

UOA

# Scientific

A student-run publication

October 2023  
Volume 3 No. 4

ISSN 2815-7893

## P1 - ARTIFICIAL INTELLIGENCE

Will AI End Scientific Thinking?

## P9 - EARLY DEVELOPMENT

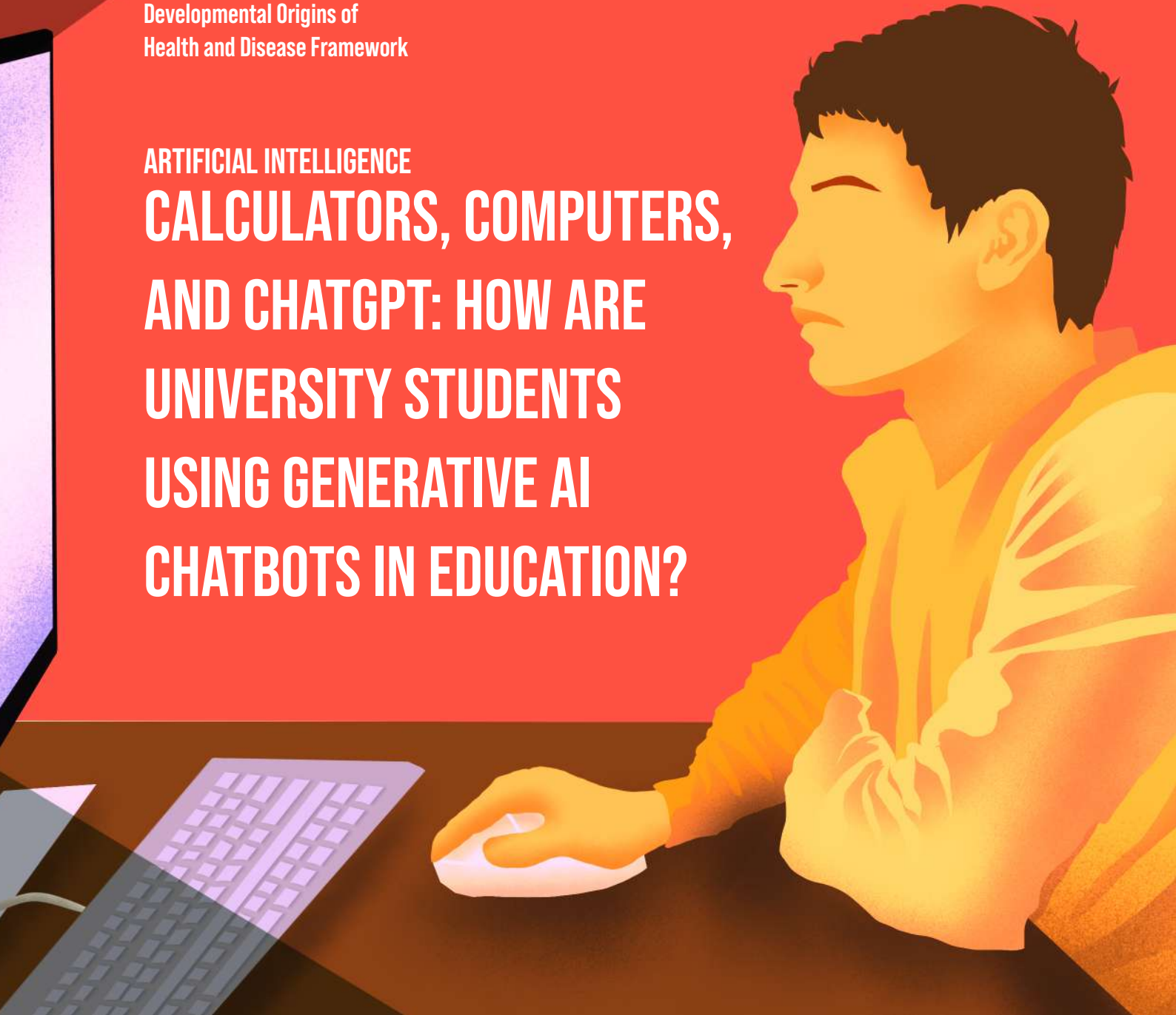
Why Early Life Matters Most: The  
Developmental Origins of  
Health and Disease Framework

## ARTIFICIAL INTELLIGENCE

**CALCULATORS, COMPUTERS,  
AND CHATGPT: HOW ARE  
UNIVERSITY STUDENTS  
USING GENERATIVE AI  
CHATBOTS IN EDUCATION?**

## P19 - COMPUTATIONAL BIOPHYSICS

Simulating the Mechanisms of Life  
Through the Lens of the Virtual World



## Editors' Note

Tēnā koutou Scientific readers! Welcome to the fourth edition of Volume 3.

With the end of semester approaching, and exams up ahead, we hope you've been taking care of yourself - and what better way to take a break than with our latest edition?

As assignments and exams occupy the forefronts of our minds, the role of artificial intelligence (AI) and ChatGPT has become a topic dominating many conversations around education. Taylor Lee explores the role of generative chatbots in education, particularly relating to our very own University, and their potential in fostering collaboration and literacy. Devin Ranasinghe examines the impact of AI on scientific thinking, practice, and analysis, while incorporating discussion on how AI is structured and discussions around some of AI's biggest questions.

Nadia Snegirev demystifies the revolutionary multi-disciplinary field of computational biophysics, while talking about the past, present, and future of this field. Noor Radhi explains the developmental origins of health and disease framework, emphasising the importance of early life development in defining the risk for non-communicable diseases, while Riya Balia writes on expectation-defying mimiviruses which challenge our assumptions about virology. Shyla Mani tackles the difficult topic of genetically modified crops, and how they may be impacting soil microorganisms. As always, we have something for everyone to enjoy!

As we wrap up another edition and another year of Scientific, we can't thank our readers, writers and all of our contributors and collaborators enough - a community is nothing without its individuals, and we've felt every bit of support our community has given our way.

Ngā mihi nui,  
Cindy Yi, Creative Officer for UoA Scientific 2023



## Will AI End Scientific Thinking? 1

This article aims to explore the impact of artificial intelligence (AI) on scientific thinking, practice, and analysis, encompassing discussions on topics such as “What is a pattern?” and “Why is AI not scientific?” while explaining the structure of AI.

Devin Ranasinghe

## 4 A Balancing Act: Managing the Impact of Genetically Modified Crops on Soil Microorganisms

How do genetically modified crops, addressing issues from pests to malnutrition, influence microbial genetics, diversity and structure? Research hints at subtle microbe shifts linked to GM crops, but causation remains unclear in the literature.

Shyla Mani

## Mimivirus: The Microbe-Mimicking Virus 7

What happens when viruses defy expectations? The discovery of giant viruses has altered our understanding of virology. This article briefly explores the genetics, evolution, and hosts of these unusual viruses, notably mimiviruses.

Riya Balia

## 9 Why Early Life Matters Most: The Developmental Origins of Health and Disease Framework

Early life development includes both prenatal and postnatal periods where interactions with the environment define disease risk. Here, we discuss the Developmental Origins of Health and Disease theory, plus interviews with Liggins Institute researchers.

Noor Radhi



## Calculators, Computers, and ChatGPT: How are University Students Using Generative AI Chatbots in Education? 14

A brief exploration and one of the first empirical studies of university students using artificial intelligence chatbots like ChatGPT in the educational context, and its potential to foster literacy and collaboration.

Taylor Lee

## 19 Simulating the Mechanisms of Life Through the Lens of the Virtual World

Computational biophysics — where has it come from, what does it look like, and where is it heading? A guide to this ‘new’ multi-disciplinary field revolutionising biology, demystifying the process and exploring the scope of what is possible.

Nadia Snegirev

## References 23

# Will AI End Scientific Thinking?

Devin Ranasinghe

Artificial Intelligence

Artificial intelligence (AI) improvements have made humans more cautious of AI. However, there seems to be one “blind spot.” That is, “What is the effect of AI on scientific thinking?”. We could call the scientific thinking process objective-abstract-imaginative thinking. To generate the above schools of thought, humans need to observe the objective world, utilise imaginative power to develop an idea, and then use abstract concepts to explain a particular phenomenon. Unlike scientific thinking, however, AI absconds core knowledge but still produces robust results. So, will the rise of AI end the era of scientific thinking?

## Part 1: The Loss of Gravity

Coming from a different background to New Zealand, I always wondered how different cultures, especially the Western world, perceive the world. One of the key features of Western thought (the Enlightenment) is how to perceive the world through scientific explanations.

One of the most prominent figures of the Enlightenment was Sir Isaac Newton, who discovered the famous laws of motion (LM) and universal gravitation (UG). LM and UG were revolutionary, as we all know, and led to the massive development of technology through classical mechanics and explained planetary motion. But what makes LM and UG scientific? Let’s examine UG further. UG is based on the observation of planetary motion in space (notably the works of Kepler). Using imagination and insights, Sir Isaac derived a **hypothesis** that actually resembles the UG law. Then, he formulated the laws using **abstraction**, especially calculus, in his famous book, *Philosophiæ Naturalis Principia Mathematica* [1]. Additionally, the laws of motion and gravity can be tested and proven by methods such as **induction** and **falsification** [2]. This process validates Newton’s laws as scientific and applicable universally (even though the theory of relativity came up with different explanations long after Newton).

Now, let’s examine how we use UG laws generally. We know UG laws by the famous equation:

$$F = \frac{Gm_1m_2}{r^2}$$

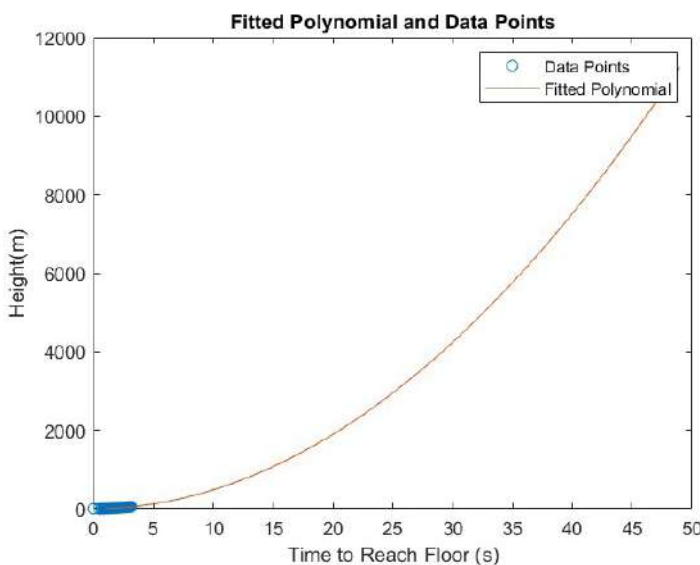


Figure 1: Polynomial fitting of height against the time taken to reach the floor generated on MATLAB. Image by Devin Ranasinghe.

We know that  $g$  (gravity) =  $GM/R^2 \approx 9.81 \text{ ms}^{-2}$  downwards (since acceleration is a vector). Then, we can derive some simple kinematic equations using the integration of acceleration to obtain the velocity and integrate again to acquire the distance. We can obtain the common kinematic equation,  $s = ut + 0.5at^2$ . As we know  $a = g$  downwards, we can calculate the time the ball took to hit the floor when the height to the floor is a known parameter.

However, there is another simple way of calculating or approximating the time the ball takes to hit the floor. Let’s consider a scenario: someone releases an iron ball from 0 to 49 m heights in 1 m increments and calculates the time taken for the ball to hit the floor at each height. We will have 50 data samples and time against the distance (height). Now, we can use methods like polynomial fitting to find a polynomial expression to fit in the table’s data.

### Glossary

**Hypothesis:** A supposition or proposed explanation made on the basis of limited evidence as a starting point for further experimentation.

**Abstraction:** Conceptualising a particular phenomenon to make it more understandable. In this context, it is by using scientific concepts and mathematics.

**Induction:** Generalising a scientific conclusion to a broader vicinity based on limited observation.

**Falsification:** For a theory to be considered scientific, it must be able to be tested and conceivably proven false. The process of trying to refute a scientific argument is called falsification.

We can utilise MATLAB functions for this. In my case, I used some data on MATLAB and obtained a six-degree polynomial expression (you may add Gaussian noise to the data). As readers, you can repeat the same on MATLAB easily. Then, we can use this polynomial function to estimate the time taken for the ball to reach the floor when it is released from 50 m. We can also use the kinematic equation ( $s = ut + 0.5at^2$ ) to calculate the time taken for the ball to reach the floor when released from 50 m. When we compare the result of the polynomial function and the kinematic equation, we can observe some errors between the data, but also a certain pattern that both data from the polynomial fitting and the kinematic equation follow.

Now, let's recall what we did and did not do. In polynomial fitting, we did not use or input the parameter for gravitational acceleration ( $g = 9.81$ ). We also did not obtain any kinematic equations. We only observe the height and the time to create a non-linear relation between the two variables. Therefore, we did not utilise UG or LM and their derivation. We only used polynomial fitting to find a relationship between the data, which happens to be non-linear. We are unaware of any relation between the mass of the iron ball and the time taken. We are also unaware of the impact on time when external forces, such as air resistance (speed and direction of wind), are exerted on the ball. We can also dispense the standard metric units like seconds, meters, and even more complex units ( $\text{ms}^{-2}$ ) and use more subjective measurements.

But do these issues matter when using the polynomial expression in real-life free-fall applications? We can add more parameters (mass and air resistance) to fit the polynomial equation and find a relation between the time and the distance (height). This is a fundamental principle of non-scientific thinking: thinking without any core knowledge (no scientific theory) of objects and how they are interrelated in the objective world [3]. It is a method of thinking that is purely based on creating non-theory relations between objects. Some can argue mathematics (abstract concepts) has been used here to create a non-linear relation. However, the mathematical operation used here is simple mathematics to find a relation between numbers. Another argument would be that this experiment belongs to the empirical observation in science. I agree that this is a scientific experiment, but what exactly is the conclusion of this experiment? Still, it absconds from a proper hypothesis fundamental to scientific thinking, which means this experiment is inspired by just day-to-day observation of free-fall objects. Also, it lacks a plausible scientific explanation of why this non-linear relation between time and height exists. This is not explained by a ground-breaking scientific discovery like the universal gravitation.

## Part 2: A Box in Black

The OceanGate's Titan implosion has been a major unfortunate incident reported by the media recently. The implosion has been attributed to different causes, from material failure to communication loss [4]. However, I was gobsmacked to learn that the Titan had made it to the Titanic wrecks in previous explorations. This fact has been overridden by the explanations of how scientific tests have not been adequately conducted during the manufacturing of the Titan. Had the Titan made it to the Titanic wrecks again, this would not have been news!

I am not implying that conducting scientific tests after the implosion is unimportant at all. But there should have been proper scientific conduct on how the Titan made it to the Titanic wrecks before, with possible failures of the submersible. Now, a **"Box in Black"** or **absence of scientific explanation** as to how the Titan explored the Titanic wrecks despite its failures that

were discovered after the implosion, is present.

The power of science can sometimes be limited by the individuals who conduct it. This is because there would be biases that many would follow unconsciously and are mostly limited to current scientific explanations and experiments. Artificial Intelligence (AI) can be applied to fill this gap where certain biases and ambiguities need to be redeemed.

Let's take a simple AI, or more precisely, the machine learning (ML) application. An image of a cat needs to be identified correctly from other animals by an ML algorithm. The first step is to acquire a collection of cat images and preprocess them to obtain data regarding rotation, flips, and colour contrast. Next, features such as shape (a cat shape), colour (different colours specific for cats), and unique features (fur, paws) would be extracted by the algorithm. Then, the ML algorithm would train the model to extract cat images from non-cat images. Humans can be involved in any of these processes; they can decide what features should be extracted. After the model is trained, it suddenly, from nowhere, can identify cat images.

Nevertheless, the beauty of ML lies in its architecture. It can occupy different architectures, but a Convolution Neural Network (CNN) would be the most appropriate one to utilise here [5]. CNN is a deep learning architecture where features are selected by feature extraction. It has several layers for functions such as preprocessing and applying non-linearity to emphasise certain features, downsampling, dropping out (retaining only prominent features), generating complex features (dense layer), and calculating the error. Previously, I mentioned humans can be involved in identifying which features need to be extracted, but in a CNN-based ML algorithm (unlike in a shallow learning algorithm, like linear regression, where features might be predefined by humans), the algorithm itself identifies them [6].

