Journal and Proceedings

of the

Royal Society
of
New South Wales

2022

Volume 155 Part 1

Numbers 485 & 486

"… for the encouragement of studies and investigations in Science Art Literature and Philosophy …"

THE ROYAL SOCIETY OF NEW SOUTH WALES

OFFICE BEARERS FOR 2022

Patron	Her Excellency The Honourable Margaret Beazley AC QC Governor of New South Wales
President	Dr Susan Pond AM FRSN DSc (UQ) FRACP FTSE FAHMS FAICD
Vice-President	Ms Judith Wheeldon AM FRSN BS (Wis) MEd (Syd) FACE
Hon. Secretary(Gen)	Mr Bruce Ramage MRSN
Hon. Treasurer	Mr John Cameron MRSN
Hon. Librarian	Mr John Hardie AM FRSN FHEA FGS MACE
Hon. Webmaster	Em. Professor Lindsay Botten FRSN PhD (Tas) FOSA FAIP FAusMS
Councillors	Professor Katherine Belov AO FRSN BSc(Hons) PhD (Macq) Professor Sean Brawley FRSN Professor Philip Gale FRSN Ms Pamela Griffith FRSN Dr Davina Jackson FRSN FRGS FRSA PhD (Kent) Professor Eric Knight FRSN DPhil (Oxon) Dr Jessica Milner Davis FRSN PhD (UNSW) Em. Professor Robert Marks FRSN FRSA PhD (Stan) Em Professor Christina Slade FRSN BA PhD (ANU) Grad Dip Ed
Hunter Branch Representative	Professor George Willis FRSN
Southern Highlands Branch Rep.	Mr Mike Jonas MRSN
Western NSW Branch Rep.	Dr David Nash FRSN
Executive Office	The Association Specialists

EDITORIAL BOARD

Website: http://www.royalsoc.org.au

The Society traces its origin to the *Philosophical Society of Australasia* founded in Sydney in 1821. The Society exists for "*the encouragement of studies and investigations in Science Art Literature and Philosophy*": publishing results of scientific investigations in its *Journal and Proceedings*; conducting monthly meetings; awarding prizes and medals; and by liaising with other learned societies within Australia and internationally. Membership is open to any person whose application is acceptable to the Society. Subscriptions for the Journal are also accepted. The Society welcomes, from members and non-members, manuscripts of research and review articles in all branches of science, art, literature and philosophy for publication in the *Journal and Proceedings*.

ISSN (printed) 0035-9173
ISSN (online) 2653-1305

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 1–4. ISSN 0035-9173/22/0101-04

Editorial: Snakes, quantum computers, and global risks

Robert Marks

Economics, UNSW Sydney, Australia

Email: robert.marks@gmail.com

A recent programme on ABC Radio National's Occam's Razor introduced me to Professor Christina Zdenek and her perspective on the danger of snakes in Australia. She provocatively called her talk, "Seven reasons why Australia is the lucky country when it comes to snakes." I asked her for a revised version for the *Journal*, which is included here.

In fact, in Australia, there are five families of land snake: the pythons, file snakes, blind snakes, the rear-fanged snakes or colubrids, and the front-fanged snakes or elapids. Blind snakes (Typhlopidae) are small, shy and rarely seen; file snakes (Acrochordidae) are aquatic and live in northern swamps and billabongs; and pythons (Pythonidae) are widespread and non-venomous. Then there are the other two: elapids and colubrids.[1]

Almost all Australian venomous snakes belong to the Elapidae family, including king browns (*Pseudechis australis*), tiger snakes (*Notechis scutatus*), and the spectacularly patterned Collett's snake (*Pseudechis colletti*). Globally, elapids include the cobras of Asia and Africa, the mambas of Africa, as well as coral snakes, sea snakes, and, in Australia, over 130 species of land and sea snakes. While many elapids have evolved venom, Australia's elapids are singular in their potency. The world's most venomous land snake, the inland taipan (*Oxyuranus microlepidotus*) is capable of killing around 250,000 mice with the venom from a single bite, according to the LD_{50} parameter.[2]

Elapids are thought to have arrived in Australia many millions of years ago as a sea snake, according to evolutionary ecologist Rick Shine from the University of Sydney. "In the case of the elapids, the ancestor that has come from Asia to Australia looks to be of a modern-day krait — a sea snake species (*Bungarus* sp.)," Professor Shine said. Today, sea kraits are among the most venomous snakes in the world, so Australia's elapids likely had a head-start in their development of potent venom.

The fifth group is the Colubridae family, or rear-fanged snakes, like the non-venomous common tree snake (*Dendrelaphis punctulatus*), brown tree-snake (*Boiga irregularis*), and the non-venomous keelback (*Tropidonophis mairii*). The colubrids are the world's most successful family of snakes, and Australia is unique in being the only continent where elapids outnumber colubrids: only ten species of colubrids occur here, according to the Queensland Museum.

The colubrids are thought to have arrived in Australia after the elapids, but there is still some debate as to exactly when. And the path that they took to get here probably explains why many of Australia's colubrids are tree specialists. "The colubrids came a

---

[1] The next four paragraphs derive from Klivert (2022). See also Sleeth et al. (2021).

[2] LD_{50} is the amount of a material, given all at once, which causes the death of 50% (one half) of a group of test animals.

lot more recently down through Asia and New Guinea," said Professor Shine. What that means is that their migration path was a heavily forested one that favoured tree-dwelling species. "There's probably a connection there where you're likely to get more tree-dwelling species." That migration pathway is the likely explanation for why many of Australia's non-venomous colubrids are climbers. But the snake that most commonly bites Australians, the eastern brown snake (*Pseudonaja textilis*), is an elapid. Read Zdenek's article.[3]

In 2019 we reported (Marks 2019) the surprising first-ever image of a black hole, about 55 million light years from Earth, at the centre of the galaxy M87. Now we have an image of a much closer black hole, Sagittarius A*, at the heart of the Milky Way, only 27,000 light years away, which appears as a faint shadow surrounded by glowing material (Kruesi and Conover 2022). The image was created by planet-spanning network of radio telescopes, known as the Event Horizon Telescope.

Last year, the inaugural Warren Prize was awarded to Simon Devitt FRSN of UTS, and he has contributed a long paper based on his work with others on the architectures being explored for building quantum computers. The sensitivity of the set-up (the risk of decoherence due to thermal noise etc.) means that much thought and effort must go in to designing these platforms. We had hoped to print Professor Devitt's paper in last December's issue, but it now appears here.

Len Fisher FRSN has a continuing interest in the challenges of global catastrophic risks. In past issues we have published Fisher (2018), in which he outlined his entry in the Stockholm "New Shape" competition in developing new ideas for the governance of such risks. The paper in this issue (joint with Anders Sandberg), reprinted from *Global Policy* (2022), continues to explore the issue of governance in the face of such risk, without taking into account such outrages as the Russian invasion of Ukraine (see below).

The 2021 Forum was held on 4 November, at Government House, Sydney. It was sponsored by the Royal Society of NSW, together with Australia's Learned Academies: Health and Medicine, Humanities, Science, Social Sciences, and Technology and Engineering. The theme was "Power and Peril of the Digital Age."

The focus of the Forum was a conversation about digitalisation and the use of data. It was framed around the future life of a child born on the day of the Forum, into a world of increasingly complex digital systems that holds great value and vulnerability.

In previous years Forum participants have delivered papers, with some discussion from the floor. These papers have then been published here, in the *Journal & Proceedings*. For last year's Forum, this was changed: in each session two separate presentations were made with a common theme, followed by a discussion between the two presenters moderated by Professor Ian Oppermann, Chief Data Scientist, NSW Government. There were five sessions in the morning, each of which is printed below. But the ten presentations were not papers.

Starting with a technological framing, the Forum explored several major aspects which would impact the journey of that child as we approached 2030 and beyond: aspects of technology, health, defence and

---

[3] An earlier paper on Australian snakes is Berncastle (1866).

security in a digital age, and the changing nature of industry as the world and society evolve.

Opened by Her Excellency, the Governor of NSW, the five themes of the sessions were: (1) Science and Technology Underpinning the Digital Age: Past, Present and Future, with Cathy Foley and Hugh Durrant-Whyte; (2) the Digital Lifetime of a Child Born Today, with Frances Foster-Thorpe and Sue Bennett; (3) Avoiding a Digital Dark Age, with Shawn Ross and Theresa K. D. Anderson; (4) the Health of Our Digital Child, with Zoran Bolevich and Louisa Jorm; and (5) the Safety and Security of our Digital Child, with Dale Lambert and Rory Medcalf AM. The discussions of the second morning do not appear here.[4]

With the advent of the COVID-19 pandemic, and the consequent lock-downs and isolations, the Society moved from face-to-face meetings to virtual gatherings in 2020. This facilitated recording meetings and subsequently putting the videos online — in our case, via YouTube. In some cases, we have used these recordings to derive printed papers (see Grant 2021, Holmes 2021, and Milner Davis 2021). In other cases we have tried to derive printed papers for publishing here, but have not succeeded, for various reasons. These include presentations by Michelle Simmons,[5] Saul Griffith,[6] Jason Sharples,[7] and Angela Moles.[8]

As always, we include abstracts from recent outstanding PhD theses. We do not choose these abstracts: selection is left up to the candidates' universities. An exception to this rule, however, is the abstract by Elena Castilla, a doctoral graduate of the Complutense University of Madrid, Spain. Dr Castilla took the initiative of sending us her abstract, suitably formatted. After some investigation, we accepted it.

Geoff Harcourt AC FRSN died last December. We include an obituary of this outstanding economist and man.

Exchanges in a time of war

The Royal Society exchanges its journal with 66 learned societies around the world, including two in Russia: one in St Petersburg (the Russian Academy of Sciences) and one in Moscow (the Library for Natural Sciences). After February 24, 2022, I wondered whether we should continue the exchanges, given the Russian invasion of Ukraine. We discussed this in Council, but decided to continue the exchanges, given that Russia is an autocracy, and so the scientists of the two institutions are not directly responsible.

My thoughts: We condemn the Russian invasion of Ukraine as a barbaric, unprovoked assault on the values of a liberal democracy and the Ukrainian people which also represents a broader threat to democracy. But we will continue our century-old

---

[4] The complete Forum can be viewed online, at https://www.youtube.com/channel/UCoyHmDj2VLkgnpm-t7sIzSQ

[5] The new field of atomic electronics, address to the Annual Dinner of the Royal Society, 10 May 2019.

[6] Our energy future, two parts. See https://www.youtube.com/watch?v=sQthtORLaFg&list=PLYFFwCGj2FIaOm9l-b_oreihzfMYI6HnS&index=7 and https://www.youtube.com/watch?v=LopP_O8_dgE&list=PLYFFwCGj2FIaOm9l-b_oreihzfMYI6HnS&index=4

[7] Extreme bushfires and the age of violent pyroconvection. See https://www.youtube.com/watch?v=cTRXkM_z-S8&list=PLYFFwCGj2FIaOm9l-b_oreihzfMYI6HnS&index=10. Also see O'Connor (2021).

[8] The RSNSW Poggendorff Lecture 2020: Are our weeds becoming new native species? See https://www.youtube.com/watch?v=dMy7VpEzEbs

exchange of information with Russian intellectuals as an important means to support freedom of thought in Russia and give encouragement to those committed to a free, open society.

In fact, when I tried to post copies of the December issue to Moscow and St Petersburg, I found that the Australian Post Office was not then accepting mail for Russia.

Important note for member subscribers

Earlier the Society moved to a new database that members and fellows use when renewing their memberships. Previously, if you were a subscriber to the paper copy of the *Journal*, this subscription and its cost were easily rolled over to the new year. For whatever reason, the new system made rolling-over your subscription difficult. As a result the numbers of member subscribers fell from 212 in 2021 to 26 in 2022. After efforts on our parts, the numbers of member subscribers has risen to over 60 — still many fewer than previously.

I hope that no-one misses out on a paper copy of this issue: if you discover too late that you no longer subscribe to the *Journal*, it is too late for this issue: we will not print extra copies. Subscribe now at the Society's Online Shop[9] for the December issue.

Housekeeping

I'd like to thank Davina Jackson, Lindsay Botten, Rory McGuire and Jason Antony for their thorough help in producing the *Journal*.

June 8, 2022.

References

Berncastle, J. (1866) On snake-bites and their antidotes, *Trans. of the Philosophical Society of N.S.W. 1862–1865* 1: 191–196.

Fisher, L. (2018) Global challenges: personal reflections on the Stockholm "New Shape" competition, *Journal & Proceedings of the Royal Society of New South Wales* 151: 213–231.

Fisher, L. & Sandberg, A. (2022) A safe governance space for humanity: necessary condions for the governance of global catastrophic risks. *Global Policy*, 00, e13030. https://doi.org/10.1111/1758-5899.13030.

Grant, S. (2021) With the falling of the dusk. *Journal & Proceedings of the Royal Society of New South Wales* 154: 211–223.

Holmes, E.C. (2021) The discovery, origins and evolution of SARS-CoV-2 (COVID-19). *Journal & Proceedings of the Royal Society of New South Wales* 154: 161–181.

Klivert, N. (2022) Think most venomous snakes don't climb? Think again. ABC 5 June 2022. https://www.abc.net.au/news/science/2022-06-05/venomous-snakes-can-climb-trees-disproving-myth/100861486

Kruesi, L. and Conover, E. (2022) We finally have an image of the black hole at the heart of the Milky Way, *Science News*, May 12.

Marks, R.E. (2019) Editorial: "The Old One does not play at dice," *Journal & Proceedings of the Royal Society of New South Wales* 152: 157–159.

Milner Davis, J. (2021) Taking humour and laughter seriously: The multi-disciplinary field of humour studies. *Journal & Proceedings of the Royal Society of New South Wales* 154: 182–200.

O'Connor, M. (2021) What it's like to fight a megafire, *The New Yorker*, 8 November.

Sleeth, M., Eipper, S., and Madani, G. (2021) Opportunistic observations of climbing behaviour and arboreality in Australian terrestrial elapid snakes (Elapidae: Hydrophiinae) *Herpetology Notes* 14: 1407–1415.

---

[9] https://members.royalsoc.org.au/rsnsw-shop/

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 5–39. ISSN 0035-9173/22/0105-35

Blueprinting quantum computing systems

Simon J. Devitt

Centre for Quantum Software and Information, University of Technology Sydney, Australia

Email: simon.devitt@uts.edu.au

Abstract

The development of quantum computing systems has been a staple of academic research since the mid-1990s when the first proposal for physical platforms were proposed using Nuclear Magnetic Resonance and Ion-Trap hardware. These first proposals were very basic, essentially consisting of identifying a physical qubit (two-level quantum system) that could be isolated and controlled to achieve universal quantum computation. Over the past thirty years, the nature of quantum architecture design has changed significantly and the scale of investment, groups and companies involved in building quantum computers has increased exponentially. Architectural design for quantum computers examines systems at scale: fully error-corrected machines, potentially consisting of millions if not billions of physical qubits. These designs increasingly act as blueprints for academic groups and companies and are becoming increasingly more detailed, taking into account both the nature and operation of the physical qubits themselves and also peripheral environmental and control infrastructure that is required for each physical system. In this paper, several architectural structures that I have worked on will be reviewed, each of which has been adopted by either a national quantum computing program or a quantum startup. This paper was written in the context of an award with the Royal Society of New South Wales, focused on my personal contributions and impact to quantum computing development, and should be read with that in mind.[1]

Introduction

The development of the second generation of quantum technology such as quantum computing platforms, quantum communication systems, quantum simulators and sensors is anticipated to create a new technological revolution, similar to the digital computing revolution of the 20th century. While significant research and development in quantum technology began in the middle of the 1990s, largely in the academic space, since 2014 we have seen an explosion in investment and consequently an explosion of progress.

Governments, multi-national corporations and the equity investment community have recognised the potential impact of quantum technology and have invested accordingly. A €1 billion European flagship (Riedel et al. 2019), the US$1.2 billion quantum initiative (Raymer and Monroe 2019), over US$10 billion in declared investment from the Chinese government (Zhang et al. 2019), a €2 billion quantum investment from the Germans as part of their COVID-19 recovery package, and others (Brennan et al. 2021), have cemented quantum technology as a major global development goal in the 21st century. Corporate investment in building, particularly, quantum computing

---

[1] In 2021, Simon Devitt was awarded the inaugural Warren Prize for research by engineers and technologists in their early to mid-careers, by the Royal Society of New South Wales.

systems is also extensive: Google, Microsoft, Amazon, IBM, Baidu, TenCent and Alibaba are just some of the major global technology firms that have established extremely strong research groups to build and deploy quantum computing systems. IBM and Amazon have already launched fee-based cloud access to several quantum computing platforms, such as Ion-Traps and Superconducting systems. Quantum startups are now ubiquitous worldwide, both in the hardware and software space, with the three largest quantum computing startups — IonQ[2], Rigetti[3] and PsiQuantum already valued over one billion dollars.[4]

The construction of a large-scale quantum computer is now a serious goal amongst national programs, multinational corporations and the equity funding community and there is a variety of physical platforms that people are developing. Which platform or platforms will end up being viable for scientifically or commercially useful computational tasks is still up for debate, but the diversity of approaches will undoubtably spur more rapid development.

Over the past 15 years, I have been involved in several results related to quantum architecture development and design, with numerous exceptional theoretical and experimental colleagues. We introduced the one of the earliest large-scale quantum computing architectures that incorporated quantum error correction, was modular in nature, and could be conceptually scaled to an arbitrary degree (Devitt et al. 2009) and have worked on architecture designs for multiple different physical systems (Oi, Devitt, and Hollenberg 2006; Stephens et al. 2008; Nemoto et al. 2014; Lekitsch et al. 2017; Mukai et al. 2020). We introduced a design for a high-performance quantum computing system that could perform distributed blind quantum computing in an error-corrected environment (Devitt, Munro, and Nemoto 2011) and designed architectures for quantum communication networks that could serve to connect distributed quantum computing systems together (Munro et al. 2010, 2012; Devitt et al. 2016).

Aside from architecture development, I have also developed parts of theoretical frameworks for how to program, implement and resource-optimise fully error-corrected quantum algorithms (Devitt, Munro, and Nemoto 2013). This includes examining the practical requirements of classically processing error-correction information from a quantum computer (Devitt et al. 2010), how to compile high-level quantum circuits into error-corrected compatible forms (Fowler and Devitt 2012; Herr et al. 2018; Herr, Nori, and Devitt 2017; Paler et al. 2014; Devitt 2016; Horsman et al. 2012), and what the formal requirements are for benchmarking quantum algorithms on practical machines (Devitt et al. 2013; Meter and Devitt 2016; Paler, Herr, and Devitt 2019).

In this paper, I specifically examine three of the scalable quantum computing blueprints that I have been involved in that have been adopted by quantum startup companies and national programs worldwide. I look specifically at three designs: one in Ion-Traps that has been adopted by the UK startup Universal Quantum (Lekitsch et al.

---

[2] https://finance.yahoo.com/quote/IONQ/

[3] https://www.rigetti.com/merger-announcement

[4] https://www.wsj.com/articles/psiquantum-raises-450-million-to-build-its-quantum-computer-11627387321

2017), one using Nitrogen Vacancies (NV) in diamond, adopted by the US-based startup Turing inc and the Austrian start-up Godel GmbH (Nemoto et al. 2014), and one in superconductors that has been adopted by the Japanese national program, Q-Leap and Moonshot (Mukai et al. 2020; Kwon et al. 2020).

The basics of quantum computing

The core operational element of a quantum computer is the quantum bit (qubit): this is a well-defined two-level quantum system that can exist in a variety of physical platforms. One of these two levels corresponds to a binary 0 state and the other to a binary 1 state (Nielsen and Chuang 2000).

The physical systems used to define these two states can be anything from the electronic levels of an ionised atom (Cirac and Zoller 1995), the polarisation state of a single photon (O'Brien 2007) or the spin state of a phosphorus atom in a silicon crystal (Kane 1998)[5].

A qubit lives in a two-dimensional complex vector space, where the general state of a qubit is given by equation 1:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle\,, \qquad (1)$$
$$|\alpha|^2 + |\beta|^2 = 1, \quad \{\alpha, \beta\} \in \mathcal{C}$$

Gate operations on qubits are defined by unitary operations on this vector space. As the modulus squared of all amplitudes sums to one — as they represent the probabilities that qubits are measured in one of the two possible basis states — unitarity is required to ensure that probabilities are conserved as quantum gates are performed.

When considering an array of qubits in a quantum computer, the total size of the vector space grows exponentially. For an N-qubit quantum computer, the complete state of the system can be described by a complex column vector of size, $2N$:

$$|\psi\rangle = \begin{pmatrix} z_0 \\ z_1 \\ z_2 \\ . \\ . \\ . \\ z_{2^N-1} \end{pmatrix}, \qquad \sum_{i=0}^{2^N-1} |z_i|^2 = 1. \qquad (2)$$

Gates are then defined as unitary matrix operations, G, of size $2N \times 2N$, where $G\dagger G = I$, and $\dagger$ is the conjugate transpose of G.

This is effectively what a quantum computer is. It is a matrix multiplier over a complex column vector of size $2N$. The inability for classical computers to simulate a quantum computer is due to the exponential scaling of this column vector. As each element in the matrix is a complex number, and each complex number requires two real numbers (doubles), the memory required to completely store the state of a quantum computer containing N qubits is (in Bytes) $2 \times 8 \times 2^N$.

Even for small qubit arrays, this scaling overtakes the memory capacity of any classical system. For example, the Google Sycamore chip contains $N = 53$ functional qubits (Arute et al. 2019). To completely store the state of Google's Sycamore chip would require 144 PetaBytes. This is essentially the argument being used in a recent IBM paper claiming a method to simulate Google's quantum supremacy has resulted in the Summit supercomputer (Pednault et al. 2019). The argument is to re-task hard-disk

---

[5] This paper inspired Michelle Simmons FRS FRSN. [Ed.]

space as virtual RAM to reach the capacity needed to store the entire complex vector representing the Sycamore quantum processor.

There are multiple techniques that can be used to approximate the behaviour of a quantum computer in classical systems and there are efficient classical algorithms for exact simulation of restricted classes of gate operations (Aaronson and Gottesman 2004; Markov and Shi 2008; Perez-Garcia et al. 2007). However, we know that the efficient simulation of a universal set of quantum gates is not possible unless fundamental conjectures of complexity theory are proven to be false (Aaronson 2013).

Qubits naturally existing in this complex vector space implies that a quantum computer only requires *N* qubits to perform the same computation, and a variety of results since the early 1990s have demonstrated that the complex vector space of quantum logic is more powerful than binary logic for a specific set of problems. This includes direct simulation of quantum systems (Lloyd 1996), solving the hidden abelian subgroup problem (Kitaev 1995) — which includes factoring large integers (Shor 1994) — solving numerous optimisation and graph problems (Lee, Santha, and Zhang 2020), and solving large sets of linear equations (Harrow, Hassidim, and Lloyd 2009).

Quantum computing is now growing into a larger and larger industry, with numerous multinational technology companies constructing and making available small quantum processors over the cloud. National initiatives being established in most major industrial countries (Roberson and White 2019; Riedel et al. 2019; Sussman et al. 2019; Zhang et al. 2019; Yamamoto, Sasaki, and Takesue 2019; Fedorov et al. 2019; Raymer and Monroe 2019; Knight and Walmsley 2019) and investment from venture capital firms into both quantum hardware and software companies now eclipsing the $1 billion scale.

Part of the development of quantum computers is asking how we blueprint large-scale quantum computing systems — to the level needed to implement the wide variety of quantum algorithms that are provably more efficient than their classical counterparts. These blueprints include the hardware design of the qubits and qubit control systems, the manner in which Quantum Error Correction (QEC) is embedded within these designs, how environmental infrastructure is deployed for these systems, and how the classical computing support structure is integrated to the quantum hardware.

The core elements of an architecture: the DiVincenzo criteria

David DiVincenzo was one of the first to enumerate a minimal list of physical requirements that are needed to build a large-scale quantum computer. In 2000 he published what are now referred to as the DiVincenzo criteria (DiVincenzo 2000): a set of five elements a quantum system must have to be, in principle, suitable for constructing a quantum computer. These are:

1. A scalable physical system with well-characterised qubits
2. The ability to initialise the state of the qubits to a simple fiducial state
3. Long relevant decoherence times
4. A “universal” set of quantum gates
5. A qubit-specific measurement capability.

These five conditions are a necessary minimal set that is required for a physical platform to be appropriate for a quantum computer, but they are by no means sufficient when building a scalable system.

From the middle of the 1990s to approximately the middle of the 2000s, there were literally dozens of physical quantum systems proposed that satisfied the DiVincenzo criteria. Many of these "architecture" proposals have slowly disappeared as researchers have realised that the details of a practical quantum architecture go well beyond these five elements.

Of the many different physical systems proposed for quantum computing, there remain eight that are under major development:

1. Superconducting qubits (Kwon et al. 2020)
2. Ion Traps (Brown, Kim, and Monroe 2016; Lekitsch et al. 2017)
3. Optical qubits (single photon and continuous variables) (O'Brien 2007; Braunstein and Loock 2005)
4. Colour centres (such as Nitrogen Vacancy centres in diamond) (Nemoto et al. 2014)
5. Quantum Dots (Jones et al. 2012)
6. Donors in Silicon (Kane 1998)
7. Neutral Atoms (Saffman 2018)
8. Anyonic Systems (Nayak et al. 2008).

A set of criteria that I generally use to define a "major" quantum computing system are:

1. There is significant funding available for the platform.
2. The platform has already demonstrated the fabrication and control of a small number of qubits (1–10) to the point where it is now somewhat routine.
3. There are experimental and/or theoretical researchers involved in a systems development that are "true believers," i.e. they are focused strongly on actually building a scalable quantum computer, rather than simply doing interesting and more foundational physics work.

Each of the eight systems listed above satisfy each of these criteria, except arguably for Anyonic systems, but the vast amount of money invested by Microsoft into this highly experimental platform necessitates its inclusion on the above list. Investment into these platforms is somewhat evenly distributed across corporations, governments, universities and startups.

Modern quantum architecture designs

Beyond the DiVincenzo criteria, architecture development and blueprinting quantum computing systems have evolved significantly over the past 10–15 years. As we are entering the era of engineering small qubit chipsets, designs for larger-scale systems are becoming more complex. Early quantum computing blueprints generally consisted of little more than identifying an appropriate two-level quantum system and describing the interaction dynamics that enabled a universal set of quantum gates, initialisation and measurement.

As we have further understood the actual necessities of a large-scale quantum computer, design blueprints across a variety of different systems have become more detailed and more sophisticated. The most notable change is the detailed introduction of Quantum Error Correction (QEC) protocols.

Quantum error correction

Fabricating and controlling physical qubits is a difficult thing to do. The coupling of individual qubits to the environment inevitably leads to loss of quantum coherence (known as decoherence). From an information-processing standpoint, this decoher-

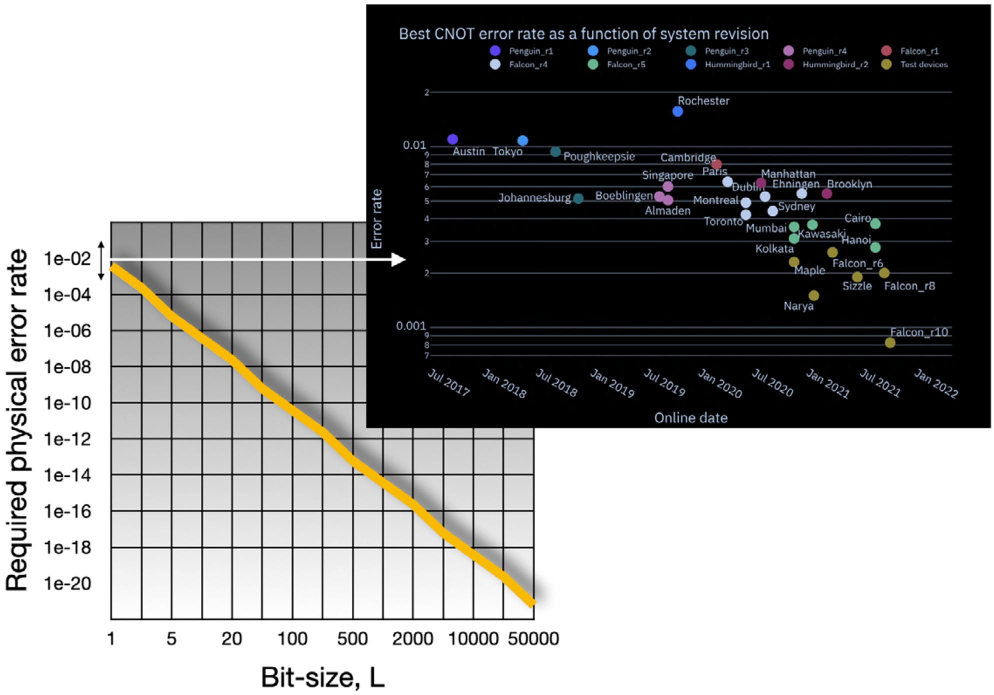


Figure 1: Plot of the required error physical qubit error rates needed to implement Shor's algorithm for various bit sizes, L, using the construction of Gidney and Ekera (2019). The insert illustrates experimental error rates over IBM's deployed superconducting systems between 2017 and 2021 (superconducting qubits were first demonstrated in the late 1990s, with error rates on the order of 5–10%). Experimental error rates only live in the upper left-hand corner of required error rates for Shor's algorithm.

ence effect introduces errors, which quickly renders the output from any quantum algorithm essentially random. Errors can come from inaccurate fabrication, bad qubit design — where the environment couples too strongly to the two-level system that defines the qubit — or could be induced by imperfect control.

While experimentalists have done an impressive job at decreasing qubit error rates in physical systems over the past two decades, physical engineering alone will not be sufficient to reduce error rates to a degree necessary for large-scale quantum algorithms to be run. Shown in Figure 1 is a plot of the physical error rates needed for qubits if you were to implement Shor's factoring algorithm.

For Shor's algorithm, we wish to be able to factor numbers that have a bit-size length, L, of $L > 1024$, as this is approximately the minimum key size used in modern implementations of RSA public key encryption. From Figure 1, you can see that this would require physical error rates on the qubits of at most $O(10^{-14})$.

Shown in the insert of Figure 1 is a plot of the physical error rates achieved in the laboratory since qubits were first fabricated and tested in the late 1990s. While significant progress has been made, the error rates achieved in the lab insert lie in the very top left-hand corner of the larger plot in terms of the physical error rates that have been demonstrated, and that experimentalists have reduced these errors from approximately 3–5% to approximately 0.1% in 20 years. This does not prove that a new, revolutionary method for qubit fabrication and control will not be made that allows physical errors to drop by a further 10 or 11 orders of magnitude, but it does suggest that physical systems are going to need some help when it comes to reducing error rates to the level that is needed to implement large-scale quantum computing.

Quantum Error Correction (Devitt, Munro, and Nemoto 2013) provides this framework, by taking high error rate, physical qubits and encoding them into a code block to form a logical qubit. This logical qubit then contains sufficient redundancy so that errors on the physical qubits can be detected and corrected without destroying or unintentionally modifying any information within the encoded block.

While there are a plethora of QEC codes to choose from, in terms of large-scale quantum architecture design there has been unarguably a preferred technique, known as the surface code (Fowler et al. 2012). The surface code encodes a single logical qubit into a two-dimensional array of physical qubits. Illustrated in Figure 2 is an encoded qubit using a distance $d = 5$ code.

Code distance is a measure that counts the minimal number of physical errors that is required to create a logical error. i.e. if we consider a logically encoded $|0\rangle L$ state, how many physical bit-flips (X-gates) are required in order to take $|0\rangle_L \leftrightarrow |1\rangle_L$? For a distance $d = 5$ code, we require five physical X-gates to induce a logical bit-flip.

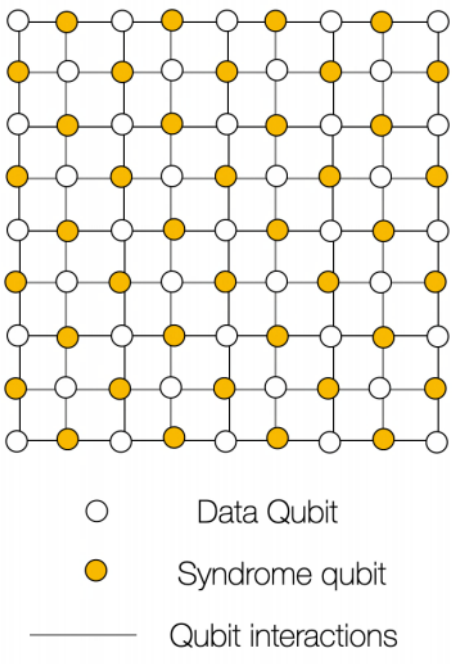


Figure 2: A 2D lattice of qubits encoding a distance surface code logical qubit. A distance code requires a lattice of $(2d - 1)^2 = 9 \times 9 = 81$ qubits and is able to correct for up to two arbitrary errors on any one of the physical qubits. Approximately half of the qubits are actually part of the encoded block (coloured white) and the others are Syndrome qubits (yellow) which are used to extract error information and are repeatedly measured. The solid likes represent the required interactions between qubits. All qubits must be interacted with their four nearest neighbours.

In quantum information we need to protect against both bit-flip gates (X-gates) and phase-flip gates (Z-gates) which can cause errors in quantum superposition, i.e.:

$$(|0\rangle + |1\rangle)/\sqrt{2} \quad \leftrightarrow \quad (|0\rangle - |1\rangle)/\sqrt{2} \, .$$

In the most general case, it is assumed that the probabilities of physical bit-flips and phase flips are identical and hence we need identical levels of protection for these two types of errors in the logical qubit. In the surface code, this is achieved by ensuring that the surface code is a square lattice.

We will not summarise the specific details of how error-correction works, there are sufficient reviews in the past 20 years to cover these topics (Devitt, Munro, and Nemoto 2013; Fowler et al. 2012). However, we will briefly discuss how the encoded error rates (logical error rates) scale.

Through direct numerical simulation of the error correction protocols on the surface code (Devitt et al. 2016), it is possible to relate the distance of the code, *d*, the physical error rate associated with each qubit in the code, p, and the failure probability of the logically encoded information, P_L, as equation 3:

$$P_L \approx C_1(C_2 p)^{\frac{d+1}{2}}. \quad (3)$$

These simulations track the effect of Pauli bit and phase errors on the quantum circuits used to implement the circuit code, apply error-decoding protocols and then estimate whether logical Pauli errors are formed. By Monte Carlo simulations of many instances, it is possible to estimate the logical failure rate of the encoded qubit, P_L.

The first thing to notice is that as $d \to \infty$, the logical error rate $P_L \to 0$ if and only if $C_2 p < 1$. If $C_2 p > 1$, the logical error rate gets worse as a function of code distance. This is what is known as the code threshold, i.e. the largest physical error rate that the code can still provide correction. If physical error rates are larger than the threshold, errors are introduced faster than they can be extracted by the error correction and logical error rates get worse instead of better.

The fault-tolerant threshold for an actual architectural model is heavily dependent on the type of error-correction code used, the type of errors that are anticipated at the physical *later* in a specific hardware system, and how a hardware model physically realises the gate operations needed to realise the chosen code. The actual threshold for the surface code is more explicitly modelled in numerous papers and has been found to be approximately (Fowler et al. 2012) under a balanced Pauli noise model, assuming an underlying architecture based on a 2D nearest-neighbour array of interacting qubits.

For the surface code, the code distance is related to the number of physical qubits in the lattice, *N*. Specifically a code of distance d requires a $N = (2d - 1)^2$ qubit square lattice. Hence we can rewrite eq. 3 as:

$$P_L \approx C_1(C_2 p)^{-\frac{\sqrt{N}+3}{4}}. \quad (4)$$

Consequently, if p is under the code threshold, the logical error rate P_L will decrease, exponentially, with the $\sqrt{N}$, which is the number of physical qubits along the edge of the 2D lattice that defines a surface code logical qubit.

This is the mechanism that allow us to reduce the logical error rate of an error-corrected qubit without having to find ways to reduce the physical error rates of constituent physical qubits. Instead of performing Shor's algorithm on physical qubits which have been engineered to have error rates on the order of $p \approx 10^{-14}$ to factor numbers larger than 1024 bits, we instead create sufficiently large encoded qubits such that Eq. 4 achieves a logical error rate that is small enough. This does not require us to change p (provided it is already below the threshold), but rather it requires us increase *N*, i.e. build more physical qubits.

Hence, large-scale quantum computation is largely an exercise in building an error-correction machine. As one of the pioneers of error-correction, Andrew Steane, once reportedly quipped:

> A quantum computer is an error-correction machine, computation is just a by-product.

This is reflected in large-scale design blueprints. The majority of the design and analysis is focused around implementing error-correction for a large number of qubits in the most effective manner possible.

It should be heavily stressed that error-correction adds a very large overhead for any quantum algorithm. This is not only due to the direct overhead of encoding logical qubits with a collection of physical qubits, but is also due to ancillary protocols needed to maintain fault-tolerance for encoded operations.

This consequently means that many algorithms that provide a clear quantum advantage over classical computers require millions if not hundreds of millions of qubits. Most of the most accurate benchmarking of quantum algorithms have occurred for factoring (Gidney and Ekera 2019; Devitt et al. 2013), quantum simulation (Reiher et al. 2017; Babbush et al. 2018) or applications using quantum random-access memories (Matteo, Gheorghiu, and Mosca 2020).

Infrastructure and control

Another major component that has become a staple of blueprinting quantum computing systems is related to the classical control systems and environmental infrastructure needed to operate specific qubit systems.

As quantum information is so fragile, and environmental decoherence can couple into a system so easily, it is often required that physical qubits need to be housed in infrastructure that carefully controls the environment in which the qubit lives. This can take multiple forms, depending on the specific system being designed.

In some systems, qubits need to be kept at incredibly low-temperature environments. This is particularly true for superconducting qubits or qubits built from phosphorus donors in silicon or qubits built from artificial atoms (quantum dots). These qubits are often characterised by energy separations that occur at microwave frequencies. At these frequencies, thermal photons from the environment can be of the same order as the energy of the qubit. Consequently they can induce qubit dynamics that are otherwise unwanted. i.e. noise and errors.

As many of these qubits are designed to operate at these energy separations, the only other choice is to make the environment as cold as possible as to suppress the production of these thermal photons, which is exactly what is done.

Dilution refrigeration systems are now commonplace around the world, which allows for the reduction of temperature in a sample chamber to the milli-Kelvin range. The issue with dilution refrigeration cooling is that it significantly constrains the number of qubits that can physically be placed inside such a refrigerator. It also constrains the amount of energy that can be routed into a refrigerator to perform qubit control, initialisation and measurement without overwhelming the ability of the refrigerator to keep the system cold. This requires significant thought as to how to overcome what are essentially engineering roadblocks to qubit systems that may contain millions or billions of physical qubits.

In other systems, such as Ion-Traps, qubits need to be placed into an essentially perfect vacuum. An ion-trap is an electromagnetic trap that can be used to confine a single charged atom. Clearly, such a trap becomes ineffective in an atmosphere, where collisions between atmospheric atoms and the trapped ion can easily eject the ion from the electromagnetic trapping potential. Consequently, Ion-Trap quantum computers are routinely pumped down to a pressure of approximately 10^{-11} Torr or better. This requires an initial pumping down of the chamber, a cooking stage above 200K to cause additional atoms stuck within the walls and other surfaces to evaporate into the chamber, and then a second pumping stage. Additionally, it is expected that for a scalable system, the vacuum system will need to be cooled to 4K to further reduce the rate at which ions are heated (Niedermayr 2015).

As with superconductors, environmental infrastructure can make scalability difficult. Having large vacuum systems that need to house millions or hundreds of millions of physical ions for long-term operation presents a difficult engineering challenge. Although, unlike superconductors, Ion-Traps have the ability to be connected with photons. This allows fibre optic coupling of independent ion-traps (Monroe et al. 2012), something that is extremely difficult to do in superconducting systems (Magnard et al. 2020).

Photonic computers — where qubits are individual photons — also have their own specific challenges. As photons travel at the speed of light, it is extremely hard to build computational chips, as you cannot spatially localise photons to do actual gate operations. While there are techniques that are used to overcome the problem that individual quantum bits are flying around your computer at 30 cm per nanosecond (Rudolph 2017), having physical qubits moving around this quickly creates significant challenges for scalability for these particular models.

Other physical platforms for quantum computing also have associated infrastructure and control issues, each of which has to be handled in its own way. Good quantum architecture designs and blueprints contain significant details with respect to achieving scalability in infrastructure and control, not just in terms of being able to fabricate a large number of individual qubits.

Fabrication and cost

The digital computing industry has unarguably been dominated by issues surrounding fabrication and cost. One of the most foundational principles in classical microprocessor development was Moore's law, an empirical prediction made in 1965 by the co-founder of Intel, Gordon Moore, who observed that the number of components on a classical microprocessor was doubling every 18 to 24 months.

