International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-10, Issue-2, Mar 2022
https://doi.org/10.55524/ijircst 2022.10.3.60
Article ID CST1031, Pages 294-299
www.ijircst.org

# A Review of Big Data in Smart Farming

**Ranjeev Kumar Chopra**

Assistant Professor, Department of Computer Applications, RIMT University, Mandi Gobindgarh, Punjab, India

Email Correspondence should be addressed to Ranjeev Kumar Chopra; ranjeevk.chopra@rimt.ac.in



**ABSTRACT-** Intelligent agro is now a growing concept that accentuates employment of ICT inside cyber-physical farmland managerial chain. The said growth is projected to be accelerated by newer advancements such as the Internet of Thing (IoT) & Mobile computing, which will enable ranchers to deploy increased machines with machine intelligence. All of that is represented by the Data Analysis phenomenon that alludes to massive volumes of data of different forms that may be gathered, processed, and used to make decisions. One purpose of current research is to get a better knowledge of the present state of Big Data processing in Precision Agriculture and even to uncover underlying accompanying psychosocial challenges that need to be addressed. A systematic technique has been used to construct a philosophical paradigm for assessment, which is often used in subsequent study on the issue. With regards to basic productivity, our research found that Data Analysis implementations in Agriculture 5.0 have an influence on this kind of agriculture distribution network. In farming activities, information systems is increasingly getting used to provide actionable analytics, dictate genuine logistical decisions, and reimagine workflows in way to build game-changing marketing strategies. As a consequence, many authors anticipate that Big Data will cause substantial alterations in the responsibilities & economic dynamics between different parties in current food distribution system ecosystems. The partner panorama exhibits a fascinating interaction involving giant virtual enterprises, venturing capitalists, &, on occasion, little entrepreneurs as well as entrants. Most municipal institutions, make accessible data available on the provision provided customers' confidentiality be respected. From a sociological aspect, the authors suggest concentrating research on management policies including such accountability and acceptable business strategies for information exchange in diverse distribution network settings.

**KEYWORDS-** Agriculture, Big Data, Machines, Smart Farming.

## I. INTRODUCTION

Farming operations very soon became more data-driven as smart equipment and detectors seems to be available on farms and agricultural data propagates in amount & breadth [1][2]. Advanced Irrigation is getting spurred forward by significant advances inside the IoT based computing systems. Whilst advanced digital irrigation focuses only on in-field unpredictability, Precision Irrigation extends beyond by establishing authority to make decisions on statistics and also the locality, incorporating meaning and scenario understanding inspired by legitimate happenings. Real-time assisting reconstruction skills were also required to execute nimble tasks, especially in instances of drastically changed operational requirements and perhaps similar circumstances. Those functions often include cognitive assistance regarding equipment implementation, administration, plus use [3–6]. The concept of Agriculture 5.0 as a cyber-physical model is characterized in Fig. 1, which suggests that the agricultural system is governed by digital smart objects via World Wide Web. Smart devices improve conventional instruments by incorporating autonomous perspective thru multiple sensing units, constructed knowledge, & the capacity to undertake unsupervised or faraway operations [7–9]. Even it is hinted in this example that computers might perform a substantial role in supervision, it is expected that perhaps the activity of people in evaluation & strategy management will be gradually supplemented via machinery, culminating in a completely unsupervised cyberphysical process [10–12]. Creatures shall continually be engaged within progression, however at an upper cognitive level, with robots doing the majority of the operational tasks.

