Big Data: Fueling the Next Evolution of Agricultural Innovation

Steven T. Sonka

Emeritus Professor, Agriculture and Consumer Economics, University of Illinois at Urbana-Champaign, Illinois, USA

`ssonka@illinois.edu`

Abstract. Application of Big Data in agriculture could both contribute to economic gain and to reduction of environmental impact. Especially at the farming level, the high cost of measuring actual operations as they occurred historically restrained decision making in the sector. Currently several sensing technologies associated with Big Data are being evaluated and adopted within the sector. Their adoption offers the opportunity to alter that historic benefit/cost relationship. Combined with advanced analytics, measurement and analysis of diverse sources of data promise to create value for sector decision makers and society. While consumers likely will continue to be the ultimate beneficiaries of such advances, the pattern by which value is captured by entities in the sector remains uncertain. Factors such as organizational collaboration and the application of rules associated with intellectual property will have significant impact upon the evolution of Big Data's implementation within agriculture.

Keywords. Farming Systems, Information Technology and Data Processing, Knowledge Economy, Agriculture-Industry Relationships.

1   Introduction

Agriculture[1]   is a vitally important sector affecting the global economy, societal well-being, and the vitality of natural ecosystems. Access to safe, nutritious, and affordable food is a goal for the citizens of all nations. In many developing countries, agricultural production employs the majority of the labor force. In more developed nations, an effective food and agricultural sector typically is a key component of the economy. Since man first tilled the soil to raise crops, agriculture has affected its supporting natural resource systems. Producing food (and other products) for current and expected future population levels are stressing those natural systems and developing means to reduce that stress is of global interest.

Innovation, especially in the last 150 years, has been an important means by which food and agricultural systems have increased productivity and fed an ever increasing global population (Borlaug, 2000; Chakraborty and Newton, 2011; Reid, 2015). Mechanization of tillage practices fueled expansion of land available for production while reducing human drudgery and labor needs. Biology focused on crop breeding

---

[1] The term agriculture often is viewed as synonymous with the farming activity. However, in this paper agriculture is viewed more broadly to encompass the entire food and agricultural system from genetics to retail. The terms, production agriculture or farming, will be used when referring to that specific subsector in the system.

increased the amount of production available from a given amount of inputs. Science applied to mitigation of the pests that affect crops and livestock and to more effective preservation of agricultural produce after harvest further ensure that food availability could expand for much of the world's population. More recently, genetic advances through application of biotechnology have been successfully employed (albeit not without controversy) and offer future potential to further contribute to human wellbeing. To be effective, however, each of these innovations had to be understood, adopted, and adapted by farmers and other managers.

In just the last few years, another source of innovation, Big Data, has captured the attention of citizens and decision makers in both the public and private sectors. While some would assert that Big Data currently is riding the crest of its "hype cycle" (Zwilling, 2014), application of Big Data has been effectively applied in numerous diverse settings. And Big Data is perceived to be as relevant for agriculture as it is for the rest of the economy, even by non-aggies. Padmasree Warrior, Chief Technology and Strategy Officer for Cisco Systems (Kirkland, 2013), believes:

> *In the next three to five years, as users we'll actually lean forward to use technology more versus what we had done in the past, where technology was coming to us. That will change everything, right? It will change health care; **it could even change farming**. There are new companies thinking about how you can farm differently using technology; sensors connected that use water more efficiently, use light, sunlight, more efficiently.*

While such potentials are exciting, it is important to remember that Big Data won't have much impact unless it too is understood, adopted, and adapted by farmers and other managers.

The purpose of this article is to explore the potential implications for Big Data and its adoption in agriculture. Because of the article's perspective on the future, its findings are necessarily speculative. The article is comprised of the following five sections:

- Key analytical concepts
- Precision agriculture; precursor to Big Data
- Likely sources of value creation
- Understanding the potential for value capture
- Summary and implications

This article's perspective is that the tools and techniques associated with Big Data offer the potential for agriculture to become significantly more effective in the pursuit of both economic and societal goals. Big Data's application can remove one of the fundamental constraints limiting agricultural managers – farmers, private sector managers, and public sector decisions makers. The constraint that the cost of measurement of actual operations typically has been significantly higher than the resulting benefits. Therefore, decisions tended to be driven by general conditions and often had a heavy bias to repeating traditional methods. Learning from actual operations was limited. The capabilities associated with Big Data offer the potential to fundamentally fuel management innovation. As will be further detailed in the paper, fully exploiting Big Data capabilities likely will require development of novel relations between firms and sectors within agriculture. This evolution could contribute to fundamental strategic

change in the sector.

2   Key analytical concepts

Big Data is a term that has received extensive exposure. However, as illustrated in Figure 1, that exposure is a relatively recent phenomenon. Prior to 2011, the Big Data term was barely of note. However, the term's usage literally exploded in 2012 and 2013. Therefore, while it is both appropriate and important to attempt to anticipate the potential impact of Big Data, that anticipation can't be based upon historic experience in the overall economy or in the agricultural sector itself. Instead, this analysis must intentionally be speculative in nature. Of course, as physicist Niels Bohr has said, "Prediction is extremely difficult. Especially about the future" (Ellis, 1970, p. 431).
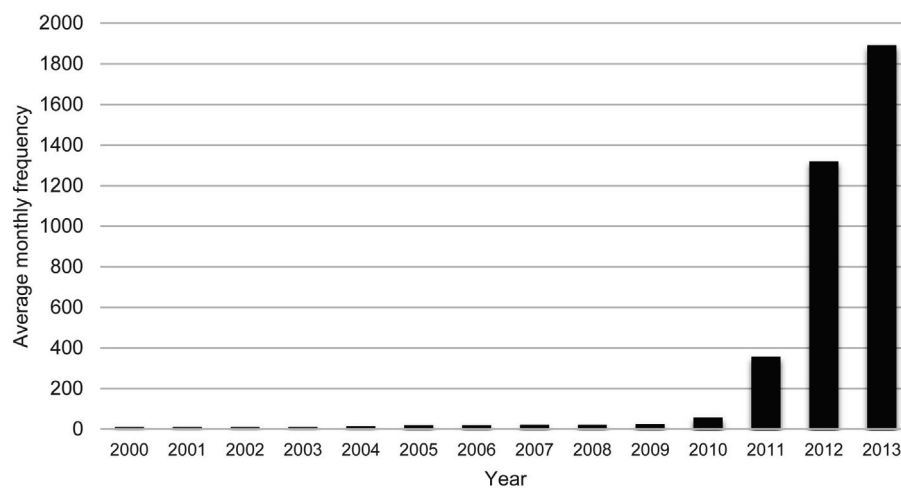

Fig. 1. Frequency distribution of documents containing the term "big data" in ProQuest Research Library (Gandomi and Haider, 2015).