This observation was not only related to the ability of fabricating more components onto an integrated circuit, but also the cost in doing so. Feature sizes and hence transistor sizes decreased exponentially at the same time as costs per transistor dropped exponentially. This is the primary reason behind the ubiquitous nature of information-processing technology worldwide.

While there have always been peripheral discussions with certain platforms about the ease and cost associated with the production of a quantum computer, these discussions have become much more serious as the technology moves out of the academic laboratories and into commercial production.

Certainly the most dominant issue that motivates discussions surrounding mass manufacturing and costs is current estimates surrounding the number of physical qubits needed to implement large-scale algorithms such as Shor's algorithm or problems in quantum chemistry. In Gidney and Ekera (2019) and Reiher et al. (2017) we see some of these estimates, which attempt to incorporate and optimise the full integration of error-correction and fault-tolerant protocols — given these resource estimates, the cost of a quantum computer to implement these algorithms can be estimated assuming various order-of-magnitude estimates for the Price Per Qubit (PPQ). These algorithms require a very large number of physical qubits, and hence cost becomes a big factor.

Table 1: Cost of various machines for well known quantum algorithms as a function of PPQ.

PPQ	Factoring (Gidney and Ekera 2019)	Nitrogenase (Reiher et al. 2017)
$1000	$20 Billion	$200 Billion
$1.00	$20 Million	$200 Million
$0.01	$200,000	$1 Million

Qubit costs can be parameterised as the PPQ, where you average over the total cost of the machine, including classical control systems, environmental infrastructure etc, as a function of the number of qubits in this system. For example, a superconducting system housing 50 actual qubits would require a $500K dilution refrigeration system and an additional $500K in microwave signal generators, niobium wiring, chip fabrication costs and classical computer control, and would have a PPQ of $20,000. At this scale, a quantum computer capable of factoring (Gidney and Ekera 2019) or the simulation of complex molecules such as nitrogenase (Reiher et al. 2017) would carry a price tag higher than the GDP of Austria ($417 billion) and Japan ($4.8 trillion). Table I illustrates the cost for factoring and quantum chemistry for various orders of magnitude for a PPQ. Even at an effective $1 PPQ, a quantum computer of sufficient size for factoring or quantum simulation would be a significant investment.

The entire quantum technology industry is founded on the precept that qubits will eventually be cheap and quantum technology will be at least as ubiquitous as large computational servers, if not as ubiquitous as mobile phones and the internet of things. To achieve this, PPQs must be reduced by at least five orders of magnitude compared to the state of the art today, and likely much much lower.

Three hardware architectures

In this section I will summarise some of the architectures that I have been involved with (Nemoto et al. 2014; Lekitsch et al. 2017; Mukai et al. 2020; Kwon et al. 2020) and the development of the key theoretical results that allow for error-corrected implementations of quantum algorithms on these, and other, hardware (Devitt et al. 2010; Fowler and Devitt 2012; Fowler, Devitt, and Jones 2013; Herr, Nori, and Devitt 2017; Horsman et al. 2012; Paler et al. 2014, 2012; Paler, Devitt, and Fowler 2016; Devitt 2016; Herr et al. 2018; Paler and Devitt 2018).

Ion-traps

Many groups have proposed large-scale ion trap quantum computers (Cirac and Zoller 2000; Kielpinski, Monroe, and Wineland 2002; Schaetz et al. 2004; Duan et al. 2004; Metodi et al. 2005; Steane 2007; Stock and James 2009; Kim and Kim 2009; Crick et al.

2010; Amini et al. 2010; Monroe et al. 2012).[6] Two themes dominate — multi-zone micro-fabricated traps, with each zone containing a handful of ions, and optical linking of traps. Optical linking is probabilistic, with a success probability of 2.2×10^{-4} the best achieved to date (Stephenson et al. 2020).

While this is expected to improve, the current state-of-the-art is far from sufficient for a practical computer (Nickerson, Li, and Benjamin 2013). Even if optical linking is improved by several orders of magnitude, this approach, while technically scalable, will result in a computer much slower than is theoretically possible. A direct interaction between ions separated by microns will generally be far faster than a multi-metre photon mediated interaction. Existing multi-zone micro-fabricated trap designs make use of control electrodes in the same plane as the trap electrodes, precluding scaling to an arbitrarily large 2-D lattice of qubits as the required number of control electrodes per unit length around the edge of the chip grows without bound. An alternative approach is required to achieve high performance and scalability.

Gven ions can be reliably transported along linear traps and through X-junctions (Walther et al. 2012; Wright et al. 2012), the core of our solution to the scaling problem is to abandon the optical linking of chips and instead align chips each containing a moderate number of X-junctions so that ions can be transported from chip to chip as though along an unbroken trap. Each X-junction under normal circumstances contains just two ions, only one of which is used as a qubit. The second ion would be used for sympathetic cooling (Rohde et al. 2001). A surface-electrode single X-junction is shown in Figure 3. The design incorporates electrodes to rotate the trap axes, enabling laser cooling of ions at any location (Stenholm 1986). This is necessary to enable characterisation of the junctions and adjustment of electrode control pulses at time of construction.

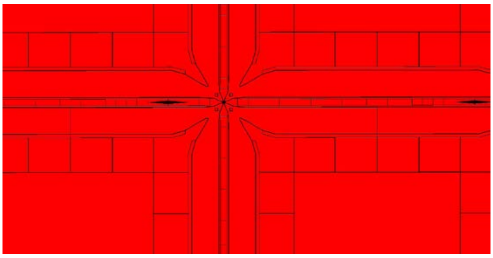


Figure 3: Schematic 5×5mm surface-electrode X-junction produced by the Sussex group in our initial design. Black diamonds represent holes above which ions are interacted and manipulation. Three optical fibres are cemented into each hole, two of which deliver laser light and the third of which is used to gather photons for measurement. A total of 52 DC connections are required per junction.

Each junction would occupy a 5 × 5 mm patch of a larger wafer and be associated with a single interaction and manipulation region. 16 of these junction patches would be bundled together to create a repeating 20 × 20 mm section (Figure 4). 25 of these sections would be fabricated on a 100 × 100 mm chip.

In addition to two Rf Voltage connections, each section also requires 48 fibre connections (3 connections per junction zone, two for the entanglement lasers one for detection) and 840 DC connections (52 connections per junction and 8 additional per section) for operation. An in-vacuum

---

[6] The work in designing an ion-trap architecture was performed with experimentalists from the University of Sussex, Google and researchers in Denmark and Germany (Lekitsch et al. 2017).

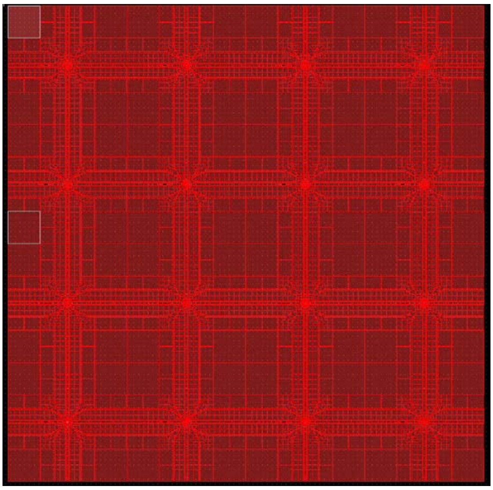


Figure 4: Schematic of a section of a larger wafer containing 16 X-junctions produced by the Sussex group in our initial design. This section would be tiled across the wafer. Each section would be electrically isolated from all others

DC system consisting of 21, 40-channel DACs (digital to analogue converters) is used to generate 840 different DC voltages, and reduces the number of DC connections required for each section to a total of 8. The 20 × 20 mm footprint is large enough to mount DACs, deserializers, amplifiers, filters, LC resonant circuit and fibres underneath the traps on vertical in-vacuum PCBs. This is shown schematically in Figure 5.

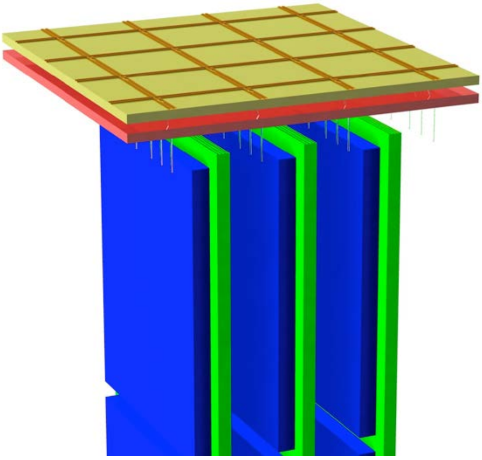


Figure 5: 20 × 20mm 16 X-junction section with chip carrier to gather the 840 DC electrode connections into neat rows of 100 × 100μm pads, 48 optical fibres, and three printed circuit boards schematically holding all required electronics (DACs, deserializers, amplifiers, filters, LC resonant circuit) within the footprint of the section.

The individual 100 × 100 mm chips also need to be accurately aligned with one another to allow shuttling of ions from one chip to the other. If a minimum alignment precision of 5μm in each direction of the rails is achieved, fast adiabatic shuttling was shown to be possible in our simulations (Lekitsch et al. 2017). To achieve this alignment accuracy, a precision machined stainless steel frame would be mounted on top of a six-axis piezoelectric positioner, possibly with the frame topped with piezoelectric pistons to permit any warping of the chip to be corrected (Figure 6). The alignment could be performed during assembly of the chips, using a laser measurement system and microscope.

Vacuum chamber

The primary design constraints on the vacuum chamber are appropriate laser access and light removal, electromagnetic shielding, unobstructed and close line of sight to all traps for characterization, clean loading of ions so the trap is not degraded over time, and possibly low operating temperature.

Given light is delivered by fibres to every X-junction, it is not feasible to always have a direct line of sight through a window out of the vacuum chamber. Instead, we propose coating the inside of the vacuum chamber

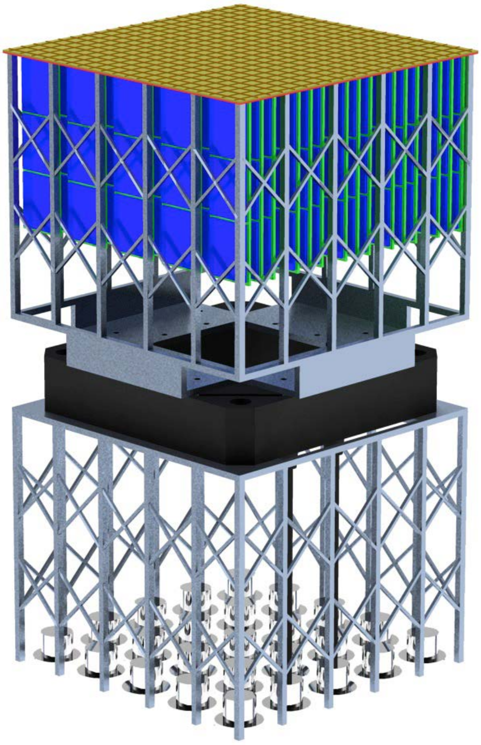


Figure 6: 100 × 100 mm chip with underside mounted electronics, precision stainless steel mounting frame, six-axis piezoelectric positioner, and array of vacuum feedthroughs. Each feedthrough must occupy no more than 20 × 20 mm and permit 48 optical fibres and 8 DC connections to pass through the vacuum chamber wall.

with a product such as Magic Black[7], which absorbs 99.99% of light at typical ion manipulation wavelengths. Excellent electromagnetic shielding and cryogenic operating temperatures have been achieved by using a thick-walled, oxygen-free, high-conductivity copper chamber in a helium-bath cryostat (Brown et al. 2011).

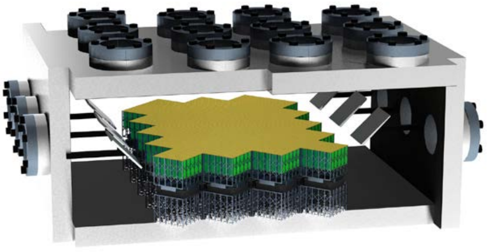


Figure 7: Repeating unit cell of a corridor of a high-performance surface code optimized ion-trap quantum computer. Windows in the side walls permit cooling laser sheets to enter. Windows in the roof next to the side walls permit exit of this light after reflection by 45° mirrors. Windows across the remainder of the roof permit characterization of the trap at time of construction.

Figure 7 shows an example of a vacuum chamber simultaneously satisfying most design constraints, with the exception of loading and 2-D scalability. Cooling lasers enter through the side walls and exit via the roof with the aid of 45 mirrors. An array of roof windows provides unobstructed and close line of sight to all traps. Figure 8 shows a modification of Figure 7 designed to also provide a separate loading region.

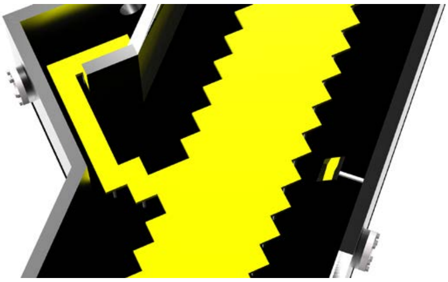


Figure 8: Vacuum chamber divided into a load region (left) and compute region (right) to control contamination. Vertical mirrors strategically attached to the walls and 45° mirrors (rectangles inside the vacuum) ensure that cooling light sheets entering through the chamber windows (rectangles outside the vacuum) can reach the vast majority of the surface of all chips. Chips deepest in the secluded region would be equipped with load slots and ion sources, enabling back-side loading.

---

[7] https://www.acktar.com/catagory/MagicBlack

2D scalability could be achieved with a honeycomb vacuum system as shown in Figure 9. The bubbles in the honeycomb corresponding to regions outside the vacuum. These bubbles ensure that the required mechanical strength of the walls does not grow with system size, and that the required alignment precision of the cooling lasers does not grow with system size. For true maintenance and assembly scalability, air locks would need to be built into the system to enable sections of the computer to be removed, replaced, or added, to increase the power of the computer. Each bubble also contains a load region of the form shown in Figure 9, to ensure that all chips are no more than a constant distance from their nearest load region.

The final blueprint design that Universal Quantum has adopted replaces a significant amount of the laser control needed to interact ions with a global microwave field pulse that is applied over all the qubits in the system. Tuning individual qubits in and out of resonance with this microwave field is achieved using a gradient magnetic field that is produced local to each individual qubit by a gradient current carrying wire that is located under zones in each trap that are used for interacting qubits (Weidt et al. 2016).

The system is ultimately designed to realise a large 2D array of qubits that can be interacted on a square grid to produce a collection of surface-code encoded logical qubits. Ion shuttling is utilised to space out the individual ions and allow ions to interact with their neighbours to the north, east, south and west across the entire machine. This is sufficient to achieve a universal error-corrected machine.

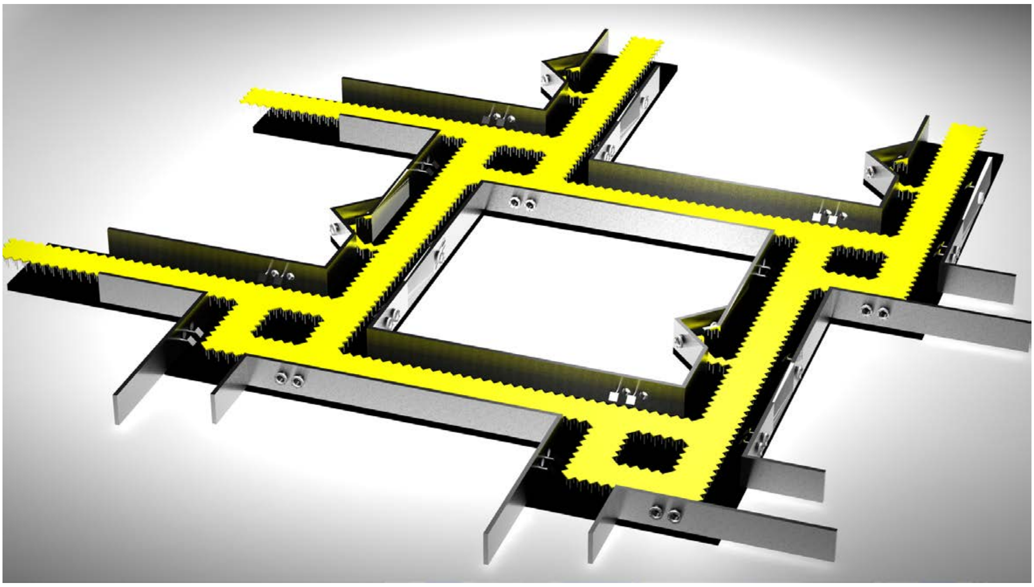


Figure 9: An infinitely extendable honeycomb vacuum system containing an arbitrarily large number of qubits. Note that optical alignment needs to be achieved only over a finite distance. Load regions are included in a regular pattern to ensure scalable ion loading.

The final size of such a honeycomb ion trap system is ultimately dictated by the number of qubits that a quantum algorithm requires. However, for large-scale applications in quantum chemistry, anywhere between 1M and 100M physical qubits will be required, depending on the exact simulation problem. This would imply an extremely large machine, at an effective 5 x 5 mm footprint per physical qubit, you would require a total area of surface traps of 5 × 5 = 25 m^2 to 50 × 50 = 2500 m^2, and with the necessary honeycomb design the actual physical size of 25–2500 m^2 of surface traps would be much larger.

Nitrogen Vacancy centres in diamond

As with Ion-traps, the ultimate goal for the design is to build a system that can faithfully create the 2D lattice of qubits to produce error-corrected logical qubits with the surface code.[8] In the case of NV diamond, we utilised a slightly different model of the surface code, known as the Raussendorf lattice (Raussendorf, Harrington, and Goyal 2006). This is due to the fact that a diamond-based quantum computing architecture uses a highly probabilistic optical connection to realise quantum gates between individual NV qubits. The Raussendorf lattice is a Measurement Based Quantum Computational (MBQC) version of the surface code. MBQC techniques are particularly useful when an underlying hardware architecture is built using probabilistic gates.

The NV diamond computer is based on qubits which are nitrogen defects within a diamond crystal. The diamond crystal itself provides what is known as a spin vacuum substrate (Aharonovich, Greentree, and Prawer 2011). Essentially, the diamond crystal itself provides the same isolation properties that an actual physical vacuum does for ion-trap technology. The operational temperature of the diamond system is 4K. While still cold, 4K cryogenic technology is far simpler than the dilution refrigeration systems needed for a 30 mK thermal environment for systems such as superconductors. 4K cryogenic technology is so advanced that we are able to effectively launch these sorts of cooling systems into space (Gehrz et al. 2007). In 2003, the Spitzer space telescope was launched by NASA. On board is 360 litre liquid helium cryostat needed to cool instrumentation to approximately 1.5K to look at faint heat signatures from astronomical objects.

Figure 10: An NV diamond chip-set that is optically coupled, placed inside a 4K helium cryostat system.

---

[8] The work in designing an architecture for quantum computation using Nitrogen Vacancy (NV) centres in diamond was performed in collaboration with the Technical University of Vienna and NTT Basic Research Labs (BRL) in Japan (Nemoto et al. 2014).

The diamond-based quantum chip-set is an array of optically coupled nitrogen-defect qubits embedded within a diamond lattice. The chip itself consists of an etched silicon base, with a ultra-thin diamond wafer "glue" on top. The diamond wafer is doped with individual nitrogen atoms separated from each other sufficiently that they don't directly interact. Individual qubits are coupled to each other using a layer of integrated silicon optics that sits above the diamond layer.

Optical pulses are sent between individual nitrogen-defect qubits to enact multi-qubit gates. These optical pulses can, in general, be weak coherent states that are easily produced. The system geometry is spaced out and optimized to allow the control structures for both the NV and optical layer to be fabricated to high accuracy.

Shown in Figure 10 is the device itself. On the right-hand side is a rendering of the microscopic detail of each chip, with multiple qubit arrays (chip-sets) connected to each other with fibre-optic connections. Shown on the bottom left is a single chip, fabricated by the Trupke group at the University of Vienna, containing a micro-cavity system [Figure 11]. On the top left is a commercially available 4K liquid-helium cryostat, similar to the device currently used in the Trupke lab at the University of Vienna.

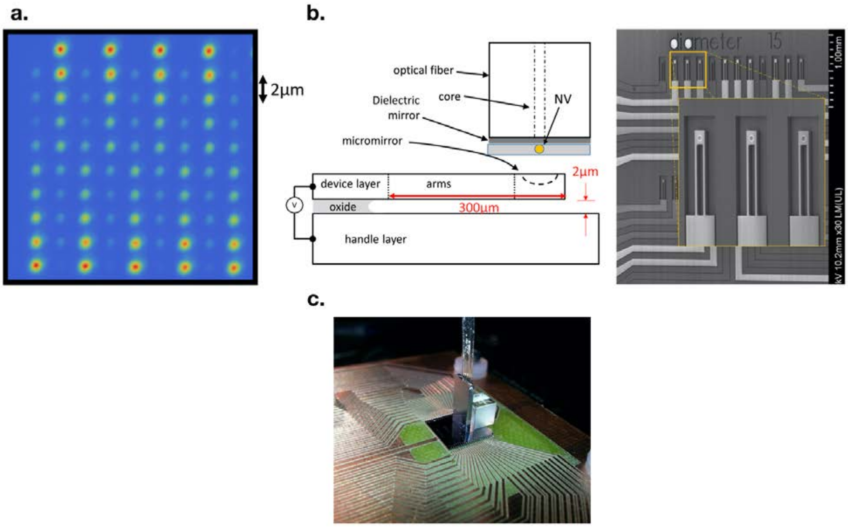


Figure 11: Photos of the three primary components of a diamond quantum chip appropriate for large-scale quantum computing. Fig a: This is a photo of a fabricated sample from Michael Trupke's group at the University of Vienna, implanted with NV-defect qubits in a 9x10 array, with each NV location separated by 2 micrometers. The intensity of each spot indicates the number of NV-centres implanted at each location, where we control the number of NV-defects implanted down each column to a single defect (the dimmest point in each column). This image also demonstrates controlled placement, with each position in our 9x10 array successfully doped. Fig b: Schematic and scanning electron microscope image of an array of micromirrors on cantilevers. The wafer-scale fabrication enables the creation of large numbers of these devices in a single fabrication process. Fig c: Photo of the cavity chip with integrated fibre optics unit connected on top. We have demonstrated controllable and selective coupling between arbitrary pairs of cavities on the chip.

The device consumes approximately 3 to 4 litres of liquid helium over the course of about 12 to 18 months. Losses are mainly though leakage in the closed loop recycling system.

The cryostat system is designed to accommodate multiple quantum chip sets. All the individual physical qubits in the system are connected to each other using integrated silicon optics and fibre connections, and multiple chips can be easily connected to allow quantum information to interact across multiple chip sets.

The NV defect is formed by first growing an ultra-high-purity diamond crystal, where each carbon atom in the lattice is the specific isotope of Carbon-12 and the lattice itself contains no other impurities. Such diamond chips are readily available commercially, with purities above 99.99%. This crystal is then doped, using low-energy ion implantation with single nitrogen atoms. At low enough densities, a single nitrogen atom will form to substitute one of the carbon atoms, and "kick out" an adjacent carbon atom, forming the Nitrogen-Vacancy defect. The diamond lattice itself produces an ultra-clean environment, negating the need for vacuum systems; providing motional stability, and eliminating a large amount of photonic and phononic noise, providing us with very stable, low-decoherence qubits.

In the diamond architecture, the electronic qubit is used as a communications mechanism to allow us to entangle multiple, isolated, qubits with optical photons via etched silicon waveguides and/or fiber optics.

Furthermore, the transition between the ground state of the electronic system and the excited state of the electronic system occurs with a photon at a wavelength of 638 nanometers. This places the transition in the optical frequency range. Unlike many other quantum architectures, the basic physics of the NV-defect provides access using photons at optical frequencies (rather than the more common microwave frequencies). Optical access to an otherwise solid-state qubit system is the key to build a distributed, modular based quantum computer that can scale arbitrarily. By using optical photons and fiber optics, we do not require direct coupling between qubits. This allows us to space out our system, have multiple parts of the computer housed in separate cooling systems and build a machine that can be expanded by simply adding more and more qubit chips as they become available.

Interacting physically separated NV-qubits

The mechanism to create entanglement between two physically separated NV-defects makes use of an optical cavity that enhances the interaction between the electronic qubit and an optical field. This is illustrated in Figure 12.

Through a well-known quantum protocol, known as dipole induced transparency, the NV-defect is placed within an optical cavity that resonates at the same frequency as the optical transition of the NV-electronic qubit, when that qubit is only in the zero state. This resonance-matching changes the reflectivity properties of the cavity such that if a photon tries to enter the cavity from outside while the atom is in the zero state, it will be reflected at the entrance to the cavity. If the electronic qubit is in the one state, the photon will enter the cavity and be absorbed or scattered. This quantum mechanical dependence of the reflectivity properties of the cavity gives us a mechanism to produce entanglement between two spatially isolated electronic qubits in two separate NV-cavity systems.

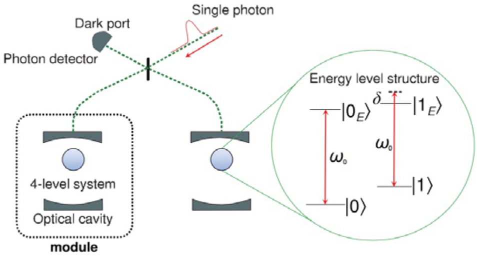


Figure 12: The optical mechanism to entangle two NV-defects. The optical transitions of the respective electronic systems is tuned to resonance with two optical cavities. This resonance condition changes the reflectivity properties of the two cavities dependent on the electronic state. If each electron is placed into a superposition of its two ground states and a single photon that is first split on a beamsplitter is sent to both cavities, there is a finite probability that they will reflect from both cavities and are detected at the photon detector. Because we don't know which cavity the photon reflected from, the resultant electronic state is entangled. The success probability for this scheme is upper bounded at 12.5% when all other components are perfect.

If we take a single photon from an external source (a highly attenuated laser, for example) and first send it through an optical beam splitter, this will place the photon into an equal superposition of heading towards NV-system one or NV-system two. If each of the electronic qubits are placed into an equal superposition of the zero and one states, then the photon will be reflected from each of the cavity systems with a probability of 50%. We then send these reflected photons back through the beamsplitter and with a photon detector, measure if we see a photon come back to us.

87.5% of the time, we do not see any photon with our detector, as the photon may have been absorbed by either one of the two NV-cavity systems or it may not come out the same port of the beam splitter corresponding to the location of the photon detector. However, 12.5% of the time, we will see a "click," indicating that the photon returned to the detector. If we see a "click," then the photon must have reflected off one of the two cavities. However, we cannot say which cavity it actually reflected from. If we cannot ascertain which cavity it reflects off, the final state must be a linear superposition of the two possibilities (photon reflected off cavity one + photon reflected off cavity two). Since the quantum mechanical state of the electron in each of the two NV-cavity systems determines if the photon reflects or not (depending on their respective qubit state), the detection of the photon will result in the electronic states of the two NV-defects being entangled. This is the primary mechanism for NV-defect entanglement.

In the case of the diamond based system, the nitrogen nucleus provides us with a protected memory space to store quantum entanglement while we repeatedly attempt to create new electron–electron entanglement using highly probabilistic quantum gates. This technique is well known in the community and it is commonly referred to as brokered graph state quantum computation.

Using the optical interface, we can attempt to establish an entanglement bond between the electronic qubits in two, physically separate, NV-cavity systems. Conservatively, assuming a connection efficiency of 1%, we require approximately 100 attempts before we are reasonably confident that a connection will be established (>50% chance of a connection after 100 attempts).

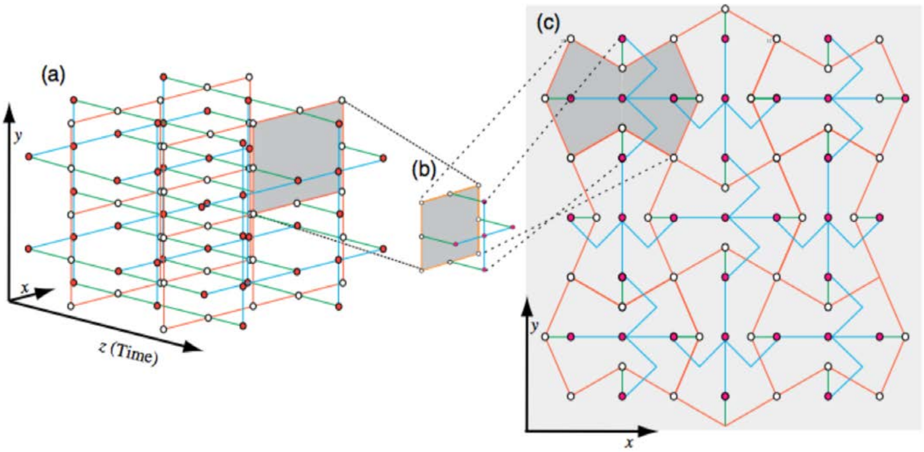


Figure 13: The diamond architecture is based on a 3D cluster state that is a universal resource state for topologically error-corrected quantum computation. Rather than creating a full 3D lattice, we create a 2D+1 lattice on a planar lattice of NV-qubits. This allows for arbitrary quantum algorithms that are protected from errors using powerful topological codes.

Given the intrinsic speed of this system, these 100 attempts can be completed in approximately 3 microseconds (Nemoto et al. 2014). If we were to try to connect a third electronic qubit after this initial successful connection, any failed attempt would destroy the entanglement we had previously established. Hence, before we attempt to connect a third electronic qubit into our entangled state, we use the hyperfine interaction between the electrons and their respective nitrogen nuclear qubits to transfer the entanglement.

If a successful entanglement bond is transferred to the nuclear qubits, this “frees” the electron qubits to be used again to probabilistically create another entanglement bond with other electronic qubits in other NV-cavity systems without the possibility of destroying the entanglement we have already created. The nuclear system acts as a protected memory (or broker) to allow for further creation of electron/electron entanglement. The protected nuclear memory allows for extremely inefficient optical connections without adversely impacting the error rates associated with the creation of NV/NV entanglement or slowing the computer down by such a degree as to make large-scale computation impractical.

Creating an entangled resource state for error-corrected computation

The basic mechanism for NV/NV entanglement can be used to create a cluster state of arbitrary size. Hence the diamond architecture can be used to build a universal resource state for large-scale, error-corrected quantum computation.

Embedding topological error-correction codes into our computational model simply requires us to produce an entangled resource state that can be used to mimic surface code, error-corrected logical qubits. This requires the creation of a 2D+1 dimensional cluster state, known as the Raussendorf lattice (Raussendorf, Harrington, and Goyal 2006). The structure of this state is illustrated in Figure 13.

The cluster state we require is abstractly a 3D state. The cross section of this state represents the size of the computer — how much space we have in terms of error-corrected qubits — while the third dimension represents time steps available in our computation — the longer the third dimension is, the more steps we have available. Creating this 3D cluster state directly would introduce a temporal scaling to our architecture — we would need more physical qubits to enact more error-corrected gates — but we do not need to create the entire third dimension in physical space. In fact we only need to create two layers at a time along the temporal axis of the cluster state. Hence the cluster we need to prepare is the 2D cross section, plus one additional layer along the temporal axis (i.e. a 2D+1 cluster state).

Figure 13c illustrates the actual physical layout of the cluster state that we need to produce, assuming each of our physical NV-defects lies on a 2D plane. We basically take two sequential cross-sections of the 3D cluster and "pancake" it down to 2D. This results in a 2D qubit layout that requires next-to-nearest-neighbour connections (the colour coding in Figure 13 is not relevant to this discussion), but the optical connectivity of this architecture allows for these longer-range connections.

Creating a cluster state with this general structure provides an appropriate universal resource state for fault-tolerant, topologically error-corrected, universal computation. The size of the computer and the strength of error correction is only related to the actual 2D size of the array shown in Figure 13c. The array in Figure 13c consists of cells in the 3D cluster state, and a large-scale quantum algorithm (for example factoring) would require arrays of the order $10{,}000 \times 10{,}000$ cells (Nemoto et al. 2014).

Because a diamond system is optically connected and cooling infrastructure is comparatively simple when compared to large vacuum systems or dilution refrigeration, a highly distributed diamond system is expected to be simpler to scale. While such a machine would still be large, it would not need to be a carefully interconnected high-vacuum systems or ultra-large dilution refrigerators.

Superconductors

The motivation behind the Japanese superconducting micro-architecture is a well-known problem within superconducting quantum computing known as "the wiring problem." This micro-architecture takes a fundamentally different structural approach to eliminate this problem without changing any other issue related to large-scale operation of a superconducting quantum computer.[9]

Superconducting qubit systems have arguably emerged as the leading platform for large-scale computing architectures. Not only have we seen significant advances in recent years in reliable fabrication and control technology, but the quality of the qubits themselves has increased by many orders of magnitude. Superconducting qubit systems have demonstrated physical error rates close to (and in some cases, below) the fault-tolerant threshold for surface-code-based error-correction techniques (Barends et al. 2014), and multi-qubit arrays have been fab-

---

[9] The third design discussed is a modified design for a superconducting quantum computer that was designed in collaboration with the group of Jaw Shen-Tsai at the Japanese national laboratories, Riken and Tokyo University of Science (Mukai et al. 2020; Kwon et al. 2020).

ricated and tested by many groups worldwide. Private investment in superconducting quantum computing technology has also exploded, with companies such as Google, IBM, Alibaba, Intel, Rigetti and others now actively and aggressively promoting and supporting the platform. The first demonstration of quantum supremacy was demonstrated with Google's 53-qubit, Sycamore superconducting processor, and IBM have deployed over 20 superconducting quantum computers through the cloud on the IBMQ network.[10]

Scaling these systems to the level needed to achieve error-corrected, commercial applications will require integrated chipsets containing of order 1000 physical qubits or more in a surface code error-corrected logical qubit and this presents certain technological challenges. One of the most significant is the so-called wiring problem. This is the fact that a large, error-corrected array of superconducting qubits requires a 2D qubit chipset that allows for nearest-neighbour couplings and that qubits within the centre of such a chip cannot be directly accessed for the fabrication of bias lines, control lines and readout machinery.

The current consensus within the superconducting community is that the control wiring for such chips should be fabricated in the third dimension, utilizing several techniques to place bias, readout and control wires orthogonal to the plane of the chipset itself. This technique has shown promise (Rosenberg et al. 2017; Foxen et al. 2017), but it is very unclear if these control fabrication techniques are compatible with maintaining high-fidelity operations. The largest concern is the ability to reduce cross-talk and control line contamination of neighbouring qubits to the degree necessary to achieve fidelities of 99% or higher across the chip.

This new micro-architecture was designed specifically to side-step this issue. We demonstrated that a pseudo-2D arrangement of superconducting qubits — completely compatible with surface code based error-correction — can be constructed in a physical bi-linear arrangement of superconducting qubits. This bi-linear array allows for each physical qubit to be biased, measured and controlled using wiring that remains in-plane with the chipset — eliminating completely the need for 3D control line fabrication.

To achieve this new architecture we introduced small air-bridges within the resonators coupling together individual qubits. These air-bridges allow us to create a criss-cross resonator design that allows us to create the pseudo-2D qubit arrangement. We demonstrated that the resonator quality and crosstalk is not adversely affected by the introduction of these air-bridges and that we can anticipate no adverse effects on the architecture by moving to this new design.

A standard superconducting quantum microarchitecture

To maintain compatibility with quantum error-correction codes, a minimal design for a superconducting chipset is a 2D nearest-neighbour interacting array. Google's Sycamore processor is a 54-qubit planar wafer design. Illustrated in Figure 6a is a schematic of the qubit (gray crosses) layout, where 2D nearest-neighbour interactions are mediated by adjustable couplers (blue boxes). The four adjustable couplers allow switchable quantum gates to be implemented between

---

[10] https://www.ibm.com/quantum-computing/systems/

a given qubit and its four nearest neighbours. In Sycamore, a 54-qubit chipset was initially designed, but one of the 54 qubits was non-functional after fabrication (white cross).

Figure 14b shows an actual photograph of the 10mm Sycamore chip. The central square region is the actual chipset, while emanating from the central processing region are control lines that are used to bias, control and measure each of the 54 qubits and control each of the 88 adjustable couplers (both built from transmons[11]).

The packaging of Sycamore (i.e. the control lines that need to be fabricated from each transmon to the boundary of the chipset) is already quite dense and complex. In the context of the surface code error-corrected qubit, an array of 54 qubits would only be sufficient to construct a distance $d = (\sqrt{N} + 1)/ 2 = 4$ logical qubit (using a standard square planar code configuration (Horsman et al. 2012)). This is a very small amount of error correction and would not be sufficient for any large-scale quantum algorithm.

If we expanded to a $N = 1521$ or $N = 2401$ chipset, corresponding to distances $d = 20$ and $d = 25$ code respectively, it would be simply too dense to have enough physical space on the chipset for bias, control and measurement lines for both the transmons acting as qubits and the transmons acting as adjustable couplers. (A $N = 1521$ qubit chipset would require over 3000 additional transmons as couplers, for a total of over 4600 transmons per chip.)

Planar wiring is therefore not a viable method to scale a microarchitecture of this type. Instead, the common method is to envisage control lines to be fabricated perpendicular to the chip plane. These three-dimensional wiring technologies generally consist of techniques such as flip-chip bonding, pogo pins, and through-silicon vias (TSVs) (Barends et al. 2014; Takita et al. 2017; Reagor et al. 2018; Chou et al. 2018; Bejanin et al. 2016; Vahidpour et al. 2017; Foxen et al. 2018; Rosenberg et al. 2017).

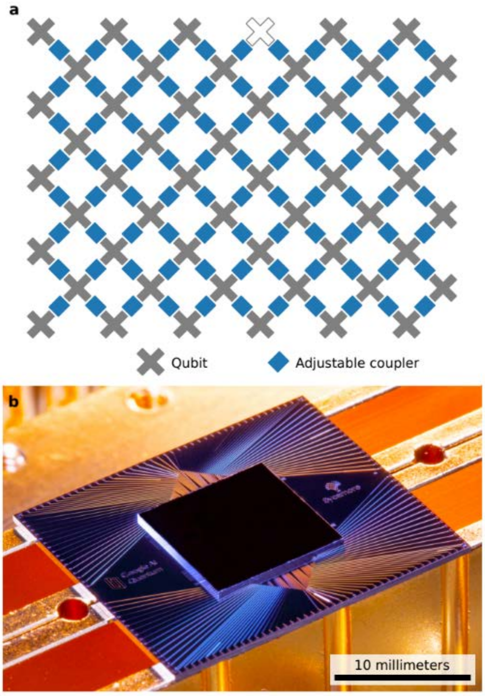


Figure 14: The Google Sycamore processor (Arute et al. 2019). Fig a represents the layout of 54 qubits (grey crosses, one white cross for a non-functioning qubit) in a 2D nearest neighbour microarchitecture. Each blue box represents a transmon that is used as an adjustable coupler. Fig b is a photo of the Sycamore chip. You can see control wires that emanate from the central square region containing the qubits.

While it is still unclear if complex 3D wiring will affect the performance of the

---

[11] A transmon is a type of superconducting charge qubit designed to have reduced sensitivity to charge noise. [Ed.]

underlying chipset — fidelities of superconducting chipsets are very much on the boundary of the threshold for what is required of error-corrected system — it is expected to be a source of fabrication and cross-talk noise that would be good if it could be eliminated completely.

The superconducting air-bridge resonator

The key to this new micro-architecture design was introducing an air-bridged superconducting resonator. An air-bridged resonator is a standard Coplanar Waveguide Resonator (CWR) that contains a break in the waveguide and a literal bridge that connects the two halves of the waveguide. Figure 15 shows a scanning electron microscope (SEM) image of a superconducting air-bridge that was fabricated by the Google group (Chen et al. 2014).

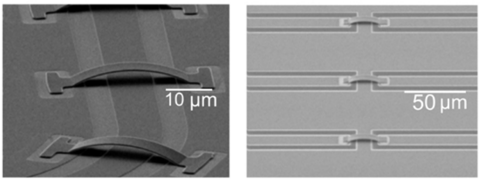


Figure 15: A scanning electron microscope (SEM) image of a series of air-bridges fabricated by the Google quantum team (Chen et al. 2014).

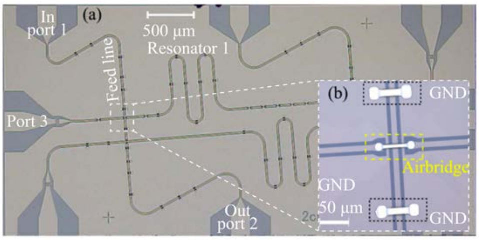


Figure 16: A SEM image of four crossed resonators, fabricated by the Tokyo University of Science group as part of this microarchitecture redesign. From (Mukai et al. 2020).

The existence of a bridge allows us to consider a cross resonator design. This is where a CWR is fabricated under a bridge that connects a second CWR fabricated in an orthogonal direction. Illustrated in Figure 16 is a SEM image of the actual device that the Tokyo University of Science team fabricated as part of this project.

