Strategic Foresight: Knowledge, Tools, and Methods for the Future

Report

Author(s):
Kohler, Kevin

Publication date:
2021-09

Permanent link:
https://doi.org/10.3929/ethz-b-000505468

Originally published in:
CSS Risk and Resilience Reports

RISK AND RESILIENCE REPORT

Strategic Foresight
Knowledge, Tools, and Methods for the Future

Kevin Kohler

Commissioned by the Federal Commission for Nuclear, Biological and Chemical Protection and the Federal Office for Civil Protection

Zürich, September 2021
Center for Security Studies (CSS), ETH Zürich

CSS
ETH Zürich

ETH zürich

Available online at: css.ethz.ch/en/publications/risk-and-resilience-reports.html

Author: Kevin Kohler
Clients: Federal Commission for Nuclear, Biological and Chemical Protection (ComNBC);
Federal Office for Civil Protection (FOCP)
ComNBC project supervision: Anne Eckhardt, President; Claire Walenda, Vice-president;
Niels Blatter, Member; César Metzger, Scientific Secretary
FOCP project supervision: Stefan Brem, Head of Risk Analysis and Research Coordination
ETH-CSS project management: Benjamin Scharte, Head of the Risk and Resilience Team;
Oliver Thränert, Head of Think Tank
Language editor: Mike Woods
Layout and graphics: Miriam Dahinden-Ganzoni

DOI: 10.3929/ethz-b-000505468

Table of Contents

Table of Figures

Table of Tables

Glossary

ARIMA: Autoregressive Integrated Moving Average

ARPA-E: US Advanced Research Projects Agency – Energy

BRICS: Brazil, Russia, India, China, South Africa

CBRN: Chemical, Biological, Radiological and Nuclear

ComNBC: Swiss Federal Commission for Nuclear, Biological and Chemical Protection

DARPA: US Defense Advanced Research Projects Agency

ESPAS: European Strategy and Policy Analysis System

FOCP: Swiss Federal Office for Civil Protection

GO-Science: UK Government Office for Science

GPT: General-Purpose Technology

IARPA: US Intelligence Advanced Research Projects Activity

IDEA: Inspire, Debate, Engage and Accelerate Action

IEA: International Energy Agency

mRNA: Messenger Ribonucleic Acid

NATO: North Atlantic Treaty Organization

NIC: US National Intelligence Council

OECD: Organisation for Economic Co-operation and Development

PAM: Policy Analysis Market

R&D: Research and Development

STEEP: Social, Technological, Economic, Environmental, Political

TUNA: Turbulent, Uncertain, Novel, Ambiguous

UAV: Unmanned Aerial Vehicles

UNDP: United Nations Development Programme

VUCA: Volatile, Uncertain, Complex, Ambiguous

1. Introduction

The COVID-19 pandemic abruptly changed the lives of almost everyone on the planet, causing more than four million recorded deaths (Ritchie et al., 2021), changing the way we travel, work, and socialize, as well as reducing the global economic output by trillions of dollars. It has highlighted the importance of monitoring and addressing low probability, high impact risks and reinforced the willingness to "think about the unthinkable" (Kahn, 1962; Chief of the Swiss Armed Forces Thomas Süssli in Thränert and Wenger, 2020). This report has been commissioned by the Federal Commission for Nuclear, Biological and Chemical Protection (ComNBC) and the Federal Office for Civil Protection (FOCP). It reviews methods in strategic foresight that can help organizations to deal with and reduce uncertainty. The report uses several examples from the Chemical, Biological, Radiological, and Nuclear (CBRN) domains and discusses some caveats, such as information hazards (box 4) that are particularly relevant to this context. However, the overview it provides can be useful across a wide range of strategic decision-making processes.

Why Strategic Foresight?

"Gouverner, c'est prévoir" (De Girardin, 1849, p. 143). The words government and governance are derived from the Latin verb *gubernare,* which means to steer. Clearly, the metaphorical steering of any firm or state requires some semblance of what lies ahead, otherwise one would be equally likely to steer into an iceberg as to avoid it.

The challenge, as Buckminster Fuller (1969) lamented, is that we live in increasingly complex and specialized societies of which nobody possesses an integrated high-level understanding. Furthermore, there are historical reasons for caution when it comes to grand deterministic visions for the future. High modernism, the combination of top-down planning and overconfidence, resulted in various misadventures in the 20th century, ranging from the forced relocation of farmers in the Soviet Union and Tanzania, to Mao's disastrous campaign to eliminate sparrows, to right-angled concrete monstrosities in planned cities (Scott, 1998). Consequently, one can empathize with former German chancellor Helmut Schmidt's dismissal that "anyone with visions should go to the doctor" (Ratcliffe, 2016) or the members of US Congress that abstained from voting on financing Samuel Morse's first electrical telegraph line in 1843 because they admittedly did not understand the technology (Huurdeman, 2003, p. 60). Nevertheless, there are strong arguments in favor of engaging in strategic foresight and not simply resigning to muddle through our technological development:

First, a lack of perceived agency is detrimental to our mental health. As humans, we want to be the captains of our own fate.

Second, we constantly have to make decisions that affect the future, whether they are informed by our best effort attempts at foresight or not. The 89 US representatives who voted in favor of Morse's telegraph line probably did not anticipate that their funding would accelerate the adoption of electron-based information and communication technology, which would become one of the most dynamic fields of research and application up to the present day. Still, we should be glad they voted yes.

Third, most technological developments, threats, and opportunities can be controlled, shaped, or changed more easily in their emerging phase. At a later stage, when there is a dominant design, the impacts of the technology can be monitored and predicted more easily. However, the sociotechnical system has much more inertia due to factors such as infrastructure and training investments, network effects, and supply chains. In science and technology studies, this is known as the Collingridge Dilemma.

Fourth, educational investment in generalist foresight remains modest. For example, whereas every student has mandatory history classes, "futures literacy" is rarely a part of the curriculum, even as an elective and despite a demand for futures-related tasks in many professions.

Fifth, systematic assessments of the success of many qualitative methods, such as scenario planning, remain difficult. However, there is evidence for measurable improvements in specific forecast areas, such as weather forecasts (box 2) and, to a more limited degree, in generalist forecasts, as shown by Philipp Tetlock's "superforecasters" (section 7.2.4).

In summary, strategic foresight has its limitations and the tractability of intellectual progress in it is not always obvious. However, its importance and neglectedness justify investing more resources to inform strategic decision-making as well as to work on improving foresight tools. Importantly, there is no inherent conflict between strategic foresight and approaches that focus on collecting more empirical data or increasing resilience in order to better cope with and learn from unforeseen events. On the contrary, foresight can help to prioritize which areas would profit from more data collection and resilience, whereas more data and the ability to learn from disasters can help to improve foresight.

Outline

There are a number of existing reports on foresight methods by other institutions and authors. This includes the United Nations Development Programme (UNDP) (2018), the European Commission (Wilkinson, 2017), the United Kingdom (Government Office for Science [GO-Science], 2017), Cuhls (2008), and Hyndman and Athanasopoulos (2021). This report aims to be accessible through the extensive use of examples and the reiteration of key points. Furthermore, the report aims to contribute to a better understanding of strategic foresight by offering a broad interdisciplinary view that touches upon the most relevant concepts from both the futures studies and forecasting literature literature. It complements and contextualizes them with background information from science and technology studies as well as risk management. As such, the report may help organizations to select the general approaches suitable for their goals and provides pointers on where to find more in-depth material so that promising approaches can be adapted to more specific contexts.

The report first provides background information on the nature of technology, innovation dynamics, and types of determinism in section 2 and on different models of risk and uncertainty in section 3. Section 4 provides an overview of types of foresight methods, disciplinary structures, and the governmental organization of strategic foresight in selected countries. Section 5 discusses horizon scanning, which is used to detect "weak signals" of potential technological and social trends. This is followed by section 6, which looks at established trends and the analysis of their impacts. Section 7 looks at forecasting methods based on human judgement and statistics. This includes various forms of aggregation (section 7.2), such as the Delphi method or prediction markets, as well as a discussion of how to evaluate forecasts (section 7.5). Section 8 discusses scenarios, including the use of creativity techniques, role play, and fiction to design, explore, and enhance them. Section 9 briefly touches upon modelling and simulation. Lastly, the conclusion (section 10) first reiterates the general use cases and limitations of different methods, and subsequently summarizes their implications for the more specific case of CBRN-threat anticipation.

2. Technology

Ever since the Neolithic Revolution and the subsequent Industrial Revolution(s), the interconnected set of human technologies has grown through capturing and harnessing additional natural phenomena as well as new combinations and configurations of existing technologies. **The very idea of progress rather than a cyclical future only became prevalent with the acceleration of technology-driven economic growth in the last centuries. As such, technology is inextricably linked to change and often dominates our visions of the future.** This section provides a brief overview of what technology is, how it evolves, and the degree to which it determines the future.

2.1 What Is Technology?

The term technology can refer to artifacts, sociotechnical systems, fields of research and application, and the entire "technium."

Technical artifacts: Specific, human-made objects that make use of natural phenomena to fulfill a particular function. For example, a bike, a printer, or a hammer. Technical artifacts themselves usually consist of multiple component systems or modules. These are self-contained units that fulfill a specific subfunction to enable the functioning of the whole artifact. Some of these components may be unique to a specific artifact but many, such as screws, wheels, motors, computer chips, or sensors, are incorporated into a wide range of artifacts. A complicated technical artifact, such as a fighter jet or a smartphone, has thousands of functional components organized into multiple hierarchical layers.

Sociotechnical systems: Clusters of one or more technical artifacts as well as people, organizational hierarchies, legal codes, and incentive structures that enable and define a system's functionality. For example, the Internet does not just rely on technical artifacts, such as computers and cables, but on common rules (e.g., TCP/IP) and people (e.g., content creators). Most complex technical artifacts would be dramatically less useful without the sociotechnical system in which they are embedded. If you could take a state-of-the-art fighter jet with you in a time machine to the Middle Ages, it would not enable military superiority, rather it would be hopelessly inoperable.

Fields of research and application: Collections or toolboxes of individual technical artifacts and practices. For example, biotechnology and artificial intelligence (AI) are umbrella terms that include technical artifacts with a variety of designs, components, and functions.

The "technium": The entire set of technical artifacts and engineering practices available to a civilization (Kelly, 2010).

None of these conceptions is inherently superior to the others. Their differences mostly stem from the fact that we use the term technology at multiple levels of analysis. A seat can be viewed as a technical artifact insofar as it has a useful standalone function to enable humans to rest in a comfortable position. However, if we look at a seat installed in a car it is just a component of a technical artifact. If we zoom out further, we can also look at the car as part of a socio-technical personal ground transport system that involves drivers, mechanics, fuel supply, manufacturers, and governing institutions that maintain common infrastructure, define and enforce traffic rules, or levy taxes to internalize negative environmental externalities. Similarly, we can view personal vehicles as one transport mode within the larger fields of transport technology or urban planning, which are themselves a part of the technium.

2.2 Adoption and Evolution

This section presents a selection of concepts related to technological growth and innovation that provide a useful background for strategic foresight.

Combinatorial evolution: The modularity of technology means that advances in a component in one technical artifact can easily be shared laterally and integrated into other artifacts that are only remotely related. This so-called "combinatorial evolution" allows for a faster spread of innovations than hereditary biological evolution (Arthur, 2009, Chapter 1). For example, Chris Anderson has called personal Unmanned Aerial Vehicles (UAV) the "peace dividend of the smartphone wars" (Pauker, 2013). This is because gyro sensors, which measure the rotation velocity in UAVs, used to be clunky and cost about 10,000 USD per piece in the 1990s. However, thanks to research and development (R&D) for smartphones, gyro sensors have been integrated into computer chips and now cost less than 1 USD. Similarly, UAVs may profit from a "peace dividend of the automated driving wars" if the automotive industry's demand for lasers to measure the distance of objects ("lidar"), results in their miniaturization and a significant decay of their price. Combinatorial evolution strengthens the case that the price and performance of components ought to be tracked to predict the evolution of a technical artifact.

S-curves: Futurists often extrapolate historical growth; however, taken too far, this can produce nonsensical results. For example, von Foerster et al. (1960) famously extrapolated superexponential population growth and predicted that the human population would reach infinity by Friday, November 13, 2026. There are no natural examples of infinite exponential growth curves. At some point, one or more bottlenecks slow growth down and there is saturation. For example, for animal populations, this would be at the carrying capacity of an ecosystem. A commonly observed growth pattern is that of a sigmoid or "s-shaped" curve, which involves a period of acceleration followed by an equally long one of deceleration (Smil, 2019). S-curves have therefore been used to approximate the adoption of technologies as well as price and/or performance improvements due to incremental innovations.

However, it is important to be aware of the limitations of assuming an s-shaped development pattern of "surface indicators". For example, if an artifact spreads to new markets or there is a new mechanism to achieve its purpose, the curve changes. The decline in newly confirmed cases of COVID-19 from mid- to late-February 2020 would have indicated a saturation of case numbers. However, the local success in suppressing the infectious disease was soon outmatched by its global spread. Similarly, Hubbert (1956) introduced the notion that the annual production of non-renewable resources follows a bell-shaped curve, which corresponds to an s-curve in terms of the total coal, oil, and gas extracted. For about forty years, it became established wisdom that the US had reached its "peak oil" production in 1970. Yet, with the adoption of the extraction technique of hydraulic fracturing, the US managed to increase its oil and gas production again in the 2010s, reaching new record highs.

Creativity: Creativity is generally understood as the ability to form something new that is somehow valuable. There are a number of myths about creativity (Burkus, 2013). Out of the big five character traits, only openness to experience has an established positive correlation with creativity. Domain-specific learned skills in fine arts, such as painting styles, are sometimes overattributed to creativity. At the same time, creativity does not require creating something entirely new, rather it is usually about adding on, improving, and combining previous concepts. As such, there is usually not a single "eureka moment" and if you have a creative idea or innovative new product, others may only see its value in hindsight. For example, Campanario (2009) discusses the resistance of the scientific community to many discoveries that have later been awarded the Nobel Prize. Lastly, constraints do not necessarily dampen creative potential. In fact, constraints such as limited material resources to fulfill a function, challenging operational environments, or artifact users with specific needs can often enhance creativity.

Incremental innovation: Incremental or sustaining innovation involves the gradual change of technologies bringing about an improvement in prices and/or performance over time. On an artifact level, this mainly works through the replacement of component materials and structural deepening with additional components added to the artifact (Arthur, 2009, Chapter 7). Additionally, economies of scale and learning decrease the per unit production cost of goods. This is also referred to as Wright's (1936) Law.

Disruptive innovation: Disruptive or transformative innovation refers to a paradigm shift in technology, business models, and value chains. Future disruptive innovation cannot be anticipated by the extrapolation of performance within the current paradigm or by monitoring innovation only within the target domain. However, this does not mean that disruptive innovation is entirely unpredictable. There have only been a few scientific revolutions that have changed our knowledge in a discontinuous way. Almost all disruptive innovation consists of incremental improvements in a niche outside of the target domain, which eventually remove the obstacles to commercial viability and allow the niche product to displace the incumbent paradigm.

For example, inventors such as Georges-Louis Le Sage, Francis Ronalds, Samuel Thomas von Sömmering, and Pavel Schilling had been experimenting with prototypes of the electrical telegraph for decades before it started to displace mail by horses, trains, and steamboats in the 1840s. Only after a reliable source of electric currents (Volta, 1800), a reliable detection of electric currents (Ørsted, 1820), and an understanding of electric transmission over large distances (Henry, 1830) were established two groups around Samuel Morse and William Fothergill Cooke managed to develop the telegraph to commercial viability. Similarly, the first vaccines based on Messenger Ribonucleic Acid (mRNA), which were developed for SARS-CoV-2 by Moderna and BioNTech, might be viewed as disruptive. However, the scientific interest in the possibility of using mRNA for therapeutics started as far back as 1961. The path to commercial viability required progress in the reliable production of mRNA (Mullis, 1984) and the replacement of a molecule that causes allergic reactions (Karikó and Weissman, 2005).

There are two archetypal modes of organizing transformative innovation (Sen, 2017). **First, there is the "island" model, the deliberate segregation of a group of innovators from operational concerns.** The goal here is to "herd nerds," i.e., to bring innovators together and decouple them from the immediate concerns and interests of operational units to limit bureaucracy and the transaction costs for research coordination. Notable examples of this model include Lockheed Martin's "Skunkworks" or Xerox's Palo Alto Research Center. **Second, there is the "bridge" model, which involves open innovation networks that connect units across organizational boundaries.** The idea here is to connect to and follow the R&D conducted at places such as universities or start-ups.

Implementation and restructuring lags: There are two principal causes of delay between technological breakthroughs and their observable economic impact as "total factor productivity growth." One is that many technologies become more useful with higher adoption rates due to economies of scale and network effects. The other is that while technologies create new "liberties of action" (Sawhney and Lee, 2005), it takes time to discover and develop newly enabled business processes, value networks, and complements. For example, the Internet represented a shift from one-to-many to many-to-many communication. Der Spiegel already had a website in 1994, whereas Facebook only launched a decade later. However, Facebook grew exponentially because its business model leverages many-to-many communication. This delayed impact is also captured in the Gartner Hype Cycle. The company that goes furthest in embracing lags may be Nintendo, whose philosophy of "lateral thinking with withered technology" explicitly focuses on new combinations of well-understood and cheap technology rather than the cutting-edge (Yokoi and Takefumi, 1997; Reynolds, 2016). The important takeaway for foresight is to not only focus on new technologies but to consider new uses of recently commercialized technologies. This is particularly relevant for foresight on adversarial threats from non-state actors, which often lack large financial resources or deep technical expertise.

General-purpose technologies: General-purpose Technologies (GPTs) are artifacts, sociotechnical systems, or fields of research and application that can be applied widely across the economy and create significant spillover effects. They are referred to as "engines of growth" (Bresnahan and Trajtenberg, 1992) because of their pervasiveness as well as their high number of innovational complementarities. The intensity of the early adoption of GPTs may be indicative of future total factor productivity growth. Suggested examples of past and emerging GPTs include the domestication of plants, the domestication of animals, the smelting of ore, the wheel, writing, bronze, iron, the water wheel, the three-masted sailing ship, printing, the factory system, the steam engine, the railway, the iron steamship, the internal combustion engine, electricity, the automobile, the airplane, mass production, the computer, lean production, the Internet, modern biotechnology (Lipsey, Carlaw, and Bekar, 2005, pp. 131–216), and AI (Klinger, Mateos-Garcia, and Stathoulopoulos, 2018). This list suggests that major innovations in energy, transport, and information and communication technology are of particular interest for long-term foresight.

Industrial Revolution(s): The term Industrial Revolution was first used to describe the development of the UK between 1760 and 1840 but has been generalized to describe the rapid onset of economic change through the application of technology. Mechanization and electrification are often counted as the First and Second Industrial Revolution. However, there is no real agreement amongst authors on how many industrial revolutions there have been and when they occurred (figure 1). Notwithstanding claims that previous industrial revolutions were "linear" whereas the current pace is "exponential" (Schwab, 2016, par. 3), the global economy has been growing exponentially for many centuries. The long-run development of the world economy is best described as an accelerating exponential growth pattern, with the fastest growth occurring between 1950 and 1973, followed by a slight deceleration (figure 2). As such, the concept of another industrial revolution is of limited value for foresight unless it is tied to specific GPTs.

Technological determinism: Technological determinism describes the ways in which technology can be an autonomous force that shapes society. **Technological politics** refers to intentional design choices. Technological decisions can be "similar to legislative acts or political foundings that establish a framework for public order that will endure over many generations" (Winner, 1980, p. 29) and a "material validation of the social order" (Feenberg, 2010). Examples include fences, surveillance technology, encryption, and the broad linear Parisian boulevards whose construction facilitated the suppression of riots.

The ideas of **technological momentum** (Hughes, 1983) and technological frames (Bijker, 1995) emphasize the inertia and constraints arising from established sociotechnological systems. This inertia emerges from the logic of sunk costs: assets have been bought, standards set, infrastructure built, employees trained, interactions routinized, and interests entrenched, all of which constrain subsequent decisions. **Technological sleepwalking** combines technological momentum with a lack of foresight, resulting in a "technics-out-of-control" theme. For example, Winner warned that if technological change happens too quickly, societies face the prospect "of going adrift in a vast sea of unintended consequences" (1977, p. 89). Collingridge (1982) advocates for more small-scale empirical experiments to reduce uncertainty as much as possible before technological momentum becomes too great.

Lastly, Dafoe (2015) states that technology enables new sociotechnical configurations, some of which confer advantages in competition. Hence, when there is sufficiently intense and prolonged economic and/or military competition sociotechnical systems **will evolve to**

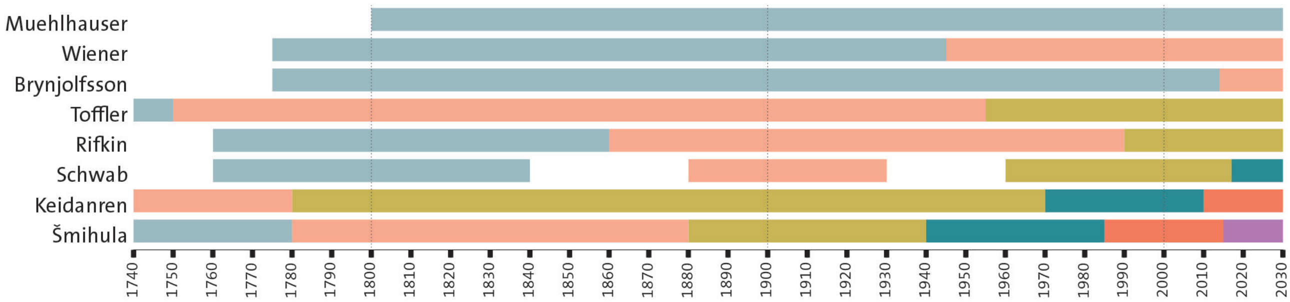

Figure 1. Timespans of supposed Industrial Revolutions, machine ages, waves, or levels of society 1740–2030.
Data from Muehlhauser (2017), Wiener (1948, pp. 27–28), Brynjolfsson and McAfee (2014), Toffler (1980, p. 14), Rifkin (2011, p. 2), Schwab (2017, pp. 6–7), Keidanren (2018, p. 5), and Šmihula (2010, p. 62).