The analysis presented here will employ three strategic concepts as a fundamental framework:

- The role of business models
- Value creation/value capture
- The resource-based theory of the firm.

Each concept will be described briefly in this section as they form the basis for the analysis presented later in the report.

2.1   The Role of Business Models

The term business model achieved extensive notoriety in the late 1990s as an outgrowth of the sudden surge of interest in e-commerce and the Internet as a business tool (Zott et al., 2011). While much of the media use of the term is not well structured, the term has important use as a means to understand the business and technological logic by which firms compete in their marketplace. As will be detailed later, the use of Big Data

tools and approaches in agriculture likely will affect the nature of competition and of inter-firm relationships. Business models that have long existed in the sector therefore will be under pressure for change.

Although media use of the business model term tends to be unstructured, recent work in the academic literature does provide useful definitions:

- A firm's business model is "a system of interdependent activities that transcends the focal firm and spans its boundaries" (Zott and Amit, 2010, p.216).
- The business model is "the heuristic logic that connects technical potential with the realization of economic value" (Cheesebrough and Rosenbloom, 2002, p.529).
- Business models consist of four interconnected elements – customer value proposition, profit formula, key resources, and key processes (Johnson et al., 2008).

The nature of business models for firms in production agriculture (farms) and those firms which support the farm sector have to a large extent been dictated by the costs of capturing and communicating data (Sonka et al., 2000). Historically the costs of data management were high relative to the direct benefits of doing so. Therefore, transaction-based interactions (employing only price and quantity information) dominate the business models both at the farm and the agribusiness level. As will be detailed in later sections of the paper, that historic cost/benefit relationship will be fundamentally altered by the application of the technologies and methods associated with Big Data. This has the potential to reshape the dominant business model employed in the sector as well.

2.2 Value creation/value capture

To be successful, innovations need to provide value to users and to do that in a way that provides incentives and compensation to the inventors (as well as returns to the business entities employing the innovation to provide goods and/or services). The processes of value creation and value capture, therefore, are key to understanding adoption of innovations. Those processes, however, have differing dynamics that should be carefully understood.

From an economic perspective, innovations are judged based upon the value that their use can provide. That use actually can be further divided into two components; use value and exchange value (Bowman and Ambrosini, 2000). Exchange value is more easily measured as it is documented as the price users pay for the goods and/or services associated with use of the innovation. Profits are the difference between the exchange value and the cost of providing those goods and/or services.

Use value, however, is the perceived benefit received by the user. For business uses, use value often can be measured. For consumer innovations, the benefits exist but tend to be subjective in nature.

Value capture is the process by which the profits earned from use of innovations accrue to the various entities involved. Customers compare benefits from use of the innovation with existing and emerging alternatives which can address the same purpose. Value capture, the realization of the exchange value, is driven by the bargaining power of

buyers and sellers (Bowman and Ambrosini, 2000).

In agriculture, the eventual beneficiaries of technological progress historically have been consumers. While innovations from Big Data may not change that outcome, the pattern by which actors in the sector are benefitted from their use is a dynamic and uncertain process.

2.3   Resource-based theory of the firm

A strategic concept, the resource-based theory of the firm, has proven useful in understanding and anticipating the dynamics of value capture in numerous settings (Barney, 1991). Relative to technology innovation, this approach focuses on the resource portfolios of effected firms.

Here the firm[2]   is considered as a bundle of resources. Some of those resources can be complements essential for successful implementation of the innovation in question. Other resources can be competitive substitutes, which serve as forces to constrain innovation or which may be rendered obsolete by innovation. In the competitive marketplace, firms which excel are those who can integrate innovative technologies with existing resources in a manner which fosters sustainable competitive advantage. Such resources are identified as:

- Valuable,
- Rare,
- Hard to imitate, and
- Have weak substitutes.

The resource-based approach is particularly intriguing relative to Big Data applications in agriculture because of the likely need for complementary resources to fully exploit the benefits of Big Data innovations. These resources reside in firms and organizations at differing levels within the sector.

3   Precision agriculture; Precursor to Big Data

This section of the paper will provide a brief overview of the precision agriculture experience. It is not intended as comprehensive assessment. It is intended to provide a sense of the evolution of precision agriculture, identify the more popular technologies employed and discuss the admittedly scanty evidence as to the economic gains from use of these innovations.

It is important to note that precision agriculture and Big Data are not synonymous. As we'll see, the current tools and techniques of precision agriculture have existed largely without Big Data concepts. However, it is hard to foresee that Big Data approaches could have significant impact without employing precision agriculture technologies. Further, some of attributes of Big Data adoption likely are foretold by the precision agriculture experience.

---

[2] For simplicity, the term, firm, is used in this discussion, even though it might be more accurate to refer to economic actors. Such economic actors could include NGO, universities, and government research entities who are and have the potential playing key roles in the evolution of Big Data in agriculture. This might particularly be the case for Big Data application in developing agricultural settings.

Precision agriculture has several dimensions; indeed the concept itself is not precisely defined. A 1997 report of the National Research Council refers to precision agriculture, "*... as a management strategy that uses information technologies to bring data from multiple sources to bear on decisions associated with crop production*". Key technologies and practices included within precision agriculture are:

- Georeferenced information;
- Global positioning systems;
- Geographic information systems and mapping software;
- Yield monitoring and mapping;
- Variable-rate input application technologies;
- Remote and ground-based sensors;
- Crop production modeling and decision support systems; and
- Electronic communications.

The term, precision agriculture, primarily has been linked to crop production. However, precision practices (and Big Data techniques for that matter) are equally applicable in animal agriculture, where georeferencing can refer to both sub areas of a field and individual animals. The tracking processes and required tools may differ but the managerial goal is still to separately manage increasingly smaller units of observation.

Farmers and agribusiness managers played a significant role in the development of precision agriculture. For example, in the mid-1990s, a group of agribusiness professionals in Champaign County, Illinois, came together to explore the opportunities associated with two emerging technologies — site-specific agriculture and that strange thing called the Internet. This group, called CCNetAg, was part of an initiative co-sponsored by the local Chamber of Commerce and the National Center for Supercomputing Applications at the University of Illinois. A voluntary enterprise, CCNetAg provided a vehicle for farmers, agribusiness managers, and university researchers to jointly explore adoption of these tools. Figure 2 depicts their expectations of a <u>then</u> future precision agriculture.

Although created some time ago, the graphic continues to depict key elements of precision farming:

- The role of georeferencing is indicated by satellites linking to the farm field.
- On the field itself, key farming operations are being directed by and are capturing digital information on:

    • Soil characteristics,
    • Nutrient application,
    • Planting,
    • Crop scouting, and
    • Harvesting.