This prototype device consisted of four resonators: two run vertically across the chip and two run horizontally. The resonators have four crossing points, where an air-bridge is used so that one resonator travels under the other. Notice that air-bridges are used in multiple locations around the CWR chip, to ensure that there is a common electrical ground. If certain conducting islands were isolated on the chipset, it would create a potential difference between regions and consequently create stray capacitances.

Along the horizontal resonators, a series of air-bridges were fabricated along its length. These air-bridges did not have a secondary resonator passing underneath and were fabricated to test whether there was any level of resonator degradation as a function of the total number of air-bridges along a resonator.

Figure 17 illustrates experimental data of the prototype air-bridged system. In Figure 17a we plot the infidelity of a resonator-mediated quantum gate between two superconducting qubits as a function of the quality factor of the resonator. The horizontal dotted line represents the error rate (infidelity) reaching the level where error correction is viable (approximately 0.7%) and the vertical line is the experimental measured quality factors of the test resonator system with between 15 and 20 air-bridges. The blue curve lies under the horizontal line when it intersects with the vertical line, meaning that a resonator containing between 15 and 20 air-bridges can still be used to enact quantum gates between two superconducting qubits at an error rate below the surface code threshold. (top) 25 mm (bottom), excluding headers and footers.

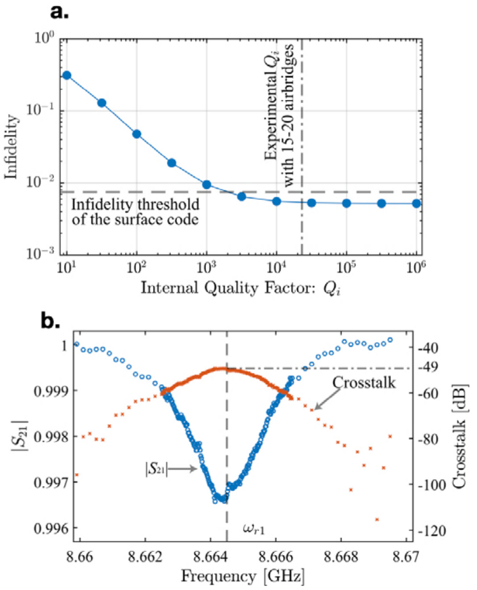


Figure 17: Fig a. illustrates the infidelity of a resonator mediated quantum gate between two superconducting qubits as a function of the quality factor of the resonator. Fig b. illustrates the resonance of the air-bridged CWR (blue) and the measured noise that contaminates the orthogonal resonator (red). From (Mukai et al. 2020).

Figure 17b is a plot of the resonance dip of the CWR (blue) and the measured crosstalk to the orthogonal resonator that passes under the air-bridge (red). At resonance, the maximum measured crosstalk was approximately −49dB. This demonstrates that resonator crosstalk for a crossed resonator system is effectively non-existent. Potential induced phase-shifts on the resonator mode during simultaneous use of the two resonators for a pair of two-qubit interactions — each interaction using one of the two resonators — still needs to be performed when the full four-qubit system is fabricated.

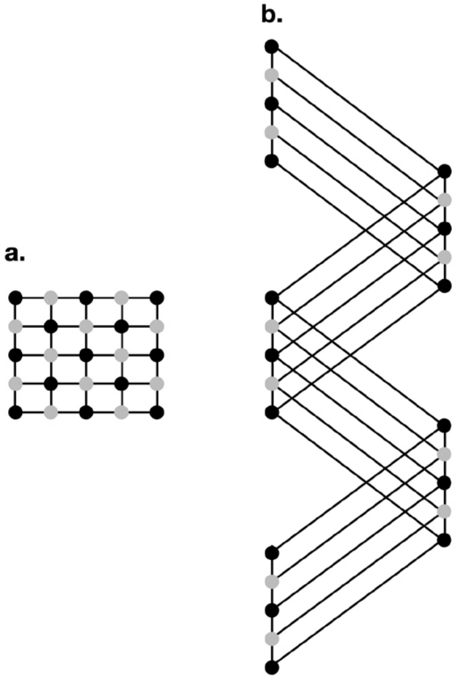


Figure 18: A new bi-linear microarchitecture formed from the standard 2D layout of qubits needed for surface code error correction. Fig a illustrates a standard logical qubit that requires a grid. Fig b illustrates the new bi-linear arrangement where couplings between columns are achieved using air-bridged crossed CWRs.

A bi-linear microarchitecture for superconducting quantum chips

Once air-bridges can be introduced on resonators, enabling the ability to cross resonators, we can redesign the standard 2D qubit layout with nearest-neighbour interactions to a bi-linear array of super-

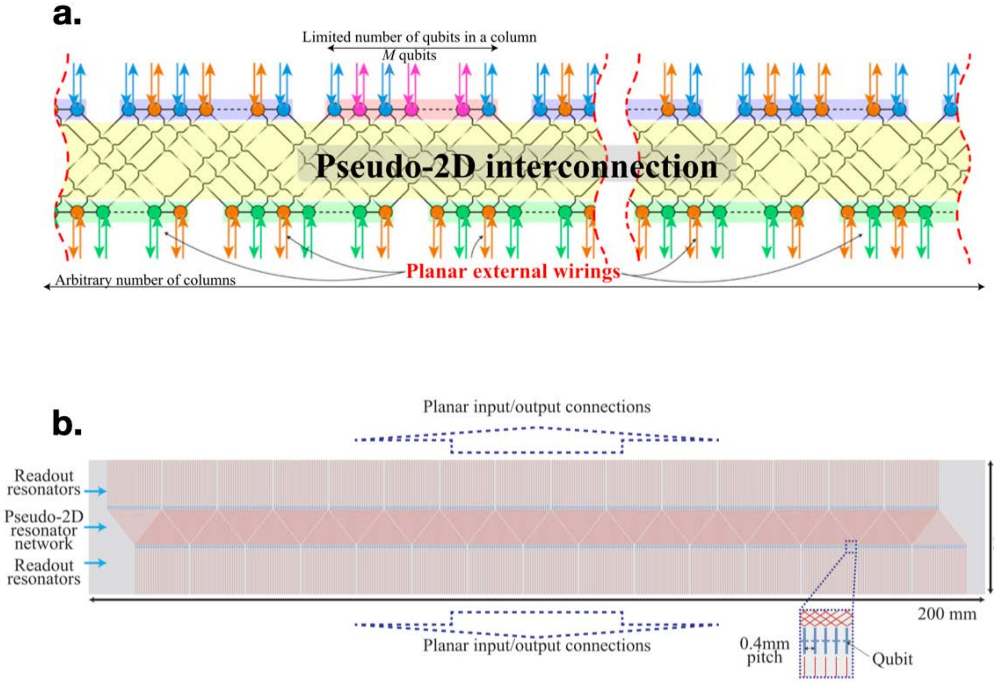


Figure 19: Physical layout of the new microarchitecture. Fig a: An arbitrarily long but fixed-width surface code can be created using a bi-linear arrangement of superconducting qubits. The fixed width of the surface code ensures that the air-bridged resonators have a finite length and number of air-bridged crossings. Each superconducting qubit can be accessed in the plane for the control, initialisation, and readout technology. Fig b: A design of a logical qubit chip consists of 30×30 physical qubits, encoding a d = 15 logical qubit. Qubits are depicted as blue crosses (inset) and resonators are as red lines. All the external input and output connections can be achieved by the conventional planer wiring technology. The resulted chip size is approximately 30 mm × 200 mm rectangular. From (Kwon et al. 2020).

conducting qubits that interact through a series of crossed resonators.

The modification to the microarchitecture is shown in Figure 18. Figure 18a shows a standard 5×5 surface code array that corresponds to a distance $d = 3$ error-correction code. As you can see, the qubits within the interior become difficult to access for fabrication of bias lines, control and readout lines.

If we take each column of this 2D array and lay them out in two distinct columns, alternating columns from the original 2D patch, we can create a bi-linear array, as shown in Figure 18b. Required qubit interactions within a column do not change, but interactions between columns now become longer range and they cross. These interactions can be achieved using air-bridged resonators.

The significant difference between these two layouts is that, in the bi-linear array, there is now lateral planar access to every qubit within the computer. Qubits that used to be buried within the centre of the 2D lattice are now placed along one of the columns in the bi-linear array, where bias, control and measurement lines can be fabricated and placed immediately alongside. This solves the 3D wiring problem by introducing a new component, the air-bridged resonator, which was demonstrated not to adversely affect the performance of a quantum gate between two superconducting qubits.

As each column is alternatively placed on each side of the bi-linear array, the number of air-bridges is dictated by the number of qubits within a column. As shown in Figure 19a, each block (both green and blue) corresponds to a single column of M qubits. The number of crossing points for air-bridges is $M - 1$ and as we can alternate on which resonator the air-bridge is actually fabricated, the total number of air-bridges on a single resonator is $(M - 1)/2$. Before discussing the total length of the superconducting chip-set, we need to detail how qubits are arranged

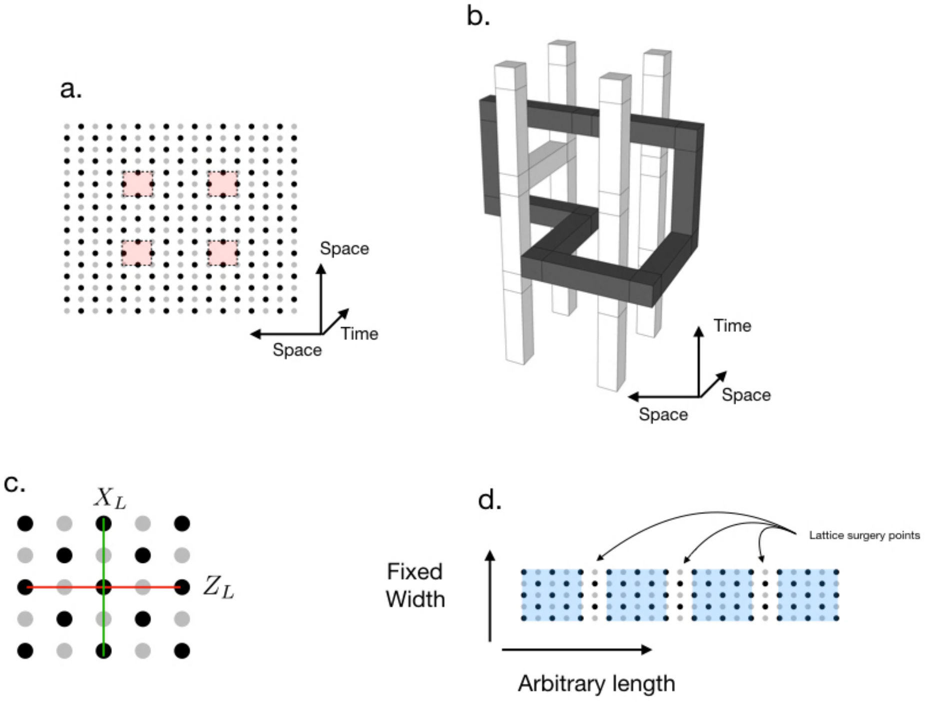


Figure 20: Braid and lattice surgery based logic on a standard 2D surface code. Fig a is how a set of two logical qubits, encoded with a distance d = 3 surface code are introduced as four pairs of defects, where a defect is defined as a region of the surface code where physical circuits are switched off (shaded in red). Fig b is a logical CNOT operation enacted over two logical qubits (the two pairs of white defects). Time is represented vertically and, as the circuit is executed, defects are moved to complete topological braids. A quantum circuit is consequently represented as a geometric figure, with a cross-section related to the number of physical qubits needed and the third axis representing the time needed to complete the logic gate. Fig c is a single planar code logical qubit at d = 3. In this situation, logical qubits are isolated from each other until logic operations via lattice surgery are enacted. Fig d illustrates the associated logical qubit layout for lattice surgery logic. Each logical qubit is illustrated as the blue shaded region, with an extra column of physical qubits used as the merge/splitting points for lattice surgery logic. Significant physical resources are saved using lattice surgery compared to defect-based encoding.

in this system so that logic operations can be performed in the error-corrected system.

In the surface code, there are two predominant models for performing error-corrected logic. The first techniques is known as topological braiding, and was the original formalism for error-corrected logic (Raussendorf, Harrington, and Goyal 2006; Fowler and Devitt 2012; Fowler, Devitt, and Jones 2013; Paler, Devitt, and Fowler 2016; Paler et al. 2012; Devitt et al. 2013).

The layout of qubits for braided logic is a full 2D array of qubits as shown in Figure 20a. In this array, we have four regions (coloured in red) where the error-correction procedures that usually occur across the lattice are simply not performed ("switched-off"). This creates holes or "defects" within the code. These defects are effectively degrees of freedom in the code-space that can be used to encode multiple qubits of logical information.

The error-correction strength of the information encoded within these defects is determined by the circumference of the defect and the separation in the lattice between defects. In Figure 20a the number of data qubits (black qubits) that circumscribe a defect region is four and the minimum number of data qubits between any two defect regions is three. Consequently the code distance used to encode the information is $d = 3$ (the minimum of the two). For various technical reasons (Raussendorf, Harrington, and Goyal 2006), a logical qubit of information is represented by two defects. Hence Figure 20a represents a computer encoding two logical qubits with a distance $d = 3$ code.

Logic operations are performed by changing the locations of these defect regions over time and tracing out a space/time geometry of world lines for these defect regions. An example of this is given in Figure 20b. The XY-plane of this diagram represents the spatial 2D lattice of Figure 20a, where the four white pillars represent four of the defects. Time is represented by the Z-direction. As the computer evolves over time, the defect regions in the lattice are moved (by switching off and on parts of the 2D lattice) to trace out the geometric structure of Figure 20b. This enacts a gate operation. In this case, a two-qubit logically encoded CNOT operation. The control qubits are represented by the two white pillars to the left, the target qubit is the two white pillars on the right. The input for the circuit is at the bottom of the image while the output is at the top. The details of how these gates actually enact logic operations can be found in several references (Fowler and Devitt 2012; Fowler, Devitt, and Jones 2013).

The main point from Figure 20a is that the computer needs to be effectively a large square 2D array of physical qubits, with defects placed throughout the array. This is not compatible with the new bi-linear array microarchitecture where we are limited to the total length of a column in the 2D lattice which dictates the number of air-bridged resonators needed by the design.

This can be solved by moving to a different methodology of fault-tolerant logic called lattice surgery (Horsman et al. 2012). In the lattice surgery model, logical qubits are simple square patches of surface code, determined by the distance of the underlying quantum code. A square $M \times M$ patch corresponds to a code distance of $d = (M + 1)/2$ [Figure 20c]. Each square patch is laid next to each other This effectively means that our arrangement of logical qubits in the computer becomes a Linear

Nearest Neighbour array (LNN) [Figure 20d]. We have a sufficient number of rows in the design to encode a single logical qubit and we have as many columns as necessary to house the total number of logical qubits in the computer.

In lattice surgery, isolated square patches of planar code are interacted along a boundary to enact multi-qubit logic gates. This reduces the overall physical resource cost of each logical qubit and several results now suggest that lattice surgery techniques will always be more resource efficient when implementing large-scale algorithms (Herr, Nori, and Devitt 2017; Litinski and Oppen 2018; Fowler and Gidney 2018).

For a single logical qubit encoded with the planar code, a square 2D array of physical qubits is needed. For a distance d quantum code, a $(2d - 1) \times (2d - 1)$ array of physical qubits is sufficient. Illustrated in Figure 20 is a distance planar code requiring 25 physical qubits with the associated logical bit-flip (XL) and phase-flip (ZL) operators illustrated. In Figure. 20d, we illustrate the same LNN logical layout of planar code logical qubits that require less physical resources than defect-based logical qubits. In Figure 20d there is an additional column of physical qubits that are spacers between each encoded qubit that is required to perform the lattice surgery operations. It should be noted that the current methods for circuit compilation using lattice surgery still assume a 2D nearest-neighbour arrangement of logically encoded qubits (Herr, Nori, and Devitt 2017; Fowler and Gidney 2018). Compilation into this pseudo-LNN logical structure will require modifications over current techniques (Herr, Nori, and Devitt 2017). However, this won't adversely impact the physical structure of this new architecture.

For a very large error-correcting code, d can be of the order of 15–20, requiring an array containing 29–39 rows of qubits with 29–39 columns, per logically encoded qubit. Consequently, for a quantum computer containing N logical qubits at distance 15 on the planar code, we would utilize an array of $29 \times (29N + (N - 1))$. Here, 29 is the number of qubits in a column, and $29N$ is the number of columns in the array for each logical qubit, and the extra factor of $(N - 1)$ is the spacing region between each logical qubit needed for lattice surgery. This would translate into a bi-linear array, as shown in Figure 19b of $N(2d-1)(2d-2) = 29 \times 30N$, with each set of air-bridged cross-resonators having at most $\lceil (2d - 1)/2 \rceil = 15$ crossings. The factor of ½ comes about due to the fact that alternate resonators can be chosen to contain an air-bridge. Hence, while 29 crossings are at required at most, a given resonator will only contain half that number of air-bridges.

In Figure 19b we illustrate a much larger array, where each individual qubit has wiring access from either above or below the bi-linear array. The cross-resonator network containing the air-bridges are contained within the centre, and allows for the bi-linear array to operate as if it were a long, rectangular 2D array of physical qubits. A system of this size would represent a single superconducting chipset, containing $30 \times 30 = 900$ physical qubits, encoded into a single error-corrected qubit. This would correspond to a distance of $d = 15$ quantum code, sufficient to correct for up to seven physical errors in each error-correction cycle. Estimates on the physical size of this chipset will be 200 mm × 30 mm.

This new micro-architecture for superconducting quantum computers effectively eliminates the problem of wiring up a massive 2D array of physical qubits. Instead, we translate the qubit chip into a bi-linear array of physical qubits that have direct lateral access for bias, control and measurement wiring. While superconducting chipsets have not yet reached the scale where a shift to this new micro-architecture is needed, as these systems scale further a new approach to the underlying structure will need to be adopted by essentially all manufacturers of superconducting quantum computers.

Designing systems of the future

As nascent quantum chipsets further progress, redesigns and additions to these blueprints will undoubtably occur. The basic building blocks for each of these quantum computing platforms have been demonstrated and there is a clear conceptual pathway to a large-scale system, capable of universal, error-corrected quantum computation.

It is anticipated that many of the challenges that lie ahead are related to the ability to further decrease error rates through improvements in fabrication and control, and solve the problems related to how we scale systems under the constraints of environmental infrastructure such as dilution refrigeration systems and vacuums.

There have been significant advancements in the construction of error-correction protocols and quantum resource optimisation. The surface code still remains the preferred technique for error-correction in experimentally realisable large-scale systems, and physical qubit resources continue to drop as theorists develop new and improved methods for error-corrected logic operations and algorithmic compilation.

The timeframe of when a fully error-corrected system will become available to implement scientifically or commercially useful quantum computing systems is still unknown, but for many platforms the initial ingredients have been demonstrated and it is becoming clear that engineering challenges and capital may be much more significant than any fundamental issues of quantum physics.

Acknowledgements

The work reviewed and referenced in this paper could not have been achieved without fantastic collaborations with theorists and experimentalists, physicists and computer scientists. I would like to acknowledge the ion-trap group at the University of Sussex, the NV-diamond groups at the University of Vienna and the Technical University of Vienna and the Superconducting quantum research teams at Toyko University of Science and Riken, Japan, for their fantastic collaborations on the blueprints reviewed. Much of the initial work on architectures and system design was performed in collaboration with K. Nemoto and W.J. Munro of the National Institute of Informatics and NTT labs, Japan, and I would like to acknowledge my many collaborators who worked with me on error-correction, compilation and optimisation for large-scale machines such as A.G. Fowler, A. Paler, D. Herr, D. Horsman, R. Van Meter, A.D Greentree, T. Tilma and A.M Stephens.

While I have attempted to cite many other results from others on architecture development in all of the systems examined in this paper, I am sure I have missed many wonderful papers. The context of this manuscript did necessitate it being “me”-centric and that in no way should be interpreted to downplay the extraordinary work that exists

within the field of quantum architectures and system design.

References

Aaronson, S. (2013). *Quantum Computing since Democritus*. Cambridge University Press.

Aaronson, S., and D. Gottesman. (2004). Improved simulation of stabilizer circuits. *Physical Review* A 70 (5): 052328–28. https://doi.org/10.1103/PhysRevA.70.052328.

Aharonovich, I., A.D. Greentree, and S. Prawer. (2011). Diamond photonics. *Nature Photonics* 5: 397. https://doi.org/10.1038/nphoton.2011.54.

Amini, J. M., H. Uys, J.H. Wesenberg, S. Seidelin, J. Britton, J.J. Bollinger, D. Leibfried, C. Ospelkaus, A.P. VanDevender, and D.J. Wineland. (2010). Scalable ion traps for quantum information processing. *New J. Phys.* 12: 033031.

Arute, F., K. Arya, R. Babbush, D. Bacon, J.C. Bardin, R. Barends, R. Biswas, et al. (2019). Quantum supremacy using a programmable superconducting processor. *Nature* 574 (7779): 505–10. https://doi.org/10.1038/s41586-019-1666-5.

Babbush, R., C. Gidney, D.W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven. (2018). Encoding electronic spectra in quantum circuits with linear t complexity. *Physical Review* X 8 (4): 041015–15. https://doi.org/10.1103/PhysRevX.8.041015.

Barends, R., J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T.C. White, et al. (2014). Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* 508 (April): 500 EP–. https://doi.org/10.1038/nature13171.

Bejanin, J.H., T.G. McConkey, J.R. Rinehart, C.T. Earnest, C.R.H. McRae, D. Shiri, J.D. Bateman, et al. (2016). The Quantum Socket: Three-Dimensional Wiring for Extensible Quantum Computing. arXiv:1606.00063.

Braunstein, S.L., and P. van Loock. (2005). Quantum information with continuous variables. Reviews of Modern Physics 77 (2): 513–77. https://doi.org/10.1103/RevModPhys.77.513.

Brennan, G., S.J. Devitt, T. Robinson, and P. Rohde. (2021). *An Australian strategy for the quantum revolution*. Australian Strategic Policy Institute Policy Brief 43.

Brown, K.R., A.C. Wilson, Y. Colombe, C. Ospelkaus, A.M. Meier, E. Knill, D. Leibfried, and D.J. Wineland. (2011). Single-qubit gate error below in a trapped ion. *Phys. Rev. A* 84: 030303.

Brown, K.R, J. Kim, and C. Monroe. (2016). Co-designing a scalable quantum computer with trapped atomic ions. *Npj Quantum Information* 2 (1): 16034. https://doi.org/10.1038/npjqi.2016.34.

Chen, Z., A. Megrant, J. Kelly, R. Barends, J. Bochmann, Y. Chen, B. Chiaro, et al. (2014). Fabrication and characterization of aluminum airbridges for superconducting microwave circuits. *Applied Physics Letters* 104 (5): 052602. https://doi.org/10.1063/1.4863745.

Chou, K.S., Jacob Z. Blumoff, C.S. Wang, P.C. Reinhold, C.J. Axline, Y.Y. Gao, L. Frunzio, M.H. Devoret, L. Jiang, and R.J. Schoelkopf. (2018.) Deterministic teleportation of a quantum gate between two logical qubits. *Nature* 561 (7723): 368–73. https://doi.org/10.1038/s41586-018-0470-y.

Cirac, J.I., and P. Zoller. (1995). Quantum computations with cold trapped ions. *Phys. Rev. Lett.* 74: 4091–94.

———. (2000). A scalable quantum computer with ions in an array of microtraps. *Nature* 404: 579–81.

Crick, D.R., S. Donnellan, S. Ananthamurthy, R.C.Thompson, and D.M. Segal. (2010). Fast shuttling of ions in a scalable penning trap array. *Rev. Sci. Instrum.* 81: 013111.

Devitt, S.J., A.G. Fowler, A.M. Stephens, A.D. Greentree, L.C.L. Hollenberg, W.J. Munro, and K. Nemoto. (2009). Architectural design for a topological cluster state quantum computer. *New. J. Phys.* 11: 083032.

Devitt, S.J., A.G. Fowler, T. Tilma, W.J. Munro, and K. Nemoto. (2010). Classical processing requirements for a topological quantum computing system. *International Journal of Quantum Information* 8: 1–27.

Devitt, S.J., W.J. Munro, and K. Nemoto. (2011). High performance quantum computing. *Progress in Informatics* 8: 49. https://doi.org/10.2201/niipi.2011.8.6.

Devitt, S.J. (2016). Programming quantum computers using 3-D puzzles, coffee cups, and doughnuts. *XRDS* 23 (1): 45–50. https://doi.org/10.1145/2983545.

Devitt, S.J., A.D. Greentree, A.M. Stephens, and R. Van Meter. (2016). High-speed quantum networking by ship. *Scientific Reports* 6 (1): 36163. https://doi.org/10.1038/srep36163.

Devitt, S.J, W.J. Munro, and K. Nemoto. (2013). Quantum error correction for beginners. *Reports on Progress in Physics* 76: 076001. https://doi.org/10.1088/0034-4885/76/7/076001.

Devitt, S.J., A.M. Stephens, W.J. Munro, and K. Nemoto. (2013). Requirements for fault-tolerant factoring on an atom-optics quantum computer. *Nature Communications* 4 (1): 2524. https://doi.org/10.1038/ncomms3524.

DiVincenzo, D.P. (2000). The physical implementation of quantum computation. *Fortschritte Der Physik* 48 (9-11): 771–83. https://doi.org/10.1002/1521-3978(200009)48:9/11%3C771::AID-PROP771%3E3.0.CO;2-E.

Duan, L.-M., B.B. Blinov, D.L. Moehring, and C. Monroe. (2004). Scalable trapped ion quantum computation with a probabilistic ion-photon mapping. *Quant. Info. Proc.* 4: 165–73.

Fedorov, A.K., A.V. Akimov, J.D. Biamonte, A.V. Kavokin, F. Ya Khalili, E.O. Kiktenko, N.N Kolachevsky, et al. (2019). Quantum technologies in Russia., *Quantum Science and Technology* 4 (4): 040501. https://doi.org/10.1088/2058-9565/ab4472.

Fowler, A.G., and S.J. Devitt. (2012). *A Bridge to Lower Overhead Quantum Computation*. arXiv:1209.0510.

Fowler, A.G., S.J. Devitt, and C. Jones. (2013). Surface code implementation of block code state distillation. *Scientific Reports* 3: 1939. https://doi.org/10.1038/srep01939.

Fowler, A.G., and C. Gidney. (2018). *Low overhead quantum computation using lattice surgery*. arXiv:1808.06709.

Fowler, A.G., M. Mariantoni, J.M. Martinis, and A.N. Cleland. (2012). Surface codes: towards practical large-scale quantum computation. *Phys. Rev.* A 86: 032324.

Foxen, B, J.Y. Mutus, E Lucero, R Graff, A. Megrant, Y. Chen, C. Quintana, et al. (2017). Qubit compatible superconducting interconnects. *Quantum Science and Technology* 3 (1): 014005. https://doi.org/10.1088/2058-9565/aa94fc

Gehrz, R.D., T.L. Roellig, M.W. Werner, G.G. Fazio, J.R. Houck, F.J. Low, G.H. Rieke, B.T. Soifer, D.A. Levine, and E.A. Romana. (2007). The NASA Spitzer Space Telescope. *Review of Scientific Instruments* 78: 011302. https://doi.org/10.1063/1.2431313.

Gidney, C., and M. Ekera. (2019). *How to factor 2048-bit RSA integers in 8 hours using 20 million noisy qubits*. arXiv:1905.09749.

Harrow, A.W., A. Hassidim, and S. Lloyd. (2009). Quantum algorithm for solving linear systems of equations. *Phys. Rev. Lett.* 103: 150502.

Herr, D., F. Nori, and S.J. Devitt. (2017). Lattice surgery translation for quantum computation. *New Journal of Physics* 19 (1): 013034. https://doi.org/10.1088/1367-2630/aa5709.

Herr, D., A. Paler, S.J. Devitt, and F. Nori. (2018). Lattice surgery on the Raussendorf Lattice. *Quantum Science and Technology* 3 (3): 035011. https://doi.org/10.1088/2058-9565/aac450.

Horsman, C., A.G Fowler, S.J. Devitt, and R. Van Meter. (2012). Surface code quantum computing by lattice surgery. *New Journal of Physics* 14 (12): 123011. https://doi.org/10.1088/1367-2630/14/12/123011.

Jones, N.C., R. Van Meter, A.G. Fowler, P.L. McMahon, J. Kim, T.D. Ladd, and Y. Yamamoto. (2012). Layered architecture for quantum computing. *Phys. Rev. X* 2: 031007.

Kane, B.E. (1998). A silicon-based nuclear spin quantum computer. *Nature* 393: 133–37.

Kielpinski, D., C. Monroe, and D.J. Wineland. (2002). Architecture for a large-scale ion-trap quantum computer. *Nature* 417: 709–11.

Kim, J., and C. Kim. (2009). Integrated optical approach to trapped ion quantum computation. *Quant. Inf. Comput.* 9: 181–202.

Kitaev, A.Y. (1995). *Quantum measurements and the Abelian Stabilizer Problem*. quant-ph/9511026.

Knight, P., and I. Walmsley. (2019). UK National Quantum Technology Programme. *Quantum Science and Technology* 4 (4):040502. https://doi.org/10.1088/2058-9565/ab4346.

Kwon, S., A. Tomonaga, G.L. Bhai, S.J. Devitt, and J.S. Tsai. (2020). Gate-based superconducting quantum computing. *Journal of Applied Physics* 129 (4), 041102

Lee, T., M. Santha, and S. Zhang. (2020). *Quantum algorithms for graph problems with cut queries*. arXiv:2007.08285.

Lekitsch, B., S. Weidt, A.G. Fowler, K. Mølmer, S.J. Devitt, C. Wunderlich, and W.K. Hensinger. (2017). Blueprint for a microwave trapped ion quantum computer. *Science Advances* 3 (2): e1601540. https://doi.org/10.1126/sciadv.1601540.

Litinski, D., and F. von Oppen. (2018). Lattice surgery with a twist: simplifying Clifford Gates of surface codes. *Quantum* 2 (May): 62. https://doi.org/10.22331/q-2018-05-04-62.

Lloyd, S. (1996). Universal quantum simulators. *Science* 273: 1073–78.

Magnard, P., S. Storz, P. Kurpiers, J. Schar, F. Marxer, J. Lutold, J.-C. Besse, et al. (2020). *Microwave Quantum Link between Superconducting Circuits Housed in Spatially Separated Cryogenic Systems*. arXiv:2008.01642.

Markov, I.L., and Y. Shi. (2008). Simulating quantum computation by contracting tensor networks. *SIAM J. Comput.* 38 (3): 963–81. https://doi.org/10.1137/050644756.

Matteo, O.D., V. Gheorghiu, and M. Mosca. (2020). Fault-tolerant resource estimation of quantum random-access memories. *IEEE Trans. on Quantum Engineering* 1: 1–13 https://doi.org/10.1109/TQE.2020.2965803

Meter, R. Van, and S.J. Devitt. (2016). The path to scalable distributed quantum computing. *Computer* 49 (9): 31–42. https://doi.org/10.1109/MC.2016.291.

Metodi, T.S., D.D. Thaker, A.W. Cross, F.T. Chong, and I.L. Chuang. (2005). A quantum logic array microarchitecture: scalable quantum data movement and computation. MICRO-38.

Monroe, C., R. Raussendorf, A. Ruthven, K.R. Brown, P. Maunz, L.-M. Duan, and J. Kim. (2012). *Large Scale Modular Quantum Computer Architecture with Atomic Memory and Photonic Interconnects*. arXiv:1208.0391.

Mukai, H., K. Sakata, S.J. Devitt, R. Wang, Y. Zhou, Y. Nakajima, and J.-S. Tsai. (2020). Pseudo-2D superconducting quantum computing circuit for the surface code: proposal and preliminary tests, *New J. of Physics* 22 (4): 043013. https://doi.org/10.1088/1367–2630/ab7d7d.

Munro, W. J., K.A. Harrison, A.M. Stephens, S.J.Devitt, and K. Nemoto. (2010). From quantum multiplexing to high-performance quantum networking. *Nature Photonics* 4 (11): 792–96. https://doi.org/10.1038/nphoton.2010.213.

Munro, W. J., A.M. Stephens, S.J. Devitt, K.A. Harrison, and K. Nemoto. (2012). Quantum communication without the necessity of quantum memories. *Nature Photonics* 6 (11): 777–81. https://doi.org/10.1038/nphoton.2012.243.

Nayak, C., S.H. Simon, A.Stern, M. Freedman, and S.D. Sarma. (2008). Non-Abelian anyons and topological quantum computation. *Reviews of Modern Physics* 80 (3): 1083–1159. https://doi.org/10.1103/RevModPhys.80.1083.

Nemoto, K., M. Trupke, S.J. Devitt, A.M. Stephens, B. Scharfenberger, K. Buczak, T. Nöbauer, M.S. Everitt, J. Schmiedmayer, and W.J. Munro. (2014). Photonic architecture for scalable quantum information processing in diamond. *Physical Review X* 4 (3): 031022–22. https://doi.org/10.1103/PhysRevX.4.031022.

Nickerson, N.H., Y. Li, and S.C. Benjamin. (2013). Topological quantum computing with a very noisy network and local error rates approaching one percent. *Nature Communications* 4 (1): 1756. https://doi.org/10.1038/ncomms2773.

Niedermayr, M. (2015). *Towards Cryogenic Surface Ion Traps*. Ph.D. Thesis. https://diglib.uibk.ac.at/ulbtirolhs/content/titleinfo/508936/full.pdf.

Nielsen, M.A., and I.L. Chuang. (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge U.P.

O'Brien, J.L. (2007). Optical quantum computing. *Science* 318 (5856): 1567. https://doi.org/10.1126/science.1142892.

Oi, D.K.L., S.J. Devitt, and L.C.L. Hollenberg. (2006). Scalable error correction in distributed ion trap computers. *Physical Review A* 74 (5): 052313–13. https://doi.org/10.1103/PhysRevA.74.052313.

Paler, A., D. Herr, and S.J. Devitt. (2019). Really small shoe boxes: on realistic quantum resource estimation. *Computer* 52 (6): 27–37. https://doi.org/10.1109/MC.2019.2908621.

Paler, A., and S.J. Devitt. (2018). Specification format and a verification method of fault-tolerant quantum circuits. *Physical Review A* 98 (2): 022302–2. https://doi.org/10.1103/PhysRevA.98.022302.

Paler, A., S.J. Devitt, and A.G. Fowler. (2016). Synthesis of arbitrary quantum circuits to topological assembly. *Scientific Reports* 6 (1): 30600. https://doi.org/10.1038/srep30600.

Paler, A., S.J. Devitt, K. Nemoto, and I. Polian. (2014). Mapping of topological quantum circuits to physical hardware. *Scientific Reports* 4 (1): 4657. https://doi.org/10.1038/srep04657.

Paler, A., S.J. Devitt, K. Nemoto, and I. Polian. (2012). Synthesis of topological quantum circuits. In *Proceedings of the 2012 IEEE/ACM International Symposium on Nanoscale Architectures*, 181–87. NANOARCH '12. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2765491.2765524.

Pednault, E., J.A. Gunnels, G. Nannicini, L. Horesh, and R. Wisnieff. (2019). *Leveraging Secondary Storage to Simulate Deep 54-qubit Sycamore Circuits*. arXiv:1910.09534.

Perez-Garcia, D., F. Verstraete, M.M. Wolf, and J.I. Cirac. (2007). Matrix product state representations. *Quantum Info. Comput.* 7 (5): 401–30.

Raussendorf, R., J. Harrington, and K. Goyal. (2006). A fault-tolerant one-way quantum computer. *Ann. Phys.* 321: 2242–70.

Raymer, M.G, and C. Monroe. (2019). The US National Quantum Initiative. *Quantum Science and Technology* 4 (2): 020504. https://doi.org/10.1088/2058-9565/ab0441.

Reagor, M., C.B. Osborn, N. Tezak, A. Staley, G. Prawiroatmodjo, M. Scheer, N. Alidoust, et al. (2018). Demonstration of universal parametric entangling gates on a multi-qubit lattice. *Science Advances* 4 (2). https://doi.org/10.1126/sciadv.aao3603.

Reiher, M., N. Wiebe, K.M. Svore, D. Wecker, and M. Troyer. (2017). Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences* 114 (29): 7555. https://doi.org/10.1073/pnas.1619152114.

Riedel, M., M. Kovacs, P. Zoller, J. Mlynek, and T. Calarco. (2019). Europe's Quantum Flagship Initiative. *Quantum Science and Technology* 4 (2): 020501. https://doi.org/10.1088/2058-9565/ab042d.

Roberson, T.M., and A.G. White. (2019). Charting the Australian quantum landscape. *Quantum Science and Technology* 4 (2): 020505. https://doi.org/10.1088/2058-9565/ab02b4.

Rohde, H., S.T. Gulde, C.F. Roos, P.A. Barton, D. Leibfried, J. Eschner, F. Schmidt-Kaler, and R. Blatt. (2001). Sympathetic ground-state cooling and coherent manipulation with two-ion crystals. *J. Opt. B: Quantum Semiclass.* Opt. 3: S34.

Rosenberg, D., D. Kim, R. Das, D. Yost, S. Gustavsson, D. Hover, P. Krantz, et al. (2017). 3D integrated superconducting qubits. *Npj Quantum Information* 3 (1): 42. https://doi.org/10.1038/s41534-017-0044-0.

Rudolph, T. (2017). Why I am optimistic about the silicon-photonic route to quantum computing. *APL Photonics* 2 (3): 030901. https://doi.org/10.1063/1.4976737.

Saffman, M. (2018). Quantum computing with neutral atoms. *National Science Review* 6 (1): 24–25. https://doi.org/10.1093/nsr/nwy088.

Schaetz, T., D. Leibfried, J. Chiaverini, M.D. Barrett, J. Britton, B. Demarco, W.M. Itano, J.D. Jost, C. Langer, and D.J. Wineland. (2004). Towards a scalable quantum computer/simulator based on trapped ions. *Appl. Phys. B* 79.

Shor, P.W. (1994). Polynomial time algorithms for discrete logarithms and factoring on a quantum computer. In Proc. *Algorithmic Number Theory. First International Symposium*, ANTS- i, 289.

Steane, A.M. (2007). How to build a 300 Bit, 1 Gflop quantum computer. *Quant. Inf. Comput.* 7: 171–83.

Stenholm, S. (1986). The semiclassical theory of laser cooling. *Rev. Mod. Phys.* 58: 699.

Stephens, A.M., Z.W.E. Evans, S.J. Devitt, A.D.Greentree, A.G. Fowler, W.J. Munro, J.L. O'Brien, K. Nemoto, and L.C.L. Hollenberg. (2008). Deterministic optical quantum computer using photonic modules. *Phys. Rev. A* 78: 032318.

Stephenson, L.J., D.P. Nadlinger, B.C. Nichol, S. An, P. Drmota, T.G. Ballance, K. Thirumalai, J.F. Goodwin, D.M. Lucas, and C.J. Ballance. (2020). High-rate, high-fidelity entanglement of qubits across an elementary quantum network. *Physical Review Letters* 124 (11): 110501–1. https://doi.org/10.1103/PhysRevLett.124.110501.

Stock, R., and D.F.V. James. (2009). A scalable, high-speed measurement-based quantum computer using trapped ions. *Phys. Rev. Lett.* 102: 170501.

Sussman, B., P. Corkum, A. Blais, D. Cory, and A. Damascelli. (2019). Quantum Canada. *Quantum Science and Technology* 4 (2): 020503. https://doi.org/10.1088/2058-9565/ab029d.

Takita, M., A.W. Cross, A.D. Córcoles, J.M. Chow, and J.M. Gambetta. (2017). Experimental demonstration of fault-tolerant state preparation with superconducting qubits. *Physical Review Letters* 119 (18): 180501–1. https://doi.org/10.1103/PhysRevLett.119.180501.

Vahidpour, M., W. O'Brien, J. Tyler-Whyland, J. Angeles, J. Marshall, D. Scarabelli, G. Crossman, et al. (2017). *Superconducting through- silicon vias for quantum integrated circuits*. arXiv:1708.02226.

Walther, A., F. Ziesel, T. Ruster, S.T. Dawkins, K. Ott, M. Hettrich, K. Singer, F. Schmidt-Kaler and U. Poschinger. (2012). *Controlling Fast Transport of Cold Trapped Ions*. arXiv:1206.0364.

Weidt, S., J. Randall, S.C. Webster, K. Lake, A.E. Webb, I. Cohen, T. Navickas, B. Lekitsch, A. Retzker, and W.K. Hensinger. (2016). Trapped-ion quantum logic with global radiation fields. *Physical Review Letters* 117 (22): 220501–1. https://doi.org/10.1103/PhysRevLett.117.220501.

Wright, K., J.M. Amini, D.L. Faircloth, C. Volin, S.C. Doret, H. Hayden, C.-S. Pai, et al. (2012). *Reliable Transport Through a Microfabricated x- Junction Surface-Electrode I on Trap*. arXiv:1210.3655.

Yamamoto, Y., M. Sasaki, and H. Takesue. (2019). Quantum information science and technology in Japan. *Quantum Science and Technology* 4 (2): 020502. https://doi.org/10.1088/2058-9565/ab0077.