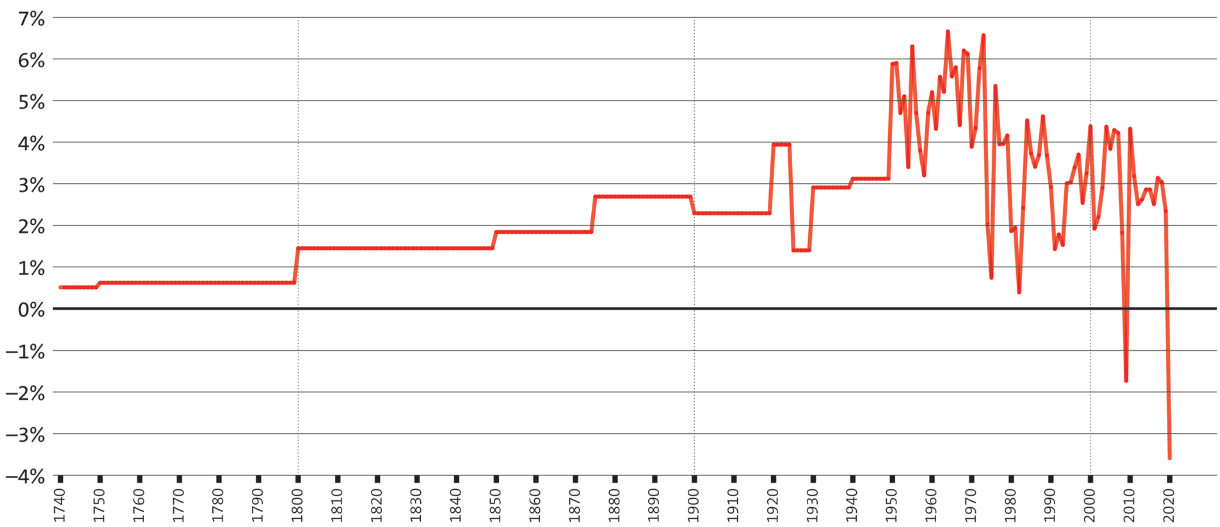

Figure 2. Annualized gross world product growth rate 1740–2020.
Data for 1700, 1750, 1800, 1850, 1875, 1900, 1920, 1925, 1930, 1940 and 1950 from DeLong (1998, p. 6), for 1951 to 1960 from Maddison (2001, p. 332), for 1961 to 2020 (World Bank, 2021a).

become more adapted to success in the economic and/or military competition. In the short run, technological determinism is "self-evidently untrue: human beings construct machines, not the reverse" Williams (2002). However, in the long run, the semi-anarchic self-help system of states (Waltz, 1979) makes it challenging to suppress technological developments on a global scale.

The larger the set of actors that have the capability to develop or deploy a technology, the higher the chance that it will be developed and deployed (Bostrom, Douglas, and Sandberg, 2016). Hence, in terms of risk reduction, it is sometimes easier to focus on global public goods that only rely on the single best unilateral or collective effort of any state or group of states rather than on mutual restraint or the weakest link amongst many or all states (Barrett, 2007, p. 20). For example, increasing global resilience to biological threats through the maintenance of a mRNA production capacity at all times (Sandbrink and Shattock, 2020; Pandemic Preparedness Partnership, 2021, p. 31) can be provided with less international coordination than the global prevention of spillover events through the closure of wet markets or more stringent safety verification in labs.

Key Takeaways for Foresight

- What we call a "technology", such as a car, often contains many technical subsystems.
- Innovation in one component of an artifact can quickly diffuse to artifacts in other domains. Hence, specialists in one domain might want to collaborate with specialists from other domains and generalists for foresight.
- "Disruption" is hard to anticipate. At best, the target function can be defined in a technologically neutral way. Subsequently, one can look for niche domains or recently enabled business models that might fulfill this function and assess their key bottlenecks.
- Creativity is an ingredient of innovation and foresight. However, contrary to popular beliefs it is not inherently tied to fine arts, a lack of constraints, or a single "eureka moment".
- Aside from future and emerging technologies, also consider that the use of technologies that have matured in the last 10 years is likely not fully explored yet.
- History suggests that potential or early-stage general-purpose technologies, such as major changes in energy, transport, as well as information and communication technology, are of particular interest for long-term foresight.
- There is a dilemma between uncertainty and control in technology governance. Over time, the impacts of technology become clearer but sunk costs make it harder to change directions.
- It is difficult to suppress the development of high-risk technology with economic and/or military benefits on a global scale and over prolonged periods of time because of the coordination that this would require.

3. Risk and Uncertainty

In the classic conception by Knight (1921), **risk refers to incertitude that can be objectively measured through scientific approaches, whereas uncertainty is incertitude that can be assessed by subjective estimates. Foresight is ultimately aimed at reducing uncertainty about the future.** However, it does not prescribe how decision-makers deal with risk and uncertainty. That is the role of risk management.

There are competing conceptualizations of risk and uncertainty. Section 3.1 provides some background on why uncertainty is an inherent attribute of futures studies, and section 3.2 provides a brief overview on four conceptions of uncertainty. However, **subsequent sections will only rely on two uncertainty concepts: the known-unknown matrix** (section 3.2.1) **and levels of uncertainty** (section 3.2.4).

Note that some authors bundle uncertainty with other terms into acronyms to describe the strategic environment as VUCA (Volatile, Uncertain, Complex, Ambiguous) or TUNA (Turbulent, Uncertain, Novel, Ambiguous). Novelty is a given in times of technological progress. Volatility and turbulence are part of complexity and uncertainty with an emphasis on thresholds and non-linear effects that can lead to chaotic behavior. Complexity is a cause of uncertainty and ambiguity refers to contested framings or value trade-offs.

3.1 Fundamental Challenges

Fundamental indeterminacy: As Laplace (1814) has written "we may regard the present state of the Universe as the effect of its past and the cause of its future". Hence, for Laplace's demon – which knows the physical laws of the Universe, is aware of the precise location of all the particles in it and possesses the intellect to analyze this data – "nothing would be uncertain and the future just like the past would be present before its eyes." However, there are some fundamental challenges to this deterministic outlook. First, there are irreversible thermodynamic processes, meaning it is not possible to physically reconstruct all past positions and momenta of particles from their current state. Second, the Universe may be fundamentally probabilistic at the level of analysis of quantum physics, as stipulated by the Copenhagen interpretation. This quantum indeterminacy would aggregate and make a deterministic understanding of the future impossible.

Kolmogorov Complexity: Scientific disciplines focusing on interactions at small scales are more exact than social sciences. Sociology is "applied psychology," psychology is "applied biology," biology is "applied chemistry," chemistry is "applied classical physics," and, at some point in the future, classical physics ought to be "applied quantum mechanics" (Munroe, 2008). The increasing levels of abstractions of physical phenomena are necessary due to the greater scope of the studied systems. They are also useful due to multiple levels of emergent phenomena, meaning that properties that no parts have on their own can emerge from scale and interconnected systems. Such properties may be described fairly accurate without the parts as microfoundations. The Kolmogorov complexity of an object can be defined as the length of the shortest computer program that produces a simulation of the object that provides a reliable approximation of the output dimensions you are interested in (Li and Vitányi, 2019, pp. vii–viii). While any given star has more atoms than all of humanity put together, simulating an approximation of a star's gravity is less complex than simulating an extended conversation between two specific humans. Since we do not rely on a highly detailed simulation of the location of atoms in the star, a lot of information can be compressed without meaningful loss. In contrast, it appears hard to reduce information about the human brain without affecting the outputs that we are interested in. Of course, a globalized socio-technical system consisting of billions of human brains and technical artifacts has a Kolmogorov complexity that is many orders of magnitude greater than that of systems that can be computed with physics formulas.

Chaos Theory: Many complex systems are "chaotic" insofar as they are dynamic systems that are highly sensitive to initial conditions. This means that the present state of the system determines its future, but due to interconnections, threshold effects, and feedback loops, the system's approximate present state does not approximately determine its future state. As small initial differences can yield widely diverging outcomes for such systems, long-term predictions of their behavior are very difficult.

Consequences for foresight: First, given the complexity of predicted systems, forecasts tend to rely on high-level abstract concepts as data points. The downside of this is that we may track surface indicators without an understanding of the relevant microfoundations that cause increases or decreases in them.

Second, for similar reasons, the scope of forecasts is usually limited to specific artifacts, fields of research, countries, or regions, even though there are interdependencies that cut across the technium and the globe. The future is not just the aggregation of separate socio-technical trends, such as cheaper DNA-testing, larger artificial neural networks, superannuation, or access over ownership models, but the interactions between all of them.

Third, foresight and forecasts about social, technological, or political developments exist within the system about which a prediction is being made. Forecasts influence our expectations of the future in the present and, thereby, our anticipatory behavior. In other words, predictions can have self-fulfilling or self-negating properties. For example, there is a bandwagon effect with regards to voting and elections (Schmitt-Beck, 2015). This means that if public predictions say that candidate Jane Doe will win an election, this increases her chances of winning it, which may in turn incentivize politically biased election forecasts. The same idea can be applied to stock-related forecasts, where the phenomenon is known as "talking your book." An example without motivated reasoning would be the forecast by Goldman Sachs (Wilson and Purushothaman, 2003) that the economies of Brazil, Russia, India, and China (BRIC) will be larger by 2040 than those of the G6 (US, Japan, Germany, UK, France, and Italy). This has contributed to the creation of the BRICS group (incl. South Africa) that was quite active in the late 2000s to early 2010s, even though its members have very little in common in terms of culture, demographics, geography, values, or governance systems.

3.2 Conceptualizations

3.2.1 Knowns and Unknowns

A simple but useful way to think about uncertainty is represented in a known-unknown matrix. Variations of the concept have existed for a long time, going back to the Persian poet Ibn Yamin (1286–1368), if not further (Verity, 2012, p. 69). However, the concept gained particular traction after it was used by former US secretary of defense Donald Rumsfeld to address missing evidence with regards to weapons of mass destruction in Iraq: "there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tends to be the difficult ones" (Rumsfeld and Myers, 2002).

Known Knowns (≈facts)	**Known Unknowns** (≈uncertainty levels 2–4)
Unknown Knowns (≈hidden knowledge)	**Unknown Unknowns** (≈uncertainty level 5)

Table 1. Known-unknown matrix. The first term refers to awareness of an issue, the second to the degree of knowledge about an issue. The uncertainty levels are explained in section 3.2.4.

For known knowns, we understand the causal links and can manage facts and requirements. For known unknowns, we have some knowledge about the probability and impact of such events. There is no universal agreement on what unknown knowns are. Some have taken the concept to mean knowledge that is repressed or embedded in traditions and rituals without its practitioners being aware of why these practices are beneficial (Henrich, 2015). For the purposes of strategic foresight, the most useful concept frames it as knowledge, which is unknown within an organization's decision-making structure, but which exists somewhere within a wider community. Lastly, unknown unknowns are events with non-zero probability that no one within an organization or the wider community anticipates.

The known-unknown matrix makes two primary distinctions: First, whether an organization has adequate knowledge about a topic and, second, the degree to which the larger scientific community understands the causal relationships involved. From a foresight perspective, we are particularly interested in shifts from unknown knowns to known knowns and unknown unknowns to known unknowns. Importantly, these are two separate processes that may require separate methods. The first one corresponds to William Gibson's famous quote that "the future is already here – it's just not very evenly distributed." (Quote Investigator, 2012) and might be addressed through more systematic information integration. The second one requires more creativity and ideation.

3.2.2 Contested Evaluations of Outcomes

Knowledge about …	… outcomes	
	not problematic	problematic
… probabilities		
not problematic	**Risk**	**Ambiguity**
problematic	**Uncertainty**	**Ignorance**

Table 2. Contrasting states of incomplete knowledge.
Adapted from Stirling, 2008, p. 99.

Stirling (2008) expands on the Knightian distinction between risk and uncertainty with a normative dimension. Risk applies to fields where past experience and scientific models foster confidence in probabilities and outcomes. Under these conditions, conventional risk assessments offer a rigorous approach. In the case of uncertainty, the available empirical information and analytical models do not constitute a definitive basis for assigning probabilities. Ambiguity is defined by disagreements over the selection, bounding, measurement, prioritization, or interpretation of outcomes. Contending ways of framing the issues may invoke ecological, economic, or social factors and may value their trade-offs differently. Finally, there is the condition of ignorance. Here, neither probabilities nor outcomes can be fully characterized. This is the domain of "unknown unknowns."

3.2.3 IRGC Model

A distinctly non-Knightian conception of risk that does not juxtapose risk with uncertainty is used by the International Risk Governance Council (IRGC) (2017, pp. 17–18) at the Swiss Federal Institute of Technology Lausanne. Renn (2008), van Asselt (2005), and Aven and Renn (2010) distinguish between four categories of risk: simple risk, uncertain risk, complex risk, and ambiguous risk. Simple risks refer to situations in which causes and effects are understood and there is no normative disagreement. In complex risks, the causal chains are difficult to establish. In uncertain risks, it is difficult to assign probabilities to the occurrence of an event and its consequences. Ambiguous risk refers to situations with disagreements on values and norms or different interpretations regarding the nature and consequences of the risks. This typology of risks can be connected to different levels of stakeholder involvement (IRGC, 2020, pp. 5–6). Simple risks can be managed through existing routines by regulatory bodies. For complex risks, scientific knowledge should be maximized through consultation of external researchers. For uncertain risks, affected stakeholders should be involved as well. Lastly, ambiguous risks also require the inclusion of civil society for societal debate.

3.2.4 Levels of Uncertainty

Another school of thinking popular at RAND, the Delft University of Technology, and the Society for Decision-Making under Deep Uncertainty (Marchau et al., 2019) argues that the Knightian distinction between risk and uncertainty is not fine-grained enough. Rather they offer a five-step spectrum, which ranges from complete certainty to complete ignorance with intermediate levels of uncertainty without any separate notion of risk (Walker et al., 2013; Figure 4).

Level 1 uncertainty represents a situation in which one admits that one is not absolutely certain. For example, one can use a sensitivity analysis of model parameters to see the impacts of small perturbations of model input parameters on the outcomes.

Level 2 uncertainty can still be defined statistically in the form of either a single prediction with a confidence interval or several predictions (scenarios) with associated probabilities.

In level 3 uncertainty, several alternatives can be enumerated and graded in terms of perceived probability. This can be trend-based scenarios based on alternative assumptions about key variables (e.g., different assumptions about Gross Domestic Product (GDP) growth).

Level 4 uncertainty refers to several possible alternatives that cannot be ranked in terms of perceived possibility. This can be due to a lack of data or understanding about functional relationships being studied and/or disagreements about perceived likelihoods.

Lastly, level 5 uncertainty or "deep uncertainty" represents recognized ignorance, and includes unknown unknowns that are not anticipated at all.

	Level 1	**Level 2**	**Level 3**	**Level 4**	**Level 5**
Context	A clear enough future (with sensitivity)	Alternate futures (with probabilities)	Alternate futures (ranked)	A multiplicity of plausible futures (unranked)	Unknown future
System model	A single system model	A single system model with a probabilistic parameterization	Several system models, one of which is most likely	Several system models, with different structures	Unknown system model
System outcomes	Point estimates with sensitivity	Several sets of point estimates with confidence intervals, with a probability attached to each set	Several sets of point estimates, ranked according to their perceived likelihood	A known range of outcomes	Unknown outcomes
Weights on outcomes	A single estimate of the weights	Several sets of weights, with a probability attached to each set	Several sets of weights, ranked according to their perceived likelihood	A known range of weights	Unknown weights

Table 3. The progressive transition of levels of uncertainty from complete certainty to total ignorance. Adapted from Walker et al. (2013).

Box 1: The Zoo of Risk and Uncertainty Metaphors

Nassim Taleb (2007) has introduced the black swan as a metaphor for unknown unknowns. The success of his book with the same title following the 2008 financial crisis inspired the emergence of an entire zoo of metaphorical animals and mythical creatures, which have been used by futurists and risk analysts to give memorable names to their preferred categorizations. A non-comprehensive list includes "grey swans" (Taleb, 2007, p. 37) "dragon-kings" (Sornette, 2009), "black elephants" (Gupta, 2009), "dirty white swans" (Zeisberger and Munro, 2010), "black turkeys" (Siegel, 2010), "red swans" (Woo, 2011), "gray rhinos" (Wucker, 2013), "black jellyfish" (Sweeney, 2015), and "green swans" (Elkington, 2020). Furthermore, already before 2007, Klinke and Renn (2002) have used Greek mythology to classify risks as "Cassandra," "Cyclops," "Damocles," "Medusa," "Pandora," and "Pythia".

In general, risk metaphors should be used in moderation. Even if the categories themselves can be meaningful, the metaphors are sometimes duplicative and require more context for comprehension than more descriptive terms. As an example, consider their added value in the following sentence: "A predictable gray rhino, in the sense that danger of a general European war was well known, but also a surprising black swan, in the sense that contemporaries seemed bewildered by its outbreak, the First World War was a true dragon-king event in terms of its vast historical consequences" (Ferguson, 2021, Chapter 3).

Key Takeaways for Foresight

- Foresight can reduce but not eliminate uncertainty due to probabilistic behavior at the smallest scales, complexity, and chaos theory.
- Forecasts and other foresight outputs within the boundaries of a predicted system can have self-fulfilling or self-negating effects.
- Some knowledge already exists outside an organization and may be integrated into it (unknown-knowns), other knowledge must be created from scratch (unknown-unknowns).
- Complex and uncertain risks require the involvement of science (and affected stakeholders), whereas civil society can primarily contribute to normative questions.
- We can reduce the level of uncertainty about the future by defining new scenarios (level 5 to level 4), ranking them (level 4 to level 3), assigning probabilities to them (level 3 to level 2), or collapsing them into one clear enough future, the robustness of which to perturbations can be tested (level 2 to level 1).

4. Strategic Foresight

There is a wide variety of similar and overlapping terms that have been used to describe methodologies for improving our understanding of the future. These include "foresight," "futurology," "futures studies," "futures research," "long-range planning," and "strategic management." The definitions and popularity of these terms have varied amongst disciplines and over time. This report uses the term strategic foresight because it enjoys popularity across disciplines, and because a wide variety of methods can be included under its umbrella. According to Coates (1985), "foresight includes qualitative and quantitative means for monitoring clues and indicators of evolving trends and developments and is best and most useful when directly linked to the analysis of policy implications. … Foresight in government cannot define policy, but it can help condition policies to be more appropriate, more flexible, and more robust in their implementation, as times and circumstances change … It is not planning – merely a step in planning" (p. 343). Figure 3 highlights this by showing a policy cycle that involves a feedback loop based on empirical evidence and an additional loop based on anticipation. The former reduces uncertainty about the past and present, whereas the latter reduces uncertainty about the future. Both inform policy formulation but should not be confused with it. For an overview on public policy strategies to cope with identified risk and uncertainties, see Li et al. (2021, pp. 526–528).

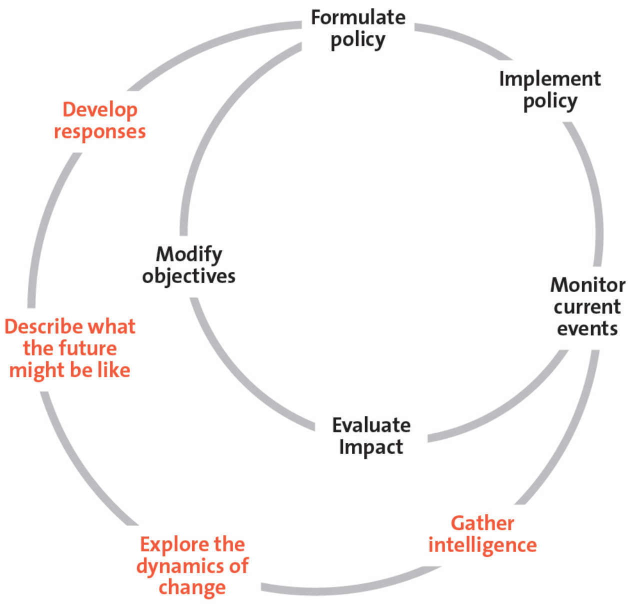

Figure 3. The futures bridge.
Adapted from Government Office for Science, 2017, p. 5.

The particular emphasis on *strategic* foresight means that the focus primarily lies on the use of foresight methods to inform higher-level and longer-term planning processes, such as the impact of geostrategic shifts on supply chain risks. Its counterpart is shorter-term foresight on a more operational level, such as demand forecasting for the next few days.

In terms of structure, this section discusses some common dimensions along which foresight methods differ, highlights disciplinary clusters, and provides examples of the organization of foresight activities in governments. The subsequent sections on horizon scanning (section 5), trend analysis (section 6), forecasting (section 7), scenarios (section 8), as well as modelling and simulation (section 9) discuss the various specific methods that fall under the umbrella term of strategic foresight with their respective weaknesses and strengths.

4.1 Types of Methods

Strategic foresight can be approached with a wide array of methods, each of which may be suitable for different purposes and contexts. Furthermore, most methods are not mutually exclusive and can be combined in various ways. The variety may appear overwhelming at first, but the sources of understanding in all of them are ultimately combinations of expert intuition, causal reasoning, and data from the past, the present, and simulated futures. Furthermore, it can help to categorize methods along different axes. For example, Popper (2008, p. 66) has introduced a "foresight diamond" with the methods arranged along two dimensions: "creativity vs. evidence" and "expertise vs. interaction" (Figure 3). The following describes four of the most useful axes along which methods can be distinguished.