The "Black Box" here mainly lies in the feature extraction and its relationship to the functions used and weights on each connection. In this ML algorithm, no individual can know why each weight was assigned to each connection between neurons to create a relationship between each feature. The features we are discussing could be much more complex, derived from surface features, such as shapes or colour. While it is possible to identify these features after execution, why they were

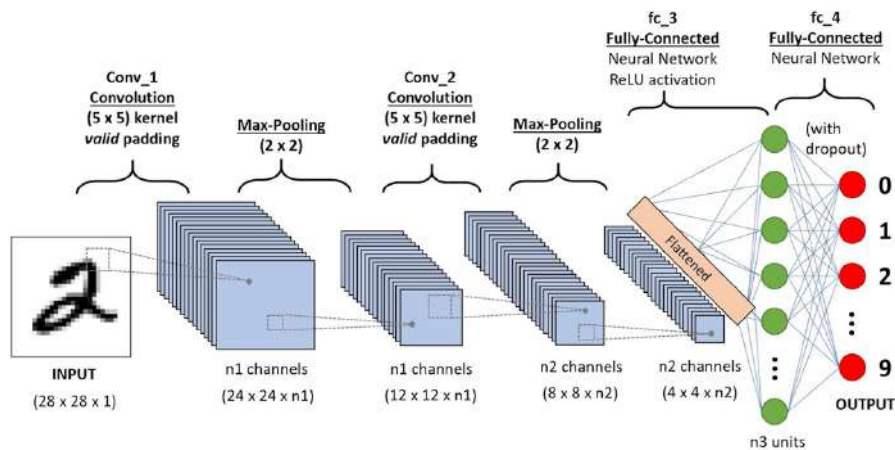


Figure 2: An architecture of CNN to identify written numbers [7]. Image by Analytic Vidya

considered more prominent is not evident. Additionally, the functions (such as ReLU, sigmoid, and tanh) employed for extracting features are paramount in deciding which features are prominent [5-6]. It is unclear why the functions remove some features while retaining others. The architecture will use different weights on each connection between neurons to identify and minimise the training error. Therefore, the prime target of the model is to minimise the error in assigning values on each connection. Complex features are just a result of reducing error rather than following specific guidelines to identify a feature. If we revert to the news on the Titan implosion, I always wonder if scientists really need to analyse the Titan based on scientific theories such as theories on fluid dynamics, pressure, material sciences, electromagnetism, and signal processing. Or else, they can just insert the data they have already acquired from the previous explorations of Titan and let an ML algorithm decide what features, say, to reinforce the hull, a submersible should comprise. Suppose the ML is input with data from previous explorations such as pressure, temperature changes, and the design deformations that occurred during exploration. Will it output a hull design based on creating complex connections between data?

### Part 3: A Doomsday for Science?

***“Artificial intelligence is growing fast, as are robots whose facial expressions can elicit empathy and make your mirror neurons quiver.”***

– Diane Ackerman

The main two aspects emphasised in Part 1 and Part 2 are,

1. Artificial intelligence (AI) is just a pattern recognition system.
2. The involvement of scientific theories is absent or extremely limited in AI.

There could be arguments, such as the AI model was developed on the basis of science and logic. While it has a regular architecture that follows, it does not mean the output of AI can be routed back to the input through its architecture. Most AI experts are now convinced that AI probably would

not be entirely understandable forever. Some precautions for AI ambiguity exist, such as explainable AI (XAI) [8]. These models consist of simple architectures that are more transparent and meaningful in interpretation. I request you to explore the topics of XAI and its recent development to realise the differences between AI and XAI.

What would be the immediate impact of AI on the science? I would say nothing because scientific theories and methods are still popular. Theoretical physics has evolved from classical physics, and AI is not substantially involved in the experimentation of theoretical physics yet, especially in quantum mechanics. AI has the capability to replace many realms of classical physics, but that is in the future’s hands to decide. It should be noted that AI is already more utilised in economic science and business environments, where more uncertainties are present [9]. As an example, AI is being used for stock price predictions where prices are highly unpredictable. However, as I mentioned in Part 1, physics theories make precise predictions of the future based on formulas. Economic theories, on the other hand, are not natural science theories like physics. They involve individual psychology, societal aspects, government interventions, and global alterations like the recent pandemic, which are highly uncertain. AI is rapidly used for economic predictions based on different atmospheres. If AI can be successfully used in such uncertainties, can’t it replace the scientific analysis of different phenomena and explain them better than scientific theories?



#### Devin Ransinghe - ME (Mechatronics)

Devin is an enthusiast of science, technology, and philosophy. He is passionate about discussing ideas related to technological improvements and their effect on human life. As an ardent reader of material related to the above fields, he tries to derive his interpretation of ideas that can sprout new perspectives.

## Explained

# A Balancing Act: Managing the Impact of Genetically Modified Crops on Soil Microorganisms

Shyla Mani

Plant Microbiology

Beneath the soil, inconspicuous microorganisms wield more influence than many of us realise. These microscopic bacteria, fungi, viruses and more dwell throughout soil ecosystems across the globe through a plethora of lifestyles [1, 2]. As key players in nutrient cycling and symbiotic partnerships with plants, they quietly sustain the delicate balance of nature [1, 2]. In parallel, genetically modified (GM) crops emerge as adaptive solutions in agriculture, tailored to modern challenges [3]. But when these genetically engineered crops are suddenly introduced, they pose a risk of altering the nearby microbiota that has adapted to survive in the presence of non-GM crops.

*A consequential query arises from this balancing act: How and where do human-engineered crops intersect with the deeply established domain of soil microorganisms? Is the impact of GM crops on soil microbes significant enough to even warrant concern?*

This article aims to delve into the connections existing between GM crops and microorganisms present in our soil. It explains why it is important to carefully study how GM crops might affect these microorganisms, if they do at all. Furthermore, it raises the question of whether additional research, collaborative efforts, and unambiguous guidelines are important to maintain a state of equilibrium between GM crops and microbes.

## Microorganisms in Soil Ecosystems

Microorganisms, although not visible to the naked eye, exert substantial influence beneath the soil environments. These diminutive organisms, including bacteria, fungi, and various other microscopic life forms, establish the fundamental framework of soil ecosystems [1, 2]. This is because they undergo essential processes that we might not always see but are necessary for the survival of all the organisms around them and the way our natural world functions [2].

In the soil, microbes (especially fungi and bacteria) engage in nutrient cycling and breaking down organic matter to release essential nutrients to be taken in by plants [4]. Additionally, these soil microbes can foster symbiotic relationships with plants, enhancing their nutrient uptake and overall health [4]. These plants then become a food source for many herbivorous and omnivorous organisms. Recognising the importance of their roles, we can see that microorganisms significantly contribute to keeping the soil and wider natural environment in balance [4, 5].

## Genetic Modification: Enhancing Crops for a Changing World

In modern agriculture, genetically modified (GM) crops have emerged as innovative solutions tailored to address evolving challenges and needs across agriculture and the environment, such as food insecurity, population growth, and the increasing demand for more nutritious alternatives to today's food [3]. These crops have had their plant DNA intentionally altered by techniques like CRISPR-Cas9 and gene insertion to enhance specific

traits [6, 7]. For instance, a crop's genetic code might be tweaked to make it more resistant to common pests, thus reducing the need for harmful chemical pesticides [6].

This targeted genetic modification isn't just the result of lab experiments but real response efforts to challenges we face today, such as increasing pest problems, climate change, and a lack of food and nutrient levels in crops for humans and other species [3, 7].

These advancements hold the key to addressing global malnutrition challenges, offering the prospect of producing heartier, more nutritious crops that can thrive in a changing and demanding environment, ultimately helping to meet the food requirements of an expanding population [3].

## Potential Impacts on Soil Microbes

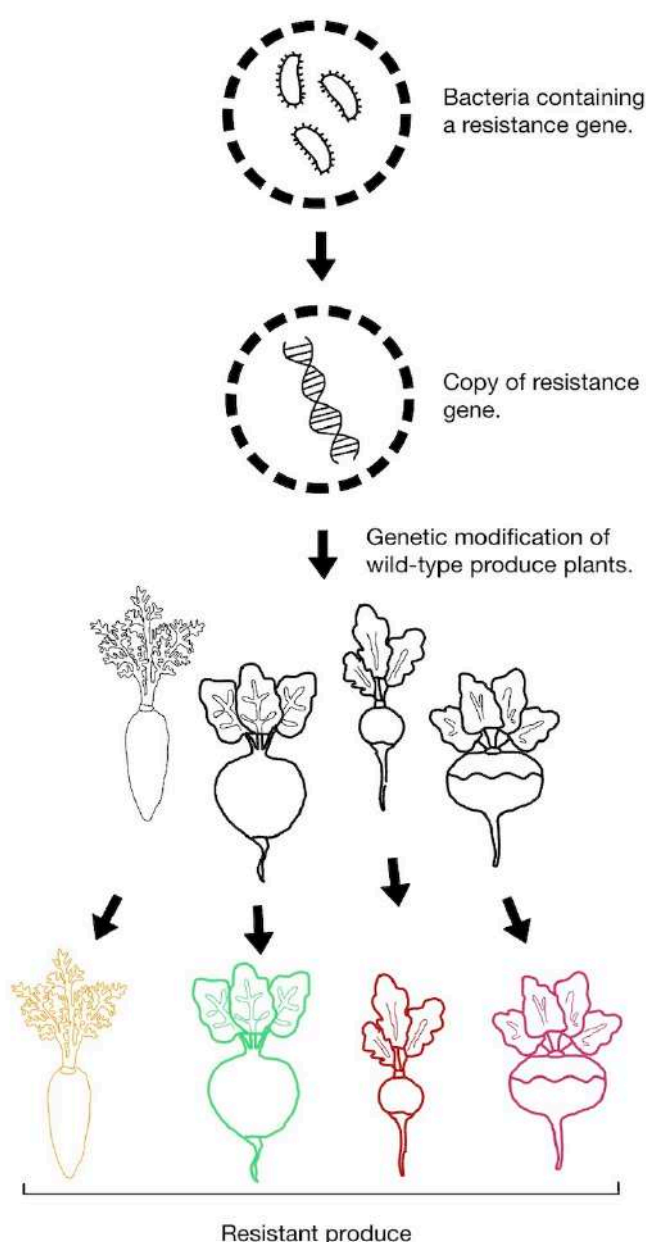
The increasing prominence of GM crop production and usage in our ever-evolving world has spurred investigations into their potential effects on soil microorganisms.

Research has unveiled that GM crops, particularly those produced to resist herbicides, can somewhat influence the composition of soil microbe communities that originally thrived in the presence of non-GM crops [2, 8]. Scientists have employed various methodologies, ranging from molecular analysis to microbial cultivation, to identify these shifts in soil microbe communities [2]. The most pronounced of these community changes are observed in the rhizosphere of GM plants (the region surrounding their roots) and during the decomposition of their plant material [1–4]. The rhizosphere plays a significant role in plant-microorganism interactions, impacting the growth and development of both parties involved [2]. Alterations in the composition of root exudates — substances released by plant roots into the rhizosphere — due to genetic modification, can tip the balance of microbial life in the soil [2].

Recent research has shown that genes

from GM crops can enter the soil through root exudates by horizontal gene transfer (HGT), which involves the transfer of genetic material between microorganisms [2]. For instance, proteins from GM plants, such as the insecticidal *Bacillus thuringiensis* toxin (Bt toxin), can persist in the soil, interacting with microorganisms and impacting their community structure [6]. The presence of the Bt toxin in GM plants, like *Gossypium* (cotton), is a result of genetic modification, entailing the insertion of specific Bt toxin-encoding genes into the plant's genome [6]. Typically derived from the bacterium *Bacillus thuringiensis*, this toxin acts as a natural pesticide targeting specific insect pests [6]. However, the introduction of Bt genes to the soil via

root exudates adds an additional layer of complexity to the impact GM crops have on soil microbiota [6]. This complexity arises from the potential for HGT of protein-coding genes between GM plants and soil microorganisms, influenced by shared DNA sequences within the microbial community [2, 6]. These proteins can persist in soil and interact with microorganisms, potentially altering their genetic characteristics and consequentially affecting their diversity, survival, and functionality [2, 6].



**Figure 1:** Illustration of Genetic Modification Process in Crop Enhancement. In the diagram, Tier 1 - 2 shows the extraction of a target resistance gene from a bacterium known to have it. Moving to Tier 2 - 3, the precise insertion of this resistance gene into the genomes of wild crop species is seen. The final tier shows genetically modified crops, colour-coded to highlight those now equipped with the inserted resistance gene. These modifications confer resistance to the specific challenges addressed by the incorporated gene.

It should be noted that whether observed changes in soil microbe communities, in response to GM crop presence, are beneficial, detrimental, or even significant to microbial survival/lifestyle remains blurry across the literature. This can depend on a range of biotic and abiotic factors in the soil, as well as the specific microbe-plant interactions at play [1]. Most studies find little to no difference between a microbial species surviving in the presence of GM crops as opposed to non-GM crops [1]. Furthermore, while studies have demonstrated observable modifications in the populations and structures of microbial communities around GM crops compared to non-GM crops, it is widely accepted that these changes are often inconspicuous against a backdrop of other natural environmental factors also affecting the plant-microbe interactions. This consensus spans numerous studies examining various soil ecosystems [1–8].

As our understanding of the relationship between GM crops and soil microbes continues to unfold, ongoing research remains indispensable. Advanced techniques, such as next-generation sequencing and DNA metabarcoding, will offer the potential for a more comprehensive understanding of how GM crops influence soil microorganisms [2].

### Sustainable Agriculture and Biodiversity in Aotearoa

In the context of Aotearoa's unique ecosystem, commitment to sustainable farming and environmental conservation, balancing genetic modification, our natural environment, and societal concerns over whether we want GM crops or not, become as important as it is challenging.

We need to find a way to navigate the complexities of producing GM crops while respecting the country's rich biodiversity [9]. This challenge goes beyond labs and academic circles as it mixes with legal and political concerns that



reflect the nation's values [3, 10]. Hence, researchers, policymakers, and legal experts must work together to figure out how GM crops fit into New Zealand's agriculture. This goal ensures that having enough food produced to fulfil rising demands also means maintaining our natural environment for future generations. It also links the possibility of more crops with safeguarding Aotearoa's native biota [10, 11].

Balancing ownership rights, protecting our native plants and animals, and giving everyone a fair chance to use new technology (especially when they have the potential to enhance our lives significantly), needs careful thought. Considering how GM crops might affect New Zealand's involvement in

international trade is also important [7, 10].

As New Zealand keeps moving forward, it is clear that the road to sustainable agriculture practices doesn't stop. This journey means that science, politics, and citizens must come together to create a world where GM crops can do well, all while keeping our "uniquely Aotearoa" natural environment in mind.



**Shyla Mani - BSc/LLB, Biological Sciences & Law**

Shyla is a 4th year student currently completing her BSc/LLB, majoring in Biotechnology & Law. She aims to combine her interests in environmental microbiology, IP law and science communication to enter the world of biotech innovation.

## Explained

# Mimivirus: The Microbe-Mimicking Virus

Riya Balia

Virology

In 1992, a microbiologist from Leeds Public Health Laboratory was sent to Bradford, a city in England [1]. His name was Timothy Rowbotham, and he was seeking the causative agent of a local pneumonia outbreak.

A sample he took from the warm waters of a cooling tower contained a microbe unlike anything anyone had ever seen before. Based on its size and resemblance to gram-positive bacteria under a light microscope, Rowbotham identified it as another pneumonia-causing bacterium and named it the Bradford coccus. It was not until 2003 that the samples reached Richard Birtles, a biochemist in France, who realised that the organism was not a bacterium; it was a giant virus [2]. This discovery led to its formal classification as *Acanthamoeba polyphaga mimivirus* (AMPV).

*Mimivirus* is not the only genus of giant viruses that has been discovered. *Megavirus* was isolated in 2010, followed by *Pandoravirus* in 2013 and *Pithovirus* the year after. More recently, *Tupanvirus* was described in 2018. The size, genome length, and complexity of these viruses are unlike standard viruses.

### Diameter

Given that the diameter of a hair strand is around 50,000 nm, it is difficult to fathom that the average width of a virus is approximately 20 nm to 100 nm [3-4]. Giant viruses, on the other hand, exhibit diameters exceeding 200 nm, with AMPV measuring 750 nm and earning its title of a “giant among giant viruses” [1]. As illustrated in Figure 1, the diameters of giant viruses are comparable to those of bacteria, which typically range from 200 nm to 2,000 nm [5]. This was one of the factors that contributed to Rowbotham’s misclassification of AMPV.