Zhang, Q., F. Xu, L. Li, N.-L. Liu, and J.-W. Pan. (2019). Quantum information research in China. *Quantum Science and Technology* 4 (4): 040503. https://doi.org/10.1088/2058-9565/ab4bea.

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 40–47. ISSN 0035-9173/22/01040-08

Australia is the lucky country when it comes to snakes

Christina N. Zdenek

Lab Manager and Post-doctoral Research Fellow at the Venom Evolution Lab at
The University of Queensland

E-mail: c.zdenek@uq.edu.au

Abstract

Despite the common perception worldwide that Australia is a dangerous and deadly place when it comes to snakes, Dr Zdenek begs to differ. Here are seven facts as to why: first, Australian snakes bolt away from humans; second, Australia has very few snakebite deaths; third, Australia has great access to excellent antivenom; fourth, Australia has the world's only snake venom detector kits; fifth, if you do get bitten, you're very unlikely to lose a limb; sixth, snakebite treatment in Australia is covered by Medicare; and, seventh, snake venom can save lives.

Introduction[1]

Australia has a global reputation as a land full of danger, where seemingly everything can kill you. Crocodiles lurk in tropical waters, large spiders creep in our bathrooms, we have venomous plants, and we share our suburbs with some of the most venomous snakes on the planet.

Snakes hold a particular fascination for many people in many cultures, often accompanied with fear (Polak et al. 2016). The bite of an Australian eastern brown snake (*Pseudonaja* spp.) can kill a human in under an hour (Allen et al. 2012). That's just one of more than 150 species of venomous snakes inhabiting the island continent across land and sea. Australian snakes are well and truly overrepresented out of the world's top 25 most venomous snakes (as measured on mice) (Broad et al. 1979).

Counterintuitively, Australia is relatively very lucky when it comes to snakes. Here are seven facts why.

Australian snakes bolt away from humans

The best way to survive a snakebite is of course not to be bitten (Zdenek 2021). Keeping your distance is the easiest way to avoid a bite.

But what if you're walking through the bush and don't see the snake? Luckily, most Australian snakes will rapidly slither away from us or are at least physically capable of doing so (Whitaker and Shine 1999) because their morphology permits it. With most Australian snakes being active foragers that pursue their prey, this renders them physically capable of rapidly escaping danger too, such as a human (predator) unknowingly walking toward them. In Australia, by staying still, you stay safe.

In contrast, sit-and-wait ambush predators such as rattlesnakes and vipers (e.g. in Mexico and Indonesia) are physically incapable of rapidly escaping danger. Instead, they have to hold their ground, relying on crypsis (avoiding detection by remaining still) to escape danger, and therefore can be

---

[1] Earlier versions of this paper appeared in *The Conversation* of 7 February 2022, and on the ABC's Radio National, Occam's Razor, 3 April 2022.

easily trodden on. They're physically incapable of rapidly bolting away due to their ambushing foraging mode and stout, short body morphology. Furthermore, these venomous snakes sense body heat via infra-red sensing pit organs on their face. In Australia, the only snakes with such heat-sensing ability are non-venomous pythons.

Australia has very few snakebite deaths

Compared to other snake-inhabiting countries, Australia has orders of magnitude fewer snakebites and related deaths (Gutiérrez et al. 2017). For example, South Africa, which has just 2.2 times the population of Australia, has 159 times the snakebite deaths (476) on average every year (Halilu et al. 2019). Every year in India, they average around 58,000 snake bite deaths (Laxme et al., 2021). By contrast, Australia has two or three snake bite deaths per year on average (Welton et al. 2016). Furthermore, this low death count is not merely due to the population size of countries, as is illustrated when controlling for population size by calculating the annual number of snakebite deaths per 100,000 inhabitants. India has up to 6.7 deaths per 100,000, whereas Australia has 0.13 snake bite deaths. This is because Australia has great access to high-quality care and treatment.

Australia has great access to excellent antivenom

Antivenom is the only specific treatment for snakebites (Williams et al. 2018). If you're unlucky enough to be bitten by a highly venomous snake, getting the antivenom as quickly as possible is vital. Luckily, antivenoms work quickly, and Australia's are of high quality. Australia actually pioneered the development of this life-saving medicine in the early 1900s (Winkel et al. 2006).

Antivenom is often produced from purified horse antibodies, after the horse has been dosed with a small amount of venom. It's well known that antivenom can sometimes cause anaphylaxis, which occurs around 10% of the time in Australia (Ibister et al. 2008). These reactions are much less common in modern antivenoms produced using Good Manufacturing Practices (Bush et al., 2015) and can be quickly reversed by adrenaline administered in a hospital.

By contrast, some other countries have alarmingly ineffective antivenoms, as well as triggering anaphylaxis over 50% of the time (Variawa et al. 2021; Williams et al. 2007). Moreover, many antivenoms in foreign countries fail to actually improve patient outcomes. Very few have actually undergone clinical trials, or even preclinical testing (Alirol et al., 2015). Australian antivenoms are regularly tested for quality control during manufacture by the pharmaceutical company Seqirus (part of CSL, Commonwealth Serum Laboratory) (Verity et al. 2021). Further illustrating the quality of Australian antivenoms, even compared to other rich nations, is the small average dose required (4 vials median dose, interquartile range 2–5 vials) (Isbister et al., 2008), compared to extremely high doses (dozens of vials) often required in the USA (Bush et al., 2012).

When available, specific (monovalent) antivenoms are superior to polyvalents because there are higher titre levels of specific antibodies against toxins from the offending species, thereby increasing effectiveness. Higher specific titre levels also reduces the foreign protein load (injection dose), thereby decreasing the chance for serum sickness (Williams et al. 2018). In many parts of the world (e.g. USA, Indonesia, and India), only polyvalents are available.

We're lucky in Australia to have five types of monovalent snake antivenoms for treating bites from our most dangerous groups of land-based snakes (Tiger Snake, Black Snake, Brown Snake, Taipan, and Death Adder) and one polyvalent, which is a mixture of the other five, used when antivenom choice is unclear. A sixth specific antivenom exists for sea snakes.

Antivenom is available at most (~750) major hospitals in Australia. For more remote regions, snakebite victims benefit from proven pressure-immobilisation snakebite first-aid (Sutherland et al. 1979), which should be applied before the Royal Flying Doctor comes to the rescue.

Some less lucky countries such as the USA do not have this snakebite first-aid option, due to the extensive cytotoxicity and necrosis of rattlesnake bites, which cause significant local tissue damage. Thus, patients may be worse off when arriving at hospital compared to Australian snakebite patients, due to the lack of a suitable first-aid measure being available.

Australia has the world's only snake venom detection kits

Using the wrong antivenom can lead to ineffective treatment, and the victim's snake identification is unreliable (Wolfe et al. 2020).

In 1979, Australia became the first country in the world to have a commercial snake venom detection kit to make quick antivenom choice more accurate (Knudsen et al. 2021). Even now, Australia is the only country with this option. This is probably because the kits are expensive to develop, and the people most in need of them are the ones least able to afford it. So, it's a small market.

Other countries must rely on more dangerous options. Either the victim brings the snake to hospital for a professional ID, or doctors have to rely on the patient's symptoms and the location where the patient was

Eastern Brown Snake, by the Author

bitten to take an educated guess as to which antivenom might work (Blaylock 2005).

This is a challenge because there can be extensive overlap of symptoms caused by venom from totally unrelated species (Feola et al. 2020). Plus, snakebite envenoming is very complicated, making years of experience treating snakebite often a prerequisite in correctly identifying the species responsible.

If you do get bitten, you're very unlikely to lose a limb

Snakebites in Australia are often painless. This is in part due to the short fangs of the most offending group of snake in Australia, our brown snakes (*Pseudonaja* spp.), which are responsible for most bites in Australia (Ibister et al. 2009), but mainly because most Australian snakes have venom which has little to no local effect at the bite site (White 1991). As such, snakebites in Australia very rarely result in amputations.

By contrast, across sub-Saharan Africa, amputation is unfortunately common (Chippaux 2011), with nearly 2400 amputations per year reported in Africa's most populous country, Nigeria (Halilu et al. 2019). Unfortunately, the people most at risk of snakebite, and losing a limb as a result, are the ones least able to afford the high treatment costs (Harrison et al. 2009).

Snakebite treatment in Australia is covered by Medicare

Antivenom can be prohibitively expensive (Zdenek et al. 2019), costing thousands of dollars per dose, making it out of reach for many people in poorer countries. But our snakebite treatment is covered by Medicare.

Our nearest neighbour, Papua New Guinea, is a snakebite hotspot. Yet many people simply do not have the money to pay for the antivenom, which can cost up to 60% of their annual income. As a result, in some areas in PNG, taipans kill more people than malaria, owing to high treatment costs and and distant clinics. This leads to there being 120 deaths per 1000 snakebites in PNG, whereas in Australia it's 1 in 1000. This massive disparity in survivability exists despite both countries sharing multiple very closely related venomous snake species, illustrating how fortunate we are in Australia.

In the USA, where, similar to Australia, relatively good healthcare is available, snakebite patients can be left with medical bills over $140,000. These exorbitant bills result from costly antivenom (over $3,000/vial), the high dose (dozens) of antivenom used on average to correct symptoms, and high daily cost in intensive care units.

Despite Australian snakes being much more venomous (drop-for-drop on lab mice) than American snakes (Broad et al., 1979), in Australia, treatment for a bite without medical evacuation may cost around $6,000 — antivenom costs $347–$2,320 per vial (Johnston et al. 2017), plus care — but this cost is covered by Medicare for permanent residents and citizens. For countries less fortunate, cheaper snakebite treatments are desperately needed (Gutierrez et al., 2011; Laustsen et al., 2017). In the venom lab[2] I manage, we are working to make snakebite treatment more affordable by testing next-generation snakebite treatments (Chowdhury et al. 2021). One compound (Varespladib (LY333013)) performed exceptionally well in our extensive pre-clinical

---

[2] The Venom Evolution Lab at The University of Queensland.

tests (e.g. Chowdhury et al., 2021b; Zdenek et al., 2020), thereby assisting the orally-administered drug to recently (15 Aug. 2021) progress to Phase 2 human clinical trials.

Snake venom can save lives

Snake venoms save hundreds of thousands of lives every year. There are six therapeutic drugs on the global market designed from snake venom toxins, with another three at least in clinical trials.

One excellent example is Captopril, a drug that was modelled off a toxin from a snake venom in South America (the Brazilian pit viper, the jararaca, *Bothrops jararaca*). Captopril lowers high blood pressure that otherwise can result in heart disease, the world's biggest killer[3] (Opie & Kowolik, 1995). Since its development, it has led to an entirely new class of drugs (ACE inhibitors) being developed to treat heart disease.

Australian venomous snakes may also contribute to the therapeutic drug market. A toxin from the venom of eastern brown snakes (*Pseudonaja textilis*) is currently being tested as a drug to reverse life-threatening bleeding complications in patients on Direct Oral Anticoagulants (DOACs) for the prevention of stroke and deep vein thrombosis (Verhoef et al. 2017). What's more, this same venom has recently been used in a gauze delivery scaffold called Snake Venom Hydrogels (Yegappan et al. 2022) as a novel rapid wound sealant. Australia's many venomous snake species hold in their venom glands a "mini drug library" for scientists to trawl through for new life-saving drugs.

Conclusion

The Australian and global perspective on Australian snakes and snakebite is largely negative and undeserving.

Australian snakes pose little risk to humans: they flee from approaching humans, their bites can usually be treated quickly, and, counterintuitively, their venom holds therapeutic promise. Furthermore, snakes play a vital role in controlling populations of introduced rats and mice — vermin that can have devastating financial impacts on crops and communities (Brown & Singleton, 2000). Ticks and fleas live on those vermin can also increase incidence of disease and pet death (Fearn et al. 2002).

Perspective and attitudes alter human behaviour (Eiksund 2009). Persecution of snakes by humans is a relatively common practice in Australia, which stems from fear. This fear, which is not evidence-based and fails to consider the points made herein, increases one's risk to snakebite and also reduces populations of an otherwise important taxon that serves as predators and prey in ecosystems.

Rather than harming snakes, we are better off appreciating and respecting Australia's wealth of venomous snakes.

References

Alirol, E., Lechevalier, P., Zamatto, F., Chappuis, F., Alcoba, G., Potet, J. (2015) Antivenoms for snakebite envenoming: What is in the research pipeline? *PLoS Negl. Trop. Dis.* 9, 1–11.

Allen G.E., Brown S.G.A., Buckley N.A., O'Leary M.A., Page C.B., Currie B.J., White

---

[3] In the late 1960s John Vane of England and Sérgio Henrique Ferreira of Brazil found that one of the viper's venom's peptides selectively inhibited the action of angiotensin-converting enzyme (ACE), which was thought to function in blood pressure regulation; the snake venom functions by severely depressing blood pressure. Captopril, an analogue of the snake venom's ACE-inhibiting peptide, was first synthesized in 1975. In 1982 Vane was awarded a Nobel Prize. [Ed.]

J., Isbister G.K., for the ASP Investigators (2012) Clinical effects and antivenom dosing in brown snake (*Pseudonaja* spp.) envenoming — Australian Snakebite Project (ASP-14) *PLOS One*, https://doi.org/10.1371/journal.pone.0053188.

Blaylock R.S. (2005) The identification and syndromic management of snakebite in South Africa, *South African Family Practice* 47(9): https://doi.org/10.1080/20786204.2005.10873288.

Broad A.J., Sutherland S.K., Coulter A.R. (1979) The lethality in mice of dangerous Australian and other snake venom, *Toxicon* 17(6): 661–664. https://doi.org/10.1016/0041-0101(79)90245-9.

Brown, P.R. & Henry, S. (2000) Impacts of house mice on sustainable fodder storage in Australia. *Hum. Conflicts with Wildl. Econ. Considerations* 6, 48–58.

Bush S.P., Seifert S.A., Pearl S.R. (2012) Continuous infusion of crotalidae polyvalent Fab (ovine). in "A Patient with Rattlesnake Evenomation." Presented at Biology of the Rattlesnakes, Tucson Az, 2011. Published in Bush S.P., Case study — rattlesnake envenomation. In: Crotaline Snakebite Management: the role of CroFab Crotalidae Polyvalent immune Fab (ovine). Btg international, inc. 2013.

Bush, S.P., et al. (2015) Comparison of F(ab')2 versus Fab antivenom for pit viper envenomation: A prospective, blinded, multicenter, randomized clinical trial. *Clin. Toxicol.* 53: 37–45.

Chippaux J.-P. (2011) Estimate of the burden of snakebites in sub-Saharan Africa: A meta-analytic approach, *Toxicon* 57(4): 586–599. https://doi.org/10.1016/j.toxicon.2010.12.022.

Chowdhury A., Zdenek C.N., Lewin M.R., Carter R., Jagar T., Ostanek E., Harjen H., Aldridge M., Soria R., Haw G., Fry B.G. (2021) Venom-induced blood disturbances by palearctic viperid snakes, and their relative neutralization by antivenoms and enzyme-inhibitors. *Frontiers in Immunology* 12, 10 June. https://doi.org/10.3389/fimmu.2021.688802.

Chowdhury, A., Lewin, M.R., Zdenek, C.N., Carter, R. & Fry, B. G. (2021) The relative efficacy of chemically diverse small-molecule enzyme-inhibitors against anticoagulant activities of African Spitting Cobra (*Naja* Sp.) venoms. *Front. Immunol.* 12: 1–9.

Eiksund, S. (2009) A geographical perspective on driving attitudes and behaviour among young adults in urban and rural Norway. *Safety Science*. 47: 529–536.

El-Aziz, T.M.A., Soares A.G., and Stockand J.D. (2019) Snake venoms in drug discovery: valuable therapeutic tools for life saving, *Toxins*, 11(10): 564. https://doi.org/10.3390/toxins11100564.

Fearn S., Robinson B., Sambono J. and Shine R. (2002) Pythons in the pergola: the ecology of "nuisance" carpet pythons (*Morelia spilota*) from suburban habitats in south-eastern Queensland, *Wildlife Research* 28(6): 573–579. https://doi.org/10.1071/WR00106.

Feola A., Marella G.L., Carfora A., Della Pietra B., Zangani P., Campobasso C.P. (2020) Snakebite envenoming a challenging diagnosis for the forensic pathologist: a systematic review, *Toxins* (Basel) 12(11): 699. https://doi.org/10.3390/toxins12110699.

Gutiérrez, J.M., León, G. & Burnouf, T. (2011) Antivenoms for the treatment of snakebite envenomings: The road ahead. *Biologicals* 39: 129–142.

Gutiérrez J., Calvete J., Habib A. et al. (2017) Snakebite envenoming. *Nat. Rev. Dis. Primers* 3: 17063. https://doi.org/10.1038/nrdp.2017.63.

Halilu S., Garba I., Hamza M., Chippaux J.-P., Kuznik A., Habib A. (2019) Snakebite burden in Sub-Saharan Africa: estimates from 41 countries, *Toxicon* 159(3): 1–4. https://doi.org/10.1016/j.toxicon.2018.12.002.

Harrison R.A., Hargreaves A., Wagstaff S.C., Faragher B., Lalloo D.G. (2009) Snake envenoming: a disease of poverty, *PLOS Neglected Tropical Diseases*. https://doi.org/10.1371/journal.pntd.0000569.

Isbister G.K., Brown S.G., MacDonald E., White J., Currie B.J.; Australian Snakebite Project Investigators (2008) Current use of Australian snake antivenoms and frequency of immediate-type hypersensitivity reactions and anaphylaxis. *Med. J. Aust.* 188(8): 473–6. https://doi.org/10.5694/j.1326-5377.2008.tb01721.x.

Isbister G.K., Duffull S.B., Brown S.G., ASP Investigators (2009) Failure of antivenom to improve recovery in Australian snakebite coagulopathy. *QJM* 102(8): 563–8. https://doi.org/10.1093/qjmed/hcp081.

Johnston, C.I. et al. (2017) The Australian snakebite project, 2005–2015 (ASP-20). *Med. J. Aust.* 207: 119–125.

Knudsen C., Jürgensen J.A., Føns S., Haack A.M., Friis Rasmus U.W., Dam S.H., Bush S.P., White J., Laustsen A.H. (2021), Snakebite envenoming diagnosis and diagnostics, *Frontiers in Immunology* 12, 2021 https://doi.org/10.3389/fimmu.2021.661457.

Laustsen, A.H., Johansen, K.H., Engmark, M. & Andersen, M.R. (2017) Recombinant snakebite antivenoms: A cost-competitive solution to a neglected tropical disease? *PLoS Negl. Trop. Dis.* 11, 1–14.

Laxme, R.R.S., Khochare, S., Attarde, S., Suranse V., Iyer A., Casewell N.R., Whitaker R., Martin G., Sunagar K. (2021) Biogeographic venom variation in Russell's viper (*Daboia russelii*) and the preclinical inefficacy of antivenom therapy in snakebite hotspots. *PLoS Negl. Trop. Dis.* 15, e0009247.

Polák, J., Sedláčková, K., Nácar, D., Landová, E. & Frynta, D. (2016), Fear the serpent: A psychometric study of snake phobia. *Psychiatry Res.* 242, 163–168.

Sutherland S.K., Coulter A.R., Harris R.D (1979) Rationalisation of first-aid measures for elapid snakebite, *Lancet* 1(8109): 183–5. https://doi.org/10.1016/s0140-6736(79)90580-4.

Variawa S., Buitendag J., Marais R., Wood D., Oosthuizen G. (2021), Prospective review of cytotoxic snakebite envenomation in a paediatric population, *Toxicon* 190: 73–78, https://doi.org/10.1016/j.toxicon.2020.12.009.

Verhoef D., Visscher K.M., Vosmeer C. et al. (2017) Engineered factor Xa variants retain procoagulant activity independent of direct factor Xa inhibitors. *Nat. Commun.* 8, 528. https://doi.org/10.1038/s41467-017-00647-9.

Verity, E.E., Stewart, K., Vandenberg, K., Ong, C. & Rockman, S. (2021), Potency testing venoms and antivenoms in embryonated eggs: An ethical alternative to animal testing. *Toxins* (Basel) 13: 1–13.

Welton R.E., Williams D.J., Liew D. (2016) Injury trends from envenoming in Australia, 2000–2013. *Internal Medicine Journal* 47(2): 170–176. https://doi.org/10.1111/imj.13297.

Whitaker P.B. and Shine R. (1999) Responses of free-ranging brownsnakes (*Pseudonaja textilis*: Elapidae) to encounters with humans, *Wildlife Research* 26(5): 689–704. https://doi.org/10.1071/WR98042.

White J. (1991) Snakebite: an Australian perspective, *Journal of Wilderness Medicine* 2(3): 219–244. https://doi.org/10.1580/0953-9859-2.3.219.

Williams, D.J., Jensen, S.D., Nimorakiotakis, B., Müller, R. & Winkel, K.D. (2007) Antivenom use, premedication and early adverse reactions in the management of snake bites in rural Papua New Guinea. *Toxicon* 49, 780–792.

Williams D.J., Habib A.G., Warrell D.A. (2018) Clinical studies of the effectiveness and safety of antivenoms. *Toxicon* 150: 1–10. https://doi.org/10.1016/j.toxicon.2018.05.001.

Winkel, K.D., Mirtschin, P. & Pearn, J. (2006), Twentieth century toxinology and antivenom development in Australia. *Toxicon* 48, 738–754.

Wolfe A.K., Fleming P.A., Bateman P.W. (2020) What snake is that? Common Australian snake species are frequently misidentified or unidentified, *Human Dimensions of Wildlife* 25(6): 517–530. https://doi.org/10.1080/10871209.2020.1769778.

Yegappan, R. et al. (2022) Snake Venom Hydrogels as a rapid haemostatic agent for uncontrolled bleeding, *Adv. Healthc. Mater.* 2200574, 22000574.

Zdenek C.N., Hay C., Arbuckle K., Jackson T.N.W., Bos M.H.A., Op den Brouw B., Debono J., Allen L., Dunstan N., Morley T., Herrera M., Gutiérrez J.M., Williams D.J., Fry B.G. (2019) Coagulotoxic effects by brown snake (*Pseudonaja*) and taipan (*Oxyuranus*) venoms, and the efficacy of a new antivenom. *Toxicol In Vitro* 58: 97–109. https://doi.org/10.1016/j.tiv.2019.03.031.

Zdenek, C. N., Youngman N.J., Hay C., Dobson J., Dunstan N., Allen L., Milanovic L., Fry B.G. (2020) Anticoagulant activity of black snake (Elapidae: *Pseudechis*) venoms:

Mechanisms, potency, and antivenom efficacy. *Toxicol. Lett.* 330: 176–184.

Zdenek C.N. (2021) How to survive a venomous snakebite — from a professional who's been bitten before. ABC Science, 30 Oct 2021.

❧

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 48–71. ISSN 0035-9173/22/01048-24

A safe governance space for humanity: necessary conditions for the governance of Global Catastrophic Risks

Len Fisher[a] **and Anders Sandberg**[b]

[a] School of Physics, University of Bristol, UK; FRSN

[b] Future of Humanity Institute, Oxford Martin School, University of Oxford, UK

[a] E-mail: Len.Fisher@bristol.ac.uk

Abstract

The world faces a multiplicity of global catastrophic risks (GCRs), whose functionality as individual and collective complex adaptive networks (CANs) poses unique problems for governance in a world that itself comprises an intricately interlinked set of CANs. Here we examine necessary conditions for new approaches to governance that take account of the known properties of CANs — especially, that small changes in one part of the system can cascade and amplify throughout the system, and that the system as a whole can also undergo rapid, dramatic and often unpredictable change with little or no warning.

Introduction[1]

Many governance schemes have been proposed for the management of global catastrophic risks, defined by Bostrom & Ćirković (2008) as situations that "have the potential to inflict serious damage to human well-being on a global scale." We argue here that most of these schemes suffer from a fatal logical flaw, in that they begin with a favoured system of governance and attempt to apply it to the world situation, rather than examining the world situation and asking what system of governance might be most appropriate. Here we analyze some of the major schemes that have been proposed, and ask how they stack up against the criteria required for governance in the face of real-world complexity.

Our argument is developed in four steps:

1. A brief review of global catastrophic risks (GCRs) and their governance
2. Conceptual framing of our social-economic-ecological world and the threats that endanger it as complex adaptive networks (CANs)
3. Analysis of the necessary conditions for the effective governance of GCRs as CANs
4. Evaluation of different proposed forms of governance in terms of those necessary conditions.

1. Principles for governance of Global Catastrophic Risks (GCRs)

Bostrom & Ćirković's definition of GCRs states that they must be "serious," but without defining this term. Here we adopt a criterion suggested by the authors themselves,

---

[1] Reprinted under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0). Original citation: Fisher, L. & Sandberg, A. (2022) A Safe Governance Space for Humanity: Necessary Conditions for the Governance of Global Catastrophic Risks. *Global Policy*, 00, e13030. Available from: https://doi.org/10.1111/1758-5899.13030

that a GCR is serious if its consequences are likely adversely to affect tens of millions of people, or to cost trillions of (U.S.) dollars.

Many current or looming events fit these conditions. We have culled a (non-exhaustive) list of those that are believed by many authors to be among the most important from the World Economic Forum *Global Risks Report* (World Economic Forum, 2020), the Global Challenges Foundation *Global Catastrophic Risks 2020* (Global Challenges Foundation, 2020), the Stockholm Resilience Centre review *Planetary Boundaries: Exploring the Safe Operating Space for Humanity* (Rockström *et al.* (2009), the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES, 2019), the Intergovernmantal Panel on Climate Change 6th Assessment Report (IPCC, 2021) and Toby Ord *The Precipice* (Ord, 2020). Our criteria for inclusion are that the risk has been identified by multiple sources as being among the most important, and that claims for its catastrophic nature are based on hard evidence.

We will show that much of the power of each threat derives from its being a component of a complex network, whose other members include the individuals, communities and environments that are under threat. The individual threats are also linked with each other to form an over-arching network whose governance must be considered as a whole.

Our non-exhaustive list comprises:

Internal threats

1. Climate change
2. Loss of biodiversity
3. Degrading environment and resource depletion
4. Food insecurity
5. Pandemics
6. Population increase and urban expansion
7. Collapse of international governance
8. Unaligned artificial intelligence
9. Cyber risks
10. Increasing polarization of societies
11. Rising disparity of income and wealth
12. Weapons of mass destruction
13. Great power war
14. Genocidal totalitarianism
15. Runaway technological disasters.

External threats

1. Asteroid impact
2. Supervolcanic eruptions
3. Geomagnetic storms generated by solar superflares.

1.1 Governance principles for GCRs

Any successful governance scheme for GCRs must take into account their variability in scope, severity and probability (Avin *et al.*, 2018). There are strong arguments (Ord, 2020) for giving high priority to existential risks, even those with relatively low probability. As an aid to prioritization, Bostrom (2013) has proposed a "rule of thumb" *maxipok* principle: *Maximise the probability of an "OK outcome," where an OK outcome is any outcome that avoids existential catastrophe.*

Bostrom points out that this principle, although superficially similar to the well-known *maximin* principle ("choose the action that has the best worst-case outcome") is in fact quite different in outcome. The maxipok principle promotes relevant action, while the maximin principle is open to the interpretation that, in the face of existential risk, "we ought all to start partying as if there were no tomorrow."

The maximin principle nevertheless has some merit for lesser, but still catastrophic, risks, so long as there is enough information for the best worst-case outcome to be reliably assessed (e.g. Bognar, 2011; Sunstein, 2019). If this is not the case, then the *precautionary principle* comes into play. The principle has been formulated in a number of different ways (references in Clarke, 2005), and may be exemplified by the closing Ministerial Declaration from the United Nations Economic Conference for Europe in 1990, which states that "When there are threats of serious or irreversible damage, lack of full scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation" (quoted in Sunstein, 2007).

The precautionary principle has been the subject of extensive philosophical and political debate (Read & O'Riordan, 2017). Failure to apply it at the start of the COVID-19 pandemic may have been responsible for many excess deaths (Basili, 2019), but its application later in the pandemic, when the dangers of the AstraZeneca vaccine came into question, may also have resulted in excess deaths (Faranda *et al.*, 2021). Clarke (2005) also points out that the precautionary principle, as commonly formulated, leads to a paradox. It suggests, for example, holding back on "risky" research in some areas. But what if that research provides the only route to an eventual solution?

Sunstein (2007) has suggested a stronger form of the principle, in the form of the *Catastrophic Harm Precautionary Principle*: "When risks have catastrophic worst-case scenarios, it makes sense to pay special attention to those risks, even when existing information does not enable regulators to make a reliable judgement about the probability that the worst-case scenarios will occur."

One way of paying special attention to catastrophic risks is what Turchin (2018) calls the "Plan A, Plan B" model. In this dual approach, Plan B is "a backup option, implemented if Plan A fails. In the case of global risks, Plan A is intended to prevent a catastrophe and Plan B to survive it … ." Turchin claims that this model has "shown its effectiveness in planning actions in unpredictable environments." Other models that make similar claims are those based on resilience (Folke *et al.*, 2010), sustainability (Burch *et al.*, 2019), or the primacy of human rights (Voeneky 2019).

A similar, but more subtle, scheme has been proposed by Cotton-Barratt *et al.* (2020) as a "Defence in Depth" against human extinction. In this scheme, three sequential layers of protection provide a defensive structure, in the manner of the concentric defences of a mediaeval castle (Faulkner, 1963). The layers here are Prevention, Response and Resilience, with the inner layer of resilience especially acting to prevent global catastrophes from becoming extinction catastrophes.

All of these schemes, and others that have been suggested in the very large literature on global catastrophic risks (cf. Baum & Handoh, 2014; Baum & Barrett, 2018; Galaz, 2019) and global systemic risks (Centeno *et al.*, 2015) come with question marks as to *when* they should be implemented and to *how* they should be implemented. Sunstein (2007), for example, admits that the Catastrophic Harm Precautionary Principle is "lamentably vague" in these regards. It does not "specify the threshold information that would trigger the principle; the role of costs; and how regulators should incorporate whatever information exists about the probability of catastrophe."

Faber (2011) offers a specific framework in response to these questions. According to this framework, schemes for the management of catastrophic risks must fulfil the following ten practical requirements:

1. facilitate modelling of the considered system such that all relevant events leading to losses may be represented together with their interdependencies
2. consistently account for the level of available knowledge as well as natural variability
3. facilitate decision making at a scale of system representation necessary to support the decisions in question
4. quantify risks in a marginal as well as a non-marginal sense; i.e., be able to represent the effect of losses due to a given event on economic growth and the living conditions for future generations
5. specifically address decision making in the situations before, during and after hazard events
6. facilitate standardised procedures for systems representations in risk assessments
7. account for information which might become available in the future and facilitate that options for future decisions are included in the decision optimisation
8. facilitate for consistent risk aggregation whereby it is ensured that the results of independently performed risk assessments can be applied to assess and manage the risk in larger context-portfolios
9. facilitate decision optimisation and the assessment of the acceptability of decisions
10. enhance risk communication and risk management documentation.

These general principles for risk management apply to all types of system. We now show how they emerge naturally as general principles for the governance of complex adaptive systems (CAS).

2. Conceptual framing of the world and GCRs as complex adaptive networks

The conceptual framing of our global socio-(economic)-ecological system (SES) as a complex adaptive network (CAN), in which the components interact in non-linear ways, with many positive and negative feedback loops, was initiated in the 1990s (Pohl, 1999). It has since been put on a firm footing (Ostrom, 2009; Schweitzer *et al.*, 2009; Levin *et al.*, 2013; Sayama *et al.*, 2013; Levin, 2019). The typical features of such a system (Chan, 2001; Helbing, 2013; Sayama *et al.*, 2013; Pattberg & Widerberg, 2019) are:

- Connectivity (the system forms a network)
- Self-organization and strong correlations dominate the system behaviour, and elements can co-evolve, based on their interactions with other elements and the environment
- Distributed control (no single centralized control mechanism, so that opportunities for external or top-down control are very limited)
- Sensitive dependence on initial conditions (a small change in one part of the system can lead to large (often unpredictable) changes in other parts). When change *does* happen, the system might show numerous

different behaviours (multiple equilibria), depending on the respective initial conditions

- Emergent order — the behaviour of the system cannot be understood or predicted just by understanding the behaviour of the individual elements (Miller & Page, 2007).

Our socio-economic-ecological world displays all of these features (Pohl, 1999; Ostrom, 2009; Schweitzer *et al.*, 2009; Levin *et al.*, 2013; Sayama *et al.*, 2013; Levin, 2019). Its individual members (people, societies, ecosystems, economies, plants, animals, oceans, atmosphere, etc) interact either directly or indirectly, and change over time as a result of these interactions. There is no central control of these interactions. A small change in one part of the system (collapse of a bank or the eating of a bat) can lead to dramatic, system-wide changes (financial collapse/pandemic). It is usually impossible to predict the long-term effects of the behaviour of the individual members of the system.

The governance system itself can be a complex system in its own right (e.g. international law (Kim & Mackey, 2014)), and is also a part of the larger complex adaptive system In terms outlined by George Soros (2013) and placed into the context of complex adaptive systems by Eric Beinhocker (2013), it is *fallible* and *reflexive*. *Fallible*, because the complexity of the world that we are trying to govern exceeds our capacity to understand it. *Reflexive*, because the governance system is an active participant in the system that it is trying to govern. Thus, any governance actions are liable to feedback and affect the governance system itself. According to Beinhocker, such a reflexive system has two additional elements that distinguish it from a normal dynamic feedback system:

- *Internal model updating*: The internal decision model of the agents [governance systems] is not fixed, but can itself change in response to interactions between the agent and its environment [the system to be governed]
- *Complexity*: The system has *interactive complexity* due to multiple interactions between heterogeneous agents, and *dynamic complexity* due to nonlinearity in feedbacks in the system.

2.1 GCRs as CASs

GCRs themselves form an interconnected network (Fisher, 2019) that has all the characteristics of a CAN (Levin *et al.*, 2013). Global warming, for example, is connected to food security, with longer growing seasons meaning that pests can increasingly survive between seasons. Our evolving choice of food, on the other hand, may affect global warming (Wilett *et al.*, 2019). Food insecurity can even drive revolution and war (Lagi *et al.*, 2011), which affect food supplies in their turn.

Each threat has an internal structure which makes it a complex network (global warming, for example, involves many interlinked chemical, physical and social processes, with multiple feedback loops). Each network also has most or all of the characteristics of a complex *adaptive* network (CAN) (Table 1). The assembly of networks also forms a super-complex adaptive network, whose governance must be considered as a whole.

A relatively clear-cut example is provided by the pandemic spreading of the COVID-19 virus, which produces CAS and CAN dynamics. As shown in Figure 1, there are

Table 1: Examples of how different GCR threats have CAN features embedded

	Connectivity	**Self-organisation**	**Distributed control**	**Sensitive dependence**	**Emergent order**
Climate change	Complex feedbacks between atmosphere, geosphere, biosphere, anthroposphere	Strong couplings, biosphere and economy respond to climate events and policy	Multilateral international political/economic system	Chaotic weather/disaster responses; system tipping points	Yes
Loss of biodiversity	Ecosystem nutrient web; ecosystem services; economy	Ecosystems and economies change as a response (adaptation, mitigation, restoration); transnational cooperation and spillovers	Species, farmers, industry, governments	Keystone species; biodiversity hotspots; multiple equilibria (e.g. kelp forest/urchin barren)	Yes
Degrading environment and resource depletion	International economy; global supply chains	Economic price responses; technological substitution and adaptation	Industry, government	Alternative economic or technology paths; innovation	Yes
Food insecurity	Connected energy-agriculture-distribution system; economy as a whole	Price shocks; farmer and societal adaptation	Crops, pests, farmers, industry, governments	Climate and pest drivers; choice of agricultural system; import/export policy choices	Yes
Pandemics	International transport network; social interaction network	Ongoing research and experience changes response; infodemics; literal evolution of pathogens	Individual and societal decisions, international organisations	Exponential pathogen growth; policy decisions; successful or unsuccessful containment	Yes
Population increase and urban expansion	Networked demographic factors (urban economics; health systems; education); culture	Demographic and urban network effects (e.g. costs of child-rearing, urban economies of scale); cultural shifts	Individual and local decision-making. Rare cases of top–down control (interacting with individual choices in complex ways, e.g. China one-child policy, planned cities)	-	Yes
Collapse of international governance	International governance norms, treaties, laws, and relationships	Cascade effects; formation or dissolution of international organisations or alliances	Sovereign states	International crisis events; formation/breakup of alliance constellations	Yes
Unaligned artificial intelligence	(Varies depending on scenario)	Self-improving technology; instrumental goal convergence	Software, programmers, companies, governments	Intelligence amplifies probability of desired goal states from low-probability states	?

(continued)

Table 1: Examples of how different GCR threats have CAN features embedded *(continued)*

	Connectivity	**Self-organisation**	**Distributed control**	**Sensitive dependence**	**Emergent order**
Cyber risks	Internet; economics of cyber actors	Technological, legal, and economic responses	Software, programmers, companies, governments	Exploit detection; technological and security regime choices; liability and insurance rules	Yes
Increasing polarisation of societies	Social media networks; social networks	Sociological group dynamics; online and offline community construction	-	-	Yes
Rising disparity of income and wealth	Economic network	Rich-get-richer-dynamics; redistribution policies	-	-	Yes
Weapons of mass destruction	(International alliance network a driver)	Arms race dynamics	(Sovereign states and UN)	Technological discovery and availability; different control regimes	?
Great power war	(International alliance network a driver)	Strategic and tactical interactions; responses in all affected parts of society	(Sovereign states)	Inciting events causing escalation; randomness of war	Yes
Genocidal totalitarianism	Hierarchical social networks	-	-	-	?
Runaway technological disasters	Varies	Varies	Varies	Varies	Varies
Asteroid impact	-	Disaster response	-	Timing and size determines location and consequences (e.g. land-based fires, urban disaster, or tsunami)	Yes (in human disaster response)
Supervolcanic eruptions	Environment; food system	Disaster response	-	-	Yes (in human disaster response)
Geomagnetic storms generated by solar superflares	Power grid networks	Disaster response	-	-	Yes (in human disaster response)

several positive feedback loops producing accelerating change and sensitivity to initial conditions, but also inhibitory feedback allowing for bi-stability and oscillation. Control is distributed among numerous actors who update their behaviour based on their partial understanding of the system. Strength of interaction can change, different subsystems can overlap, and external factors can feed into the dynamics unexpectedly.

Beyond this simple model, COVID-19 has had obvious outside knock-on effects, such as the cancellation of sports tournaments, closure of restaurants, restriction of travel,

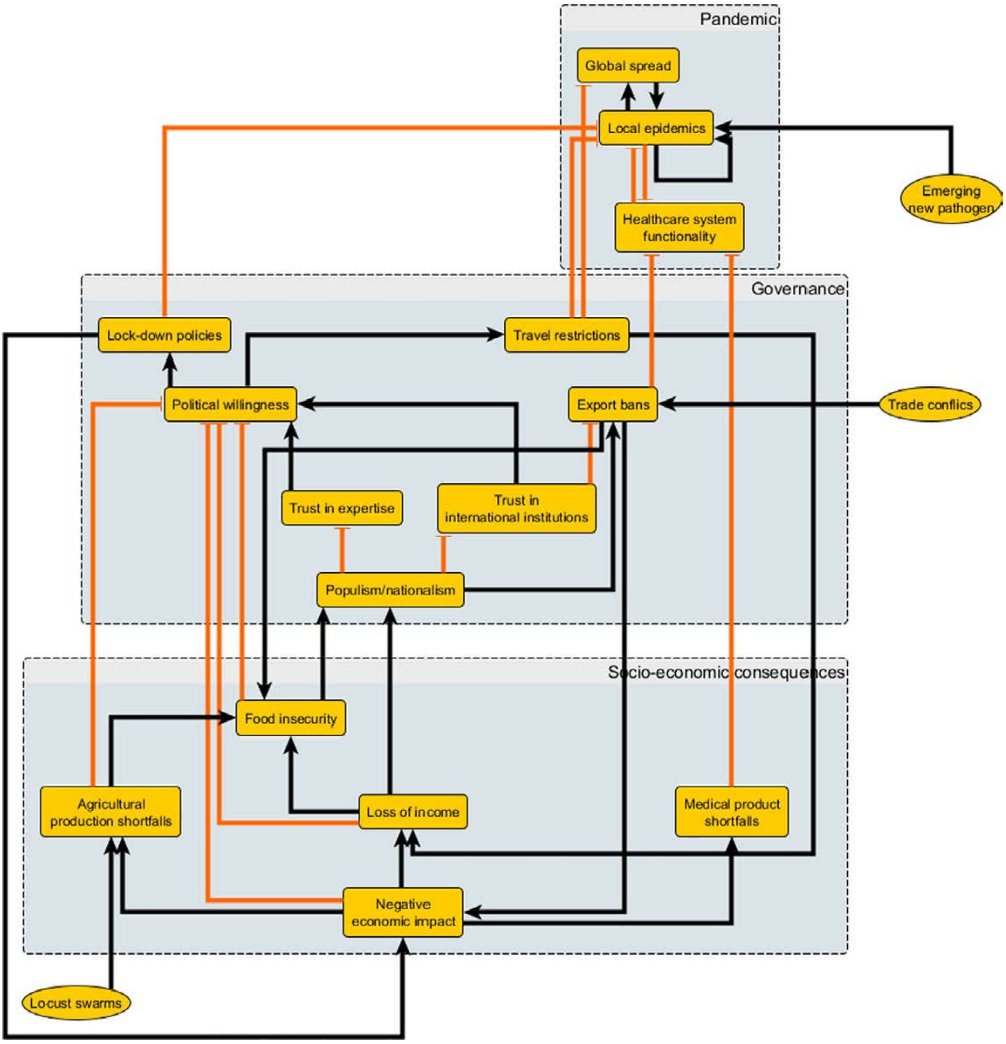


Fig.1. Simple CAN feedback model of part of the COVID-19 pandemic system. Black lines indicate amplifying impact, red lines indicate inhibiting impact. Figure 1 suppresses the spatial and organizational dimensions: most factors are actually clusters of linked but separate (sub/inter)national factors.

social isolation, and loss of income for many small businesses (Haleem *et al.*, 2020). But there are many less obvious connections, including reduced carbon dioxide emissions (Anjum, 2020), an increase in endangered sea turtle nesting and hatchling survival as beaches remain clear of people and rubbish (Luscombe, 2020), and interlinked disruptions to the global economy that could even lead to the reversal of globalization and large consequent shifts in the economic power base (Baldwin & di Mauro, 2020).

