

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339180663>

Bio-informational futures: The convergence of artificial intelligence and synthetic biology

Article in EMBO Reports · February 2020

DOI: 10.15252/embr.202050036

CITATIONS

5

READS

296

3 authors:



Thom A Dixon

Macquarie University

8 PUBLICATIONS 21 CITATIONS

[SEE PROFILE](#)



Natalie Curach

Bioplatforms Australia Ltd.

17 PUBLICATIONS 467 CITATIONS

[SEE PROFILE](#)



Isak S. Pretorius

Macquarie University

354 PUBLICATIONS 20,568 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Cell wall-dependent phenotypes [View project](#)



Constructing a synthetic *Saccharomyces cerevisiae* pan-genome neo-chromosome [View project](#)

Bio-informational futures

The convergence of artificial intelligence and synthetic biology

Thom A Dixon¹, Natalie C Curach^{1,2} & Isak S Pretorius¹

Biofoundries are the place where synthetic biology and artificial intelligence are converging into technology platforms with the capacity to create synthetic organisms at a massive scale. This will not just generate new approaches to and solutions for intractable policy problems, but biofoundries stand to transform society itself. The confluence of the life sciences and the information sciences gives rise to future scenarios that can be broadly categorised as *bio-informational futures*. It is therefore important to recognise and evaluate their transformational potential for the coming decade when synthetic biology applications in commercial, government and research organisations will become increasingly visible to the public. Scientists, technologists, practitioners and policy makers need to work closely with biofoundry operators and the public to shape and guide this bio-informational future.

Biological foundries

Genome foundries or biofoundries use high-throughput modular laboratory equipment within a highly automated design-build-test-learn workflow to design and construct genetically reprogrammed organisms for research and biotechnology. Examples of academic and industrial biofoundries and biofoundry research workflows include the MIT Foundry, Gingko Bioworks, Amyris and the London Biofoundry. These can be single-site locations or globally integrated laboratory solutions, whereby multiple sites each play a specific role within an industrial or research workflow. Generally, biofoundries are typified not by their location, the equipment used or the data generated, but by their automated workflow for the sole purpose of

engineering biological systems. They open up avenues of scientific inquiry and industrial commercialisation that have previously been prohibitively expensive, overly resource-intensive or time-constrained.

.....
“The confluence of the life sciences and the information sciences gives rise to future scenarios that can be broadly categorised as bio-informational futures.”

This commentary will frame the convergence of synthetic biology and artificial intelligence in the biofoundry around five policy challenges: Healthy People, Prosperous Economies, Resilient Societies, a Secure Planet and Innovative Technologies (Fig 1). The domain of *Healthy People* includes research focussed on health outcomes, integrated health care and wellness. *Resilient Societies* comprises innovations that enable and enhance ethical, just and inclusive communities. The domains of *Prosperous Economies* and *Secure Planet* encapsulate the deployment of biofoundry capabilities to strengthen economic productivity and promote prosperity in a sustainable way. *Innovative Technologies* refers to the technologies, systems, designs and practices that will enable further advances in synthetic biology and artificial intelligence.

Convergence in the Biofoundry

These technologies naturally come together in the biofoundry. The *design-build-test-learn* cycle employed in synthetic biology is

well suited for machine learning (Fig 2), and the use of well-characterised organisms has proven highly amenable to integrating artificial intelligence. By way of example, the yeast *Saccharomyces cerevisiae* is a standard model organism for research and industry with a nearly infinite metabolic design space and a much smaller solution space, that is the number of metabolic designs that can be realised. Quality checking whether a given human design sits within the solution space is the kind of problem artificial intelligence is very good at. Currently, biofoundry processes presume an industrial workflow in which humans still create and evaluate designs. Increasingly, however, biofoundries employ artificial intelligence at each step in the *design-build-test-learn* cycle.

.....
“In the world’s most sophisticated biofoundries, human involvement is already being extricated from repetitive workflows.”