### Genome Length

Normal genome lengths of viruses range from 2 kb to 305 kb [7-8]. SARS-CoV-2, the virus that has disrupted our lives for the last few years, is approximately 29.8 kb [9]. As a shocking comparison, AMPV has one of the largest viral genomes sequenced to date at approximately 1,200 kb [10]. To put this into perspective, the genome size of AMPV surpasses that of some bacteria, as bacterial genome lengths range from 600 kb to 8,000 kb [11].

### Glossary

**Horizontal gene transfer:** Movement of genetic information between organisms, independently of the parent-to-offspring genetic transfer mechanism.

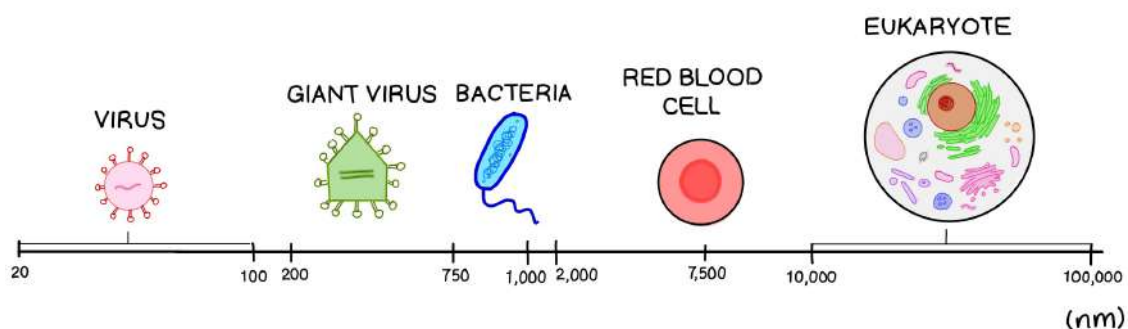
**Phylogenetic analysis:** Study of the evolutionary development of an organism using its genetic information.

**Myeloid cells:** Blood cells originating from the bone marrow that will mature into adult blood cells like monocytes, neutrophils, basophils, etc.

**Interferon:** Antiviral proteins involved in innate immune responses.

### Genetic Complexity

The collective genomes of the virus kingdom encode an average of 48 proteins per genome [12]. This number is skewed by giant viruses, such as AMPV, which encode around 1,000 proteins [13]. Many of these 1,000 proteins are uncharacterised or have cellular functions that have not been observed before in other viruses. A virus belonging to the *Mimiviridae* family, *Cafeteria roenbergensis* virus, has demonstrated this [14]. *C. roenbergensis* virus displays similar protein sequences from other domains of life. A significant portion, around 45% of its protein sequence, bears a resemblance to sequences typically found in eukaryotic cells, while 22% of



**Figure 1:** Relative sizes of biological entities on a logarithmic scale, measured in nanometers. Drawings are not to scale [6].

its protein sequence is comparable to that of bacterial cells; the rest is from viruses. Similar divisions in protein origin apply to other viruses within the *Mimivirus* genus [14].

These observations blur the boundaries between viruses and cellular life and compel us to explore theories regarding the evolutionary origins of giant viruses.

### Evolution

The reductive and expansive models are the two main hypotheses for the evolution of giant viruses [14]. According to the reductive model, the genetic downsizing of an ancestral cell leads to dependency on other cells. The presence of cellular functions (DNA repair machinery and protein synthesis components, with the exception of ribosomes) in nearly all giant viruses is consistent with this model. The expansive model suggests that genomically smaller ancestral viruses underwent gene duplication and **horizontal gene transfer** to evolve into current giant viruses. This model agrees with metagenomic studies that reveal gene exchanges between giant viruses and a variety of organisms. The expansive model is generally more accepted within the scientific community.

Some scientists believe that giant viruses constitute an independent domain of life alongside Archaea, Bacteria, and Eukarya. However, this theory remains highly controversial, primarily because **phylogenetic analysis** can be biased by horizontal gene transfer or inappropriate methodologies [15].

The idea of a virus evolving in size is fairly comical. Picture one rolling

down the street and eating everything in its path as it expands to the height of traffic lights! Fortunately (or unfortunately for fans of science fiction), giant viruses do not haunt our planet quite in that way. So, what exactly do they do?

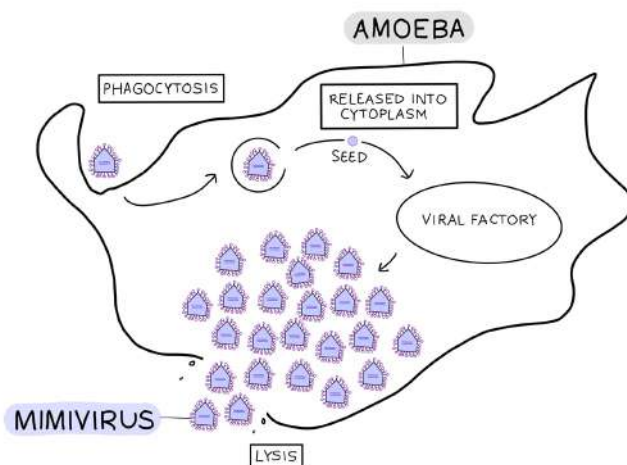
### Hosts

The primary hosts of mimiviruses are amoebas, which are unicellular organisms that have the ability to alter their own shape. As shown in Figure 2, the virus enters by phagocytosis, releases its seed in the cytoplasm, exploits the amoeba replication factory, and exits by cell lysis. Alas, the reach of mimiviruses extends well beyond amoebas. This genus has been detected in oysters, insects, monkeys, cattle, and humans [15].

As AMPV was discovered among bacteria that could potentially cause pneumonia, the question of whether it is a human pathogen naturally arises. Reports regarding the presence of AMPV antibodies in human serum have increased, and its genetic sequence has been found in gut microbiota and faeces [15-17]. Research has also demonstrated the ability of AMPV to replicate in **myeloid cells** and interfere with our **interferon** system [18]. Even so, the answer is not definite, as the position of giant viruses in human disease is an emerging area of research.

### Conclusion

Viruses are the most prevalent biological entities in the world [19]. Small but ubiquitous, their effects and mechanisms have been intensely researched for over a century. The discovery of giant viruses, however, has revealed complexities that challenge our previous assumptions. Although they have an ancient history, for virologists, giant viruses are the newest kid on the block. Accordingly, we should continue studying their genetics, evolution, and effects on humans to develop a deeper understanding of their role on our planet.



**Figure 2:** Replication cycle of mimiviruses in amoebas [16].



#### Riya Balia - BSc, Biomedical Science

Riya is a second-year Biomedical Science student specialising in infection and immunity. Aside from her interest in all things immunology, she is passionate about space, art, and fresh fruit ice cream.

## Explained

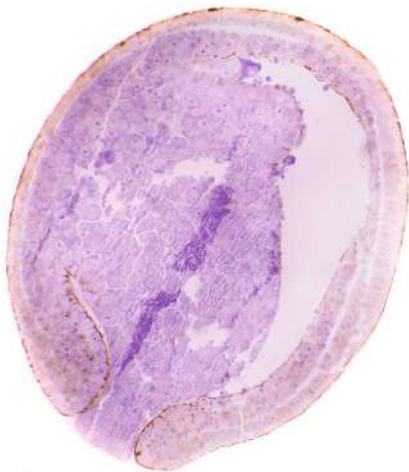
# Why Early Life Matters Most: The Developmental Origins of Health and Disease Framework

Noor Radhi

## Early Development

Early life is a stage we have all been through and come out of, where we have one shot to develop our body tissue, a functional immune system, and adapt to the complex extrauterine environment. The ability of a single cell, of size in microns, to generate a complete human being through continuous growth, feedback, and environmental cues is a miracle. It is literally a matter of life or death. Meanwhile, the capacity for this process to continue without much interference in coordinated and complex steps defends its marvel. This period of life includes both prenatal and postnatal development stages and is extremely important in setting up the neonate for the remainder of its life [1].

In this article, we will go through these early stages of life and discuss the Developmental Origins of Health and Disease framework, an important theory gaining traction in biomedical literature for its importance in influencing the neonate's future health outcomes based on environmental exposures in early life. I have recruited some of the Liggins Institute's accomplished researchers to examine these concepts and explore the ongoing research at the university aimed at tackling the problems that may arise during this time.



**Figure 1:** Microscopic cross-section of a frog blastula undergoing gastrulation. Image from iStock.

## Prenatal Development

Prenatal (also known as antenatal) development is the period of development from conception until birth. It is broken into three stages: the germinal stage, the embryonic stage, and the fetal stage [1]. The germinal stage describes the rapid and coordinated cell divisions occurring after fertilisation through cleavage, blastulation forming a blastocyst and surrounding trophoblast layer, and implantation onto the endometrial lining of the uterus. Trophoblast cells will give rise to the future placenta, and the

blastocyst will give rise to the embryo [2]. The embryonic stage starts from week three post-fertilisation and lasts around six weeks. During this stage, the embryo undergoes gastrulation (see Figure 1), allowing cells to form the three germ layers: endoderm (future nervous system), endoderm (future respiratory and digestive tracts), and mesoderm (future musculoskeletal systems), which further differentiate and migrate to form the distinguishable embryo structure and organs [1]. Finally, the fetal stage describes week nine until birth, where the placenta is fully formed and functional, providing the fetus with oxygen and nutrients from the mother's bloodstream and allowing the removal of waste products [1].

The Renaissance Man himself, Leonardo Da Vinci, made drawings of an embryo in the womb in his notebooks from around 1511 following his experience with dissections guided by the anatomist Marcantonio della Torre in 1506 [3]. His drawing contains meticulous and minute details of the placenta with associated blood vessels and the umbilical cord attached to a fetus in a breech position (Figure 2).

It stands to reason that clinically measuring and identifying these stages of development is conducive to ensuring the healthy development and growth of the embryo to recognise any defects or concerns. It is noted that from as early as the 1740s, the age of the embryo was determined through embryonic length by taking measurements such as head circumference and crown-rump length [4]. Since then, new methods such as embryonic age by determining the gestational age have been employed, which describes the time between conception and birth. However, the concept of embryonic staging was used to develop the Carnegie Collection, a collection of catalogued human embryos which provided the basis for the Carnegie stages, first categorised by Franklin P. Mall and further developed by his successors at the Carnegie Institution of Washington after whom the staging system is named. The stages are a numbered system from 1 to 23, classifying an embryo based on characteristics such as morphology, age, and size [5]. A recent 2023



**Figure 2:** The fetus in the womb sketch by Leonardo Da Vinci taken from his notebooks. Image obtained from Dreamstime. Da Vinci wrote alongside his drawing, “Breathing is not necessary because it is vivified and nourished by the life and food of the mother... And one and the same soul governs these two bodies, and desires, fears and pains are common to this creature as to all other animated parts” [3]. While our understanding has developed far beyond some of the primitive descriptions here, the understandings gained from simple observation alone are quite remarkable.

review by researchers and healthcare professionals in the Netherlands critically appraised publications utilising various staging values and concluded that Mark Anthony Hill’s 2007 paper on Carnegie staging [6] should be set as the gold standard of embryological staging [5].

This prenatal period of life is where congenital abnormalities can arise and occur in upwards of 2% of pregnancies [7]. These abnormalities can be caused by genetic factors such as single-nucleotide variants, aneuploidy, and copy number variants [8] and include abnormalities such as congenital diaphragmatic hernia, neural tube defects, and hypoplastic left heart syndrome [9]. They can also be caused by environmental exposure to teratogens, compounds that cause birth abnormalities and are encountered by the fetus in utero following exposure to the mother in food, water, through the skin or inhalation [10]. A classic example of a teratogen causing birth abnormalities was the use of thalidomide in the late 1950s for pregnant women to treat morning sickness. Used across 46 countries, thalidomide was observed to contribute to the development of congenital abnormalities, including damage to fetal organs such as the

eyes, ears, heart, and the birth of babies with severely deformed limbs [11].

## Postnatal Development

Postnatal development describes the time of life after birth and the transition to the extrauterine environment. During this transition, there are changes in the circulatory system, ventilation, and metabolism. Ventilation through the lungs begins in the baby’s first breath, the heart’s right-to-left chamber shunts such as the foramen ovale closes, and metabolism ramps up, leading to increased glucose requirements [12]. The period of postnatal development is characterised by the stages of infancy, childhood, and puberty, where growth is largely dependent on nutrition and growth hormones [13].

The delicate interplay between the mother and her neonate in utero, as well as postpartum through breastfeeding, provides the neonate with the immunity and microbiome to face the complex and pathogen-rich outside world. Breastmilk, particularly colostrum, the thick milk produced by the mother in the first seven days after giving birth, is rich in nutrients, sugars, and fats but also contains stem cells, epithelial cells, and immune cells such as T cells and antibody-producing B cells [14]. These indispensable immune factors provide the baby with an immediate first line of defence, which is skewed towards the pathogens and microbes the mother has faced and, therefore, that the baby will encounter in its first few years.

## The Developmental Origins of Health and Disease (DOHaD) Hypothesis

I sat down with a few academics at the Liggins Institute to hear their insight about the DOHaD theory and discuss their research related to early life. The Liggins Institute is located at the Faculty of Medical and Health Sciences and is an internationally recognised research centre focused on the role of early life and the nutritional environment in non-communicable diseases in the context of genetics, translational science, and the determinants of health.

I had a chat with Professor Mark Vickers, whose work at the Liggins Institute is focused on DOHaD and the association between poor maternal nutrition and the development of obesity and type 2 diabetes. He also works on reversal therapies for suboptimal early life environments that cause metabolic disturbances in neonates to assess later disease risk. Below is what he had to say during

our interview.

**How would you simply explain the DOHaD theory with all of your experience in the field?**

“Basically, DOHaD means that a poor early life environment can predispose you to an increased risk of developing a range of diseases in adult life, which could be obesity, diabetes, all the common diseases we see today. A good example of that is poor maternal nutrition, so if a mother has a poor diet during pregnancy and early infancy during the period of lactation, that can predispose her infant to a range of illnesses in later life. The most common term for that is early life developmental programming. There is also a lot of evidence now for the health of the father impacting the health of his children as well. So, for example, if a father is obese at the time of conception, he can programme health outcomes in his child. What’s also important is that a lot of these programming effects cross generations, so a single adverse pregnancy, for example, not only affects the mother and her child, but also the next generation after that as well, meaning two generations of impacts. So you can see how a cycle of diseases such as obesity and diabetes keeps being transmitted across generations. There is enormous evidence around the programming of disease due to a poor early life environment.

Another common term is ‘The First Thousand Days’, which is pregnancy and the first two years of life. This sets the foundation for the child’s entire life course, and once an individual is programmed to have an increased risk of disease, if they then have a poor lifestyle after birth (in terms of poor diet, lack of exercise), that further exaggerates the disease risk across their life course. And a lot of it makes intuitive sense; for example, if a mother is undernourished during pregnancy, that can affect the development of the kidney or pancreas, and once the child is born and develops, they may have issues related to kidney function, high blood pressure, insulin resistance, and so forth.

The trouble is that a lot of these diseases, once they are manifest during

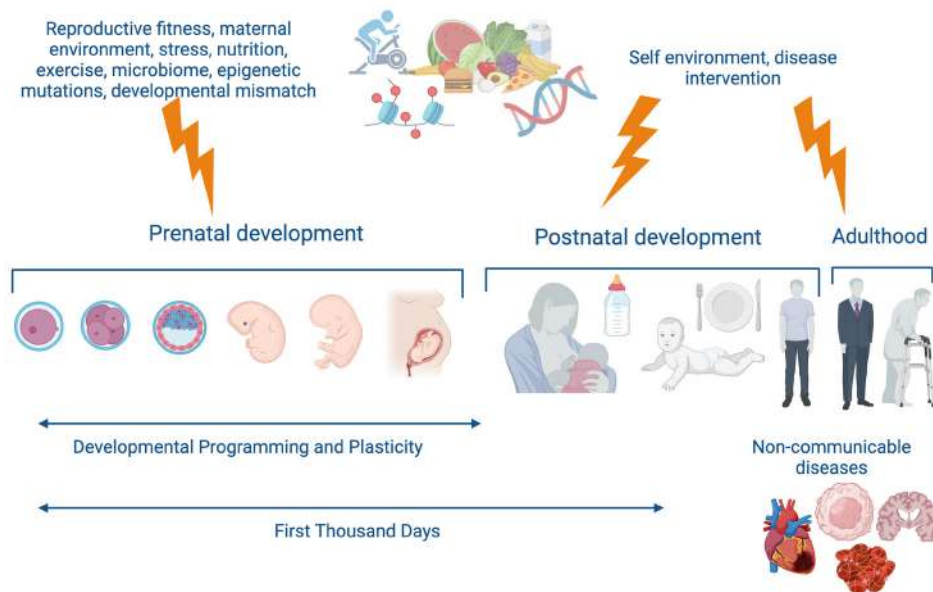
adulthood, do not have a lot of treatment strategies that work effectively, so the best place to target is right early at the origins of where those diseases are occurring. The trouble with the DOHaD work is that despite it being around for several decades now, the uptake and communication of it as a concept is still quite poor.