The COVID-19 pandemic has also generated considerable mistrust in governance structures across the world, whose behaviour has had to change in response (Garrett, 2020). Early overconfidence in some cases was later used as evidence of lack of knowledge among authorities, leading to maladaptive public or institutional responses, such as reluctance to use masks, to vaccinate or to increase testing capacity.

The consequences of these responses fed back into the systems in the form of greater infection rates, leading to further cycles of more vigorous action (such as compulsory mask-wearing) and stronger public responses to these actions (such as public demonstrations). These are examples of reflexivity in action, and reinforce our point that the governance of GCRs cannot be considered in isolation, but only in the overall context of governance of CANs, and especially an awareness that ongoing feedback loops are always liable to offer a potential for instabilities.

3. Necessary conditions for the governance of GCRs

3.1 Unsuitability of current governance structures

As the examples above reveal, GCRs constitute a unique challenge to governance. Klinke (2014) argues that "the key peculiarities of global risks — complexity, scientific uncertainty, and socio-political ambiguity — are … generic features" and that "there is a lack of a broader societal and political consensus of how to handle this kind of insecurity." Silja Voeneky (2019) offers many concrete examples, from artificial intelligence to gene editing, and points out that "Thus far, no international treaty on existential and global catastrophic risks and scientific research exists" and that, in general "international treaty law is not sufficient to govern these research areas."

Tom Pegram and Julia Kreienkamp (2019) argue that the major problem is that legacy governance structures, such as the UN Security Council or the General Agreement on Tariffs and Trade, are designed for the administration of *complicated* problems, which "may have many components, but the relationships between the components are fixed and clearly defined" so that "a rules-based governing framework is appropriate to establish order and control" because "cause and effect relationships are linear such that … we can identify a clear cause for each observed effect and predict system-level outcomes of each change."

Complicated problems, they say, are however quite different from *complex* problems, where "The relationship between cause and effect is non-linear and effects are usually the result of several interacting causes. Due to feedback loops, we cannot establish clear cause-and-effect relationships or predict system-level outcomes."

Andy Haldane, then chief economist of the Bank of England, made this point in a speech delivered to the Peterson Institute for International Economics (Haldane, 2017). He showed that the global financial system behaves as a CAS, and that "Complex systems exhibit tipping points, with small changes in parameter values capable of moving the system from stability to collapse. In complex webs, the failure of two identical-looking banks can have very different implications for financial system stability. *The radical uncertainty in such complex webs generates emergent behaviour which can be near-impossible to predict, model and estimate*" [our emphasis].

Haldane went on to argue that traditional governance systems, which are based on prediction, modeling and estimating, are ill-suited to the governance of the world's financial networks, and that a new approach must be sought. The same argument applies to GCRs.

Adriana Abdenur (2020), writing for the Global Challenges Foundation, argues that "rather than inventing new governance mechanisms from scratch, the most effective and legitimate route for dealing with unknown (or little understood) risks is to strengthen the existing global governance system." We believe that this latter approach, unfortunately still being used by many governments and international organizations such as the UN (to which Abdenur is an advisor), is ill-conceived in principle and dangerous in practice.

3.2 A fresh start: key conditions for effective governance of GCRs

We argue that human society *does* need new governance mechanisms, better suited to handling the catastrophic risks that it now faces. We examine here the necessary conditions for the governance of such risks in the light of their behaviour as CANs, and then analyze the types of governance system best adapted to implementing those principles.

Our list derives from our considerations of GCRs as CANs. We have identified five necessary conditions for their governance. These may not be sufficient, and indeed there may be more, but these five at least are necessary for effective governance.

3.2.1 Recognition

Successful governance must consist in maximizing the chances of the best outcomes while preparing for the worst. An effective governance system must be "epistemically humble" about what it can predict and control. Unfortunately, human nature seeks certainty (Kruglanski & Orehek, 2012), which means that incentives in governance have generally favoured avoiding uncertainty, and that politicians and other decision-makers have tended to overclaim their degree of control. The feedback following inevitable failure is another example of both fallibility and reflexivity in governance.

The first and obvious requirement for the effective governance of GCRs is *recognition* that the traditional goals of certainty and control are not generally achievable (Makridakis & Taleb, 2009). In particular, the risks involved are not usually susceptible to traditional methods of top-down governance, the governance system itself forms part of the network (Kooiman, 2003), and the governance system may even be a threat to stability on its own account (Keohane, 2001).

This is the opposite of the traditional concept of "legibility," the approach of viewing a system to be governed in simplistic, orderly terms that make it governable (Scott, 1999). In real life, this still-common approach (reflected in the common political demand to provide explanations that can fit on a single sheet of paper):

1. looks at a complex and confusing reality;
2. fails to understand the subtleties of how the complex reality works;
3. attributes that failure to the irrationality of the system being looked at;
4. comes up with an idealized version of how it *ought* to look; and
5. uses authoritarian power to impose that vision, demolishing the old reality (Rao, 2010).

Scott provides many real-life examples; the reader can no doubt furnish more of this very common approach to governance, which is exemplified by the history of changing approaches to mask-wearing during the Covid-19 pandemic, with politicians frequently imposing simplistic "solutions" on what is a confused and complex reality (McConnell & Stark, 2021).

3.2.2 Flexibility and speed

Because CANs can undergo rapid, irreversible, dramatic change with little or no warning, effective governance requires *flexible, rapid decision-making processes* that can respond to and cope with such changes.

Ashby's Law of Requisite Variety (Ashby, 1958) suggests that this can only be achieved if the governance system has more potential variety than the system to be governed. Peters *et al.* (2019) argue that this need not be the case, and point to simple strategies such as that of Balinese rice farmers ("copy your most successful neighbours") that have enabled them to survive the vicissitudes of politics and war over centuries. Gigerenzer and his group (2001) have provided evidence for the success of such simple ("heuristic") approaches. Perhaps Ashby's Law should be replaced by the not-quite-equivalent "The only way to control your destiny is to be more flexible than your environment" (Dawson, 2012). Requisite variety is just one way to achieve such flexibility, but a more effective way may be to concentrate on just a few key issues or decision points where change can be implemented rapidly.

Rate factors are certainly important in many cases, especially when one part of the system cannot keep up with the rate of change in another part and loses the previous relation to it. One example is soil carbon-temperature feedback, where rapid warming causes CO_2 release, and possibly the collapse of thermohaline circulation in the deep ocean (Ashwin *et al.*, 2012). In governance itself, there are numerous examples when governance does not or cannot keep up with change or overshoots change, as with the governance of climate change (Victor, 2011), and the resistance to "lockdown" measures in some parts of the United States during the COVID-19 pandemic (Sevastopulo & Shubber, 2020; Pellis *et al.*, 2020).

A useful illustrative example is offered by Simon Levin (2019). "Many corals and barnacles," he says, "have evolved rigid structures that resist strong flows, whereas the bull kelp bend with the flows. In our societies, as in the marine environment, rigid design and robust components may work best over the short term; but a flexible adaptive component, either bending with the flow or involving replaceable components, can prolong persistence. The right balance between them varies from organism to organism, and from strategy to strategy."

Rate factors become important in a different way when considering the speed at which computer-aided decision making can take place. "Speed-ups appear to pose a serious challenge to human ability to control technological processes due to growing gaps of speed between computation and control ("cybernetic gaps") and challenges to setting the goals they are optimizing for due to gaps of speed between computation and the human world ("ethical gaps"), in turn posing a profound challenge to governance systems that are themselves to some extent hybrid human- computational systems suffering internal speed gaps" (Sandberg, 2019).

3.2.3 Integrated monitoring and action

Successful application of Ashby's Law (or any simpler version) requires the *ability to monitor* the ongoing behaviour of the network and its interactions and to *act* on this information. Clearly not everything can be known, but it is important at least to capture key features that can serve as a guide to action.

For example, if we can predict that something (e.g. atmospheric CO_2 concentrations) will have an effect (climate change), then we can focus governance on that something. As an extreme example, if all contacts of infected persons in a pandemic can be traced, they can be isolated and the spread of the disease brought under control (Keeling *et al.*, 2020).

Monitoring the structure of the network itself can also help with effective governance. It may help to avoid tipping points (known technically as *critical transitions*) through guiding changes in the organization of the connections, as could have been the case with the global financial crisis of 2008 (May & Haldane, 2011). Even under conditions of deep uncertainty, monitoring can still be valuable in setting limits on the number and type of scenarios that need to be considered (Walker *et al.*, 2010).

Reflexivity may appear to be a fundamental limit to monitoring, causing an infinite regress of considering the consequences of monitoring. However, many existing engineering systems accurately take into account their own predictions using e.g. adaptive control theory and Bellman's equations (Bellman, 1961). This is possible because they typically do not aspire to perfection, merely a high level of practical optimality. The reflexivity problem is by no means easy, but it is not unsolvable if one is willing to work with approximations.

Some of the deepest uncertainties can occur when stochastic internal variability triggers a shift in the state of a system. There may be a complete lack of warning (Lenton, 2013), and actions during rapidly changing situations (such as the occurrence of a new pandemic) must be taken "on the hoof." Integrated monitoring and action are especially important during such scenarios.

Sometimes, however, there can be warning signs. Bifurcation tipping points, for example, are often preceded by *critical slowing down* (Scheffer *et al.*, 2009)), where the system becomes more and more sluggish in its response to small perturbations and disruptions. It is important to monitor and respond to such warning signs before a "runaway" situation develops. This can require significant changes in governance culture. As the history of actions to cope with climate change has shown (Harrison &

Geyer, 2019), it can be difficult to persuade policy-makers to take warning signals seriously until it is too late. Also, the interpretation of some early-warning signs may be subject to the prosecutor's fallacy: "conditionally selecting systems known to experience a transition of some sort and failing to account for the bias that this introduces" (Boettiger & Hastings, 2012).

Another change in culture concerns care in the use of metrics. Once an indicator is made into a policy target, it can lose the information content that qualifies it to play its role as an indicator (Newton, 2011). This effect (known as Goodhart's Law) is particularly relevant to the governance of CANs, since indicators and the system reciprocally affect each other (Manheim, 2016, 2018). Therein lies the problem, since "Complex systems can only be managed using metrics, and once the metrics are put in place, everyone is being incentivized to follow the system's logic, to the exclusion of the original goals. If you're not careful with your metrics, you're not careful with your decisions. And you can't be careful enough" (Manheim, 2018). A prime example is the failure of the algorithm for modifying UK examination results in 2020 (Hao, 2020).

These various *caveats*, however, are not arguments against the use of integrated monitoring and action as a support for effective governance. They illustrate, rather, the importance of using the information gained in a precise and accurate manner.

3.2.4 Cooperation and coordination

It hardly needs saying that achieving the necessary monitoring and action requires *cooperation* and *coordination* at individual, group and international levels. The principles underlying effective cooperation have been the subject of numerous studies, with action often being sadly restricted by Underdal's "Law of the Least Ambitious Programme" (Victor, 2006), which says that action tends to be restricted by the least enthusiastic party.

Cooperation and coordination are nevertheless necessary for the governance of GCRs, since flexibility and speed are generally unachievable without them. They are especially important in three key areas:

1. taking actions that change the system to meet goals (e.g. reducing greenhouse gas emissions to mitigate climate change (Mattoo & Subramanian, 2013; Victor, 2016; Mason *et al.*, 2017)
2. taking actions that reduce uncertainty, both in practical terms (e.g. government guarantees, insurance (Louaas & Picard, 2020)) and in terms of community perceptions (Wachinger *et al.*, 2013; Kuhlemann, 2019))
3. steering the system away from tipping points (Galaz *et al.*, 2016) (e.g. reducing the reproduction number R to below 1 so as to stop the spread of a pandemic (Nouvellet *et al.*, 2021).

3.2.5 Resilience and preparedness

Finally, effective governance of global systemic risks needs to recognize that unexpected or unpredictable systemic change is always on the cards, and that dealing with such change requires *preparedness* for situations when change becomes inevitable.

When it comes to complex adaptive systems, effective preparedness for sudden change involves *investment* in resilience, which may mean investment in restoring the *status quo* and/or investment in adapt-ing to new situations (Carpenter *et al.*, 2012; Fisher, 2015).

An example of the former is resilience planning for global catastrophic biological threats such as pandemics, biological weapons and synthetic biological risks. According to Luby & Arthur (2019), resilience planning should occur at multiple levels and take several forms, including having distributed systems (e.g. urban gardens and urban farms) to provide essential food, water and power, since these are far less susceptible to cataclysmic point failure than are completely centralized systems.

Implicit in Luby and Arthur's proposal is the idea that resilience should involve protection of the current system and an eventual return to normality. This may not

Table 2: Selected proposed governance systems assessed in terms of our five necessary conditions

Proposal	**Recognition**	**Flexibility and speed**	**Integrated monitoring and action**	**Cooperation and coordination**	**Resilience and preparedness**
Act local; think global (Clemens, 2013)		+		+ (in part)	
Dispersed authority (Brosig, 2019)		+			
Multiple plausible futures (Maier et al., 2016)	+	+	+		+
Scenario planning *via* ensemble forecasting (Lempert, 2002)	1(?)		+	+	
Resilience thinking (Berkes, 2007; Folke, 2019; Folke et al., 2010)	+	+	+	+ (but not far enough)	+
Balance between positive and negative feedback; control v emergence (Choi et al., 2001)	+	+	+	+	+
Adaptive management (Allen et al., 2011)	(Depends on situation)				
Reframing decision theory for CAS (Bankes (2002))	+?	+?	+?	+?	+?
Adaptive policies for handling deep uncertainty (Walker et al., 2010)	+	+	+ (implicit)	+	+
Decision theory plus threshold approach (Polasky et al., 2011)	+		+		+
Dynamic Adaptive Policy Pathways (Haasnoot et al., 2013; Kwakkel et al., 2016)	+	+	+	+	+
Orchestrating Interactions Between Institutions (Haas 2019)				+	
Catalytic Probes (Harrison and Geyer, 2019)			+		
Sensitive Intervention Points (Farmer et al., 2018)	+	+?	+	+	+

always be possible, however, or even desirable (cf. Kareiva & Fuller. 2016), and resilience may need to involve the capacity to adapt and transform (Carpenter *et al.*, 2012).

ALLFED (The Alliance to Feed the Earth in Disasters) has considered a number of options with regard to the provision of food in the event of a natural disaster such as a massive volcanic explosion that fills the atmosphere with dust and blocks out the sunlight necessary for normal plant growth. Stockpiling, microbial electrosynthesis, scaling of greenhouse crop production to low sunlight scenarios, and the use of microbial protein are just some of the scenarios under consideration (Baum *et al.*, 2015).

Importantly, and especially because the most serious GCRs are so unpredictable, the investment in either case must be made ahead of time. Persuading those in power of this necessity is, perhaps, the most difficult problem of all.

4. Potential systems of governance

It is clear that most, if not all, current governance systems do not and cannot meet the necessary criteria as outlined above. The reasons for this have been spelled out by a number of authors (e.g. Duit & Galaz, 2008; Young, 2017; and especially papers in Galaz, 2019). Here we examine some of the major alternative governance systems that have been proposed, and ask how they stack up against our five conditions (see Table 2).

4.1 Close fits to necessary conditions

We find that three of the proposed sets of governance principles (Control *v.* Emergence, Adaptive Policies for Handling Deep Uncertainty, and Dynamic Adaptive Policy Pathways) fulfil all five of our necessary conditions, while two others (Resilience Thinking and Sensitive Intervention Points, SIPs) come very close. Here we examine them in greater detail.

4.1.1 Balance between positive and negative feedback; control v emergence (Choi et al.*, 2001)*

Thomas Choi and his colleagues point out that supply chain networks are often CANs that "emerge," rather than resulting from purposeful design by a single entity, The problems of their management/governance are thus similar in principle to those of other CANs, including GCRs, which can similarly emerge from a combination of circumstances, rather than a single identifiable cause.

The major problem identified by Choi *et al.* is selecting an appropriate balance between control and emergence. "The emergent patterns in a supply network," they argue, "can much better be managed through positive feedback, which allows for autonomous action. [But] allowing too much emergence can undermine managerial predictability and work routines [while] imposing too much control detracts from innovation and flexibility."

This general balance between control and emergence could provide a foundation for the governance of GCRs, and is compatible with our five necessary conditions.

Those in power must *recognize* that perfect certainty and control are not achievable. Continuous monitoring and consequent action are necessary to maintain the dynamic balance between control and emergence, as is flexible, rapid decision-making. Cooperation between planners and those who are responsible for implementing plans is essential. And allowance must also be made for the possibility of unexpected situations.

4.1.2 Adaptive policies for handling deep uncertainty (Walker et al.*, 2010)*

"Deep uncertainty" is defined as "The condition in which analysts do not know or the parties to a decision cannot agree upon (1) the appropriate models to describe interactions among a system's variables, (2) the probability distributions to represent uncertainty about key parameters in the models, and/or (3) how to value the desirability of alternative outcomes" (Lempert *et al.*, 2003).

The history of most, if not all, GCRs shows that they fit this description. Policy makers have a choice of how to respond. Apart from burying their heads in the sand, or maintaining a belief in an over-arching dogma and/or an ability to control, there appear to be three sensible (not necessarily exclusive) options (see Leusink & Zanting, 2009)):

- Resistance: plan for worst possible case or future situation
- Resilience: Whatever happens, make sure you can recover quickly
- Adaptation: Prepare to change the policy, in case conditions worsen.

Adaptive policies provide the flexibility required by our necessary conditions. As discussed by Walker *et al.*, they may be purposeful (planned adaptation, autonomous adaptation) or timed (anticipatory adaptation, reactive adaptation). In both cases, adaptive policies fit with our five necessary conditions. They recognize that perfect certainty and control are not achievable. By their very nature, they require integrated monitoring and action to enable flexible, rapid decision-making, and cooperation and coordination to implement those decisions over appropriate time scales. And they are able to incorporate investment in resilience and preparedness.

4.1.3 Dynamic Adaptive Policy Pathways (Haasnoot et al.*, 2013; Kwakkel, Haasnoot & Walker, 2016)*

Dynamic Adaptive Policy Pathways are a refinement of Walker *et al.*'s adaptive policies, incorporating the idea of a flexible strategic vision. They are "based on the concept that, in light of deep uncertainties about the future, one needs to design dynamic adaptive plans. Such plans contain a strategic vision of the future, commit to short-term actions, and establish a framework to guide future actions. [They are] a fusion of adaptive policymaking and adaptation tipping points."

As Figure 2 shows, they incorporate all of our necessary conditions, some directly (Recognition (1,2), monitoring and action (10), resilience and preparedness (7) and flexible, rapid decision-making (4a, 4b)), with cooperation and coordination being necessary for effective implementation of the whole process.

4.1.4 Resilience thinking (Berkes, 2007; Folke et al.*, 2010; Folke, 2019)*

Investment in resilience is one of our key conditions, but some authors believe that it can be taken further to form the foundation for governance of social-ecological systems. Here we examine whether this approach might also be appropriate for the governance of GCRs.

The underlying concept in resilience thinking is that of transformability across multiple scales. Resilience in this context (Folke *et al.*, 2010) is: "the capacity of a SES[2]

---

[2] Or any CAN — the Authors.

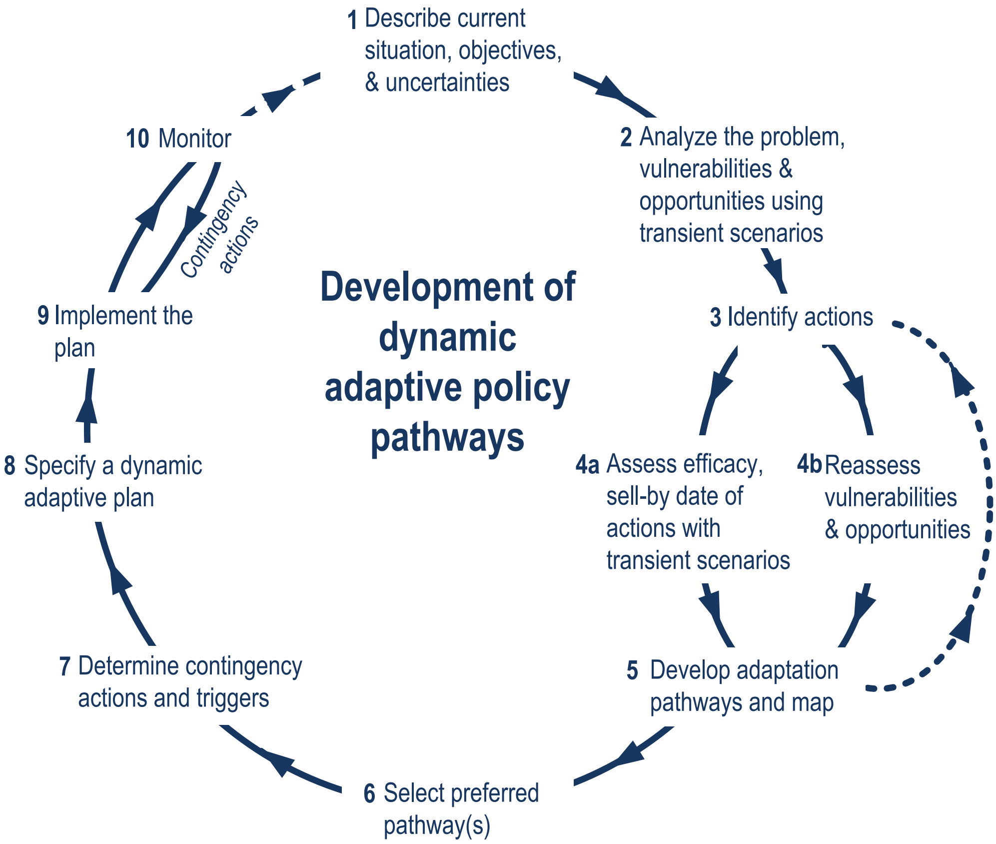


Fig. 2 Development of dynamic adaptive policy pathways (from Haasnoot et al., 2013 with permission)

to continually change and adapt yet remain within critical thresholds. Adaptability is part of resilience. It represents the capacity to adjust responses to changing external drivers and internal processes and thereby allow for development along the current trajectory (stability domain). Transformability is the capacity to cross thresholds into new development trajectories. Transformational change at smaller scales enables resilience at larger scales. The capacity to transform at smaller scales draws on resilience from multiple scales, making use of crises as windows of opportunity for novelty and innovation, and recombining sources of experience and knowledge to navigate social-ecological transitions."

Governance in this context consists of finding "ways to foster resilience of smaller more manageable SESs that contribute to Earth System resilience and to explore options for deliberate transformation of SESs that threaten Earth System resilience."

A number of strategies have been proposed for enhancing resilience in CASs (e.g. Duit, 2015; Sellberg *et al.*, 2018; Crépin, 2019). These include fostering ecological, economic and cultural diversity; planning for changes that are likely to occur; fostering learning; and communicating the societal consequences of recent changes. These strategies, and the basic concept, certainly fit our conditions 1) and 5). Remaining within stability domains also requires that conditions 2) and 3) be met. It is not so clear whether resilience thinking requires cooperation and coordination (one may envisage situations where built-in resilience through law or custom does not particularly require cooperation), but nevertheless we may consider resilience thinking to be a serious option for the governance of GCRs.

4.1.5 Sensitive Intervention Points (Farmer et al. *2019)*

SIPs are points (in time, function, or place) where "an intervention kicks or shifts the system so that the initial change is amplified by feed-back effects that deliver outsized impact."

Clearly the use of SIPs for governance requires that our conditions 2) to 4) be met. Monitoring and subsequent action are obviously essential, as is flexible, rapid decision-making and cooperation and coordination on time scales compatible with the changes to be induced.

It is possible, however, to visualize a governance system whose leaders believe in the possibility of top-down control and predictability of outcomes, but who could nevertheless use SIPs as a tool for governance. Without the recognition of GCRs as CANs, however, the effectiveness of the interventions would be a matter of luck, and interventions could even backfire (as with the introduction of cane toads for pest control in Australian canefields). Our condition 1), then, is not strictly necessary, but becomes highly desirable.

Governance solely by the use of SIPs does not strictly require investment in resilience and preparedness either (our condition 5), but such investment is highly desirable on more general grounds.

Overall, SIPs offer a very useful tool that fits our conditions 2) to 4), but which may best be used to facilitate other approaches to the governance of GCRs, particularly in the implementation of dynamic adaptive pathways.

Conclusions

We have established necessary and enabling conditions for the governance of GCR, and have examined a broad set of policy proposals in the light of these conditions. We find that *Adaptive Policies for Handling Deep Uncertainty*, as proposed by Walker *et al.* (2010), provides the most promising approach, with a *Balance Between Positive and Negative Feedbacks, Dynamic Adaptive Policy Pathways, Resilience Thinking*, and the use of *Sensitive Intervention Points* providing suitable enabling tools.

We are not aware of any existing governance system that fulfils these conditions, and argue that a totally new approach to the governance of GCR is required. This must be based on the recognition of the nature of GCRs as CANs, and of the known properties of CANs — especially that they possess emergent properties that are more than the sum of their parts, and that they are liable to sudden, unpredicted (and often unpredictable) system-wide change.

We add here that there is one further practical question. This is that *enabling conditions* must be found which will facilitate transition to the new form of governance. These conditions are *processes* that must be possible within any governance system that fulfils the five necessary conditions.

Three processes are particularly important:

- The incorporation of "bridging organizations" to connect governance levels and spatial and temporal scales (Folke, 2019)
- The evocation and maintenance of trust (Prieser & Woermann, 2019)
- Complexity leadership (Nooteboom & Teismann, 2019).

We will discuss these processes in detail, and whether they need to be modified for societies with different cultural values (Ruck *et al.*, 2020), in a subsequent paper.

Acknowledgements

LF acknowledges with gratitude the continued support of the Head and staff of the School of Physics, University of Bristol, UK. For this research and manuscript preparation, AS received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No 669751). This paper is also a contribution to the work of the Oxford Martin Programme on the Post-Carbon Transition.

Endnote

The American psychologist Frank Knight (1921) drew a distinction between *risk* ("decision situations in which probabilities are available to guide choice") and *uncertainty* ("decision situations in which information is too imprecise to be summarized by probabilities") (Runde, 1998). The "risks" that are encompassed in the phrase "global catastrophic risks" might better be described in Knightian terms as "uncertainties," since often we have no means of assessing their probabilities, or whether there are additional scenarios that we have not considered, or even been able to consider. The phrase "global catastrophic risks" is, however, now firmly embedded in the literature, and we will stay with it, clarifying where necessary any ambiguity with the Knightian meaning.

References

Abdenur, A.E. (2020) *Global Governance and Global Catastrophic Risks: Is the United Nations ready for the Anthropocene?* Global Challenges Foundation Report.

Anjum, N.A. (2020) Good in the worst: COVID-19 restrictions and ease in global air pollution. Preprints 2020, 2020040069 https://doi.org/10.20944/preprints202004.0069.v1.

Ashby, W.R. (1958) Requisite variety and its implications for the control of complex systems, *Cybernetica* 1: 83–99.

Ashwin, P., Wieczorek, S., Vitolo, R. & Cox, P. (2012) Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system, *Philosophical Transactions of the Royal Society* 370: 1166–184.

Avin, S., Wintle, J.W., Weitzdörfer, J., Heigeartaigh, S.Ö., Sutherland, W.J. & Rees, M.J. (2018) Classifying global catastrophic risks, *Futures* 107: 20–26.

Baldwin, R. & di Mauro, B.W. (eds) (2020) *Economics in the Time of COVID-19* (London: CEPR Press).

Bankes, S.C. (2002) Tools and techniques for developing policies for complex and uncertain systems, *Proceedings of the National Academy of Sciences of the USA* 99: 7263–7266.

Basili, M. (2019) The CoVid-18 pandemic outbreak: The precautionary principle and institutional failures, *Mercato Concorrenza Regole, Società editrice il Mulino* 3: 475–483.

Baum, S. D. and Barrett, A. M. (2018) Global catastrophes: the most extreme risks, in V. Bier (ed.), *Risk in Extreme Environments: Preparing, Avoiding, Mitigating, and Managing*. New York, NY: Routledge, pp. 174–184.

Baum, S.D., Dekenberger, D.C., Pearce, J.M., Robock, A. & Winkler, R. (2-15) Resilience to global food supply catastrophes, *Environment, Systems and Decisions* 35: 301–313.

Baum, S.D., & Handoh, I.C. (2014) Integrating the planetary boundaries and global catastrophic risk paradigms, *Ecological Economics* 107: 13–21.

Beinhocker, E.D. (2013) Reflexivity, complexity, and the nature of social science, *Journal of Economic Methodology* 20: 330–342.

Bellman, R.E. (1961) *Adaptive Control Processes: A Guided Tour*. Princeton: Princeton Univ. Press (republished in Princeton Legacy Library series, https://doi.org/10.1515/9781400874668.

Berkes, F. (2007) Understanding uncertainty and reducing vulnerability: lessons from resilience thinking, *Natural Hazards* 41: 283–295.

Boettiger, C. & Hastings, S. (2012) Early warning signs and the prosecutor's fallacy, *Proceedings of the Royal Society* B279: 4734–4739.

Bognar, G. (2011) Can the maximin principle serve as a basis for climate change policy? *The Monist* 94: 329–348.

Bostrom, N. (2013) Existential risk prevention as global priority, *Global Policy* 4: 15–31.

Bostrom, N. & Ćirković, M.M. (2008) *Global Catastrophic Risks*. Oxford: Oxford Univ. Press.

Brosig, M. (2019) Restricted complexity: a middle path between post-modern complexity theory and positivist mainstream IR, in Orsini, A. *et al.*, *Complex Systems and International Governance*, *International Studies Review*. pp. 1015–1019. https://doi.org/10.1093/isr/viz005.

Burch, S. *et al.* (2019) New directions in earth system governance research, *Earth System Governance* 1, 100006.

Carpenter, S.R., Arrow, K.J., Barrett, S., Biggs, R., Brock, W.A., Crépin, A-S., Engström, G., Folke, C., Hughes, T.P., Kautsky, N., Li, C-Z., McCarney, G., Meng, K., Måler, K-G., Polasky, S., Scheffer, M., Shogren, J., Sterner, T., Vincent, J.R., Walker, B., Xepapadeas, A., & Zeeuw, A-D. (2012) General resilience to cope with extreme events, *Sustainability* 4: 3248–259. https://www.doi.org/10.3390/su4123248.

Centeno, M.A., Nag, M., Patterson, T.S., Shaver, A. and Windawi, A.J. (2015) The emergence of global systemic risk, *Annual Review of Sociology* 41 (1): 65–85.

Chan, S. (2001) *Complex Adaptive Systems, ESD.83 Research Seminar in Engineering Systems* https://stuff.mit.edu/afs/athena/course/other/esd.83/OldFiles/www/notebook/Complex%20Adaptive%20Systems.pdf.

Choi, T.Y., Dooley, K.J. & Rungtusanatham, M. (2001) Supply networks and complex adaptive systems: control versus emergence, *Journal of Operations Management* 19: 351–366.

Clarke, S. (2005) Future technologies, dystopic futures and the precautionary principle, *Ethics and Information Technology* 7: 121–126.

Clemens, W.C. (2013) *Complexity Science and World Affairs*. New York: SUNY Press.

Cotton-Barratt, O., Daniel, M. & Sandberg, A. (2020) Defence in depth against human extinction: prevention, response, resilience, and why they all matter, *Global Policy* 11: 271–282.

Crépin, A.-S. (2019) Complexity, resilience and economics, in Galaz, V. (ed.) *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar, pp. 166–187.

Dawson, R. (2012) The law of requisite variety: why flexibility and adaptability are essential for success, blog post. https://rossdawson.com/blog/the-law-of-requisite-variety-why-flexibility-and-adaptability-are-essential-for-success/.

Duit, A. (2015) Resilience thinking; lessons for public administration, *Public Administration* 94, 364–380. https://doi.org/10.1111/padm.12182.

Duit, A. & Galaz, V. (2008) Governance and complexity — emerging issues for governance theory, *Governance: An International Journal of Policy, Administration and Institutions* 21: 311–335.

Faber, M.H. (2011) On the governance of global and catastrophic risks. *International Journal of Risk Assessment and Management* 15: 400. https://doi.org/10.1504/IJRAM.2011.043698.

Faranda, D., Alberti, T., Arutkin, M., Lembo, V. & Lucarini, V. (2021) Interrupting vaccination policies can greatly spread SARS-CoV-2 and enhance mortality from COVID-19 disease: The AstraZeneca case for France and Italy, *Chaos* 31, 041105, https://doi.org/10.1063/5.0050887.

Farmer, J.D., Hepburn, C., Ives, M.C., Hale, T., Wetzer, T., Meay, P., Rafaty, P., Srivastav, S. & Way, R. (2018) Sensitive intervention points in the transition to the post-carbon society, *Science* 364: 132–134.

Faulkner, P.A. (1963) Castle planning in the fourteenth century, *Archaeological Journal* 120: 215–235.

Fisher, L. (2015) More than 70 ways to show resilience, *Nature* 518: 35.

Fisher, L. (2019) Global challenges: personal reflections on the Stockholm "New Shape" competition, *Journal & Proceedings of the Royal Society of New South Wales* 151: 213–231.

Folke, C. (2019) Governing for emergence in social-ecological systems, in Galaz, V. (2019) *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar, pp. 24–37.

Folke, C., Carpenter, S.R., Walker, B., Scheffer, M., Chapin, T. & Rockström, J. (2010) Resilience thinking: integrating resilience, adaptability and transformability. *Ecology and Society* 15: 20–28.

Galaz, V., Österblom, H., Bodin, Ö & Crona, B. (2016) Global networks and global change-inducing tipping points, *International Environmental Agreements* 16: 189–221.

Galaz, V. (2019) *Global Challenges, Governance, and Complexity: Applications and Frontiers.* Cheltenham: Edward Elgar.

Garrett, L. (2020) COVID-19: the medium is the message, *The Lancet* 395: 942–943.

Gigerenzer, G., Todd, P.M. & the ABC Research Group (2001) *Simple Heuristics That Make Us Smart*. Oxford: Oxford Univ. Press.

Global Challenges Foundation (2020) *Global Catastrophic Risks 2020* https://globalchallenges.org/wp-content/uploads/Global-Catastrophic-Risks-2020-Annual-Report.pdf.

Haas, P. (2019) IR Encounters Complexity, in Orsini, A. et al., *Complex Systems and International Governance*, *International Studies Review* 1–30. https://doi.org/10.1093/isr/viz005.

Haasnoot, M., Kwakkel, J., Walker, W.E. & ter Maat, J. (2013) Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world, *Global Environmental Change* 23: 485–498.

Haldane, A. (2017) Rethinking financial stability, speech delivered to the conference, *Rethinking Macroeconomic Policy IV*, Peterson Institute for International Economics, Washington, D.C., U.S.A. https://www.bankofengland.co.uk/speech/2017/rethinking-financial-stability.

Haleem, A., Javaid, M. & Vaishya, R. (2020) Effects of COVID 19 pandemic in daily life, *Current Medicine Research and Practice* 10: 78–79. https://doi.org/10.1016/j.cmrp.2020.03.011.

Hao, K. (2020) The UK exam debacle reminds us that algorithms can't fix broken systems, *MIT Technology Review* https://www.technologyreview.com/2020/08/20/1007502/uk-exam-algorithm-cant-fix-broken-system/.

Harrison, N. & Geyer, R. (2019) The complexity of governance of climate change, in Orsini, A. et al. (2019) *Complex Systems and International Governance, International Studies Review*. pp. 1–30. https://doi.org/10.1093/isr/viz005.

Helbing, D. (2013) Globally networked risks and how to respond, *Nature* 497: 51–59.

IPBES (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, https://www.ipbes.net/system/tdf/spm_global_unedited_advance.pdf?file=1&type=node&id=35245

IPCC (2021) *Climate Change 2021: The Physical Science Basis*. https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/.

Kareiva, P. & Fuller, E. (2016) Beyond resilience: how to better prepare for the profound disruption of the Anthropocene, *Global Policy* 7: 107–118.

Keeling, M.J., Hollingsworth, T.D. & Read, J.M. (2020) The efficacy of contact tracing for the containment of the 2019 novel coronavirus

(COVID-19), *MedRxiv preprint*. https://doi.org/10.1101/2020.02.14.20023036.

Keohane, R.O. (2001) Governance in a partially globalized world, *American Political Science Review* 95: 1–13.

Kim, R.E. & Mackey, B. (2014) International environmental law as a complex adaptive system, *International Environmental Agreements* 14: 5–14.

Klinke, A. (2014) Postnational discourse, deliberation, and participation toward global risk governance, *Review of International Studies* 40: 247–275.

Knight, F.H. (1921) *Risk, Uncertainty and Profit*. Chicago: Univ. of Chicago Press.

Kooiman, J. (2003) *Governing as Governance*. London: SAGE Publications.

Kruglanski, A.W. & Orehek, E. (2012) The need for certainty as a psychological nexus for individuals and society, in *Extremism and the Psychology of Uncertainty* (M.A. Hogg & D.L. Blaylock, eds.). London: Blackwell.

Kuhlemann, K. (2019) Complexity, creeping normalcy and conceit: sexy and unsexy risks, *Foresight* 21: 35–52.

Kwakkel, J., Haasnoot, M. & Walker, W. (2016) Coping with the wickedness of public policy problems: Approaches for decision-making under deep uncertainty, *Journal of Water Resources Planning and Management* 142: 1–5. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000626.

Lagi, M., Bertrand, K.Z. & Bar-Yam, Y. (2011) The food crises and political instability in North Africa and the Middle East, arXiv:1108.2455v1 [physics.soc-ph] 11 Aug 2011.

Lempert, R.J. (2002) A new decision sciences for complex systems, *Proc. of the National Academy of Sciences of the USA* 99: 7309–7313.

Lempert, R., Popper, S. & Bankes, S. (2003) *Shaping the Next One Hundred Years: New Methods for Long-term Policy Analysis*, MR-1626-RPC, The RAND Pardee Center.

Lenton, T.M. (2013) What early warning signs are there for environmental shocks? *Environmental Science & Policy* 27S, S60–S75.

Leusink, A. & Zanting, H.A. (2009) Towards a trade-off framework for climate proofing the Netherlands, with experience from 4 case studies, https://edepot.wur.nl/15291 (in Dutch).

Levin, S., Xepapadeas, T., Crépin, A-S., Norberg, J., DeZeew, A., Folke, C., Hughes, T., Arrow, K., Barrett, S., Daily, G., Ehrlich, P., Kautsky, N., Mäler, K-G., Polasky, S., Troell, M., Vincent, J.R. & Walker, B. (2013) Social-ecological systems as complex adaptive systems: modeling and policy implications, *Environment & Development Economics* 18: 111–132.

Levin, S. (2019) The architecture of robustness, in Galaz, V. (2019) *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar, pp. 16–23.

Louaas, A. & Picard, P. (2020) Optimal insurance coverage of low-probability catastrophic risks, hal-02875534 https://hal.archives-ouvertes.fr/hal-02875534/file/Insurability%20of%20low%20probability%20catastrophic%20risks.pdf.

Luby, S. & Arthur, R. (2019) Risk and response to biological catastrophe in lower income countries, in *Global Catastrophic Biological Risks*, pp. 85–105.

Luscombe, R. (2020) Florida: endangered sea turtles thrive thanks to Covid-19 restrictions, *The Guardian*. 19 April. https://www.theguardian.com/us-news/2020/apr/19/florida-leatherback-turtles-coronavirus-beaches.

Maier, H.R., Guillaume, J.H.A., van Delden, H., Riddell, G.A., Haasnoot, M. & Kwakkel, J.H. (2016) An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling and Software* 81: 154–64.