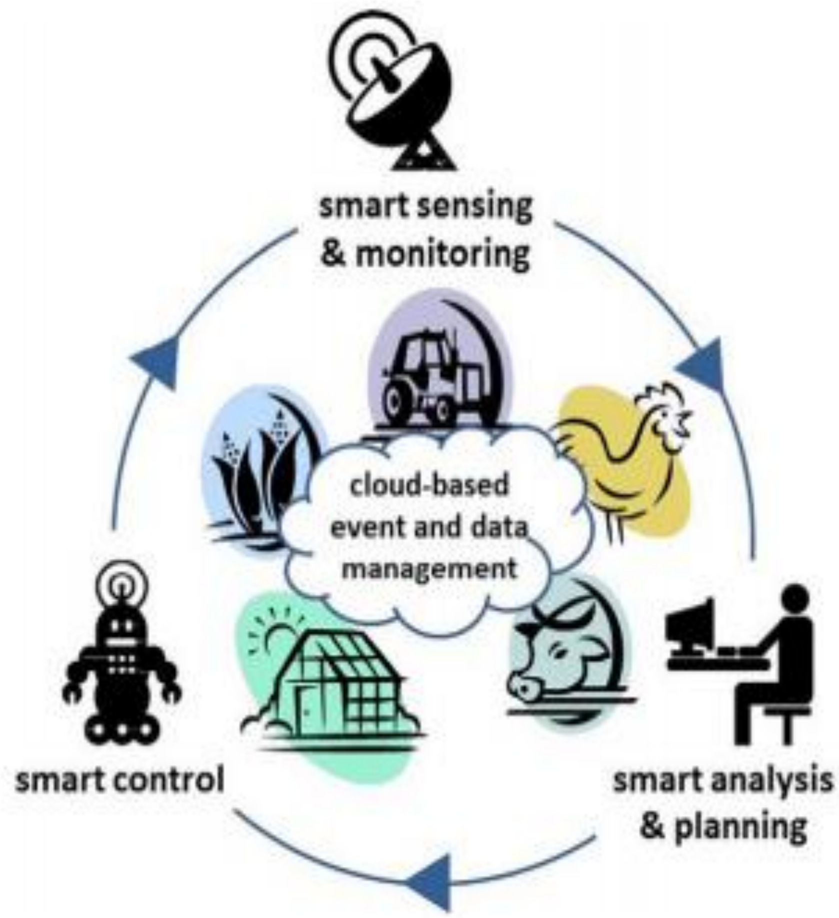


Fig. 1: Illustrates cycle of Intelligent Agro in cyber-physical management [13].

Devices usually are outfitted using a range of sensing instruments that collect knowledge from respective surroundings and utilize it to direct the equipment' conduct [14][15]. Characteristics Of gigantic information perform a significant, mutually reinforcing purpose in the advanced: equipment are outfitted with a range of technology to detect knowledge in immediate surroundings and utilize that information to control the robots' behaviour [16][17]. It might vary beyond simple control schemes to advanced deep neural learning to implement crop protection strategy [18–20]. This is improved by combining it with other Massive Information resources such as climate or marketplace information, along with references from several of the similar farms. Because of the fast pace of change in this field, it's difficult to provide a unified definition of Big Data (BD), although it's usually defined as data sets that are so big or complicated that conventional data processing applications are insufficient. Big data necessitates a collection of methods and technologies, as well as new kinds of integration, in order to extract insights from large, complicated information. Big Data refers to information assets with such a large volume, velocity, and diversity that their translation into value necessitates the use of specialized technology and analytical techniques. Characteristics Of gigantic information perform a significant, mutually reinforcing purpose in the advanced: equipment are outfitted with a range of technology to detect knowledge in immediate surroundings and utilize that information to control the robots' behaviour. The FAIR principle, which states that data should be Findable, Accessible, Interoperable, and Reusable, is emphasized by the Data FAIRport project, which stresses the more operational component of Big Data. This also indicates the significance of metadata, or information about information (e.g. time, location, standards used, etc.) [21–23].