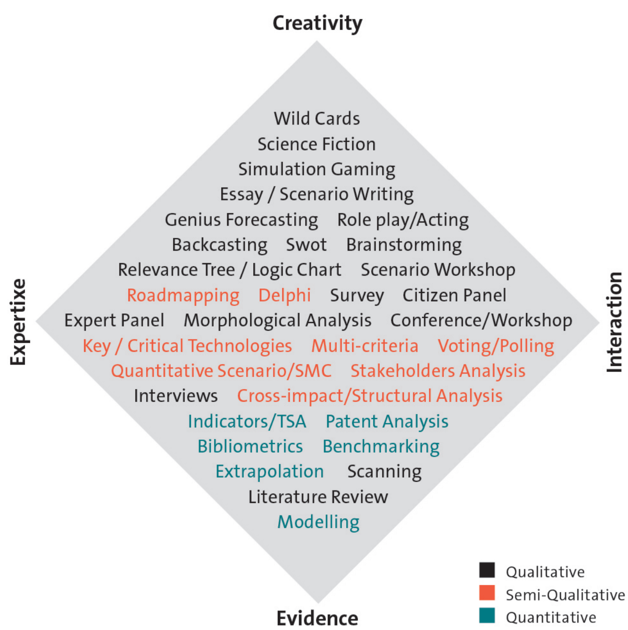

Figure 4. The foresight diamond.
Reprinted from Popper (2008, p. 66).

Quantitative vs. qualitative: This cleavage has substantial overlaps with distinctions between evidence-based vs. creativity-based or structurally closed vs. structurally open methods. However, there are crucial differences. For example, a literature review is both a qualitative and evidence-based method. Quantitative methods are particularly used in horizon scanning and shorter-term forecasts. Qualitative methods are particularly useful for normative contexts, the interaction between complex organizations, or exploring the consequences of scenarios and structural trends.

Exploratory vs. normative: Methods can be exploratory in the sense that they are aimed at finding plausible futures and identifying interesting interactions in them. Examples of exploratory methods include horizon scanning, the definition of emerging issues, and role play. Other methods may be better described as predictive, such as using time-series analysis to forecast future data points. However, in both exploratory and predictive approaches, the value judgement about the preferability of certain outcomes is not an explicit part of the methodology. In contrast, normative methods, such as "backcasting," explicitly focus on finding and examining preferred futures rather than the most probable futures. The resulting normative scenario may then be used to find actions by different actors to help make this vision a reality.

Participatory vs. expert-based: Traditionally, most futures methods focus on bringing together experts with academic or institutional credentials. With the advent of prediction platforms, it has also become feasible to select experts based on their forecasting track record. In contrast, participatory futures directly involve anyone that is interested in participating in exploring potential futures. Their aim is often to democratize long-term thinking and create collective images of the future that turn into collective actions and behaviors in the present.

Short-term vs. long-term: Quantitative methods, such as time-series analysis, are used more in short-term contexts, whereas some creativity-oriented techniques, such as science fiction prototyping, are mainly suitable for exploration of longer-term futures. However, many methods can be employed for different time horizons. The degree of long-term thinking mainly differs between policy areas. Nuclear waste is probably the issue with the longest socially accepted time horizon. For example, the US Department of Energy has set up the Human Interference Taskforce to design pictograms that deter human intrusions into waste repositories for about 10,000 years. The German Bundestag went even further and passed a law for nuclear waste repositories to last a million years. Climate change is another topic with a long time horizon, including policy goals such as limiting global warming to 2 degrees Celsius by 2100. In energy and transport, it is still common to look about 30 years into the future. In contrast, in fast-paced areas such as AI and genome editing, foresight often only focuses on the next 2 to 10 years. Foresight that focuses on forecasting events, such as elections, is even more short-term and usually only looks ahead for about two years.

4.2 Disciplinary Clusters

In addition to the nominal "futures" community that focuses on methods across application areas, there are several disciplinary research and practice communities that are involved in future-focused activities. The following list provides an overview. However, it is not exhaustive, and the discussed disciplinary clusters can have overlaps. For example, while Philipp Tetlock is a political scientist who has done substantive work on improving the training of intelligence analysts, he is also a professor at the Wharton Business School. CBRN threat anticipation is primarily part of the civil protection and military clusters.

Futures (everyone): Futures studies, futures research or just "futures" is the cluster of research that originated in the pursuit of the idealistic, critical, and humanistic goals pioneered by Ossip Flechtheim's (1971) futurology. In the late 1960s and the 1970s, futures studies received a large boost mainly due to environmental concerns and the seminal Club of Rome study Limits to Growth (Meadows et al., 1972). As such, it is associated with normative foresight. At the same time, the use of the plural "futures" indicates the turn away from forecasting and a single clear enough future towards the use of multiple scenarios. Today, futures studies is its own discipline within the social sciences and mainly associated with the use of qualitative methods.

Military (states): Strategy originally stems from a military context, where it denotes the overall plan and means to achieve an outcome. Militaries have a long tradition of war games which aim to anticipate interactions on the battlefield. In the West, military (grand) strategy and technology foresight gained particular traction after the Second World War as the United States geared up for a long strategic competition with the Soviet Union. In 1948, the US Air Force founded the RAND Corporation to help formulate policies and strategies for the Cold War. In the 1950s, RAND used and advanced game theory, scenarios, as well as the Delphi method. Intelligence services also employ political analysts that make predictions about threats and foreign political developments in a national security context.

Civil protection (states): After the end of the Cold War, there has been a shift towards broader definitions of national security and a more comprehensive approach to risk management by states. For instance, the Sendai Framework for Disaster Risk Reduction suggests a multi-hazard approach to disaster risk management

(United Nations Office for Disaster Risk Reduction, 2015, pp. 10–11) and the Organisation for Economic Co-operation and Development (OECD) (2014, pp. 4–5; 2018a, pp. 28–29) even recommends an all-hazards approach. In order to compare risks, national risk assessments define reference scenarios and assess their probability. For naturally recurring hazards, such as earthquakes, this is usually done based on long-term historical data. Societal and technological hazards require more foresight activities. The most common timeframe of disaster risk assessments in the OECD (2018b, p. 28) is to look five years ahead.

Economics (states): Econometric models have a long tradition in GDP forecasting, with the caveat that these models often capture technological developments in catch-all terms such as total factor productivity. The societal impact of technology is discussed in terms such as skill-biased technological change, general-purpose technologies, and technological unemployment. The understanding that adoption speed and intensity of new technologies is decisive for the economic success of states, has led to a push for technology foresight as a public policy tool to improve civilian science, technology and innovation systems in Europe in the 1990s (Miles, 2010).

Business Management (firms): Future-focused business management literature has been developed since the 1950s by scholars such as Igor Ansoff (1965), Michael Porter (1997), and Peter Drucker (2014), using terms such as "strategic management," "innovation management," and "change management." This includes the development of Shell's 2x2 scenarios, horizon scanning, business war games, as well as various forms of industry and demand forecasting. Key ideas include disruptive innovation and business model innovation. Global consulting firms have also built up futures-related teams as part of their service in assisting companies and public sector actors with strategic management.

Finance (firms): The stock market can be framed as a prediction market, in which value investors, such as Warren Buffett (the "Oracle of Omaha"), bet on future dividends and valuations of companies. Stock analysts regularly predict whether a specific stock will rise or fall. Similarly, venture capital firms have a stake in understanding future trends and business models. The insurance and re-insurance industries offer risk transformation services for most hazards, which means firms need to realistically assess their likelihoods and impacts to be competitive and profitable. Note that in finance the term "futures" refers to contracts that obligate parties to buy or sell an asset at a predetermined date and price.

Urban planning (cities): Foresight methods are used in urban planning as a lot of transport infrastructure is expensive and has long lifecycles. At the same time, urban planning often aims to implement normative visions of the future, such as smart, green, pedestrian-friendly, or egalitarian cities. Currently, there is also a high modernist renaissance, with massive projects for state-planned cities in Saudi Arabia (Neom), China (Xiong'an), and Indonesia (East Kalimantan).

4.3 Governmental Structures

This section provides a non-exhaustive overview of the main foresight activities in selected countries and international organizations. Governments differ in the intensity and centralization of their foresight efforts, as well as their degree of collaboration with independent think tanks and research institutes. Foresight tools can be applied to sectoral issues within departments or agencies as well as to cross-cutting issues in a whole-of-government manner. Political systems that directly integrate strategic foresight into the processes of their executive branch, such as the UK or Singapore, have a more direct causal path from foresight results to high-level decision-making than decentralized systems, such as Switzerland and Germany.

4.3.1 Europe

European Union: Foresight efforts at the European Commission are conducted by the Directorate-General for Research and Innovation, which is responsible for EU policy in that area; the Institute for Prospective Technological Studies in Seville, which is part of the Joint Research Centre, the Commission's science and knowledge service; and Inspire, Debate, Engage and Accelerate Action (IDEA), an advisory service of the Commission, formerly known as the European Political Strategy Centre, which is intended to provide innovative and forward-looking ideas. The President of the European Commission, Ursula von der Leyen, has assigned Maroš Šefčovič as Vice-President-designate for Interinstitutional Relations and Foresight. He initiated an annual foresight report (European Commission, 2021) and is setting up the new EU Network of Strategic Foresight, which will bring institutions from the EU and member states together.

In the European Parliament there is the Panel for the Future of Science and Technology. It consists of Members of the European Parliament and is responsible for external expert assessments of the impact of technologies for the use in parliamentary committees. Furthermore, the European Parliamentary Research Service has a Scientific Foresight Unit and a Global Trends Unit, which has published a "Global Trendometer" since 2016. The European Strategy and Policy Analysis System (ESPAS) is an EU inter-institutional strategic foresight network, that provides a framework for collaboration between the EU institutions engaged in the topic. It hosts the research library Open Repository Base on International Strategic

Studies (ORBIS), organizes an annual conference; and publishes its “Global Trends Report” each legislative cycle (ESPAS, 2019).

United Kingdom: The GO-Science futures team advises the UK prime minister and supports embedding strategic foresight tools across the Civil Service to support policymaking. The futures team provides a toolkit (GO-Science, 2017), an emerging technologies database, a trend deck, megatrends, foresight reports, and a guide to futures thinking and foresight. It also collaborates with the Cabinet Office through the Horizon Scanning Programme Team and supports the regular horizon scanning meeting of the departments, in which permanent secretaries consider the long-term impact of key futures topics. The Ministry of Defence’s think tank, the Developments, Concepts and Doctrine Centre (2018), publishes its Global Strategic Trends report that looks 30 years into the future. Furthermore, in 2020, the government launched a cross-departmental reputation-based prediction platform called Cosmic Bazaar (section 7.2.3). Outside of government, the lottery-funded Nesta foundation conducts many foresight activities.

The UK is also a center of longtermist thinking. This includes the Future of Humanity Institute in Oxford, the Centre for the Study of Existential Risk in Cambridge, and the All-Party Parliamentary Group for Future Generations, which aims to raise “awareness of long-term issues, explore ways to internalise longer-term considerations into decision-making processes, and create space for cross-party dialogue on combating short-termism in policy-making.” In 2019, the Well-being of Future Generations Bill was introduced to Parliament. If passed, this piece of legislation would require public bodies to publish future generations impact assessments and establish a commissioner for future generations.

France: France Stratégie, formerly called the *Commissariat General au Plan* and the *Centre d’analyse stratégique,* is the foresight unit of the French prime minister. It largely focuses on the impact assessment of public policies and coming reforms. The *Délégation à l’Aménagement du Territoire et à l’Action Régionale* (delegation for regional planning and action) is the country’s half publicly funded regional foresight and policy agency and focuses on regional aspects of national economic plans. In the parliament, there is the *Délégation sénatoriale à la prospective* (senate delegation for foresight), which is in charge of reflecting on social and economic transformation. Notable private futures organizations include Bertrand de Jouvenel’s Futuribles and Michel Godet’s *Cercle d’Action pour la Prospective.*

Germany: There is no centralized system for strategic foresight in Germany. However, there is the German Council of Economic Experts *(die fünf Wirtschaftsweisen),* a small body of five economic advisors that has provided an annual report for the government since 1963. Emerging regulatory issues are addressed in the relevant departments or through a network of independent organizations located outside of governmental structures, such as the Institute for Futures Studies and Technology Assessment, the Fraunhofer Institute for Systems and Innovation Research, or consulting companies like *Z_punkt*. Germany is an outlier amongst OECD countries in that it scores high on the EU Global Innovation Indicator but low on the value of its national foresight studies (Meissner, 2012, p. 911). In a military context, the *Bundeswehr* Office for Defence Planning has a dedicated future analysis section.

Switzerland: Strategic foresight is organized in a decentral fashion in Switzerland. Most forward-looking studies are conducted within federal departments (e.g., Federal Office for Spatial Development, 2016). The strategic management support section of the Federal Chancellery (2011, 2015) used scenarios and trend analysis to provide the Federal Council with interdepartmental outlook reports for the next 10 to 15 years in previous legislative periods. However, this format has been retired in favor of invited essays on priorities in agenda setting (Federal Chancellery, 2018). The technology foresight program of *armasuisse Science and Technology* uses automated horizon scanning, workshops, and board games. It regularly publishes a “Deftech Scan” and offers a research assistance tool that searches for keywords in previous works by more than 60 future-oriented organizations. The FOCP (2019, 2020a) is responsible for the national risk analysis of disasters and emergencies and produces a risk report as well as a risk register. Governmental foresight activities may also be assisted by extraparliamentary expert commissions, such as the ComNBC or the Commission for Pandemic Preparedness and Response.

Outside of government, the Center for Security Studies (CSS) at ETH Zürich produces reports on strategic trends (Haas et al., 2020), civil protection trends (Hauri et al., 2020), and organizes cyber war games. It has also recently published an edited volume on The Politics and Science of Prevision (Wenger et al., 2020). The Geneva Centre for Security Policy offers a course on strategic foresight, and the grassroots think-tank *foraus* has developed an online platform for participatory futures. In terms of general futures and trend research, there is also the Gottlieb Duttweiler Institute; *avenir suisse;* and swissfuture, the association of Swiss futurists. Finally, the Switzerland-based World Economic Forum produces a variety of future-oriented products with a global outlook, including the Global Risks Report.

4.3.2 North America

United States: In general, governmental foresight in the US is organized in a decentralized fashion in government agencies. The Government Accountability Office operates the Center for Strategic Foresight, and the Federal Foresight Community of Interest, which is a forum for federal employees from various agencies, think tanks, and industries to network, learn, analyze, develop, and communicate foresight methods and best practices. The National Intelligence Council (NIC) is responsible for the mid-term and long-term strategic thinking of the US Intelligence Community, which includes the NSA, CIA, FBI, DIA, and so on. Since 1997, it has published an unclassified strategic assessment report called Global Trends every four years. The report describes key trends and uncertainties for the next 20 years (NIC, 2021). The Intelligence Advanced Research Projects Activity (IARPA) addresses cross-agency challenges in the US intelligence community and, among other things, holds geopolitical forecasting challenges. In-Q-Tel is a venture capital firm set up in support of the intelligence community's needs. On the defense side, the Defense Advanced Research Projects Agency (DARPA) is the country's R&D agency for creating and avoiding strategic surprises. Its sister agency at the Department of Energy is the Advanced Research Projects Agency – Energy (ARPA-E). Additionally, there are some decentralized forward-looking structures in the military, such as the Strategic Foresight and Futures Branch at the US Air Force Warfighting Integration Capability.

The US also has various private national security think tanks with significant foresight programmes, including RAND, the Institute for the Future, the Hudson Institute, or the Atlantic Council. Another private research organization, the Santa Fe Institute is known for its particular focus on complex systems.

4.3.3 Asia

Singapore: The Centre for Strategic Futures is a governmental think tank that focuses on blind-spots, pursues open-ended long-term futures research, and experiments with new futures methodologies. The center is also part of the Strategy Group in the Prime Minister's Office, which drives whole-of-government strategic planning by identifying key priorities and emerging issues over the medium to long term.

4.3.4 International Organizations

OECD: The Office of the Secretary-General of the OECD has its Strategic Foresight Unit. In 1990, the OECD launched an International Futures Programme, which engaged in foresight activities. These focused on emerging sectors that could or should become of interest to OECD member economies, including outer space, risk management, and futures thinking in education. In 2013, the OECD (2020) established the Government Foresight Community, which holds an annual meeting.

NATO: The Allied Command Transformation of NATO operates the Long-Term Military Transformation program to anticipate and prepare for the future. NATO (2017; 2018) also produces a report called Strategic Foresight Analysis, which discusses the security environment over the long-term horizon, and another known as the Framework for Future Alliance Operations, which focuses more narrowly on challenges and opportunities for NATO armed forces.

United Nations: The Global Centre for Public Service Excellence, which is jointly funded by the UNDP and the Singaporean government, uses foresight tools such as board games to help civil servants address the Sustainable Development Goals. The UNDP (2018) has also created the Foresight Manual for the 2030 Agenda. The Paris-based United Nations Educational, Scientific and Cultural Organization (UNESCO) has a dedicated futures literacy team that promotes the inclusion of futures thinking and methods as skills that should be included in school curricula. It also administers a broader Global Futures Literacy Network and organizes a Futures Literacy Summit. For the UN General Assembly 2021 the Secretary-General António Guterres prepared a report that contains an emphasis on long-term intergenerational thinking and the representation of future generations. This includes a proposed a "Futures Lab", which will support states and other authorities to build capacities and exchange good practices to enhance longtermism, and a Special Envoy for Future Generations (UN, 2021a, p. 45).

5. Horizon Scanning

Horizon scanning is the systematic and repeated examination of "weak signals" (Ansoff, 1975) hinting at future developments, threats, or opportunities. It is not about predicting the future or exploring what-if scenarios in-depth. Rather, horizon scanning focuses on finding and analyzing existing early indicators of potential changes external to the organization.

The Harvard Business School professor Francis Aguilar (1967) coined the term "environmental scanning" for the overall process of looking for trends and forces of change external to a firm. **In this report, the term horizon scanning is understood as a subset of environmental scanning that focuses on weak signals of emerging issues and potential future trends rather than established trends** (Miles and Saritas, 2012). This section discusses early indicators for potential changes and section 6, which is called trend analysis, looks at established forces of change. However, be aware that these two scanning terms may be used interchangeably in other reports, including a previous report on horizon scanning by the CSS (Habegger, 2009).

The primary purpose of horizon scanning can be framed as a **reduction of unknown knowns** in the known-unknown-matrix (section 3.2), with the goal of reducing strategic surprise. Horizon scanning is generally organized along the bridge model of innovation (section 2.2) and spans **across disciplinary and departmental boundaries.** This is because disruptive innovation often works through "redomaining" and advances in components of one artifact can easily proliferate to artifacts in other domains. The output of horizon scanning is often a regularly updated **list of emerging issues,** which can inform scenario building and other foresight activities.

Exploratory horizon scanning is the open-ended, hypothesis-generating search for emerging issues. Conversely, **issue-centered horizon scanning** starts with a range of established or potential emerging issues. This second type specifically looks for signals which weaken or reinforce the case to monitor a newly identified emerging issue or the ongoing development of established emerging issues. These approaches are complementary. For instance, a project can start with exploratory scanning, which generates a list of signals. From this list, potential emerging issues can be selected and subsequently scanned in a more targeted fashion (Amanatidou et al., 2012).

Scanning for weak signals often relies on the idea that there are different stages of innovation. According to Martino (2003), the first stage is basic research, the second is applied research, the third is development, the fourth is application, and the fifth is social impact. Methods such as bibliometrics and patent analysis aim to capture developments before the application and impact stages.

5.1 Bibliometrics

Bibliometric analysis describes quantitative measurements to evaluate and monitor the body of published academic research. Bibliometrics help to understand the past and identify what countries, institutions, journals, authors, and keywords are most influential in some topical area. As a structural analysis tool, it can highlight shifts in a field over time and help to detect otherwise hidden current trends and patterns that merit further attention. The interpretation of bibliometrics is often augmented with a visual representation of the network topology.

Popular software-tools for bibliometric analyses include:

- **VOSviewer:** The Visualization of Similarities Viewer (VOSviewer) is a commonly used piece of software for creating, visualizing, and exploring scientific maps from bibliometric data (Van Eck and Waltman, 2010).
- **BibExcel:** This piece of freely available software is developed and used for the purpose of extracting bibliographic information such as frequency counts for cited references, the field of cited references, and co-citation analysis. (Persson et al., 2009).

The most commonly used databases for academic citation data are:

- **Web of Science:** This research tool is operated by Clarivate and offers subscription-based access to databases that provide comprehensive citation data for various academic disciplines.
- **Scopus:** This service is operated by Elsevier and offers subscription-based access to databases that provide comprehensive citation data for various academic disciplines.

5.2 Patent Analysis

Patent analysis uses patent data, which is freely available in most countries, to analyze a specific industry or technology sector. It is used to assess the competitive position of firms, to avoid copyright infringement, and to identify technology trends. Filing patents can be a costly and time-consuming process. As such, patent filings indicate that there is optimism in the future economic benefits of an artifact or field of research. The patent growth curve of fields of research and applications often roughly follows an s-shape (Andersen, 1999). A key limitation to both bibliometrics and patent analysis is that researchers in the defense industry will often not disclose their findings publicly.

The **World Intellectual Property Organization (2016) offers a fairly comprehensive manual** for open-source patent data analytics. There are also many commercial patent analysis services, including those provided by the Swiss Federal Institute of Intellectual Property (2021) and many private consulting firms.

5.3 Investments

Another approach is to track investments into basic and applied R&D. This data is not as readily available as bibliometrics and patents, but it is nevertheless an early indicator for areas of technological progress. Venture capital investments are also an early indicator. These focus on attempts at commercialization, which are closer to the third stage (development) in Martino's model. There are commercial providers of start-up analytics. As tracking investments in the entire market is time consuming, it can also make sense to only focus on the activities of organizations with large R&D budgets and/or impressive track records in innovation. For public R&D grants, this could include DARPA and ARPA-E. An example of an early-stage capital provider would be In-Q-Tel, the venture capital firm of the US intelligence services. In Europe, research institutions with significant budgets that might be of interest may include the two federal institutes of technology in Switzerland or the Max Planck and Fraunhofer institutes in Germany.