Makridakis, S. & and Taleb, N. (2009) Decision making and planning under low levels of predictability, *International Journal of Forecasting* 25: 716–733.

Manheim, D. (2016) Goodhart's Law and why measurement is hard, *Ribbonfarm* https://www.ribbonfarm.com/2016/06/09/goodharts-law-and-why-measurement-is-hard/.

Manheim, D. (2018) *Building Less Flawed Metrics*, MPRA Paper No. 90649, posted 21 Dec 2018 https://mpra.ub.uni-muenchen.de/98288/.

Mason, C.F., Polasky, S. & Tarui, N. (2017) Cooperation on climate-change mitigation, *European Economic Review* 99, 43–55.

Mattoo, A. & Subramanian, A. (2013) A "Greenprint" for international cooperation on climate change, *World Bank Development Research Group Policy Research Working Paper* 6440 https://openknowledge.worldbank.org/bitstream/handle/10986/15581/wps6440.pdf;sequence=1.

May, R.M. & Haldane, A.G. (2011) Systemic risk in banking ecosystems, *Nature* 469: 351–355.

McConnell, A. & Stark, A. (2021) Understanding policy responses to COVID-19; the stars haven't fallen from the sky for scholars of public policy, *Journal of European Public Policy* 28: 1115–1130.

Miller, J.H. & Page, S. (2007) *Complex Adaptive Systems*. Princeton: Princeton Univ. Press.

Newton, A.C. (2011) Implications of Goodhart's Law for monitoring global biodiversity loss, *Conservation Letters* 4: 264–268.

Nooteboom, S. & Teisman, G. (2019) Leadership and complexity: can individuals make differences in complex systems?, in Galaz, V. (2019) *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar, pp. 239–273.

Nouvellet, P., …., & Donnelly, C.A. (2021) Reduction in mobility and COVID-19 transmission, *Nature Communications* 12: 1090 https://doi.org/10.1038/s41467-021-21358-2.

Ord, T. (2020) *The Precipice*. Oxford: Oxford Univ.Press.

Ostrom, E. (2009) A general framework for analyzing sustainability of social-ecological systems, *Science* 325: 419–422.

Pattberg, P. & Widerberg, O. (2019) Studying global governance as a complex system: a network perspective, in Orsini, A. et al. (2019) *Complex Systems and International Governance*, *International Studies Review*. pp. 1–30. https://doi.org/10.1093/isr/viz005.

Pegram, T. & Kreienkamp, J. (2019) *Governing Complexity: Design Principles for Improving the Governance of Global Catastrophic Risks*. Global Governance Institute Research Policy Brief. London: Univ. College London.

Pellis, L., Scarabel, F., Stage, H.B., Overton, C.E., Chappell, L.H.K., Lythgoe, K.A., Fearon, A., Bennett, E., Curran-Sebastien, J., Das, R., Fyles, M., Lewkowicz, H., Pang, X., Vekaria, B., Webb, L., House, T. & Hall, I. (2020) Challenges in control of Covid-19: short doubling time and long delay to effect of interventions, arXiv:2004.00117v1 [q-bio.PE] 31 Mar.

Peters, B.G., Pierre, J. & Galaz, V (2019) Simple solutions for complexity? in Galaz, V. (2019) *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar, pp. 188–210.

Pohl, A. (1999) Some notions of complex adaptive systems and their relationship to our world, InterSymp-99: Advances in collaborative decision-support systems for design, planning and execution (Baden-Baden, Germany, Aug 2–7), 9–24. https://digitalcommons.calpoly.edu/cadrc/34/.

Polasky, S. et al. (2011) Decision-making under great uncertainty: environmental management in an era of global change, *Trends in Ecology and Evolution* 26: 398–404.

Prieser, R. & Woermann, M. (2019) Complexity, philosophy and ethics, in Galaz, V. (ed.) *Global Challenges, Governance, and Complexity: Applications and Frontiers*. Cheltenham: Edward Elgar, pp. 38–62.

Rao, V. (2010) A big little idea called legibility, https://www.ribbonfarm.com/2010/07/26/a-big-little-idea-called-legibility/.

Read, R. & O'Riordan, T. (2017) The precautionary principle under fire, *Environment: Science and Policy for Sustainable Development* 59: 4–15.

Rockström, J. et al. (2009) Planetary boundaries: exploring the safe operating space for humanity, *Ecology and Society* 14. https://www.jstor.org/stable/26268316.

Ruck, D.J., Matthews, L.J., Kyritsis, T., Atkinson, Q.D. & Bentley, R.A. (2020) The cultural foundations of modern democracy, *Nature Human Behaviour* 4: 265–269.

Runde, J. (1998) Clarifying Frank Knight's discussion of the meaning of risk and uncertainty, *Cambridge Journal of Economics* 22: 539–546.

Sandberg, A. (2019) There is plenty of time at the bottom: the economics, risk and ethics of time compression, *Foresight* 21: 84–99. https://doi.org/10.1108/FS-04-2018-0044.

Sayama, H., Pestov, I., Schmidt, J., Bush, B.J., Wong, C., Yamanoi, J. & Gross, T. (2013) Modeling complex systems with adaptive networks, *Computers and Mathematics with Applications* 65: 1645–1664.

Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., van Nes, E.H., Rietkerk, M. & Sugihara, G. (2009) Early-warning signs for critical transitions, *Nature* 461: 53–59.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A. & White, D.R. (2009) Economic networks; the new challenges, *Science* 325: 422–425.

Scott, J. (1999) *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale: Yale Univ. Press.

Sellberg, M.M., Ryan, P., Borgström, S.T., Norström & A.V., Peterson, G.D. (2018) From resilience thinking to resilience planning: lessons from practice, *Journal of Environmental Management* 217: 906–918.

Sevastopulo, D. & Shubber, K. (2020) Trump cheers as anti-lockdown protests spread, *Financial Times* https://www.ft.com/content/c8f6f413-39c4-47ce-b1ff-0e02969cb612.

Soros, G. (2013) Fallibility, reflexivity, and the human uncertainty principle, *Journal of Economic Methodology* 20: 309–329.

Sunstein, C.R. (2007) The Catastrophic Harm Precautionary Principle, *Issues in Legal Scholarship Symposium: Catastrophic Risks: Prevention, Compensation and Recovery*.

Sunstein, C.R. (2019) Maximin, *Harvard Law School Discussion Paper 1020*. https://ssrn.com/abstract=3476250.

Turchin, A. (2018) Approaches to the prevention of global catastrophic risks, *Human Prospect* 7 N2: 53–65.

Victor, D. (2006) Toward effective international cooperation on climate change: Numbers, interests and institutions, *Global Environmental Politics* 6: 90–103.

Victor, D. (2011) *Global Warming Gridlock: Creating More Effective Strategies for Protecting the Planet*. Cambridge: Cambridge Univ. Press.

Victor, D.G. (2016) What the framework convention on climate change teaches us about cooperation in climate change, *Politics and Governance* 4: 133–141.

Voeneky, S. (2019) Human rights and legitimate governance of existential and global catastrophic risks, https://ssrn.com/abstract=3363552.

Wachinger, G., Renn, O., Begg, C. & Kuhlicke, C. (2013) The risk perception paradox — implications for governance and communication of natural hazards, *Risk Analysis* 33: 1049–1065.

Walker, W.E., Marchau, V.A.W.J. & Swanson, D. (2010) Addressing deep uncertainty using adaptive policies, *Technological Forecasting & Social Change*.

Wilett, W., et al (2019) Food in the Anthropocene: the EAT-Lancet commission on healthy diets from sustainable food systems, *The Lancet Commissions* 393: 447–492.

World Economic Forum (2020) *The Global Risks Report* (15th Edition) https://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf.

Young, O. (2017) *Governing Complex Systems. Social Capital for the Anthropocene*. Cambridge, MIT Press.

⁂

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 72–73. ISSN 0035-9173/22/01072-02

2021 Royal Society of NSW and the Learned Academies Forum: "Power and Peril of the Digital Age"

Her Excellency the Honourable Margaret Beazley AC QC

Opening Address

To all of you watching, in New South Wales and far beyond, I am delighted to join you in this opening session of the Royal Society of New South Wales and Learned Academies Virtual Conference. This morning I am speaking to you from my office here at Government House which stands on Gadigal land, and by the beautiful Gadigal waters of Sydney Harbour.

As we consider the unrelenting pace of the Digital Age over these next days, I pause to acknowledge the wisdom, culture and continuing connection to lands, waters and communities of First Nations people who have lived in the Sydney region for at least 60,000 years.[1]

Downstairs in the Main Hall here at the House, there is a large portrait of the first President of the Royal Society of New South Wales, founded 200 years ago and initially known as the Philosophical Society of Australasia. The subject of that portrait was an avid astronomer, my predecessor, Sir Thomas Brisbane, the 6th Governor of New South Wales.

It is an honour to continue this long line of Vice Regal patronage of the Royal Society, whose mission to enrich our lives through knowledge and inquiry is vigorously pursued with intellectual rigour.

The theme of this year's conference is exciting and, in many ways, confronting. People do not necessarily handle change well. Indeed, according to the eponymous Mr Google, change sits on top of the list of the top ten fears people have. Uncertainty ranks fifth. And that is in everyday life. Yet, as John F. Kennedy put it so well: *Change is the law of life. And those who look only to the past or* [*live only in the*] *present are certain to miss the future.*[2]

I commend the Society and Learned Academies Forum for this year's forward looking theme — *Power & Peril of the Digital Age* — and for the dynamic way they have sought to focus our deliberations.

We have entered a new age that has already changed and will continue to change the world and, I would add, most likely, civilisation as we know it.

The Bureau of Statistics estimates there is one birth in Australia every 1 minute and 45 seconds. So, a beautiful little bundle of joy and potential has been born whilst I have been speaking to you — born into a world of increasing digital complexity. A world, as you are going to explore, which brings both power and peril.

With a few clicks on your phone, you can access nearly all of human knowledge that exists in our time. The baby's name may have been a family favourite — or the parents might have googled 'baby names', along with some 6.9 billion searches made on the same day, using the services of a company

---

[1] https://sydneylivingmuseums.com.au/exhibitions/gadigal-place

[2] Address in the Assembly Hall at the Paulskirche in Frankfurt, June 25, 1963

that is reputed to hold a 92.7 per cent share of the global search market.[3]

As the baby grows, their immunisation status will be recorded on an app, their toddlers' clothing will be ordered online and delivered by drone, perhaps. In 2030, aged 8, the child will travel to school on an autonomous school bus, perhaps. Twenty years on, in 2050 aged 28, this tech-savvy young person will have every aspect of health care delivered remotely, will sit through international conferences in a room full of holograms, almost undoubtedly.

However, this little person, born today, also has to live in an actual world.

In 2019, the CSIRO published its Australian National Outlook[4], exploring multiple potential futures for our Nation for the next 40 years. The aim was holistic. It was to: *help Australians continue to enjoy the best quality of life available to any nation, and for future generations to have access to even better opportunities.*[5]

The report recognised the challenges:

- the Rise of Asia
- Technological Change
- Climate Change and Environment
- Demographics
- Trust, and
- Social Cohesion.

The report modelled two scenarios: a *Slow Decline* — should we fail to adequately address these issues; and an "*Outlook Vision*" that called for decisive action and a long-term view to achieve positive outcomes. The modelling demonstrated that *the difference between these two scenarios is large and worth fighting for.*[6]

Whilst the immediate future of our little birthday baby is in our hands, the longer-term future is not so certain. This uncertainty underlines that what we are discussing here over these two days is intensely pragmatic and significant.

I congratulate the Forum organisers on this collaboration between the Society and the Australian academies of Health and Medical Sciences; the Humanities; Science; Technology and Engineering; and the Academy of the Social Sciences in Australia, which creates a deep well of expertise to discuss this future which is so "worth fighting for" to get it right.

My thanks to Dr Susan Pond, President of the Royal Society of New South Wales and Chair of the Planning Committee for the Forum; the representatives of the five Learned Academies who formed the Program Committee; our two lead speakers: Australia's Chief Scientist, Cathy Foley, and NSW Chief Scientist and Engineer, Hugh Durrant-Whyte; and our Moderator, Dr Ian Oppermann, NSW Chief Data Scientist and Industry Professor, University of Technology Sydney.

It is my pleasure to officially open the Royal Society of New South Wales and Learned Academies Forum for 2021.

---

[3] https://www.salesforce.com/au/blog/2021/09/seo-best-practices.html#:~:text=Google%20is%20projected%20to%20hit,for%20a%20lot%20of%20information

[4] https://www.csiro.au/en/work-with-us/services/consultancy-strategic-advice-services/csiro-futures/australian-national-outlook

[5] https://www.csiro.au/-/media/Showcases/ANO/ANO2_MainReport_WEB_190614.pdf

[6] https://www.csiro.au/-/media/Showcases/ANO/ANO2_ExecutiveSummary_190613.pdf

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 74–75. ISSN 0035-9173/22/01074-02

2021 Royal Society of NSW and the Learned Academies Forum: "Power and Peril of the Digital Age"

Government House, Sydney
4 November 2021

Preamble[1]

We are at a moment in time when we must acknowledge and address the inexorably rising tide of data use and digital services. History will categorise the early decades of the 21st Century as the digital age, the age of prodigious development and use of digital technologies that enable us to transfer and access information easily and swiftly.

So much so that digital interaction is a defining characteristic of modern human life. Societies, economies, and political processes are infused and connected by the ubiquitous use of smart machines and software that process and communicate information to us in ways that would have been unimaginable just a few years ago. The pace of digitalisation was already fast by the end of 2019, before COVID-19 emerged.

The pandemic broke through cultural barriers and enabled implementation of digital strategies in a matter of days or weeks rather than years. Digital technologies and supercomputer simulation are central to dealing with the pandemic itself, as well as being the primary driver of productivity in almost every other aspect of society.

Companies, governments, and organisations across the world are increasingly taking advantage of the benefits associated with data analytics and simulation, artificial intelligence, and the Internet of Things to solve problems never solved before, to undertake projects in five days that would have taken five years.

Problems such as those embodied in the United Nations General Assembly's Sustainable Development Goals[2] and their achievement by 2030. Tangible benefits include greater social connectivity, learning opportunities, information access and usage, versatile working and transport, greater access to entertainment, and new forms of banking and finance.

Unlocking the power of the digital age also brings perils associated with concerns about data security, state-based and transnational crime and terrorism, complexity, privacy, social disconnection, media manipulation, distortion of the truth, communities left behind, national defence and market vulnerabilities, and outstripping rule-making and regulatory structures.

This year, the Royal Society of NSW in partnership with Australia's Learned Academies — Health and Medicine, Humanities, Science, Social Sciences, and Technology and Engineering — has chosen "Power and Peril of the Digital Age" as the theme for our annual Forum. Our goal is to have a grown-up conversation about digitalisation and the use of data. It will be framed around the

---

[1] See video of the Forum at https://www.youtube.com/watch?v=-xFGnpyAZUY&list=PLYFFwCGj2FIY-6HdbblwSrxwPw05Nf29l&index=1

[2] https://sdgs.un.org/goals

future life of a child born on the day of the Forum, 04 November 2021. This child will be born into a world of increasingly complex digital systems that holds great value and vulnerability.

Starting with a technological framing, the Forum will explore several major aspects which will impact the journey of that child as we approach 2030 and beyond. We will explore aspects of technology, health, defence and security in a digital age, and the changing nature of industry as the world and society evolve.

Finally, our annual Forum will be a call to arms for the host Societies to focus on challenges identified during these two days that must be addressed for Australia to remain a prosperous, successful, and safe democracy in the digital age.

◆

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 76–83. ISSN 0035-9173/22/01076-08

Science and technology underpinning the digital age: past, present and future

Cathy Foley[a], **Hugh Durrant-Whyte**[b]

[a] Chief Scientist of Australia
[b] New South Wales Chief Scientist and Engineer

Moderator and Rapporteur: Ian Oppermann, Chief Data Scientist, NSW Government; Industry Professor, University of Technology, Sydney

Cathy Foley

I'm speaking from the lands of the Gadigal people of the Eora nation. It's great to be able to talk to you about how digitalisation is going to affect us; but I look at it from the past as well as the present and into the future. Let's go back and think about the past. It seems very natural to assess change in my own lifetime. I was born in the 1950s, and in that decade, Australia was filled with an enormous amount of innovation. This laid the groundwork for so much that's come since then. For example, the School of the Air started in 1951; it was cutting-edge radio education at the time. Australia's first computer, the CSIRAC, provided computing services to all the CSIRO from 1951 to 1955. It was built in the late 1940s by Trevor Pearcey, Maston Beard and Geoff Hill. In 1955, Australia benefited from mass distribution of the polio vaccine, which cut an annual number of about 1,500 cases to just 125 in the first year alone. This seems sort of reminiscent of what's happening now.

In the mid-1950s, my school was only beginning to teach senior science to girls. Today, the School of the Air relies on video conferencing and has a sophisticated studio setup in Alice Springs. We've lately seen and should remember that for many months, most Australian children have also been remotely learning via video conferencing systems.

We've learnt a lot about digital deliveries of education. A defining feature of our society today is our sheer, unbelievable connectivity: with ubiquitous smartphones, computers and wi-fi in our households and in our pockets. We all probably have a mobile phone within reach right now. With that connectivity comes a level of computing power and information access that we've just never seen before. And it's not just local or personal connectivity; but our ability to speak instantly to others around the world, regardless of the location or time zone. There was a major acceleration of remote networking solutions in response to the pandemic.

Seventy years ago, we were rolling out the polio vaccine; now we are all concerned with the COVID vaccines, which were developed in nine months compared to the usual ten years to develop a vaccine. Last night we saw the Prime Minister's Science Prize awarded to Eddie Holmes,[1] who in early 2020, when the pandemic was beginning, openly shared his knowledge of the genome sequence of

---

[1] See Holmes E. C. (2021) The discovery, origins and evolution of SARS-CoV-2 (COVID-19). *Journal & Proceedings of the Royal Society of New South Wales* 154:161–181 [Ed.]

the virus, thus enabling laboratories to develop successful vaccines.

So where are we now? The future isn't written, obviously, and it's hard to really draw exact trends from the past and the present into the decades ahead. Yet I do see things in the areas of health improvements, connectivity and new ways to access the world around us. But there is a common thread above all — the incredible and ubiquitous impacts of digitalisation, which have given this century an entirely new paradigm — the digital world.

Throughout this conference, we are talking about a child who is born now and how digital technologies will impact their lives. I think it's important for us to give that person a name. I went digging around and realised that we had a ready-made whole family. The Jetsons. This television series aired in 1962, and it was set in 2062. That means that George Jetson was scheduled to be born in 2022; not today, but close enough. If we want to put a name to our child of the future, why not George Jetson?

Before you call me out, I do realise that the Jetsons is an old show. And of course, it's got very outdated gender roles, but it did lead me to thinking about things. So if George Jetson will come into the world next year, we know that he will live in a time of rapid change, tremendous connectivity and massive technological upheaval. So how do we train him to live in a world with technologies that don't exist yet? How do we prepare today's babies for careers that are still unnamed? And of course, when will we invent the Jetsons' flying car?

My key takeaway is that digital technologies are rapidly evolving way beyond some of the science fiction dreams that we had yesterday. So I'm going to consider future industries that are real and not make-believe. Every one of them relies on current and future technologies that require us to grow them if we want to see success.

The first technology is hydrogen; shipping our sunshine overseas as the newest clean renewable energy. Another is space: the Australian Space Agency is aiming to triple the size of the Australian space industry by 2030, and with that bring along 20,000 new jobs in artificial intelligence and machine learning. By 2024, only three years away, the number of these jobs is expected to grow by a quarter of a million. Digital technicians will be the largest workforce in Australia within the next few years, We're also seeing huge potential for the emerging quantum technology industry. I want to talk about that a bit more in a moment.

But even with those industries already underway, the future is much bigger than that. I'm hoping to throw some ideas out that came to me when I started wondering about the next decade or two.

The first idea is brain-machine interfacing: biological computers. Next are new ways of working in the virtual world, and the virtual commute. When digitally enabled home working is the norm, how will we adapt to this enormous change in our life patterns? Where do we draw the line between work and leisure? And will it be normal to work on a different continent to your company's head office? Will there be a head office? We've seen already that COVID massively accelerated the discussion that had happened before.

Next is automation. Will we not actually own cars in the future, but instead will we buy into a consortium or just rent a self-driving car that we call up on demand? And thinking about defence and security, what will wars of the future be like? Will there be a human pulling the trigger, or can our social media algorithms be weaponised?

Among all these hypotheticals, we need to consider digital ethics and social licence. These questions and all these future technologies won't go anywhere on their own. Whenever we talk about new technologies, we need to think about what I call science plus. We need to have science plus engineering. We need to have science plus user design and the user interface. Science plus the right business model, because we've seen that new digital technology often brings a whole new way of engaging economically. There's also science plus the regulation and policy settings that are needed for social licence. And, of course, marketing, which allows us to recognise what's available. And they all need to be talking to each other all of the time.

I want to talk about an area that I am championing: quantum. I think it's the next big industry for Australia. These technologies can do things such as accelerate drug and materials development for health care, enhance national security and support defence, increase productive mineral exploration and water resource management for mining and other sectors. Improve secure communications to industries like space and create optimisation processes, say for finance and logistics. And it's happening already, with noisy, intermediate scale quantum computers that are available via the cloud. We're seeing Airbus designing wings with these computers and Deutsche Bank using them to do develop transport efficiency algorithms.

But progress doesn't just happen by accident. Let's look at Moore's law, which suggests that the number of components in an integrated circuit doubles every 12 months. This observation was made by Gordon Moore in a paper in 1965.[2] It's a remarkably insightful paper. It also forecast that integrated circuits would lead to such wonders as home computers or at least terminals connected to a central computer; automatic controls for automobiles and personal portable communications equipment. Moore's observation was not driven by a particular scientific or engineering necessity, but it was a reflection that matched just how things happened.

The silicon chip industry took note of Moore's law and adopted it as a goal, a target for the whole industry to hit. As a consequence, we've seen that, if we want to get the most out of digital technologies, we usually need to give ourselves stretch targets to achieve. And this is something which I'm really observing right now as we're seeing a plethora of new quantum technologies being pushed out from research into industry and then just taking off.

To finish, things are changing rapidly and the challenge and the opportunity are there for us to take. In terms of social licence, we should make sure that any digital technologies we adopt, or have thrust upon us, are ones that really will make a difference for good.

Hugh Durrant-Whyte

Over the last 40 years, I've been working in the field of AI, artificial intelligence, and particularly its applications in robotics and autonomy. So I will reflect on the past and future of AI and what its implications could be; noting that AI receives an awful lot of

---

[2] Moore, G. E. (1965), 'Cramming more components onto integrated circuits' *intel.com*. Electronics Magazine, 19 April.

press coverage and is often underpinned by a lot of interesting and new technologies.

It is useful to start by understanding what AI is. It has been around a very long time. Frank Rosenblatt invented the first neural network back in 1953. So AI, as a field in neural networks, has been around now for nearly 70 years. I remember reading my first book on AI: The introduction to AI by Patrick Winston,[3] who was the first director of the AI Lab at MIT. He said, and I use this quote a great deal: "What is AI? AI is anything we currently cannot do. When we know how to do it, it's called an algorithm."

What he was really trying to say is there's nothing special about AI, different from what other sides of computer science do. And yet we've heard a lot of what I would call grandstanding around AI. Here's a particular example I sometimes like from Vladimir Putin: "AI is the future for all humankind. Whoever is the leader in this sphere will be the ruler of the world." And in contrast to that, Andrew Ng, who is a professor at Stanford and a founder of some very famous AI companies, said in *Wired*, "I worry about AI superintelligence in the same way I worry about overpopulation on Mars." It will happen one day, but in such a distant future, he's no longer concerned about it.[4]

Pragmatically, when you think about AI, it is not pixie dust, it is not some kind of magic. It is basically data. It is algorithms and the way that those are put together to solve problems and applications. I think the interesting future for AI is the great applications that are being really rethought about and the kinds of new discoveries in science and the changes that will come in the lifetime of the child who is born today. I do want to emphasise the kind of difference or disparity between what the experts in the field know about AI and what I think at the moment.

A survey that was conducted at some of the most prestigious AI conferences showed how long experts in the field think it will take for AI to reach the level of human intelligence. Interestingly, more than half of the people in the field think it will not happen in 100 years. Yes, let me repeat that: most experts think AI will not match human intelligence during the lifetime of a child born now. That's quite important because although we think of AI as intelligent, it isn't yet by a long, long way. OK? I have a little cartoon, which reflects some of my conversations with the general public. In this illustration, Wally is basically saying I built an MVP, a minimum viable product, and the pointy-head boss says, Well, that's just a block of wood. And Wally says no, I call it artificial intelligence, and the pointy-head boss says, what's his middle name? And Wally says, well, it's shy like people. And the boss says it has emotions. So you get my picture here … sometimes I think we're a little bit credulous about what AI can actually do.

And I also mention Michael Jordan, who is the most highly cited researcher in AI. He has often said that it really is not a science or anything special. It's an engineering problem and we're still at the very early stages in terms of how we actually build significant AI systems. So while, as Jordan says, the science fiction discussions about AI and superintelligence are fun, they're a distraction.

---

[3] Winston (1976) *Artificial Intelligence* Addison-Wesley.

[4] Andrew Ng: Why 'Deep Learning' is a mandate for humans, not just machines, *Wired*, 2015.

There's not been enough focus on how we build large-scale machine learning systems that really work; that deliver value to us as humans and also that do not amplify inequities. There have been 'AI winters' where people who had hyped AI gradually realised that it wouldn't work. And I suspect that we're near another AI winter now. While there have been lots of predictions for what things like neural networks will do, in the last two or three years, there've been a lot of other papers on the fundamental limitations on those kinds of approaches. I think it'll be a long time before we'll see a self-driving car in the city of Sydney without a steering wheel. Probably not in my lifetime is the answer.

So what is the AI future? AI is already changing things, but it is changing things where we've got what we call weak AI, lots of data decisions which are just yes or no. Some very predictable outcomes. What we're not really good at yet is strong AI, where there is very little data, where the decisions to be made are not obvious, and where there are very high levels of uncertainty. Truthfully, we don't even really have the mathematics or the understanding, about to how to build algorithms to manage that kind of problem.

Rod Brooks, a very famous guy in the autonomy area, says just about every successful deployment of AI is used for one of two experiences. Either there's a person in the loop or the cost of failure for a blunder is very low.[5] So we see AI in areas where, say, I'm predicting what things you will buy on the internet. While that's fine, we do not see AI out there driving cars through the city of Sydney at this point. And we probably won't in the future; we still have people sitting in the driver's seat for good reasons.

So there will be impacts of AI in autonomy and automation, and there will be impacts in job replacement and elimination of work. Possibly the most important thing is that AI is beginning to revolutionise science and society; in my view, in a very, very positive way.

Three examples show what's possible. Last September, DeepMind, a company in London owned by Google, announced that they'd won a competition to predict protein structure using an algorithm called Alphafold, which uses some AI techniques called reinforcement learning, deep reinforcement learning. This is basically applied statistics. It's not anything magic; pixie dust. But it is really interesting that Alphafold can predict with precision every protein in the human body. That is absolutely transformational. It will revolutionise medicine. And I don't think people in the medical community realise what this has just done. This has changed the whole way we will think about medicine, about health, about synthetic biology, everything in the future. So AI genuinely will have a transformational impact on discovery in this space.

There is also a robot system that I worked on; using machine learning, AI, to make discoveries in minerals. This is an area which people would not have considered 10 years ago — but now there's so much data out there and modelling information has been transformed. And there are systems out there like Obsidian, which our team developed, which will predict with accuracy the depth and geological mineralisation right across an entire continent. And there are

---

[5] Brooks, R. (2021) An Inconvenient Truth About AI. AI won't surpass human intelligence anytime soon, *IEEE Spectrum*, 29 Sept.

now many companies out there doing very similar things. Again, this will transform geology and mineral discovery.

The third case shows how we use AI to understand complex human problems. This example is from my wife, who works in this area. It's about trying to understand the drivers for disengaged and vulnerable youth and trying to predict the life courses of those people to understand what impact you can have by different types of interventions. This is not a trivial problem if there are 500 factors available for predicting what a person will do in the future. Then there are 200 to 500 possible combinations of these factors. So there are more ways or models for how a person will progress through their life than there are atoms in the universe. At the moment, our standard AI techniques cannot manage that many combinations, and indeed, in this example, all the top 100 combinations produce exactly the same predictions.

So in fact we do not know what we're doing at all in this area. But nevertheless, new types of mathematics, new types of AI, are driving our ability to use data to make informed decisions about ways that we can improve the human condition. That will have a terrific impact on the lifetime of this person born today.

Let me dispel the dystopia. I think that the Googles, the Facebooks, all of these sorts of things, are a distraction at this point. What will really happen with AI is it will transform science. It will transform discovery. It will transform the way that we work in this world; all, in my view, to positive effect.

Discussion

Prof Oppermann: These were two very different perspectives on the impacts of science and technology to a future world. Cathy highlighted some technological advancements because of the quantum leap, which will change the way we do things, yet it's important to understand the limits of what's possible. It's important for us, as we think about this future environment to consider what is possible, what's likely and in fact is not possible. Understanding what's not just is as important as being able to understand the likely scenario. Hugh, let's go back to you. You said that AI will transform the way we do discovery. What do you mean by that?

HD-W: So I think the best example is, in fact, the Alphafold one, you know, people have been trying to predict protein folding for 50 years because understanding the structure of proteins tells you what they will do and how they will work, and therefore what proteins you need to tackle if you're building a vaccine, what protein to design in lots of other areas and so on. AI is now able to predict every single protein, the structure of every single protein in the human body. What's more, it can predict the structure of any protein from any sequence of amino acids. So the whole process now of discovering vaccines, of discovering mechanisms, of discovering new types of ways that life works and so on, can be tackled using these new techniques. It is transformational.

Prof Oppermann: And does it have implications for the nature of research? The nature of investigation is the future of research. An algorithm sorts through data and puts out something interesting, and then researchers say that we will take it further or … ?

HD-W: Absolutely not. I think that's part of the challenge. In the traditional way of doing science, you get a pile of data and then search through the data to confirm or deny a hypothesis you already have. Now, the whole point about discovery is, on the contrary, to really understand the area of models, e.g. the way things fold as proteins, the way geology works as geology, the way humans evolve in terms of the environment they work in and so on, and use data to understand which of those most likely explains the data. Discovery is quite different from hypothesis testing. And I think there needs to be a transformation in science to really begin to adopt these approaches. I'm not sure science is as ready for it as the machine learners are.

Prof Oppermann: Cathy, what do you think quantum really can do and what can quantum not do, what will quantum not change? What is the arc of influence of quantum knowledge?

CF: I don't think I can answer where it can go or where it can't. Right now, we're beginning to see the evolution of laboratory projects where we were able to control quantum states, which weren't possible before, to build machines that are able to maintain them in a stable state for periods of time. But how do we create something that can do computing for us, which allows us in some cases, to do it faster? In other cases, we can actually do computations in ways which we can't currently do, which is what they call the "quantum advantage" or "quantum supremacy." At this stage, we've only had occasional demonstrations of that ability or that idea.

So there's a wide range of views as to whether that will be possible in three years, or never. Because some people feel that a quantum computer will never work; they think that you just can't get to a point where you can control quantum states, where you can do multiple calculations at the same time — which does seem like an exhilarating capacity.

Regardless, we're now seeing the ability of new technologies to control quantum states, which allows us to do better sensing. It's allowing us to do some acceleration with even our classical computers; to do computations faster. And in theory, when we're hearing about the range of possibilities in human nature … where we have choices and can choose new pathways … it's allowing us to be more than atoms of the universe.

If we were able to get a quantum computer to operate, we would have the advantage of dealing with big questions such as climate; a major topic at the moment. Climate modelling is limited by the computation power you've got in your supercomputers. So, if we were able to do more complex modelling, that would give us the potential to have sort of a way of modelling reality in a way that might be closer to the truth because you've got to remember any model is only as good as the design of the parameters you choose.

Quite often we think that models are everything but, hopefully, when we have more powerful computers, we will be able to compute models that are closer to reality. So when you ask the question of where the boundaries are, it's impossible to say. We've seen over history that when you have a lot of people who are focused on a problem putting their brains into it together, collectively, we actually see pretty significant shifts in the hardware side of things.

Prof Oppermann: What we're trying to do today is helping create a model to help us think through the next decades: 2030 to 2050 and beyond. One question from our audience is about the bias associated with AI training on white males in Western countries … a very famous facial recognition example. Let me ask our panellists: When will we understand biased algorithms and their potential harms? When will we have enough protections and understanding of how to use data similar to the way we use electricity?

HD-W: There's a big difference between AI as an algorithm and the data you put into it. AI as an algorithm is not really biased or unbiased or anything, really. It's not ethical or unethical. It's an algorithm. But it is difficult if you take an algorithm, that say classifies things, and you try it on a set of data, which comprises only one class, then you will end up with an answer which basically classifies everything into that class … if you see what I mean.

It's the data part that really raises the ethics issues rather than the algorithm. My wife is a statistician and she looks at this and she says, well, statisticians have been dealing with bias for so long, it's not funny. You know, we need to sample the complete space in a way that's representative of the space that we wish to make decisions on. If we are to end up with a regression function, a clustering algorithm which genuinely will lead to good outcomes in the future, it's not rocket science. One thing we really need to do if we're going to address bias in AI is to make sure everyone gets trained in statistics when they're at school. Then people might appreciate the way data drives decision-making processes. For example, some people published a paper where they tried to distinguish between criminals and non-criminals, and it turned out that they had trained their method to detect whether subjects were smiling or not. Those who weren't smiling were from prisons. My point is that it's good statistics that drives accuracy more than anything else.

CF: I'm a big fan of stats and I'm really concerned that we aren't teaching it enough. We're seeing too many statistics schools being closed down in universities; that's a big flag for me. There's a group of people who are really keen to get more statistics into our education system, and I think that's absolutely the way forward. We need to understand more precisely how we're making decisions and who we are considering. Is it human nature or the way we socialise? We have this truckload of unconscious bias. To have a logical way to cut through that bias is absolutely critical.

We must add transparency. A lot of things which are being done via AI systems, behind search engines or social media, involve some decision making which doesn't give us a choice or knowledge of what parameters they are using. To work towards more transparency, we need (a) understanding, so there's education, and (b) transparency, so we know what's going on. Another one is to have ethical constraints that we as a society agree on. How are we going to manage these emerging technologies and the opportunities they offer, but also make sure that they're used in ways that allow us all to feel safe and that there isn't bias that means that some are disadvantaged.

Prof Oppermann: We've been really honoured today to hear these perspectives from Australia's Chief Scientist and New South Wales Government's Chief Scientist and Engineer.

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 84–90. ISSN 0035-9173/22/01084-07

Digital lifetime of a child born today

Frances Foster-Thorpe[a], **Sue Bennett**[b]

[a] Executive Director Shaping Futures, NSW Department of Premier and Cabinet
[b] Deputy Director and Connected Child Co-Leader, ARC Centre of Excellence for the Digital Child, University of Wollongong

Moderator and Rapporteur: Ian Oppermann, Chief Data Scientist, NSW Government; Industry Professor, University of Technology, Sydney

Frances Foster-Thorpe

I join you from both a futures unit in New South Wales and a national team that's focused on how to bring together data from across Australia's governments to improve outcomes for people with disabilities. Today, I'll focus first on the digital life of the child from the perspective of the work that is currently underway in governments to improve the data that would be available for different phases of the child's life. Looking towards 2050, different efforts will need to be made to respond to the expectations that governments are hearing from communities.

For the first nine years of the life of a child who is lucky enough to be born in Australia today, the focus for governments is making sure that data is collected and then used in a way that connects the child and their family to supports that they need to thrive. There are so many potential ways that data and digital systems could be used, it seems helpful to break this down into three key information needs that we will work towards in government.

The first is that data is made available to those working directly with children and to parents, families, caregivers, in order that they can more accurately assess a child's needs. Early on, that information might be available to say, nurses and parents to help assess the child's development and whether they're hitting milestones and then perhaps later to an early educator or teacher in kindergarten and Year 1 to see how they are progressing at preschool and school. Anyone who's had a child born into their family in the last 10 years would know just how extensive the information is about the milestones and the micro milestones that each child is now expected to reach.

The second information need is data that's available to the system responsible for providing services and supports to a child and to their families so that they can see whether there's a match between those needs and the steps available. This question isn't just about the service available in the area where a child and their family live. It involves much more nuanced questions to determine whether that service is appropriate to that child's needs. So if a child has a particular type of disability, for example autism, is there a service available that's inclusive and specialist to the particular needs of that child? There's also a need for public services to understand how children and families are making their way through different systems. This requires what we often refer to as pathways analysis.

And the third information need is making data available to those analysing and researching the evidence base about how to best meet the needs of a child and their family. This is about connecting needs, services, supports and funding to the outcomes that a child, or cohort of children, can achieve.

A child born today will benefit from the much better understanding that we've obtained globally about the power of early intervention to change pathways that children are on. This is not just about government services — although I mostly think about data in those terms — it's also, crucially, about what I call informal support. This includes supports that come from family and community or supports that a family might privately pay for, as well as the value for a child of being born into an inclusive community or a democracy. Although these are more intangible things, they are really important to life outcomes. Even though we have learnt a great deal about early intervention over recent decades, I think governments come from the perspective, as do most researchers in the sector, that there is so much more to learn; particularly where children have specific needs.

For example, a child born with a particular developmental delay in a culturally diverse family. What are their particular needs? I'm focusing on the rich insights that can come from linking data from service systems and from survey data, without identifying any individual child. Government data is often subject to privacy-preserving and de-identification techniques. This is about understanding the particularities of how children and different cohorts of children move through pathways in life. There are also data systems that identify particular children and families, and they're really important to things like child protection outcomes. Australian governments have done lots of work over the past five to 10 years, to really improve those data systems that support the needs I talked about.

If this child is born in about half of Australian states or territories, they are likely to be born into a jurisdiction that has the kind of data linkage that I was talking about to some extent. But overall, it will be patchy. In the other half, they will not have access to all the service systems that are serving them. Nowhere is there really high-quality, nationally linked data. Australia is comparatively behind in the way that governments can access data systems to inform their own services and have provided access to others, such as researchers or those that provide services, as well as individuals.

We're behind comparable countries like New Zealand, the United Kingdom and Canada, basically, because we have an unusual federal split of who funds different services. Not only are we a federation, with many coordination challenges, but we're a federation where key services provided to children and families are provided across different levels of government. If you think about the first two months of this child's life, they're likely to be born into a hospital system provided by a state government. But when they go home, they'll be serviced by community nurses, also probably provided by a state government, but potentially born into an Aboriginal community, perhaps by a national Aboriginal medical service provider that will then go to a GP.

Fortunately for the child born today, they are born into a country that is massively accelerating its efforts to provide the kind of data systems that I've spoken about today. But I wanted to look forward to when the child is a teenager and then a young adult, and point out expectations that we're hearing from communities. One is that these information systems are not just the governments'. They are also for people who are making decisions about their own lives. It's not just a one-way transfer of data that governments hold, but it also could enable a rich conversation about the information that communities have about their own outcomes and what's important to them. Also, communities now demand to have a voice in how the data is being used. And we really see this, particularly in the Closing the Gap agreement recently signed by governments, where there's a commitment to sharing much more data, but also a commitment to sharing decision-making about the data.

Sue Bennett

By way of introduction, I'm going to start a story; the story of a child born today and what they're likely to experience as they grow up surrounded by technology. It's clear to all of us that a baby born today is born into an increasingly digital world. For many decades, new innovations in digital technologies have been changing the ways we live our lives, from the ways we work and learn, to the ways we communicate, the ways we socialise, the way we interact with government services such as education and health, and the ways we spend our leisure time. We were well aware of this before COVID, but perhaps we're even more aware of it now. Technology has enabled many of us to learn and work safely from home, to maintain contact with our families, friends and communities, order essentials and other things online, all while being restricted from our usual in-person activities. Through information provision across the internet contact-tracing apps, we've been able to manage some of the risks of the pandemic. Health professionals who've been able to provide care remotely have done so through the help of technology.

Although we've seen some of the downsides, such as the spread of misinformation, technology has been vital in helping us get through this so far.

So, what now for the 800 or so babies who will be born today in Australia? Well, although we might say those babies are just coming into the world for the first time now, most of them already have a digital footprint.

That footprint will have begun in a number of ways. The most familiar one is an ultrasound photo shared through social media, together with accompanying comments from family and friends. Perhaps parents used a pregnancy app to track progress in anticipation of today, and even before that, parents might have used a fertility app to help them. They can save each interaction on social media or with an app producing data. Some of that data is information that we've entered ourselves — numbers, words, images — and some of that data is automatically generated, say, from settings such as locations and time stamps.

All this data can be analysed and used in real time or cumulatively, and it can be bundled in new and unexpected ways. It can be used to deliver information such as targeted advertising or to make recommendations to individuals, and it can be used to build profiles that power the algorithms and models that were mentioned earlier. And it can be sold on to third parties, and all of this before Day 1 in the lives of babies born today.