## II. LITERATUE REVIEW

Between Feb 2009 & Feb 2014, we examined literature to answer criticism inquiries stated in the Outline. The study period was chosen for practical reasons, taking into account the circumstance that the information is a relatively new phenomena; it wasn't anticipated to be any references prior to 2009. We utilized two inclusion criteria for the literature search, in addition to the publication period: 1) publishing of the whole article; 2) applicability of review topic There were several prohibiting measures cast-off: 1) articles written in languages other than English; 2) courses that are exclusively concerned with technical design. A systematic method was used to conduct the literature review. This was accomplished in three stages. The both datasets were selected for the extensive exposure of appropriate collected works as well as sophisticated bibliometric features like related literature or citation suggestions. 613 peer-reviewed papers were found in these two databases. By selecting sections that addressed the study topics, they were examined for relevance. When seeing books, we utilized mainly searching tool in order to find paras that included chief terms, & thereafter visualize the content to determine if they might be linked to the study inquiries. Four researchers conducted the screening, with each evaluating around 149 courses & communicating the results with many of others using the locus supervision program EndNote X7. Accordingly, 19 were deemed very important, while 95 were deemed somewhat pertinent. The other paper too was deemed irrelevant since they only indirectly addressed Big Data, so were not read or analyzed further. We discovered a limited amount of books, which may be explicated by the fact that Smart Agriculture is reasonably fresh ideas. Usages, in particular, are constantly developing & are unlikely to be taken into consideration in studied papers that usually are often behind. As a result, we chose to incorporate grey literature in our analysis. We searched for periodicals, blogging sites, and many of the related sites in English using different search engines. This resulted in three reports, 225 magazine pieces, 319 blog posts, and 19 tweets. The relevancy of each of the 319 blogs was determined by looking at the heading and paragraphs that included the searching keywords. Promising doublings were also eliminated. As a consequence, a short list of 29 blogs was created, each of which was assessed after further reading. As a consequence, 9 blogs have been identified as providing content that is relevant to our approach. The relevancy of 193 magazine publications was equally assessed based on the title and phrases containing the search keywords. The end outcome is a set of 19 papers once duplicates are removed. We then went through these 25 articles again to assess them further. As a result, 9 articles have been identified as potentially useful for further investigation.

### *A. Conceptual Framework:*

For the above research, a cognitive architecture has been advanced to afford a structured grouping of disputes & conceptions for socioeconomic assessment of Business Intelligence systems in Precision Agriculture. Amongst few of the difficult aspects of certain solutions is that typically have to have collaboration from a wide range of parties who serve varied responsibilities in the knowledge logistics system. As a consequence, the structure blends statistical methods & studies into logistic access control. Integrated supply systems are built up of actors that cooperate together forwards and backwards to give advantage to customers. The concept of a production process, which is a structure of interrelated processes that generally adds quality to a commodity and function, has been the foundation of supply systems. In big data operations, the production process includes a series of events generally begins with data collecting and ends with knowledge commercialization. The reputed network planning moral foundation consists of three interrelated facets: system design, operational functions, & managerial characteristics. The system design is made up of partner firms and the links across the firms. Operational functions include acts that provide a certain real worth solution for the customer. The administrative factors that allow the channel's commercial processes to be interconnected & controlled are referred to as administrative aspects. There are two aspects to access control: technical and administration. Regarding the goals, the design was converted to networking for Business Intelligence systems in Precision Agriculture, as illustrated in Figure. 2.

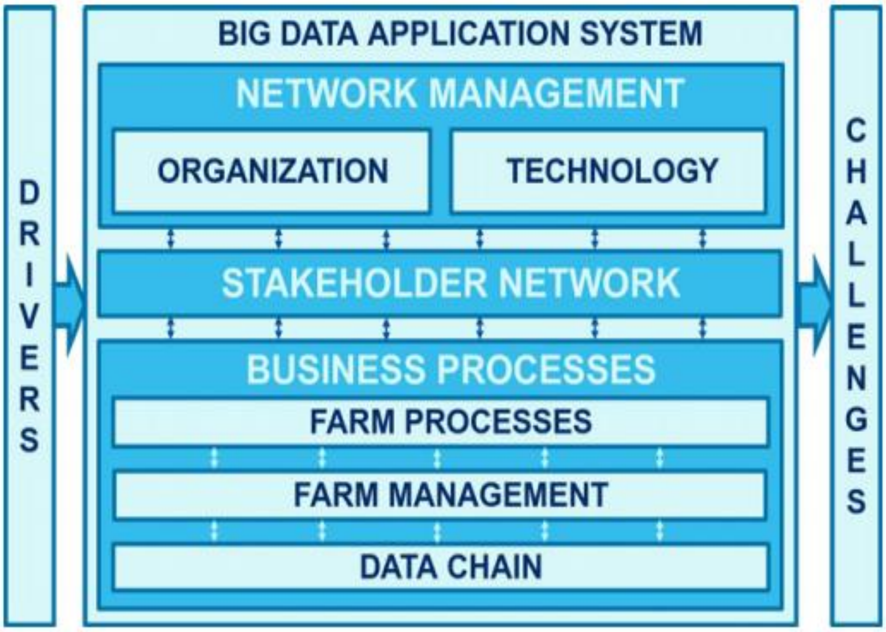


Figure. 2: Intangible structure for the analysis of collected works.

