

# Biodata Sciences Infancy: Applications Expectations and Trends

Ewa Liwarskan\*

Department of Environmental Engineering, University of Lodz Technology, Politechniki, Poland

## Introduction

Biology has developed into a quantitative scientific field and is becoming increasingly digitalized. The widespread use of high-throughput technological platforms in biological research, which enable the systematic collection, cleaning, storage, and integration of millions of data points on genes, proteins, and other biological moieties from thousands of tissues and organisms, is a key driver. It is not an exaggeration to say that the strategic maximization of data is the future of biological (and biomedical) research in clinical samples settings. Big biodata is not a far-off dream. Not only have we already entered the age of big bio-data, but biological data are also being generated and accumulated at an ever-increasing rate. The unraveling of the human genome took several years and involved a large number of teams, costing approximately \$2.7 billion from 1990 to 2003. By 2016, the same experiment no longer costs more than \$1,500 and only takes an afternoon in a single lab. Similarly, whereas initially it took a global consortium five years to map a single tomato genome, 150 distinct tomato genomes can now be completed within a year. The Expression Atlas and the proteomics identification (PRIDE) database are two examples of big data management systems that have emerged as a result of the big bio-data landscape.

## Decription

In order to gain a better understanding of diseases and improve one's health, it is necessary to make use of the growing amount of big bio-data. Data generation, management, analysis, visualization, and interpretation issues should take center stage, necessitating a paradigm shift in mindset and expertise. Additionally, improved algorithm efficiency and advancements in database management platforms are required to deal with larger data volumes. When a lot of data is collected, problems with veracity and complexity arise as well. These problems need to be fixed as soon as possible. The challenges facing conventional fields like computational biology and bioinformatics are now greater than ever. In today's technological landscape, data science and artificial intelligence (AI) have already acted as innovation drivers in industries like finance and business. Instead of working in operations behind the scenes, data scientists direct the transformation of data into useful insights. In financial technology (fintech), AI-driven algorithmic trading and stock recommendation systems are two examples. In engineering, automated engine design, system maintenance, and robotics are also examples. We predict that alongside the rapid and extensive generation of biological data, a new type of data science that will specifically address domain-specific issues relevant to biology will emerge, given the

recent data explosion and concomitant advances in data science in other fields like business, finance, and computing. This subset of data science is referred to as "bio-data science (BDS)" by us.

It is incorrect to consider BDS additively in terms of the disciplinary cores. BDS is more than its components. Data storytelling is frequently compared to data science. Additionally, these peculiarities must be carefully taken into account when interpreting the data in order to tell a compelling story. To put it another way, the core disciplinary areas of BDS need to work together. Proteins operate as functional units known as protein complexes rather than independently, illustrating the significance of domain and synergy with statistics. In order for a complex to function, its components must tightly co-express in order for the complex to form. However, a fundamental assumption of many statistical tests is that each gene or protein operates independently when we interpret a matrix of gene or protein expression from a purely statistical perspective. This indicates that, despite the fact that the genes under consideration are not independent of one another (for instance, two proteins in a protein complex typically have correlated expression profiles), we correct for the total number of genes when attempting to reduce false positive rates. Overcorrection as a result of independence assumption results in a loss of statistical power. Corrections based on the potential number of protein complexes that could be formed would have been more reasonable in such instances. As a result, the biological domain not only generates the inquiries that need to be answered, but it also provides constraints that must be comprehended and taken into account in order to construct reliable models [1-3].

## Conclusion

There are three main areas of study in BDS: Computer science, mathematics, and statistics are all included in the application domain of biology (Figure 1). Questions regarding the biological origin, such as the cause of a disease or comprehending the diagnostic utility of an inferred biomarker, are the focus of the biology core area. Problem-solving, dealing with repetition (such as repeatedly running the same algorithm on large subsets of data), and resolving data storage issues are all topics that fall under the purview of the core area of computer science. Modeling, data summarization, and normalization are all topics that fall under the purview of the core area of mathematics and statistics. Although descriptive and exploratory statistical data analysis is not unique to BDS, it is an important part of biostatistics and, to a lesser extent, bioinformatics. However, BDS focuses more on making predictions with new technology that uses AI/ machine learning (ML) on big data [4,5].

## References

1. Weis, Judith S., and Daniel Sol, eds. *Biological invasions and animal behaviour*. Cambridge University Press, 2016.
2. Bartomeus, Ignasi, Montserrat Vilà, and Luís Santamaría. "Contrasting effects of invasive plants in plant-pollinator networks." *Oecologia* 155 (2008): 761-770.
3. Bartuszevige, Anne M., and Bryan A. Endress. "Do ungulates facilitate native and exotic plant spread?: Seed dispersal by cattle, elk and deer in northeastern Oregon." *J Arid Environ* 72 (2008): 904-913.
4. Bjerknes, Anne-Line, Ørjan Totland, Stein Joar Hegland, and Anders Nielsen.

\*Address for Correspondence: Ewa Liwarska, Department of Environmental Engineering, University of Lodz Technology, Politechniki, Poland, E-mail: Ewa.Liwarska11@lboro.ac.uk

**Copyright:** © 2022 Liwarska E. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Received:** 02 November 2022, Manuscript No. jbabm-23-87235; **Editor assigned:** 04 November 2022, PreQC No. P-87235; **Reviewed:** 16 November 2022, QC No. Q-87235; **Revised:** 21 November 2022, Manuscript No. R-87235; **Published:** 28 November 2022, DOI: 10.37421/2168-9768.2022.14.356

"Do alien plant invasions really affect pollination success in native plant species?." *Biol Conserv* 138 (2007): 1-12.

5. Botero, Carlos A., Franz J. Weissing, Jonathan Wright and Dustin R. Rubenstein. "Evolutionary tipping points in the capacity to adapt to environmental change." *PNAS* 112 (2015): 184-189.

**How to cite this article:** Liwarska, Ewa. "Biodata Sciences Infancy: Applications Expectations and Trends." *J Bioanal Biomed* 14 (2022): 356.