The ultimate goal of using artificial intelligence is the automation of designs based on previous learning cycles. Synthetic biology’s design and solution space is simply too large for human comprehension and artificial intelligence could potentially narrow these spaces down to a number that can be efficiently generated and tested at a scale unattainable by human effort alone. Moreover, deploying artificial intelligence may offset the cognitive biases of human designers. Ultimately, this will have a positive feedback on assembly processes, by flagging sub-optimal processes before they are deployed. Microsoft’s *Station*

¹ARC Centre of Excellence in Synthetic Biology, Macquarie University, Sydney, NSW, Australia. E-mail: sakkie.pretorius@mq.edu.au

²Bioplatforms Australia, Macquarie University, Sydney, NSW, Australia

DOI 10.15252/embr.202050036 | EMBO Reports (2020) e50036

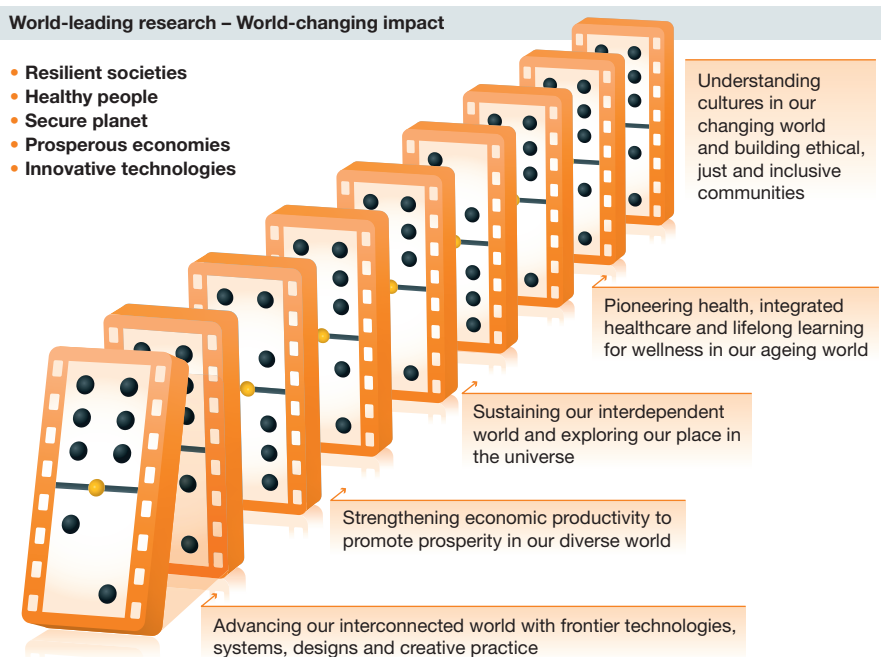


Figure 1. Today's grand challenges.

There are five main domains where the confluence of synthetic biology and artificial intelligence will potentially have the biggest impact. These domains are *Healthy People*, *Resilient Societies*, *Prosperous Economies*, a *Secure Planet* and *Innovative Technologies*.

B, an artificial intelligence enabled laboratory solution, shows what these platforms look like [1]. Codexis and Zymergen already use their proprietary software and artificial intelligence capability to design novel proteins for desired functionality. Enabling artificial intelligence to conduct multiple testing cycles with autonomous design changes between each cycle will further reduce human effort and involvement and could lead to novel design solutions.

In the world's most sophisticated biofoundries, human involvement is already being extricated from repetitive workflows. More time is spent on designing project objectives and less time on designing biofoundry processes thanks to programs such as Riffyn, Benchling and Antha by Synthace. The *human-in-the-loop* approach is transitioning to a *human-on-the-loop* approach, and in specific cases, a *human-out-of-the-loop* approach. Commercial enterprises have understood and driven this transformation for more than a decade and have generated value through leveraging proprietary data with proprietary algorithms. These commercial examples indicate what can be achieved when life is reimagined as information, and information reimagined as

life around the *design-build-test-learn* pathway. If synthetic biology is broadly defined as the engineering of biology, bio-informatics can be broadly defined as the information-managed engineering of biology. In this sense, it is different to bioinformatics because it is not just about understanding properties of organisms in the real world. Rather, it is about using information architectures to design, create and deploy novel biological devices in real-world scenarios.

Approaching grand challenges

As a platform technology and as a disruptive mode of scientific practice, synthetic biology is already addressing many economic and policy challenges from the well-known cases of synthetic artemisinin and biofuel, to transformative changes in agriculture and novel therapies. The domestication of yeast fundamentally altered society thousands of years ago, and the spinning jenny heralded the advent of industrialisation, similarly engineering biology represents a new change for agriculture and industry. Synthetic biology is the industrialisation of the life sciences at the genetic level, at the molecular level—nanotechnology that works.