We have done surveys recently looking at adolescents because they are the next generation of parents, so we want them to know about DOHaD to optimise their pregnancy and the health of their child across their life course. But the understanding of DOHaD as a concept and the ‘the first thousand days’ is remarkably poor, so we need to do a far better job at communicating those ideas moving forward.

There are ways of intervening and preventing, such as through maternal nutritional supplements. The Liggins has done some work looking at fish oil supplements for mothers, for example. So there are some ways to prevent some of this programming happening, but it has not translated to practice yet.”

**Following the process of educating people and improving understanding, what are the actionable steps you would like to see?**

“A lot of it almost comes down to common sense. My area of expertise is around nutrition, and it comes back to a balanced diet. We know that at both ends of the nutritional spectrum,



**Figure 3:** Overview of early life development and its effects on non-communicable disease risk in later life. Created with BioRender.

if a mother is underweight or overweight during pregnancy, that will lead to poor outcomes for the offspring, which are made worse if that child is then exposed to high-fat diets or high-sugar diets. We also know that this developmental programming also programmes appetite control, so it can make you prefer high-fat foods over other foods and so forth. So it's about this awareness of the impact of the early life environment that can impact you poorly for the rest of the life course of your child and the next generation."

### **Molecular Mechanisms Underpinning DOHaD**

The DOHaD hypothesis describes the concept that early life and the environment during this period influence the lifelong risk of non-communicable diseases (NCDs). It states that certain developmental pathways in utero as well as after birth impact both the susceptibility and resilience of individuals to future health challenges [15]. The DOHaD hypothesis was first introduced by David Barker in 1986 through his paper, which connected the rise in heart disease with neonatal mortality in England and Wales, in which he postulated this was caused by poor nutrition in early life [16]. Scientists have since run with it, taking it from a theory that covered meiosis and gametogenesis in the fetal period to the whole of prenatal and postnatal development, including maturation into adolescence [17].

There are certain molecular mechanisms that underpin the DOHaD framework, including epigenetics. Epigenetic modifications are heritable changes in gene expression through mechanisms that do not directly alter the DNA sequence and include methylations of DNA or modification of histones and the chromatin structure. These work by changing the accessibility of the genome by altering its architecture and, therefore, its ability to be read and expressed [18].

A current and rising body of literature presents epigenetic modifications as a mechanism behind the programming and changes involved in DOHaD [19-22]. For example, obesity and its consequent NCDs, such as type 2 diabetes and cardiovascular disease, are rising concerns in modern societies. While obesity is a multifactorial and complex condition, recent efforts in understanding the gene-environment interactions involved have revealed the role of epigenetics in the development of obesity and its associated diseases. Through epigenome-wide association studies (EWAS), hyperglycaemia, type 2 diabetes, and obesity have been associated with certain DNA methylation differences at loci involved in metabolism and inflammation, including ABCG1, PHOSPHO1, SOCS3, SREBF1, and TXNIP [23-26]. These events involved in the pre-conception period of the mother and father, as well as the in-utero environment, have successfully predicted the future risk of acquiring type 2 diabetes [23]. To clarify, the DOHaD paradigm does not set out to predict that an individual will acquire a disease based on certain epigenetic barcodes but rather describes the altered risk and susceptibility of encountering altered metabolic or molecular pathways that may give rise to a certain NCDs based on quantifiable and perceptible changes in early life.

Epigenetic modifications are also a way by which the fetus can undergo 'developmental plasticity', a process where the fetus adjusts its phenotype to the environment using environmental cues such as nutrient availability, toxic exposures, stress, and microbiome composition [15, 20]. This process allows the fetus to adapt and survive in a changing environment from a single genotype but also avoids the necessity of genetic evolution in survival, conferring an evolutionary advantage [20]. It then stands to reason that the environment in early life can indeed play a role in altered

phenotypic expression and, therefore, health outcomes.

This process, while helpful in early life, can also conversely lead to increased disease risk in future life, substantiated by the theory of 'developmental mismatch'. This describes a future environment different to the one predicted by the environmental cues that were present in utero and early life, where the adaptations now serve as a disadvantage. Examples where different environmental conditions in early life or later life can cause mismatch include an unbalanced diet and body composition of the mother, increased energy-rich foods, and reduced physical activity of the neonate in later life [27].

Epigenetic modifications owing to the environment can also include those due to the interactions and interplay between the fetus and the placenta. The placenta is a short-lived organ in the maternal uterine tissue attached to the fetus through the umbilical cord [28]. Prenatal exposures such as maternal smoking, phthalate plastic chemicals, persistent organic pollutants, and psychosocial stressors can lead to changes in placenta epigenetics [28]. A 2019 study with a cohort of over 400 mothers and their neonates found that maternal smoking was associated with over 71 different methylation sites, seven of which influenced infant birth weight [29]. Other studies found that these methylation events were involved in preterm birth, birth length, birth weight, and head circumference [30]. Exposure to air pollution through particles less than 2.5µm in diameter is negatively associated with organ and vasculature development and positively associated with atopic dermatitis [28].

### **The Prenatal Corticosteroid Story in Preterm Babies**

Sir Graham Liggins, after whom the Liggins Institute is named, was a Kiwi obstetrician. While working with pregnant sheep, he made the observation that infusing the sheep with corticosteroids allowed the survival of prematurely born babies. Between 1969 and 1972, he worked on a clinical trial in humans with Ross Howie and together demonstrated that administration of prenatal corticosteroids to pregnant mothers with indications of giving birth prematurely crosses the placenta and leads to a 50% reduction in respiratory complications such as respiratory distress syndrome caused by premature birth. Before this discovery, the prognosis for preterm babies born before 32 weeks of gestational age was bleak, with the

majority dying of uninflated lungs [31]. Their results were published in 1972 in "A Controlled Trial of Antepartum Glucocorticoid Treatment for Prevention of the Respiratory Distress Syndrome in Premature Infants" and have since set the standard of care for preterm babies worldwide [32]. Their work was instrumental in saving the lives of thousands of babies and paved the way for a drastic improvement in preterm baby prognoses, now being used routinely by doctors around the world.

I had a chat with Professor Justin O'Sullivan, the Director of the Liggins Institute, whose work is based on genomic research in the Newborn Genomics programme to drive precision medicine, diagnosis, and treatment of NCDs through methods like molecular biology and bioinformatics. His group also works on the human microbiome and its relationship with human disease and gene regulation. We spoke about the work going on at the Liggins Institute and how this has been translated into the healthcare system.

### **What does current ongoing work at the Liggins look like?**

"If you work your way across the Liggins, we have people who have specialties in everything from maternal-fetal specialists to neonatal specialists; we have endocrinologists, nutritionists, and basic scientists. So, there is a real spread of expertise across the whole institute. It is not really about individual expertise that we all bring in; the key things the institute does can be thought of in terms of preclinical research, clinical research, and then translation science.

So, as examples, we have clinical trials that are being done in gut bugs, where we are doing fetal microbiome transfers with young children and studying the impact of that. With translation science, we get lots of guidelines in terms of taking findings that we make to the public and using them in the healthcare system directly. We also have the trio sequencing programme, which we are setting up in newborn genomics, looking at genetic sequencing of chronically or critically ill children and their parents in the neonatal and pediatric intensive care, so that is a translational project where we bring science across to the clinical. And then there are other studies like animal studies investigating cancer, and we also have a farm where large studies are done on sheep.

Other things that researchers find and are finding at the Liggins are very important. This includes the dextrose gel studies that Jane Harding and her team did, where babies are born with hypoglycemia because they have been dependent on their mother for so long, and the switchover to be dependent on themselves is very important. Some babies take a while for their glucose metabolism to kick in properly because they have to switch on all the right genes, and while they do that, they become hypoglycemic, which can be associated with damage. But the way to fix it is simple: you just take a bit of this dextrose gel, pop it on a finger, and stick it inside their mouths so that the baby can absorb the sugar and will no longer be hypoglycemic. A simple solution for a common problem. This was part of the Sugar Babies study in 2010 and it is one of those studies that went from a research environment in a clinical trial and has been translated right through to clinical practice."

### **Concluding Remarks**

To summarise, scientific knowledge is continuously growing to support the importance of not just prenatal development but also the environment and maternal and paternal health leading up to conception. There are roles for nutrition, genetics, and the cellular microenvironment in setting up the risk of NCDs. While substantial evidence exists for the validity and importance of DOHaD, such as ongoing work at the Liggins Institute, uptake and public understanding are limited and need to be targeted to allow translation of this knowledge to communities so that we can start to see improved health outcomes.

## **Acknowledgments**

I would like to thank Professor Mark Vickers and Professor Justin O'Sullivan for their time in the interviews and for helping me write and explain such overwhelming topics of DOHaD and the corticosteroid story, as well as the research work at the Liggins Institute.



### **Noor Radhi - BSc, Biomedical Science**

Noor is a third-year biomedical science student specialising in immunology and cancer biology & therapeutics. She is particularly interested in cancer pharmacology and exploring the driving factors of autoimmune diseases. She is the current head of outreach for Scientific and vice-president of SAMS.



Interview

# Calculators, Computers, and ChatGPT: How are University Students Using Generative AI Chatbots in Education?

Taylor Lee

Artificial Intelligence

What are we able to do with a computer? With a calculator? With an artificial intelligence (AI) chatbot? Certainly, they are all powerful tools based on human input and interaction. These technological mediums change the nature of our interactions and abilities, with educators and students usually being the earliest robust adopters of new technologies. Chatbots, at their core, are interactive **agents** that respond to user interaction in a conversational way [1].

## Early Chatbots in Education

From 1964 to 1966 at MIT, Joseph Weizenbaum created ELIZA, the very first **natural language processing** chatbot and artificial psychotherapist [2]. Since sparking a new communication between humans and machines, other programs in AI-based education, such as SCHOLAR in 1970 [3] and OpenAI's ChatGPT in 2021 [4], have followed.

ChatGPT uses large text datasets from the internet, learning patterns and relationships in the text data to predict and provide the best answer in conversation. It uses reinforcement learning from human feedback (RLHF) for enhanced capabilities. ChatGPT has a multifaceted usage as AI has demonstrated the potential to help learners identify knowledge gaps through conversations and time-efficient tasks. It shapes current technologies and their capabilities much more than in the past.

Computer scientist Alan Turing debated whether machines could think, famously seeking to distinguish between machine and human [5]. This debate remains relevant as conversations surface around changes in education. In 2001, Eugene Goostman—a chatbot portrayed as a 13-year-old Ukrainian boy with grammatical errors and general knowledge—was created, proving that in some regards it could pass the Turing Test [6]. In 2023, ChatGPT has the power to break the Turing test [7]. However, it has been argued that a new test should be created to measure AI intelligence.

All students interviewed have consented and remain anonymised.

## Current Studies

Current studies exploring AI usage in education have focused on how computing instructors plan to adapt or prevent ChatGPT and other AI coding assistants by designing “AI proof” or encouraging usage on assignments, and any challenges and opportunities the specific tools present [8]. However, it has not accounted for diverse academic backgrounds nor focused on students.

## Chatbots in Education Today

While the introduction of AI chatbots and **generative AI** has been growing, so too is this a revolutionary time for existing standards in education. At universities like the University of Auckland, there have been Canvas banners, resources for teaching, and emails about acknowledging, determining, and

understanding the advancements of generative AI, often warning students about permitted uses and cheating [9-10].

To understand student usage and perspectives about AI chatbots in education, I conducted five diverse academic interviews with individuals from humanities to STEM backgrounds at the University of Auckland. All participants were interviewed separately. As of the published date, there are little to no empirical studies about how university students are using AI tools such as ChatGPT. Here is a summary of my findings so far.

ID	Study	Degree	Months	Usage frequency per week	Main AI tool usage	Prior experience with AI tools
P1	Mathematics, Computer Science	Undergraduate	14	7 days	ChatGPT 4, Bing AI, Google Bard	experience creating and using AI tools, assignments and lectures
P2	Civil Engineering	Masters	9	7 days	ChatGPT 3.5, Bing AI	experience using AI tools for language and idea enhancement, assignments
P3	Engineering Science	PhD	12	5 days	ChatGPT 3.5, Copilot	experience using AI tools for code generation, research
P4	Business, Law	Undergraduate	6	2 days	ChatGPT 3.5	experience using AI tools for idea generation, for their one quiz, assignments
P5	Education, English	Undergraduate	1	2 days	ChatGPT 3.5	experience using AI tools for language and idea enhancement, for their one quiz, assignments

*Note.* Months indicate how long they have been using AI generative tools in their education on average.

**Table 1:** Participant’s profile using AI generative tools in an educational context. Months indicates how long they have been using AI generative tools in their education on average.

All students interviewed use ChatGPT, two use Bing AI and two use code-generation tools like GitHub Copilot. They use these when they have an assignment and in their workflow about four times a day. One student interviewed has paid for the premium version of ChatGPT due to its customisability and advanced models. To cover the costs, they split it among friends.

The main motivation for using ChatGPT compared to other tools, according to the students, is because it encompasses all aspects of generation, rather than one. While it is convenient to use other chatbots when built into a browser or product in such tools as Bing AI and Copilot.

## Bridge the Gap - Ideate and Extend with AI

All students interviewed said they do not use AI tools to do their assignments for them. Instead, they said they use it to ask questions about existing materials such as gaps in lectures or knowledge, as an idea or code generator, and to add more details to their work.

P3, an Engineering Science student, codes and comments with a basic framework. “I know the algorithm and the logic. I write comments about what the program needs to do. Then the code generator fills in the rest and

I fix whatever area I need to.”

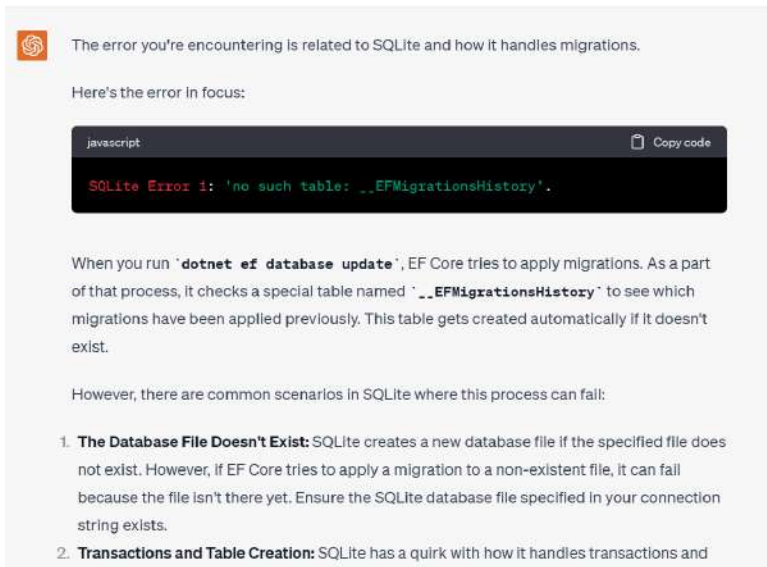
P4, a Business and Law student, says they have ideas but struggle to put them into words. “I put everything down and ask it to make sense. Like when I am coming up with arguments when I ask it to show it from a financial perspective.”

P2, a Civil Engineering student, also uses it for the same purposes. “I feel like AI is quite different from humans. It sounds too robotic and formal. I usually don’t copy everything, I only copy things I usually like.”

P5, an Education and English student, uses ChatGPT. “When I am researching or reading something too long, if I really can’t figure it out and don’t have a lot of time, I use ChatGPT. It gets the key details and polishes my words. I think it’s important to use your brain rather than always rely on a tool.”

```
at Microsoft.EntityFrameworkCore.Design.Operations.UpdateDatabase.<>c__DisplayClass10.  
at Microsoft.EntityFrameworkCore.Design.Operations.OperationBase.Execute(Action) [Active]  
SQLite Error 1: 'no such table: __EFMigrationsHistory'.
```

Why



The error you're encountering is related to SQLite and how it handles migrations.