- The layers that underlie the farm field represent the notion that visual mapping would allow the farmer, and the farmer's advisors, to see meaningful correlations to inform future decisions.
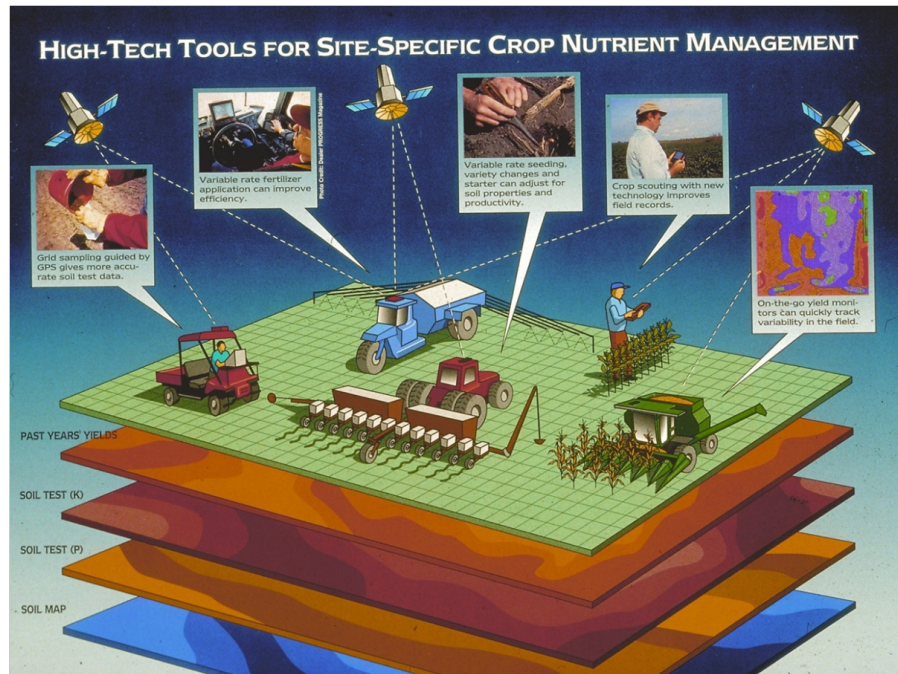
Fig. 2. A mid-1990s view of precision farming from the CCNetAg group (Sonka and Coaldrake, 1996).

Since 1997, technologies have advanced, although the general categories remain relevant. For example, auto-steer capabilities on farm machinery have become much more prevalent. And active, detailed measurement of the planting process (recording where "skips" occur) is now feasible. Further, the ability to monitor the status of farm machinery as it operates is now paired with electronic communications to signal when machine operations are out of acceptable bounds.

While there have been many publications describing precision agriculture, reports with independent evaluation of the economics of adoption are much less numerous. One means to assess whether there are net benefits of a technology is to monitor its marketplace adoption. For several years the Center for Food and Agricultural Business at Purdue University and CropLife magazine have surveyed agricultural input suppliers regarding the adoption of precision agriculture. Focused primarily on the Midwest and Southern regions, this work is a particularly useful assessment of the technology's application. From the 2015 report, Figures 3 and 4 provide evidence of adoption for key precision agriculture practices (Erickson and Widmar, 2015).

The crop input dealers who provided input for this study are uniquely well positioned to understand and report on adoption of these technologies. Their firms provide inputs (fertilizer, pesticides, and seeds) and services to producers evaluating and adopting precision agriculture.
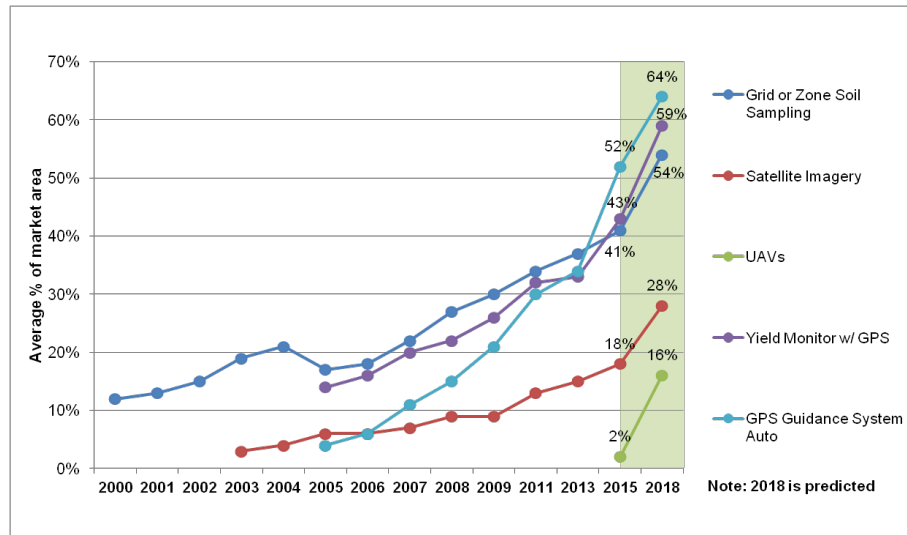
Fig. 3. Estimated market area using precision services over time (adapted from Erickson and Widmar, 2015).

Early interest in precision agriculture focused on site-specific application of inputs and on use of yield monitors. As shown in Figure 3, grid sampling, a practice associated with site-specific lime and fertilizer application, is currently employed on about 2 out of 5 crop acres. Increased coverage to a majority of acres is expected by 2018. Similar adoption rates (43% and 59%) are noted for GPS-assisted yield monitors. Over the last decade, use of GPS guidance systems has increased rapidly, to a current use estimated to exceed 50%. Continued strong growth to 2018 is expected. The use of satellite imagery and UAVs as tools to support crop production is more recent. Current use affects 18% and 2% of acreages, respectively. Interesting, acreage covered by UAVs is expected to increase eightfold, to 16%, in just three years.

Figure 4 describes a relatively consistent adoption pattern for VRT (variable rate technology) practices. In the early 2000s, adoption was at single digital levels. Since then, steady increases in the extent of acreage covered have occurred. However, the most utilized practice, application of lime, is only now achieving coverage on 41% of the total acreage. These patterns also are interesting because of the very different price regimes that existed for corn and soybeans over these 15 years. When output prices were low prior to 2008, the driver for adoption likely was cost reduction. Possibly, increasing yields was a more significant factor in recent years when prices were higher.