5.4 Looking at the Fringes

While horizon scanning often focuses on technology, it is also applicable to social and political issues. According to Molitor (2003), there is an **information lifecycle of emerging issues which can be used to monitor for precursors to changes in public policy,** also referred to as "emerging issues analysis." He argues that while many fringe political causes never take off, almost all mainstream political causes have started as fringe causes. Thus, horizon scanning should also **look for agents of change and activist groups outside of the mainstream.** Molitor also argues that future social phenomena are often first found in relatively wealthy urban contexts. This is plausible insofar as the high population density in such environments allows for the formation of communities with shared interests that may have a low prevalence in the overall population. Given that the Internet is a "gigacity" in terms of accessibility, it supports differentiation according to interests that is multiple orders of magnitude greater than physical cities. The social news forum Reddit is particularly useful for monitoring subcultures, as it explicitly consists of interest-based subfora called subreddits and has about half a billion monthly active users. For example, the subreddit r/Biohackers has over 30,000 subscribers.

5.5 Issues Papers

Another common approach are formats that allow subject matter experts to share their knowledge about developments at the edge of their field. This includes interviews with experts and stakeholders, workshops, or conferences. The company Shell has developed a specific two-step process for identifying main themes of today's discussion as well as emerging issues, and then summarizing them in an issues paper.

The Shell approach starts with the seven standardized questions listed below. The intention of these is to gather the opinions of stakeholders on strategic issues in a given policy area and highlight insight into agreement or conflict about ways forward. The interviews are conducted under the Chatham House rule, meaning that information can be shared without being attributed to specific people.

1. If you could speak to someone from the future who could tell you anything about [the focus topic], what would you like to ask?
2. What is your vision for success?
3. What are the dangers of not achieving your vision?
4. What needs to change (systems, relationships, decision making processes, culture for example) if your vision is to be realised?
5. Looking back, what are the successes we can build on? The failures we can learn from?
6. What needs to be done now to ensure that your vision becomes a reality?
7. If you had absolute authority and could do anything, is there anything else you would do?

Following the interview stage, the team selects and clusters key quotes to illustrate strategic issues and choices around the policy area. The aggregated issues paper highlights main themes and emerging issues in the policy area as well as conflicting views and expectations.

5.6 Further Approaches

On the more quantitative side, text mining and topic modelling of scientific and other publications is sometimes used to highlight the groupings between papers and look at their content. In the specific context of biological threats other sources of early signals might include clinical trial databases, virus monitoring, discussion and publications of threat actors, as well as early warning networks. For example, the companies BlueDot and Metabiota, as well as the HealthMap at Boston Children's Hospital, use a mix of machine learning and expert assessment to find and map news articles about transmissible diseases (Kohler and Scharte, 2020). It is also possible to scan for threats in a systematic way at an even earlier stage through the broad environmental sampling of animals to track viruses before they may create a spillover event. Under the Bush and Obama administrations, the US Agency for International Development ran the PREDICT program, which collected more than 100,000 samples from animals and found nearly 1,000 novel viruses, including a new Ebola virus (Kelly et al., 2017). Importantly, to avoid increasing the likelihood of spillover events or accidental releases of sampled viruses from labs, such efforts only make sense when they conform with strict biosafety standards. There are also hopes that **increasingly cheap and rapid tests** for viruses in humans will allow for much higher sampling volumes in the future, which in turn would enable more fine-grained nowcasting and forecasting (section 7) of the spread of transmissible diseases (Mina et al., 2020).

Beyond all these methods, it is also useful to recall that organizations will have already conducted some form of horizon scanning on almost all topics (e.g., table 4). Hence, rather than starting from scratch, **it makes sense for organizations to look for and aggregate the signals and emerging issues that have already been identified by others.** This is possible through desk research, such as through the use of the foresight-oriented search engines from *armasuisse* or ESPAS. Horizon scanning can also form the basis for international institutional collaboration. Examples of this would include joint horizon scanning projects between the UK, Denmark, and the Netherlands as well as between Australia and New Zealand.

<5 years	5–10 years	>10 years
• Access to biotechnology through outsourcing • Crops for changing climates • Function-based design in protein engineering • Philanthropy shapes bioscience research agendas • State and international regulation of DNA database use	• Agricultural gene drives • Neuronal probes expanding new sensory capabilities • Distributed pharmaceutical development and manufacturing • Genetically engineered phage therapy • Human genomics converging with computing technologies • Microbiome engineering in agriculture • Phytoremediation of contaminated soils • Production of edible vaccines in plants • The rise of personalized medicine such as cell therapies	• Bio-based production of materials • Live plant dispensers of chemical signals • Malicious use of advanced neurochemistry • Enhancing carbon sequestration • Porcine bioengineered replacement organs • The governance of cognitive enhancement

Table 4. Emerging issues identified by a bioengineering horizon scan. Adapted from Kemp et al. (2020).

6. Trend Analysis

In the statistical sense, a trend refers to prolonged increases or decreases in one or several metrics. The term is also used more broadly in the sense of structural shifts that are expressed in, though not limited to, specific quantitative indicators.

Trend analysis is a common approach used by companies and governments to understand and adapt to changes in their external environment. In contrast to horizon scanning, this approach does not look for weak signals but at established ongoing dynamics. **Fads** or **crazes** are ephemeral trends that usually last less than a year, such as the popularity of certain fashion styles, diets, pieces of music, dances, or online challenges. Trends can also inspire **countertrends,** subcultures, and phenomena that are in explicit opposition to the overall trend, such as digital detoxes as a reaction to digitalization. **"Megatrends"** is a term used for macrotrends that have a significant scope, impact, and persistence. Another term, **"driving force,"** has been used inconsistently. Some claim that drivers of change are phenomena which can be influenced through one's strategic choices and are hence closer to institutional environments than trends (Saritas and Smith, 2011). Others use driving force as a synonym for the word "trend." Others still claim that while trends have a clear direction, driving forces refer to important metrics that can both increase or decrease, such as the price of oil (Cairns and Wright, 2018). Lastly, some use the term "drivers" to describe the underlying causes of trends. In order to avoid confusion, this report uses the term **"key variables"** for important metrics that may change direction (section 3.2.4; section 8.2) and **causes** for the underlying factors leading to trends.

The analysis of trends, including their impacts and underlying causes, is valuable because our ability to forecast events gets worse as the forecasting horizon increases. Tetlock and Gardner (2016, p. 244) found that the accuracy of expert predictions in geopolitics and economics declined towards being no better than chance for five years and beyond. However, it is possible to identify more long-term structural factors that influence the future independent of individuals and events. One such distinction was introduced by the French historian Fernand Braudel (1949). His work distinguishes between **evental time** *(courte durée)*, which consists of events, politics, and individuals with names; **social time,** which is comprised of long-term social, economic, and cultural shifts; and **geographic time** *(longue durée)*, which denotes slow changes in the environment. For example, Xi Jinping's rise to paramount leader of the People's Republic of China falls under evental time. However, the return of China as a great power is the product of social time and structural factors, such as the country's enormous human capital and its increasing integration into global markets since the late 1970s.

Hodgson and Sharpe (2007, p. 128) offer a visualization that reflects a similar distinction. For them, the stream of events that occurs is often the surface expression of deeper trends and structures. Optimally, we are able to observe not only a longer-term statistical pattern and its impacts but also understand the factors that cause a trend.

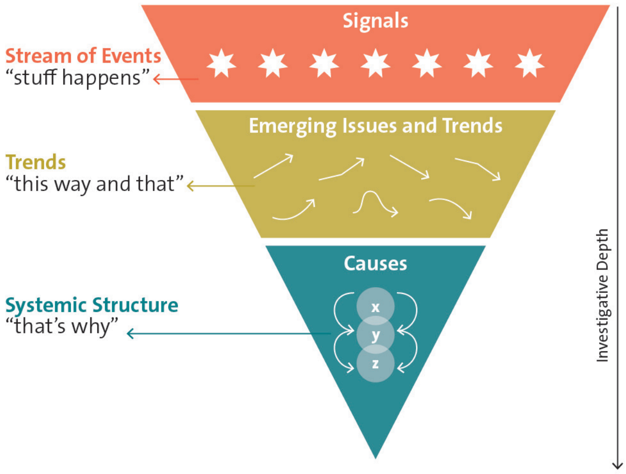

Figure 5. Visualization of signals, trends, and causes.
Adapted from Hodgson and Sharpe (2007, p. 128); United Nations Development Programme (2018, p. 27).

6.1 Trend Identification

There is no universal formula for trend identification and selection. The search for trends can be exploratory but it is often hypothesis driven. **Trends are usually identified based on databases, indexes, literature reviews, and expert workshops, which might be guided by a list of predefined categories in which relevant trends might fall.** Trends can be selected based on their expected impact on the organization, the interest in them amongst the internal and external stakeholders of an analysis, as well as lists of emerging issues that highlight the potential acceleration of existing or newly emerging trends. Of course, a complex trend, such as the rise of China, cannot be fully captured by a single indicator or even in a single index. Nevertheless, a claimed trend should still be backed up by a relevant indicator or index to distinguish it from conjecture. In the case of China's rise, this could be the evolution of its share of the global GDP or a comprehensive power index, such as the Asia Power Index by the Lowy Institute (Lemahieu and Leng, 2020).

STEEP: The STEEP framework is part of the environmental scanning literature. It lists a number of dimensions in which an organization exists and in which it makes sense to look for relevant trends. STEEP stands for **social, technological, economic, environmental, and political.** Aguilar (1967) first introduced the acronym as ETPS, which did not include an environmental dimension. Various authors reshuffled the acronym to PEST or STEP and others started to introduce a large variety of additions. For instance, STEEPV includes values as a dimension, whereas PESTEL adds a legal dimension to the framework. The unpronounceable acronym STEBNPDILE includes five additional dimensions: business methods, natural resources, demographics, international, and legal. There are many more. However, as with risk metaphors, the number of environmental scanning acronyms can be more confusing than helpful. STEEP is mostly used to look for trends; however, the framework can also be used to guide searchers for weak signals or for key variables for scenarios.

Data sources: Good portals for reliable data on a broad variety of STEEP topics include **Our World in Data (2021), the OECD (2021), the World Bank (2021b), the United Nations (2021b), with its various agencies, and national statistical offices,** such as the Swiss Federal Statistical Office. Aside from broad open data portals, there are often databases, datasets, and reports dedicated to specific topics. For example, the EM-DAT database (Centre for Research on the Epidemiology of Disasters, 2021) and the Sendai Monitor (United Nations Office for Disaster Risk Reduction, 2021) focus on disaster statistics, and the World Values Survey (Haerpfer et al., 2020) is the most comprehensive dataset on how global sociopolitical attitudes are evolving. For complex trends, indexes can often be good resources. For example, the KOF Globalisation Index (Gygli et al., 2019) is a good indicator to highlight the dynamics of globalization.

Interpreting data: There are many caveats when it comes to deducing underlying trends from data. Here are three: First, shifts in the public perception of a problem do not have to correspond to actual shifts in its severity. Second, it is worth paying attention to changing definitions and standards in data collection and interpretation. For instance, when a problem decreases, this tends to lead to a broader definition of the issue. Third, it can be deceiving to focus only on data on outcomes. This would be the case in contexts where we expect a few low probability events to cause a large share of an overall impact over a long period. This is particularly relevant for CBRN threats. Alternative approaches to find underlying trends could include looking at "close-call" incidents or examining the development of factors such as vulnerability, exposure, stockpiled weapons, and capacities to respond.

6.2 Trend Causes

There is no universally applicable cognitive schema to derive the causes of trends. The most promising approaches are literature reviews and interviews with experts. The deceptively named causal layered analysis (section 6.2.1) is more useful for a discourse analysis that can help to construct scenarios. The following are some informal questions that may help to guide research on trend causes.

- **Is this a subtrend of a larger trend?** This question is important because large trends have many sectoral subtrends. For example, digitalization has correlating subtrends, such as the digitalization of healthcare, the increasing number and damage of cyberattacks, or the rise of e-commerce, which in turn may have their own subtrends. While sector-specific factors certainly do play a role in each trend, the main underlying causes of digitalization – such as the decreasing costs of computing power and data storage, the increased availability of data, and the increased penetration and bandwidth of Internet connections – have downstream effects on all subtrends.
- **What are the economic incentives for the trend?** This question is relevant because there is a long-run selection pressure in favor of socio-technical configurations that convey economic and/or military advantage over other configurations (section 2.2). For example, urbanization has benefits in the form of income that appears to scale superlinearly with urban density and the corresponding increased social connectivity (Bettencourt et al., 2007). Meaning, if a city's size doubles, per capita wages rise by an average of about 15 per cent.
- **What are the current and historical bottlenecks for the trend?** This question is useful because many trends are not caused by incentives that are fundamentally new. Instead, they are caused by the removal of bottlenecks that change the limits of what is possible and economically viable. For example, historical population density was bottlenecked by factors such as food production capacity per area, the transport and storage capacities for food and energy, housing density, and epidemics (Smil, 2017).

6.2.1 Causal Layered Analysis

Causal layered analysis is a "critical" futures research method developed by Sohail Inayatullah (1998). It is meant to get to the bottom of things by asking workshop participants to describe phenomena on four hierarchical layers. The first layer consists of events, issues, and quantitative trends. The second layer is concerned with economic, cultural, political, and historical factors that explain quantitative data. The third layer is concerned with the worldview that legitimates them. The fourth and lowest layer consists of collective metaphors and myths.

The term causal layered analysis is a misnomer for a layered qualitative discourse analysis box because the lower layers are not the key causes of development on the upper layers. As such, the causal layered analysis cannot be used as a tool for identifying systemic structures that drive long-term technological trends. As discussed in section 2.2 on technological determinism, in the long run there are selection pressures beyond the social construction of technology. However, it can be used as a tool for identifying linguistic structures in a qualitative discourse analysis, and it can be useful when generating scenarios by expanding their range and richness.

6.3 Trend Impact Analysis

The goal of trend impact analysis is to explore what new opportunities and threats emerge with an extrapolated trend and what secondary and tertiary effects it may have on specific fields. This can be done with various methodologies including **literature reviews, workshops, focus groups, expert surveys, panels, and scenario approaches.** The latter are discussed in the separate section 8. The following are short descriptions of selected techniques that are used in the context of analyzing impacts of a specific scenario or trend.

Futures Wheel: The Futures Wheel is a mind mapping technique developed by Jerome Glenn (2009) that aims to visualize first- and second-order implications of a change. A selected trend or scenario is added to a note or post-it and placed in the center of the wheel. Subsequently, participants derive first-order consequences and place them in a circle around the center. In a following round, the second-order consequences are brainstormed and put in a circle that surrounds the previous one.

Cross-impact analysis: This set of approaches was originally developed in the context of Delphi surveys (section 7.2.2) and is determined by expert judgment. "Cross-impact" refers to the coefficient according to which one event or trend raises or lowers the occurrence probability of another event or trend.

Multi-criteria analysis: Multi-criteria analysis is an umbrella term for methods that assess a continuum or a discrete set of policy options by explicitly taking into account more than one dimension of interest and the interplay between multiple, often contrasting, objectives (e.g., economic growth, air quality, and public transport). The objectives can be weighed differently and are represented by at least one criterion (e.g., GDP growth rate, change of carbon dioxide emissions, and public transport accessibility in an area). The performance of each option regarding each criterion is then assessed on a regularized scale (e.g., 1 to 100).

6.4 Megatrends

The term "megatrends" was coined by the American futurist John Naisbitt, who predicted ten longer term shifts in his 1982 bestseller of the same name. Specifically, he predicted transitions from 1) industrial society to information society; 2) forced technology to high tech; 3) national economy to global economy; 4) short term to long term; 5) centralization to decentralization; 6) institutional help to self-help; 7) representative democracy to participatory democracy; 8) hierarchies to networking; 9) North to South; and 10) binary to multiple options. **The general idea of megatrends is that they are persistent for a longer period than average trends and that they affect many people.** However, what exactly constitutes a megatrend is inadequately defined and most books on the topic offer little methodology for trend selection. After analyzing megatrend publications, Slaughter (1993) concluded bluntly: "So what is the motivation for such works? In a word, marketing. The apparently authoritative identification of such trends is a symbolically (but not substantively) powerful resource which can be sold to anyone seeking a competitive 'edge' during difficult times. Their popularity within business environments and elsewhere is undoubtedly due to the way that they interpret and simplify the world. In other words, 'megatrends' provide a largely false sense of security, a way of gaining a seriously distorted impression of 'the big picture' without critical thinking or further effort" (pp. 831–832).

Slaughter's criticism is good to keep in mind for all types of publications on trends. At the same time, this is not to say that trend studies are of no use. Trend studies cannot claim to be comprehensive, and some of the trends that they do identify may not persist. Nonetheless, analyzing trends can still provide relevant insights. This is also true for Naisbitt's megatrends, of which at least some were fairly accurate. Most notably, his predicted shift from a national economy to a global economy.

Megatrends are used in management and public sector consulting to highlight developments and signal internal expertise to the outside. In finance they are also used to offer forward-looking thematic investment portfolios to clients. Examples of megatrend publications from consulting firms include KPMG (2014), PwC (2016), Deloitte (2017), and EY (2020). In finance, BlackRock (2021) and MSCI (2021) offer index funds that track companies operating in megatrend-themed areas such as clean energy, digital security, the digital economy, digital health, urbanization, ageing societies, and future mobility. SoftBank has created the world's largest venture capital fund, the one hundred billion USD Vision Fund, based on the notion that it wants to invest in the technological singularity, the point at which artificial general intelligence surpasses human capabilities in (almost) all economically valuable tasks. SoftBank expects this to happen by the mid-21st century (Bork and Heuer, 2019).

7 Forecasting

Forecasts are predictions about the future, often but not exclusively focusing on one variable of interest at a certain future date. They are based on a mix of human judgement and statistics. Due to their numeric output, predictions are particularly suitable for aggregation and evaluation.

"It is difficult to make predictions, especially about the future" (Quote Investigator, 2013). However, the fact that we cannot accurately forecast events in open-ended systems over longer time periods should not lead us to dismiss forecasting in general. **One important advantage that forecasts have over scenarios and similar methods is that they are falsifiable, which enables their systematic evaluation,** even if the results might be disappointing. The key, as Tetlock and Gardner (2016, pp. 88–89) argue, is to find forecasting questions in "the Goldilocks zone," meaning at the right level of predictability. Karl Popper (1972, Chapter 6) has introduced a popular metaphorical dichotomy between orderly clock-like systems that can be solved through reduction and "highly irregular, disorderly, and more or less unpredictable" cloud-like systems. If a predicted metric is too "clock-like," we can just calculate the system's future state. For example, we do not need a forecast to guess at what time the sun will rise in Zurich tomorrow. However, if the predicted metric is too "cloud-like," forecasts do not provide any useful information. For instance, trying to predict in 2021 who will be the US president in 2044 is pointless.

This section distinguishes between **judgmental forecasting, which involves more informal qualitative methods,** and **statistical forecasting, which uses more formal quantitative methods.** The suitability of these approaches depends on the system for which a forecast is being made. There are situations for which no data is available, meaning that judgmental forecasting is the only viable approach. This can for example be the case for emerging sociotechnical questions. Vice versa, there are situations where there is a lot of good and stable data and stand-alone statistical forecasting works significantly better than judgmental forecasting. An example here could be the recommender systems on Amazon or Netflix. At the same time, the boundaries between judgmental and statistical forecasting are not always clear-cut. Mixed approaches include the use of statistical forecasts as anchors that are adjusted using judgment; judgmental forecasts on which statistical analyses are performed; or the independent generation of statistical and judgmental forecasts and their subsequent combination.

In general, the scope of statistical forecasts is narrower and more short-term than that of judgmental forecasts. However, along with the increased availability of data and computing power, the application area of statistical forecasting has been expanding. Hence, testing where the boundaries of its application currently lie can be worthwhile. Further, the complexity and reducibility of forecasting questions have not always corresponded with our intuitions. It is not without some irony that weather forecasts, the objects of which are literally clouds, have consistently improved (see box 2), whereas the track record of other predictions (e.g., sociopolitical) is much more mixed. The crowd appears to be harder to predict than the cloud.

7.1 Judgmental Forecasting

7.1.1 Expert Intuition

The reliability of expert intuition depends on the task environment, practice, and feedback. According to Kahneman and Klein (2009), **expert intuition can be trusted in a "high-validity" environment which has stable relationships between objectively identifiable cues and subsequent events or the outcomes of possible actions.** Examples of high-validity fields include medicine and firefighting, whereas forecasts of the future value of individual stocks or long-term predictions of political events are made in a near zero-validity environment. Kahneman and Klein further point out that "fractionation" of skill is a problem. Observers and experts can find it challenging to determine the boundaries of expertise, meaning professionals may sometimes be called upon to make judgments in areas in which they have no real skill. For example, the professional ability to play a sport or to build an artifact is not the same as the ability to forecast the longer-term development of the industry to which that sport or artifact belongs to.

Forecasts based on subjective judgment also reflect the limitations and **biases of human cognition.** For instance, the results of expert surveys can be inconsistent based on the framing of forecasting questions. For example, an aggregate prediction of top AI experts for when AI systems will achieve human levels of performance on all human tasks was substantially earlier than their forecast for when AI systems will achieve parity with humans on the task of conducting AI research (Grace et al., 2018). However, logically, the latter is a subset of the former.

7.1.2 Analogy

An analogy describes two domains (e.g., events, time periods, or technologies), whose constituent elements have a common relationship pattern. With an analogy, the reasoner's initial functional understanding of relationships in the source domain is projected onto the target domain, thereby allowing for new inferences about the target domain (Holyoak, 2012, p. 234). According to the structure-mapping theory of Gentner (1983), an analogy is different from a literal similarity, in which the source and the targets share many attributes as well as relationships.