When the spinning jenny was invented in 1764 by James Hargreaves, the ways in which it was used were just as important as the invention itself. *Taylorism* and *Fordism* were words to describe how the newly minted mechanical industries used these and other tools of mass production. *Lean methodology* is how we characterise these concepts today. Similarly, what differentiates the life sciences' laboratories of tomorrow is not what tools exist, it is how they are used, and what they are used for. The changes in thinking that deployed the power of steam are the same changes occurring in biofoundries today. Biofoundries are reimagining productivity in a transformative way, and they are not just seeking a marginal increase in output or a marginal decrease in unit cost. This has implications across multiple economic sectors.

In the domain of *Healthy People*, synthetic biology is making advances in the treatment of cancer, the production of vaccines and analgesics, and gene therapy. In the domain of preserving a *Secure Planet*, synthetic biology develops new crop varieties, it enables the deployment of biosensors that can monitor the environment, and synthetic biology's ongoing research on artificial photosynthesis as a source of sustainable energy will likely be transformative [2].

However, the deployment of *Innovative Technologies* can give rise to new challenges. With this in mind, synthetic biologists have worked hand-in-hand with bioethicists, philosophers, journalists and law experts since the early days of the discipline. The challenge of protecting and promoting a *Resilient Society* dictates the need to constantly engage with the public and to ensure that a robust framework of science communication maintains a two-way flow of information. The language that is used to describe synthetic biology and its applications is important; the breadth of techniques employed, coupled with a common negative connotation with the word *synthetic*, has seen the adoption of the more generic term *engineering of biology*. Indeed, the language used in the promotion of consumer products derived from synthetic biology commonly uses the term *biologically derived*. As synthetic biology continues to evolve, the language of the discipline needs to develop in an inclusive collaboration with different publics who will be affected by the translation of synthetic biotechnologies.