Here's the error in focus:

```
javascript Copy code  
SQLite Error 1: 'no such table: __EFMigrationsHistory'.
```

When you run `dotnet ef database update`, EF Core tries to apply migrations. As a part of that process, it checks a special table named `__EFMigrationsHistory` to see which migrations have been applied previously. This table gets created automatically if it doesn't exist.

However, there are common scenarios in SQLite where this process can fail:

- The Database File Doesn't Exist:** SQLite creates a new database file if the specified file does not exist. However, if EF Core tries to apply a migration to a non-existent file, it can fail because the file isn't there yet. Ensure the SQLite database file specified in your connection string exists.
- Transactions and Table Creation:** SQLite has a quirk with how it handles transactions and

**Figure 1:** A coding question and suggestions related to database query using ChatGPT 4

## AI as a 24/7 Tutor and Coworker

AI guidance is often personalised and tailored to the situation, pulling from multiple sources at once. When asked how they would describe their relationship with AI, the P2 and P5 said they see AI as a tutor, while P3 and P4 said it's like a teammate or coworker. P1 said it is both like a coworker and a tutor.

P1, a Mathematics and Computer Science student, also said it helps 24/7, being an always accessible tool. For general uses and knowledge, they said it is better to use ChatGPT, as illustrated in [Fig. 1]. However, if it is something specific to the course, they ask for help, but rarely so. "It's not realistic to ask a real human for help 24/7. It's better to post on Piazza but that doesn't mean the answer is going to appear in an instant." The P4 and P5 agree about general and specific knowledge usage.

P4 further stated they find their relationship with AI as a launching point for inspiration. "I ask it to generate some ideas based on what I am supposed to be researching. For example, in my Law paper, I need ideas for a different perspective for an argument. I say 'What would be another argument for this topic. This is for my Law essay.' However, if it's something more specific that only the other people who took the paper understand, that's when I ask my classmates[...]"

P1 also found it to be more cost effective and would rather spend \$20 per month than \$20 per hour using a tutor. "It teaches me what I need to know and also corrects me at the same time." As a tutor themselves, P3, an Engineering Science student says they see AI as a coworker. "We are writing code together. I am writing comments, I am fixing what it is writing, and I am asking questions about the code."

## Improving Time Efficiency and Language Acquisition

All students said it has saved them time, since it is similar to a **search**

**generation summary tool** or provides them ideas they would not have otherwise considered.

P3 said, "I am very pro AI. I use it as a tool to write code efficiently and quickly. It speeds up the process and saves me time from searching through the internet since it gathers everything in one place. I don't necessarily use it to generate something from scratch. Instead, I get it to fix whatever code I have already written."

When it comes to last minute assignments, P2 found it helpful to have ChatGPT. "I had a 4,000 word essay due in two days. I asked questions like, 'How would you reword what I wrote?' and more information about the Treaty of Waitangi, as illustrated in [Fig. 2]. I was luckily able to submit it on time."

P2 and P5 said their language acquisition skills have improved since ChatGPT imitates having a conversation as a human might.

P5 remarked, "My sentence structure is not very good, so I use it to ask what certain words mean when I am reading an assignment or to restructure and correct my sentences in a certain way. I feel more confident and can understand things I learn from the class." P2

## Glossary

**Agent:** a computer program or machine that is designed to perceive and interact with its environment, make decisions, and take reactive or proactive actions based on specific goals.

**Natural language processing:** teaching a machine how to put words and sentences together with grammar.

**Generative artificial intelligence:** a machine that learns from data to generate its own content, based on what it has learned.

**Search generation summary tool :** a tool that summarises search engine results and other training data, rather than acting as another search engine

**Hallucination:** also known as confabulation or artificial hallucination is when there is a confident response by an AI that provides false statements.

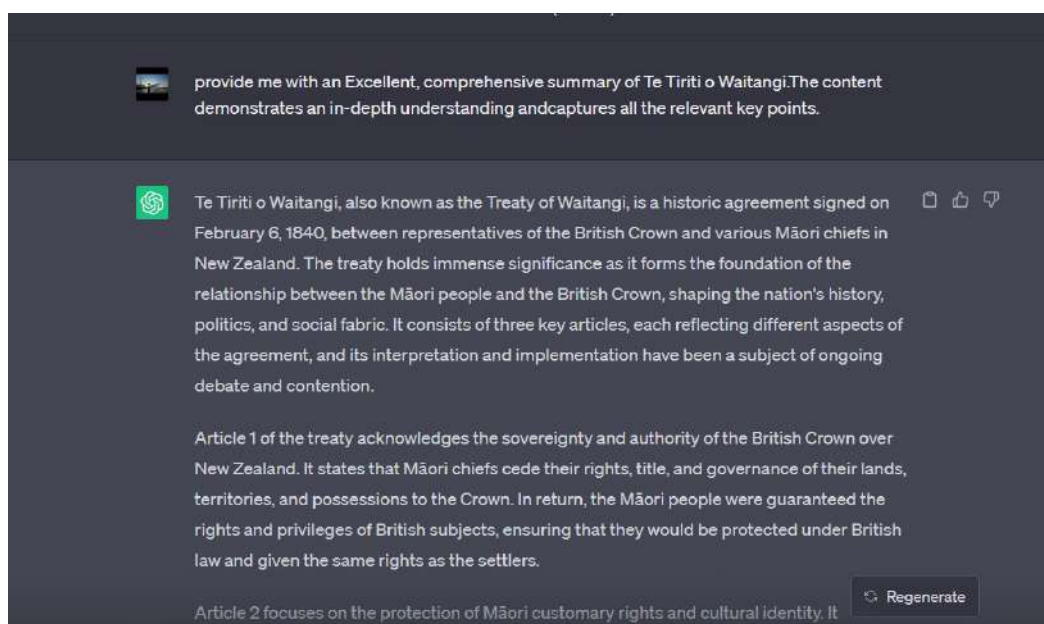


Figure 2: Summarise key findings about Te Tiriti o Waitangi

had similar experiences saying, “I think that it has improved my language skills, especially since English is not my first language. I know how to write properly even though the information may be fake or too formal.”

### Relying on Critical Thinking Rather Than AI

Two students out of the five interviewed said they have used ChatGPT while taking a quiz on at least one occasion. Two students have used it to generate code for them. Despite using it, they said the information was not always accurate or reliable and that they scored very low. As psychologist Abraham Maslow said, “If the only tool you have is a hammer, it is tempting to treat everything as if it were a nail” [11].

Some limitations the students noticed were that they **hallucinate** with fake information, like in text-to-text generation where the program generates text based on a given prompt, but it doesn’t always get it right., AI often corrects itself only after a human points out the error and explains why it’s wrong. Not mentioned by the students, but another noted limitation is that AI will be skewed towards Westernised English and Spanish-speaking audiences in the ChatGPT 3.5 model, performing low-quality output and more errors.

In their Business paper, P4 said, “I heard that other classmates were planning to use ChatGPT during a quiz, and it felt like they had an advantage. I obviously felt bad about using it but I didn’t want to miss out. It’s more advanced for general knowledge. However, for anything specifically referenced in lectures, it’s more challenging for AI to pick up on that.”

In order to check ChatGPT’s accuracy, students have compared it with lecture slides, provided readings, and Google searches. P4 said they also

***“It’s not really the AI; it’s the human using it. Everything has advantages and disadvantages. It’s about how we will regulate and instruct students to use these tools. We need to learn how to use them.”***

– Maths and Computer Science Undergraduate Student

talk to other classmates and Piazza to see if they are on the right track. “Some of my classmates go straight off ChatGPT, but I think it’s smarter to fact-check and look into my notes.”

P3 expressed their concerns about the overreliance on ChatGPT, especially in introductory-level papers. In an introductory Engineering class they teach, “about 200 out of 1000 were caught cheating because they all had the same coding style. We compared coursework and tests to see if anything was different. Of course, tests are time-based, so there may be more pressure.”

P3 also emphasised that in introductory coding papers, the purpose is to learn to code. “I think it’s especially tough when teaching students who don’t know how to code. They aren’t learning by copying code. It might be different if they understand the algorithm and data structure, but I think that’s a whole different discussion.”

### Regulating AI Usage in Classrooms

All participants said there is a need to regulate AI usage and that there is no point in restricting its use since no one will stop using it. This dilemma between human and AI regulation and capabilities can be best described as a mind-body problem. Ordinary matter (“body”) is inherently passive, simply responding to outside forces. However, human beings (and other living things) are not passive; they have autonomy and action that machines don’t necessarily have [12].

P1 said, “We need to accept and regulate it. It’s not really AI. It’s the human using it. Everything has advantages and disadvantages. It’s about

how we will regulate and instruct students to use these tools. We need to learn how to use them." P1 also said the university can integrate how to use AI responsibly into the required Academic Integrity paper.

P3 said the current way the university is regulating AI is not working. "The lecturers say they are using AI detection but it doesn't actually work. It's very far from being accurate. We have to manually check the code or see our analytics, and compare students' work to one another."

P1, P2, and P3 said there should be an age restriction for AI usage. P1 proposed, "Only people who have self-control and education about how to use AI-powered tools should be able to use them. Maybe educate younger students early on."

However, it could be argued that banning it in classrooms does not mean students will not have access outside. Therefore, incorporating ChatGPT in class could prevent its misuse. P3 said one of their papers asked students to write a prompt to the code screenshot, to educate students about prompt generation.

As of now, there are no AI-specific laws in New Zealand. However, the Fair Trading and Harmful Digital Communications Acts and obligations under the Tiriti o Waitangi [13] still apply. They all take a neutral and principle-based approach, meaning the same rights and protections apply to AI tool usage. Challenges and opportunities for development remain, which could be implemented even outside the educational system.

### Educators as Teachers, Moderators, and Moral Agents

Our emphasis on humans should be deliberate as communication extends further than a signal to send and receive a message. Students agreed that the richness of a person cannot be replaced by a robot. As technology continues to advance, AI usage should not be discouraged but rather embraced responsibly.

P4 shared how their lecturer encouraged students to use AI as a tool. "In my Information Systems paper, literally in the first lecture, my lecturer was like, I encourage you to use AI instead of avoiding the entire conversation altogether. Yes, I know you can use it to cheat, but that isn't going to help you out in the long run. You need to learn how to use it to your advantage, like as a study tool, a tool for research. It's like your second Google."

P3 who assists in one of the introductory papers said that students have told them they feel more comfortable talking to an AI than a lecturer. "I know some lecturers who have been trying to change that dynamic. They use AI as a starting point and then build off what ChatGPT wrote. Being caught using ChatGPT should not be something that's necessarily feared, depending on how they used it."

Mathematician and educator Seymour Papert's vision encourages computer-aided instruction - making the computer teach the child; the

***"There's something different about human emotional connection. Teachers are not there to just assign work. They are there to help students grow a love for learning and make them excited."***

**– Education and English Undergraduate Student**

students interviewed similarly agreed. "The child programs the computer and, in doing so, both acquires a sense of mastery over a piece of the most modern and powerful technology and establishes an intimate contact with some of the deepest ideas from science, from mathematics and from the art of intellectual model building" [14].

P1 said AI has been helpful to fill in any gaps and that educators could also make sure students are going on the right track. "Oftentimes, I have to ask very specific questions about what I am asking for. For example, my COMPSCI 230 lecturer was hard to understand. However, asking AI to explain concepts in a more straightforward way was helpful. Then I confirm it to see if I am on the right track through the teaching assistants later."

Similar to theoretical physicist Richard Feynman's teaching philosophy [15], P5 said teachers will never be replaced. "There's something different about human emotional connection. Teachers are not there to just assign work. They are there to help students grow a love for learning and make them excited."

### Conclusion

These findings capture a snapshot of 2023, as educators start to discuss best practices and literacy around AI usage. Using these findings can lead to more conversations about how we can work together in effective, equitable, and ethical ways, and how we can shape the future of our education. The question remains: How will we use AI to unlock further positive potential?



#### **Taylor Lee - BSc, Computer Science, Psychology, Sustainability**

Taylor is an undergraduate student studying computer science, psychology, and sustainability under the Science Scholars programme. She is currently interested in human behaviour related to artificial intelligence and sustainable digitisation.

# Simulating the Mechanisms of Life Through the Lens of the Virtual World

Nadia Snegirev

Computational Biophysics

**A new competitor has entered the race for understanding biological systems at the molecular level. A little later to the game compared to other more established techniques, but with the power of big data and high throughput, it might just revolutionise the game as we know it — think enzyme design from scratch, reproducing the cellular environment, and determining structure and function from just a sequence. Computational biophysics — where has it come from, what does it look like, how can we trust our results, and where is it heading?**

## Setting the Scene for Computational Biophysics

Reducing the real world into systems which can be modified, changed, and precisely controlled is a common approach in science for developing theories and understanding about real-world systems [1]. This kind of reductionist approach is common throughout science, particularly in biology, where we try to reduce a system down to the key elements needed to reproduce its behaviour [2]. For any experiment or model, the level of detail you go into, the size and complexity of the system, and what you include or leave out is an important part of the design [3-5].

We acknowledge through the discussion of *in vivo*, *in vitro*, and now *in silico* models that we are working with abstractions of reality. However, what makes *in silico* different is that we work directly with a mathematical description of the system [6]. In the case of computational biophysics, the simulations predate the advent of computers, with the simulation of a simple reaction occurring by hand back in 1936 [7]. The development of computation has simply allowed the required calculations to be done at both speed and scale.

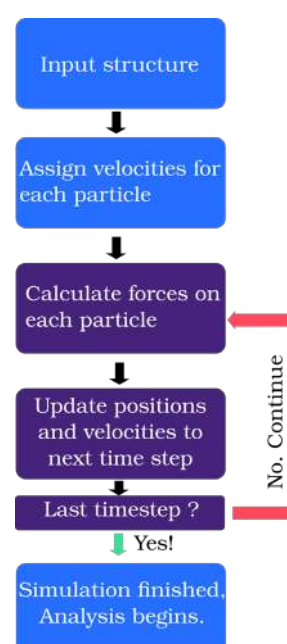
This brings me to the field of Computational Biophysics, which requires a bit of breaking down. The “bio” in biophysics indicates that our focus is biological systems, studied using physics-based theory. The computational part comes in since the systems of equations formed are distributed across the computer’s hardware and then solved. Then we use mathematical and statistical methods to analyse the data generated, typically done with Python in our research group.

Firstly, we input a starting structure of the protein to be studied, obtained via modelling or experiments like X-ray crystallography, NMR, or Cryo-EM. When experimental structures for proteins are not yet resolved, the structures can now be obtained through AI, using AlphaFold [8] or similar tools [9]. Each particle in the experimental or modelled structures is then assigned velocities, based on the temperature at which the system is simulated. The forces between the atoms in the system are then determined using a set of parameters defined for each atom pair. These parameters are sourced from a force field, with different force fields performing better for different types of systems [10]. Some examples of these include AMBER [11], CHARMM [12], UFF [13] and OPLS [14]. Now that we have forces, masses, positions, and velocities, we can set up differential equations for motion. Integrating over time allows us to move forward in time, updating the velocities and positions at each step and repeating the process. This is done via a molecular dynamics (MD) engine, which can run via the command line or in some cases with a graphical user interface. Some of the more popular ones include AMBER, CHARMM, GROMACS, LAMMPS, and NAMD. Many of these are free and open source, which assists with accessibility.

This all involves a careful balance of accuracy, size, and complexity of simulation length, alongside matching the restraints of the available computational resources [15]. The simulation stops when we have enough data for a particular process we want to observe. We say that we have ‘sampled enough’ of the space that our molecule would explore, and that our results roughly would not change if we ran the simulation repeatedly from slightly different starting conditions (e.g. different velocities) [15-16].

From here, the analysis begins. Focusing on the general behaviour of the system, we do not typically focus on any moment in time, but rather the set of conformations the molecule explores through the simulation [17]. This captures if there is enough sampling for your system and the amount of error associated with this.

Following this, a good place to start is at the atomic level, investigating the types of



**Figure 1:** Flowchart describing steps of running a molecular dynamics simulation. Figure by Nadia Snegirev.

interactions, distances, angles, and dihedrals of atoms or groups of atoms. A combination of what you observe and what has previously been studied will give you an indication of what to focus on, and where there might be interesting trends. Then stepping back to the bulk observables, like the proportion of secondary structure present, end-to-end distances measuring compaction of protein, radius of gyration and root mean square fluctuations (RMSF) – both measuring the degree of dynamic behaviour/change in positions over time but with a slightly different method. This is by no means a comprehensive list of the possible types of analysis, but a start for basic analysis; the type also changes for the system of study and the process of interest. These can be investigated via the tools of the MD engine of your choice [18], through other software packages [19], or purpose-built scripts.