Media and marketing attention sometimes blur distinctions between precision agriculture and Big Data. Some communications seem to suggest that Big Data is just an updated buzzword for precision agriculture practices. That is not the case.
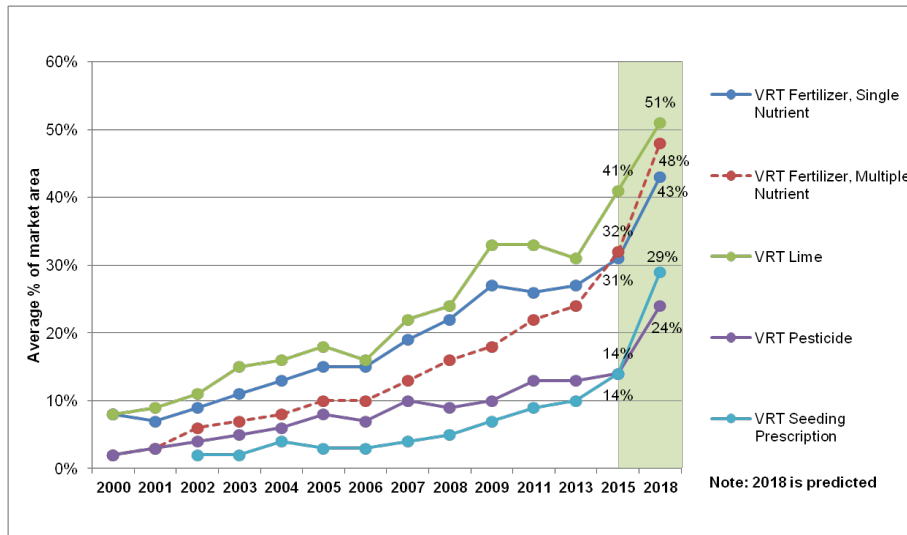
Fig. 4. Estimated market area using VRT technology over time (adapted from Erickson and Widmar, 2015).

Figure 2 above can be used to identify key differences:

- While a useful picture, that graphic does focus our attention on the individual field. The volume characteristic of Big Data requires observations from many, many farm fields to be effective. Discerning the interrelated effects of soil type, several nutrients, and seed variety requires data dispersed over time and space.

- While the farmer has several types of precision data from each field, additional sources of data naturally reside and originate beyond the fencerow. Achieving the Big Data's variety characteristic requires access to that broader set of information.

- Precision agriculture employs comparisons across field map layers as its dominant method of analysis. The effect of a single factor, such as a blocked tile line or a buried fencerow, often is observable from a map. However, identifying complex interactions across several production factors and multiple years requires much more sophisticated tools. Analytics is a major differentiating feature of Big Data.

- As noted previously, precision agriculture has had 20+ years of experience. Aggregating all the digital information collected from yield monitors and site-specific input operations would result in an extremely large set of data. However, that data currently is located on innumerable thumb drives, disk drives, and desktop computers. Effective analysis won't be possible unless/until that data can be accessed and aggregated. The associated organizational issues of doing that will be discussed in a later section of this article.

Both precision agriculture and Big Data arise from the advent and application of information and communication technologies. As noted previously, they are not

synonymous. That said, it is hard to foresee that Big Data approaches will have significant impact without employing the data generated by precision agriculture practices.

4   Likely sources of value creation

Big Data generally is referred to as a singular entity. It is not! In reality, Big Data is much more a capability than it is a thing. It is the capability to extract information and insights where previously it was economically, if not technically, not possible to do so. Advances across several technologies are fueling the growing Big Data capability. These include, but are not limited to computation, data storage, communications, and sensing. The growing ability of analysts and managers to exploit the information provided by the Big Data capability is equally important.

Although of relatively recent origin, numerous attempts have been made to define Big Data. For example:

- The phrase "big data" refers to large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future (The National Science Foundation, 2012).
- Big Data shall mean the datasets that could not be perceived, acquired, managed and processed by traditional IT and software/hardware tools within a tolerable time (Chen et al., 2014)
- Big Data is where the data volume, acquisition velocity, or data representation [variety] limits the ability to perform effective analysis using traditional relational approaches or requires the use of significant horizontal scaling for efficient processing (Cooper and Mell, 2012).
- Big Data is high-volume, -velocity, and -variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Gartner IT Glossary, 2012).

The purpose of this section of the paper is to move beyond those definitions to explore how application of Big Data could foster the creation of value in agriculture. Three pathways to value creation are identified. Application of tools to measure and monitor agricultural activities – at extremely low cost – is the first. Data analytics which can integrate data from diverse sources to generate novel insights is the second. The third factor focuses on external pressures to better monitor agricultural activities which, in so doing, create sources of data that potentially can lead to strategic change.

4.1   Dimensions of Big Data

Three dimensions (Figure 5) often are employed to describe the Big Data phenomenon: Volume, Velocity, and Variety (Manyika et al., 2011). Each dimension presents both challenges for data management and opportunities to advance business decision-making. These three dimensions focus on the nature of data. However, just having data isn't sufficient. Analytics is the hidden, "secret sauce" of Big Data. Analytics, discussed later, refers to the increasingly sophisticated means by which useful insights can be

fashioned from available.

"*90% of the data in the world today has been created in the last two years alone*" (IBM, 2012). In recent years, statements similar to IBM's observation and its emphasis on volume of data have become increasingly more common.
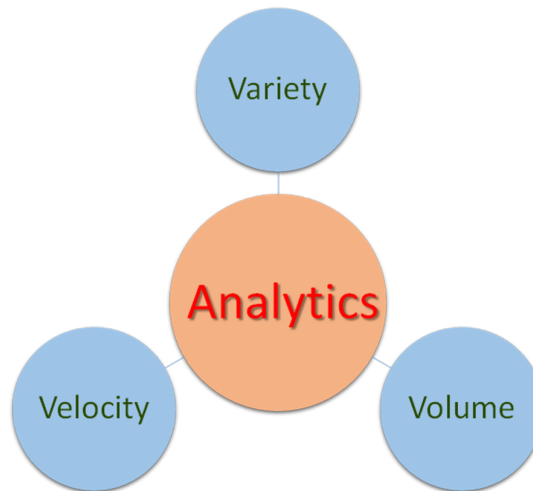

Fig. 5. Dimensions of Big Data.

The Volume dimension of Big Data is not defined in specific quantitative terms. Rather, Big Data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective; with no single standard of how big a dataset needs to be to be considered big. And that standard can vary between industries and applications.

An example of one firm's use of Big Data is provided by GE — which now collects 50 million pieces of data from 10 million sensors everyday (Hardy, 2014). GE installs sensors on turbines to collect information on the "health" of the blades. Typically, one gas turbine can generate 500 gigabytes of data daily. If use of that data can improve energy efficiency by 1%, GE can help customers save a total of $300 billion (Marr, 2014)!

The Velocity dimension refers to the capability to acquire, understand, and respond to events *as they occur*. Sometimes it's not enough just to know what's happened; rather we want to know what's happening. We've all become familiar with real-time traffic information available at our fingertips. Google Map provides live traffic information by analyzing the speed of phones using the Google Map app on the road (Barth, 2009). Based on the changing traffic status and extensive analysis of factors that affect congestion, Google Map can suggest alternative routes in real-time to ensure a faster and smoother drive.

