The Oklahoma Cooperative Extension Service
Bringing the University to You!

The Cooperative Extension Service is the largest, most successful informal educational organization in the world. It is a nationwide system funded and guided by a partnership of federal, state, and local governments that delivers information to help people help themselves through the land-grant university system.

Extension carries out programs in the broad categories of agriculture, natural resources and environment; home economics; 4-H and other youth; and community resource development. Extension staff members live and work among the people they serve to help stimulate and educate Americans to plan ahead and cope with their problems.

Some characteristics of Cooperative Extension are:

- The federal, state, and local governments cooperatively share in its financial support and program direction.
- It is administered by the land-grant university as designated by the state legislature through an Extension director.
- Extension programs are nonpolitical, objective, and based on factual information.
- It provides practical, problem-oriented education for people of all ages. It is designated to take the knowledge of the university to those persons who do not or cannot participate in the formal classroom instruction of the university.
- It utilizes research from university, government, and other sources to help people make their own decisions.
- More than a million volunteers help multiply the impact of the Extension professional staff.
- It dispenses no funds to the public.
- It is not a regulatory agency, but it does inform people of regulations and of their options in meeting them.
- Local programs are developed and carried out in full recognition of national problems and goals.
- The Extension staff educates people through personal contacts, meetings, demonstrations, and the mass media.
- Extension has the built-in flexibility to adjust its programs and subject matter to meet new needs. Activities shift from year to year as citizen groups and Extension workers close to the problems advise changes.

FAPC-209

Robert M. Kerr Food & Agricultural Products Center

FOOD TECHNOLOGY FACT SHEET

Adding Value to OKLAHOMA

405-744-6071 • www.fapc.biz • fapc@okstate.edu

December 2017

Big Data and Opportunities for Agriculture and Food Industry

Nurhan Dunford
FAPC Oil/Oilseed Chemist

The size of the digital data that can be stored, processed and easily accessed has exploded since the inception of personal computers in the1970s, then the World Wide Web in the 1990s and later social media in the 2000s. Mobile phones, online shopping, social networks, electronic communication, Global Positioning System (GPS), and instruments and machinery equipped with data collection and storage capabilities all generate enormous amount of data during their everyday operations.

Due to its increasing use, the Oxford English Dictionary recently included the term "Big Data" in its content and described it as "extremely large data sets that may be analyzed computationally to reveal patterns, trends and associations." Still, what constitutes Big Data is not well established. Some consider a minimum of 1 terabyte (storage capacity equals to about 1,500 CDs or 220 DVDs or enough space to store about 16 million Facebook photographs) is the minimum size to meet the criteria for Big Data. Nonetheless, the minimum size qualifier for the Big Data is changing rapidly due to technological developments.

Adaption of any disruptive technology, such as Big Data, is daunting. Fortunately, new data mining techniques are being developed by computer scientists, statisticians and mathematicians. For example, a new approach, artificial intelligence (AI), refers to detection and exploitation of patterns in data. However, businesses still face challenges in exploiting Big Data and making investment decisions because of the demanding data processing speed, interpretation, quality, security and privacy, and shortage of qualified data scientists.

The worth of any data is defined by its completeness, accuracy, consistency, objectivity, articulation and truthfulness. The unreliability and uncertainty concealed in the source diminish the value of the available data. For example, data collected from social media and other Internet sources may be subjected to consumer sentiments and could be unreliable and uncertain due to subjectivity of human opinions. Fortunately, analytics and statistical tools and techniques have been developed to deal with uncertainty and unreliability in Big Data.

Big Data provides tremendous opportunities for creation of new businesses, development of new products and services, and improvements in existing manufacturing and business operations. Significant cost savings, better decision making, and higher product and service quality can be achieved using Big Data analytics. This fact sheet highlights a few examples of Big Data use in food and agricultural industries and emerging trends and concerns in the field.