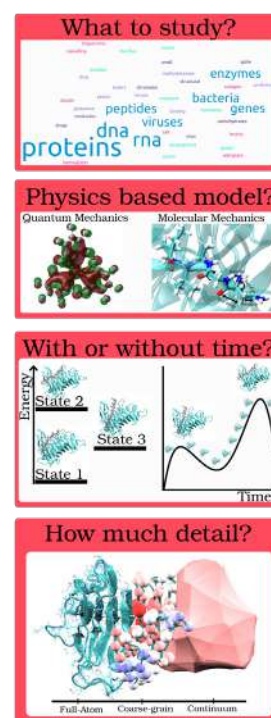
### The 'Why' and 'How' of Computational Biophysics Today

There has been expansive development in the field of computational science due to increases in computational power and the development of frameworks, which have made running MD simulations much more accessible. This has opened up the field and allowed a broader range of hypotheses to be tested using computational simulations, with longer simulations and studying larger systems now possible.

We cannot talk about computational performance without first talking about the role of Graphical Processing Unit (GPU) computing and how these increases have revolutionised the field.

GPU-accelerated computing started to be used in computational chemistry in 2007, when NVIDIA released Compute Unified Device Architecture (CUDA) [21]. The principle of GPU computing is that larger jobs can be split into smaller tasks and run in parallel, since the GPU contains hundreds of cores (CPUs have much fewer cores) [22]. It's like giving a worksheet out in a tutorial. The worksheet is completed much more quickly if you assign each student a particular question to complete in parallel with their classmates. This works well for scientific computing (like molecular simulation), which is data-intensive. As a result, there have been hundred-fold reductions in computational cost for molecular simulations, allowing larger system simulations and longer timescales [21]; performance has been observed to accelerate by 10-100 times [23].

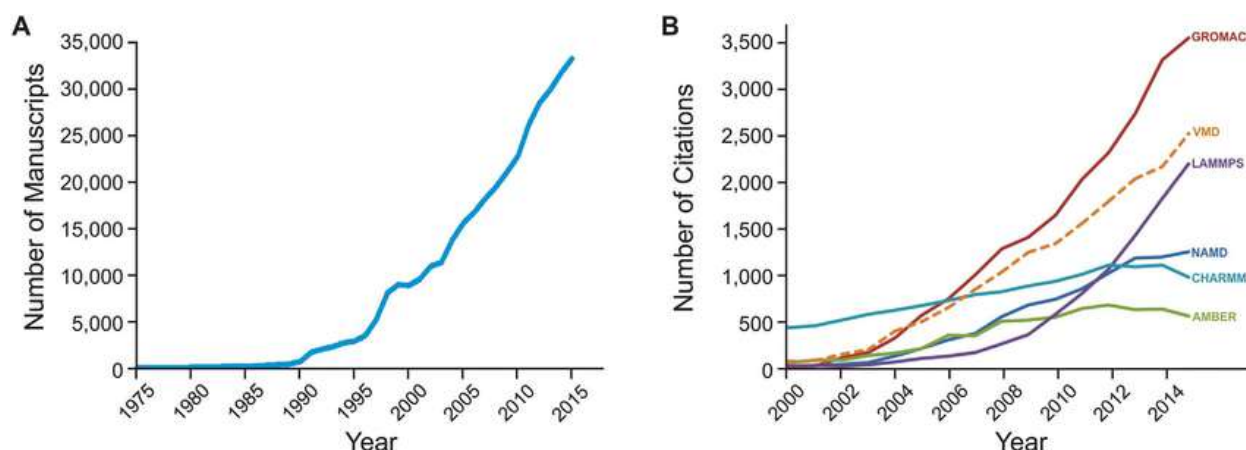
A quick introduction to the variety of infrastructure available nowadays is seen in Figure 4.



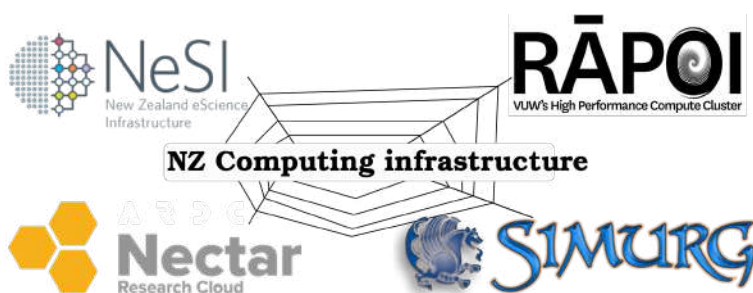
**Figure 2:** Schematic explaining the different choices involved in simulating biological systems and the considerations a researcher might need to make. Figure by Nadia Snegirev.

High performance computing (HPC) consists of a series of identical machines that are connected locally with a high-speed network. These are typically the least accessible, potentially requiring institutional subscriptions. Of note is the Anton supercomputer, a purpose-built molecular dynamics infrastructure, which can now simulate 1  $\mu$ s/day for a system of a million atoms [24].

On the other hand, Grid computing is made up of many different devices using software to communicate results between them. An example of this is GPUGRID, where people can



**Figure 3:** Graphs showing the increase in articles being published which incorporate MD simulations, and the growth in use of MD engines [20].



**Figure 4:** Schematic summarising some of the infrastructure available to New Zealand researchers, access depending on university affiliation and subscription. Figure by Nadia Snegirev.

volunteer GPU time on their own personal device when they are not using it [25]. This network, in its second year of operation, has become one of the most computationally powerful computing projects, allowing many simulations to run and helping to further understand protein folding.

Cloud computing is a relatively new development, operating on a pay-per-use model and utilising virtual machines [26]. You typically connect to this via the internet. Many features of cloud computing are appealing and can increase the accessibility of running molecular simulations, including allowing resources to be used more effectively, reducing idling time of resources, overcoming the expenses related to setting up a local computing cluster, and easy scaling of computational jobs and tasks. The administration and maintenance of the network falls on the vendor, as opposed to the user, which is the case for a locally run network.

Simulating life is so much more feasible now than when the first simulation of a protein was just 8.8 picoseconds in length [27]. It is also much more accessible with the movement towards open-source software for both running simulations and doing analysis; new types of infrastructure have also reduced the barrier for entry [28].

### How Can You Trust Your Model and Your Results?

So, how do you know if you can trust your model – if your predictions and results are grounded in reality? A common saying in computer science is “Garbage in, Garbage out,” meaning results are limited by the quality of the input given to the model. The same principles apply here, and there are many ways in which error is introduced into the model [16]. Firstly, how accurately the force field describes the interatomic interactions. Significant variation can be observed between force fields, vastly influencing the accuracy of the results. Secondly, the amount of sampling as previously discussed, which is linked to the convergence of the system and thus the statistical validity of the results [3]. Additionally, the quality of the simulation software: the innate accuracy and consistency of running simulations with this software. An example of this is numerical instability due to rounding and other processes, and these errors propagate throughout a simulation [28]. Finally, how well the method was implemented: how well your simulation represents the real physical system, and the validity of the underlying assumptions. Overall, you aim to minimise or eliminate these errors through various approaches. However, it is an impossible task to try to quantify all these individual errors and their contribution to the overall accuracy of the simulation.

The best method to determine the accuracy of simulations is by comparison to experiments. Single-molecule techniques are invaluable here as they work on the same size scale as simulations. Some examples include: single-molecule Forster energy resonance transfer (smFRET),

where you measure distance-related energy transfer between two different labelled groups; and Raman spectroscopy, where your target is immobilised on a special surface, which enhances the signal enough to allow the detection of single molecule events [29]. Otherwise, more general techniques can be used to observe bulk properties of the system; in the scope of traditional biochemical and chemical techniques, for example, investigating enzyme processivity [30].

### Emerging Research Areas

This is an ever-evolving research field, moving fast with technological advancement. So, where to next? What are some predicted future directions?

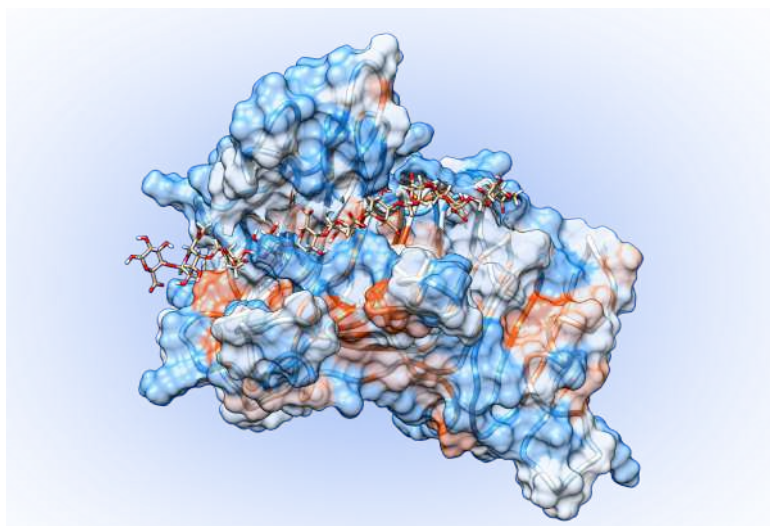
#### Simulating Large!

We are currently limited to the size of systems that can be simulated. Overcoming this would allow us to simulate a much greater diversity of systems, considering much larger complexes and structures, as well as their environments. Much progress has been achieved in being able to simulate protein folding – the formation of tertiary structures from amino acid sequences. Beyond tertiary structures, simulating the association between folded monomers into larger functional assemblies would open a window into understanding the function of large multimeric molecules [32]. The one billion atom benchmark was broken in 2019 with the simulation of an entire gene locus, containing 83 kilobases of DNA with 427 nucleosomes, to investigate condensed DNA structure [33].

#### Moving Towards a More Realistic Cellular Environment

We can think even larger, aiming to replicate the cellular environment with processes like relevant structural proteins, molecular crowding, and the cell wall/membrane. Simulating the interaction between molecules and the cell membrane has shown to be possible, with the example of polymyxin B1 in *Escherichia coli* (E.coli) [34]. Traditionally, the cellular environment is represented by the ‘solvation’ of the system – either implicitly as a uniform set of conditions, or explicitly with solvent molecules like water and ions added into the simulation box. However, in an actual cellular environment, much more diverse species are present, leading to molecular crowding. Molecular crowding can influence the equilibrium states for an enzyme as well as the diffusion of substrates to the active site of the enzyme, particularly if a substrate is large [35]. Hence, this will likely influence the kinetic rates for a given enzyme and free energy





Artwork by Nadia Snegirev.

landscape. However, aiming to add the molecular environment back into the system may mitigate some of the advantages of MD simulations, as the understanding of a process may be obscured by the ‘noise’ produced by other species in the simulation.

### Getting Kinetic Information from Simulations

As mentioned earlier, the majority of information which can be obtained from MD simulations is thermodynamic information, considering the relative energy differences between states. The typical output is a free energy surface for a given variable of interest, which describes a process/motion of the molecular system of study. On the other hand, kinetic information focuses on the transition between two or more energetic states and the time/rate of transition. This information is challenging to obtain as this energy barrier can be very high, meaning you might only see a single kinetic event in a long simulation. A system may also get ‘stuck’ in a particular conformation of lower energy, stopping you from seeing a whole kinetic event or multiple events. Even worse, these kinetic events happen on longer timescales than one is able to simulate [36]. This makes observing a kinetic event challenging, and even more challenging to sample enough of a particular kinetic event to make the results statistically valid.

An approach to doing this is by brute force: running many or very long simulations. Markov State Modelling was used to study Barnase-Barstar protein association, which is a diffusion limited process [37]. Thousands of simulations were run for a total duration of 2 milliseconds. This also used an algorithm to select a favourable starting structure to help speed up mapping of the conformational space. From all this information, a network of all the states was constructed with associated transition rates between the two states. Luckily, other methods are in development to achieve this using some ‘enhanced’ sampling methods [38].

### Protein Sequence Design

Modifying the amino acid sequence can be used to understand and design function. This is typically done through simulating mutants and single

residue substitutions – changing a residue crucial to function and observing a decreased or enhanced activity. This is a much simpler task using computational methods, such as changing lines of code, compared to achieving this experimentally – a nontrivial task that takes significant time.

MD simulations are a prime technique for engineering new proteins to improve and understand function. Protein stability can be enhanced, allowing proteins to be stable in more extreme conditions like high temperatures, high salt or solvent concentrations, or extreme pH. The use of new solvents can assist in lowering the impact of a chemical process, in combination with using biocatalysts like enzymes for a reaction. However, one challenge is the stability of enzymes in non-aqueous environments, limiting their application. The use of MD simulations showed that in ionic liquids, water essential for the reaction was being removed from the surface of the enzyme, inhibiting function. Hence, a proposed strategy is modifying the surface of the enzyme to improve compatibility [39]. Enzymes can also be engineered to accept different substrates, perform particular reactions, or even perform new reactions entirely; an example is Kemp Eliminase being engineered to perform a non-natural reaction [40].

Overall, this is a highly dynamic area of research, pushing the boundaries of what is possible and accelerating existing research. I will leave you with some of the quotes printed out after running a simulation, a tradition across many different simulation software that covers pop culture, science, and history.

**GROMACS reminds you: “If my PhD doesn’t allow me to be right on the internet, what is it even good for?” (Martin Vögele)**

**GROMACS reminds you: “Don’t bring an anecdote to a data fight.” (Molly Hodgdon)**

**GROMACS reminds you: “A robot will be truly autonomous when you instruct it to go to work and it decides to go to the beach instead.” (Brad Templeton)**



#### Nadia Snegirev - BSc (Hons), Computational Biophysics

Nadia is a current Honours student passionate about multidisciplinary research, which is how she ended up in computational biophysics. Her other passions include equity and advocacy, working to support a thriving and inclusive scientific community through her role as a SciSA Co-President.

# References

## Will AI End Scientific Thinking?

- [1] Tibees, "Reading Newton's Principia Mathematica by candlelight," YouTube. Jun. 17, 2019. [Online]. Available: <https://www.youtube.com/watch?v=2DBeFqc6c8Y>
- [2] S. Mcleod and S. Mcleod, "Karl Popper: Theory of Falsification," Simply Psychology, Jul. 2023, [Online]. Available: <https://www.simplypsychology.org/karl-popper.html#:~:text=The%20Falsification%20Principle%2C%20proposed%20by,by%20observing%20a%20black%20swan.>
- [3] L. Spinney, "Are we witnessing the dawn of post-theory science?" The Guardian, Jan. 09, 2022. [Online]. Available: <https://www.theguardian.com/technology/2022/jan/09/are-we-witnessing-the-dawn-of-post-theory-science>
- [4] K. Hamilton, "Here's what we know about OceanGate's sub that tours Titanic—Using 1 button," Forbes, Jun. 19, 2023. [Online]. Available: <https://www.forbes.com/sites/katherinehamilton/2023/06/19/heres-what-we-know-about-oceangates-sub-that-tours-titanic-using-1-button/?sh=242c33353de9>
- [5] M. Gurucharan, "Top 12 Commerce Project Topics & Ideas in 2023 [For Freshers]," upGrad blog. <https://www.upgrad.com/blog/basic-cnn-architecture/>.
- [6] "Create simple deep learning neural network for classification - MATLAB & Simulink example - MathWorks Australia." <https://au.mathworks.com/help/deeplearning/ug/create-simple-deep-learning-network-for-classification.html>
- [7] P. Ratan, "What is the Convolutional Neural Network Architecture?," Analytics Vidhya, Jan. 2021, [Online]. Available: <https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/>
- [8] Wikipedia contributors, "Explainable artificial intelligence," Wikipedia, Sep. 2023, [Online]. Available: [https://en.wikipedia.org/wiki/Explainable\\_artificial\\_intelligence](https://en.wikipedia.org/wiki/Explainable_artificial_intelligence).
- [9] S. Gibbons, "2023 Business Predictions As AI And Automation Rise In Popularity," Forbes, Feb. 02, 2023. [Online]. Available: <https://www.forbes.com/sites/serenitygibbons/2023/02/02/2023-business-predictions-as-ai-and-automation-rise-in-popularity/?sh=358a74cc744b>

## A Balancing Act: Managing the Impact of Genetically Modified Crops on Soil Microorganisms

- [1] K. E. Dunfield and J. J. Germida, "Impact of genetically modified crops on soil- and Plant-Associated microbial communities," Journal of Environmental Quality, vol. 33, no. 3, p. 806, Jan. 2004, doi: 10.2134/jeq2004.0806.
- [2] Z. Guan, S. Lu, Y.-L. Huo, Z. Guan, B. Liu, and W. Wei, "Do genetically modified plants affect adversely on soil microbial communities?," Agriculture, Ecosystems & Environment, vol. 235, pp. 289–305, Nov. 2016, doi: 10.1016/j.agee.2016.10.026.
- [3] A. Spök, T. Sprink, A. C. Allan, T. Yamaguchi, and C. Dayé, "Towards social acceptability of genome-edited plants in industrialised countries? Emerging evidence from Europe, United States, Canada, Australia, New Zealand, and Japan," Frontiers in Genome Editing, vol. 4, Aug. 2022, doi: 10.3389/fgeed.2022.899331.
- [4] X. Zhao, Y. Jiang, Q. Liu, H. Yang, Z. Wang, and M. Zhang, "Effects of Drought-Tolerant EA-DREB2B transgenic sugarcane on bacterial

communities in soil," Frontiers in Microbiology, vol. 11, May 2020, doi: 10.3389/fmicb.2020.00704.