For analysts interested in retailing, anticipating the level of sales is important. Brynjolfsson and McAfee (2012) report on an effort to monitor mobile phone traffic to infer how many people were in the parking lots of a key retailer on Black Friday — the start of the holiday shopping season in the United States — as a means to estimate retail sales.

Variety, as a dimension of Big Data, may be the most novel and intriguing of these three characteristics. For many of us, data referred to numbers meaningfully arranged in rows and columns. For Big Data, the reality of "what is data" is wildly expanded. The following are just some of the types of data available to be converted into information:

- Financial transactions
- The movement of your eyes as you read this text
- "Turns of a screw" in a manufacturing process
- Tracking of web pages examined by a customer
- Photos of plants
- GPS locations
- Text
- Conversations on cell phones
- Fan speed, temperature, and humidity in a factory producing motorcycles
- Images of plant growth taken from drones or from satellites
- Questions

4.2   Data variety requires low cost measurement

"You Can't Manage What You Don't Measure!" is a phrase attributed to both Peter Drucker and W. Edwards Deming. This phrase is as applicable to farmers as it is to managers at Toyota or Amazon (Brynjolfsson and McAfee, 2012). The relationship between measurement and the ability to make improved decisions is critically important in understanding the potential for Big Data to affect agricultural management.

The author of this paper grew up on a small farm in the Midwest region of the United States and, throughout his career, has learned extensively from farmers in the United States and globally. With apologies for a small digression, let me use personal experience to focus on the linkage between measurement and management. Growing up on a farm, the linkage between what could be measured and our ability to improve performance was straightforward. In those days, we had to carry the, hopefully, full milking machine from the cow to the milk tank and it was fairly easy to know which cows were producing more. And because there were less than 20 cows in the herd, it also was possible to remember those higher producing cows and give them an extra portion of grain.

On this same farm, about 120 egg producing chickens were housed in a building, with ample room to roam outdoors as well. Eggs were collected twice a day. Performance of the entire of group was observable. Knowledge that could lead to improved performance of individual birds, however, was not observable. Technically, it might have been possible to establish a production system where measurement of individual bird performance would have been available. However, the economics of egg production at that time didn't justify the costs of such a system.

The important point to stress here is that the desire to link measurement of outcomes and management actions in farming is not new. However, the economics of measurement (the cost of measurement versus the benefits of doing so), given the available technology, inhibited my father and other farmers from capturing and

exploiting more data.

4.3   Variety as a key

Suddenly (at least in agricultural measurement terms), the "what is data" question – the variety dimension of Big Data – has new answers. Figure 6 provides a visual illustration of the change. In its upper left hand corner, we see data as we are used to it – rows and columns of nicely organized numbers. The picture in the upper right hand corner is of a pasture in New Zealand. Pasture is the primary source of nutrition for dairy cows in that country and supplemental fertilization is a necessary economic practice. The uneven pattern of the forage in that field is measured by a sensor on the fertilizer spreader to regulate how much fertilizer is applied – as the spreader goes across the field. In this situation, uneven forage growth is now data.

The lower left hand corner of Figure 6 shows the most versatile sensor in the world – individuals using their cell phone. Particularly for agriculture in developing nations, the cell phone is a phenomenal source of potential change – because of both information sent to those individuals and information they now can provide. And as illustrated in the lower right hand quadrant of Figure 6, satellite imagery can measure temporal changes in reflectivity of plants to provide estimates of growth (RIICE, 2013). The picture is focused on rice production in Asia.
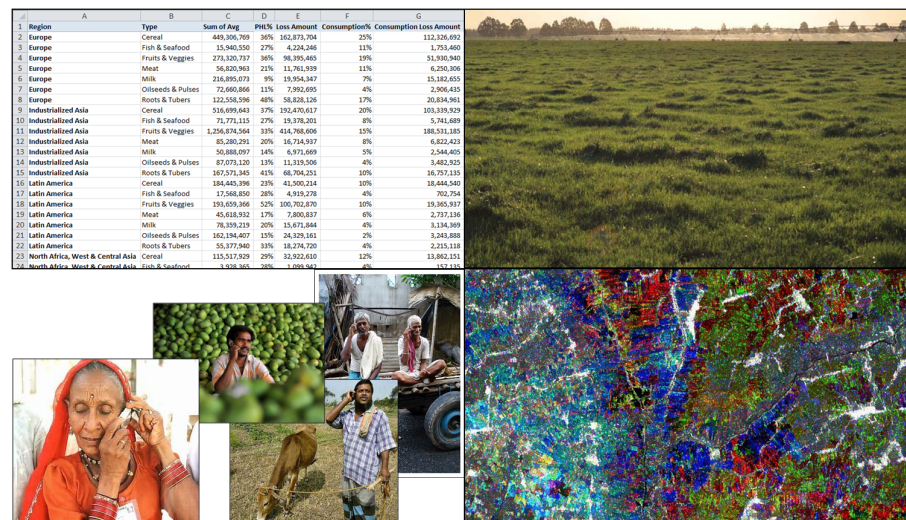

Fig. 6. A few sources of data3.

While satellite imagery is one source of remotely sensed data, recent years have seen a pronounced increase in the capabilities and interest in Unmanned Aerial Systems (UAS) as a source of data for agriculture. There are numerous ongoing efforts to transform UAS technology originally focused on military purposes to applications supporting production agriculture. "Universities already are working with agricultural

---

3 Graphics courtesy of: agrioptics.co.nz; T. Abdelzaher, Champaign, IL.; Mock, Morrow & Papendieck; International Rice Research Institute.

groups to experiment with different types of unmanned aircraft outfitted with sensors and other technologies to measure and protect crop health" (King, 2013). Example applications include:

- Monitoring of potato production (Oregon State University)
- Targeting pesticide spraying on hillside vineyards (University of California, Davis)
- Mapping areas of nitrogen deficiency (Kansas State University)
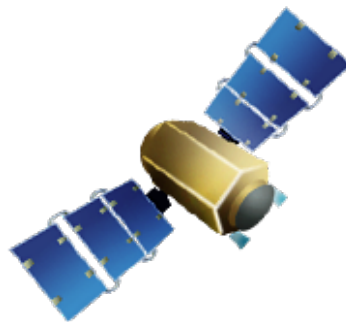- Detecting airborne microbes (Virginia Polytechnic Institute and State University)

Fig. 7. Unmanned Aerial Systems offer low cost data acquisition[4].

Those specific examples are only a few of the numerous experiments and demonstrations being conducted to identify cost effective means to employ UAS technology (Figure 7). UAS capabilities offer flexibility and potentially lower cost relative to the use of even small manned aircraft. Development efforts are being conducted globally; however, it is likely that initial commercial application will occur where higher value crops dominate.

4.4   Analytics

Access to lots of data, generated from diverse sources with minimal lag times, sounds attractive. Managers, however, quickly will ask, what do I do with all this stuff? Without similar advances in analytic capabilities, just acquiring more data is unlikely to have significant impact within agriculture.

