Big Data

Agricultural Cyberbiosecurity Education Resource Collection

Authored by David Smilnak, Dr. Anne Brown, Dr. Joseph Simpson, Jaylan Day, and Dr. Hannah H. Scherer

What is Big Data?

Big data is everywhere. It's in your smartphone, the cash register you checked out from, your online purchases, health records, etc. That's why this topic is so hard to wrap our minds around. Big data is the continuously growing **volume, velocity, and variety** of information. What we do with that information is the other part. Professionals in careers like data analytics, data engineering, and risk analysis, among others, collect data and look for patterns and trends. Those patterns give us insights that allow us to predict what might happen next.

Applying those predictions is the next step. That's why online retailers show you ads for products you are considering buying. The data that is out there highlights trends. You may be in the market for a new car. Your web searches and other data come together to suggest to an algorithm that you want a specific car. Then you start seeing ads for that car. This might seem unnerving when coming from a retailer. What about doctors? They use this wide amount of information to help predict outbreaks. Police departments use it to solve crimes. The idea of big data is to uncover useful trends in all areas of everyday life.

Figure 1. Big Data 2267x1146 white, Camelia, Boban, https://commons.wikimedia.org/wiki/File:BigData_2267x1146_white.png, Creative Commons Attribution-Share Alike 3.0 Unported License

Big Data in Biological Sciences

In biology, the interactions between cells in our bodies represent a complex network of data sharing and outcomes. Often, we can quantify and use basic principles of mathematics and physics to predict and analyze the processes and ways cells communicate, interact, and function. The field of **Computational Science** uses these data interactions to create profiles and atomistic-level data to understand biological phenomena. This entire field of biology uses computer science, mathematics, data science, and biology to inform and expand our basic understanding of biological processes.

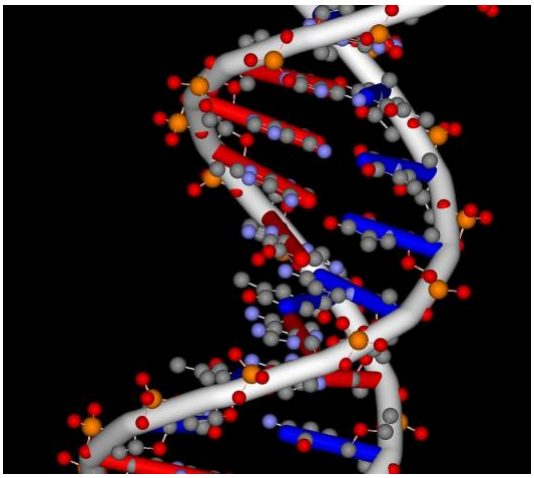

Figure 2. "DNA" by ghutchis is licensed under CC BY-ND 2.0.

Recently, we've seen a real-time example of how big data can contribute to biological science. The COVID-19 crisis collected a huge amount of data. That data was compiled and analyzed to track the pandemic. Some health agencies have even allowed you and me access to it through a phone app. Getting alerts about your individual risk of exposure is a direct impact of big data in biological science that you can see today.

Big Data in Agriculture

Agriculture today is far more advanced than you may realize. Big data is used throughout the industry. Farmers and ranchers use computers for everyday tasks, including ordering supplies, tracking crop and irrigation schedules, checking commodity prices, and keeping an eye on weather conditions. Farmers and ranchers have a history of observing changes on their farms and tracking them. Digitizing this data allows us to predict future changes and collaborate more quickly with others.

On the farm, **precision agriculture** is becoming more common. Precision agriculture uses measurements such as **LiDAR** maps of fields, soil moisture, crop stress, and nutrient levels to efficiently manage a field. These measurements are used to apply herbicides, pesticides, and fertilizers on a plant-by-plant basis. This can increase food production and decrease waste. In addition, we can use satellite networks to guide tractors, cell phones to run irrigation systems, and computer programs to tell us the best time to harvest. Big data in agriculture is going to become increasingly important as the global population increases and climate change progresses.

Figure 3. Sita Kumari, a farmer, uses mobile phone apps to enhance her yields and get access to markets and labor" by CGIAR System Organization is licensed under CC BY-NC-SA 2.0."

Figure 4. An autonomous tractor by John Deere. "Our Future" by adamthelibrarian is licensed under CC BY-NC-SA 2.0.

