Agri Articles
(e-Magazine for Agricultural Articles)
Volume: 04, Issue: 01 (JAN-FEB, 2024)
Available online at http://www.agriarticles.com


# The Advent of Big Data and AI in Public Health

(*Dr. Brindha S)
PhD Scholar, Department of Veterinary Public Health and Epidemiology,
College of Veterinary Sciences, GADVASU, Ludhiana, Punjab, India
*Corresponding Author's email: brindhasundar18@gmail.com

## Abstract
The era of Big Data, characterized by the need for substantial computational resources to handle the ever-increasing volume and complexity of data, has embraced cyber-physical systems, cloud computing, and the Internet of Things (IoT), collectively known as Industry 4.0. These large-scale data processing systems often involve significant levels of process automation. The global focus on Big Data has been coupled with the rise of artificial intelligence (AI) in diagnostics and decision-making. AI, exemplified by software programs like the Generative Pre-trained Transformer version 3 (GPT-3), can simulate a context-sensitive response or a conversation with a human user in natural language. AI models can contribute to nearly all real-world use cases, from customer service to personal assistance. They can be integrated into existing websites and mobile apps via their APIs, performing tasks ranging from answering frequently asked questions to providing personalized recommendations and grammar corrections. From a public health perspective, AI-based applications can offer accessible, cost-efficient, and interactive solutions for health education and promotion. They can assist in the self-management of chronic illnesses, provide access to remote or automated health services, screenings, diagnosis, and therapy, and offer emotional support for mental health issues.

## Introduction
The era of Big Data is characterized by the need for substantial computational resources to handle the ever-increasing volume and complexity of data from diverse sources, including the internet and remote sensor networks. This data, which can be structured, semi-structured, or unstructured, often involves intricate interrelationships that are syntactic, semantic, social, cultural, economic, and organizational. Furthermore, epistemic uncertainties, such as data corruption by noise and artefacts, data entry errors, duplicate records, missing records, and incomplete digital records, can hinder decision-making processes for both humans and computers. The Big Data culture has embraced cyber-physical systems, cloud computing, and the Internet of Things (IoT), collectively known as Industry 4.0. These large-scale data processing systems often involve significant levels of process automation. The global focus on Big Data has been coupled with the rise of artificial intelligence (AI) in diagnostics and decision-making, spurred by recent advances in computer technology, the rapid growth of social media, and the declining cost of microelectronics in connected devices.AI, exemplified by software programs like the Generative Pre-trained Transformer version 3 (GPT-3), can simulate a context-sensitive response or a conversation with a human user in natural language. AI models can contribute to nearly all real-world use cases, from customer service to personal assistance. They can be integrated into existing websites and mobile apps via their APIs, performing tasks ranging from answering frequently asked questions to providing

personalized recommendations and grammar corrections. From a public health perspective, AI-based applications can offer accessible, cost-efficient, and interactive solutions for health education and promotion. They can assist in the self-management of chronic illnesses, provide access to remote or automated health services, screenings, diagnosis, and therapy, and offer emotional support for mental health issues. This paper explores these possibilities and their implications for the future of public health.

## Capabilities of Artificial intelligence (AI)

Artificial intelligence (AI) and natural language processing (NLP) are two rapidly evolving fields that have the potential to revolutionize many aspects of our lives, including healthcare. AI research has led to the development of models capable of interpreting radiographs, detecting irregular heartbeats via smartwatches, identifying reports of infectious diseases in media, determining cardiovascular risk factors from retinal images, and discovering new applications for existing medications. These models are trained on large volumes of controlled, labelled, and structured data, and their ability to provide continuous, rapid analysis of data could significantly alter how we promote health and manage diseases. AI systems have the potential to review and categorize the approximately 1.3 million research articles indexed by PubMed each year, monitor social media posts for signs of mental illness or outbreaks of foodborne illness or influenza, and interact with individuals seeking health-related information. NLP, a subset of AI, focuses on developing algorithms and models that can use language in the same way humans do. It is commonly used in virtual assistants and search engines, and it allows for the analysis and extraction of information from unstructured sources, automation of question answering, and sentiment analysis and text summarization. Given that natural language is the primary means of knowledge collection and exchange in public health and medicine, NLP is crucial for unlocking the potential of AI in biomedical sciences. Modern NLP platforms are built on models refined through machine learning techniques, which involve a model, data, a loss function to measure how well the model fits the data, and an algorithm for training the model. Recent advancements in these areas have led to improved NLP models powered by deep learning, a subset of machine learning. Innovations in different types of models, such as recurrent neural network-based models (RNN), convolutional neural network-based models (CNN), and attention-based models, have enabled modern NLP systems to capture and model more complex linguistic relationships and concepts than simple keyword searches. This has been facilitated by vector-embedding approaches that encode words before feeding them into a model, recognizing that words exist in context and treating them as points in a conceptual space rather than isolated entities. The performance of these models has also been enhanced by transfer learning, which involves using a model trained for one task as the starting point for training on a related task. Advances in hardware and increases in freely available annotated datasets have further boosted the performance of NLP models. New evaluation tools and benchmarks are helping to expand our understanding of the type and scope of information these new models can capture.