Analytics and its related, more recent term, data science, are key factors by which Big Data capabilities can actually contribute to improved performance in the agricultural sector. Data science refers to the study of the generalizable extraction of knowledge from data (Dhar, 2013). Tools based upon data science are being developed for implementation in the sector, although these efforts are at their very early stages.

The associated concept of analytics similarly is maturing and its use refined (Davenport, 2013; Watson, 2013). Analytic efforts can be categorized as being of one of three types:

- Descriptive efforts focus on documenting what has occurred,

---

[4] Graphic courtesy of: Microsoft Corporation

- Predictive efforts explore what will occur, and
- Prescriptive efforts identify what should occur (given the optimization algorithms employed).

One tool providing predictive capabilities was recently unveiled by the giant retailer, Amazon (Bensinger, 2014). This patented tool will enable Amazon managers to undertake what it calls "anticipatory shipping", a method to start delivering packages even before customers click "buy". Amazon intends to box and ship products it expects customers in a specific area will want but haven't yet ordered. In deciding what to ship, Amazon's analytical process considers previous orders, product searches, wish lists, shopping-cart contents, returns, and even how long an Internet user's cursor hovers over an item. Analytics and its related, more recent term, data science, are key factors by which Big Data capabilities actually can contribute to improved performance, not just in retailing, but also in agriculture.

In agriculture, as in most fields, descriptive efforts have been most common and even those are relatively infrequent. Within production agriculture, knowing what has occurred – even if very accurately and precisely – does not necessarily provide useful insights as to what should be done in the future.

Production agriculture is complex, where biology, weather, and human actions interact. Science-based methods have been employed to discern why crop and livestock production occurs in the manner in which they do. Indeed, relative to the Big Data topic, it might be useful to consider this as the "small data" process.

The process starts with lab research employing the scientific method as a systematic process to gain knowledge through experimentation. Indeed the scientific method is designed to ensure that the results of an experimental study did not occur just by chance (Herren, 2014). However, results left in the lab don't lead to innovation and progress in the farm field. In the United States, the USDA, Land Grant universities, and the private sector have collaborated to exploit scientific advances. A highly effective, but distributed, system emerged where knowledge gained in the laboratory was tested and refined on experimental plots and then extended to agricultural producers.

In agriculture, therefore, knowledge from science will need to be effectively integrated within efforts to accomplish the goals of predictive and prescriptive analytics. Even with this additional complication, the potential of tools based upon emerging data science capabilities offers significant promise to more effectively optimize operations and create value within the agricultural sector.

4.5   Public pressures to better monitor agriculture

Beyond its direct economic impact, society has intense interest in the social and environmental effects of the agricultural sector. Food safety and security are of public interest in every society. Interest in mitigating negative environmental impacts of agricultural operations is increasingly of interest and that interest is not constrained to just citizens in developed nations. In addition to public sector interest, some consumer segments express interest and concern regarding the practices and methods employed to produce food. Therefore, in addition to public sector-based regulation, documentation as to practices employed is increasingly being required by the private sector by food manufacturers and retailers.

Interestingly, technological innovations, such as those noted previously, have potential to provide much better evidence as to these societal and environmental effects. These include both tools to more precisely measure and monitor as well as analytical methods to better understand and predict effects.

At first blush, managers tend not to welcome additional constraints, whether from public or private sources. However, there can be an interesting "unintended consequence" effect when information is captured digitally. That digital information, which might not have been captured otherwise, now becomes available for analysis. As we saw in the early days of the 1990s knowledge economy, unintended insights can be developed from digital data captured for other purposes (Sampler, 1997; Shapiro and Varian, 1999). Application of those insights can drive strategic change in effected industries.

5   Understanding the potential for value capture

To be attractive, prospective innovations have to display the potential to create value in the marketplace. The longer run economic effect of adopting innovations, however, is determined by value capture, the distribution of resulting benefits to consumers and among the firms within the value chain affected by the innovation. The prior section identified three interrelated pathways by which the technologies and application of Big Data can potentially create value for consumers, society and to the sector's economic entities. This section will explore the concept of value capture relative to the adoption of Big Data within agriculture.

Identification of potential value creation typically is more straightforward than is predicting the pattern and extent of value capture. Historically, food consumers have been the eventual beneficiaries of technology adoption in production agriculture. Even if that remains the likely long-run outcome, the allocation of net benefits among the sector's economic actors is of key interest. As noted in an earlier section of the paper, a strategic concept called resource-based theory of the firm has proven useful in understanding and anticipating the dynamics of value capture (Bowman and Ambrosini, 2000). Particularly in the context of Big Data in agriculture, the resources needed to create and capture value often will not reside within one firm. Therefore, new business models that enable collaboration across firm boundaries likely will be needed. Implementation of these business models could allow application of Big Data tools and techniques to be powerful and sustainable sources of competitive advantage.

From a manager's strategic perspective, therefore, implementation of effective Big Data based innovations is attractive. Within agriculture some of the data comprising these systems likely will come from external sources (for example, weather data, environmental regulatory filings, and futures market price movements). Other systems, however, will be based upon data generated from activities internal to the operations of firms in the food and agribusiness sector. Although that data often will be analyzed in combination with external data, firms will need access to internal data to effectively compete. Therefore, data access, based upon current operations, represents a resource of critical potential importance and is a starting point for this analysis.

Figure 8 provides a high level view of the key subsectors within agriculture that has

proved useful for consideration of future competitive dynamics relating to Big Data. The genetics subsector is separately identified here because of its linkages with Big Data. A number of firms in that category have capabilities to operate as input suppliers as well. The input supply category refers to providers of equipment, seed, fertilizer, and chemicals to farmers as well as providers of financial and managerial services. The production agriculture segment is comprised of farming firms, which can range from low-resourced, smallholders to family corporations to subsidiaries of major corporations. The 1st handler segment refers to firms which aggregate, transport and initially process agricultural produce but do not directly market to consumers. The final segment relates to food manufacturers and retailers. These types of activities are combined here because of their common interest in employing Big Data tools to better understand consumers.

From a strategic perspective, it is important to stress that Big Data tools already are extensively employed, particularly at both "ends" of the sector. Firms at the food manufacturing and the food retailing levels expend considerable resources to continually develop a better understanding of consumers. Insights gained through application of Big Data analytics can allow managers both to anticipate and respond to consumer concerns. Far upstream in the sector, bioinformatics and other Big Data tools are employed to accelerate research and development processes, advancing genomic capabilities of the sector. Figure 8 identifies, at a general level, key interests that "naturally" reside within each subsector and have the potential to be important within Big Data applications.