Connection to Cyberbiosecurity

Cyberbiosecurity is concerned with protecting the infrastructure, data, and products of biologically dependent sectors of industry. This includes things like proprietary seeds, health records, irrigation systems, and much more. Sectors including agriculture, biomedicine, and defense are generating and using more data than ever before. That data allows these sectors to work more effectively and do amazing things. However, it also provides a new target for attack and abuse. What makes this different from other computer attacks is the direct impact on people, animals, and plants.

One area where cyberbiosecurity and big data intersect is food security. In developing countries, famine is a major concern. The United Nations has used big data to create projects like the Global Pulse Initiative. This initiative uses a wide variety of data to predict famine and direct resources to curb the impact on human life.

Glossary

Computational Science: The art of creating mathematical models of a scientific process.

Precision Agriculture: Technology that maximizes the efficiency of a farm but is not critical to farming operations.

LiDAR: A detection system that works on the same principle as radar, recording the amount of time between an emitted electromagnetic wavelength

and receiving its reflection to measure distance, but uses light from a laser.

The 3 V's of Big Data:

- **Volume:** The amount of data produced and/or available for analysis.
- **Velocity:** The speed at which we create, process, and receive data.
- **Variety:** The different types of data, i.e. spatial, temporal, discrete, and continuous.

References

University of Wisconsin. (2015, May 18). *What Is Big Data?* University of Wisconsin Data Science Degree. https://datasciencedegree.wisconsin.edu/data-science/what-is-big-data/

Dajose, L. (2019, June 04). Biology and Big Data. *Caltech Magazine, 82*(2), 24-29. https://magazine.caltech.edu/post/biology-and-big-data

Shekhar, S., Schnable, P., LeBauer, D., Baylis, K., & VanderWaal, K. (2017). Agriculture Big Data (AgBD) Challenges and Opportunities from Farm to Table: A Midwest Big Data Hub Community Whitepaper. *White paper for the US National Institute of Food and Agriculture.*

Ramsey, F., & Seyyedhasani, H. (2021). Cyber attacks in agriculture: Protecting your farm and small business with cyberbiosecurity.

Murch, R. S., So, W. K., Buchholz, W. G., Raman, S., & Peccoud, J. (2018). Cyberbiosecurity: An Emerging New Discipline to Help Safeguard the Bioeconomy. *Frontiers in Bioengineering and Biotechnology*, 39. https://doi.org/10.3389/fbioe.2018.00039

UN Global Pulse—Big Data for Development and Humanitarian Action. (2018, May 31). https://www.unglobalpulse.org/

Additional Resources

The Top 10 Highest-Paying Big Data Careers **https://www.northeastern.edu/graduate/blog/highest-paying-big-data-careers/**

About the authors

This resource was developed by faculty and students at Virginia Tech: David Smilnak, Ph.D. Candidate in the Department of Agricultural, Leadership, and Community Education; Dr. Anne Brown, Assistant Professor in the Department of Biochemistry and University Libraries; Dr. Joseph Simpson, Assistant Professor in the Pamplin College of Business; Jaylan Day, Undergraduate Student in the Department of Chemistry; Dr. Hannah H. Scherer, Associate Professor and Extension Specialist in the Department of Agricultural, Leadership, and Community Education.

Acknowledgments

This resource was developed, in part, with the help and expertise of the Center for Advanced Innovation in Agriculture (CAIA) through a CAIA-Graduate Student Association writing workshop held in 2022. Jordan Wilson contributed to an earlier version of this work.

This work was supported, in part, through the CCI Southwest Virginia Node Cyberbiosecurity Seed Grant program and the USDA National Institute of Food and Agriculture, Women and Minorities in Science, Technology, Engineering, and Mathematics Fields (WAMS) Grants Program, award #2020-38503-31950. Technical and publishing support was provided by the Open Education Initiative of the University Libraries at Virginia Tech.

About this project

Cyberbiosecurity is an emerging field that focuses on creating security measures for digital aspects of our food and agriculture systems, creating a structure and opportunity for a safe food system that can meet the large needs of a growing population and world. This educational resource was developed as part of a project to support formal and non-formal agricultural educators in integrating cyberbiosecurity topics and research-based strategies for engaging middle-school-aged girls in STEM into their educational programs.

The entire resource collection can be accessed here: https://doi.org/10.21061/cyberbiosecurity

The project is an outreach effort of the Virginia Tech Center for Advanced Innovation in Agriculture.

COLLEGE OF AGRICULTURE AND LIFE SCIENCES
CENTER FOR ADVANCED INNOVATION IN AGRICULTURE
VIRGINIA TECH

Did you know that you can customize and share your version of this Open Educational Resource?