## Future scopes

Public health is a field dedicated to achieving optimal health outcomes across various populations. This is accomplished by developing and implementing interventions that target preventable causes of poor health. The success of these interventions hinges on several factors. First, it's crucial to effectively quantify the burden of disease or disease risk factors in the population. This involves gathering data on the prevalence and impact of diseases and health conditions within a specific population. Second, it's important to identify groups that are disproportionately affected or at risk. This requires a thorough understanding of the social determinants of health, including socioeconomic status, education, physical environment, employment, and social support networks, among others. By identifying these at-risk groups, public health professionals can tailor interventions to meet the specific needs of these populations. Third, determining best practices is a key component of public health. This involves identifying the most effective prevention or therapeutic strategies for a given health issue. Best practices are typically informed by a robust body of scientific research and are continually updated as new evidence emerges. Finally, measuring outcomes is a critical part of any public health intervention. This involves tracking the progress of interventions over time and assessing their impact on population health. Outcome measures can include changes in disease prevalence, improvements in health behaviors, reductions in health disparities, and improvements in quality of life, among others. The PICO (patient/problem, intervention/exposure, comparison, outcome) concept is a key strategy in this evidence-based decision-making process. It provides a structured framework for formulating research questions and conducting systematic reviews of the scientific literature. The PICO framework helps ensure that decision-making in public health is grounded in the best available evidence. In today's digital age, the rate of information production and publication is unparalleled. This includes a wide range of sources, such as scientific literature, technical reports, health records, social media posts, surveys, registries, and other documents. The sheer volume of information available presents both challenges and opportunities for public health. This is where Natural Language Processing (NLP) comes in. NLP is a field of artificial intelligence that provides a powerful tool for quickly analyzing vast amounts of text. It involves teaching computers to understand, interpret, and generate human language in a way that is both meaningful and useful. NLP opens new avenues for research and evidence-based decision making in public health. It can aid in identifying populations, interventions, and outcomes for disease surveillance, prevention, and health promotion. For instance, NLP platforms can detect specific features in unstructured data like medical records or social media posts, enhancing surveillance systems with real-world evidence. NLP methods have even been used to predict depression before it was recorded in medical records. This kind of predictive modelling can help public health professionals intervene earlier and potentially prevent the onset of disease. Real-time text mining of scientific publications using NLP can provide timely, evidence-based recommendations during critical times, such as outbreaks. This can help ensure that public health responses are informed by the most current body of evidence.

## References


1. Benke, K., & Benke, G. (2018). Artificial intelligence and big data in public health. *International journal of environmental research and public health*, *15*(12), 2796.
2. Jungwirth, D., & Haluza, D. (2023). Artificial intelligence and public health: an exploratory study. *International Journal of Environmental Research and Public Health*, *20*(5), 4541.
3. Thiébaut, R., & Thiessard, F. (2018). Artificial intelligence in public health and epidemiology. *Yearbook of medical informatics*, *27*(01), 207-210.
4. Panch, T., Pearson-Stuttard, J., Greaves, F., & Atun, R. (2019). Artificial intelligence: opportunities and risks for public health. *The Lancet Digital Health*, *1*(1), e13-e14.