Agricultural operations occur across time and space. Therefore, the logistics of providing inputs, production, and aggregating output consume considerable resources. Advances in information and communication technology combined with Big Data analytics offer the potential to reduce the amount of resources needed. Deadweight loss is a term that describes system inefficiencies that can be reduced by enhanced coordination within and between firms. Even in advanced agricultural settings, reduction of deadweight loss is perceived to be an attractive potential use of Big Data innovations.

In this context, deadweight loss refers to the processes by which inputs and outputs are delivered (when and where). A more intriguing issue for many is whether application of Big Data can fundamentally alter decision making as to "what" should be done. Can we further optimize the biology of agricultural production, especially in the context of the larger food and agricultural system? Earlier it was noted that new sensing technologies offer the potential to monitor and document what actually occurs as agricultural production takes place. The resulting data potentially would be available at never before levels of detail, in terms of time and space, and at low-cost. Further, analytic capabilities could combine diverse sources of data to discern previously unknown patterns and provide insights not available previously.

A result of application of these innovations would be optimization of agricultural production systems, simultaneously reducing its environmental impact and improving profitability. There are two interrelated factors that need to be addressed in considering the possible evolution of this optimization:

- Production agriculture involves biologic processes subject to considerable

uncertainty. Therefore, even if one knows exactly <u>what</u> occurred in one production season and what actions would have optimized performance under those circumstances, that information may not be a good predictor of what actions should be undertaken in the next season. Agricultural science is devoted to discerning the why of agricultural production. That science will need to be integrated within Big Data techniques to truly optimize system performance.

- In most systems of agricultural production today, even the knowledge of what occurred doesn't necessarily reside within one organization. Further, as was noted for precision agriculture, individual entities at the production level typically don't have the scale to produce sufficient data nor to have the capabilities needed to analyze that data.

Because of these two factors, collaboration across organizational boundaries will be required to fully exploit the potential benefits of Big Data's application to agriculture. A host of factors, beyond technological effectiveness, will influence the speed and extent of this exploitation. These relate to intellectual property and competitive dynamics as well as the magnitude of economic benefits available. Such factors are not insurmountable and can be viewed as much as opportunities as they are impediments. How they are resolved, however, will have a major impact on Big Data's eventual contribution to performance within agriculture.
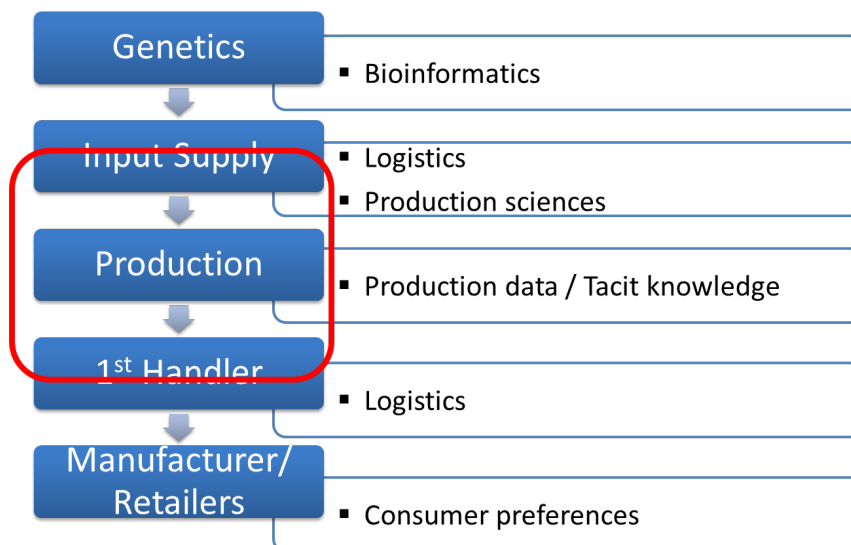

Fig. 8. Subsectors and their key strategic interests relating Big Data.

6   Summary and implications

Big Data capabilities have emerged in recent years as potential "game changers" that could affect economies and societies in profound, although somewhat uncertain, ways. Those potentials extend to economic, social, and environmental performance of food and agricultural systems as well. Although it is very early days in terms of Big Data

adoption and agriculture, expectations already have been altered and investment in research, development, and testing of associated technologies is occurring.

Although necessarily speculative, this article explores the potential impact of Big Data in the context of the agricultural sector. While noting some of the technologies associated with potential Big Data implementation, a decision-making lens is adopted as the primary conceptual tool for this exploration. The reason for doing this is the belief that use of Big Data capabilities will have primary impact by altering decision-making processes relating to:

- Adoption and implementation of new technologies,
- Management of on-going operations, and
- Execution of existing and new relationships:

  - Among competing and collaborating firms
  - Between suppliers and customers
  - With customer and non-customer stakeholders.

To be economically attractive, innovations have to display the potential to create value in the marketplace. The longer run economic effect of adopting innovations, however, is determined by value capture, the distribution of resulting benefits to consumers and within the value chain affected by the innovation. Value capture is heavily influenced by the resource portfolios of effected firms. These strategic concepts, the business model, the resource-based theory of the firm, value creation and value capture, are employed here to frame the exploration of Big Data's potential effects.

Two interrelated questions are addressed in the context of potential strategic change driven by Big Data innovations. Specifically, if such change does occur:

- What would be the likely source of change?
- Who (in the context of economic entities) would be the likely change agents?

Historically, the geographic, time, and economic dimensions of agriculture have constrained the decision making capabilities of sector managers. Although managers desired to be able to measure the impact of their decisions and actions, typically the cost of measurement exceeded the benefits of doing so. Innovations, many of which are integral within a broad perspective of Big Data, now offer the potential to fundamentally alter that benefit/cost dynamic and in so doing foster the potential for value creation in the sector.

Three interrelated forces are identified as likely change agents driving value creation as (if) Big Data capabilities are applied in agriculture:

- Extensive implementation of low-cost sensor capabilities will allow managers to measure actual operation of systems and more effectively respond both in "real-time" and in planning future operations.
- The application of advanced analytics will provide insights that support improved decision making.
- Societal and business motivations will increasingly require more extensive monitoring in response to requirements imposed by the public sector or by customers. Because the associated data will be digital, prior experience indicates that additional use of that data can drive strategic change extending beyond the original intent.

The forces just identified offer the potential for value creation which can provide benefits to consumers, society and to the sector's economic entities. As is typically the case, identification of the forces for potential value creation is more straightforward than is anticipating the pattern of future value capture. Historically, consumers have been the eventual beneficiaries of adoption of technology in production agriculture. Even if that remains the likely long-run outcome, the allocation of net benefits among the sector's economic actors is of key interest.

Without attempting to predict that allocation, a number of key factors of interest can be detailed. It is important to note that Big Data capabilities already are being employed within the food and agribusiness sector. Firms at the retail and manufacturing level are aggressively monitoring social media and other data sources to better understand and serve consumers. Bioinformatics has become an essential tool for firms providing genetic resources for crops and livestock. In addition to direct application of the resulting information, linkages with associated partners at other levels of the sector offer the potential for further economic and social gains.