According to the United Nations Food and Agriculture Organization, there are about 1 billion people who do not have enough food to eat. Global hunger is one of the fundamental and moral challenges facing humanity. Considering world population is expected to reach about 9 billion by year 2025, food security is the most important problem that needs to be managed. Food security is achieved "when all people at all times have access to sufficient, safe, nutritious food to maintain a healthy and active life." Big Data can be helpful in achieving this goal. Big Data already is being used in humanitarian food security initiatives since 2009.

The United Nations (UN) has been leading the "Global Pulse" initiative, which is an international collaboration allowing UN agencies to use Big Data to monitor global socioeconomic crises such as famine, droughts and conflicts, and respond in a timely manner and gain real-time feedback on how well policy responses are working.

One of the critical factors for attaining food security is ensuring food availability, which largely depends on production. Technological tools supporting sustainable food production at the farm level are vital to achieve high yields at affordable cost.

The agricultural industry is using Big Data to help farmers make decisions that will increase yields and deliver safe, nutritious food to communities around the world. Predictive models developed using Big Data identify best management practices for achieving the best crop and livestock performance under various environmental conditions. The models, based on the most advanced machine learning algorithms and rooted in comprehensive and reliable datasets, provide the most accurate predictions. Such datasets include numerous weather and soil measurements as well as corresponding plant or animal performance assessments under various management regimes over several years.

During the recent decades, farmers have been introduced to precision agriculture, which is a farming management approach measuring and responding to field variability often using GPS tracking systems. Precision agriculture allows farmers to measure, hence, know more about their fields, crops and operations, and directly translate that knowledge into improved decision making and potentially better performance and higher yield. About four years ago, a survey of soybean farmers indicated that implementation of this technology resulted in a rapid payback, 5 percent savings on seed, fertilizer and chemicals expenses, about 16 percent crop yield increase and 50 percent reduction in water use. Tools used in precision agriculture relies on data collection and Big Data analytics. Agricultural production practices (inputs, i.e. type and amount of seed, fertilizer and pesticides applied, yield etc.), weather patterns and soil chemistry are monitored at the farm level and collected data are stored digitally. Advanced Big Data analytics facilitate precise management of agricultural operations and making better predictions and smarter decisions based on data. Farmers can target more-effective interventions using this technology.

Today, farm machinery equipped with digital systems that collects, retrieves and analyzes data and provide feedback on agronomic practices and yield estimates are readily available and broadly used on farms. Wi-Fi enabled barns allow farmers to manage their operations using software developed based on Big Data platforms. This technology can increase production yields while minimizing the adverse effects of agriculture on environment. The experts believe new technologies based on Big Data have the potential to double the agricultural production in the near future.

Ag-Analytics (https://ag-analytics.org/) is an example of an open-source, open-data platform for agricultural and environmental finance, insurance and risk management information. A number of data visualization web tools have been developed to help the users of this source. A few examples of these tools are follows:

- Crop Yield/Weather/Climate Tool: Users can quickly and easily view historical average yields for a variety of crops grown in the USA and climate and weather data. An interactive map allows users to select and evaluate desired regions.
- Dairy Margin Protection: This is a decision tool that can be used to calculate insurance premiums and forecasted Dairy Margin Protection Program payments and milk price based on the state, the annual milk production and amount of base production being insured. The tool updates daily based on current prices.
- Crop Insurance Premium Calculator: This is the only publicly available premium calculator for estimating Federal Crop Insurance Rates. It updates continuously, is simple to use and has a fast interface for generating quotes.
- Commodity Futures: Displays the futures of certain commodities either for continuous or selected contracts.
- Spot Price Interpolation: Provides information on commodity spot and future prices.

Sustainable farming has become a new carrier path for some young entrepreneurs to capitalize on the consumer demand for locally grown foods. These are usually small farmers with limited resources including limited capital and management, organization and record keeping expertise. Most small farmers sow, weed, water, harvest and hope to make some money at the end of the season, but do not know if their operations are really sustainable.