- [5] F. D. Provenza, J. J. Villalba, J. P. Haskell, J. W. MacAdam, T. C. Griggs, and R. D. Wiedmeier, "The value to herbivores of plant physical and chemical diversity in time and space," Crop Science, vol. 47, no. 1, pp. 382–398, Jan. 2007, doi: 10.2135/cropsci2006.02.0083.
- [6] W. Tian, X. Yi, S. Liu, C. Zhou, and A. Wang, "Effect of transgenic cotton continuous cropping on soil bacterial community," Annals of Microbiology, vol. 70, no. 1, Oct. 2020, doi: 10.1186/s13213-020-01602-8.
- [7] S.S.-E.-A. Zaidi, A. Mahas, H. Vanderschuren, and M. M. Mahfouz, "Engineering crops of the future: CRISPR approaches to develop climate-resilient and disease-resistant plants," Genome Biology, vol. 21, no. 1, Nov. 2020, doi: 10.1186/s13059-020-02204-y.
- [8] A. Tsatsakis et al., "Environmental impacts of genetically modified plants: A review," Environmental Research, vol. 156, pp. 818–833, Jul. 2017, doi: 10.1016/j.envres.2017.03.011.
- [9] M. J. Oliver, "Why we need GMO crops in agriculture," PubMed Central (PMC), Dec. 01, 2014. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6173531/>
- [10] "Genetic Modification: The New Zealand Approach," Ministry for the Environment, Jun. 2004.
- [11] "Genetic modification," Ministry for the Environment, Oct. 2001.

## Mimivirus: The Microbe-Mimicking Virus

- [1] J. L. Van Etten, "Giant Viruses", American Scientist, Feb. 06, 2017. <https://www.americanscientist.org/article/giant-viruses>
- [2] B. Bryson, "Microbial You" in The Body: A Guide for Occupants. Penguin Random House LLC. 2019.
- [3] L. Metrus and K. McCarthy, "How to determine your hair strand type and density," Byrdie, Aug. 2022, [Online]. Available: <https://www.byrdie.com/find-your-hair-type#:~:text=%22Hair%20with%20a%20small%20diameter,diameter%20of%20about%20120%20microns.%22>
- [4] J. Louten, "Virus Structure and Classification," Essential Human Virology, pp. 19–29, May 2016, Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7150055>
- [5] S. Aryal, "Different size, shape and

- arrangement of bacterial cells," *Microbiology Info.com*, Aug. 2022, [Online]. Available: <https://microbiologyinfo.com/different-size-shape-and-arrangement-of-bacterial-cells/#:~:text=Size%20of%20Bacterial%20Cell&text=The%20average%20diameter%20of%20spherical,0%20%C2%B5m>.
- [6] R. Balia. Adapted from Figure 1 in "Types of Microorganisms | Microbiology," [courses.lumenlearning.com](https://courses.lumenlearning.com/microbiology/).
- [7] C-R. Huang and S. J. Lo, "Evolution and Diversity of the Human Hepatitis D Virus Genome," *Advances in Bioinformatics*, vol. 2010, pp. 1–9, Feb. 2010, doi:<https://doi.org/10.1155/2010/323654>
- [8] K. V. Chaitanya, "Structure and Organization of Virus Genomes," *Genome and Genomics*, pp. 1–30, Nov. 2019, doi: [https://doi.org/10.1007/978-981-15-0702-1\\_1](https://doi.org/10.1007/978-981-15-0702-1_1)
- [9] R. A. Khailany, M. Safdar, and M. Ozaslan, "Genomic characterization of a novel SARS-CoV-2," *Gene Reports*, vol. 19, p. 100682, Jun. 2020, doi: <https://doi.org/10.1016/j.genrep.2020.100682>
- [10] A. Villalta et al., "The giant mimivirus 1.2 Mb genome is elegantly organized into a 30-nm diameter helical protein shield," *eLife*, vol. 11, p. e77607, Jul. 2022, doi: <https://doi.org/10.7554/eLife.77607>
- [11] "GenomeSize-an overview | ScienceDirect Topics," [www.sciencedirect.com](https://www.sciencedirect.com/topics/medicine-and-dentistry/genome-size#:~:text=Bacterial%20genome%20size%20ranges%20from). <https://www.sciencedirect.com/topics/medicine-and-dentistry/genome-size#:~:text=Bacterial%20genome%20size%20ranges%20from>
- [12] T. K. Mohanta, A. K. Mishra, Y. K. Mohanta, and A. Al-Harrasi, "Virtual 2D mapping of the viral proteome reveals host-specific modality distribution of molecular weight and isoelectric point," *Scientific Reports*, vol. 11, no. 1, Oct. 2021, doi: [10.1038/s41598-021-00797-3](https://doi.org/10.1038/s41598-021-00797-3).
- [13] M. Legendre, S. Santini, A. Rico, C. Abergel, and J.-M. Claverie, "Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing," *Virology Journal*, vol. 8, no. 1, Mar. 2011, doi: <https://doi.org/10.1186/1743-422x-8-99>.
- [14] N. Brandes and M. Linial, "Giant Viruses—Big Surprises," *Viruses*, vol. 11, no. 5, p. 404, Apr. 2019, doi: <https://doi.org/10.3390/v11050404>.
- [15] P. Colson, B. La Scola, A. Levasseur, G. Caetano-Anollés, and D. Raoult, "Mimivirus: leading the way in the discovery of giant viruses of amoebae," *Nature Reviews Microbiology*, vol. 15, no. 4, pp. 243–254, Apr. 2017, doi: [10.1038/nrmicro.2016.197](https://doi.org/10.1038/nrmicro.2016.197).
- [16] R. Balia. Adapted from Figure 2 in J. S. Abrahão et al., "Acanthamoeba polyphaga mimivirus and other giant viruses: an open field to outstanding discoveries," *Virology Journal*, vol. 11, no. 1, Jun. 2014, doi: [10.1186/1743-422x-11-120](https://doi.org/10.1186/1743-422x-11-120).
- [17] F. P. Dornas et al., "Detection of mimivirus genome and neutralizing antibodies in humans from Brazil," *Archives of Virology*, vol. 162, no. 10, pp. 3205–3207, Jun. 2017, doi: [10.1007/s00705-017-3455-5](https://doi.org/10.1007/s00705-017-3455-5).
- [18] L. C. F. Silva et al., "A resourceful giant: APMV is able to interfere with the human type I interferon system," *Microbes and Infection*, vol. 16, no. 3, pp. 187–195, Mar. 2014, doi: [10.1016/j.micinf.2013.11.011](https://doi.org/10.1016/j.micinf.2013.11.011).
- [19] K. Moelling and F. Broecker, "Viruses and evolution – Viruses First? A personal perspective," *Frontiers in Microbiology*, vol. 10, Mar. 2019, doi: [10.3389/fmicb.2019.00523](https://doi.org/10.3389/fmicb.2019.00523).
- cell-based models of human embryo development," *Seminars in Cell & Developmental Biology*, vol. 131, pp. 44–57, Nov. 2022, doi: <https://doi.org/10.1016/j.semcd.2022.05.023>.
- [3] "Leonardo da Vinci (Vinci 1452-Amboise 1519) - The fetus in the womb; sketches and notes on reproduction," [www.rct.uk](http://www.rct.uk). <https://www.rct.uk/collection/919102/the-fetus-in-the-womb-sketches-and-notes-on-reproduction>
- [4] J. M. Tanner, "A History of the Study of Human Growth.," *Isis*, vol. 73, no. 3, pp. 455–456, Sep. 1982, doi: <https://doi.org/10.1086/353075>.
- [5] S. Flierman, M. Tijsterman, M. Rousian, and B. S. de Bakker, "Discrepancies in Embryonic Staging: Towards a Gold Standard," *Life*, vol. 13, no. 5, pp. 1084–1084, Apr. 2023, doi: <https://doi.org/10.3390/life13051084>.
- [6] M. A. HILL, "Early Human Development," *Clinical Obstetrics and Gynecology*, vol. 50, no. 1, pp. 2–9, Mar. 2007, doi: <https://doi.org/10.1097/grf.0b013e31802f119d>.
- [7] C. Rydberg and K. Tunón, "Detection of fetal abnormalities by second-trimester ultrasound screening in a non-selected population," *Acta Obstetrica et Gynecologica Scandinavica*, vol. 96, no. 2, pp. 176–182, Nov. 2016, doi: <https://doi.org/10.1111/aogs.13037>.
- [8] A. Marouane, R. A. C. M. Olde Keizer, G. W. J. Frederix, L. E. L. M. Vissers, W. P. de Boode, and W. A. G. van Zelst-Stams, "Congenital anomalies and genetic disorders in neonates and infants: a single-center observational cohort study," *European Journal of Pediatrics*, vol. 181, no. 1, pp. 359–367, Aug. 2021, doi: <https://doi.org/10.1007/s00431-021-04213-w>.
- [9] M. H. Wojcik and P. B. Agrawal, "Deciphering congenital anomalies for the next generation," *Molecular Case Studies*, vol. 6, no. 5, p. a005504, Aug. 2020, doi: <https://doi.org/10.1101/mcs.a005504>.
- [10] A. Flamier, S. Singh, and T. P. Rasmussen, "A standardized human embryoid body platform for the detection and analysis of teratogens," *PLoS ONE*, vol. 12, no. 2, Feb. 2017, doi: <https://doi.org/10.1371/journal.pone.0171101>.
- [11] N. Vargesson, "The teratogenic effects of thalidomide on limbs," *Journal of Hand Surgery (European Volume)*, vol. 44, no. 1, pp. 88–95, Oct. 2018, doi: <https://doi.org/10.1177/1753193418805249>.
- [12] S. U. Morton and D. Brodsky, "Fetal Physiology and the Transition to Extrauterine Life," *Clinics in Perinatology*,

### Why Early Life Matters Most: The Developmental Origins of Health and Disease Framework

- [1] D. Crawford, "Biological basis of child health 2: introduction to fertilisation, prenatal development and birth," *Nursing children and young people*, vol. 32, no. 3, pp. 32–41, Apr. 2020, doi: <https://doi.org/10.7748/ncyp.2020.e1245>.
- [2] B. Sozen, Deniz Conkar, and J. V. Veenvliet, "Carnegie in 4D? Stem-

vol. 43, no. 3, pp. 395–407, 2016, doi: <https://doi.org/10.1016/j.clp.2016.04.001>.

- [13] P. C. Hindmarsh, M. P. P. Geary, C. H. Rodeck, J. C. P. Kingdom, and T. J. Cole, "Factors Predicting Ante- and Postnatal Growth," *Pediatric Research*, vol. 63, no. 1, pp. 99–102, Jan. 2008, doi: <https://doi.org/10.1203/PDR.0b013e31815b8e8f>.
- [14] C. Kalbermatter, N. Fernandez Trigo, S. Christensen, and S. C. Ganai-Vonarburg, "Maternal Microbiota, Early Life Colonization and Breast Milk Drive Immune Development in the Newborn," *Frontiers in Immunology*, vol. 12, May 2021, doi: <https://doi.org/10.3389/fimmu.2021.683022>.
- [15] M. Penkler, M. Hanson, R. Biesma, and R. Müller, "DOHaD in science and society: emergent opportunities and novel responsibilities," *Journal of Developmental Origins of Health and Disease*, vol. 10, no. 3, pp. 268–273, Nov. 2018, doi: <https://doi.org/10.1017/s2040174418000892>.
- [16] D. Barker, "Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales," *The Lancet*, vol. 327, no. 8489, pp. 1077–1081, May 1986, doi: [https://doi.org/10.1016/s0140-6736\(86\)91340-1](https://doi.org/10.1016/s0140-6736(86)91340-1).
- [17] K. Suzuki, "The developing world of DOHaD," *Journal of Developmental Origins of Health and Disease*, vol. 9, no. 3, pp. 266–269, Sep. 2017, doi: <https://doi.org/10.1017/s2040174417000691>.
- [18] D. E. Handy, R. Castro, and J. Loscalzo, "Epigenetic Modifications: basic mechanisms and role in cardiovascular disease," *Circulation*, vol. 123, no. 19, pp. 2145–2156, May 2011, doi: <https://doi.org/10.1161/circulationaha.110.956839>.
- [19] T. Bianco-Miotto, J. M. Craig, Y. P. Gasser, S. J. van Dijk, and S. E. Ozanne, "Epigenetics and DOHaD: from basics to birth and beyond," *Journal of Developmental Origins of Health and Disease*, vol. 8, no. 5, pp. 513–519, Sep. 2017, doi: <https://doi.org/10.1017/s2040174417000733>.
- [20] F. Lynch, S. Lewis, I. Macciocca, and J. M. Craig, "Epigenetics and DOHaD: how translation to predictive testing will require a better public understanding," *Journal of Developmental Origins of Health and Disease*, vol. 13, no. 4, pp. 424–430, Oct. 2021, doi: <https://doi.org/10.1017/s2040174421000568>.
- [21] D. J. Barker, "The fetal and infant origins of adult disease," *BMJ*, vol. 301, no. 6761, pp. 1111–1111, Nov. 1990, doi: <https://doi.org/10.1136/bmj.301.6761.1111>.
- [22] P. D. Gluckman, M. A. Hanson, C. Cooper, and K. L. Thornburg, "Effect of In Utero and Early-Life Conditions on Adult Health and Disease," *New England Journal of Medicine*, vol. 359, no. 1, pp. 61–73, Jul. 2008, doi: <https://doi.org/10.1056/nejmra0708473>.
- [23] J. C. Chambers et al., "Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study," *The Lancet Diabetes & Endocrinology*, vol. 3, no. 7, pp. 526–534, Jul. 2015, doi: [https://doi.org/10.1016/S2213-8587\(15\)00127-8](https://doi.org/10.1016/S2213-8587(15)00127-8).
- [24] T. Dayeh et al., "DNA methylation of loci within ABCG1 and PHOSPHO1 in blood DNA is associated with future type 2 diabetes risk," *Epigenetics*, vol. 11, no. 7, pp. 482–488, Jun. 2016, doi: <https://doi.org/10.1080/15592294.2016.1178418>.
- [25] E. W. Demerath et al., "Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci," *Human Molecular Genetics*, vol. 24, no. 15, pp. 4464–4479, May 2015, doi: <https://doi.org/10.1093/hmg/ddv161>.
- [26] K. J. Dick et al., "DNA methylation and body-mass index: a genome-wide analysis," *Lancet (London, England)*, vol. 383, no. 9933, pp. 1990–1998, Jun. 2014, doi: [https://doi.org/10.1016/S0140-6736\(13\)62674-4](https://doi.org/10.1016/S0140-6736(13)62674-4).
- [27] K. M. Godfrey, K. A. Lillycrop, G. C. Burdge, P. D. Gluckman, and M. A. Hanson, "Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease," *Pediatric research*, vol. 61, no. 5 Pt 2, pp. 5R10R, 2007, doi: <https://doi.org/10.1203/pdr.0b013e318045bedb>.
- [28] S. Lapehn and A. G. Paquette, "The Placental Epigenome as a Molecular Link Between Prenatal Exposures and Fetal Health Outcomes Through the DOHaD Hypothesis," *Current Environmental Health Reports*, vol. 9, no. 3, pp. 490–501, Apr. 2022, doi: <https://doi.org/10.1007/s40572-022-00354-8>.
- [29] A. Cardenas, S. M. Lutz, T. M. Everson, P. Perron, L. Bouchard, and M.-F. Hivert, "Mediation by Placental DNA Methylation of the Association of Prenatal Maternal Smoking and Birth Weight," *American Journal of Epidemiology*, vol. 188, no. 11, pp. 1878–1886, Sep. 2019, doi: <https://doi.org/10.1093/aje/kwz184>.
- [30] T. M. Everson et al., "Placental DNA methylation signatures of maternal smoking during pregnancy and potential impacts on fetal growth," *Nature Communications*, vol. 12, no. 1, Aug. 2021, doi: <https://doi.org/10.1038/s41467-021-24558-y>.
- [31] E. R. Norwitz and J. A. Greenberg, "Beyond antenatal corticosteroids: what did mont liggins teach us?," *Reviews in obstetrics & gynecology*, vol. 3, no. 3, pp. 79–80, Jan. 2010.
- [32] G. C. Liggins and R. N. Howie, "A Controlled Trial of Antepartum Glucocorticoid Treatment for Prevention of the Respiratory Distress Syndrome in Premature Infants," *Pediatrics*, vol. 50, no. 4, pp. 515–525, Oct. 1972, doi: <https://doi.org/10.1542/peds.50.4.515>.