### *B. Data Chain:*

The data chain is the series of events that lead from data collection through decision-making and data marketing. It encompasses all of the actions required to handle data for farm management. The key stages in this chain are shown in Figure 3. The data chain, as an essential component of business operations, must have a technological layer that collects raw data and transforms it to information, as well as a business layer that makes choices and extracts value from data services and business intelligence. In each step, the two layers may be intertwined, forming the foundation of what has been known as the "data value chain."

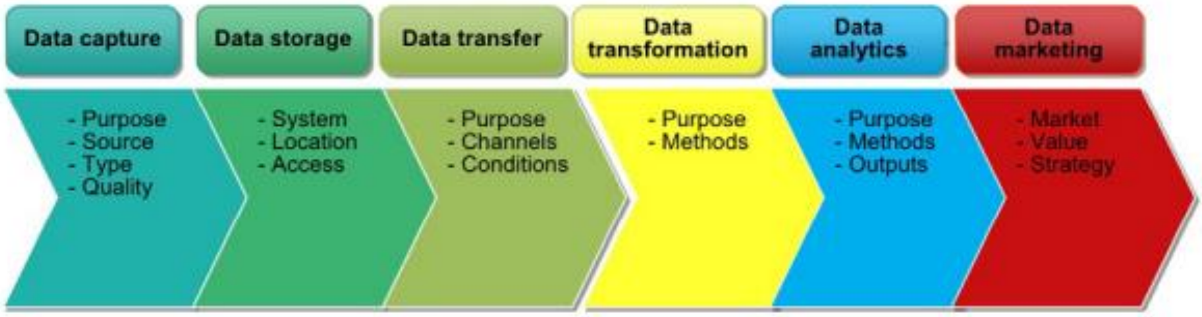


Fig. 3: The structure of usages of Big Data.

## III. DISCUSSION

### *A. Big Data Role in High-Tech Agriculture:*

It's a growing tendency to see use of Big Data and related methods and methodologies to cultivation as a big potential for technology stack application, investment, and creation of extra value in the agri-food sector. Big data applications in agriculture aren't only for primary construction; they too are important for increasing adeptness of the whole process & addressing etable security issues. Currently, the major of big data usage addressed in the literature occur in Europe and North America. However, given the increasing interest and attention exhibited in the literature, the number of applications in other nations, such as China, is anticipated to quickly increase. BD is the centre of sophisticated, game shifting commercial process at a size that makes traditional method of bootlegging and cleaning all of it into a data warehouse obsolete. Benchmarking, sensor deployment and analytics, predictive modelling, and utilizing improved models to control crop failure risk and increase feed efficiency in animal production are all possibilities for Big Data applications in agriculture

.Table 1: Illustrates the generalized factors running improvement of Big Data in Intelligent Agriculture.

| Push factors | Pull factors |
|---|---|
| • **General technological developments**<br>- Internet of Things and data-driven technologies<br>- Precision Agriculture<br>- Rise of ag-tech companies<br><br>• **Sophisticated technology**<br>- Global Navigation Satellite Systems<br>- Satellite imaging<br>- Advanced (remote) sensing<br>- Robots<br>- Unmanned Aerial Vehicles (UAVs)<br><br>• **Data generation and storage**<br>- Process-, machine- and human--generated<br>- Interpretation of unstructured data<br>- Advanced data analytics<br><br>• **Digital connectivity**<br>- Increased availability to ag practi-tioners<br>- Computational power increase<br><br>• **Innovation possibilities**<br>- Open farm management systems with specific apps<br>- Remote/computer-aided advise and decisions<br>- Regionally pooled data for scientific research and advise<br>- On-line farmer shops | • **Business drivers**<br>- Efficiency increase by lower cost price or better market price<br>- Improved management control and decision-making<br>- Better local-specific management support<br>- Better cope with legislation and paper work<br>- Deal with volatility in weather conditions<br><br>• **Public drivers**<br>- Food and nutrition security<br>- Food safety<br>- Sustainability<br><br>• **General need for more and better information** |

To summarize, Big Data is used to forecast future agricultural outcomes, drive real-time operational choices, and reinvent business processes for quicker, more creative action and game-changing business models. In the future, decision-making will be a complicated combination of human and machine elements. Farming's scope and structure are anticipated to alter as a result of big data. While there are concerns about whether algorithms will be able to replace farmers' expertise. The forces that run the growth of Big Data and Intelligent Agriculture are summarized in Table 1.