Firms operating at the input supply, production agriculture, and first handler levels of the sector are beginning to explore Big Data application. Employing Big Data for management and logistics purposes has the potential to reduce costs and to improve economic performance.

Optimization of the biology of production agriculture is a beguiling potential with extensive potential benefits. A few "farming" organizations do have the scale of operation which could justify development and application of Big Data capabilities. More generally, the information resources needed to move towards optimization reside within multiple organizations. The most numerous of these are individual farming operations. Typically, however, some combination of firms at the input supply, service provision and output handling/processor level also will have key elements of the needed information resources. Future decisions to shape effective business models for firms operating in these domains will determine the ultimate value capture dimensions of Big Data's application in agriculture.

Acknowledgements

7   References

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*. 17: 99-120.

Barth, D. (2009). The bright side of sitting in traffic: Crowdsourcing road congestion data. Retrieved February, 2, 2016, from Google's Official Blog Web site: https://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html

Bensinger, G. (2014). Amazon wants to ship your package before you buy it. *The Wall Street Journal*. Retrieved February, 2, 2016, from the Wall Street Journal Web site: http://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it/

Borlaug, N. E. (2000). Ending world hunger: The promise of biotechnology and the threat of antiscience zealotry. *Plant Physiology, 124*(2), 487-490.

Bowman, C., & Ambrosini, V. (2000). Value creation versus value capture: Towards a coherent definition of value in strategy. *British Journal of Management, 11*, 1-15.

Brynjolfsson, E. & McAfee, A. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60-68.

Chakraborty, S., & Newton, A. C. (2011). Climate change, plant disease and food security: An overview. *Plant Pathology, 60*(1), 2-14.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications, 19*(2), 171-209.

Cheesebrough, H. W., & Rosenbloom, R. S. (2002). The Role of the Business Model in Capturing Value from Innovation. *Industrial and Corporate Change, 11*, 533-534.

Cooper, M., & Mell, P. (2012). Tackling Big Data [PDF document]. Retrieved February 2, 2016, from National Institute of Standards and Technology Web site: http://csrc.nist.gov/groups/SMA/forum/documents/june2012presentations/fcsm_june2012_cooper_mell.pdf

Davenport, T. H. (2013). Analytics 3.0. *Harvard Business Review, 91*(12), 64-72.

Dhar, V. (2013). Data Science and Prediction. *Communications of the ACM, 56*(12), 64-73.

Ellis, A. K. (1970). *Teaching and learning elementary social studies*. Boston: Allyn & Bacon.

Erickson, B. & Widmar, D.A. (2015). *2015 Precision Agricultural Services Dealership Survey Results*. West Lafayette, Indiana: Purdue University. Retrieved February 2, 2016, from Center for Food and Agribusiness, Purdue University Web site: http://agribusiness.purdue.edu/resources/2015-precision-dealership-survey-results

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*, 137-144.

Gartner IT Glossary. (2012). *Big Data*. Retrieved June, 5 2015, from Gartner Web site: http://www.gartner.com/it-glossary/big-data

Hardy, Q. (2014). G.E. opens its Big Data platform. Retrieved February 2, 2016, from The New York Times Web site: *http://bits.blogs.nytimes.com/2014/10/09/ge-opens-its-big-data-platform/?_r=2*

Herren, R. V. (2014). *The Science Of Agriculture A Biological Approach 4th Ed*. Independence: Cengage Learning.

IBM. (2012). *What is big data?* Retrieved November 22, 2015, from IBM Web site: https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html

Johnson, M. W., Christenson, C. C., & Kagermann, H. (2008). Reinventing Your Business Model. *Harvard Business Review*, *86*(12), 50-59.

King, R. (2013). Farmers Experiment with Drones. Retrieved February 2, 2016, from *The Wall Street Journal* Web site: http://blogs.wsj.com/cio/2013/04/18/farmers-experiment-with-drones/

Kirkland, R. (2013). Connecting everything: A conversation with Cisco's Padmasree Warrior. Retrieved February 2, 2016 from the McKinsey & Company Web site: http://www.mckinsey.com/Insights/High_Tech_Telecoms_Internet/Connecting_everything_A_conversation_with_Ciscos_Padmasree_Warrior?cid=disruptive_tech-eml-alt-mip-mck-oth-1305

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. Retrieved February 2, 2016 from McKinsey & Company Web site: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

Marr, B. (2014). How GE is using Big Data to drive business performance. Retrieved February 2, 2016, from SmartDataCollective Web site: http://www.smartdatacollective.com/bernardmarr/229151/how-ge-using-big-data-drive-business-performance

National Research Council. (1997). *Precision Agriculture in the 21st Century: Geospatial and Information Technologies in Crop Management*. Washington, D.C.: National Academies Press. Accessed October 30, 2015. http://www.nap.edu/catalog/5491/precision-agriculture-in-the-21st-century-geospatial-and-information-technologies

National Science Foundation. (2012). Core techniques and technologies for advancing big data science & engineering. Retrieved February 2, 2016, from the National Academies Press Web site: http://www.nsf.gov/pubs/2012/nsf12499/nsf12499.htm

Reid, J. F. (2015). The impact of mechanization on agriculture. *The Bridge, 41*(3), 22-29. Retrieved from: https://www.nae.edu/File.aspx?id=52553

Remote sensing-based information and insurance for crops in emerging economics [RIICE]. (2013). About RIICE. Retrieved February 2, 2016, from International Rice Research Institute Web site: http://www.riice.org/about-riice/

Sampler, J. L., (1997). Redefining Industry Structure for the Information Age. *Strategic Management Journal*. 19:343-356.

Shapiro, C. & H. R. Varian. (1999). *Information Rules*. Boston: Harvard Business School Press.

Sonka, S. T., & Coaldrake, K. F. (1996). Cyberfarm: What does it look like? What does it mean? *American Journal of Agricultural Economics*, *78*(5), 1263-1268.

Sonka, S. T., Lins, D. A., Schroeder, R. C., & Hofing, S. L. (2000). Production Agriculture as a Knowledge Creating System. *International Food and Agribusiness Management Review*, *2*(2), 165-178.

Watson, H. J. (2013). The Business Case for Analytics. *Bized, 12*(3), 49-54.

Zott, C. & Amit, R. (2010). Designing your Future Business Model. *Long Range Planning*, *43*,216-226.

Zott, C., Amit, R., & Massa, L. (2011). The Business Model: Recent Developments and Future Research. *Journal of Management*, *37*(4), 1019-1042.

Zwilling, M. (2014, July 7). Where is your technology in the Gartner Hype Cycle? Retrieved February 2, 2016, from Startup Professionals Musings Web site: http://blog.startupprofessionals.com/2014/07/where-is-your-technology-in-gartner.html