Analogies are widely used in the context of assessing the long-term impact of new technologies. However, relational reasoning is also misleading, as it biases the reasoner to ignore or overlook important ways in which the new is not like the familiar. According to Khong (1992, pp. 20–21), analogies are cognitive devices that help policymakers with up to six analytical tasks. First, they can help define the nature of the situation confronting the decision-maker. Second, they can help to assess the stakes. Third, they can provide policy prescriptions. Fourth, they can predict the chances of success of policy options. Fifth, they can evaluate the moral rightness of policy options. Sixth, they can warn about dangers associated with a policy option.

In the context of politics, some scholars have argued that analogies are a strictly instrumental tool, one which is used to rationalize and advocate for preexisting policy preferences. History is a "grab-bag from which each advocate pulls out a 'lesson' to 'prove' his point" (Fairbank, 1966). However, others have pointed to examples in which politicians relied on historical analogies to perform analytical functions and make sense of policy dilemmas (e.g., Snyder and Diesing, 1977, Chapter 4). Khong's analysis of US decision-making in the Vietnam War shows that analogies have also been used extensively in private settings (Khong, 1992, p. 61). Here, analogies helped to inform secondary characteristics of policy choices and inundated decision makers regarding contradictory evidence (Khong, 1992, p. 224).

7.1.3 Reference Class

A forecasting problem can be viewed as part of a reference class of events, whose historical outcomes can be assessed to give a baseline probability. For example, this may be what share of incumbent presidents have won reelection in a country. According to Tetlock and Gardner (2016, p. 279), this "outside view" is usually the right starting point for judgmental forecasts, one which can then be updated based on the particulars of the case, the "inside view." Almost all questions can be framed as part of at least one reference class; however, there may be substantial ambiguity when there are several possible reference classes that indicate different base rates.

7.2 Aggregated Judgement

Maybe the most famous "wisdom of the crowds" example stems from Francis Galton (1907), who observed that the median guess of the weight of an ox at an agricultural fair in Plymouth was within 1 per cent of its true weight. As such, the idea that aggregate predictions are more accurate than individual predictions is quite old. However, in most forecasting contexts, except for voter or consumer intention polling, the aggregation is applied to expert rather than layman judgement. Furthermore, many-to-many communication through the Internet has enabled new forms of organizing collective intelligence, such as virtual labor markets, tournament-based collaboration, and open collaboration (Prpić et al., 2015). In forecasting, these new possibilities have led to reputation-based prediction tournaments and prediction markets with exchange-traded bets on political outcomes.

7.2.1 Expert Survey

Providing experts with a set of questions and aggregating their predictions is a simple and common form of aggregate judgement. In contrast to representative polls, the participants are selected based on expertise in a topic area rather than based on socio-demographical characteristics (e.g., education, income). Experts can, for example, be selected based on a citation count in a field or their participation at a certain workshop or conference. Their answers are usually given an equal weight, but they can also be performance-weighted based on a self-rating, peer rating, or a set of calibration questions. For instance, an equal-weighted average of peer forecasts performed better than any individual forecast on a set of Bloomberg survey forecasts of US economic data (Qu et al., 2019). For issues involving a considerable range of predictions, it often makes sense to use the median value as the mean is sensitive to outliers.

The main weaknesses of expert surveys are that they are subject to selection bias, such as which experts are asked and answer, and to all types of human biases. As such, they can produce different results based on the framing of questions. Furthermore, especially for interdisciplinary topics, it can be hard to identify who has expertise in the first place. The key strengths of expert surveys are that they are easy to conduct and that simple aggregated forecasts often perform better than individual forecasts.

7.2.2 Delphi Method

The Delphi technique was developed at RAND in the late 1950s and early 1960s. Its name is derived from the ancient Greek Oracle of Delphi. The oracle's answers were usually highly ambiguous; however, the journey to and from Delphi provided decision-makers with a strategy retreat to discuss a problem at length. **The Delphi method is meant to produce a consensus expert forecast,** while reducing factors that lead to distortions in face-to-face expert conferences, such as dominating characters or the reluctance of someone to change their opinion in public. Its key characteristic is the use of several iterations of anonymous questionnaires that include median forecasts and reasoning. There have been many variations on the Delphi method, some of which involve quantitative simulation models, in-depth interviews, or group meetings. Similarly, variations of the Delphi method have been applied to scenario construction rather than prediction. Below is a description of a classic Delphi study. See Gordon (2009) for a more detailed account.

Finding the participants: Panelists are not meant to be representative of the larger population but selected for relevant expertise. They are identified through relevant publications on the subject as well as recommendations from institutions and other experts. The list of participants can also be cross-checked with a list of relevant subfields to the question to ensure that all relevant occupational backgrounds are covered. Additional participants can be added by asking professors to nominate bright students or by publicly asking for answers on preliminary questions and selecting the most convincing responders. Most Delphi studies include 15 to 35 experts, although there is no upper limit.

Four questionnaire rounds: In the first round, participants might be asked to provide their judgment on which dates they expect specific milestones with regards to the development of a technology. The subsequent analysis would identify the range of opinions, and the second round would only be addressed to those holding extreme positions. These individuals are asked to reassess their positions and provide their reasoning for them. In the third round, the group judgement and the synthesized reasons for extreme positions are presented to everyone. Participants are asked to reassess their opinion and to potentially refute presented arguments. In the final round, the updated consensus and the new arguments are presented, along with a last reassessment request.

A weakness of the Delphi method is that it can take months, as one needs to wait for expert answers in each round. Some scholars have also expressed skepticism towards the initial ideas that the Delphi method can remove the negative effects of face-to-face interaction and that an expert consensus is more likely to be accurate than an individual forecast. The key strength of the Delphi method is its ability to crystallize the reasons for disagreements between experts.

7.2.3 Reputation-based Prediction Platforms

On prediction platforms, participants are neither representative of demographics nor of academic credentials on a topic. Instead, they are self-selected based on an interest in forecasting and/or the topic area. The primary incentive to participate in prediction platforms is not money; however, one often exists in the form of some kind of points or ranking system that reflects the prediction record of users. Further, the platforms may offer prizes for the best forecaster. Notable reputation-based prediction platforms include the following:

- **Good Judgement Project:** The Good Judgment Project was co-created by Philipp Tetlock and began in July 2011 as part of an IARPA competition. Good Judgement Inc. is the commercial spin-off from the project. It operates Good Judgement Open, a public forecasting tournament with questions ranging from geopolitics to finance to sports. The company also provides access to its most successful forecasters as a service (section 7.2.4).
- **Metaculus:** The company Metaculus was founded in 2015 by US data scientists and offers a platform for probability predictions to binary questions, numerical-range predictions, and date-range predictions. In 2020, the firm's site introduced its bi-weekly Bentham prize for the most valuable user contributions. It also created several one-off prizes, including the Li Wenliang prize for forecasts related to the COVID-19 pandemic and the Salk prize for vaccine research and distribution.
- **Forecast:** The Forecast app lets people use virtual points to trade on future events and outcomes. Its beta version was launched in June 2020. In October 2020, Facebook's New Product Experimentation team opened the app up to all users in the US and Canada. Due to Facebook's massive user base, the project might help to popularize reputation-based predictions if it is continued and integrated into the social media platform.
- **Foretell:** Foretell is a pilot project launched by Georgetown's Center for Security and Emerging Technology that focuses on questions relevant to technology-security policy. It has a particular focus on US-China politics, AI, and information technology.
- **Cosmic Bazaar:** Cosmic Bazaar was launched in April 2020 by the UK government to supplement traditional forms of analysis and prediction. It includes 1,300 forecasters from 41 different government departments and several allied countries. One aim of the project is to identify persistently successful forecasters that could be quickly mobilized to answer questions in a crisis. A second aim is to increase "cognitive diversity," as under pseudonymity a junior data scientist may contest the predictions and reasoning of a senior ambassador without regard to bureaucratic hierarchies (Joshi, 2021).

7.2.4 Superforecasters

The term "superforecaster" was coined by Tetlock in the context of the Good Judgement Project. To qualify as a superforecaster, an individual has to score in the top 2 per cent of the reputation-based Good Judgement platform. As Tetlock and Gardner (2016, p. 104) show, forecasting has clear elements of skill, insofar as the correlation between the performance in one year to the next across all forecasters is 0.65 and about 70 per cent of superforecasters were able to maintain their status year-to-year. This performance-based selection of online volunteers has performed very well compared to other approaches in prize competitions organized by IARPA. Good Judgement Inc. also offers superforecasting-as-a-service for financial services, governments, and the energy sector. This includes a paid subscription service and the possibility to have superforecasters answer custom questions in private.

According to Tetlock and Gardner (2016, pp. 191–192), the composite characteristics of the modal superforecaster includes a cautious and humble philosophical outlook, active open-mindedness and intellectual curiosity, feeling comfortable with numbers and probabilities, not being wedded to any idea or agenda, valueing diverse views, as well as possessing grit and a growth mindset. The authors have also composed ten commandments for aspiring superforecasters: 1) Triage questions according to predictability. 2) Break seemingly intractable problems into tractable sub-problems. This is also called "fermiization" after physicist Enrico Fermi. 3) Strike the right balance between inside and outside views. First, anchor with the outside view. Second, adjust this using the inside view. 4) Strike the right balance between underreacting and overreacting to evidence. 5) Look for the clashing causal forces at work in each problem. 6) Strive to distinguish as many degrees of doubt as the problem permits, but no more. In other words, be as precise as possible in probabilistic forecasts. 7) Strike the right balance between timidity and overconfidence, between prudence and decisiveness. 8) Look for the errors behind your mistakes but beware of hindsight biases. 9) Bring out the best in others and let them bring out the best in you. Foster a culture of perspective taking, precise questioning, and constructive disagreement. 10) Master the error-balancing bicycle. Practice, practice, practice! 11) "Don't treat commandments as commandments" (Tetlock and Gardner, 2016, pp. 277–285).

7.2.5 Prediction Markets

Individuals who sell their talks about the future to companies and individuals should also be able to directly monetize their privileged insights about the future, as long they are based on publicly available information. Strictly speaking, both the regular stock market and sports betting are prediction markets. However, what is usually meant by a prediction market is an exchange traded market of bets on sociopolitical outcomes. These are mostly binary options describing, for example, whether a specific candidate will win or lose an election. Exchange trading means that the market price of a contract adjusts dynamically based on the participants' predictions rather than having fixed odds offered by a bookmaker. The market value of predictions is then seen as an indicator of the probability of these events. The money incentives participants to be as accurate as possible as one can earn money based on better information. Prediction markets have produced forecasts with lower prediction errors than conventional forecasting techniques. Notable prediction markets include:

- **Policy Analysis Market:** In 2003, DARPA was setting up the Policy Analysis Market (PAM) to test the ability of markets to forecast geopolitical trends. However, the experiment quickly turned sour after two US senators held a press conference in which they decried PAM as a "terror market" in which people could bet on terrorist attacks. As this occurred shortly after the September 2001 attacks, which were accompanied by rumors about people buying put options on airlines before the event happened (Poteshman, 2006), this quickly spiraled into a moral panic. Politicians and newspapers alike condemned the project in harsh terms, and it was cancelled within days. In reality, PAM was focused on geopolitical questions rather than terror attacks (Hanson, 2007). Furthermore, whereas it is theoretically possible to earn money on the regular stock market through shorting attack targets (see e.g., Troianovski, 2017), PAM had restricted participation and would have involved small financial sums.
- **PredictIt:** PredictIt is a research project of Victoria University of Wellington in New Zealand, which had already run the prediction market iPredict from 2008 to 2016. It primarily offers bets on who wins political elections and had about 100,000 active traders as of March 2020. PredictIt offers a data sharing program for members of the academic community.
- **Betfair:** Betfair is the world's largest online betting exchange. It mainly offers sports bets; however, it also offers an increasing range of political bets.
- **Blockchain markets:** There are several attempts to build decentralized prediction markets. Augur is a platform built on the Ethereum blockchain in which users can create their own prediction market for virtually any question including politics.

A weakness of prediction markets is that they often have a fairly limited liquidity, which can undermine their efficiency. Further, regulatory requirements can make it complicated to set one up. If prediction markets are increasingly taken into account in public discussions as election forecasts, there is also an incentive to manipulate them. Lastly, public relations risks with regards to "terrorism market" accusations have to be taken into account.

In contrast, prediction markets have a good track record against individual predictions and even against aggregate predictions. In theory, the monetary incentives should help to minimize motivated reasoning and make individuals more willing to admit to uncertainty.

Play money markets: Some projects aim to use the dynamics of prediction markets without requiring a purchase or payment from participants. Instead, participants get a certain amount of play money per prediction round to bet on outcomes, which can later be converted into real prize money. There is no substantial difference between play money markets and reputation-based prediction platforms that offer occasional prize money. However, because some play money markets are very clearly framed as prediction markets, they are still listed here. **Hypermind** is a market that has existed since 2000 and offers sponsored competitions and other services to private sector firms. **Replication Markets** offers prize money for bets on whether research claims, mainly in social and behavioral sciences, will replicate or not. Its data is analyzed as part of a DARPA program. The **Intelligence Community Prediction Market** is a closed market for US government employees and contractors with security clearance. It has been run by IARPA since 2010 and allows participants to use non-monetary points to buy and sell shares of answers to intelligence questions.

Prize tournaments: Aside from platforms with reputation-based communities or exchange-traded markets, there are also one-off prize tournaments. The most notable of these have been organized by IARPA, including the **Aggregative Contingent Estimation** program, two **Geopolitical Forecasting Challenges,** and the **Hybrid Forecasting Competition.** For an overview of the various intelligence community forecasting programs by DARPA, IARPA, and the CIA, see Horowitz et al. (2021, p. 13).

7.3 Automated Judgement

Judgmental bootstrapping: Judgmental bootstrapping or "policy capturing" describes the process of inferring how successful experts make their forecasts and transforming this into a model with explicit rules (Armstrong, 2001). This approach starts with interviews of a diverse set of experts regarding what information they rely on for their forecasts and why. The resulting list of variables is then reduced to a manageable number based on the plausibility of causal effects. Subsequently, the judgements of successful experts are used to quantify what mix of these variables best correspond to the value of their predictions. As data becomes available on actual outcomes, the bootstrapping model should be recalibrated to improve the estimates and help experts reexamine their use of information. Experts may still rely on more subtle information that is not captured in the bootstrapping model. However, the model profits from consistency, whereas experts may rely on slightly different decision-models for each forecast. Bootstrapping can be a solution for complex problems for which the alternative is to rely on no or much less experienced experts.

Expert systems: Expert systems are related to judgmental bootstrapping. However, whereas the latter focuses on inferring the procedures of experts, expert systems aim to represent processes directly by being based on knowledge about methods and the domain (Collopy, Adya and Armstrong, 2001). To acquire the knowledge to create a rule-based expert system, researchers rely on a broad array of sources. These can include expert interviews, case studies, judgmental bootstrapping, recordings of experts at work, and literature reviews. The recording of experts is particularly useful when they might lack an awareness of their own processes. Expert systems might be useful for semi-structured problems with a lack of historical data.

7.4 Statistical Forecasting

Statistical forecasting is particularly popular in business management, often focusing on sales data. Short-term demand forecasts are often used for scheduling of personnel, production, and transportation. Medium-term forecasts inform the purchase of raw materials, the hiring of employees, and investments into equipment. Long-term forecasts are used to inform strategic decision-making.

The relevant data can usually be visualized on a time plot, with the predicted quantity on the x-axis and time on the y-axis. The most common time series patterns include the following:

- **Trends:** In statistical forecasting, trends are understood more narrowly than in trend analysis (section 6) and refer to sustained increases or decreases in a metric. Trends do not have to be linear.
- **Seasonal patterns:** A seasonal pattern occurs when a time series is affected by seasonal factors such as the time of the year or the day of the week. Seasonality is always of a fixed and known period.
- **Cyclic patterns:** A cycle occurs when the data exhibits rises and falls that are not of a fixed frequency. These fluctuations are usually due to economic conditions and are often related to the "business cycle." The duration of these fluctuations is usually at least two years.

Time series can be decomposed and **seasonally adjusted** in order not to conflate general trends and seasonal patterns. Subsequently, the most common approach is to use some type of **regression** to compute a formula that has a small prediction error on the training set. The accuracy of point forecasts is then evaluated on a **test set** or initially on a validation set and subsequently on a test set. These are subsets of the dataset that are kept back during the training in order to avoid overfitting the prediction algorithm to the data points. A set of methods that has been shown to improve business forecasting is called **exponential smoothing.** Here, the relative weights of past observations for the predictions decay exponentially as the observations get older. Another highly popular approach is to use **autoregressive integrated moving average (ARIMA)** models (Box et al., 2015), which adapt the time series to a suite of standard structures.

Large tech firms have open-sourced **forecasting toolboxes** for R and Python. Notable examples include:

- **Prophet:** Prophet is an open-source package for R and Python, which was released by Facebook's (2021) data science team. It fits non-linear trends with yearly, weekly, and daily seasonality as well as holiday effects.
- **CausalImpact:** This open-source R package released by Google (2021) helps to estimate the effect of a designed intervention by predicting how a time series would have evolved counterfactually without the intervention.
- **AnomalyDetection:** This open-source package for R, released by Twitter (2021), helps to detect anomalies in a wide variety of contexts and in the presence of seasonality and an underlying trend.

A comprehensive overview of statistical forecasting is provided in **Rob Hyndman and George Athanasopoulos' open-source textbook *Forecasting: Principles and Practice* (2021).**

7.4.1 Machine Learning

Machine learning is a subfield of AI and refers to data analysis methods that learn to build decision-making models based on training data. In recent years, machine learning has been successfully applied to a large variety of classification and short-term prediction tasks, including spam emails, financial fraud detection, board games, computer vision, and natural language processing. Agrawal, Goldfarb and Gans (2018) call current AI "prediction machines." Indeed, most applications can be framed as a form of prediction. For example, a large neural network that has been trained on large amounts of texts and can produce human-like text in a conversation, such as GPT-3 (Brown et al., 2020), operates by predicting the next pieces of text based on its training data.

Machine learning is particularly successful in contexts where there is a lot of good and stable data, a clear goal, a clear application perimeter, and relatively static adversaries. Unfortunately, these characteristics are not shared by complex open systems, such as those concerning geopolitical questions. However, AI can still play a part in strategic foresight. For example, its increasing ability to go through vast amounts of videos, sound recordings, and text may be useful in horizon scanning (section 5). Furthermore, the speed at which neural networks have evolved in the last years mean that the scope of their usefulness must be reassessed on a regular basis.

7.4.2 Tournaments

The M-competitions, named after their initiator Spyros Makridakis, are a series of time series forecasting competitions that aim to improve empirical evidence and advance the theory and practice of statistical forecasting (Makridakis et al., 1982; Makridakis et al., 1993; Makridakis and Hibon, 2000; Makridakis et al., 2020). The most recent competition, M4, included 100,000 time series on topics such as finance, industry, and demographics. The intervals between successive observations were mostly monthly, quarterly, or yearly. Interestingly, the first three M-competitions all found that the use of more sophisticated and complex methods did not seem to improve forecasting accuracy. The M4 competition was less clear on how complex methods compared with simple ones. Pure machine learning approaches fared fairly poorly; however, hybrid approaches, which use both traditional statistics and elements of machine learning, performed well.

From a complexity viewpoint, it must be mentioned that there are many factors that contribute to surface indicators, such as tourist arrivals or quarterly earnings. Predictions based on the historical development of a single indicator alone are inherently limited and one should be cautious in generalizing the relative performance of approaches in M-competitions to other prediction tasks.

7.5 Evaluation

There are multiple reasons why the systematic evaluation of predicted outcomes against actual outcomes matters.

First, experts often suffer from a self-serving memory and hindsight bias. For example, Tetlock asked political experts in 1988 for a probability estimate that the Communist Party would lose its monopoly on power in the Soviet Union in the next five years. When he returned to the experts after the collapse of the Soviet Union and asked them what probability they had put down, they misremembered their own estimate by recalling a number that was on average 31 percentage points higher than the number they had stated in 1988 (Tetlock and Gardner, 2016, p. 184). Hence, forecasts should be recorded.

Second, judgmental forecasts are often ambiguous, which can lead to misunderstandings and allow experts to interpret vagueness in their forecasts in their favor. For example, in intelligence estimates the ambiguity of verbal probability expressions, such as “likely”, has led to a drive towards standardizing probability ranges that correspond to verbal expressions despite push back from intelligence analysts.

Third, many forecasters do not systematically evaluate past predictions. Hence, they will likely overestimate their forecasting ability. Furthermore, there may be a difference between internal and external assessments. For example, the futurist Ray Kurzweil (2010) concludes that 78 per cent of his technology predictions for 2009 turned out to be entirely correct and 86 per cent to be at least essentially correct. In contrast, an independent assessment judged 27 per cent to be true and 15 per cent to be weakly true (Armstrong, 2013). For Kurzweil’s 2019 predictions, the same independent assessors found 12 per cent of his predictions to be true and 12 per cent to be weakly true (Armstrong, 2020).

Box 2: Weather Forecasting

Forecasting meteorological phenomena, including natural hazards such as hurricanes, is very complex but grounded in an understanding of how to conceptualize the most important interactions at small scales. Given the availability of more data and more capable supercomputers, the performance of weather and hurricane forecasts has improved consistently over the decades. Indeed, governments may even set performance targets for the forecasts of their meteorological agency (Federal Department of Finance, 2019, p. 137). This stands in contrast to predictions about technological, social, or political phenomena.