What might we expect through the next five, 10, 15 years and beyond? There will be much more sharing of photos and videos, and real-time calls online to stay in touch. As young children grow and explore the world, touch-screen technologies will become increasingly part of their lives. They'll be playing games and consuming and creating content through smartphones and tablets. As homes become smart homes, households are purchasing voice-activated smart devices that respond to our requests.

A great recent example you might have seen in the media was of a family who discovered that their young child was having conversations with their smart speaker. Add internet-enabled toys: including, say, internet-connected teddy bears that can playback sounds recorded on a smartphone; figurines with motion sensors and face recognition that become part of gameplay; and programmable robots from the simple to the sophisticated.

And now, as we move our attention beyond the home, early childhood education and care services are using technology to connect children, families, carers, educators and other professionals supporting children. These are online platforms allowing communication and sharing of information. Health services and health records are becoming increasingly digitised in a similar way, with clinics and services collecting, storing and sharing information from the earliest stages of a child's life. And as a child grows older, this continued recording, generating, storing and sharing of data will occur across schooling through the official platforms for learning and records of assessment through homework and education apps. It expands through health-monitoring apps like wearable devices that monitor physical activity and nutrition and manage medical conditions.

Over time, with growing independence, babies born today will be making their own choices about engaging with technologies that we know about now. And whatever technologies come next, I'll stop here with the description because I'm sure we'll all have a picture in our minds about the lives ahead for these babies.

I'll leave us to ponder some of the big questions. How can we make the most of the power of technology and the data generated to make our lives better? How can we head off some of the perils that we're already alert to and be ready to face new ones? How can we balance protections, especially for the most vulnerable in our society, with personal choice and freedoms?

Discussion

Prof Oppermann: It's challenging for us to get away from the topic of data because this is how we see the world around us; also something we're increasingly generating as we use online services. Sue, you raised the point about an internet-enabled teddy bear, which seems slightly mind-blowing. A question from our audience is how might we effectively coordinate and scale co-design of our digital technologies? If we are constantly keeping people-centred design, people-focused outcomes at the heart of everything we do, that should be a defence or protection against doing the wrong thing, inadvertently doing things which are not for the net benefit of people in the long term, but scaling that people-centred design is challenging. Sue, how would you effectively go about coordinating and scaling co-design for digital technologies?

SB: Well, we do very little co-design of technology with the people who will end up using it. And if we're talking about children … sometimes people have a deficit view and think children don't really know what's good for them, they don't know what they need or want. A lot of products that are created for children are created by people who have not even worked with children; who often produce something that has not been tested on children and who then don't have an iterative cycle of obtaining and using feedback to improve the relevant technologies. As we domesticate technology, we often change the way that a design is envisaged. So, I think we have to fundamentally change our mindset and build it into the model of interaction. With that will come a cost. It'll cost in time and additional work, but it's overdue. Absolutely.

Prof Oppermann: I love the expression of "domesticating technologies." Frances, there are more questions coming in. I'm going to offer you the choice of either talking about co-design or about the importance of language and how the semantics that we use really matter.

FF-T: Maybe. I'll try to answer both, again talking on behalf of governments. Any data, government data and digital initiatives, need to start with co-designing what it is that we're trying to produce in terms of what we'd call insight tools. For example, we produced the National Disability Data Asset Pilot even though we didn't yet have all of the data to power those tools, and what we heard from the people with disability that we co-designed with (although I hasten to add that were not children and young people) was that they didn't just want a bunch of data insights about them put online for others to read. They wanted themselves to see the insights that can be generated by the kinds of data that government holds.

They wanted to understand the stories that people living with disability hold, to understand: Do those data insights resonate in my life or not? And what's the missing part of our picture? Increasingly, communities are saying it's not just about data, it's about information, knowledge and stories and bridging the quite antiquated idea that you have quantitative insights and you have qualitative insights; knowing that now you can have insights that are developed through a much more sophisticated dialogue between the quantitative and qualitative data. That can be really effectively done through designing the technology whereby we share those insights.

Prof Oppermann: Now I asked our first speakers, Hugh and Cathy, when will we reach the point that we actually know and understand how to appropriately use data, acknowledging biases, acknowledging potential harms, acknowledging all of the challenges. Hugh gave a very strong answer around bias and the skills that statisticians developed around bias. But Sue raised the point about internet-connected teddy bears where there are data sets being generated, interactions being analysed, which we wouldn't otherwise have expected.

Frances, over to you: Where do you think we are in terms of our ability to handle data appropriately? When do we get to the point where we know how to use data safely?

FF-T: Oh, that is a tricky one. I think we use data very safely within governments where we have had a lot of experience with data within a particular service system … where we've been using it for decades. And we're very aware of the limitations of that data. But what I think Hugh and Cathy were talking about is much more sophisticated, linked, and integrated data sets and much more sophisticated methodologies to use that data. We should be open in saying we're relatively early in that process. And that's not just about understanding the data itself, it's also about that process of triangulation I referred to; of saying — even once you understand the data very well — that's only part of the picture. What are your other sources of insights? So I think that governments have heard those expectations from communities, and we are busily figuring out what are the systems that need to be in place to systematically learn the power and the limits of data and the insights that we should be developing.

Prof Oppermann: And now to the parallel question of appropriate use. It may seem innocuous to have a teddy bear connected to the internet, but there are some real challenges there. How do we reframe the conversation around what we should be doing? How do we tell what is appropriate use?

SB: We need to have much greater transparency around the data collected and how it's used. Most of us are aware that data is being collected about us and it's being used in particular ways, but it's actually invisible to us. It's hidden, it's obscured. And that's deliberate because data is a valuable, commercial commodity. But that value is lost to us because we're engaged in relationships where we're required or expected to give it away in exchange for a service or product. When can we have a level of ownership over our data (even as we've given it away), and have some control and some say in how it's used? Inherent in providing people with a data-driven service is actually understanding what they want and allowing them to bring their interpretations to say: "Actually, you've got this bit wrong" or "No, I don't want you to use this and I don't want you to use these things together."

It strikes me that in so much of our lives, we're giving away something that's incredibly valuable to us, and I wonder how we can put some of the control back and build some new social norms about how we expect not just government but also commercial providers to engage with our data in responsible ways. These are moral and ethical questions and dialogues that we really need to have as a society.

Prof Oppermann: Frances, let me ask you about the moral and ethical dialogue. Is there a way that we can fast-forward and build those frameworks so that we can build our touchstone for appropriate use of data?

FF-T: I think we owe a great deal to the thinking that Aboriginal communities have done about institutionalising that dialogue. I'd say we should absolutely have frameworks. But what matters to everyone is that those frameworks are used every day and that the dialogue can evolve theses over time that have community trust and social licence. We are early on in building this understanding, but I think that the Closing the Gap national agreement that was signed by governments in 2021 is very interesting in that it agrees to shared decision-making about using data where there is that institutionalised dialogue between a given community or organisation and government. I should say that it is much more difficult for governments to handle data about individuals because of the way that data systems are configured; it is not a matter of just saying, "I'd like to opt out." That's almost a separate problem set. But that idea of institutionalising dialogue about acceptable uses of data is where participatory data stewardship thinking or data sovereignty thinking being developed by communities and in dialogue with governments.

Prof Oppermann: I just want to ask you one question. I really want you to dig deep and think about 2050. What are the most important topics that we're discussing about the life of today's child in 2050?

SB: Well, I think there's one word: inequalities. Despite all of the advances that we've made, although some inequalities fade away, some persist, new ones arise and continue. So I think we need to look at how those inequalities are arising and we need to have our eyes very much on inclusion.

Prof Oppermann: Frances, 2050 — what's the most important topic we're discussing?

FF-T: How are we empowering a 31-year-old now, a young adult or a mature adult, to not make decisions just about their own lives, but to engage with their communities and as a citizen, about the issues that matter to them?

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 91–97. ISSN 0035-9173/22/01091-07

Avoiding a digital dark age

Shawn Ross[a], **Theresa K. D. Anderson**[b]

[a] Professor of History and Archaeology and Director, Digitally Enabled Research, Office of the Deputy Vice-Chancellor, Macquarie University

[b] Social Informaticist, Connecting Stones Consulting; Research Fellow, School of Information Sciences, University of Illinois at Urbana-Champaign

Moderator and Rapporteur: Ian Oppermann, Chief Data Scientist, NSW Government; Industry Professor, University of Technology, Sydney

Shawn Ross

As a historian and an archaeologist, I have a long-standing interest in both dark ages and the social impacts of technological innovation. I'd like to start by posing the questions: What is a Dark Age and What is the Digital Dark Age?

In my own work, we often discuss dark ages as products of a systems collapse, something that produces a less complex society than the one that preceded it. In a dark age, communities are smaller, poorer, more isolated, with less economic specialisation and social differentiation, and without the resources to support literacy, education or technology. Even technologies that may have been supported by preceding ages. Denizens of a dark age may look back and wonder at the accomplishments of their predecessors. So, the loss of, rejection of, modern information technology would likely cause such a dark age, as it's a pillar of modern complexity. Eliminating today's technology isn't an option. We need to learn how to live with it. And some concepts related to dark ages, or the aspects of societies that allow them to avoid dark ages, may serve as metaphors. At least they might help a child born today understand and thrive in our digital age. Addressing this problem briefly has been a challenge. So I want to present three vignettes, each ending with a provocation.

The first vignette I've called "digital city and pathologies of interaction." Interaction is a defining characteristic of human life and a key component of the complexity that supports dynamic societies. Cities are, in essence, the original technology of interaction. For 6000 years, city were the engines of exchange, wealth generation, cultural experimentation, technological innovation. When we're looking back at the past, the decline of cities is often taken as a hallmark of the dark age.

In better times, people move into cities. In fact, until about 1900, people had to move into cities because, for all of their advantages, plague, famine and other problems meant that deaths in cities always outnumbered births. Indeed, it took millennia to develop the ancillary technologies and practices around hygiene, sanitation, transportation and logistics to make life in cities reasonably secure. Those demographic limits to urbanisation have only recently been lifted, and for the first time now, most people live in cities.

As the abstract that was prepared for this forum observed, digital interaction is a defining characteristic of modern human life. Digital interaction is essentially our new urbanisation, and it incubates its own social contagion. We're all familiar with the discourse around misinformation and fake news, and Jonathan Haidt and others have also explored how the combination of social media and mobile devices has led to increased depression of teenagers.[1] That's just one example that I've come across recently. I'd end with the provocation that we've got to prepare a child born today to navigate this unsanitary online environment and encourage the development of technologies and practices to curb the spread of social contagion. At the same time, though, we should recognise the generative power of the interconnectivity.

I tend to focus on negative things, and I was reminded in the runup to this forum by my co-presenter, Theresa, that we shouldn't overlook the positive parts as well. That interconnectivity, supercharged by the online world, is a place that, like physical cities, fosters exchange and experimentation.

So, my second vignette relates to knowledge, power and magic. A stereotypical view of a dark age involves a monk sitting in a scriptorium copying manuscripts, and there may be a little bit of truth behind this image. With lower literacy rates, access to information can become constrained to a small class of specialists doing things that most don't understand. That monastery image brought to mind Arthur C. Clarke saying that any sufficiently advanced technology is indistinguishable from magic.[2] Now I'm a professor at a university, I and my colleagues think of our students as digital natives. But in my experience, their digital literacy is relatively limited. Russell Kirsch (1929–2020), an early computing pioneer, complained that modern devices are so elegant and easy to use that they subtly direct people towards becoming passive consumers of technologies, instead of using technology to do new things. Douglas Rushkoff, a programmer, goes even further, arguing that programming is the new literacy of the digital age and that the illiterate are being directed by the technology and those who have mastered code.[3] So, my provocation here is that the world is being divided between those who are digitally literate and those who are not, with the former accumulating power. A child born today risks becoming a passive consumer of technology that might as well be magic. To avoid technological serfdom, that child must become digitally literate, and such literacy would also enable that child to become a creator, using digital tools to do things that no one's done before. Considering the accelerating pace of change, learning how to learn technologies is the key to that. Vernor Vinge's *Rainbows End* is an interesting science fiction take on it.[4]

Finally, I'd like to talk about technological transformations and their discontents, something that's close to my heart as an ancient historian. History is marked by transitions in information technology, from orality to literacy, scroll to codex, manuscript to print, print to mass media,

---

[1] Haidt, J. (2018). The Coddling of the American Mind: How Good Intentions and Bad Ideas Are Setting Up a Generation for Failure, co-written with Greg Lukianoff. New York City: Penguin Press.

[2] Clark A.C. (1962) *Profiles of the Future: An Inquiry into the Limits of the Possible*, Harper & Row.

[3] Rushkoff, D. (2010) *Program or be Programmed: Ten Commands for a Digital Age*, Soft Skull Press.

[4] Vinge, V. (2006) *Rainbows End*, Tor Books.

to hypertext to social media; all of which created opportunities, challenges and anxieties. For example, on the cusp of literacy, Plato's Athens used, but distrusted, the written word. Recourse to writing allowed systematic and sustained consideration of the world in ways that had been impossible in an oral society. At the same time, Athenians worried about sophists and rhetoricians separating persuasion from truth. Misinformation has a long history. Athens fostered sophisticated approaches to assessing truth, but also saw a moral panic that included the execution of Socrates. Others have explored similar transitions later: Marshall McLuhan, most famously, in *The Gutenberg Galaxy*.[5] I'd end with the provocation that a child born today will live in a world shaped by new and unforeseen information technologies.

Today's child would benefit from learning what they can from these previous transitions, whether the principles of logic and argumentation brought to bear against sophism in antiquity, or the media literacy training that facilitated understanding and critique of mass media advertisers. Understanding such transitions would foster the sort of awareness that would help that child to realise the latent possibilities of new information technologies, which are usually used for a time in old ways before their potential is fully reached. I look forward to discussing these matters further later in the session.

Prof Oppermann: Shawn, with the aspect of inequality and talking about our future digital world like an ancient city. I think these are unhygienic places, a really powerful concept. And I'm hoping that no one gets put to death associating with our new technological uses. Now I'm going to pass to Teresa Anderson. She is a social informaticist from Connecting Stones Consulting and someone whom I've worked with in a range of different areas, in particular talking about trusted or trustworthy use of data or trusted environments.

Theresa K. D. Anderson

I'm speaking today from Geawegal land. In my work, I draw on socio-technical histories to help me understand our engagements with emerging technologies. Imagining the past as the present and the present as the future, I'd like to share with you some snapshots from that past. To begin, I want you to picture yourself on the streets of Paris in the 1880s. With the industrial revolution well underway, this is an age of public museums and of technological evangelism, world fairs and expositions like the 1889 Universal Exposition, which celebrate the latest advances in science and technology. Every technology in our lives today was at some point new and emerging. Imagine being one of the visitors walking on the exposition's grand terraces, through the gardens and exhibition hall, getting a glimpse into technologically enabled futures. We'll drop in now to the Paris Exposition of Electricity held in 1881. Electricity is so central to our lives today, especially our digital way of working, that it's hard to imagine life without it. But imagine what it was like before electricity was a commonplace household commodity. And now imagine experiencing the rooms that you're seeing here filled with electric lights, bathing the walls with colours that you've never seen before. You've never seen a windowless room illuminated

---

[5] McLuhan, M. (1962) *The Gutenberg Galaxy: The Making of Typographic Man*, Univ. of Toronto Press.

this way, and you can hardly believe that it is even possible. And now imagine sitting at a table with a device called a reading light. One of the other inventions that was presented at this exhibition was the theatre phone. It's a device that would allow you to listen to opera and theatre performances over telephone lines. It was a subscription service, so you paid for the privilege. It was also considered a very fashionable, stylish amusement.

Of course, new ways of communicating bring with them new ways of work and new kinds of workers who are needed to keep the system up and running. An advert from an American telephone company describes female telephone operators working behind the scenes as "weavers of speech." They were referred to as domestic machines. It is an operator who is holding the wires in her hands as if working yarn on a loom. Here is a little extract from this 1915 ad: "Upon the magic looms of the Bell System, tens of millions of telephone messages are daily woven into a marvellous fabric, representing the countless activities of the busy people. Day and night, an invisible hand shift the shuttles to and fro, weaving the thoughts of men and women into a pattern which, if it could be seen as a tapestry, would tell a very dramatic story of our business and our social life. Out of sight of the subscribers, these weavers of speech sit silently at the switchboards, swiftly and skilfully interlacing the cords, which guide the human voice over the country in all directions."

The theme of The Invisible Worker is an important one that I want us to keep in mind when we're thinking about the future of our child. Another thing to keep in mind is wonder. Today, it might be hard to imagine a world without the telephone. But as a young girl, I had a chance to experience just that when I was visiting the mountain villages that my parents grew up in in Central Europe. My uncle was one of the first in the area to have a phone installed in his home. And so I witnessed someone using a phone for the first time, and it really did feel like magic. But I also remember some of the older women in the village warning me to stay away from the humming electricity towers for fear of unknown spirits that were lurking there. This notion of fear and fascination is not new. This is very much a part of our human history. On one hand, it's something that brings great allure — this idea of having power and connection to magical creations. But there's also often a concern about what we might have unleashed, very much like Pandora's Box.[6]

This co-mingled sense of worry and wonder is also very present in the writing of Jules Verne. His visits to the displays of electricity in submarines at the Paris Exposition of 1867 were what inspired him to write *Twenty Thousand Leagues under the Sea* (1870). And yet, while his work inspires us to dream and to imagine, it also reveals his concern about humans becoming disconnected because of the tools that they create. Curiously, Verne's editor refused to publish his dystopian portrait of Paris in the 20th century[7] because he felt that it presented an unrealistic view of the future; yet it was similar to the dark age that was portrayed in

---

[6] "Artificial intelligence began with an ancient wish to forge the Gods," Pamela McCorduck (2004) *Machines who Think*, Routledge.

[7] Verne, J. (1994) *Paris au XX^e siècle*, Hachette.

E.M. Forster's 1909 short story, *The Machine Stops*.[8] The scariest of these tales, the scariest aspect is really just how closely it resembles our present day.

I want to give you a sense of children in this history and some different experiences of childhood from this socio-technical past. Children line up at the 1904 World's Fair in St. Louis, where one of the latest innovations on offer was an ice cream cone. Those who were present, like the Lyons children, could actually savour this delight. We know a lot about this family's history because it is recorded.[9] Their father was a prominent attorney and both his sons went on to become prominent attorneys in Missouri themselves. We don't know much about their sister, Mildred, but we do at least know the name of the little girl who is savouring an ice cream cone. We know very little about another child whose childhood is very different. She's experiencing it as part of the machinery of a textile mill.

These images remind us of both the promises and the challenges of innovation and that they are not inevitable, but neither is the dark age that Foster portrayed as an inevitable, foregone conclusion. So as the child whom we're thinking about today moves into the future, what is it that we can do to prepare them so that they're protected, fortified and inspired to venture ever closer into the dark woods that are up ahead? What is it that we can do to prepare them for digital and enabled futures that are taking shape around them? How can we equip them to embrace these challenges with a sense of adventure and wonder?

It takes a village to raise a child, and I would argue it takes a socio-technical infrastructure to raise this digital and AI world. So, to have a say in shaping this world, our child first and foremost needs opportunity. We must do better to break the cycles of disadvantage that are so profoundly shaping some of our children's futures. They must have the opportunity to learn to develop strong digital adult literacy and be exposed to knowledge. This includes contested knowledge and contested histories. The child is going to need to have the right to step into their future and experience it with all of its promises and perils. Building resilience along the way, through experimentation and learning and with these tools in their backpack, I would then feel pretty confident that this child will have a voice in shaping their socio-technical future.

Discussion

Prof Oppermann: There's a question from the audience that I will pose to both of you: "We need to challenge the assumption that we need to just live with the technologies created by Big Tech and which benefit shareholders. How might we align incentives so that the technologies we create are made with people in mind and so digital literacy and capability building is implicit in the making process?"

SR: Well, that's a hard one. One point bothers me, not only when I see my students but in my role as Director of Digitally Enabled Research at Macquarie. Researchers are using technologies as passive consumers and the commercial products they use con-

---

[8] Forster, E.M. (1909) The Machine Stops, *The Oxford and Cambridge Review*, https://librivox.org/the-machine-stops-by-e-m-forster-2/

[9] Missouri History Museum Lyons Family Collection Archive.

strains their thinking, shape what they do and their approach to work and put them on a certain path. We need enough technical literacy to understand the software we're being fed, what it does, what the alternatives are and recognising that infrastructure matters. There is a cartoon about a professor saying to a student: "Maybe you should use open-source software to write your paper" and then the student has a tirade: "But I just want to do things that are easy and fast." Then 10 years later, the student comes back and says: "Facebook is running our life" and the professor holds his head in his hands.

Does infrastructure matter? I see a real difference between students and colleagues who get on the commercial track versus the open-source track, which requires a lot more co-creation involvement in the product. If you want to do something a little different, you make your own contribution or you contact the maintainers and co-create something together with them.

I guess educating to a level of digital literacy and including in that fostering and encouragement that infrastructure matters, then you can become task-dependent either yourself, or socially on a larger scale.

TA: The infrastructure we're talking about is both human and technical, and often dominates conversations. Now, there is so much focus on the technical and on the technologies and the power that's needed for the structures that we lose sight of all the human aspects, the social structures and the ways that we can encourage learning and experimentation. Co-creation is an opportunity that more and more people are asking for. Moving forward, it's important that we empower all people and particularly the child born today. We need to make sure that they see that they have the opportunities, capacities and literacies to challenge and to question and to participate. And alongside that, we need to be training designers of tools to appreciate that that's a non-negotiable, that this is no longer a "nice to have" item in the process of design. One way that I've thought to approach training data scientists and technologists is to get them appreciating that this is the new way forward for them. Having both those aspects in future will be critical for impacting and making change.

Prof Oppermann: Another audience question: "Is there an assumption that the people really have a say in how technology develops? Doesn't it just happen to us?"

TA: It often appears that technology is allowed to happen to us; it feels like it's beyond our control. But if you look at the way a lot of tools end up being used, very often there is a co-evolving engagement that happens. It is getting increasingly difficult to see that co-evolving opportunity when you look at the power that resides in a few very large global production houses. That is critical to keep in mind.

I don't think we can assume that it's automatic or that it's a given that people have a say. But I think increasingly there is an expectation from people who see so many opportunities around them that they should be able to have a voice in that. How much you can speak to that and in the kind of ways that you can engage really depend on where you're sitting; and the part of the world that you're in, the sort of technology expertise you have. It certainly is not an equally distributed opportunity at the moment. But I've never believed the idea that the computer says no and I have no way of making the computer do anything else.

SR: What you see, over and over again, is that technologies are invented and escape into the wild. Someone used the phrase "domesticated' earlier and suggested that people use technologies in unexpected and unintended ways. So, a technology is not always used according to the intention of the creators. And some infrastructures, technologies of creation, have emerged over the past 10, 20, 30, 40 years — things like social programming platforms like GitHub or GitLab — that can facilitate the creation of software. In the physical realm, things like inexpensive 3D printing and computer-controlled milling machines have emerged and open a lot of doors.

Prof Oppermann: Theresa, I was really struck by your description of the Lyons family contrasting with the girl in the factory. What do those images look like in 2050? What do they tell us?

TA: I'm really scared because I feel like we're going to see new versions of those contrasting images. I don't think enough attention is paid to the inequity that is creating one kind of future for privileged children, as opposed to children who are born into poverty and disadvantage and have to face that. So, that is the terror lurking ahead of us.

SR: To me, as a parallel image, is the Google or Amazon programmers on one side doing that and someone who's an Amazon Mechanical Turk where you get fed little pieces of work online and you sit there, slaving away, getting paid five cents an item. That's how I would update the 19th century divide.

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 98–103. ISSN 0035-9173/22/01098-06

Health of our digital child

Zoran Bolevich[a], **Louisa Jorm**[b]

[a] Chief Executive, eHealth NSW

[b] Professor, Faculty of Medicine and Foundation Director, Centre for Big Data Research in Health, University of New South Wales

Moderator and Rapporteur: Ian Oppermann, Chief Data Scientist, NSW Government; Industry Professor, University of Technology, Sydney

Zoran Bolevich

Our digital child who is born in New South Wales today is likely to be born to a mother aged between 30 and 34. There is a one in three chance that it will be born to a mother who was not born in Australia. We have a very high likelihood, more than 90 per cent, of having a normal birth, so the child will stay in hospital for two to three days prior to going home. Relative to the rest of the world, this child also will be very fortunate to be born in Australia, a developed country with one of the world's best health systems. It will have a life expectancy of 85 years if a girl and 81 years if a boy. However, if Indigenous, this child's life expectancy will be around eight years less than a non-Indigenous child born today.

New South Wales Health, like other organisations in our health system, is committed to ensuring that a child born today is on the way to a long and healthy life. But for our children to thrive, we need to support their development and learning from conception onwards. This will improve their health and wellbeing in the present, but also into the future; giving them longer and healthier lives. This is the essence of NSW Health's First 2000 Days strategy. The first 2000 days of life, from conception to the age of five, are critical to physical, cognitive, social and emotional health over a lifetime.

Early life experiences during this time will determine whether a child is more likely or less likely to experience good health and wellbeing into adulthood. A child's development score at just 22 months can serve as an accurate predictor for educational outcomes at 26 years of age. The first 2000 days are crucial, and our strategy recognises the power of technology as a key enabler. Digital technology already underpins the delivery of services, and it provides us with the data to understand if we are delivering the services to families and children who need to have the best start in life.

In recent times, we have seen an increased uptake of virtual care and telehealth, making a tangible difference. For instance, virtual allied healthcare complements face-to-face care and is vital in exceptional situations where there are access issues for families because of the long distances they have to travel. It's part of a blended model of care, with selected appointments held face-to-face, virtual appointments to reduce stress and various existing regular appointments. In some parts of the state, this is already a reality.

Outcomes that will predict a visiting child's health journey started before they were born; which is why a digital health record commences during pregnancy. New South Wales Health has led the National

Digital Child Health co-laboratory that harmonises all the traditional paper-based pregnancy and child health records into one single digital health record. The ability for health information systems to share data lies in common terminology, clinical information modelling and standardised technical language.

We now can enable real time interactive sharing of data centred on children and their families. The digital health record is a starting point for a lifelong digital health system, which will follow and serve this person throughout their lifetime. By 2030, when our digital child is at school, it will have a comprehensive digital health record that both the child and parents will be able to use on their smart phones. This record is likely to include health and genomic data, combined with relevant information from education and social science, making it possible to predict, diagnose and treat various health or developmental problems more precisely and personally than ever before.

The child and their family will be accessing many routine health care services virtually from the comfort of their home, and their interactions with the health system might be primarily through digital channels. AI-driven care navigators will assist the family with their health care choices and suggest the best place for them to meet their needs. And if this includes, for example, a visit to a hospital emergency department, their arrival will be anticipated, and they will be visited by clinical staff who are already familiar with the child's circumstances and the relevant healthcare issues.

Digital channels of communication between health care providers and the family will continue after their hospital episode, with information following the patient and proactive advice and support services being provided. Wearable devices used by family members will be able to transmit biometric data, which will be analysed in real-time by autonomous, machine-learning algorithms and provide predictive and proactive advice for families and a designated health care team. The child and family will connect with new kinds of healthcare workers and professionals who will be equipped with advanced digital tools such as AI-driven decision support systems and real-time analytics. Training and development of healthcare professionals is likely to evolve towards more cross-disciplinary approaches — with deep inclusion of technology and data analytics subjects in the training curriculum. This might be combined with a stronger infusion of customer service and customer experience skills, as well as human-centred design and agile problem-solving methodology.

As a legacy of the pandemic era of the early 2020s, the health system is likely to be supported by strong population health and disease surveillance capabilities, as well as hopefully the sovereign vaccine manufacturing and distribution capacity. While many of these advances will improve health outcomes and safety and quality of care, there will be an ongoing debate about the ethics of using and sharing personal health information, especially genomic data and device-generated data. The meaning of patients' consent and how to give it effect might be of great interest. Significant concerns are likely to be raised about the "digital divide," which might result in greater health inequalities if left unaddressed. Quite simply, one's health outcomes may depend on a person's ability to access digital services and interact with them.

By 2050, our digital child will be entering the fourth decade of its life. Hopefully, it will live in a world where climate change has been to some extent controlled and where global geopolitical fault lines are not as deep as the ones that we are witnessing now. If this optimistic scenario becomes a reality, it is likely that this person will benefit from the many biomedical and technological advances made in the previous three decades. While this person is likely to be in good health themselves, they are likely to have access, if needed, to gene therapy, robotics technology, nanotechnology, bionics and bio-artificial organs. Digital technology, which was strongly used in previous decades, will be considered mature or even obsolete. Artificial intelligence is likely to be a mainstay of healthcare, with many routine tasks and activities delegated to autonomous agents.

Ethical questions, related to human machine interfaces, end-of-life issues and global health inequalities, might become very significant. Impacts on the health workforce are difficult to predict in this scenario. It is likely that AI and robotics will replace, or reduce demand for, some types of health care workers. But other health care workers might gain prominence, requiring creativity, as well as humanistic, compassionate care and ethical decision-making. And, of course, there will be immediate technical knowledge required to design and oversee the healthcare technology, which might be widely deployed to support the provision of care to an increasingly ageing population. Machines will become an integral part of patient care.

While it is hard to predict what the health of this child might be influenced by in 2050, we do know what will determine many of their health outcomes right now and that is the child's first 2000 days. For our children to thrive, we need to make sure that they have access to all the services they need at the time they need them, which is as early as possible. We need to support them to build not only digital literacy, but digital health literacy so that they can be active participants in their health care.

The NSW Government's Bright Beginnings initiative aims to give them the best start in life. The good news is that we can make a massive difference to the future health of our digital child by taking the right policy decisions and making the right investments, today.

Louisa Jorm

I'm speaking today from the land of the Gadigal people of the Eora nation. The health of the child born today is going to be influenced by some very major environmental forces, and I think it's timely today, with COP26 going on in Glasgow, to remember that the child born today will live in a world with rapid environmental degradation unless some big decisions and changes are made now. The child potentially will be living in an environment of poor air quality, increasing storms, major weather events, disasters and, as we've seen over the last couple of years, pandemic events. Wearing a mask is probably going to be a day-to-day occurrence for our digital child.

There is a growing body of evidence around the degree of digitalisation, in particular the use of personal mobile devices impacting on child health. We've got an increasing prevalence of child obesity and evidence that short-sightedness and other vision problems are more common among children and teenagers. Also, very unfortunately, there is some evidence that overuse or over-reliance on digital devices is changing the social-emotional development of children and their wellbeing; potentially resulting in a whole range of mental health problems.

So there are some real challenges that we know are already facing our children today that they will also experience possibly even more intensely than our current generation of children. However, on the upside, technologies are offering many new ways to improve and maintain health. Some interesting examples include smart nappies that are now going into production, keeping our digital child comfortable and notifying the parent's smartphone when there's something that needs to be changed. We've got smart bottles that can monitor a child's nutritional intake, and we even have smart dummies which can remotely monitor a child's temperature, to be read on a parent's phone. There's a whole plethora of these child-related digital technologies, health-related technologies that are now coming onto the market. Many are from start-ups which are making big claims but have not always got to real-world implementation and production. Yet many of these innovations will arrive over the next 20 to 30 years.

I also think that the digital child born today will over the next 10 to 20 years be accompanied by a personal health avatar. There is a debate among health professionals about whether you call this a personal health avatar or a digital twin. There seems to be consensus that what we can do with existing health data will not really provide the sort of detail that would make this a true digital twin. So, a personal health avatar is what we're calling it, at the moment. It is a something, like an app, that will accompany the child and will bring together its records from a whole range of different sources, and basically simulate it as a human being and simulate its responses to health-related interventions. We already have the beginning of this, in our electronic, closely controlled, My Health Record. But over the next 10 to 20 years, the amount of data that are held in those records from a whole range of sources will become much richer. Zoran mentioned genomics information. We'll also be bringing in sensor information, social media information, information from wearables and sensors, and information from all components of the health system that currently are quite difficult to integrate across sectors.

Having such a digital avatar will enable personalised treatments to be developed for the child, but it will also allow treatments to be tested or simulated on the child before they're actually put into practice, potentially resulting in much better targeting of expensive or potentially harmful therapies and avoidance of safety issues with therapies. I'm quite confident that we will have something resembling that by 2030. However, again, others have talked about some of the potential dark sides. Zoran mentioned the digital divide.

There is a current campaign in New Zealand called "leave no one behind", which is trying to address digital exclusion. I think we in Australia also need to take this sort of initiative. Not only is there a divide in whether or not people can access devices, have the money to pay for devices, or have the knowledge and skills to use them and to interact with them, but there's also the potential for algorithmic bias in AI technologies. A prime example of the potential for this is that around half of the current published scientific papers, which apply machine learning to hospital-based electronic medical records, actually use a single dataset which comes from a hospital in Boston in the USA. A very large number of AI algorithms have been built using this one dataset, and it comes from a population that looks nothing like the Australian population. Obviously it does not include any Indigenous Australians or many other population groups that we have here. There's a really urgent need for our own Australian electronic medical record data to be made much more available for research and development of AI algorithms. If we don't produce our own health database, we potentially will inadvertently implement things that may result in increasing inequalities and potential harm to unsurveyed members of the population. Any algorithm is only as good as the data that it has been fed; it cannot predict the future for, or adapt to, individuals who were not present in the training data.

To end on a slightly positive note, there is potential for digital health technologies to help address inequalities. There is an example, called Kushi Baby, that came originally from Yale, but is now implemented in Rajasthan, India. This is basically a near-field communication tag, which is used to store compressed vaccine records in an offline and battery-free format. Children are taking their health records with them in lovely little tags around their necks.

These sorts of technologies can allow countries that have been slow to adopt digital systems, who are less wealthy, to leapfrog over the paper-based records and complex electronic medical records that we now have in Australia, to far more personally controlled, low-cost and effective devices like these.

To sum up, there will be many challenges for our child, not least those relating to climate change and environmental degradation. There will also be a lot of very exciting developments emerging with technology and in particular, AI. But we need to ask a big question in 2030 and in 2050: Who's been left behind?

Discussion

Prof Oppermann: The first audience question is for Zoran: Is there a risk that the likely use of historical data is going to make the outcome inevitable by shaping decisions which lead to a self-fulfilling prophecy?

ZB: It's really important that the government health agencies have a human judgment and self-disciplined sense of the whole decision-making process, which is why it is important to look for changes in workforce development and health.

Prof Oppermann: If we wind forward and think about the richness of diagnostic, genetic, behavioural and epigenetic data, how do you envisage the interplay between all of these different sorts of data about us and about our health and our environment? And what about self-sovereign identity in e-health?

LJ: It's a real double-edged sword, isn't it? The idea that we should control our data and consent to all the uses of our data. But then, if we don't have representative data about whole populations that include all members of that population, you could end up with biased algorithms. So how does one balance that? As a researcher who works with data, I believe that we don't make enough use of the data we already have. I believe there are increasing ways that we can safeguard and protect privacy. But there's always going to be a risk of a privacy breach if large-scale individual records are being used even in aggregated ways. So I guess it's a risk management framework that we have to work with. I do strongly believe that people should have a degree of control over the use of their data and that we probably need far more dynamic ways of people being able to provide consent for specific uses of their data than they currently have.

I also think we're going to be in trouble unless we make our data more openly accessible for research and development of algorithms. One interesting area is use of public sector data by the private sector and who should share in the benefits from that. An earlier speaker mentioned the potential dollar value of data. I feel very strongly that if data are being generated by services in the public sector, any commercial returns on that data need to fund returns to the public, both in terms of potential health benefits, but also financial benefits.

Prof Oppermann: Zoran, when do you think we will have sufficient control systems, processes in place so that we're not having this conversation about appropriate use of data anymore? When we will be more focused more on the use of the data, rather than concerns around protections of data.

ZB: I see some options and opportunities with blockchain technology that could help with the present digital process. I think some governance processes are maturing and becoming more and more seamless and more permissive, but in a safer way, which would be a positive way forward. What's still missing? In the Senate context, this needs more engagement with the community, and I'm all for an honest conversation about the pros and cons; what's okay and what's not. How can we, together with the community, come up with the right balance and discover the future together? I think that's the testament, and it needs more work.

Prof Oppermann: Louisa, in 2050, what is no longer an issue? What are we spending our time discussing in the world of health?

LJ: I hope that data privacy is no longer an issue and that we will have developed the safeguards needed so that people feel comfortable with how their data are being used. What are we going to be discussing? Probably, what is the role of the health professional going forward? They're going to be working in a very, very different health system, with much more done in automated ways.

Prof Oppermann: Zoran, the same questions to you: What is no longer an issue? What are we really focused on?

ZB: I hope that this whole issue around privacy and confidentiality will be overcome.

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 104–112. ISSN 0035-9173/22/010104-09

Safety and security of our digital child

Dale Lambert[a], **Rory Medcalf**[b]

[a] Chief of Cyber and Electronic Warfare Division, Defence Science and Technology Group, Australian Department of Defence

[b] Professor and Head of the National Security College, Crawford School of Public Policy, Australian National University

Moderator and Rapporteur: Ian Oppermann, Chief Data Scientist, NSW Government; Industry Professor, University of Technology, Sydney

Dale Lambert

I'm speaking to you today from the land of the Kaurna people in Adelaide about security of our digital child. The Industrial Age is a two-tier system comprising a human domain and a physical domain in which people directly control their physical industrial machines. The Information Age expands this to a three-tiered system in which an information domain now separates the human and physical domains. In the Information Age, people issue commands to information environments and expose information environments that now directly control the physical machines.

Our digital child is born as the Information Age eclipses the Industrial Age. Society is now totally reliant on information environments. This creates an unprecedented opportunity because we are constructing a digital representation of our physical information and human domains and making those representations accessible to anyone anywhere. But it also creates a security threat because our society's physical information and human infrastructure are now totally reliant on information environments. A physical domain is now totally reliant on the information environment, and this includes our critical physical infrastructure. So, if someone controls our society's information environments, then they also control our society's physical industry.

Our information systems are also totally reliant on information environments and their algorithms, and these are susceptible to algorithmic warfare or cyber-attacks. So, if someone controls our society's information environments and algorithms, then they control our society's information. And our truth is also totally reliant on information environments. The Correspondence Theory of truth assigns truth based on correspondence with the world, but this now comes through digital images of the world that can be manipulated by image and video editors. The Coherence Theory of truth is the science truth based on coherence of opinion. But this never comes through social media that can be manipulated by fake news. So, if someone controls our society's information environments, then they control our society's truth.

What can we conclude? Well, if someone controls our information environments, they control our physical industry, they control our society's information, and they control our society's truth. In short, if someone controls our society's information environments, then they have total control of our society. Information environments are now contested. They have become the new theatre of warfare in which the participants can be nation-states, crime syndicates, ransomware syndicates, lone wolves or insiders. Human conflict now includes information warfare conducted in and through our information environments by digital ghosts within the machine.

Today's birth of our digital child corresponds with the Information Age of warfare, superseding Industrial Age warfare. By the time our digital child becomes an adult in 2039, our digital adult security will depend on automated contests conducted by digital ghosts within our machines and information environments. The security of our digital adult will critically rely on appropriate ethics and trust being embedded within the digital ghost.

So what happens as a digital adult moves beyond the Information Age? I contend that our digital adult will enter what I'm calling the Virtual Age. The computer games industry and artificial intelligence community will combine to deliver immersive technologies beyond entertainment — to become mainstream in the commerce, health, education, defence and other sectors. Why would this happen? Well, computer science has reduced the communication gulf between machines as electrical and electronic devices on the one hand and the human uses laden with rich conceptualisation, on the other hand by incrementally automating human conceptualisation within machines.

If the Virtual Age continues this application of the automation principle, then the computer gaming and AI communities will deliver us virtual people, virtual societies and virtual environments. For some of you, this might seem like a fanciful suggestion. In the Defence, Science and Technology Organisation [DSTO], we started building prototypes of such things back in 2000. In this project, someone can have a conversation with a virtual person. There is a psychological architecture underpinning the virtual person, and the virtual person can display a range of emotions. Our prototype supported agreement protocols that allowed societies to dynamically form from collections of real and virtual people, and included animated virtual environments that could represent both real and imagined worlds.

By 2063, our digital adult has reached middle age. The Information Age offered the opportunity to see something else by digitally connecting us with the wider world. But the Virtual Age goes further by offering the opportunity to be something else, by digitally experiencing real and imagined worlds. A digital middle-aged person can be their physical self or their virtual self. They can be someone else, perhaps an Indigenous Person. They can understand the mathematical curve by riding it like a rollercoaster. Or they could experience what it is like to be a DNA molecule.

But with opportunity comes threat. The threat of the Virtual Age depends on what we want to protect and secure. How should we balance our digital person's life between the real and virtual worlds, including, for example, the threat of virtual addiction? What virtual truths should we countenance when we can create virtual worlds in which conventional physics and psychological social norms need no longer apply? And what are the rights and status of virtual people? These are just a few of the many issues that will arise.