### *B. Occupational Progressions:*

#### *a) Farm Process:*

Agro Big Data is well recognized for its heterogeneity (24). The topic of the data gathered (i.e., what is the data about) and the methods by which data are produced are examples of data heterogeneity. Planting, spraying, materials, yields, in-season imaging, soil types, weather, and other activities are among the data gathered in the field or on the farm. There are three types of data creation in general:

- Process-mediated (PM).
- Machine-generated (MG).
- Human-sourced (HS).

Agro business that tops and observes commerce activities of concern, such as buying inputs, feeding, planting, applying fertilizer, accepting an order, and so on, provide PM data, or conventional business data. Industries, referencing tables, & linkages, as well as the information that defines their context, are common features of PM data. Traditional business data, which is often organized and kept in relational database systems, makes up the overwhelming bulk of what IT manages and processes in both operational and business information systems.

#### *b) Farm Management:*

Large Data, big or little, is still data, as Big Data watchers remind out. To get the most out of it, it has to be controlled and analyzed (25). Wireless networks, the Internet of Things, and cloud computing are basically just ways to get data and create Big Data. The ultimate goal of Big Data is to extract the information or insight that it contains or enables. Without Big Data analytics, agricultural Big Data will be useless. To get Big Data analytics, data from many sources must be combined into "data lagoons." Data quality problems are likely to occur throughout this procedure as a result of data mistakes and duplications. A sequence of procedures on the raw data, as illustrated in Fig. 4, may be required to guarantee data quality. The so-called data rich, information poor (DRIP) issues have been prevalent since the emergence of large-scale data collections or warehouses. The DRIP dilemma has been alleviated by the Big Data approach, which has liberated information in a way that may enable educated - but not always defensible or legitimate - judgments or choices. Thus, by addressing data quality problems with data quantity, data access limitations with on-demand cloud computing, causative analysis with correlative data analytics, and model-driven with evidence-driven applications, we may partially overcome data quality issues.

## IV. CONCLUSION

A literature study on Big Data applications in Smart Farming was performed in this article. It was determined in Section 2 that there are presently few references in peer-reviewed scientific publications. As a result, there was no way to conduct a valid quantitative analysis. Furthermore, results from grey literature may not be as scientifically sound as those found in peer-reviewed journal publications. Articles from grey literature, on the other hand, may be regarded as being susceptible to public examination and therefore fairly trustworthy since they are publicly accessible. As a result, we believe that publications from grey literature have added to the knowledge base. Aside from that, a lot of time and effort was invested into creating an analytical framework that may be utilized in future studies that take a more quantitative approach.

The analytical framework was created from the viewpoint of a chain network, with a focus on network management among the parties involved. In future study, it may be useful to look at this topic from a broader viewpoint of innovation. The same may be said about ethical elements of an invention, such as Big Data, which might be investigated more in the future. The potential of Big Data in agriculture is appealing, but the difficulties listed above must be solved in order for Big Data applications to gain traction. Although there are technological problems to be addressed, we suggest focusing first on the governance concerns that have been discovered and designing appropriate business models since they are the most limiting factors at the moment.