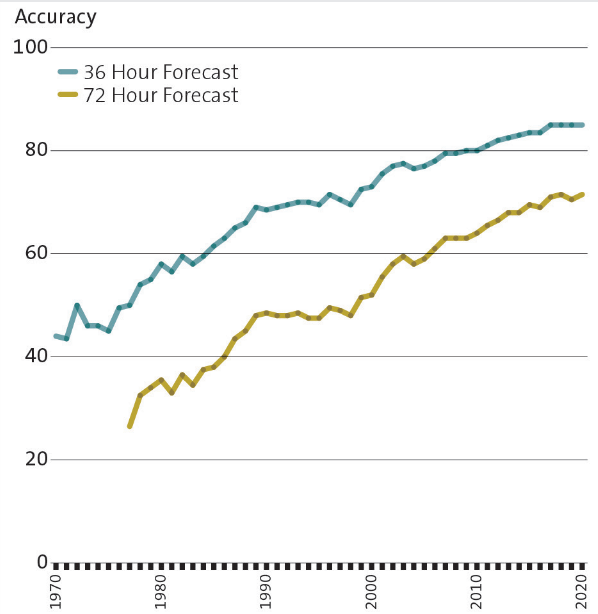

Figure 6. The 500-hPa skill in U.S. National Centers for Environmental Prediction 36- and 72-h global forecasts between 1955 and 2020.
Measured as ratio of gradients (forecast to observed). Reprinted from National Centers for Environmental Prediction (2018).

Box 3: IEA Predictions Have Been too Persistent

The International Energy Agency (IEA) is a Paris-based intergovernmental organization that was formed by OECD states in the aftermath of the 1973 oil crisis. In its annual World Energy Outlook, the agency has included forecasts for the adoption of solar photovoltaic capacity since 2006. As the record shows, the IEA has consistently predicted that solar adoption will persist at the level it is at the time of publishing. However, the actual trajectory of solar adoption has been much faster, mainly due to consistent price decays and fast adoption in China.

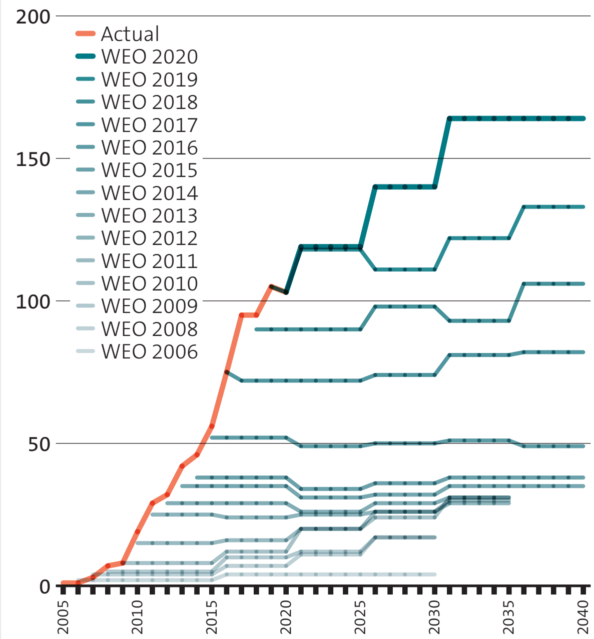

Figure 7. Newly Installed Solar Capacity per year – predicted by IEA and actual.
Adapted from Evans (2019; 2020).

Fourth, forecasting can be a skill of its own. Being an expert in a field and predicting its future are two related but separate skills. Generalists with forecasting expertise may outperform subject matter experts in subject matter forecasts. **Fifth, the forecasting track record of experts is almost entirely disconnected from public discourse.** Media organizations usually reach out to individuals for public comments and forecasts on a topic based on subject-matter expert databases, institutional credentials, and the predictability and strength of opinions. Depending on the topic and a medium’s interests and viewers, this might be a conscious choice for “policy-based evidence” by picking an expert whose opinion predictably aligns with pre-existing policy preferences. As Tetlock (2016, p. 72) notes, media organizations also have a preference for strong opinions.

Metrics: There are various metrics to verify forecasts, many of which are application area specific. For example, in meteorology, the S1-score is used for mean-sea level pressure forecasts. This section discusses two generic metrics. The Brier Score is the standard evaluation metric for probabilistic sociopolitical forecasts. The F-Score is not commonly used in sociopolitical forecasting. However, it is a popular tool in binary classification tasks with imbalanced data sets. It highlights the challenge of verifying forecasts for low probability, high impact risks, including CBRN-attacks.

7.5.1 Brier Score

First proposed by Glenn Brier (1950), the Brier score is applicable to probabilistic predictions for a set of mutually exclusive discrete outcomes. The set of possible outcomes can be binary or involve multiple categories. Each assigned probability must be in the range of zero to one and the probability of all possible outcomes must sum to one. The Brier score then measures the mean squared difference between the predicted probability assigned to possible outcomes and the actual outcomes for a set of predictions. The lower the Brier score is for a set of predictions, the better the predictions.

Calibration: Calibration refers to the degree to which a stated certainty in a predicted outcome or interval corresponds to observations of actual outcomes. Most individuals are overconfident. This means that across several forecasting questions less than 80 per cent of observed outcomes will fall within a predicted confidence interval of 80 per cent. Similarly, less than four out of five binary predictions with a stated confidence level of 80 per cent turn out to be correct. **Calibration training can help individuals to have a more accurate level of internal confidence.** The fastest way to improve calibration is with “confidence quizzes” that rely on statistics about the past that are not commonly known. This is because such quizzes do not require any waiting time for feedback on predictions of future events (Hernandez, 2017; Hernandez, n.d.).

7.5.2 F-Score

The binary classification of data points (e.g., cancer or no cancer) results in four basic combinations of the actual data category and the predicted category: True positives, true negatives, false positives, and false negatives (Table 5). The F-Score is a formula to address class imbalance in binary classification, particularly in issue areas with comparatively few true positives.

	Actual +	Actual -
Predicted +	True Positive	False Positive
Predicted -	False Negative	True Positive

Table 5. Prediction error matrix.

Precision and Recall: Many real classification problems have a class imbalance. For example, the actual positive conditions of cancer are only small fractions of the population in mammograms. Accuracy denotes the fraction of correct predictions from all predictions. Achieving **high accuracy is not very meaningful on imbalanced data sets.** Always predicting the negative condition is already sufficient to achieve this. **Therefore, we need to rely on precision and recall.** Precision reflects what share of predicted positives turn out to be true positives. It is calculated as true positives divided by the sum of true and false positives. Recall refers to the share of actual positives that have been correctly identified. It is calculated as true positives divided by the sum of true positives and false negatives. The F1-score is the harmonic mean of precision and recall, which is 2*Precision*Recall/(Precision+Recall). The higher the F1-Score, the better the predictions. The score favors similar values for precision and recall as this is when the product of two numbers is the largest compared to their sum. The lowest possible value is zero, which happens if either the precision or the recall is zero.

Box 4: Example of Brier Score vs. F1-Score
Let us hypothetically assume there are 20,000 catalogued mutations of a pandemic virus during a two-year pandemic. Of these mutations, four turn out to make the virus substantially more dangerous than previously existing variants in terms of transmissibility, severity, or immunity. For simplicity’s sake, the unrealistic assumption in this example is that all variants have one mutation. Asked to assess new mutations based on their understanding of the viral genome and limited clinical evidence, virologist A always predicts that mutations will not make the virus more dangerous. In contrast, virologist B also disregards 19,980 mutations but identifies 20 mutations that create variants of interest, including the 4 which will turn out to be variants of concern based on subsequent clinical evidence. Virologist A would perform better on accuracy and the Brier score than virologist B. However, a fire alarm that never rings is useless. In contrast, Virologist B has reduced the set of candidates by 99.9 per cent without false negatives. The F1-Score reflects this. Virologist A has a recall of 0 and therefore a F1-Score of 0. Virologist B has a precision of 0.2 and a recall of 1, resulting in an F1-Score of 0.33.

8. Scenarios

Scenario planning, scenario thinking, or scenario analysis are interchangeable terms used for what-if explorations of possible or preferable futures without aiming to quantify the likelihood of these situations. Scenarios can be enhanced through role-play and science fiction.

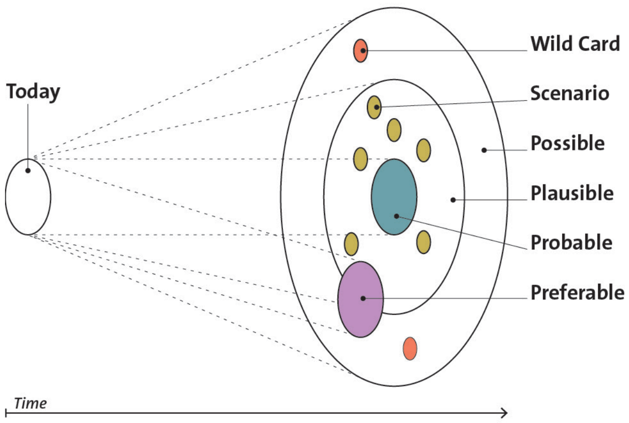

Figure 8. The futures cone – a visualization of preferable, plausible, and possible futures.
The y- and z-axes represent arbitrary dimensions. The potential values along these dimensions at an arbitrary future time are assumed to have a bell curve distribution with probability peaks in the center of the circle on the right. Adapted from Hancock and Bezold (1994, p. 25); Voros (2003, p. 16).

Scenarios with low probability and high impact are also referred to as "wild cards" in foresight (Petersen and Steinmüller, 2009). Scenarios can be built based on previous results from horizon scanning, trend analysis, and forecasting. However, scenario building can also make use of creativity techniques or the identification of key variables within a scenario workshop.

8.1 Creativity Techniques

In order to turn unknown unknowns into known unknowns (section 3.2.1) or from level 5 to level 4 uncertainty (section 3.2.4), we need to think about new combinations of technologies and their social embeddings. One set of approaches to stimulate creativity in workshops is brainstorming methods, such as the Futures Wheel (section 6.3). Another simple yet useful technique to structure the process is the use of exhaustive comparison.

Morphological box: A morphological box is a multidimensional matrix to explore all conceivable combinations along key dimensions of a selected issue. The method begins with the identification and definition of key dimensions. We then attempt to come up with something for every possible combination of these dimensions, even when particular suggestions appear to be nonsensical. Finally, we evaluate the different combinations to find what appear to be the most useful options and see if we can integrate them into a scenario. For example, if we are looking for future use cases of technologies, we might want to have a list of key technologies, business model patterns, and personas as dimensions. Key technologies may be derived from sources such as horizon scanning, trend analysis, or governmental list of key technologies. A useful source for business model patterns could include the collection of 55 such models by Gassmann et al. (2015). Personas are fictional characters which represent the needs, characteristics, life circumstances, and problems of current or future groups of customers and users (Adlin and Pruitt, 2010). Personas can be useful to make scenarios more engaging and light-hearted. They are often used in the context of design fiction and science fiction prototyping (section 8.5).

To make the use of a morphological box more engaging, it can be played as a card game. For example, a hypothetical game could include three types of cards: 1) key future technologies; 2) application contexts consisting of industries or societal issues; and 3) business models. A way to play the game would be for all players to draw several technology and business model cards. One player would draw an application context card and ask all the other players to submit at least one technology and business model card that could have an impact on the given context. Once all the cards are submitted, the players would explain their selection and the person who drew the application context card would decide which is the best answer. The player who gives the best answer gets to keep the application card as a point. To start a new round, the winner of the previous one would draw a new application card.

8.2 Classic Matrix

The classic Shell scenario framework developed by Peter Schwartz. It is an iterative process focused on identifying uncertainties that have strong consequences for key stakeholders in a qualitative **2x2 matrix** with 4 scenarios. However, the framework can also be extended into a 2x2x2 cube, with a total of 8 scenarios. For this process, participants should first develop a long list of **key variables** in a policy area. This list should then be shortened through participant voting, until two critical uncertainties have been chosen. The extremes of each are also described on a matrix. Subsequently, the participants develop appealing and memorable names for the quadrants. They also craft narratives –potentially involving characters, winners and losers, or crisis and response strategies – for paths into each quadrant of the matrix as well as descriptions of how the world could shift from one quadrant to another.

The resulting scenarios can be integrated with strategic decision-making through **wind tunneling** (Van der Heijden, 2004). Wind tunneling means creating a matrix consisting of existing internal strategies or policy options (e.g., marketing) and external environments (the four scenarios). The goal is to evaluate whether current strategies are robust across all scenarios or whether there is a need to look for and prepare alternative strategies.

Box 5: Information Hazards
Information hazards can be defined as "the dissemination of (true) information that may cause harm or enable some agent to cause harm" (Bostrom, 2011). Brainstorming about novel ways in which CBRN-material might be deployed or red teaming (section 8.4) to find illegitimate way of accessing such material can create information hazards. In cybersecurity, such concerns have led to a well-established norm of responsible vulnerability disclosure, where security researchers give affected companies time to patch vulnerabilities before they make them public. In biosecurity, these concerns have been discussed intensely for dual use research, which can be used to better understand pathogens but that might also be weaponized by bad actors or increase the risk of accidental spillovers. Examples include research on mousepox, horsepox, gain-of-function influenza, and botulinum toxin H (Esvelt, 2018; Lewis et al., 2019). Due to information hazards, participatory futures approaches are not particularly suitable for CBRN threat contexts. Further, non-disclosure agreements for participants as well as responsible vulnerability disclosure to stakeholders should be considered in relevant formats, such as ideation and red teaming.

8.3 Backcasting

Backcasting is a normative foresight approach introduced by John Robinson (1990). Rather than aiming to look forward to the future from today, this approach looks back from a preferable future to today.

Backcasting first involves an attempt to define what constitutes a preferable and plausible future scenario (figure 7). This is called "visioning," and its goal is not just to generate a list of preferable attributes or personal visions but **one coherent preferable scenario as a shared vision for the group or organization.** Polak (1973) has argued that a collective positive vision of the future is important for subsequent flourishing. However, according to psychology research only positive expectations appear to correlate positively with effort and success, whereas fantasizing about positive futures without a clear path to them appears to correlate negatively with effort and success (Oettingen and Mayer, 2002). Therefore, it is important not to stop at the vision, but to "backcast" from there and to develop a step-by-step list of necessary and contributing factors that can make this end-state a reality. For this process, Bengston et al. (2020) suggest using a **backcasting wheel,** inspired by the futures wheel (section 6.3). The preferable future is in the center and is then sequentially disaggregated into key dimensions of success: Signposts that indicate you are on the right path to success; obstacles and opportunities to reaching the signposts; and, lastly, concrete management actions that you can take to regarding the obstacles and opportunities.

Scenario network mapping is a method developed by Dennis List (2005) that aims to explore the links between several possible futures through a series of workshops. While backcasting is only one of several approaches used in scenario network mapping, it can fulfill a similar purpose as backcasting. The workshops start with the definition of milestones and key components of a technology. Once this is done, participants vote on the importance of the milestones and definitions using dots. Second, the milestones and components are connected in event trees in two subgroups once through backcasting, and once from the present to the future. Third, both event trees are merged and the participants can vote on the importance of different paths linking the present to the future as well as on normative aspects. For detailed instructions on how to conduct a scenario network mapping workshop, see the original process described by List (2005, pp. 402–404) and an adapted version by Gruetzemacher and Paradice (2019, pp. 75–77).

While backcasting is associated with normatively desirable futures, the same methodology can also be used for undesirable futures. In business, this is known as a **pre-mortem analysis.** Similarly, civil protection uses **reference scenarios of hazards and threats** to develop possible responses. For example, together with the Spiez Laboratory, the ComNBC (2011, pp. 18–20) has developed a set of 14 reference scenarios for CBRN protection. Furthermore, the FOCP (see e.g., 2020b; 2020c; 2020d) has developed 44 hazard dossiers with scenarios that inform and accompany its national risk analysis.

8.4 Role Play

The roles that people play affect their behavior. For example, Cyert, March, and Starbuck (1961) gave participants in a study the role of "cost analyst" or "market analyst." After having been given these roles, the participants made substantially different forecasts based on the same data. Similarly, Babcock et al. (1995) showed that assigning roles in bargaining influences individual judgements of fairness in favor of the interests of that role. Perhaps even more importantly, the interaction of complex organizations with large decision-choice sets is difficult to model through theory and without empirical data. Hence, Parson (1996) notes that **role play is useful when: 1) key outcomes depend on the interacting decisions of multiple agents; 2) decision choice sets are ambiguous or poorly known; and 3) large numbers of complex organizations may be required to work together.**

The decision-options of players in a role play can be **unstructured, semi-structured, or structured.** Furthermore, there are **cooperative games,** with a scripted hazard or adversary, and **adversarial games,** with human opponents. "Serious" games originated in the military context with **wargames** that explored the dynamics of military conflicts and strategies. Originally, wargames strictly focused on the movements of and fighting between two military forces with defined assets and within a limited territory. Nowadays, wargames can be more holistic and include diplomacy as well as elements of armed conflict that are less geographically bound, such as nuclear deterrence and cyber operations. Wargames can be conducted in various forms, including board games and card games. There are also serious computer games, which are discussed below under modelling and simulation (section 9.2). Lastly, wargames have been adapted to business management, where they enable the exploration of strategic interactions between a company and its competitors in a market.

Militaries and businesses engaged in cybersecurity and critical infrastructure sometimes rely on **red teaming** and penetration testing to find and close vulnerabilities. A red team denotes a group with an assigned adversary role either in a wargame or while the targeted organization continues with regular day-to-day operations. The latter is a common approach to find security loopholes to get unauthorized access to sensitive data or material, and subject an organization's plans and assumptions to analysis and challenge. Beyond formally assigned red teams, some also use the term red team mindset to describe critical thinking that challenges groupthink. The **UK Ministry of Defence (2021) offers a fairly comprehensive guide on how to organize red teaming.**

Lastly, in the context of civil protection, cooperative **crisis exercises** are used to explore decision options and institutional overlaps and gaps regarding reactions to large-scale hazards.

An example of a wargame can be found in RAND's Defense of the Baltics (Shlapak and Johnson, 2016). In this series of table-top games which were conducted in 2014 and 2015, RAND explored the possibility of a near-term Russian invasion of the Baltics. The games showed that Russian forces would reach Baltic capitals before reinforcements could arrive. As a result, NATO created the Very High Readiness Joint Task Force to be able to get boots on the ground within a much shorter time period. The telecom provider Swisscom was the first major company in Switzerland to voluntarily and explicitly subject itself to stress tests from a red team in its own Computer Security Incident Response Team. Zhang and Gronvall (2020) offer an overview of some historical uses of red teams with regards to biological threats. The quadrennial *Strategische Führungsübung* (strategic leadership exercise) of the Swiss federal government and the *Sicherheitsverbundübung* (security network exercise) of the Swiss federal government and the Cantons are examples of cooperative crisis exercises. Scenario competitions in which multiple teams compete for the best response to a scripted scenario are another form of cooperative games. An example of this would be the annual Cyber 9/12 Strategy Challenge in Geneva, organized by the Geneva Centre for Security Policy and the Atlantic Council.

As there are no real-world examples of the use of nuclear weapons in a conflict between nuclear powers, wargames have been used in this area. For example, the most realistic nuclear wargame played by the US government during the Cold War was Proud Prophet, which took place in June 1983. The outcome of the fictional two-week game was the massive destruction of the US, Europe, and the Soviet Union, with about one billion deaths, which may have contributed to a reduction in the nuclear threat rhetoric by the Reagan administration in subsequent years (Bracken, 2012, pp. 85–89). Role play in the form of crisis exercises is also used to prepare for biological threats from natural or accidental spillover events. For example, the John Hopkins Center for Health Security has co-organized various table-top exercises, including Dark Winter (2001), Atlantic Storm (2005), Clade X (2018), and Event 201 (2019). In Switzerland, two notable pandemic crisis exercises were the *Strategische Führungsübung 2005* and the *Sicherheitsverbundübung 2014* (Hauri et al., 2020, pp. 54–55, 71–72).

8.5 Fiction

Design fiction: Design fiction uses fictional artifacts and scenarios with actors, environments, and parameters that are unusual, unrealistic, or provocative in order to stimulate reflection, critical questions, and problem finding (Dunne and Rabby, 2013). Design fiction can provide an anti-environment that can reduce and/or reverse preexisting intuitions and feelings on a topic, which can help individuals and organizations to focus on first principles and potentially challenge internalized norms and assumptions. Design fiction often centers on social issues; however, it can also be a playful way to explore scientific laws (e.g., Munroe, 2014; Sandberg, 2018). An example of a design fiction question would be: "should we build robots that eat meat?" This approach may be particularly suited for normative questions, as there is evidence that human moral reasoning is often a post hoc rationalization of pre-existing feelings on a topic and that humans sometimes fail to find any rational argument to explain their moral beliefs (Haidt, 2001).

Science-fiction prototyping: Science-fiction prototypes are stories that explore the implications and uses of future technologies through fictional technological capabilities that are based on real science and technology. Johnson (2011, pp. 25–29) offers a five-step process to create a science-fiction prototype. First, pick the technology, science, or issue that should be explored with the prototype. Embed it in a world with people and locations. Second, introduce one or multiple scientific inflection points in your chosen topic. Third, explore the ramifications of the new science on the world you have described. How does it change people lives? Does it create new opportunities and threats? Fourth, describe the human inflection point. What needs to happen to adapt and fix new problems? Does the technology need to be modified? Is there a new area for experimentation or research? Fifth, what did we learn? What could have been done differently in the scenario? What ramifications has the prototyping shown for the present and how has it changed your outlook?

As an example, NATO's Allied Command Transformation (2016) has commissioned a series of futurist short stories from the consulting firm SciFutures. The resulting anthology, Visions of Warfare 2036, imagines how technological progress and other trends could impact future operations. It also includes discussion questions that each story raises. Similarly, in 2019, the French Defense Innovation Agency launched the Red Team Défense (2021), which is composed of science-fiction writers along with scientific and military experts. Its goal is to anticipate technological, economic, societal, and environmental scenarios that could generate conflicts between 2030 and 2060.

Science-fiction prototypes are usually short stories that look far ahead without necessarily attempting to provide a holistic picture of the future. However, hard science-fiction works that are set relatively close to the present, such as Ghost Fleet: A Novel of the Next World War (Singer and Cole, 2015), can also explore a specific scenario of interest with substantial realism.