### **Calculators, Computers, and ChatGPT: How are University Students Using Generative AI Chatbots in Education?**

- [1] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, vol. 2, pp. 100033–100033, Jan. 2021, doi: [10.1016/j.caeai.2021.100033](https://doi.org/10.1016/j.caeai.2021.100033).
- [2] J. Weizenbaum, "ELIZA — a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 26, no. 1, pp. 23–28, Jan. 1983, doi: [10.1145/357980.357991](https://doi.org/10.1145/357980.357991).
- [3] L. Guo, D. Wang, F. Gu, Y. Li, Y. Wang,

- and R. Zhou, "Evolution and trends in intelligent tutoring systems research: a multidisciplinary and scientometric view," *Asia Pacific Education Review*, vol. 22, no. 3, pp. 441–461, May 2021, doi: 10.1007/s12564-021-09697-7.
- [4] OpenAI, "Introducing ChatGPT," [openai.com](https://openai.com/blog/chatgpt). <https://openai.com/blog/chatgpt> (accessed Aug. 30 2023).
- [5] A. Turing, "Computing Machinery and Intelligence," *Mind*, vol. LIX, no. 236, pp. 433–460, 1950, doi: <https://doi.org/10.1093/mind/lix.236.433>.
- [6] J. Schofield, "Computer chatbot 'Eugene Goostman' passes the Turing test," *ZDNET*, Jun. 08, 2014. <https://www.zdnet.com/article/computer-chatbot-eugene-goostman-passes-the-turing-test/> (accessed Aug. 30, 2023).
- [7] C. Biever, "ChatGPT broke the Turing test — the race is on for new ways to assess AI," *Nature*, vol. 619, no. 7971, pp. 686–689, Jul. 2023, doi: <https://doi.org/10.1038/d41586-023-02361-7>.
- [8] S. Lau., G. Philip, "Teaching Programming in the Age of ChatGPT," O'Reilly Media, Jul. 18, 2023. <https://www.oreilly.com/radar/teaching-programming-in-the-age-of-chatgpt/> (accessed Aug. 30, 2023).
- [9] University of Auckland Education Office, "Generative AI in learning and teaching," [teachwell.auckland.ac.nz](https://teachwell.auckland.ac.nz/news/chatgpt-ai/). <https://teachwell.auckland.ac.nz/news/chatgpt-ai/> (accessed Aug. 30 2023).
- [10] University of Auckland Education Office, "The use of generative AI tools in coursework," [teachwell.auckland.ac.nz](https://teachwell.auckland.ac.nz/resources/assessment/ai-tools-in-coursework/). <https://teachwell.auckland.ac.nz/resources/assessment/ai-tools-in-coursework/> (accessed Aug. 30 2023).
- [11] A. H. Maslow, *The Psychology of Science: A Reconnaissance*. New York, NY: Harper & Row, 1966, ISBN: 9780809261307.
- [12] M. M. Waldrop, *The Dream Machine: J.C. Licklider and the Revolution That Made Computing Personal*. San Francisco, CA: Stripe Press, 2018, pp. 66, ISBN: 9781732265110.
- [13] Office of the Prime Minister's Chief Science Advisor, "Why is regulating AI such a challenge?," [pmcsa.ac.nz](https://www.pmcsa.ac.nz/2023/07/13/why-is-regulating-ai-such-a-challenge/). <https://www.pmcsa.ac.nz/2023/07/13/why-is-regulating-ai-such-a-challenge/> (accessed Aug. 30 2023).
- [14] S. Papert, *Mindstorms: Children, Computers, and Powerful Ideas*. New York, NY: Basic Books, 1980, ISBN: 0465046274.
- [15] R. P. Feynman, "Surely You're Joking, Mr. Feynman!": *Adventures of a Curious Character*. W. W. Norton & Company, 2018, pp. 165–6.
- Integrated Assessment, vol. 4, no. 1, pp. 5-17, 2003/03// 2003, doi: 10.1076/iaij.4.1.5.16466.
- [6] B. Di Ventura, C. Lemerle, K. Michalodimitrakis, and L. Serrano, "From in vivo to in silico biology and back," *Nature*, vol. 443, no. 7111, pp. 527-533, 2006, doi: 10.1038/nature05127.
- [7] M. Karplus, "Molecular dynamics of biological macromolecules: A brief history and perspective," *Biopolymers*, vol. 68, no. 3, pp. 350-358, 2003, doi: 10.1002/bip.10266.
- [8] J. Jumper et al., "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583-589, 2021, doi: 10.1038/s41586-021-03819-2.
- [9] R. Wu et al., "High-resolution de novo structure prediction from primary sequence," Cold Spring Harbor Laboratory, 2022.
- [10] O. Guvench and A. D. MacKerell, "Comparison of Protein Force Fields for Molecular Dynamics Simulations," in *Molecular Modeling of Proteins*, A. Kukol Ed. Totowa, NJ: Humana Press, 2008, pp. 63-88.
- [11] U. Doshi and D. Hamelberg, "Reoptimization of the AMBER Force Field Parameters for Peptide Bond (Omega) Torsions Using Accelerated Molecular Dynamics," *The Journal of Physical Chemistry B*, vol. 113, no. 52, pp. 16590-16595, 2009/12/31 2009, doi: 10.1021/jp907388m.
- [12] A. D. MacKerell, N. Banavali, and N. Foloppe, "Development and current status of the CHARMM force field for nucleic acids," *Peptide science*, vol. 56, no. 4, pp. 257-265, 2000, doi: 10.1002/1097-0282(2000)56:4<257::AID-BIP10029>3.0.CO;2-W.
- [13] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, "UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations," *Journal of the American Chemical Society*, vol. 114, no. 25, pp. 10024-10035, 1992, doi: 10.1021/ja00051a040.
- [14] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," *Journal of the American Chemical Society*, vol. 118, no. 45, pp. 11225-11236, 1996, doi: 10.1021/ja9621760.
- [15] A. Grossfield and D. M. Zuckerman, "Quantifying uncertainty and sampling quality in biomolecular simulations," *Annu*

### Simulating the Mechanisms of Life Through the Lens of the Virtual World

- [1] H. Andersen, "The history of reductionism versus holistic approaches to scientific research," *Endeavour*, vol. 25, no. 4, pp. 153-156, 2001.
- [2] K. F. Schaffner, "The Watson-Crick model and reductionism," *The British Journal for the Philosophy of Science*, vol. 20, no. 4, pp. 325-348, 1969.
- [3] A. Grossfield and D. M. Zuckerman, "Quantifying uncertainty and sampling quality in biomolecular simulations," *Annual reports in computational chemistry*, vol. 5, pp. 23-48, 2009/01/01/ 2009, doi: 10.1016/S1574-1400(09)00502-7.
- [4] H. M. Regan, M. Colyvan, and M. A. Burgman, "A taxonomy and treatment of uncertainty for ecology and conservation biology," *Ecological Applications*, vol. 12, no. 2, pp. 618-628, 2002/04// 2002, doi: 10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2.
- [5] W. E. Walker et al., "Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support,"

Rep Comput Chem, vol. 5, pp. 23-48, Jan 1 2009, doi: 10.1016/S1574-1400(09)00502-7.

- [16] W. F. van Gunsteren and A. E. Mark, "Validation of molecular dynamics simulation," *The Journal of chemical physics*, vol. 108, no. 15, pp. 6109-6116, 1998.
- [17] P. J. Flory, *Statistical Mechanics of Chain Molecules*. New York: Interscience Publishers, 1969.
- [18] W. F. van Gunsteren and H. J. Berendsen, "Groningen molecular simulation (GROMOS) library manual," *Biomos, Groningen*, vol. 24, no. 682704, p. 13, 1987.
- [19] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," *Journal of Computational Chemistry*, vol. 32, no. 10, pp. 2319-2327, 2011, doi: 10.1002/jcc.21787.
- [20] J. V. Ribeiro et al., "QwikMD – Integrative Molecular Dynamics Toolkit for Novices and Experts," *Scientific Reports*, vol. 6, no. 1, p. 26536, 2016, doi: 10.1038/srep26536.
- [21] M. Pandey et al., "The transformational role of GPU computing and deep learning in drug discovery," *Nature Machine Intelligence*, vol. 4, no. 3, pp. 211-221, 2022, doi: 10.1038/s42256-022-00463-x.
- [22] J. D. Owens, M. Houston, D. Luebke, S. Green, J. E. Stone, and J. C. Phillips, "GPU Computing," *Proceedings of the IEEE*, vol. 96, no. 5, pp. 879-899, 2008, doi: 10.1109/JPROC.2008.917757.
- [23] J. E. Stone, J. C. Phillips, P. L. Freddolino, D. J. Hardy, L. G. Trabuco, and K. Schulten, "Accelerating molecular modeling applications with graphics processors," *Journal of Computational Chemistry*, vol. 28, no. 16, pp. 2618-2640, 2007, doi: 10.1002/jcc.20829.
- [24] D. E. Shaw et al., "Millisecond-scale molecular dynamics simulations on Anton," presented at the Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, Portland, Oregon, 2009. [Online]. Available: <https://doi.org/10.1145/1654059.1654126>.
- [25] I. Buch, M. J. Harvey, T. Giorgino, D. P. Anderson, and G. De Fabritiis, "High-Throughput All-Atom Molecular Dynamics Simulations Using Distributed Computing," *Journal of Chemical Information and Modeling*, vol. 50, no. 3, pp. 397-403, 2010, doi: 10.1021/ci900455r.
- [26] J.-P. Ebejer, S. Fulle, G. M. Morris, and P. W. Finn, "The emerging role of cloud computing in molecular modelling," *Journal of molecular graphics & modelling*, vol. 44, pp. 177-187, 2013, doi: 10.1016/j.jmgm.2013.06.002.
- [27] J. A. McCammon, B. R. Gelin, and M. Karplus, "Dynamics of folded proteins," *Nature*, vol. 267, no. 5612, pp. 585-590, 1977, doi: 10.1038/267585a0.
- [28] S. A. Hollingsworth and R. O. Dror, "Molecular Dynamics Simulation for All," *Neuron*, vol. 99, no. 6, pp. 1129-1143, Sep 19 2018, doi: 10.1016/j.neuron.2018.08.011.
- [29] R. A. Lippert et al., "A common, avoidable source of error in molecular dynamics integrators," *The Journal of chemical physics*, vol. 126, no. 4, p. 046101, 2007, doi: 10.1063/1.2431176.
- [30] H. Miller, Z. Zhou, J. Shepherd, A. J. M. Wollman, and M. C. Leake, "Single-molecule techniques in biophysics: a review of the progress in methods and applications," *Reports on Progress in Physics*, vol. 81, no. 2, p. 024601, 2018, doi: 10.1088/1361-6633/aa8a02.
- [31] L. M. Kent, T. S. Loo, L. D. Melton, D. Mercadante, M. A. K. Williams, and G. B. Jameson, "Structure and Properties of a Non-processive, Salt-requiring, and Acidophilic Pectin Methyltransferase from *Aspergillus niger* Provide Insights into the Key Determinants of Processivity Control," *Journal of Biological Chemistry*, vol. 291, no. 3, pp. 1289-1306, 2016, doi: 10.1074/jbc.m115.673152.
- [32] J. Zhu et al., "Protein Assembly by Design," *Chemical Reviews*, vol. 121, no. 22, pp. 13701-13796, 2021/11/24 2021, doi: 10.1021/acs.chemrev.1c00308.
- [33] J. Jung et al., "Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations," *Journal of Computational Chemistry*, vol. 40, no. 21, pp. 1919-1930, 2019, doi: 10.1002/jcc.25840.
- [34] C. Pedebos, I. P. S. Smith, A. Boags, and S. Khalid, "The hitchhiker's guide to the periplasm: Unexpected molecular interactions of polymyxin B1 in *E. coli*," *Structure*, vol. 29, no. 5, pp. 444-456.e2, 2021, doi: 10.1016/j.str.2021.01.009.
- [35] L. Heo, Y. Sugita, and M. Feig, "Protein assembly and crowding simulations," *Current Opinion in Structural Biology*, vol. 73, p. 102340, 2022/04/01/ 2022, doi: <https://doi.org/10.1016/j.sbi.2022.102340>.
- [36] J. M. L. Ribeiro, S.-T. Tsai, D. Pramanik, Y. Wang, and P. Tiwary, "Kinetics of Ligand-Protein Dissociation from All-Atom Simulations: Are We There Yet?," *Biochemistry*, vol. 58, no. 3, pp. 156-165, 2019, doi: 10.1021/acs.biochem.8b00977.
- [37] N. Plattner, S. Doerr, G. De Fabritiis, and F. Noé, "Complete protein-protein association kinetics in atomic detail revealed by molecular dynamics simulations and Markov modelling," *Nature Chemistry*, vol. 9, no. 10, pp. 1005-1011, 2017, doi: 10.1038/nchem.2785.
- [38] Y. Yonezawa, H. Shimoyama, and H. Nakamura, "Multicanonical molecular dynamics simulations combined with Metadynamics for the free energy landscape of a biomolecular system with high energy barriers," *Chemical Physics Letters*, vol. 501, no. 4, pp. 598-602, 2011/01/07/ 2011, doi: <https://doi.org/10.1016/j.cplett.2010.11.061>.
- [39] S. Pramanik, G. V. Dhoke, K.-E. Jaeger, U. Schwaneberg, and M. D. Davari, "How To Engineer Ionic Liquids Resistant Enzymes: Insights from Combined Molecular Dynamics and Directed Evolution Study," *ACS Sustainable Chemistry & Engineering*, vol. 7, no. 13, pp. 11293-11302, 2019, doi: 10.1021/acssuschemeng.9b00752.
- [40] D. Röthlisberger et al., "Kemp elimination catalysts by computational enzyme design," *Nature*, vol. 453, no. 7192, pp. 190-195, 2008, doi: 10.1038/nature06879.



# SCIENCE WRITERS WANTED

The *Scientific* publishes science articles about student research, areas of interest, explainers, and opinion.

We are always accepting submissions. Scan the QR code or email us if you are interested in appearing in our next edition!



### Fun Fact

Platypuses are well-known for being one of the only egg-laying mammals - puggles (baby platypuses) are born (i.e. hatch from their leathery eggs) in a semi-fetal state, much like marsupial mammals such as kangaroos.

### Closing Comments

Just like that, another year of Scientific has concluded! Thanks again to all of our readers, writers, and the executive team, and we hope you've enjoyed what we've created for 2023.

While 2023 may be wrapping up, we've still got one more edition for Volume 3 coming in 2024 - keep your eyes peeled, curiosity piqued, and we'll see you next year.

Until then, Ka kite anō!

### Contact Us

Website - [uoascientific.com](http://uoascientific.com)  
Email - [scientific.review.uoa@gmail.com](mailto:scientific.review.uoa@gmail.com)  
Instagram - [uoascientific](https://www.instagram.com/uoascientific)  
Facebook - [uoascientific](https://www.facebook.com/uoascientific)  
Twitter - [@uoascientific](https://twitter.com/uoascientific)  
LinkedIn - [The University of Auckland Scientific](https://www.linkedin.com/company/the-university-of-auckland-scientific)

This magazine has been supported as a student initiative by the Faculty of Science at the University of Auckland. Articles in this publication are not peer reviewed or endorsed by the faculty and may reflect personal opinion that is not that of the university.



I'm a magazine, but I have feelings too. Don't throw me away!