D, Ramos R, Diaz L, Alvarez FP, et al. 2022. Environmental and anthropogenic factors associated with the likelihood of detecting *Salmonella* in agricultural watersheds. *Environ. Pollut.* 306:119298

Tyson GH, Tate HP, Abbott J, Tran TT, Kabera C, et al. 2016. Molecular subtyping and source attribution of *Campylobacter* isolated from food animals. *J. Food Prot.* 79(11):1891–97

Úbeda MA, Hussein WB, Hussein MA, Hinrichs J, Becker TM. 2016. Acoustic sensing and signal processing techniques for monitoring milk fouling cleaning operations. *Eng. Life Sci.* 16(1):67–77

Vangay P, Steingrimsson J, Wiedmann M, Stasiewicz MJ. 2014. Classification of *Listeria monocytogenes* persistence in retail delicatessen environments using expert elicitation and machine learning. *Risk Anal.* 34(10):1830–45

Wallhäußer E, Hussein WB, Hussein MA, Hinrichs J, Becker TM. 2011. On the usage of acoustic properties combined with an artificial neural network: a new approach of determining presence of dairy fouling. *J. Food Eng.* 103(4):449–56

Wallhäußer E, Hussein WB, Hussein MA, Hinrichs J, Becker T. 2013. Detection of dairy fouling: combining ultrasonic measurements and classification methods. *Eng. Life Sci.* 13(3):292–301

Wang H, Ceylan Koydemir H, Qiu Y, Bai B, Zhang Y, et al. 2020. Early detection and classification of live bacteria using time-lapse coherent imaging and deep learning. *Light Sci. Appl.* 9:118

Wang H, Cui W, Guo Y, Du Y, Zhou Y. 2021. Machine learning prediction of foodborne disease pathogens: algorithm development and validation study. *JMIR Med. Inform.* 9(1):e24924

Wang J, Yue H, Zhou Z. 2017. An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network. *Food Control* 79:363–70

Wang L, Ting JSL, Ip WH. 2013. Design of Supply-chain Pedigree Interactive Dynamic Explore (SPIDER) for food safety and implementation of Hazard Analysis and Critical Control Points (HACCPs). *Comput. Electron. Agric.* 90:14–23

Weis CV, Jutzeler CR, Borgwardt K. 2020. Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. *Clin. Microbiol. Infect.* 26(10):1310–17

Weller D, Brassill N, Rock C, Ivanek R, Mudrak E, et al. 2020a. Complex interactions between weather, and microbial and physicochemical water quality impact the likelihood of detecting foodborne pathogens in agricultural water. *Front. Microbiol.* 11:134

Weller DL, Love TMT, Belias A, Wiedmann M. 2020b. Predictive models may complement or provide an alternative to existing strategies for assessing the enteric pathogen contamination status of northeastern streams used to provide water for produce production. *Front. Sustain. Food Syst.* 4:561517

Weller DL, Love TMT, Wiedmann M. 2021a. Comparison of resampling algorithms to address class imbalance when developing machine learning models to predict foodborne pathogen presence in agricultural water. *Front. Environ. Sci.* 9:701288

Weller DL, Love TMT, Wiedmann M. 2021b. Interpretability versus accuracy: a comparison of machine learning models built using different algorithms, performance measures, and features to predict *E. coli* levels in agricultural water. *Front. Artif. Intell.* 4:628441

WHO (World Health Organ.). 2015. *WHO Estimates of the Global Burden of Foodborne Diseases*. Geneva: WHO

WHO (World Health Organ.). 2021. WHO steps up action to improve food safety and protect people from disease. *WHO*. **https://www.who.int/news/item/07-06-2021-who-steps-up-action-to-improve-food-safety-and-protect-people-from-disease**

Yang C-C, Jun W, Kim MS, Chao K, Kang S, et al. 2010. Classification of fecal contamination on leafy greens by hyperspectral imaging. *Sens. Agric. Food Qual. Saf. II*. **https://doi.org/10.1117/12.851069**

Yang J, Jensen BBB, Nordkvist M, Rasmussen P, Pedersen B, et al. 2018. Anomaly analysis in cleaning-in-place operations of an industrial brewery fermenter. *Ind. Eng. Chem. Res.* 57(38):12871–83

Yang M, Liu X, Luo Y, Pearlstein AJ, Wang S, et al. 2021. Machine learning-enabled non-destructive paper chromogenic array detection of multiplexed viable pathogens on food. *Nat. Food* 2(2):110–17

Zhang P, Cui W, Wang H, Du Y, Zhou Y. 2021a. High-efficiency machine learning method for identifying foodborne disease outbreaks and confounding factors. *Foodborne Pathog. Dis.* 18(8):590–98

Zhang S, Li S, Gu W, Den Bakker H, Boxrud D, et al. 2019. Zoonotic source attribution of *Salmonella enterica* serotype Typhimurium using genomic surveillance data, United States. *Emerg. Infect. Dis.* 25(1):82–91

Zhang W, Valencia A, Chang N-B. 2021b. Synergistic integration between machine learning and agent-based modeling: a multidisciplinary review. *IEEE Trans. Neural Netw. Learn. Syst.* In press

Zhou L, Zhang C, Liu F, Qiu Z, He Y. 2019. Application of deep learning in food: a review. *Compr. Rev. Food Sci. Food Saf.* 18(6):1793–811

Zieliński B, Plichta A, Misztal K, Spurek P, Brzychczy-Włoch M, Ochońska D. 2017. Deep learning approach to bacterial colony classification. *PLOS ONE* 12(9):e0184554

Zoellner C, Jennings R, Wiedmann M, Ivanek R. 2019. EnABLe: an agent-based model to understand *Listeria* dynamics in food processing facilities. *Sci. Rep.* 9:495