9. Modelling and Simulation

Modelling is the definition of a multicomponent formula representing a natural or social phenomenon. Simulation is the calculation of outcomes based on modelling assumptions. Modelling and simulation is at the intersection of forecasting and scenarios. Its outputs can be viewed as conditional, "if this, then that" predictions.

Modelling and simulation could have also been listed as model-based forecasting. As discussed in section 7 on forecasting, human intuition, causal reasoning, and historical data are the sources based on which we predict things. In judgmental forecasting, we usually rely on informal intuition assisted by formal causal reasoning and data. However, we can also try to build a formal model based on human judgement (section 7.3). In statistical forecasting (section 7.4), the focus lies on building handcrafted or trained (section 7.4.1) models that explain future data points primarily or exclusively based on historical data points. This section focuses on models that are mainly based on **causal reasoning informed by domain specific knowledge.** It is listed separately because the uncertainty in values or weights of modelling terms is often **expressed in multiple quantitative reference scenarios.**

In a CBRN context, modelling has been used extensively to predict the epidemiological progression of the COVID-19 pandemic. A classic SIR-model predicts infections, hospitalizations, and deaths based on the effective reproduction number, age-adjusted infection fatality rates, and the share of susceptible (S), infected (I) and recovered or removed (R) persons in a population (Hethcote, 2000; Keeling and Rohani, 2011). Both the reliance and the refusal to rely on epidemiological models to inform public policy during the corona crisis have been discussed with some intensity (e.g., Fuller, 2020; Lipsitch, 2020; Ioannidis, 2020). Some models have been accused of systematically underestimating the impact of COVID-19 (Piper, 2020), whereas other models have been accused of using worst case assumptions to place additional pressure on politicians to act (Dowideit and Nabert, 2021). Either way, such discussions have strengthened the case for **continuously verifying model performance** (Chin et al., 2020). Another noteworthy application of modelling in the CBRN context concerns nuclear weapons. Indeed, the use of computer simulations has been successful in reducing a reliance on the physical testing of such weapons.

9.1 Computational Social Science

Modelling and simulation is increasingly applied to social processes and phenomena, for example, to improve crowd management. The most ambitious project in simulating large, complex systems so far may have been the Living Earth simulator. It was a proposed massive computer simulation intended to bring together interactions of all aspects of life, including economic activities, climate, and other physical processes (Paolucci et al., 2012). In the very long run, it is conceivable that there will be increasingly large and sophisticated simulations that aim to understand and predict the evolution of human civilization (Bostrom, 2003). However, for the foreseeable future, it should be noted that modelling and simulation still has limited applications when it comes to anticipating future technologies, new behavior, or tactics by threat actors.

9.2 Simulation Games

Some scholars strictly reserve the term simulation for self-contained systems that faithfully aim to represent real processes without human intervention (Klabbers, 2009). Others are more open to a hybrid concept of simulation games (Wilson et al., 2009, p. 219). Computer games that aim for realism can be viewed as such a hybrid in the sense that they are still a form of role play. However, as computer game players do not have to calculate the outcomes of their actions themselves, computer games can be based on a much more complicated modelling of reality than board games.

10 Conclusion

In this report, we have examined various approaches and methodologies to foresight, with occasional examples from the CBRN field. The goal of this report is not to prescribe or proscribe any approach. Instead, it is to inform the process of deciding what methodologies are fit for purpose. A few key points highlighted by the report:

- Horizon scanning aims to find leading indicators for future developments, and it profits from the organization of networks that cross disciplinary and institutional boundaries.
- The short-term forecasting of technological and political events contains clear elements of skill. Prediction challenges are an attractive approach to such questions because they establish individual track records.
- Longer-term (five years or more) forecasting of events is not feasible. However, structural shifts and their implications can and should be analyzed.
- Scenarios are a tool to explore the sequences, consequences, and interplay between different actors in a possible future that has been identified as relevant.

Whenever possible, the document provides links to implementation guidelines for methods. Further instructions for many qualitative approaches may be found in the toolkit by GO-Science (2017). This report has aimed to provide an in-depth background section as relevant knowledge on technology and uncertainty tends to be scattered across disciplines. Unfortunately, it has been beyond the scope of this document to add similar contextual information about risk and uncertainty management that follows on from foresight activities. Still, it is worth highlighting that it is not only important to find the right method. It is also crucial to consider which specific decision-making processes are informed through foresight activities and in what format.

Application to the CBRN Context

Together with the ComNBC, the CSS will create a complementary report to this one that is tailored to the needs of CBRN-protection in Switzerland. Still, there are a few high-level points that can be highlighted in this generic report.

- As outliers are likely to create most consequences, there should be caution in deducing trends based on the annual or even decadal impacts (e.g., damages, injuries, deaths) of CBRN accidents and attacks. This does not mean that trend analysis is not suitable to a CBRN-context, but that it should consider factors such as the proliferation and price performance changes of new tools; the number of people with access to materials and tools; the number of small events or near accidents; spillover events; and capacities to respond.
- Aggregated judgmental forecasting methods such as Delphi surveys or superforecasting services may be suitable for CBRN threats, albeit with caveats. Aside from questions on outcomes, these might include questions on risk factors with more regular feedback, such as the number of biosafety or chemical laboratories, or the number of reported spillover events. The use of prediction markets would include a public relations risk.
- Workshops which explore ways of using and combining technology for novel reference scenarios can profit from creativity techniques. However, due to information hazards, closed workshops with invited participants are more advisable than ones that would involve an open and participatory approach.
- Red teaming is useful to find vulnerabilities in terms of access to materials. However, it is important that vulnerabilities are disclosed in a responsible fashion.
- For the main identified threat scenarios, crisis exercises can help improve the complex interplay of responding organizations in an incident.

Lastly, while the above points mainly focus on the national level, it is important to recall that many **CBRN threats are of global concern.** The COVID-19 pandemic is a painful reminder that we all live on the same "Spaceship Earth" (Buckminster Fuller, 1969) and that the preservation of the health of others can be in our own best interest (International Monetary Fund, 2021). Strengthening the funding, verification, and foresight capabilities on an international and global level is crucial for sustainably reducing CBRN risks.

11 Bibliography

Adlin, Tamara, and John Pruitt, *The Essential Persona Lifecycle: Your Guide to Building and Using Personas* (Burlington: Morgan Kaufmann, 2010).

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, *Prediction Machines: The Simple Economics of Artificial Intelligence* (Boston, MA: Harvard Business Review Press, 2018).

Aguilar, Francis J., *Scanning the Business Environment* (New York: Macmillan, 1967).

Allied Command Transformation, *Visions of Warfare: 2036*, (Norfolk, VA: Allied Command Transformation, 2016).

Amanatidou, Effie, Maurits Butter, Vicente Carabias, Totti Könnölä, Miriam Leis, Ozcan Saritas, … , and Victor van Rij, "On Concepts and Methods in Horizon Scanning: Lessons from Initiating Policy Dialogues on Emerging Issues," *Science and Public Policy* 39:2 (2012), 208–221.

Andersen, Birgitte, "The Hunt for S-shaped Growth Paths in Technological Innovation: A Patent Study," *Journal of Evolutionary Economics* 9:4 (1999), 487–526.

Ansoff, H. Igor, *Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion,* (New York: McGraw-Hill Companies, 1965).

Armstrong, J. Scott, "Judgmental Bootstrapping: Inferring Experts' Rules for Forecasting," in: J. Scott Armstrong (ed.), *Principles of Forecasting* (Boston, MA: Springer) 171–192.

Armstrong, Stuart, "*Assessing Kurzweil: The Results,*" lesswrong.com, 16.01.2013.

Armstrong, Stuart, "*Assessing Kurzweil Predictions about 2019: The Results,*" lesswrong.com, 06.05.2020.

Arthur, W. Brian, *The Nature of Technology: What It Is and How It Evolves,* (New York: Free Press, 2009).

Aven, Terje, and Ortwin Renn, *Risk Management and Governance: Concepts, Guidelines and Applications* (Berlin: Springer Science & Business Media, 2010).

Babcock, Linda, George Loewenstein, Samuel Issacharoff, and Colin Camerer. "Biased judgments of fairness in bargaining." *The American Economic Review* 85:5 (1995), 1337–1343.

Bengston, David N., Lynne M. Westphal, and Michael J. Dockry. "Back from the future: the Backcasting wheel for mapping a pathway to a preferred future." *World Futures Review* 12:3 (2020), 270–278.

Bettencourt, Luís M., José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B. West "Growth, Innovation, Scaling, and the Pace of Life in Cities," *Proceedings of the National Academy of Sciences,* 104:17 (2007), 7301–7306.

Bijker, Wiebe E., *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change* (Cambridge, MA: MIT Press, 1995).

BlackRock, "*Investing in the Future with Megatrends,*" blackrock.com, 2021.

Bork, Henrik, and Steffan Heuer, "*The Extraordinary Vision behind Vision Fund,*" rolandberger.com, 23.07.2019.

Bostrom, Nick, "Are We Living in a Computer Simulation?" *The Philosophical Quarterly* 53:211 (2003), 243–255.

Bostrom, Nick, "Information Hazards: A Typology of Potential Harms from Knowledge," *Review of Contemporary Philosophy* 10 (2011), 44–79.

Bostrom, Nick, Thomas Douglas, and Anders Sandberg, "The Unilateralist's Curse and the Case for a Principle of Conformity," *Social Epistemology* 30:4 (2016), 350–371

Box, George, Gwilym Jenkins, George Reinsel, and Greta Ljung, *Time Series Analysis: Forecasting and Control,* 5th ed (Hoboken, NJ: John Wiley & Sons, 2015).

Bracken, Paul, *The Second Nuclear Age: Strategy, Danger, and the New Power Politics* (New York: Macmillan Publishers, 2012).

Braudel, Fernand, *La Méditerranée et le monde méditerranéen à l'époque de Philippe II* (Paris: Librairie Armand Colin, 1949).

Bresnahan, Timothy, and Manuel Trajtenberg, "General Purpose Technologies: 'Engines of Growth?'" *National Bureau of Economic Research Working Paper Series* 4148, (1992). https://www.nber.org/papers/w4148.pdf

Brier, Glenn W., "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review* 78:1 (1950) 1–3.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, …, and Dario Amodei, *"Language Models Are Few-shot Learners,"* arxiv.org, 2020.

Brynjolfsson, Erik, and Andrew McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (New York, NY: W. W. Norton & Company, 2014).

Buckminster Fuller, R., *Operating Manual for Spaceship Earth* (New York: EP Dutton & Co, 1969).

Bush, Vannevar, *Science: The Endless Frontier* (Washington DC: United States Government Printing Office, 1945)

Burkus, David, *The Myths of Creativity: The Truth about How Innovative Companies and People Generate Great Ideas* (Hoboken, NJ: John Wiley & Sons, 2013)

Cairns, George, and George Wright, *Scenario Thinking: Preparing Your Organization for the Future in an Unpredictable World,* 2nd ed (London: Palgrave Macmillan, 2018).

Campanario, Juan Miguel, "Rejecting and Resisting Nobel Class Discoveries: Accounts by Nobel Laureates," *Scientometrics* 81:2 (2009), 549–565.

Centre for Research on the Epidemiology of Disasters, *"Emergency Events Database (EM-DAT): The International Disaster Database,"* emdat.be, 2021.

Chin, Vincent, Noelle I. Samia, Roman Marchant, Ori Rosen, John P. A. Ioannidis, Martin A. Tanner, and Sally Cripps, "A Case Study in Model Failure? COVID-19 Daily Deaths and ICU Bed Utilisation Predictions in New York State," *European Journal of Epidemiology* 35:8 (2020), 733–742.

Coates, Joseph F., (1985) "Foresight in Federal Government Policy Making," *Futures Research Quarterly* 1:2 (1985) 29–53.

Collingridge, David, *The Social Control of Technology* (London, UK: Pinter, 1982).

Collopy, Fred, Monica Adya, and J. Scott Armstrong. "Expert systems for forecasting." In *Principles of Forecasting,* 285–300. (Boston, MA: Springer, 2001)

Cornish, Edward, *Futuring: The Exploration of the Future* (Bethesda, MD: World Future Society, 2004).

Cuhls, Kerstin, *Methoden der Technikvorausschau – eine internationale Übersicht,* isi.fraunhofer.de, 2008. (Karlsruhe: Fraunhofer IRB Verlag, 2008).

Cyert, Richard M., James G. March, and William H. Starbuck. "Two experiments on bias and conflict in organizational estimation." *Management Science* 7:3 (1961), 254–264.

Dafoe, Allan, "On Technological Determinism: A Typology, Scope Conditions, and a Mechanism," *Science, Technology, & Human Values* 40:6 (2015), 1047–1076.

David, Paul A., "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *The American Economic Review* 80:2 (1990), 355–361.

De Girardin, Émile, (1849). *Les cinquante-deux,* Vol.1 (Paris: Michel Lévy Frères, 1849).

Deloitte, *"Beyond the Noise: The Megatrends of Tomorrow's World,"* deloitte.com, 2017.

DeLong, J. Bradford, *"Estimates of World GDP, One Million B.C. – Present",* delong.typepad.com, 1998.

Dowideit, Anette, and Alexander Nabert, *"Innenministerium spannte Wissenschaftler für Rechtfertigung von Corona-Maßnahmen ein,"* welt.de, 07.02.2021.

Drucker, Peter, *Innovation and Entrepreneurship* (Abingdon, Oxon: Routledge, 2014)

Dunne, Anthony and Fiona Raby, *Speculative everything: design, fiction, and social dreaming* (Cambridge, MA: MIT Press, 2013)

Elkington, John, *Green Swans: The Coming Boom in Regenerative Capitalism,* (Austin, TX: Fast Company Press, 2020).

Esvelt, Kevin, "Inoculating Science against Potential Pandemics and Information Hazards," *PLoS Pathogens* 14:10 (2018), e1007286.

European Commission, *"2021 Strategic Foresight Report: The EU's capacity and freedom to act,"* ec.europa.eu, 2021.

European Strategy and Policy Analysis System, *"Global Trends to 2030: Challenges and Choices for Europe,"* espas.secure.europarl.eu, April 2019.

EY, *"Are You Reframing Your Future or Is the Future Reframing You?,"* assets.ey.com, 2020.

Evans, Simon, *"'Profound Shifts' Underway in Energy System, Says IEA World Energy Outlook"* carbonbrief.org, 13.11.2019.

Evans, Simon, *"Solar is now 'cheapest electricity in history', confirms IEA"*, carbonbrief.org, 13.10.2020.

Facebook. *"Prophet: Forecasting At Scale,"* facebook.github.io, 2021.

Fairbank, John K., *"How to Deal with the Chinese Revolution,"* nybooks.com, 17.02.1966.

Federal Chancellery. (2011). *"Perspektiven 2025: Lage- und Umfeldanalyse sowie Herausforderungen für die Bundespolitik,"* bk.admin.ch, 2011.

Federal Chancellery, *Perspektiven 2030: Lage- und Umfeldanalyse sowie Herausforderungen für die Bundespolitik,* bk.admin.ch, 2015.

Federal Chancellery, *Die Schweiz 2030: Was muss die Politik heut anpacken? 77 Antworten,* bk.admin.ch, 2018.

Federal Commission for Nuclear, Biological and Chemical Protection, "Empfehlungen der Eidgenössischen Kommission für ABC-Schutz an den Bundesrat zur Vorsorge und Bewältigung von ABC-Ereignissen," babs.admin.ch, 2011.

Federal Department of Finance, *Band 2A - Voranschlag 2020 mit IAFP 2021–2023 der Verwaltungseinheiten Teil I (Behörden und Gerichte, EDA, EDI, EJPD, VBS),* efv.admin.ch, 2019.

Federal Institute of Intellectual Property, *"Strategic Patent Analysis,"* ige.ch, 2021.

Federal Office for Civil Protection, *Katastrophen und Notlagen Schweiz: Katalog der Gefährdungen* (Bern: Federal Office for Civil Protection, 2019).

Federal Office for Civil Protection, *Katastrophen und Notlagen Schweiz 2020: Bericht zur nationalen Riskoanalyse*. (Bern: Federal Office for Civil Protection, 2020a).

Federal Office for Civil Protection, *"KKW-Unfall,"* babs.admin.ch, 2020b.

Federal Office for Civil Protection, *"Unfall B-Betrieb,"* babs.admin.ch, 2020c.

Federal Office for Civil Protection, *"Unfall C-Betrieb,"* babs.admin.ch, 2020d.

Federal Office for Spatial Development, *"Perspektiven des Schweizerischen Personen- und Güterverkehrs bis 2040,"* are.admin.ch, 2016.

Feenberg, Andrew, *Between Reason and Experience: Essays in Technology and Modernity* (Cambridge, MA: MIT Press, 2010).

Ferguson, Niall, *Doom: The Politics of Catastrophe* (London: Penguin Press, 2021)

Flechtheim, Ossip. K, *Futurologie: der Kampf um die Zukunft* (Verlag Wissenschaft und Politik, 1971)

Fuller, Jonathan. *"Models v. Evidence,"* bostonreview.net, 2020.

Galton, Francis, "Vox Populi," *Nature,* 75 (1907) 450–51.

Gassmann, Oliver, Karolin, Frankenberger, and Michaela Csik, *The Business Model Navigator: 55 Models that Will Revolutionise Your Business* (FT Publishing International, 2015)

Gentner, Dedre, "Structure-mapping: A Theoretical Framework for Analogy," *Cognitive Science* 7:2, (1983), 155–170.

Glenn, Jerome, "The Futures Wheel," in: Jerome Glenn and Theodore Gordon (eds.), *Futures Research Methodology – Version 3.0* (Washington DC: The Millennium Project, 2009).

Google, *"CausalImpact,"* google.github.io, 2021.

Gordon, Theodore, "The Delphi Method", in: Jerome Glenn and Theodore Gordon (eds.), *Futures Research Methodology – Version 3.0* (Washington DC: The Millennium Project, 2009).

Government Office for Science, *"The Futures Toolkit: Tools for Futures Thinking and Foresight across UK Government,"* assets.publishing.service.gov.uk, November 2017.

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans, "When Will AI Exceed Human Performance? Evidence from AI Experts," *Journal of Artificial Intelligence Research* 62 (2018), 729–754.

Gruetzemacher, Ross, and David Paradice, "Toward Mapping the Paths to AGI," in: Patrick Hammer, Pulin Agrawal, Ben Goertzel, and Matthew. Iklé (eds.), *Artificial General Intelligence: 12th International Conference, AGI 2019, Shenzhen, China, August 6–9, 2019 Proceedings* (Cham, Switzerland: Springer Nature, 2019), 70–79.

Gupta, Vinay, *"On Black Elephants,"* howtolivewiki.com, 2009.

Gygli, Savina, Florian Haelg, Niklas Potrafke, and Jan-Egbert Sturm, "The KOF Globalisation Index – Revisited," *The Review of International Organizations,* 14:3 (2019), 543–574.

Habegger, Beat, *"Horizon Scanning in Government: Concept, Country Experiences and Models for Switzerland," Center for Security Studies (CSS) at ETH Zurich,* css.ethz.ch, 2009.

Haerpfer, Christian, Ronald Inglehart, Alejandro Moreno, Christian Welzel, Kseniya Kizilova, Jamie Diez-Medrano, ..., and Bi Puranen (eds.) *"World Values Survey Wave 7 (2017–2020)",* worldvaluessurvey.org, 2020.

Hauri, Andrin, Kevin Kohler, Benjamin Scharte, and Andreas Wenger, "Pandemievorbereitung in der Schweiz: Krisenvorsorge als laufender Lernprozess," in: Oliver Thränert and Benno Zogg (eds.), *Bulletin 2020 zur schweizerischen Sicherheitspolitik* (Zurich: Center for Security Studies (CSS) at ETH Zurich, 2020) 47–79.

Haidt, Jonathan, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108:4 (2001), 814–34.

Hancock, Trevor, Clement Bezold, "Possible Futures, Preferable Futures," *The Healthcare Forum Journal,* 37:2 (1994), 23–29.

Hanson, Robin, "The Policy Analysis Market (A Thwarted Experiment in the Use of Prediction Markets for Public Policy),". *Innovations: Technology, Governance, Globalization,* 2:3 (2007), 73–88.

Henrich, Joseph, *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter,* (Princeton, NJ: Princeton University Press, 2015).

Hernandez, Danny, *"How Our Company Learned to Make Better Predictions about Everything,"* hbr.org, 15.05.2017.

Hernandez, Danny, "A Foreseeable Future," aforeseeablefuture.com, n.d.

Hethcote, Herbert, "The Mathematics of Infectious Diseases," *SIAM Review* 42:4 (2000), 599–653

Hodgson, Tony, and Bill Sharpe, "Deepening Futures with System Structure," in: Bill Sharpe and Kees van der Heijden (eds.), *Scenarios for Success* (Hoboken, NJ: John Wiley & Sons, 2007), 121–144.

Holyoak, Keith, "Analogy and Relational Reasoning," in: Keith Holyoak and Robert Morrison (eds.), *The Oxford Handbook of Thinking and Reasoning* (Oxford: Oxford University Press, 2012), 234–259.

Horowitz, Michael, Julia Ciocca, Lauren Kahn, and Christian Ruhl, *"Keeping Score: A New Approach to Geopolitical Forecasting,"* global.upenn.edu, 2021.

Hubbert, M. King, *Nuclear Energy and the Fossil Fuel* (New York: Drilling and Production Practice, American Petroleum Institute, 1956).

Hughes, Thomas, P., *Networks of Power: Electrification in Western Society, 1880 – 1930* (Baltimore, MD: Johns Hopkins University Press, 1983).

Huurdeman, Anton, *The Worldwide History of Telecommunications* (Hoboken, NJ: John Wiley & Sons, 2003)

Hyndman, Rob J., and George Athanasopoulos, *Forecasting: Principles and Practice,* 3rd ed, otexts.com, 2021. OTexts.com/fpp3

Inayatullah, Sohail, "Causal Layered Analysis: Poststructuralism as Method," *Futures* 30:8 (1998), 815–829.