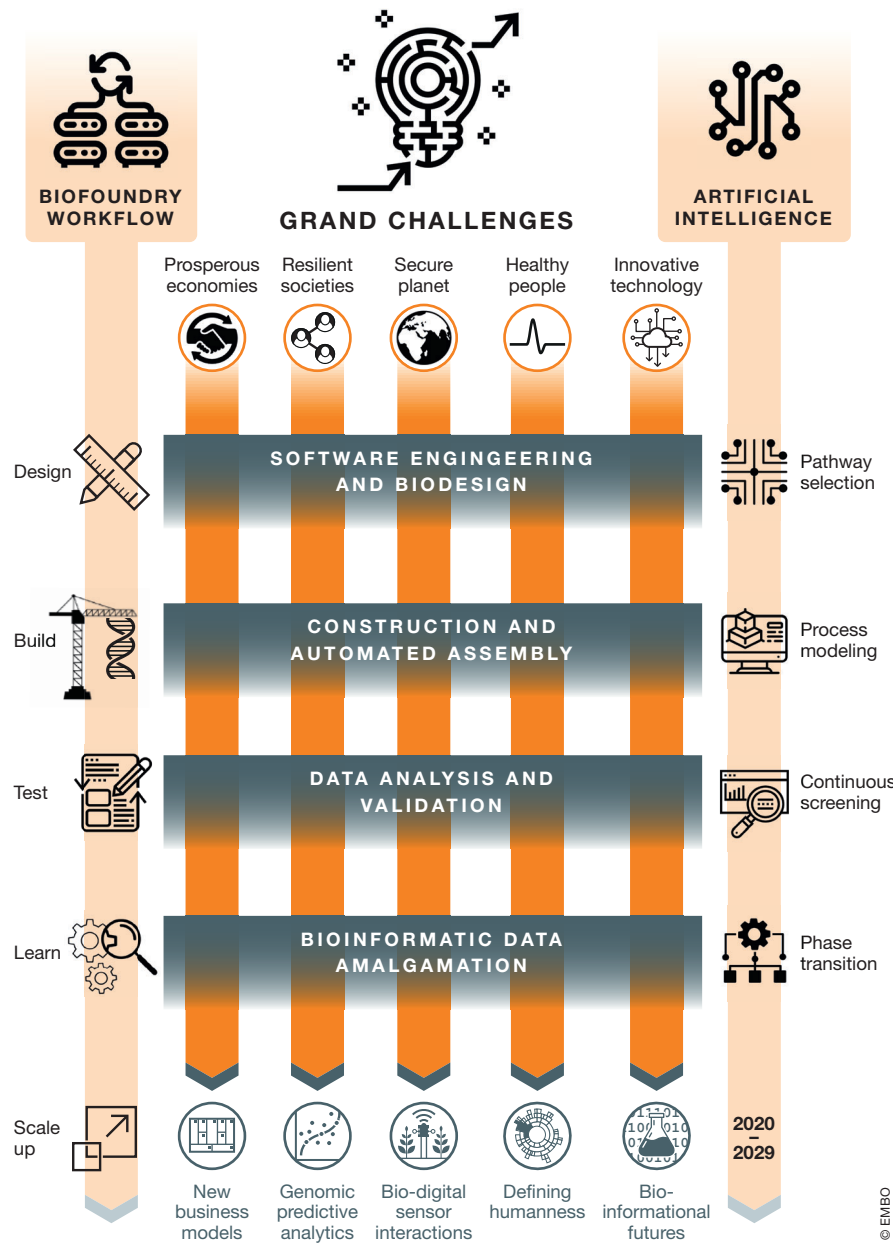


Figure 2. Convergence of synthetic biology and artificial intelligence in the biofoundry. Artificial intelligence applications enable efficiency gains in the design-build-test-learn cycle that can assist synthetic biology practitioners in meeting the grand challenges of the 21st century.

This has three parts: the language used to describe biological devices and synthetic biology processes; the language used to engage different publics in the ethics and development of synthetic biology; and the language used for engaging with stakeholder groups that are fundamentally opposed to synthetic biology. An overreliance on computing and engineering metaphors may only serve to build higher barriers between practitioners and opposition. If the language of synthetic biology can engage with and

draw from oppositional discourse, it is more likely to maintain and sustain the discipline through the decades.

Finally, all of these grand challenge domains are bound up within the conceptual framework of *Prosperous Economies*. The business models and value propositions that underpin synthetic biology could create a world of great inequity and further concentrate resources in the hands of a few. Alternatively, it could contribute to a marked increase in global equality. The instalment

of collaborative business ecosystems and developing circular economies are positive steps towards an equal distribution of economic value. The development of microbial strains that will turn waste into inputs and feedstocks is just one example of how synthetic biology enables such a circular economy. It is possible that the phase change to synthetic biology-enabled agriculture will result in the fragmentation of today’s multinational agricultural corporations that are bound to a traditional mode of centralised leadership and governance. If every owner of a small-scale subsistence farm across the globe can build a fermenter and convert their agricultural waste into high-value chemicals, it could upend traditional modes of governance and business models. However, the way in which genetic information is treated by countries, legal systems, researchers and corporations will be a major factor in deciding this trajectory.

Synthetic biology is not just enabling new ways of imagining solutions to intractable problems, it is creating new concerns for governments around the world [3]. To date, the community has been proactive in addressing these concerns. Looking ahead, there is little doubt that responsible innovation needs to form the bedrock of the new way of doing things in biofoundries.

Benefits and impacts

Commercial bio-informational platforms are enabling a phase change in the creation, curation and custodianship of genetic information. Similar to the development of artificial intelligence ethics frameworks around the world [4], nations need also be thinking about developing synthetic biology ethics frameworks. What are the solution spaces of life that humans can create within? How do we define whether or not something that could be alive should be alive?

An additional complexity is the changing state of ethics over time. Should the global, scientific or commercial community come together and define an ethics framework for synthetic biology, it needs to have an adaptive and evolutionary mechanism built in. This is no different to how companies implementing artificial intelligence need to be linked into mechanisms that forewarn shifts in societal expectations. Regular structural reviews of emerging technology ethics frameworks need to become the norm else products developed and deployed by the

corporate sector will be placed at risk. Consumer environments are increasingly based on unwritten yet dynamic ethics frameworks, and there is little doubt that the market's moral landscape will shrink and expand reacting to societal standards. Yesterday's mechanisms may not work when it comes to monitoring and regulating the real-world implementation of bio-information [5].