AgSquared (http://www.agsquared.com/) is a web-based interface that is designed to help farmers make decisions about their operations. A farmer can create a plan, calculate how many seeds and how much space the farm needs, when the crops need to be harvested and even keeps up with organic certification if needed. For example, if a farmer enters the planting date for a crop, the software creates a task reminding him/her when to weed the field and keeps a searchable detailed record of which problems arose, what was done and how much time it took to grow the plant and solve the problems. Detailed records for organic certification that trace crops from seed to harvest also can be created using the software. Over time, the data collected and stored in the program give small farmers an overview of whether their operations are efficient, profitable and sustainable in the long term. The technology is flexible enough to be used by backyard gardeners, small farmers growing tens of different crops in a season and commodity growers who manage hundreds of acres of two or three crops.

Local Orbit (https://localorbit.com/) is another program that can help large-scale buyers and small- and mid-sized suppliers to streamline the work of sourcing, selling, and delivering local and sustainable food, and connecting regional food systems. Consolidation of orders and payments, tracking and analysis of data, and supporting the communications and logistics needs of food supply chains are some of the features of this tool.

Traditional food safety data such as national monitoring data are well-structured but relatively limited and not harmonized between regions. The World Health Organization (WHO) has recently implemented the Big Data approach to improve decision making in food safety. The WHO uses the following Big Data definition: "The emerging use of rapidly collected, complex data in such unprecedented quantities that terabytes (1012 bytes), petabytes (1015 bytes) or even zettabytes (1021bytes) of storage may be required." WHO helped to develop "FOSCOLLAB" (http://www.who.int/foodsafety/foscollab/en/), which is a food safety platform integrating structured and non-structured data from multiple sectors such as animal, agriculture, food, public health and economic indicators.

Food industry generates large amounts and many different types of data every day such as records of processing and distribution conditions, microbial and chemical test results generated by safety and quality assurance laboratories, research findings and records kept by government agencies about foodborne illness outbreaks and nutrient content information. Some of the chain grocery stores and restaurants have been collecting large volumes of data on transportation temperature, shelf life, and food consumption and distribution, and analyzing this information by IBM Big Data Analytics. For example, in a traditional food management system, internal cooking temperatures of rotisserie chickens are measured about 10 times by health officers and 100 times by private investigators in one month, while a new Sustainable Paperless Auditing and Record Keeping (SPARK) system can handle about 1.4 million temperature measurements during the same period (http://corporate.walmart.com/_blog_/sustainability/20150112/how-walmarts-spark-keeps-your-food-fresh). Adaptation of such tracking systems allow companies to quickly recall the affected food from the distribution chain when something goes wrong. This data is available for exploring to gain further benefits for businesses and consumers and help to understand what triggers contamination and the spread of disease.

There are several other examples of Big Data applications in food safety. For example, the areas with an increased incidence of toxins and pathogens can be identified before entering the food chain by monitoring the conditions of crops in the field. Models that are developed using the data collected from farm fields in combination with environmental and meteorological data are already in use for predicting the contamination of the mycotoxin on wheat and the presence of *Listeria monocytogenes* on crops.

Although Big Data present remarkable opportunities in many areas including food security and agriculture, there are valid concerns this technology can challenge the accepted ethical and social norms. Capability of digital technologies to track behavior and capture data has rapidly increased. Nevertheless, understanding the ethical implications of Big Data is lagging behind its applications. Informed consent, ownership, privacy, objectivity and gaps created between those who can afford to implement the technology and those who lack the necessary resources are some of the major concerns.

Big companies involved in agriculture are investing heavily in technologies and tools for collecting farm-level data. These technologies potentially could advance inequities between farmers and large chemical, seed and machinery manufacturers and suppliers because of the lack of legal and regulatory framework to ensure farmers' access to the information collected from their fields and the technologies developed using the data. Technologies supporting particular agricultural systems of production could favor some farmers' operations at the expense of others and promote one brand versus other similar products. If the data collected from a farmer's field fall into the wrong hands, it can potentially be used against the farmer.

Development and maintenance of the agricultural dataset standards ensuring fairness, accessibility, interoperability and reusability are critical for successful use of Big Data. While promising on many fronts, Big Data and the tools and findings originating from it are provoking a host of ethical concerns. There is no question more work needs to be completed to establish a legal and regulatory frame around ownership, privacy and shared benefits stemming from Big Data analytics.