Rory Medcalf

I'm joining you from Canberra, from the traditional lands of the Ngunnawal people. I want to complement those great remarks by Dale and look at, I guess, a broader picture of what the security environment could look like for the so-called digital child growing up in Australia in the years and the decades ahead.

I'll start briefly with that fundamental question: What is security? It's one of those words that we all think we know what it means — but that we've all got different conceptualisations of. If you go back to its very roots, it's really about a state of mind. Yes, it's about physical protection, but there is no such thing as absolute security. In fact, if you take the Latin origin of security, it literally means without care — no worries. The security of next generations requires achieving a kind of world view or a perspective where people can engage confidently with risk — they cannot achieve absolute security. That will apply both to individuals and our society, indeed Australia as a nation state.

Let's look at the horizon of risk for Australians over the next few decades. It's very easy to be gloomy about this, when you look at the horizon of risk that we see right now — everything from pandemic through great power challenges, China's use of its coercive power, the use and misuse of technology, the continuing risk of terrorism, threats to social cohesion and of course, the overarching threats and risks of the impacts of climate change.

We shouldn't be complacent about any of this, but we should also bear in mind that the last 20 to 30 years have probably been the anomaly. You know, really, the generations growing up in the late 20th century and the beginning of this century had — I hate to say it — almost a long holiday from the historical traumas that most earlier Australians experienced.

Remember the experience of the world wars in the first half of the 20th century? Remember the shadow of the Cold War for much of the second half? So, what is the horizon of risk for our digital child for the next few decades and moving to mid-century? And what are some of the opportunities for society and governments in mitigating that risk? Just a few things to get you thinking about.

First, a lot of the risks we can already see on the horizon of the next ten years are going to be very, very influential in shaping our security environment for much of the rest of the century. There is the question of how states behave in a very competitive international environment. The tensions around, particularly the way that China is using its growing power in our region, the Indo-Pacific, and globally. How is the United States in particular responding? But how will Australia and other countries in our region — India, Japan, Indonesia and others—respond to these tensions? The risk of coercion, military force, even war, but also the more prevalent day-to-day risks in a competitive environment that Dale spoke about — the use of technology by states for strategic advantage, the use of investment and critical infrastructure for strategic advantage.

We're going to see generations growing up with that shadow to come to terms with and much more direct engagement, I think, in the idea of national security than we've seen for many years. There's also the domestic dimension. Even if, internationally, states succeed in managing their differences without confrontation or war, there will be ongoing threats to our sovereignty and our economy. What about the security picture at home?

We've already heard quite a lot about engagement with the information environment, with digital technologies. I think there will be a loss of innocence; that new generations will automatically recognise that their connection to the economy and to the information ecosystem is going to be a source of risk and security anxiety. But hopefully, government and society can engender a new maturity in engaging with that risk so that individuals grow up with a very strong sense of awareness about protecting their privacy, protecting their political freedoms and protecting their engagement with democratic institutions.

There will be risks, I think, to Australia's social cohesion. We've got to remember what a grand experiment a multicultural, federated Australia actually was in the history of our region and the world. The challenge there will be a tension between individuals wanting to simply get on with their lives, as opposed to individuals recognising the need to engage more actively with the political process, to be engaged in society, in politics, to protect those democratic institutions that really have allowed so much individual freedom to flourish in Australia.

And the risks to that social cohesion could come from foreign states seeking to interfere in political processes, particularly the Russian interference in the US elections in recent years. They could come from dissatisfied elements within our own society, the challenge to accept the notions of truth, the rise of coordinated misinformation, disinformation, political violence; terrorism is a fact of life in many countries today and needs to be managed and kept in perspective. We can't let fears of terrorism dominate our daily lives, but we do need an effective national security response.

And there are risks to our social cohesion more broadly. There is a need to protect privacy and political freedoms, but also to inculcate a greater sense of responsibility for our collective future and collective destiny. Those are going to be the kind of challenges that policymakers, but also communities, indeed parents, are going to face in preparing new generations for the challenges ahead. I'm not all gloomy about this, even if it's hard not to sound that way when you're focusing on security.

We've got to think also about the extraordinary capacity of Australia. This is a nation that is often not always mobilised as the sum of its parts. The ultimate challenge for our political class and for politically mobilised communities will be to rebuild a greater sense of common purpose in a democratic Australia. Education will be absolutely key here. As a parent myself, seeing new generations of school age thinking critically about the world, engaging with science and evidence, I think that we still have enormous potential in this country to meet these challenges. We have to go forward with our eyes open, and that will be the challenge for the digital child.

Discussion

Prof Oppermann: Two very interesting presentations and quite different perspectives on issues related to safety and security. Dale, I am going to ask you a question first: What do we mean by trust in a digital environment? Is it, for example, that we believe the system has our best interests at heart? What does trust really mean in a future digital world?

DL: For me, it means the system is respecting our intent, and this goes very close to what Rory was saying about state of mind. You really want implementation in the system to represent the state of mind that we want to have as our values in Australia. I mentioned having to embed trust and ethics within the machine. We need to do that because of the time frames at which things happen. So, for example, I implemented a system for sweeping defence that had to make decisions in two milliseconds. There is no chance that a person is going to be able to be involved in that sort of decision-making. So, it's really important that we take our concepts of trust and ethics and start embedding them inside computerised environments in order to maintain control, if you like, over the information systems that are actually controlling everything now.

Prof Oppermann: Rory, I am going to ask you the more general question of the interplay between encryption and the ability to survey or sense the world around us. What do you think the consequences are of limiting access to encryption and or the interplay between privacy and security?

RM: It's a great question, and one that really frustrates policymakers. There are tensions that we've got to navigate here as a liberal democracy. On one hand, we shouldn't have any illusions that by restricting the ability of our own security establishments to access technologies, to basically co-opt the private sector, for example, in accessing data, that somehow we're going to achieve complete privacy or complete protection of our liberal democratic values. There's a very competitive international environment. Whatever constraints we put on our own security agencies, to sometimes compromise civil liberties or compromise privacy in the interests of national security, there will be authoritarian powers out there who have absolutely no such compunctions. I always find it strange that we'll have people who understandably at one level are really concerned about surveillance by intelligence agencies operating under the rule of law in a democratic system, but at the same time, they're very happy to share pretty much their entire personal data with commercial entities that might have relationships with authoritarian states.

We need to find a new balance here, and I think that balance is going to be struck through constant political scrutiny through engagement with the political process — parliamentary committees and so forth. Being able to really challenge intelligence agencies to justify the powers that they have, but at the same time, to be very open about the trade-offs we make.

I use examples such as the dreadful terrorist attack in Christchurch some years ago by an Australian national, but also terrorist plots that are regularly being frustrated or uncovered. It's going to be very difficult to talk about protection of privacy when being able to access encrypted data would have prevented such attacks. This is going to be a constant challenge, and we're going to have to have very open conversations in the political process about it.

Prof Oppermann: Dale, in the world of security, it's often stated that people are some of the weakest points in the security of systems. So how can we help individuals, old and young, to have more skills in security through the next decades?

DL: The short answer is education, obviously. And we have things like the Australian Signals Directorate's "Essential Eight" that people should practise, but it's more complicated than that. It's about exposing the extent to which people are vulnerable in the things they do. For example, in the previous question about encryption, people may not appreciate the fact that even when you have encrypted data, you can analyse the packet flows of the encrypted data inside the communication networks, and you stand a very good chance of understanding what they are doing.

One example is about Adversarial Machine Learning, which is where you insert data into the data stream very deliberately to cause things to appear and disappear in the outcome from the machine-learning algorithm. You can control the conclusions of the algorithm by injecting data into the system. There's all this stuff that most people probably aren't aware of. Part of the game is making those things more exposed so people understand what the vulnerabilities are. On the human side, we have a team of psychologists in our organisation who can run a bunch of instruments that will assess someone's vulnerability to these things, based on their personality types. This has been used in various organisations. We don't use it to go in and say "sack this person because they're a risk," but it has been used to give an overall profile of what the risk is like within a particular organisation. So, the results of the individuals are masked, but the overall summary of the risk of that agency is revealed.

Prof Oppermann: One of the things about cybersecurity is we don't ever arrive at a cyber-secure environment. It's an ongoing activity because the world is changing around us and it's a very dynamic environment. I want to get back to a question around trust and ethics. Does it require data and computer engineers to have a better understanding of ethics, the rule of law and other foundational values? That's the audience question, but as a follow on … how do we ensure, over the decades of the future, that we are doing appropriate things with technology, specifically thinking about defence and security?

We have heard about the countries taking moratoria on autonomous weapons, for example. These are technologies which we will not use. The ethics is very clear in a situation like that — until you have tensions with another country that won't take that moratorium. But between those black-and-white cases, how do we ensure that we're doing appropriate things with technology over the coming decades? Is it a simple matter of making sure that computer engineers have ethics and rule-of-law training before they're released on bits and algorithms? Or is it something else?

RM: Broadly, it's a self-answering question. It's an important question. I think the short answer is absolutely in technology design; whether that design is occurring in Australia or other democracies or whether it's technologies we're making use of.

Of course, ethical principles need to be raised and addressed in the design phase. But it's not enough to put the onus on design, on engineering. It's the use of technology every day that's going to require ethical decisions and that ethical sensibility needs to be instilled in policy leaders and, frankly, in ordinary citizens as we make our own decisions about using technology.

Ironically, militaries are often ahead of the curve. Ethical training and awareness in militaries is often much greater than in other parts of the policy apparatus or civil society in democracies, or certainly in the private sector, because militaries are making life-and-death decisions every day. And when they get it wrong or do it wrong, as we've seen in current scrutiny of behaviour by a very small number of special forces in Afghanistan, it becomes a major national scandal. So, ethical sensibilities have got to become mainstreamed.

DL: Would I leave it up to the computer engineers? Absolutely not. I'm a huge fan of multidisciplinary approaches to things, and when it comes to something like embedding ethics in machines, you need to understand ethics — and ethics is not like any other discipline. Ethics isn't just one thing that people understand. There are different schools of thought. There are Aristotelian ethics, there are consequentialist ethics, there are Kantian ethics.

So, part of the challenge here is not just how you code this stuff up — it's also what kind of ethics that you want to have in your system. This is quite a serious question, and I think it's a question that should be addressed globally — a bit like we're doing with climate change at the moment, where we're trying to get the world community to share an understanding. As we move forward in time — as our digital child grows up — this is something we also want the world community to embrace so that we have a coherent policy approach across the nations.

Prof Oppermann: Another question has come in, which asks you: "Is there a difference between trust and trustworthiness, when it comes to digital systems?"

DL: Yes, probably. Trustworthiness is a term that's increasingly used in the military context. It's really asking: Is this equipment going to do what I expect it to do? Part of that is about trying to protect your equipment from cyber-attacks and things of that ilk. Trust, for me, is a broader issue. As I talked about before, it goes very much to Rory's concept of security as a state of mind. I see that as a much more human and intent-driven activity.

Prof Oppermann: The last audience question is a statement that there is a fundamental risk that encrypted data which is stolen today will eventually be decrypted with tomorrow's quantum computers and the harms will come tomorrow. Do you see this as an issue? And if so, what should we do about this today?

RM: I think that that's a reasonable question. What, how and when will we actually see quantum encryption and quantum decryption realised? It's been much promised, but it's still some way off. But just because data is encrypted doesn't necessarily make it more important. There's an enormous amount of unencrypted material out there, or open-source information out there, that is potentially incredibly useful to a future adversary or malign actor if they can aggregate it and make sense of it, particularly with AI and machine learning. So, for example, my university some years ago had a major cyber breach that was publicly reported. Other universities, corporations, government agencies, individuals, have their data taken all the time. And many of us will say, Well, so what? It doesn't matter. Who cares if you know whether it's a cybercriminal or whether it's someone sitting in an office in a government building in Shanghai reading everything that's on my screen. Well, it's going to be useful somehow, some way. So we must get much better at that; not only because of general digital hygiene, but understanding that other governments or organisations building a complete life picture of you as an individual can be particularly dangerous in the long term. We should all remember that our own data is incredibly valuable. It's who we are. It's not only our privacy, but it's our future careers, for young people who may wish to work in government, for example, or have careers in security agencies or politics. Every piece of your digital footprint now could be used against you in future.

DL: The idea that tomorrow's quantum computers are going to completely overcome the encryption situation isn't quite right. There's a thing called Shor's algorithm, which sits in a class which is a bounded quantum polynomial — that sits higher than polynomial, but it doesn't get you to the full exponential non-polynomial (NP) complexity. I mention this is for the computer scientists who are listening.

I remember a really interesting conversation with someone from a foreign intelligence agency about their fingerprints remaining in the digital world so that they can at some point be uncovered for something they've been doing, maybe in 10, 20, or 30 years' time. It's quite a worry for them because they might do something at a time when they can't be detected. But if it's being recorded through some sort of digital mechanism, they might well be uncovered at some point in the future. It's a really interesting consideration and hopefully, it might stop some people doing some nefarious things because they'll have second thoughts about it.

Prof Oppermann: One of the points you touched on was with the use of fingerprints, and historically we've used biometrics for a whole lot of different sorts of authentication purposes. If they lose their significance as individual identifiers, then we have to rethink a lot of traditional ways of identifying and securing. Let me ask you the unfair questions: In 2050, what's no longer an issue, and what are we really focused on?

RM: The second part is much easier to answer than the first. I would never go so far as to say that the kind of terrorism that we've been dealing with — quite obsessed with over the past 20 years — will no longer be an issue, but I do think that we're on a pathway to putting it in context; to recognising that terrorism is essentially a criminal activity, politically motivated violence. I think though that the risk of terrorism and violent extremism is going to be background noise in our national security debate permanently. We need to deal with bigger issues, such as Australia's resilience and sovereignty in a pretty contested region.

Relatively speaking, Australia will probably be a less powerful country than we are now. That's a very sobering thought. There is a need to draw up the connections between national resilience and the security of our energy supplies, the sustainability of our society, the environmental sustainability to connect all of that with the idea of national defence and our security in a competitive world. That's going to be the big question.

I guess the good news is that the barriers between security and economics will break down. The bad news is it's going to be very hard to turn that into a practical policy agenda that the government can operationalise. So it's going to be tough, but I think citizen engagement in national security will be a big part of the solution.

DL: I'm hoping that we won't have to worry about information warfare by that time because we will have had some sort of stabilisation and an international agreement around what's reasonable and what's not reasonable in terms of opportunity in that time frame. I talked about three different ages: The industrial age, information age, and virtual age. It's easy to underestimate how important the virtual age will be because it basically lets human ingenuity off the leash. You're no longer restricted by a lot of the physical constraints and things that we've all grown up with. This is an opportunity space and it's really up to us how imaginative we can be in exploiting that opportunity.

* * *

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, p. 113. ISSN 0035-9173/22/010113-01

Thesis abstract

Sex and stress: Is stress both a mediator and a consequence of sex reversal in the central bearded dragon (*Pogona vitticeps*)?

Meghan Castelli

Abstract of a thesis for a Doctorate of Philosophy submitted to the University of Canberra

Among vertebrates, sex-determining systems are diverse and range on a continuum from entirely genetic sex determination (GSD) to purely environmental sex determination (ESD). The central bearded dragon (*Pogona vitticeps*) possesses a heterogametic (ZZ male/ZW female) system of genetic sex determination, but high egg incubation temperatures induce sex reversal, in which ZZ genotypic males develop as female. The biological mechanism by which temperature is translated into a sexual outcome is not fully understood in reptiles but is proposed to involve the vertebrate stress axis, a highly conserved environmental sensory mechanism which generates physiological responses to stress through glucocorticoid hormone production. Here I demonstrate using developmental transcriptomes and chemical manipulation experiments that the stress axis is unlikely to mediate sex reversal, and instead find evidence for the involvement of oxidative stress responses or circadian rhythm regulation in high-temperature sex reversal. The relevance of these molecular studies is contextualised by a range-wide study of sex reversal and population genetic structure which demonstrates the lack of a clear relationship between climate and sex reversal in the wild. I have proposed that the threshold temperature for sex reversal (and thus the underlying genetic network which determines the threshold) varies across the landscape, having evolved in response to higher average incubation temperatures in warmer regions of the species' range. Together, these studies demonstrate that not only is temperature sex reversal in reptiles a molecular process likely driven by sensory mechanisms other than the stress axis, but it also has complex evolutionary dynamics in the wild.

Dr Meghan Castelli
Postdoctoral research fellow
School of Science, Technology & Engrg
University of the Sunshine Coast

E-mail: meghan.castelli@gmail.com

URL: https://researchprofiles.canberra.edu.au/en/studentTheses/sex-and-stress-is-stress-both-a-mediator-and-a-consequence-of-sex

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 114–115. ISSN 0035-9173/22/010114-02

Thesis abstract

Robust statistical inference for one-shot devices based on divergences

Elena Castilla

Abstract of a thesis for a Doctorate of Philosophy submitted to Complutense University of Madrid, Madrid, Spain

The reliability of a product, system, weapon, or piece of equipment can be defined as the ability of the device to perform as designed for, or, more simply, as the probability that the device does not fail when used. Engineers assess reliability by repeatedly testing the device and observing its failure rate. Certain products, called "one-shot" devices, make this approach challenging. For this kind of device, one can only know whether the failure time is either before or after a specific inspection time, and consequently the lifetimes are either left- or right-censored, with the lifetime being less than the inspection time if the test outcome is a failure (resulting in left-censoring) and the lifetime being more than the inspection time if the test outcome is a success (resulting in right-censoring). An accelerated life test (ALT) plan is usually employed to evaluate the reliability of such products by increasing the levels of stress factors and then extrapolating the life characteristics from high stress conditions to normal operating conditions. This acceleration process will shorten the life span of devices and reduce the costs associated with the experiment. The study of one-shot devices from ALT data has been developed considerably recently.

On the other hand, in the last decades the use of divergence measures in the resolution of statistical problems has reached a remarkable relevance in the areas of parametric estimation and parametric tests of hypotheses, together with many nonparametric uses. Particularly, the family of density power divergences is known for its robustness.

Most of the results concerning one-shot devices are based on maximum likelihood estimation (MLE), which is well-known to be efficient, but also non-robust. Therefore, testing procedures based on MLE face serious robustness problems. Along this thesis, robust estimators and tests are developed based on density power divergences for one-shot device testing.

The thesis proceeds as follows. In Chapter 2, we assume the one-shot devices with lifetimes having exponential distribution with a single-stress relationship. In Chapter 3, the exponential distribution with multiple-stress relationship is considered, generalizing the results in Chapter 2. Next, in Chapter 4 and Chapter 5, we consider the situation when lifetimes follow, respectively, a Gamma and a Weibull distribution with non-constant shape parameters. In Chapter 6, similar procedures are applied to other distribution functions, such as Lindley and lognormal. Chapter 7 develop robust inference for one-shot device testing under the proportional hazards assumption and, in

Chapter 8, we consider the competing risk model (assuming different possible causes of failure) with exponential lifetimes. Chapter 9 summarizes the main results of previous chapters and gives some ideas about future work. Appendices A and B briefly present some other results, which have also been obtained by the candidate during her Ph.D. studies.

Dr Elena Castilla
Department of Applied Mathematics
Universidad Rey Juan Carlos
Móstoles (Madrid), 28933
Spain

E-mail: elena.castilla@urjc.es

URL: https://eprints.ucm.es/id/eprint/67473

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, p. 116. ISSN 0035-9173/22/010116-01

Thesis abstract

Parturescence: a post qualitative inquiry into women's opportunities for transcendence and transformation through births

Ella Kurz

Abstract of a thesis for a Doctorate of Philosophy submitted to the University of Canberra

Childbirth, including the challenging and destabilising parts, is not just for the production of a baby, but is also a site for women's transcendence and transformation. Where this goes unacknowledged, women's possibilities are reduced. In order to open up women's possibilities through birth, in this midwifery thesis, I problematise dominant constructions of childbirth with the aim of theorising birth differently. By applying processes of post qualitative inquiry to collected data, as well as maternity and feminist literature, my practice of midwifery, and my own childbearing I move beyond the dominant constructions of birth, and think differently about the constitution and subjectivity of birth and maternity care subjects. I theorise birth as an opportunity of 'becoming other'; the rematerialisation of subjectivity, which I term 'parturescence.' I situate women's transcendence and transformation at one end of the spectrum of possible experiences of parturescence and explore how they are constituted by the intra-action between the materiality of birth and midwifery 'with woman' care. Finally, I discuss how the development of the theory of parturescence can enable all of us, childbearing women and their support people, midwives and everyone involved in the provision of maternity care, to take actions which increase women's opportunities for transcendence and transformation through birth.

Dr Ella Kurz
Research Assistant
University of Canberra
Bruce, ACT 2617

E-mail: ella.kurz@canberra.edu.au

URL: https://researchprofiles.canberra.edu.au/en/studentTheses/parturescence-a-post-qualitative-inquiry-into-womens-opportunitie

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 117–118. ISSN 0035-9173/22/010117-02

Thesis abstract

The development and implementation of youth mental health outcome measures within measurement feedback systems

Benjamin Kwan

Abstract of a thesis for a Doctorate of Philosophy submitted to the University of Canberra

This thesis by published works provides an original contribution to the knowledge about mental health outcome measures used with young people aged 12 to 25 years. It discusses how such measures can be developed and implemented as a clinical tool in day-to-day practice. The main goal was to develop a routine outcome measure that was suitable for youth mental health settings that could be implemented into measurement feedback systems. The research aimed to achieve this goal by firstly identifying the gap in existing mental health outcome measures used with young people and, secondly, by examining how these measures were being used clinically. A further aim of this research was to examine the psychometric properties of a new routine outcome measure, MyLifeTracker, in youth mental health settings. It aimed to determine developmentally appropriate clinically significant change indexes, expected change trajectories, and early warning signals for MyLifeTracker, to provide clinicians with clinically useful and evidence-based benchmarks. Lastly, the research aimed to explore factors affecting the use of MyLifeTracker in measurement feedback systems across youth mental health settings.

The research used a quantitative methodology comprising four research papers, each addressing a specific research aim. The first research paper was a systematic review of general mental health outcome measures for young people aged 12 to 25 years. It identified how these measures track change and if they had been used in feedback monitoring. The second and third papers examined the use of MyLifeTracker by clients receiving support from headspace youth mental health services across Australia. The second paper explored the reliability, validity, and sensitivity to change of MyLifeTracker for young people, across gender and age groups. The third paper determined clinically significant change indexes by gender and age groups by comparing participants from headspace services to an Australian representative community sample of young people. Expected change trajectories were also determined for the clinical group using growth curve modelling. The fourth paper reports findings from a survey of 210 clinicians from headspace centres about their use of MyLifeTracker, specifically exploring three processes of measurement feedback systems: looking at MyLifeTracker before the session, using MyLifeTracker in treatment planning, and providing feedback of MyLifeTracker scores to clients.

The systematic review identified 29 different outcome measures used with young people, however, no measures were explicitly designed for this age group. Only two measures were found to be used by clinicians in measurement feedback systems in

this age range. Findings from the review led to the recommendation that measures be explicitly designed for this age group that are suitable for routine outcome monitoring. The second paper demonstrates that MyLifeTracker provides a psychometrically sound mental health outcome measure for routine use with young people. The measure has been incorporated into an electronic system for headspace services that routinely tracks session-by-session change and produces time-series charts for ease of use and interpretation. The third paper provides clinical benchmarks for MyLifeTracker, further supporting the use of the measure in measurement feedback systems. Lastly, the fourth paper reports the different levels of use of MyLifeTracker in a measurement feedback system and highlights the factors that increase clinicians' use for each process.

The thesis supports the use of mental health outcome measures to be used not only for assessing service effectiveness and quality assurance, but also as a clinical tool to support decision making and treatment planning by clinicians and clients. It provides support for brief and easy to use measures that are meaningful for clinicians and clients. The clinical benefits of measurement feedback systems are becoming more widely known and have become part of the agenda for the future progression of psychotherapy. The thesis targets a developmental age group that has high rates of clinical deterioration, treatment drop-out and missed appointments, and who may be quite responsive to feedback monitoring. The thesis concludes by offering a range of targeted strategies that can support the future implementation of outcome measurement feedback systems into practice.

Dr Benjamin Kwan
Clinical Psychologist

E-mail: ben.kwan1@outlook.com

URL: https://researchprofiles.canberra.edu.au/en/studentTheses/the-development-and-implementation-of-youth-mental-health-outcome

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022, pp. 119–121. ISSN 0035-9173/22/010119-03

Obituary: Geoffrey Harcourt AC FRSN (1931–2021)

Peter Kriesler[a], Tim Harcourt[b], Robert Marks[a*]

[a] Economics, UNSW Sydney, Australia

[b] Institute for Public Policy and Governance, University of Technology, Sydney

*Corresponding author. Email: robert.marks@gmail.com

Geoff and his twin brother John were born in Melbourne on the 27th of June 1931. Geoff died on 7th December 2021, aged 90. Geoff attended the University of Melbourne where he flourished while studying accounting and economics, achieving first class honours, and subsequently a M. Com. While at the University of Melbourne, he was exposed to the economics of Keynes and the Cambridge School, which exerted a profound influence on all his subsequent work.[1]

Geoff was awarded a PhD scholarship to study at Cambridge, where he was supervised by Nicholas Kaldor and Ronald Henderson. He absorbed the atmosphere and the intellectual stimulation of being among the great Cambridge economists. Joan Robinson, in particular, was an important influence. Geoff attended her lectures, and closely studied her magnum opus, *The Accumulation of Capital* (1956), which had a deep effect on his subsequent development; Geoff, with Prue Kerr, would edit and write the introduction for the third edition of her book. He was awarded a PhD in 1957 for his thesis, "A study of the implications of the use of historical-cost accounting procedures to set prices and dividends, and levy taxes in a period of inflation."

Cambridge became Geoff's centre of gravity. He returned often, before moving there permanently in 1982 to take up a University Lectureship. In 1990 he become a Reader in the History of Economic Theory, until retiring in 1998, when he was made an Emeritus Reader. He was a Fellow at Jesus College during that time, and its President from 1988–89 and 1990–92. Meanwhile, he was appointed to a lectureship at the University of Adelaide in 1958, and then to a personal Chair in 1967, and Professor Emeritus in 1988. After his retirement from Cambridge, he became a visiting Professorial Fellow and then an Honorary Professor at the School of Economics at the University of New South Wales, where he spent his last decades. In the light of this, it is not surprising that Geoff regarded himself as "an Australian patriot and a Cambridge economist" (Harcourt 1995).

Geoff received many significant awards during his life. In 2016 he became a Fellow of the Royal Society of NSW and was active in our Forums. In 2018 he was made a Companion of the Order of Australia (AC) for his "eminent service to higher education as an academic economist and author, particularly in the fields of Post-Keynesian economics, capital theory and economic thought." He received honorary doctorates from universities in the UK, Europe, and Australia. He was a founding Fellow of the Academy of the Social Sciences in Australia. He was sometime President of the Economic Soci-

---

[1] This obituary draws on Kriesler (2022) and Harcourt (2022); Tim Harcourt is Geoff's son.

ety of Australia and New Zealand. In 1962 he had helped found *Australian Economic Papers*, and was its editor for two decades. He had numerous honours from learned societies in the US, UK, and Europe. He was without doubt the most eminent economist in the Royal Society of NSW since William Stanley Jevons, co-founder of microeconomics in the nineteenth century (Castles 2016, Marks 2016).

Geoff was driven by a strong commitment to social justice, which also informed his academic and policy work. He had a lifelong commitment to equity and equality, working towards alleviating poverty and against social and racial discrimination. He also had a great love of sports, both as player and spectator, particularly Australian football and cricket. In honour of his passion for Australian football, many of his papers are written in four quarters.

Geoff had a gift of putting people at ease, talking to anyone, from the Crown Prince to awestruck students, while displaying his mischievous sense of humour. He was genuinely interested in everyone he met. Geoff loved jokes.

Geoff's contribution to economics, both theory and policy, was outstanding. In over 30 books and 400 articles, numerous lectures, seminars, and interviews he had a significant impact on the discipline. Geoff made economics more humane, and humanised the "dismal" science. His contributions to economics covered a broad range of areas from esoteric pure theory to applied policy, always with the aim of trying to make the world a better place. He provided original insights, and was able to explain difficult and complex ideas in an accessible form, while often showing his mischievous sense of humour.[2]

In his important article and subsequent book on the Cambridge capital controversies (1972), Geoff provided a masterful guide to one of the most technical debates in economics; Cambridge University Press is publishing a 50th anniversary edition of the book later this year, with a new preface by Geoff, and afterwords by Avi Cohen and Tiago Mata.

The capital theory controversies were a series of debates in the 1950s–1970s on high theory, between economists mainly based in Cambridge, England (Joan Robinson, Piero Sraffa, Luigi Pasinetti) and at MIT in Cambridge, Massachusetts (Paul Samuelson, Robert Solow[3]). Although the debates were ostensibly about the problem of measuring capital, they were ultimately about the nature and meaning of capital, and the question of the most appropriate way to analyse a contemporary capitalist economy. Geoff unravelled the debates so that they became intelligible, with both clear style and humour, evident particularly in the chapter and section titles, such as the section pointing to an error by Kaldor titled, "Excuse me, Professor Kaldor, but your slip is showing."

Geoff was a founder and major contributor to post-Keynesian economics throughout his intellectual life, culminating in his books, *The Structure of Post-Keynesian Economics* and the two-volume *The Oxford Handbook of Post-Keynesian Economics*. He humanised economics by making it more accessible through his clear, and often humorous, writing style and

---

[2] Groenewegen & McFarlane (1990, pp. 196–199) give a good summary of Geoff's contributions until then.

[3] Both of whom became Nobel Laureates.

in his many biographical essays. These went behind the mask of economists to reveal not only their economic insights but also the person who developed those insights.

Bolstering his theoretical contributions lies the importance of policy, on which Geoff not only wrote copiously, but also acted by advising governments and commenting on contemporary issues. He believed that academics in general, and economists in particular, have a duty to advocate policies which would lead to a better world. During the Whitlam government, Geoff and other Adelaide economists devised The Adelaide Plan, an incomes policy that laid the foundations for the ACTU-ALP Prices and Incomes Accord of the Hawke-Keating Labor Government.

Associated with writing and advocacy was his belief in a "need for direct action if other more orthodox means proved ineffective" (Harcourt 2011, p. 124), which Geoff displayed with his important role in Australia's anti-Vietnam War movement. Despite being an unofficial advisor to the ALP, the only time he got close to an official government position was during the Whitlam government, when the Treasurer, Jim Cairns, offered him the position of either Governor of the Reserve Bank or Secretary of the Treasury. He declined both.

Everyone who knew Geoff values the wonderful human being he was, as well as being a world-class economist. He enriched the lives of everyone around him. He was a true scholar and gentleman, and the world is so much a better place because of him.

References

Castles, I. (2016) The curious economist: William Stanley Jevons in Sydney, *Journal & Proceedings of the Royal Society of New South Wales* 149: 51–58.

Groenewegen, P. and McFarlane, B. (1990) *A History of Australian Economic Thought*. Routledge, London.

Harcourt, G.C. (1969) Some Cambridge controversies in the theory of capital. *Journal of Economic Literature* 7: 369–405.

Harcourt, G.C. (1972) *Some Cambridge Controversies in the Theory of Capital*. Cambridge U.P.

Harcourt, G.C. (1995) Recollections and reflections of an Australian patriot and a Cambridge economist. *Banca Nazionale del Lavoro Quarterly Review* 48: 225–54.

Harcourt, G.C. (2006) *The Structure of Post-Keynesian Economics: the Core Contributions of the Pioneers*. Cambridge U.P.

Harcourt, G.C. (2011) Post-Keynesian theory, direct action and political involvement. *Intervention* 8(1): 117–128.

Harcourt, G.C. and Kerr, P. (2009) *Joan Robinson*. Palgrave Macmillan, Basingstoke.

Harcourt, G.C. and Kriesler, P. (2013) (eds.) *The Oxford Handbook of Post-Keynesian economics. Volume 1: Theory and Origins*, and *Volume 2: Critiques and Methodology*. Oxford U.P., New York.

Harcourt, T. (2022) Geoff Harcourt (27.6.1931 to 7.12.2021) *Australian Economic Papers*, May, https://doi.org/10.1111/1467-8454.12261.

Kriesler, P. (2022) Geoffrey Harcourt, *Royal Economic Society Newsletter*, April, pp. 22–24.

Marks, R.E. (2016) William Stanley Jevons, Fellow of the Philosophical Society of N.S.W., 1856–1859, *Journal & Proceedings of the Royal Society of New South Wales* 149: 59–64.

Robinson, Joan (2013) *The Accumulation of Capital* (3rd ed.) in the Palgrave Classics in Economics Series. Palgrave Macmillan, Basingstoke.

◆

Journal & Proceedings of the Royal Society of New South Wales, vol. 155, part 1, 2022,
pp. 122–123. ISSN 0035-9173/22/010122-02

Royal Society of New South Wales

Awards 2022

Each year, the Society makes a number of awards, mainly in the field of science. They are among the oldest and most prestigious awards in Australia and are summarised below. Full details for each award, including the procedure for nomination, may be found by clicking on the webpage https://royalsoc.org.au/awards. A nominator does not need to be a member or fellow of the Society.

Nominations for all awards close on 30 September 2022.

Archibald Ollé Prize

The Archibald Ollé Prize of $500 is given from time to time, at the discretion of the Council, to the to the author who has submitted and had accepted the best single-author paper to the Society's *Journal*.

The Clarke Medal

The Clarke Medal is awarded each year for distinguished research in the natural sciences conducted in the Australian Commonwealth and its territories. The recipient may be resident in Australia or elsewhere. The fields of geology, botany and zoology are considered in rotation.

For 2022, the medal will be awarded in Geology.

Clarke Memorial Lecture

The Clarke Memorial Lecture is delivered each year by the most recent winner of the Clarke Medal.

Edgeworth David Medal

The Edgeworth David Medal is awarded each year for distinguished contributions by a young scientist under the age of thirty-five (35) years on 1 January in the year in which the medal is awarded, for work done mainly in Australia or its territories, or contributing to the advancement of Australian science.

History and Philosophy of Science Medal

The Society's History and Philosophy of Science Medal is awarded each year to recognise outstanding achievement in the History and Philosophy of Science with preference being given to the study of ideas, institutions, and individuals of significance to the practice of the natural sciences in Australia.

James Cook Medal

The James Cook Medal is awarded periodically for outstanding contributions to both science and human welfare in and for the Southern Hemisphere.

Liversidge Lecture

The Liversidge lectureship is awarded biennially for research in chemistry. The lecture is presented in conjunction with the Royal Australian Chemical Institute (RACI). The lecture will be published in the *Journal & Proceedings of the Royal Society of New South Wales*.

Poggendorff Lecture

The Poggendorff Lecture is awarded every two to three years for research in plant biology and, more broadly, agriculture.

Pollock Memorial Lecture

The Pollock Memorial Lectureship has been awarded approximately every four years since 1949 and is sponsored by the University of Sydney and the Society in memory of Professor J.A. Pollock, Professor of Physics at the University of Sydney (1899–1922) and a member of the Society for 35 years.

The Royal Society of New South Wales Citation

The Royal Society of New South Wales Citation is awarded to a Member or Fellow of the Society who has made significant contributions to the Society, but who has not been recognised in any other way. The Awards Committee considers nominations made by a Member or Fellow. A maximum of three Citations may be awarded in any one year.

The Royal Society of New South Wales Medal

The Society's Medal is awarded from time to time to a member of the Society who has made meritorious contributions to the advancement of science, including administration and organisation of scientific endeavour and for services to the Society.

The Royal Society of New South Wales Scholarships

Three scholarships of $500 plus a complimentary year of Associate Membership of the Society are awarded each year in order to acknowledge outstanding achievements by young researchers in any field of science in New South Wales. Applicants must be enrolled in their first higher degree as research students in their first or second year, in a university or at CSIRO in either NSW or the ACT (on 1 January of the year of nomination) and have completed an undergraduate degree within NSW or the ACT.

Walter Burfitt Prize

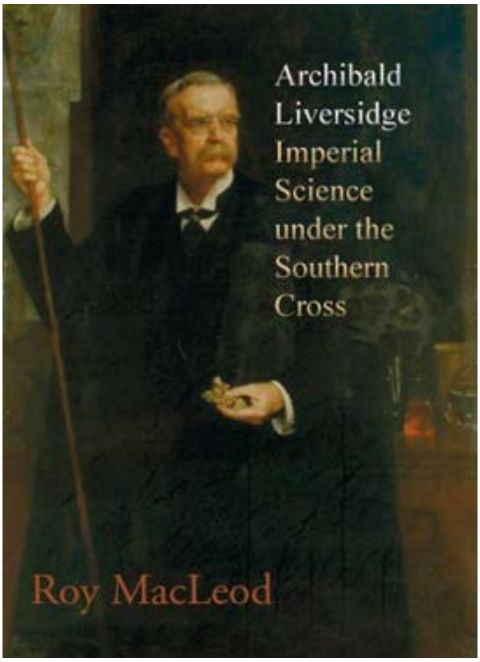
The Walter Burfitt Prize consists of a bronze medal, awarded every three years for research in pure or applied science, deemed to be of the highest scientific merit. The winner must be a resident in Australia or New Zealand. The papers and other contributions must have been published during the previous six years for research conducted mainly in these countries. Nominations for the Walter Burfitt prize will next be sought in 2022.

Warren Prize (Medal & Lecture)

The Warren Prize is awarded to recognise research of national or international significance by early- or mid-career engineers and technologists. The research must have originated or have been carried out principally in New South Wales.

Archibald Liversidge:
Imperial Science under the Southern Cross

Roy MacLeod
Royal Society of New South Wales, in association with Sydney University Press
ISBN 9781-9208-9880-9



When Archibald Liversidge first arrived at the University of Sydney in 1872 as Reader in Geology and Assistant in the Laboratory, he had about ten students and two rooms in the main building. In 1874, he became Professor of Geology and Mineralogy and by 1879 he had persuaded the University Senate to open a Faculty of Science. He became its first Dean in 1882.

In 1880, he visited Europe as a trustee of the Australian Museum and his report helped to establish the Industrial, Technological and Sanitary Museum which formed the basis of the present Powerhouse Museum's collection. Liversidge also played a major role in establishing the *Australasian Association for the Advancement of Science* which held its first congress in 1888.

This book is essential reading for those interested in the development of science in colonial Australia, particularly the fields of crystallography, mineral chemistry, chemical geology and strategic minerals policy.

To order your copy, please complete the Liversidge Book Order Form available at:

https://royalsoc.org.au/images/pdf/McLeod_Liversidge_Order_Form_Rev.pdf

and return it together with your payment to:

The Royal Society of NSW,
(Liversidge Book),
PO Box 576,
Crows Nest NSW 1585,
AUSTRALIA

or contact the Society:

Phone: +61 2 9431 8691
Fax: +61 2 9431 8677
Email: info@royalsoc.org.au

The Royal Society of New South Wales

Information for authors

Details of submission guidelines can be found in the on-line Style Guide for Authors at: https://royalsoc.org.au/society-publications/information-for-authors

Manuscripts are only accepted in digital format and should be e-mailed to:
journal-ed@royalsoc.org.au

The templates available on the Journal website should be used for preparing manuscripts. Full instructions for preparing submissions are also given on the website.

If the file-size is too large to email it should be placed on a USB drive (or other digital media) and posted to:

The Honorary Secretary (Editorial),
The Royal Society of New South Wales,
PO Box 576,
Crows Nest, NSW 1585
Australia

Manuscripts will be reviewed by the Editor, in consultation with the Editorial Board, to decide whether the paper will be considered for publication in the Journal. Manuscripts are subjected to peer review by at least one independent reviewer. In the event of initial rejection, manuscripts may be sent to other reviewers.

Papers (other than those specially invited by the Editorial Board) will only be considered if the content is either substantially new material that has not been published previously, or is a review of a major research programme. Papers presenting aspects of the historical record of research carried out within Australia are particularly encouraged. In the case of papers presenting new research, the author must certify that the material has not been submitted concurrently elsewhere nor is likely to be published elsewhere in substantially the same form. In the case of papers reviewing a major research programme, the author must certify that the material has not been published substantially in the same form elsewhere and that permission for the Society to publish has been granted by all copyright holders. Letters to the Editor, Discourses, Short Notes and Abstracts of Australian PhD theses may also be submitted for publication. Please contact the Editor if you would like to discuss a possible article for inclusion in the Journal.

The Society does not require authors to transfer the copyright of their manuscript to the Society but authors are required to grant the Society an unrestricted licence to reproduce in any form manuscripts accepted for publication in the Journal and Proceedings. Enquiries relating to copyright or reproduction of an article should be directed to the Editor.

Volume 155 Part 1 2022 Numbers 485 & 486

CONTENTS

Refereed Papers

Reprinted Paper

The Royal Society of N.S.W. and Five Academies Forum:
Power and Peril of the Digital Age

Obituary

The Royal Society of New South Wales
P.O. Box 576
Crows Nest, NSW 1585, Australia

ISSN 0035-9173

info@royalsoc.org.au (general)
editor@royalsoc.org.au (editorial)
www.royalsoc.org.au
www.facebook.com/royalsoc

Published June 2022