.....

“Biofoundries are reimagining productivity in a transformative way, they're not just seeking a marginal increase in output or a marginal decrease in unit cost.”

.....

The screening of orders for toxin or pathogen gene sequences as conducted by the International Gene Synthesis Consortium (IGSC) is one such approach. Yet, not all synthesis companies are part of the consortium; moreover, screening can only be conducted for genetic parts known to have misuse potential. A multi-layered series of security protocols are required to guarantee higher efficacy of screening. It is essential all synthesis companies participate in and actively promote such protocols, and confidentially share newly discovered dual-use parts. As confidential screening lists grow in size, their value as a cyber target increases for those seeking dual-use bio-information, or the opportunity to sell that information to interested buyers. Collaboration ensures all actors can maintain high levels of confidence in each other's activities, while open communication between international biofoundries—both commercial and research—avoids dual-use material being unintentionally sold to unvetted buyers.

Genome project-write

Genome Project-Write (GP-write) is based at the Centre of Excellence for Engineering Biology in the United States (<https://engineeringbiologycenter.org>). Formally announced on 2 June 2016, GP-write leverages two decades of work on synthetic biology and artificial gene synthesis to conduct whole-genome engineering of human cell lines and other organisms of agricultural and public health significance [6]. The capabilities that will be created, and the discoveries that will

arise from undertaking this project, will go a long way to influencing the reception of synthetic biology by publics around the globe.

The project team for GP-write is highly conscious of this, and their 2018 meeting involved presentations on the ethical, legal and social implications (ELSI) of the project. There are lawyers, ethicists, philosophers, scientists and policy and biosecurity experts involved at this early stage. Importantly, GP-write takes inspiration from the Synthetic Yeast Project and draws from their statement of ethics and governance [7]. This statement acknowledged that self-regulation is important as scientists conducting the research can continually assess and identify sources of concern. Self-regulatory steps can then mutually inform institutional and governmental oversight.

GP-write seeks to reduce the cost of engineering and testing large genomes in cell lines by more than 1,000-fold within 10 years. The use of artificial intelligence in biofoundries will be essential to achieving this goal. GP-write will perhaps do more than any other project to define the solution space of human life and in doing so some uncomfortable facts are likely to arise, not least how similar the model human organism is to much other life. In particular, GP-write will have a large impact on how synthetic biology and artificial intelligence interact with human genomic information in the biofoundry. Defining how much of the human genome's solution space is open for investigation and engineering is a societal question reliant on ethics' standards around the globe. The conduct of GP-write and its spin-off technologies may force the hand of state and international regulators to confront questions of societal, ethical and cultural significance.

.....

“Defining how much of the human genome's solution space is open for investigation and engineering is a societal question reliant on ethics' standards...”

.....

It is therefore critical that GP-write takes place in an atmosphere of international cooperation. The long-term consequences cannot be likened to other moonshot

scientific and commercial projects because GP-write will potentially demarcate the solution space (and the regulatory space) of the model human organism. There is great value in the use and reuse of human genetic data in the context of synthetic biology and artificial intelligence, including the correlation of genetic information with offline and online digital information on health, lifestyle, social status and so on. GP-write's findings may become another layer that can be used to interrogate and enhance health and lifestyle data that has already been enriched by and linked to genetic information.

.....

“At a time when it has never been more critical for international cooperation to assess and regulate emerging technologies, it is increasingly difficult to reach international consensus.”

.....

The application of insights from GP-write may represent a challenge for a world that is already experiencing competing approaches to the Internet and data, competing visions over the future of technology and heightened geopolitical tension. At a time when it has never been more critical for international cooperation to assess and regulate emerging technologies, it is increasingly difficult to reach international consensus. GP-write must navigate this world without allowing the project to descend into a competitive race that would inevitably deteriorate current practices of responsible innovation. International cooperation is essential to ensuring responsible innovation.

In the domains of *Prosperous Economies, Resilient Societies, a Secure Planet, Healthy People* and *Innovative Technologies*, the reference object is always human. If GP-write results in spin-off technologies that enable humans to live longer healthier lives, what does this mean in relation to the planet's finite resources? What does this mean for economies around the globe? Who will capture the economic benefits of these spin-off technologies, how will they be distributed, at what cost, and to whom? Knowing how to navigate these competing interests in a world of bio-information is a challenge GP-write needs to prepare for.