## REFERENCES

1. Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big Data in Smart Farming – A review. Agricultural Systems. 2017.
2. Kumar AS, Kumar DS, N MCR| SK. Impacts of Big Data on Smart Farming. Int J Trend Sci Res Dev. 2018;
3. Hirawat A, Taterh S, Sharma TK. A dynamic window-size based segmentation technique to detect driver entry and exit from a car. J King Saud Univ - Comput Inf Sci. 2021;
4. Sharma TK. Enhanced butterfly optimization algorithm for reliability optimization problems. J Ambient Intell Humaniz Comput. 2021;
5. Gupta H, Varshney H, Sharma TK, Pachauri N, Verma OP. Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction. Complex Intell Syst. 2021;
6. Gupta H, Kumar S, Yadav D, Verma OP, Sharma TK, Ahn CW, et al. Data analytics and mathematical modeling for simulating the dynamics of COVID-19 epidemic—a case study of India. Electron. 2021;
7. Sharma K, Goswami L. RFID based Smart Railway Pantograph Control in a Different Phase of Power Line. In: Proceedings of the 2nd International Conference on Inventive Research in Computing Applications, ICIRCA 2020. 2020.
8. Khatri M, Kumar A. Stability Inspection of Isolated Hydro Power Plant with Cuttlefish Algorithm. In: 2020 International Conference on Decision Aid Sciences and Application, DASA 2020. 2020.
9. Ghai W, Kumar S, Athavale VA. Using gaussian mixtures on

triphone acoustic modelling-based punjabi continuous speech recognition. In: Advances in Intelligent Systems and Computing. 2021.

10. Hahanov V, Gharibi W, Man KL, Iemelianov I, Liubarskyi M, Abdullayev V, et al. Cyber-physical technologies: Hype cycle 2017. In: Cyber Physical Computing for IoT-driven Services. 2018.
11. Khanna R, Verma S, Biswas R, Singh JB. Implementation of branch delay in Superscalar processors by reducing branch penalties. In: 2010 IEEE 2nd International Advance Computing Conference, IACC 2010. 2010.
12. Kumar Gola K, Chaurasia N, Gupta B, Singh Niranjan D. Sea lion optimization algorithm based node deployment strategy in underwater acoustic sensor network. Int J Commun Syst. 2021;
13. Barmpounakis S, Kaloxylos A, Groumas A, Katsikas L, Sarris V, Dimtsa K, et al. Management and control applications in Agriculture domain via a Future Internet Business-to-Business platform. Inf Process Agric. 2015;
14. Schönfeld M v., Heil R, Bittner L. Big Data on a Farm—Smart Farming. In 2018.
15. Gupta B, Gola KK, Dhingra M. HEPSO: an efficient sensor node redeployment strategy based on hybrid optimization algorithm in UWASN. Wirel Networks. 2021;
16. Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. Int J Inf Manage. 2015;
17. Oussous A, Benjelloun FZ, Ait Lahcen A, Belfkih S. Big Data technologies: A survey. Journal of King Saud University - Computer and Information Sciences. 2018.
18. Gupta N, Jain A, Vaisla KS, Kumar A, Kumar R. Performance analysis of DSDV and OLSR wireless sensor network routing protocols using FPGA hardware and machine learning. Multimed Tools Appl. 2021;
19. Gupta N, Vaisla KS, Jain A, Kumar A, Kumar R. Performance Analysis of AODV Routing for Wireless Sensor Network in FPGA Hardware. Comput Syst Sci Eng. 2021;
20. Kumar A, Jain A. Image smog restoration using oblique gradient profile prior and energy minimization. Front Comput Sci. 2021;
21. Stovold RGH, Adams M, Maier SW, Donald G, Gherardi S, Broomhall M. Using Pasture Growth Rate Data in a National Agricultural Drought Assessment Monitoring Tool. In 2009.
22. Solanki MS, Sharma DKP, Goswami L, Sikka R, Anand V. Automatic Identification of Temples in Digital Images through Scale Invariant Feature Transform. In: 2020 International Conference on Computer Science, Engineering and Applications, ICCSEA 2020. 2020.
23. Goswami L, Kaushik MK, Sikka R, Anand V, Prasad Sharma K, Singh Solanki M. IOT Based Fault Detection of Underground Cables through Node MCU Module. In: 2020 International Conference on Computer Science, Engineering and Applications, ICCSEA 2020. 2020.
24. Hesse A, Bertulat S, Heuwieser W. Survey of work processes on German dairy farms. J Dairy Sci. 2017;
25. Kaloxylos A, Eigenmann R, Teye F, Politopoulou Z, Wolfert S, Shrank C, et al. Farm management systems and the Future Internet era. Comput Electron Agric. 2012;