International Monetary Fund, *"New $50 Billion Health, Trade, and Finance Roadmap to End the Pandemic and Secure a Global Recovery,"* imf.org, 01.06.2021.

International Risk Governance Council, *Introduction to the IRGC Risk Governance Framework: Revised Version 2017.* (Lausanne: EPFL International Risk Governance Center: 2017).

International Risk Governance Council, "Involving Stakeholders in the Risk Governance Process," *EPFL International Risk Governance Center,* 2020.

Ioannidis, John. *"The Totality of Evidence,"* bostonreview.net, 2020.

John Hopkins Center for Health Security, *"Dark Winter,"* centerforhealthsecurity.org, 2001.

John Hopkins Center for Health Security, *"Atlantic Storm,"* enterforhealthsecurity.org, 2005.

John Hopkins Center for Health Security, *"Clade X,"* centerforhealthsecurity.org, 2018.

John Hopkins Center for Health Security, *"Event 201,"* centerforhealthsecurity.org 2019.

Johnson, Brian, *Science Fiction Prototyping: Designing the Future with Science Fiction,* morganclaypool.com, 2011. (Williston, VT: Morgan Morgan & Claypool, 2011)

Joshi, Shashank, *"How Spooks Are Turning to Superforecasting in the Cosmic Bazaar,"* economist.com, 15.04.2021.

Kahn, Herman, *Thinking about the Unthinkable.* (New York: Horizon Press, 1962)

Kahneman, Daniel, and Klein, Gary, "Conditions for Intuitive Expertise: A Failure to Disagree," *American Psychologist* 64:6 (2009), 515–526.

Keeling, Matt J., and Pejman Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton, NJ: Princeton University Press, 2011)

Keidanren, *"Society 5.0: Co-Creating the Future",* keidanren.or.jp, 2018.

Kelly, Kevin, *What Technology Wants* (London: Penguin, 2010).

Kelly, Terra R., William B. Karesh, Christine Kreuder Johnson, Kirsten V. K. Gilardi, Simon J. Anthony, Tracey Goldstein, ..., and Jonna A K Mazet, "One Health Proof of Concept: Bringing a Transdisciplinary Approach to Surveillance for Zoonotic Viruses at the Human-wild Animal Interface," *Preventive Veterinary Medicine, 137* (2017), 112–118.

Kemp, Luke, Laura Adam, Christian R Boehm, Rainer Breitling, Rocco Casagrande, Malcolm Dando, ... and William J. Sutherland, "Point of View: Bioengineering Horizon Scan 2020," *eLife,* 2020, 9, e54489.

Khong, Yuen F., *Analogies at War: Korea, Munich, Dien Bien Phu, and Vietnam Decisions of 1965* (Princeton, NJ: Princeton University Press, 1992).

Klabbers, Jan, "Terminological ambiguity: Game and simulation." *Simulation & Gaming* 40:4 (2009), 446–463.

Klinke, Andreas, and Ortwin Renn, "A New Approach to Risk Evaluation and Management: Risk-based, Precaution-based, and Discourse-based Strategies," *Risk Analysis* 22:6 (2002), 1071–1094.

Klinger, Joel, Juan Mateos-Garcia, and Kostas Stathoulopoulos, *"Deep learning, deep change? Mapping the development of the Artificial Intelligence General Purpose Technology,"* arxiv.org, 2018.

Knight, Frank. H. *Risk, Uncertainty and Profit* (Boston, MA: Houghton Mifflin, 1921)

Kohler, Kevin, and Benjamin Scharte, "Integrating AI into Civil Protection" *CSS Analyses in Security Policy,* 260 (2020).

KPMG, *"Future State 2030: The Global Megatrends Shaping Governments,"* assets.kpmg February 2014.

Kurzweil, Ray, *"How My Predictions Are Faring,"* kurzweilai.net, October 2010.

Laplace, Pierre-Simon, *Essai philosophique sur les probabilités.* (Paris: Courcier, 1814)

Lemahieu, Hervé, and Alyssa Leng, *Lowy Institute Asia Power Index: Key Findings 2020*, think-asia.org, 2020.

Lewis, Gregory, Piers Millett, Anders Sandberg, Andrew Snyder-Beattie, and Gigi Gronvall, "Information hazards in biotechnology," *Risk Analysis* 39:5 (2018), 975–981.

Li, Yanwei, Araz Taeihagh, Martin de Jong, and Andreas Klinke, "Toward a Commonly Shared Public Policy Perspective for Analyzing Risk Coping Strategies," *Risk Analysis* 41:3 (2020), 519–532.

Li, Ming, and Vitányi Paul, *An Introduction to Kolmogorov Complexity and Its Applications* (Cham, Switzerland: Springer, 2019).

Lipsey, Richard G., Kenneth I. Carlaw, and Clifford T. Bekar, *Economic Transformations: General Purpose Technologies and Long-term Economic Growth* (Oxford: Oxford University Press, 2005).

Lipsitch, Marc. *"Good Science is Good Science,"* bostonreview.net, 2020.

List, Dennis, "Scenario Network Mapping: The Development of a Methodology for Social Inquiry" (PhD thesis, University of South Australia, 2005).

Maddison, Angus, *The World Economy: A Millennial Perspective,* (Paris: OECD Publishing, 2001).

Makridakis, Spyros, Allan Andersen, Robert Carbone, Robert Fildes, Michèle Hibon, Rudolf Lewandowski, ..., and Robert Winkler, "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. *Journal of Forecasting* 1:2 (1982), 111–153.

Makridakis, Spyros, Chris Chatfield, Michèle Hibon, Michael Lawrence, Terrence Mills, Keith Ord, LeRoy F. Simmons, "The M2-competition: A Real-time Judgmentally Based Forecasting Study," *International Journal of Forecasting* 9:1 (1993), 5–22.

Makridakis Spyros, and Michèle Hibon, "The M3-Competition: Results, Conclusions and Implications," *International Journal of Forecasting* 16:4 (2000) 451–476.

Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos, "The M4 Competition: 100,000 Time Series and 61 Forecasting Methods," *International Journal of Forecasting* 36:1 (2020), 54–74.

Marchau, Vincent A., Warren E. Walker, Pieter Bloemen, and Steven W. Popper, *Decision Making under Deep Uncertainty: From Theory to Practice* (Cham, Switzerland: Springer Nature, 2019).

Martino, Joseph P., "A Review of Selected Recent Advances in Technological Forecasting," *Technological Forecasting and Social Change* 70:8 (2003), 719–733.

Meadows, Donella H., Dennis L. Meadows, Jørgen Randers, and William W. Behrens, *The Limits to Growth* (New York. Universe Books, 1972).

Meissner, Dirk, "Results and Impact of National Foresight-studies," *Futures* 44:10 (2012) 905–913.

Miles, Ian, "The Development of Technology Foresight: A Review," *Technological Forecasting and Social Change* 77:9 (2010) 1448–1456.

Miles, Ian, and Ozcan Saritas, "The Depth of the Horizon: Searching, Scanning and Widening Horizons," *Foresight* 14:6 (2012), 530–545.

Mina, Michael J., C Jessica E Metcalf, Adrian B. McDermott, Daniel C. Douek, Jeremy Farrar, and Bryan T. Grenfell, "Science Forum: A Global lmmunological Observatory to Meet a Time of Pandemics," *eLife,* 2020 9, e58989.

Ministry of Defence, *Global Strategic Trends: The Future Starts Today* 6th ed., assets.publishing.service.gov.uk, 2018.

Ministry of Defence, *Red Teaming Handbook 3rd ed*. (London: Ministry of Defence UK, 2021).

Molitor, Graham. T., "Molitor Forecasting Model: Key Dimensions for Plotting the "'Patterns of Change,'" *Journal of Futures Studies* 8:1 (2003), 61–72.

MSCI, *"Megatrends,"* msci.com, 2021.

Muehlhauser, Luke *"There was only one industrial revolution,"* lukemuehlhauser.com, 2017.

Munroe, Randall, *"Purity,"* xkcd.com, 2008.

Munroe, Randall, *What If? Serious Scientific Answers to Absurd Hypothetical Questions.* (Boston, MA: Houghton Mifflin Harcourt, 2014)

Naisbitt, John, *Megatrends: Ten New Directions Transforming Our Lives* (New York: Warner Books, 1982).

National Centers for Environmental Prediction *"NCEP Operational Forecast Skill,"* emc.ncep.noaa.gov, 2018.

National Intelligence Council, *"Global Trends 2040: A More Contested World,"* dni.gov, 2021.

NATO, *Strategic Foresight Analysis: 2017 Report, 2017*. act.nato.int, 2017.

NATO, *Framework for Future Alliance Operations: 2018 Report*, act.nato.int, 2018.

OECD, *"Recommendation of the Council on the Governance of Critical Risks"* 2014.

OECD, *Assessing Global Progress in the Govcernance of Critical Risks*. (Paris: OECD Publishing, 2018a).

OECD, *National Risk Assessments: A Cross Country Perspective*. (Paris: OECD Publishing, 2018b)

OECD, *Government Foresight Community Annual Meeting Report 2020: Strategic foresight for future-ready public policy* (Paris: OECD Publishing, 2020).

OECD, "OECD Data," data.oecd.org, 2021.

Oettingen, Gabriele, and Doris Mayer, "The Motivating Function of Thinking about the Future: Expectations versus Fantasies," *Journal of Personality and Social Psychology* 83:5 (2002), 1198–1212.

Our World in Data, "Research and Data to Make Progress against the World's Largest Problems," ourworldindata.org, 2021.

Pandemic Preparedness Partnership, *"100 Days Mission to Respond to Future Pandemic Threats,"* assets.publishing.service.gov.uk, 2021.

Paolucci, Mario, Donald Kossmann, Rosaria Conte, Paul Lukowicz, Panos Argyrakis, Ann Blandford, ..., and Dirk Helbing, "Towards a Living Earth Simulator," *The European Physical Journal Special Topics,* 214:1 (2012), 77–108.

Parson, Edward. A., "What Can You Learn from a Game?" in: Richard Zeckhauser, Ralph L. Keeney, and James K Sebenius (eds.), *Wise Choices: Games, Decisions, and Negotiations.* (Boston, MA: Harvard Business School Press, 1996), 233–252.

Pauker, Ben, *"Epiphanies from Chris Anderson,"* foreignpolicy.com, 29.04.2013.

Persson, Olle, Rickard Danell, and Jesper W. Schneider, "How to Use Bibexcel for Various Types of Bibliometric Analysis," in: Fredrik Åström, Rickard Danell, Birger Larsen, and Jesper Schneider (eds.), *Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at His 60th Birthday* (International Society for Scientometrics and Infometrics, 2009), 9–24.

Petersen, John, and Karlheintz Steinmüller, "Wild Cards," in: Jerome Glenn and Theodore Gordon (eds.), *Futures Research Methodology – V3.0* (Washington DC: The Millennium Project, 2009).

Piper, Kelsey, *"This Coronavirus Model Keeps Being Wrong. Why Are We still Listening to It?"* vox.com, 02.05.2020.

Polak, Fred, *The Image of the Future* (Amsterdam: Elsevier Scientific Publishing Company, 1973).

Popper, Karl, *Objective Knowledge: An Evolutionary Approach* (Oxford: Clarendon Press, 1972).

Popper, Rafael, "How Are Foresight Methods Selected?" *Foresight* 10:6 (2008), 62–89.

Porter, Michael E., *"Competitive Strategy," Measuring Business Excellence* 1:2 (1997), 12–17.

Poteshman, Allen M., "Unusual Option Market Activity and the Terrorist Attacks of September 11, 2001," *The Journal of Business* 79:4 (2006), 1703–1726.

Prpić, John., Araz Taeihagh, and James Melton, "The Fundamentals of Policy Crowdsourcing," *Policy & Internet* 7:3 (2015), 340–361.

PwC, *"Five Megatrends and Their Implications for Global Defense & Security,"* pwc.com, November 2016.

Qu, Ritong, Allan Timmermann, and Yinchu Zhu, "Do Any Economists Have Superior Forecasting Skills?" *Centre for Economic Policy Research,* November 2019. CEPR Discussion Paper No. DP14112.

Quote Investigator, *"The Future Has Arrived – It's Just Not Evenly Distributed Yet,"* quoteinvestigator.com, 24.01.2012.

Quote Investigator, *"It's Difficult to Make Predictions, Especially about the Future,"* quoteinvestigator.com, 20.10.2013.

Ratcliffe, Susan, (Ed.) "Helmut Schmidt 1918–2015. German Statesman; Chancellor of the Federal Republic of Germany 1974–82," in: *Oxford Essential Quotations* 4th ed. (Oxford: Oxford University Press, 2016).

Red Team Défense, "Red Team Défense: les scénarios de la saison 1 dévoilés," 2021.

Renn, Ortwin, *Risk Governance: Coping with Uncertainty in a Complex World* (London/Sterling, VA: Earthscan, 2008).

Reynolds, Daniel, "The Vitruvian Thumb: Embodied Branding and Lateral Thinking with the Nintendo Game Boy," *Game Studies* 16:1 (2016).

Rifkin, Jeremy, *The Third Industrial Revolution; How Lateral Power is Transforming Energy, the Economy, and the World* (London: Palgrave MacMillan, 2011).

Ritchie, Hannah, Edouard Mathieu, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, ..., and Max Roser, *"Coronavirus (COVID-19) Deaths"*, ourworldindata.org, 2021.

Robinson, John. B. "Futures under Glass: A Recipe for People who Hate to Predict," *Futures* 22:8 (1990), 820–842.

Sandberg, Anders, "Blueberry Earth," arxiv.org, 2018.

Sandbrink, Jonas, and Robin Shattock, "RNA Vaccines: A Suitable Platform for Tackling Emerging Pandemics?" *Frontiers in Immunology* 11:608460 (2020)

Saritas, Ozcan, and Jack E. Smith, "The Big Picture – Trends, Drivers, Wild Cards, Discontinuities and Weak Signals," *Futures* 43:3 (2011), 292–312.

Sawhney, Harmeet, and Seungwhan Lee, "Arenas of innovation: Understanding new configurational potentialities of communication technologies," *Media, Culture & Society* 27:3 (2005), 391–414.

Schmitt-Beck, Rüdiger, "Bandwagon Effect," in: Gianpietro Mazzoleni (ed) *The International Encyclopedia of Political Communication* (New York: John Wiley & Sons, 2015).

Schwab, Klaus, *"The Fourth Industrial Revolution: what it means, how to respond"*, weforum.org, 2016.

Schwab, Klaus, *The Fourth Industrial Revolution* (New York, NY: Crown Business, 2017).

Scott, James C., *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven, CT: Yale University Press, 1998).

Sen, Avery, "Island + Bridge: How Transformative Innovation Is Organized in the Federal Government," *Science and Public Policy,* 44:5 (2017), 707–721.

Shlapak, David, and Michael Johnson, *"Reinforcing Deterrence on NATO's Eastern Flank: Wargaming the Defense of the Baltics,"* rand.org, 2016.

Siegel, Laurence, "Black Swan or Black Turkey? The State of Economic Knowledge and the Crash of 2007–2009," *Financial Analysts Journal* 66:4 (2010), 6–10. 10.2307/25741280.

Singer, Peter W., and August Cole, *Ghost fleet: A Novel of the Next World War* (Boston, MA: Houghton Mifflin Harcourt, 2015).

Slaughter, Richard A., "Looking for the Real 'Megatrends,'" *Futures* 25:8 (1993), 827–849.

Šmihula, Daniel, "Waves of technological innovations and the end of the information revolution," *Journal of Economics and International Finance* 2:4, (2010) 58–67.

Smil, Vaclav, *Energy and Civilization* (Cambridge, MA: MIT Press, 2017)

Smil, Vaclav, *Growth: from Microorganisms to Megacities* (Cambridge, MA: MIT Press, 2019)

Snyder, Glenn H., and Paul Diesing, *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises.* (Princeton, NJ: Princeton University Press, 1977)

Sornette, Didier, "Dragon-Kings, Black Swans and the Prediction of Crises," *International Journal of Terraspace Science and Engineering* 2:1 (2009), 1–18.

Stirling, Andy, "Science, Precaution, and the Politics of Technological Risk: Converging Implications in Evolutionary and Social Scientific Perspectives," *Annals of the New York Academy of Sciences* 1128:1 (2008), 95–110.

Sweeney, John, "Infectious Connectivity: Affect and the Internet in Postnormal Times," in: Jenifer Winter and Ryota Ono (eds.), *The Future Internet* (Cham, Switzerland: Springer International Publishing, 2015), 107–123.

Taleb, Nassim N. *The Black Swan: The Impact of the Highly Improbable* (New York: Random House, 2007).

Tetlock, Philip. E., and Dan Gardner, *Superforecasting: The Art and Science of Prediction,* (New York: Random House, 2016).

Thränert, Oliver, and Andreas Wenger, "'Versuchen, auch das Undenkbare zu denken,' Interview mit Armeechef Thomas Süssli," in Oliver Thränert and Benno Zogg (eds.) *Bulletin 2020 zur schweizerischen Sicherheitspolitik* (Zurich: Center for Security Studies (CSS) at ETH Zurich, 2020), 13–27.

Toffler, Alvin, *The Third Wave* (New York, NY: Bantam Books, 1980).

Troianovski, Anton, *"Dortmund Bomb Suspect Attacked Soccer Team to Make $1 Million From Stock Drop,"* wsj.com, April 2021.

Twitter, *"AnomalyDetection R Package,"* github.com, 2021.

United Nations, *"Our Common Agenda: Report of the Secretary-General,"* un.org, 2021a.

United Nations, "UNdata," data.un.org, 2021b.

United Nations Development Programme, *"Foresight Manual: Empowered Futures for the 2030 Agenda,"* undp.org, 2018

United Nations Office for Disaster Risk Reduction, *"Sendai Framework for Disaster Risk Reduction 2015–2030,"* preventionweb.net, 2015.

United Nations Office for Disaster Risk Reduction, *"Measuring Implementation of the Sendai Framework,"* sendaimonitor.undrr.org, 2021.

Rumsfeld, Donald, and Richard Myers, *"DoD News Briefing – Secretary Rumsfeld and Gen. Myers,"* archive.defense.gov, 2002.

Van der Heijden, Kees, *Scenarios: The Art of Strategic Conversation* 2nd ed (Hoboken, NJ: John Wiley & Sons, 2004).

Van Eck, Nees J., and Ludo Waltman, "Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping," *Scientometrics* 84:2 (2009), 523–538

Verity, Julie, *The New Strategic Landscape: Innovative Perspectives on Strategy* (New York: Macmillan, 2012).

von Foerster, Heinz, Patricia M. Mora, and Lawrence W. Amiot, "Doomsday: Friday, 13 November, AD 2026," *Science* 132:3436 (1960), 1291–1295.

Walker, Warren, Robert Lempert, and Jan Kwakkel "Deep Uncertainty". In: Gass S.I., Fu M.C. (eds) Encyclopedia of Operations Research and Management Science. (Boston, MA: Springer, 2013)

Waltz, Kenneth N., *Theory of International Politics* (Long Grove, IL: Waveland Press, 1979).

Wenger, Andreas., Jasper, Ursula, and Myriam Dunn-Cavelty, *The Politics and Science of Prevision: Governing and Probing the Future* (London: Routledge, 2020).

Wiener, Norbert, *Cybernetics: Or Control and Communication in the Animal and the Machine* (Cambridge, MA: Technology Press, 1948).

Williams, Rosalind, *Retooling: A Historian Confronts Technological Change* (Cambridge, MA: MIT Press, 2002).

Wilkinson, Angela, *"Strategic Foresight Primer,"* espas.secure.europarl.eu, 2017.

Wilson, Dominic, and Roopa Purushothaman, "Dreaming with BRICs: The Path to 2050," *Goldman, Sachs & Company,* 2003. Global Economics Paper No. 99.

Wilson, Katherine, Wendy Bedwell, Elizabeth Lazzara, Eduardo Salas, C Shawn Burke, Jamie Estock, ..., and Curtis Conkey, "Relationships between Game Attributes and Learning Outcomes: Review and Research Proposals," *Simulation & Gaming* 40:2 (2008), 217–266.

Winner, Langdon, "Do Artifacts Have Politics?" *Daedalus,* 109:1 (1980), 121–136.

Woo, Gordon, *"Epilogue: Black and Red Swans,"* in: *Calculating Catastrophe* (London: Imperial College Press, 2011), 322–324.

World Bank, *"GDP growth (annual %),"* data.worldbank.org, 2021a.

World Bank, *"World Bank Open Data,"* data.worldbank.org, 2021b.

World Intellectual Property Organization, *"The WIPO Manual on Open Source Patent Analytics,"* wipo-analytics.github.io, 2016.

Wright, Thomas P., "Factors Affecting the Cost of Airplanes," *Journal of the Aeronautical Sciences* 3:4 (1936), 122–128.

Wucker, Michele, *"Thinking Ahead with YGL Michele Wucker,"* youtube.com, 05.04.2013.

Yokoi, Gunpei, and Makino Takefumi, 横井軍平ゲーム館 [Gunpei Yokoi Game House] (Tokyo: ASCII Media Works, 1997)

Zeisberger, Claudia and David Munro, *"Dirty White Swans: Could Unexpected Extreme Events Put You Out of Business?,"* publishing.insead.edu, 2010.

Zhang Lisa and Gigi Gronvall, "Red Teaming the Biological Sciences for Deliberate Threats," *Terrorism and Political Violence* 32:6 (2018), 1225–1244.

About the Author

Kevin Kohler is a researcher in the Risk and Resilience Team at the Center for Security Studies (CSS) at ETH Zürich. His research interests include the use of information and communication technologies in disaster risk management as well as the long-term trajectory and politics of digital technologies.

CSS
ETH Zürich

The **Center for Security Studies (CSS) at ETH Zürich** is a center of competence for Swiss and international security policy. It offers security policy expertise in research, teaching and consulting. The CSS promotes understanding of security policy challenges as a contribution to a more peaceful world. Its work is independent, practice-relevant, and based on a sound academic footing.