Global Biofoundries alliance

With all this in mind, the recently announced Global Biofoundries Alliance (GBA) [8] will be an international locus for exchanging information, setting standards and the collaborative targeting of grand projects. The Alliance will be an essential enabler for creating and codifying international norms, well before they get to a United Nations' negotiating table.

“As the bio-informational era renders itself visible to non-technical publics, self-regulation has never been more important.”

This, of course, includes material transfer agreements in both the real and the digital worlds [9] to enable biofoundry collaboration that mimics astronomy's Square Kilometre Array (SKA). In both the SKA and GBA, one can identify a decentralised infrastructure entity built from modular parts across multiple continents and enabled by high volume data transfer. The GBA working groups are the forerunners for setting and harmonising international standards in a whole range of areas, not least the teaching and training of the next generation of biofoundry operators. It is within this network of global biofoundries that the agri-industrial era evolves into the bio-informational era. They will also offer state-of-the-art readiness and response in

relation to unforeseen global pandemics or biosecurity events.

Since the 1975 Asilomar Conference, self-regulation has been an important endeavour in the life sciences. As the bio-informational era renders itself visible to non-technical publics, self-regulation has never been more important. Collaborative problem solving within the context of harmonised international ELSI approaches needs to characterise this new era. The grand challenges of *Prosperous Economies, Resilient Societies, a Secure Planet, Healthy People* and *Innovative Technologies* are without borders. Their solutions should be too. It is with this in mind that we consider the branching possibilities of our bio-informational future.

Acknowledgements

The authors gratefully acknowledge the feedback of Dr Jonathon Symons on an early draft of the manuscript. External support for Macquarie University's Synthetic Biology initiative is acknowledged from *Bioplatforms Australia*, the New South Wales (NSW) *Chief Scientist and Engineer*, and the NSW Government's *Department of Primary Industries*. Australian Government funding through its investment agency, the *Australian Research Council*, towards the Macquarie University-led *ARC Centre of Excellence for Synthetic Biology* is gratefully acknowledged. The authors also gratefully acknowledge research contributions from the synthetic biology team leaders at Macquarie University: Drs Thomas Williams, Heinrich Kroukamp, Hugh Goold, Roy Walker, Niël van Wyk and Prof Ian Paulsen.

References

- Smalley E (2019) Microsoft makes splash in AI-enabled lab solutions. *Nat Biotechnol* 37: 832–834
- Wintle B, Boehm C, Rhodes C, Molloy J, Adam L, Carlson R, Dando M, Doubleday R, Drexler E, Edwards B et al (2017) A transatlantic perspective on 20 emerging issues in biological engineering. *eLife* 6: e30247
- Dixon T (2019) Mapping the potential impact of synthetic biology on Australian foreign policy. *Australian Journal of International Affairs* 73: 270–288
- Walsh T, Levy N, Bell G, Elliott A, Maclaurin J, Mareels I, Wood F (2019) *The effective and ethical development of artificial intelligence: an opportunity to improve our wellbeing*. Melbourne, Vic.: Australian Council of Learned Academies
- National Academies of Sciences, Engineering, Medicine (2018) *Biodefense in the age of synthetic biology*. Washington, DC: National Academies Press
- Boeke JD, Church G, Hessel A, Kelley NJ, Arkin A, Cai Y, Carlson R, Chakravarti A, Cornish VW, Holt L et al (2016) The genome project-write. *Science* 353: 126–127
- Sliva A, Yang H, Boeke JD, Mathews DJH (2015) Freedom and responsibility in synthetic genomics: the synthetic yeast project. *Genetics* 200: 1021–1028
- Hillson N, Caddick M, Cai Y, Carrasco JA, Chang MW, Curach NC, Bell DJ, Le Feuvre R, Friedman DC, Fu X et al (2019) Building a global alliance of biofoundries. *Nat Commun* 10: 2040
- Kahl L, Molloy J, Patron N, Matthewman C, Haseloff J, Grewal D, Johnson R, Endy D (2018) Opening options for material transfer. *Nat Biotechnol* 36: 923